INVESTIGATION OF QUEUING AND RELATED SYSTEMS WITH THE PHASE METHOD

Rolf A. Schassberger
University of Calgary
Calgary, Alberta, Canada

ABSTRACT

The method of phases - as invented by Erlang - can be extended to yield an approximation scheme for the calculation of quantities associated with random processes that do not necessarily involve any phase-type distributions. It is the purpose of this paper to familiarize the reader with this scheme by means of two queuing examples. The first example - the queue M/G/1 - is chosen to show how the scheme yields results that have been previously obtained in many different ways. The second example - a priority queue - has received only little attention in the literature, and new results are derived in this paper.

1. THE PHASE METHOD

The phase method as a tool for the investigation of queuing systems has been invented by Erlang and has since been widely used. The method was designed to allow the use of Markov chain techniques in situations where this alley of approach is not obvious. Erlang achieved the goal by representing (or approximating) non-exponential non-exponential lifetimes by a sum of some fixed number of independent exponential phases. Gaver [3] extended this idea to allow sums of random length, in this way widening the range of admissible lifetime distributions to any with a distribution function of the type

\[ F(t) = \sum_{n=0}^{\infty} f_n E^n_{\mu}(t), \]

where \( f_n \) is an arbitrary discrete probability distribution and \( E^n_{\mu}(t) \) the n-fold convolution of the exponential distribution with parameter \( \mu \).

Schassberger [8] has noticed that the scope of this method is virtually unlimited due to the fact that any lifetime distribution function can be obtained as the weak limit of distribution functions of the phase type (1.1). Precisely, we have the

Theorem: Let \( F(t) \) be a distribution function concentrated on \([0,\infty)\) and let, for \( \mu > 0 \),

\[ F_\mu(t) = F(0) + \sum_{k=1}^{\infty} (F(k\mu) - F((k-1)\mu)) E^k_{\mu}(t). \]

Then, for all continuity points of \( F(t) \),

\[ \lim_{\mu \to 0} F_\mu(t) = F(t). \]

This theorem - for a proof see Schassberger [11] - may invite us to try the following recipe for the analysis of stochastic processes: if lifetime distributions occur in the definition of a given process, substitute some or all of these by phase-type distributions, thus obtaining an approximating process; analyse the approximating process via a phase process that should have emerged through the above substitutions; finally carry out a limiting analysis towards results for the original process.

It has been amply demonstrated (Schassberger [9], [10], [11]) that for a large variety of queuing systems this recipe works very well. In particular, the transient behaviour of G/G/1, G/M/s, G/G/1 with different priority disciplines, and feedback systems with round-robin and foreground - background scheduling has been successfully studied in this way. While in all these examples the investigation of the approximating process can be based on widely known classical concepts and methods of analysis, the limiting step is most economically performed using the more recent weak convergence theory of probability measures. Very roughly speaking, one views the original and the approximating processes as probability measures \( P \) and \( P_a \), say, respectively, on some suitable function space \( S \), say. One then establishes weak convergence \( P_a \Rightarrow P \), defined to hold if

\[ \int_S f \, dP_a \Rightarrow \int_S f \, dP \]
for some class of functions $f$ mapping $S$ into the reals. Finally, special such functions will render (1.2) the desired statement, such as a statement about certain moments or Laplace transforms. For the analyst who does not bother about proving weak convergence, the limiting step boils down to obtaining $\int f \, dp$ from $\int f \, P_\alpha$, which in all our examples is trivial.

It is the purpose of this paper to familiarize the reader with this extended phase method by means of two examples. The first example is chosen such as to enable us to derive well known results in this manner, the second example will present new results.

2. THE $M/G/1$ EXAMPLE

The familiar $M/G/1$ queue makes a convenient example for the demonstration of our ideas. We choose the virtual waiting time process \( \{W_t\}; \ t \geq 0 \), well known to be Markovian, as the object of our study.

Takacs has popularized the idea of taking the point of view that upon the arrival of a customer a random experiment is carried out to determine right away this customer's eventual service time. Then \( W_t \) can be interpreted as the sum of unexpended service times facing the server at time \( t \). If the distribution function of the service times is

\[
F_{\beta}(x) = \sum_{k=0}^\infty b_k E_k(x),
\]

(2.1)

i.e. of phase type, then we may adopt the point of view that at arrival times \( k \) independent \( \beta \)-phases are joining the system (with probability \( b_k \)). Each of these will keep the server busy for a random time period, and if \( K(t) \) is the number of not completely processed phases facing the server at time \( t \), if the distribution function of the service times is

\[
B(x) = \sum_{k=0}^\infty b_k E_k(x),
\]

(2.1)

then

\[
P\{W(t) \leq x\} = \sum_{k=0}^\infty P(K(t) = k) E_k(x).
\]

(2.2)

and we may interpret this as having initially \( k_0 \) phases in the system. Now observe that the process \( \{K(t); t \geq 0\} \) is Markovian with stationary transition probabilities \( P_{ik}(t), i, k = 0, 1, 2, \ldots \). Were these known, we could obtain

\[
P\{W(t) \leq x|K(0) = k_0\} = \sum_{k=0}^\infty P_{k_0 k}(t) E_k(x)
\]

(2.3)

If originally we want a general service time distribution function \( B(x) \), we may obtain it as \( \mu = \mu_{k_0} \) in such a way that (2.2) remains true, then \( E_{\mu}(\beta) \) converges to the desired initial condition for the waiting time \( W(0) \). Finally, therefore, if in (2.3) we let \( \mu = \mu_{k_0} \) as described, we may end up with an expression for \( P\{W(t) \leq x|W(0) = x_0\} \), where now \( W(t) \) is the virtual waiting time for the \( M/G/1 \) queue with that general service time distribution function \( B(x) \).

Let us follow this route. The Kolmogoroff forward system for the functions \( P_{k_0 k}(t) \) is given by

\[
\frac{d}{dt} P_{k_0 k}(t) = -v P_{k_0 k}(t) + \mu P_{k_0 k+1}(t) + \frac{\lambda}{k_0} b_{k-1} P_{k_0 k-1}(t)
\]

for \( k, k = 0, 1, 2, \ldots, t \geq 0 \),

\[
v = \begin{cases} \lambda + \mu & \text{for } k > 0, \\ \lambda & \text{for } k = 0, \end{cases}
\]

and with initial conditions

\[ P_{k_0 k}(0) = \delta_{k, k_0} \]

Introducing the Laplace transforms

\[
\hat{P}_{k_0 k}(\theta) = \int_0^{\infty} e^{-\theta t} P_{k_0 k}(t) \, dt, \quad \theta > 0,
\]

into this system yields the system of linear equations

\[
(\theta + \nu) \hat{P}_{k_0 k}(\theta) = \delta_{k, k_0} + \nu \hat{P}_{k_0 k+1}(\theta) + \frac{\lambda}{k_0} b_{k-1} \hat{P}_{k_0 k-1}(\theta)
\]

(2.4)

Now, with \( W(t) \) as in (2.3), we observe that, for \( Re(\theta) \geq 0 \),

\[
\int_0^\infty \int_0^{\infty} e^{-\theta t - \nu \tau} \, d_\tau P(W(t) \leq x|K(0) = k_0) \, dt
\]

\[
= \sum_{k=0}^\infty \hat{P}_{k_0 k}(\theta) \frac{1}{(1 + \theta \nu)} = 1(\theta, \nu; k_0)
\]

(2.5)

From (2.4) we derive the relation

\[
1(\nu, \theta; k_0)(\theta - w + \lambda - \lambda b(w)) = 1 + \nu \mu_{k_0} - \nu \hat{P}_{k_0 k_0}(\theta),
\]

(2.6)

where \( b(w) \) is the Laplace-Stieltjes transform of the distribution function (2.1). There is a well known function \( \phi(\theta) \) with \( |\phi(\theta)| \leq 1 \), see Cohen [1], p.625 such that for \( w = \theta + \lambda - \lambda \phi(\theta) \) the factor \( (\theta - w + \lambda - \lambda b(w)) \) of \( 1(\theta, \nu; k_0) \) in (2.6) disappears. Hence we obtain

\[
P_{k_0 0}(\theta) = \frac{1}{\theta + \lambda - \lambda \phi(\theta)} \left[ 1 + \theta + \lambda - \lambda \phi(\theta) - \kappa_0 \right]^{-k_0}
\]

and

\[
1(\nu, \theta; k_0) = \frac{1}{\theta + \lambda - \lambda \phi(\theta)} \left[ 1 + \theta + \lambda - \lambda \phi(\theta) - \kappa_0 \right]^{-k_0}
\]

(2.7)

Letting \( \mu = w \) and \( k_0 = k_0 \) as described we finally arrive at the well known \( M/G/1 \) results.
might observe that we could deal with the problem of valid without any restrictions for $B(x)$. Our variety of other queuing systems would be a matter of becomes clear that corresponding proofs for a large Thus it seems that the upshot of our method is simply a Basically, the decisive feature of the method lies in With respect to some more discussion of our method we using the phase method can safely forget about these valid without any restrictions for $B(x)$.

Our derivation of these results is carried out under the assumption that the transform in (2.5) converges to the one in (2.8) as $\mu,k \to \infty$. This fact, of course, requires a proof, which we will, however, not give here. Instead, we refer the reader to the papers by Kennedy (116) and Whitt (112), which contain a proof just along the lines outlined in section 1. From these papers it becomes clear that corresponding proofs for a large variety of other queuing systems would be a matter of routine. The analyst who just wants to produce results using the phase method can safely forget about these proofs.

With respect to some more discussion of our method we might observe that we could deal with the problem of calculating the $P(W(t) \leq x | W(0) = x)$ by directly considering the forward equation for the process $(W(\tau); \tau \geq 0)$, instead of that for the process $(K(\tau); \tau \geq 0)$. We would have to impose some restrictions on $B(x)$ as well, could then derive (2.7) and (2.8) under these restrictions, and could finally try to argue that (2.7) and (2.8) are valid for all $B(x)$.

Thus it seems that the upshot of our method is simply a special way of restricting the class of distribution functions $B(x)$ with which to start the analysis. Indeed, from a purely theoretical point of view, this is correct. It is the practical aspects of our approach that seem to be its important features.

Basically, the decisive feature of the method lies in the substitution of a problem involving a non-discrete set of states by one with a discrete such set. Where, as in our example, the result is a Markov chain problem, the analyst can resort to very well known methods and results, or, to put it the other way around, more people can be expected to proceed easily with the analysis as could be for most other approaches. In [1] we have demonstrated that, with the phase method as a tool, the analysis of many queuing systems becomes more or less a routine process, mechanised to such an extent that boredom sets in. Despite this psychological drawback we dare to state here a maxim: if you have to analyse a queuing system, first try the phase method.

In the next section we will come up with some more evidence for the usefulness of our maxim.

### 3. A Priority Queuing Example

We consider a single server who is facing two independent input streams of customers. One stream consists of high priority customers and is Poisson with parameter $\lambda_1$. The other stream, consisting of low priority customers, is a renewal stream with distribution function $A(t), A(0) = 0$, for the interarrival times. All customers are served, and the server proceeds in the standard preemptive resume fashion.

Our objective is it to provide a basis for a detailed study of this system, to be achieved by the calculation of the transition probabilities of a certain Markov process describing the system. Previous studies have been carried out by Sahin [7] and [8] and in [4], ahin investigates the virtual waiting time of low priority customers (if it exists). These results and many more will easily follow from ours.

The process to be studied is $\{W_1(t), W_2(t), Y(t), N(t); t \geq 0\}$, where $W_1(t), W_2(t)$ is the workload of the server due to customers of high (low) priority, $Y(t)$ the waiting time for the next arrival of a low priority customer, and $N(t)$ the number of low priority customers that arrived in $[0,t)$. If we assume from now on that the service times of high (low) priority customers form an i.i.d sequence of random variables with distribution function $B_1(t)$ ($B_2(t)$) and that these sequences are independent of each other as well as of the input streams, then a little reflection reveals that the above process is Markovian. We seek a useful representation of its transition probabilities.

In applying the phase method in very much the same way as for the M/G/1 example, we assume temporarily the $B_i(t)$ of phase-type, i.e.

$$B_{1i}(t) = \sum_{k=0}^{\infty} b_{ik} e^{-\lambda_i t} t^k,$$

In addition we assume for the time being the same for $A(t)$, i.e.

$$A(t) = \sum_{m=0}^{\infty} a_m E_\lambda(t),$$

and we may use the interpretation that at his time of arrival a low priority customer initiates the arrival procedure of the next low priority customer, this procedure consisting of the completion of $m$ successive independent $\lambda$-phases (with probability $a_m$). Thus, the next low priority customer arrives at the time of completion of the last of these phases.

Let $K_i(t), i=1,2$, be the number of $\mu_i$-phases present in the system and $M(t)$ the number of $\lambda$-phases to be completed for the current arrival procedure. Then the process $\{K_1(t),K_2(t),M(t),N(t); t \geq 0\}$ is Markovian with discrete state space and stationary transition probabilities and, in analogy to (2.3), we have

$$P(W_1(t) \leq x_1, W_2(t) \leq x_2, Y(t) \leq y, N(t) = n \mid K_1(0) = k_1, K_2(0) = k_2, Y(0) = y_0, N(0) = n_0) = k_1.$$
The discussion around (2.3) similarly applies here and suggests the study of the discrete-state process. The forward system for the functions (3.4) is - suppressing the starting state in the notation - given by

\[
\frac{d}{dt} p(k_1, k_2, m, n; t) = -(\lambda + \mu) p(k_1, k_2, m, n; t) + \nu_1 p(k_1 + 1, k_2, m, n; t) + \nu_2 p(k_1, k_2 + 1, m, n; t) + \lambda p(k_1, k_2, m + 1, n; t)
\]

(3.5)

\[
\nu = \begin{cases} 0 & k_1 > 0 \\ \nu_2 & k_1 = 0, k_2 > 0 \\ \nu_1 & k_1 = 0 = k_2 \\ 0 & k_1 > 0 \end{cases}
\]

\[
\nu^* = \begin{cases} 0 & k_1 > 0 \\ \nu_2 & k_1 = 0 \end{cases}
\]

\[
p(\nu^*, \nu^*, -1; \nu, \nu) = 0
\]

Introducing Laplace transforms \( \hat{p}^r(\nu^*, \nu^*, -1; \nu, \nu) \) as in (2.4), we obtain from (3.5) for \( \lambda > 0, |s| \leq 1, \text{Re}(s) \geq 0 \) and \( \text{Re}(s) \geq 0 \) the relation

\[
(\lambda - \omega_1 - s + \lambda_1 (1 - b_1(s_1)))p(\omega_1, \omega_2, s, z; \theta) = \left( 1 + \frac{1}{\nu^*} \right)^{-k_1} \left( 1 + \frac{1}{\nu^*} \right)^{-k_2} \left( 1 + \frac{s_m^*}{\lambda_1} \right) \left( \frac{s_m^*}{\lambda_1} \right)^{-\omega_1} \left( \frac{s_m^*}{\lambda_1} \right)^{-\omega_2}
\]

(3.6)
where we now talk about the original process. With this result we will have laid the basis for a detailed study of our queuing system, encompassing such quantities as actual waiting times, queue lengths, occupation times, etc. The guidelines for such a study are contained in Schassberger [11], chapter II.

Due to limited space, our analysis of (3.6) will be somewhat narrative. We list a few well known facts: If $z_1(w)$ is the Laplace-Stieltjes transform of the busy period in a $M/G/1$ queue defined by $\lambda_1$ and $B_1(t)$, and if

$$\psi(w) = w + \lambda_1(1 - z_1(w)),$$

then, for $w_1 = w = \psi(\theta - s)$, the factor of $1$ in (3.6) becomes $0$ (Cohen [1], p 625). Furthermore (Cohen [1], p 246)

$$\frac{1}{\psi(w)} = \int_0^\infty e^{-st} \Pr(W_1(t) = 0|W_1(0) = 0)dt,$$

where $W_1(t)$ is the workload in the above $M/G/1$ queue (and identical with the high priority workload in our system), and $b_2(\psi(w))$ is the Laplace-Stieltjes transform of the completion time for a low priority service in our queue (Jaiswal [5], p 85). Finally, in $0 \leq \Re(s) \leq \theta$ and for $|z| \leq 1$, we have

$$1 - ze^{-\theta(s - z)} = (1 - f^+(s,z;\theta))(1 - f^-(s,z;\theta)),$$

where, for instance,

$$f^+(s,z;\theta) = \mathbb{E}[z_1^{-}\theta Z_1]-\mathbb{E}[z_2^{-}\theta Z_2],$$

$Z_1$ and $Z_2$ being busy and idle period, respectively, and $N'$ being the number of customers served in a busy period, for a $G/G/1$ queue defined by arrival distribution $a(s)$ and service distribution $b_2(\psi(\theta - s))$ (Schassberger [11], chapter II).

Purely for the sake of simplicity we now restrict ourselves to the case $k_0 = k_2 = 0$, $m = 1$, $n = 0$. Putting $w_1 = w_2 = \psi(\theta - s)$ in (3.6) we obtain the relation

$$0 = -\left(1 + \frac{\theta}{\lambda}(1 - f^+(s,z;\theta))(1 - f^-(s,z;\theta))\right)\lambda g(\phi(\theta - s),\psi(\theta - s),z;\theta)$$

$$-\left(1 + \frac{\theta}{\lambda}\right)f^+(s,z;\theta) + \frac{1}{\phi(\theta - s)} - \frac{1}{1-e^{\theta(s,z;\theta)}} = \int_0^\tau e^{-\theta(t,z;\theta)}dt,$$

valid in $0 \leq \Re(s) \leq \theta$ and for $|z| \leq 1$. Letting

$$\frac{1}{1 - f^+(s,z;\theta)} = \int_{\tau=0}^\tau e^{-\theta(t,z;\theta)}dt$$

(renewal theory), we may write

$$\frac{1}{1 - f^+(s,z;\theta)} = u^+(s,z;\theta) + u^-(s,z;\theta)$$

with

$$u^+(s,z;\theta) = \int \int e^{-\theta(t-t')-\theta \tau} \Pr(W_1(t) = 0|W_1(0) = 0)dt dt'$$

and

$$u^-(s,z;\theta) = \int \int e^{-\theta(t-t')-\theta \tau} \Pr(W_1(t) = 0|W_1(0) = 0)dt dt'$$

We rewrite (3.7) to yield

$$\left[1 + \frac{\theta}{\lambda}(1 - f^-(s,z;\theta))\right]\phi(\theta - s) = u^-(s,z;\theta) - \left(1 + \frac{\theta}{\lambda}\right)(1 - f^-(s,z;\theta))\lambda g(\phi(\theta - s),\psi(\theta - s),z;\theta)$$

This relation is still valid in $0 \leq \Re(s) \leq \theta$, only, but the left and right-hand sides are now analytic in $\Re(s) > 0$ and $\Re(s) < \theta$, respectively. Moreover, as can be readily seen, both sides are bounded in the mentioned regions of analyticity. Hence, by Liouville's theorem, both sides represent one and the same constant $c(z;\theta)$. But $c(z;\theta) = u^-(\lambda,z;\theta)$, and thus

$$k(s,z;\theta) = \frac{\lambda g(\phi(\theta - s),\psi(\theta - s),z;\theta)}{1 - f^-(\phi(\theta - s),z;\theta)}$$

These intermediate results are interesting in themselves. We proceed, however, deriving from (3.6) the relation

$$0 - w_1 - s + \lambda_1(1 - b_1(w_1))(1(w_1,w_2,s,z;\theta) - 1(w_1,s,z;\theta)) = -(1 - za(s)b_2(w_1))\lambda g(w_1,w_2,z;\theta) - \lambda g(w_1,w_1,z;\theta)$$

($k(s,z;\theta)$ is the relevant root of $w - \phi(\theta - s) = 0$).

These intermediate results are interesting in themselves. We proceed, however, deriving from (3.6) the relation

$$0 - w_1 - s + \lambda_1(1 - b_1(w_1))(1(w_1,w_2,s,z;\theta) - 1(w_1,s,z;\theta)) = -(1 - za(s)b_2(w_1))\lambda g(w_1,w_2,z;\theta) - \lambda g(w_1,w_1,z;\theta)$$

with $u^+(s,z;\theta)$ and $u^-(s,z;\theta)$.

$$= -\left(1 - za(s)b_2(w_2)\right)\lambda g(w_1,w_2,z;\theta) - \lambda g(w_1,w_1,z;\theta)$$

$$= -(w_1 - w_2)(h(w_1,w_2,z;\theta) - h(w_1,s,z;\theta)).$$

$$+ za(s)(b_2(w_2) - b_2(w_1))\lambda g(w_1,w_1,z;\theta).$$

$$$$

324/5 ITC 7
Letting $w_1 = \phi(\theta-s)$, we obtain

$$0 = \frac{1}{\psi(\theta-s) - w_2} \left[ \lambda g(\phi(\theta-s), \psi(\theta-s), z; 0) - \lambda g(\phi(\theta-s), \psi(\theta-s), z; 0) \right]$$

$$+ \frac{1}{1 - z(s)b_2(w_2)}$$

$$+ \frac{za(s)}{1 - z(s)b_2(w_2)} \left[ \lambda g(\phi(\theta-s), \psi(\theta-s), z; 0) \right]$$

valid in $0 \leq \Re(s) \leq \theta$. We will be able to write the last expression in (3.15) as the sum of two functions $v^+(w_1, s, z; \theta)$ and $v^-(w_1, s, z; \theta)$, these being analytic in $\Re(s) > 0$, $\Re(s) < 0$, respectively. Moreover, we will find that $v^+ = 0$ as $s \to +\infty$, $v^- = 0$ as $s \to -\infty$. Once more applying the argument following (3.11) we then arrive at the results

$$h(w_1, s, z; 0) = k(s, z; 0) - v^+(w_1, s, z; \theta)(1 - z(s)b_2(w_2)),$$

(3.16)

$$\lambda g(w_1, w_1, w_1; 0) = \lambda g(w_1, w_1, z; 0) - (w_1 - w_2)v^-(w_1, w_1, z; 0),$$

which enable us to write down the desired expression for $\lambda(w_1, w_1, w_1; z; \theta)$. Now, if $T_1, T_2, \ldots$ are the arrival times of the low priority customers ($T_1$ is distributed according to $(1 + x)^{-1}, T_n = \sum_{i=1}^{n-1} x_i$ for $n > 1$, according to $a(s)$), then

$$\lambda g(w_1, w_1, w_2; 0) = \frac{Z_1(v)}{n} \left[ \frac{e^{-\theta t_n}}{w_1 (T_n -) - w_2 (T_n -)} \right]_1^\infty (3.18)$$

as follows via arguments discussed in Schassberger [11], chapter I6. Using the representation

$$b_2(w) - b_2(w_2) = \int \int \int e^{-w(t-t')}b_2(t')dt' dt,$$

(see Doetsch [2], p 35), we can therefore write

$$za(s) = \frac{b_2(\phi(\theta-s)) - b_2(w_2)}{\phi(\theta-s) - w_2} \lambda g(\phi(\theta-s), \psi(\theta-s), z; \theta)$$

$$= - \frac{\lambda g(\phi(\theta-s), \psi(\theta-s), z; 0)}{n} \int \int \int e^{-sy} b_2(t') dt' dt, \quad t_k = 0, x_k = 0$$

$$- (\phi(\theta-s) - w_2) t - t' - k \lambda = (\phi(\theta-s)x_k)$$

$$e^{-\theta t_n} - w_1 (T_n -) + w_2 (T_n -) \leq x_k dt \int \int \int \int e^{-\theta t_n} - w_1 (T_n -) + w_2 (T_n -) \leq x_k dt \int \int \int \int$$

Here $Z_1(v)$ is the distribution function belonging to $z_1(w)$. If $v^+$ and $v^-$ are obtained from this expression by integrating only over the regions $y - t - x_k - v > 0$, $y - t - x_k - v \leq 0$, respectively, then $v^+$ and $v^-$ are as claimed above. Moreover, taking (3.13) and (3.18) into account, we can claim that $v^+$ and $v^-$ are expressed in terms of known quantities. Thus, our objective has been gained.

The formalities of the limiting procedure are now very simple, just as in the M/G/1 example. If all quantities that allow it are being interpreted as quantities belonging to the original process, then (3.9) and (3.10) remain unchanged, (3.12) and (3.13) become

$$k(s, z; 0) = \sum_{n=0}^{\infty} t_n \int \int \int e^{-(\theta-t)(s, z; 0)}$$

$$\sum_{n=0}^{\infty} t_n \int \int \int$$

respectively, the expressions for $v^+$ and $v^-$ with initial conditions as in (3.19) - remain unchanged, (3.16) and (3.17) - using (3.18) with initial conditions again as in (3.19) - remain unchanged, etc. For (3.19) and (3.20) we used the fact $u^-(\lambda, z; \theta) = 0$ as $\lambda \to \infty$, which can easily be seen to hold from (3.8) and (3.10). We refer again to the paper by Whitt [12], where the validity of these formal limiting steps has been established in a general context. We are going to close the analysis of this queuing system with a few remarks concerning various consequences of our results.

Let us begin with demonstrating the possibility of deriving from our seemingly very complicated results more simple ones. From (3.19) we have
We make all these assumptions and proceed to check the validity of (3.21). Let \( Z_1(n) \) be distributed as the sum of \( n \) independent random variables, each of them with the distribution of \( Z \). Then
\[
\frac{1}{1 - f^{+}(s,1;\theta)} = \int_{0}^{\infty} \mathbb{E}[s \xi^{n} - s \xi^{n-1}] d\xi
\]
and hence
\[
u^{+}(s,1;\theta) = \int_{0}^{\infty} e^{-\theta t} P(W_1(t) = 0 | W_1(0) = 0) \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[\xi^{n}] d\xi d\xi d\xi
\]
and
\[
u^{+}(0,1;\theta) = \int_{0}^{\infty} e^{-\theta t} P(W_1(t) = 0 | W_1(0) = 0) \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[\xi^{n}] d\xi d\xi d\xi
\]
We obtain
\[
u^{+}(0,1;\theta) = \int_{0}^{\infty} e^{-\theta t} P(W_1(t) = 0 | W_1(0) = 0) \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[\xi^{n}] d\xi d\xi d\xi
\]
and
\[
u^{+}(0,1;\theta) = \int_{0}^{\infty} e^{-\theta t} P(W_1(t) = 0 | W_1(0) = 0) \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[\xi^{n}] d\xi d\xi d\xi
\]
We leave it to the reader to continue here and merely notice that in steady state the actual waiting time of a low priority customer for the beginning of service is distributed as the initial busy period of an M/G/1 queue defined by \( \lambda_1 \) and \( B_1(t) \), if this queue starts under a workload distributed as \( W_1 + W_2 \). Instead of this actual we may wish to study the corresponding virtual waiting time. If \( W_1(t) \) is distributed as the high (low) priority workload in steady state - existence of a limiting distribution here assumed - then (3.6) yields the relation
\[
\frac{1}{1 - f^{+}(0,1;\theta)} \mathbb{E}[e^{-\theta \xi}] = \int_{0}^{\infty} e^{-\theta t} P(W_1(t) = 0 | W_1(0) = 0) \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[e^{-\theta \xi}] \int \mathbb{E}[\xi^{n}] d\xi d\xi d\xi
\]
and the desired distribution can now be obtained exactly 
as for the actual waiting time, if $\hat{W}_1 + \hat{W}_2$ is replaced 
by $\hat{W}_1 + \hat{W}_2$.

Evidently the phase method has proven to be an efficient 
 tool for the study of queuing systems.

REFERENCES

 North-Holland, Amsterdam, 1969.


in queuing processes 

[4] Hooke, J.A. A priority queue with low-priority 
 arrivals general 


 queue Journ. Appl. Prob. 9, 
 370 - 381, 1972.

[7] Sahin, I. Some stochastic systems with 
 secondary inputs. 
 Cage Western Univ., 1970.

[8] Schassberger, R.A. A note on the method of stages, 
 Research paper No. 53, Dept. of 
 Mathematics, University of Calgary 
 1968.

[9] Schassberger, R.A. On the waiting time in the 
 queuing system GI/G/1. 

[10] Schassberger, R.A. On the work load process in a 
 general preemptive resume priority 

 Wien.

[12] Whitt, W. The continuity of queues 