THE GENERAL DISTRIBUTIONAL LITTLE'S LAW AND ITS APPLICATIONS

Dimitris Bertsimas and Daisuke Nakazato

OR 247-91 March 1991

The general distributional Little's law and its applications

Dimitris Bertsimas *   Daisuke Nakazato †

March 1991

Abstract

We generalize the well-known Little's law $E[L] = \lambda E[W]$ to a distributional form: $L \stackrel{d}{=} N_a^*(W)$, where $N_a^*(t)$ is the number of renewals up to time t for the equilibrium arrival process, L is the number in the system (or in the queue) and W the time spent in the system (or in the queue) under FCFS. We provide two proofs of the result. In the process we generalize a well known theorem of Burke on the equality of pre-arrival and post-departure probabilities. By offering very simple proofs of several known as well as new results in a variety of queueing systems, we demonstrate that the distributional law has important algorithmic and structural applications and leads to a certain unification of queueing theory.

1 Introduction

One of the most fundamental results in queueing theory is that the number of customers in the system (or queue), denoted by L, and the waiting time W of a

*Dimitris Bertsimas, Sloan School of Management and Operations Research Center, MIT, Cambridge, Ma 02139.

†Daisuke Nakazato, Operations Research Center, MIT, Cambridge, Ma 02139.

customer in the system (or queue) obey Little's law ([15]) that under a wide variety service disciplines

$$E[L] = \lambda E[W],$$

and λ is the arrival rate. In a highly unnoticed paper, Haji and Newell [9] address the issue of relating the distributions of L and W. Although their result is, in our opinion, both fundamental and important, it has remained widely unnoticed for the last two decades even by top queueing theory researchers. As an example, Keilson and Servi [13] have recently proved a distributional form under FCFS for the special case of a Poisson arrival process. They proved that if $G_L(z) = E[z^L]$ and $\phi_W(s) = E[e^{-sW}]$ then

$$G_L(z) = \phi_W(\lambda - \lambda z).$$

In this paper, we offer two proofs of the relation of the distributions of L and W, when the arrival process is a general renewal process and the queueing discipline is FCFS. We also remark that the result holds even for more general nonrenewal arrival processes. Although we derived the distributional law independently as a generalization of Keilson and Servi [13], as we have already pointed out, the result is not new; we believe, however, that the paper makes a contribution for the following reasons:

1. The paper makes an important structural result of queueing theory widely known to the research community.

2. We offer two proofs of the result. The first proof, which is a simple probabilistic proof from first principles, is similar to the one in [9]. Our second proof, which is new, offers more insight on the relations of pre-arrival, post-departure, general time probabilities and the waiting time. Moreover, it is the natural matrix geometric generalization of the proof technique of Keilson and Servi [13].

3. We generalize a well known theorem of Burke on the equality of pre-arrival and post-departure probabilities for stochastic processes with unit jumbs, to more general processses. Although this generalization is of independent interest, it was the key for our second proof of the distributional law.

4. Most importantly, we examine several interesting applications of the distributional law and offer very simple proofs to a variety of known as well as new results. This is particularly important, since it shows that the distributional law is a quite powerful tool for the analysis of several models in queueing theory.

The paper is structured as follows. In section 2 we present the first simple probabilistic proof of the distributional law and describe certain systems in which the result is applicable. In section 3 we present our second proof for the case in which the interarrival distribution is Coxian. This section also includes the generalization of Burke's theorem. The last section contains some algorithmic and structural consequences of the distributional law for a wide variety of queueing systems.

2 The distributional law

Consider a general queueing system, whose arrival process is an arbitrary renewal process. Let $\alpha(s)$ be the Laplace transform of the interarrival distribution, with arrival rate $\lambda = -1/\dot{\alpha}(0)$.

Let $N_a(t)$ be the number of renewals up to time t for the ordinary renewal process (where the time of the first renewal has the same distribution as the interarrival time).

Let $N_a^*(t)$ be the number of renewals up to time t for the equilibrium renewal process (where the time of the first renewal is distributed as the forward recurrence time of the arrival process).

The distributional law can be stated as follows:

Theorem 1 *Let a given class C of customers have the following properties:*

1. *Arrivals of class C form a renewal process whose interarrival time has a transform distribution* $\alpha(s)$.

2. *All arriving customers enter the system (or the queue), and remain in the system (or the queue) until served, i.e. there is no blocking, balking or reneging.*

3. *The customers are served one at a time and leave the system (or the queue) in the order of arrival (FCFS).*

4. *New arriving class C customers do not affect the time in the system (or the queue) for previous class C customers.*

Then, in steady state, the waiting time W of the class C customers in the system (or the queue) and the number L of the class C customers in the system (or queue) are related in distribution by:

$$L \stackrel{d}{=} N_a^*(W). \tag{1}$$

In other words, the c.d.f. $F_W(t)$ *of W and the generating function* $G_L(z)$ *of L satisfies the following relation:*

$$G_L(z) = \int_0^\infty K(z,t)\, dF_W(t) \tag{2}$$

where the kernel is the generating function of the equilibrium renewal process. i.e.

$$K(z,t) = \sum_{n=0}^{\infty} z^n \Pr[N_a^*(t) = n]. \tag{3}$$

It is well known (see Cox[7]) that the Laplace transform of the renewal generating function $K(z,t)$ *is given by*

$$K^*(z,s) = \int_0^\infty e^{-st}\, K(z,t)\, dt = \frac{1}{s} - \lambda \frac{(1-z)(1-\alpha(s))}{s^2(1-z\alpha(s))}. \tag{4}$$

If the Laplace transform pdf $\phi_W(s)$ *of the waiting time is known, then*

$$G_L(z) = \frac{1}{2\pi\sqrt{-1}} \oint K^*(z,s)\phi_W(-s)\, ds \tag{5}$$

where for a fixed z the contour contains all singularities of $K^*(z,s)$ *but not* $\phi_W(-s)$.

Proof

We define τ to be a random observation epoch, i.e. an observer starts observing the system at a random time τ. Let τ_n the arrival time of the nth class C customer in the system (or queue) and W_n his the waiting time in the system (or queue). It is important to number the customers within the class C in the right order. That is, within the class C, the customer who is numbered 1 is the customer who arrived most recently, i.e. the customer at the end of the queue. The customer who has the highest ordinal number n is the customer getting served (or at the head of the queue). Therefore τ_n, W_n are ordered in the reverse time direction.

Let $T_1^* = \tau - \tau_1$, i.e. T_1^* is distributed as the forward recurrence time of the arrival renewal process, and $T_n = \tau_{n-1} - \tau_n$, $n > 1$ is the interarrival time for all $n > 0$.

The key observation for the proof is the following. When an observer coming to the system at a random moment τ sees at least n customers from class C, the nth most recently arrived customer among the class C is still waiting at that moment τ of the observation (see also figure 1), i.e.

$$L \geq n \text{ } if \text{ } and \text{ } only \text{ } if \text{ } W_n > \tau - \tau_n.$$

Note that we have used here assumptions 2 and 3. Therefore

$$\Pr[L \geq n] = \Pr[W_n > \tau - \tau_n].$$

Now, because of assumptions 3 and 4, W_n and $\tau - \tau_n = T_1^* + \sum_{i=2}^{n} T_i$ are independent. Indeed, every person arriving after time τ_n joins the queue *after* the nth customer and therefore each of these arrivals does not affect the waiting time W_n of that nth customer under assumptions 3 and 4. Also, in steady state $W_n \rightarrow W$. Hence we have

$$\Pr[L \geq n] = \int_0^\infty \Pr[T_1^* + \sum_{i=2}^{n} T_i < t] \, dF_W(t),$$

which essentially proves the theorem.

We next compute generating functions. As usual, $\Pr[L = n] = \Pr[L \geq n] - \Pr[L \geq n+1]$

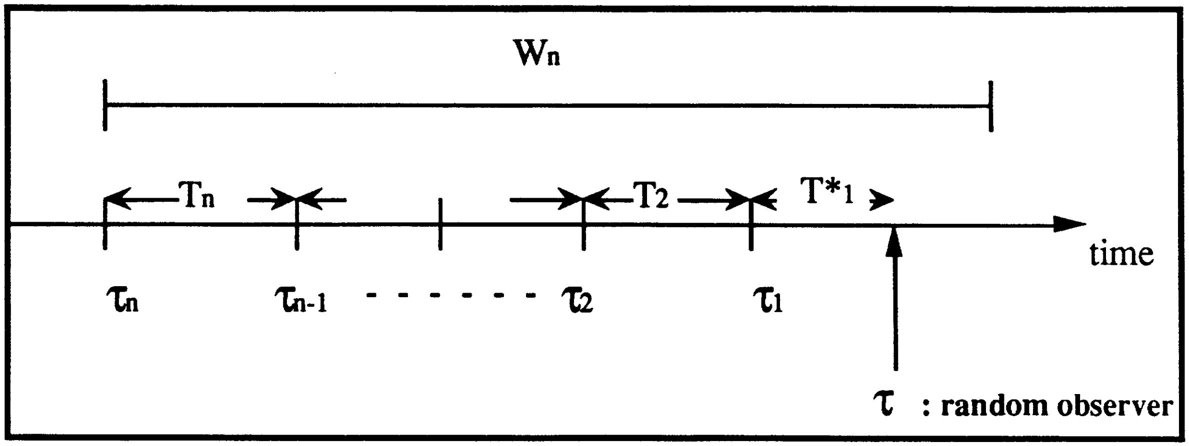

Figure 1: An illustration of Little's law.

and thus

$$G_L(z) = \sum_{n=0}^{\infty} z^n \Pr[L = n] = \int_0^{\infty} K(z,t)\, dF_W(t) \tag{6}$$

where

$$\begin{aligned} K(z,t) &= \Pr[T_1^* \geq t] + \sum_{n=1}^{\infty} z^n \left\{ \Pr[T_1^* + \sum_{i=2}^{n} T_i < t] - \Pr[T_1^* + \sum_{i=2}^{n+1} T_i < t] \right\} \\ &= \sum_{n=0}^{\infty} z^n \Pr[N_a^*(t) = n]. \end{aligned}$$

By taking the Laplace transform of $K(z,t)$, we have

$$\begin{aligned} \int_0^{\infty} e^{-st} K(z,t)\, dt &= \frac{1}{s}\left(1 - \lambda \frac{1-\alpha(s)}{s} + \sum_{n=1}^{\infty} z^n \lambda \frac{1-\alpha(s)}{s} \left(\alpha(s)^{n-1} - \alpha(s)^n\right)\right) \\ &= \frac{1}{s} - \lambda \frac{1-\alpha(s)}{s^2}\{1 - z\frac{1-\alpha(s)}{1-z\alpha(s)}\} \\ &= \frac{1}{s} - \lambda \frac{(1-z)(1-\alpha(s))}{s^2(1-z\alpha(s))}. \end{aligned}$$

Finally we use the inverse Laplace transform formula for the kernel,

$$K(z,t) = \frac{1}{2\pi\sqrt{-1}} \oint e^{st} K^*(z,s)\, ds$$

where the contour contains all singularities of $K^*(z,s)$ for a fixed z. We then have

$$G_L(z) = \frac{1}{2\pi\sqrt{-1}} \int_0^{\infty} \oint e^{st} K^*(z,s)\, ds\, dF_W(t).$$

Since $\phi_W(s) = \int_0^\infty e^{-st} dF_W(t)$, we get

$$G_L(z) = \frac{1}{2\pi\sqrt{-1}} \oint K^*(z,s)\phi_W(-s)\, ds.$$

Note that any one of the singularities of $K^*(z,s)$ need not coincide with a singularity of $\phi_W(-s)$, since we can always perturb z. □

The distributional form of Little's law (1) has the the following intuitive interpretation. In steady state, the time average number of class C customers has the same distribution as the number of class C arrivals, arriving according to the equilibrium renewal process, during the waiting time.

Important remarks:

1. Note that in order to derive (1), (2) and (3) only assumptions 2, 3 and 4 were used and not assumption 1. The renewal character of the arrival process (assumption 1) was only used to derive (4).

   As a result, the result also holds for general arrival point processes, not necessarily renewal. An example of a nonrenewal arrival process arises naturally in tandem queues, where the departure process of one queue is the arrival process of another queue. In Bertsimas and Nakazato [1] we study the characteristics of departure processes from a GI/G/1 queue and compute the kernel $K(z,t)$ appearing in the distributional law.

2. Note that if the arrival process is Poisson, then it can easily be proved that

   $$K(z,t) \;=\; e^{-\lambda(1-z)t}.$$

   Substituting into (1) we obtain Keilson and Servi's result [13]:

   $$G_L(z) \;=\; \phi_W(\lambda \;-\; \lambda z).$$

3. Let L^-, L^+ be the number in the system (or the queue) just before an arrival or just after a departure respectively for a system that satisfies the assumptions

of theorem 1. The number of customers left behind in the system (or in the queue) by a departing customer a, is exactly the number of customers that arrived during the time customer a spent in the system (or in the queue), because the queue discipline is FCFS. As a result,

$$L^+ \stackrel{d}{=} N_a(W)$$

and since the number of customers changes by one

$$L^+ \stackrel{d}{=} L^-.$$

As a result,

$$G_{L^+}(z) = \int_0^\infty K_o(z,t)\, dF_W(t) \tag{7}$$

where the kernel is the generating function of the ordinary renewal process, i.e.

$$K_o(z,t) = \sum_{n=0}^{\infty} z^n \Pr[N_a(t) = n]. \tag{8}$$

It is well known (see Cox[7]) that the Laplace transform of the renewal generating function $K_o(z,t)$ is given by

$$K_o^*(z,s) = \int_0^\infty e^{-st}\, K_o(z,t)\, dt = \frac{1-\alpha(s)}{s(1-z\alpha(s))}. \tag{9}$$

If, in addition, the arrival process satisfies the ASTA property (see Melamed and Whitt [16]), then

$$L^+ \stackrel{d}{=} L^- \stackrel{d}{=} L \stackrel{d}{=} N_a(W),$$

Note that the ASTA property is more general than PASTA and as a result, we again obtain Keilson and Servi's [13] result.

2.1 Systems for which the distributional law holds

We emphasize that the distributional law holds for a wide variety of settings; these include the time and the number *in the queue* of a heterogeneous service priority

GI/G/s system for each priority class. It does not hold, however, for the time and the number *in the system*, since assumption 3 is violated (overtaking can occur). It also holds for both the number in the system and the number in the queue of a GI/G/1 system and a GI/D/s system with priority classes. Moreover, it holds for the total sojourn time and the total number of a certain class in a queueing network without overtaking (such as G/G/1 tandem queues). Other applications include $GI/G/1$ with vacations and exhaustive service but only one priority class.

3 A generalization of a theorem of Burke and its application to the distributional law

In this section we will present another proof of the distributional Little's law for the case when the interarrival time is a general Coxian distribution, i.e. it belongs to the class of distributions with rational Laplace transform. This proof is a natural extension of Keilson and Servi [13] proof technique to general arrival processes with rational Laplace transforms.

Our strategy for proving theorem 1 is the following.

1. We introduce the Coxian distribution as a representation of a general distribution.

2. We observe that when the interarrival time is Coxian distributed, the arrival process is a special case of a phase renewal process. Using the uniformization technique, a phase renewal process is interpreted as an imbedded Markov Chain at Poisson transition epochs.

3. We generalize Burke's theorem by relating the post-departure and pre-transition probabilities. Since by Wolff's PASTA theorem, an observer at a Poisson transition epoch sees time averages, we are thus able to relate the post-departure and the general time probabilities.

4. We relate the post-departure probabilities to the waiting time distribution.

5. We finally combine the last two relations and thus prove theorem 1.

3.1 The Coxian distribution

The general Coxian class C_n was introduced by Cox [6]. In this section we will consider systems with Coxian arrival processes. We conceive of the arrival process as an arrival timing channel (ATC) consisting of M consecutive exponential stages with rates $\lambda_1, \lambda_2, ..., \lambda_M$ and with probabilities $p_1, p_2, ..., p_M = 1$ of entering the system after the completion of the 1st, 2nd, ... Mth stage. We remark that as soon as a customer in the ATC enters the system a new customer arrives at stage 1 of the ATC. The stage representation of the Coxian distribution is presented in figure 2. Note that this stage representation of the Coxian distribution is purely formal in the sense that the branching probabilities p_i can be negative and the rates λ_i can be complex numbers. The mixed generalized Erlang distribution is a Coxian distribution, where we assume that the probabilities p_i are nonnegative and the rates λ_i are reals.

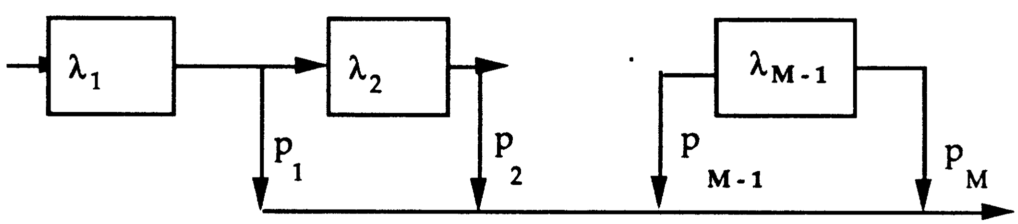

Figure 2: The Coxian class of distributions

Let $a_k(t)$ the pdf of the remaining interarrival time if the customer in the ATC is in stage $k = 1, \ldots, M$. Therefore, $a(t) = a_1(t)$ is the pdf of the interarrival time. For notational convienience we will drop the subscript for $k = 1$. Also $\frac{1}{\lambda}$ denotes the mean interarrival time.

Let $\alpha_k(s)$ be the Laplace transform of $a_k(t)$.

Let $a_i^j(t)$ be the probability to move from stage $i \le j$ of the ATC to stage j during the interval t without having any new arrival.

We will also use the notation:

$\vec{a_1}(t) = (a_1^1(t), \ldots, a_1^M(t))'$, $\vec{a_k}(t) = (0, \ldots, a_k^k(t), \ldots, a_k^M(t))'$.

$\vec{\alpha_k}(s)$ denotes the Laplace transforms of $\vec{a_k}(t)$.

$\vec{e_j} = (0, \ldots, 1, \ldots, 0)'$, $\vec{1} = (1, \ldots, 1, \ldots, 1)'$.

By introducing the following upper semidiagonal matrix A_0 and the dyadic matrix A_1:

$$A_0 = \begin{bmatrix} \lambda_1 & -(1-p_1)\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & -(1-p_2)\lambda_2 & \ddots & \vdots \\ \vdots & \ddots & & \ddots & \vdots \\ \vdots & & & \lambda_{M-1} & -(1-p_{M-1})\lambda_{M-1} \\ 0 & \cdots & \cdots & 0 & \lambda_M \end{bmatrix},$$

$$A_1 = \begin{bmatrix} -p_1\lambda_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ -\lambda_M & 0 & \cdots & 0 \end{bmatrix},$$

we can express compactly the transforms defined above as follows:

$$\vec{\alpha_k}'(s) = \vec{e}_k'(Is + A_0)^{-1},$$

$$\alpha_k(s) = -\vec{e}_k'(Is + A_0)^{-1}A_1\vec{e}_1 = \sum_{r=k}^{M} p_r\lambda_r\alpha_k^r(s) = \sum_{r=k}^{M} p_r\lambda_r \frac{\prod_{i=k}^{r-1}(1-p_i)\lambda_i}{\prod_{i=k}^{r}(s+\lambda_i)},$$

$$\alpha(s) = -\,\mathrm{trace}((Is + A_0)^{-1}A_1),$$

thus the interarrival pdf becomes

$$a(t) = -\,\mathrm{trace}(e^{-A_0 t}A_1).$$

3.2 Uniformization of the input process

We will consider queueing systems with input process forming a renewal process with interarrival time distribution being Coxian, which is a phase renewal process. We first observe that

$$(Is + A_0 + zA_1)^{-1} = (Is + A_0)^{-1} + \frac{z}{1 - z\alpha_1(s)} \begin{bmatrix} \alpha_1(s)\vec{\alpha_1}'(s) \\ \vdots \\ \alpha_M(s)\vec{\alpha_1}'(s) \end{bmatrix},$$

since for every pair of matrices C of full rank and D of rank 1, $(C + D)^{-1} = C^{-1} - \frac{C^{-1}DC^{-1}}{1+\mathrm{trace}(C^{-1}D)}$. By expressing this in real time we obtain

$$e^{-(A_0+zA_1)t} = \begin{bmatrix} a_1^1(t) & \cdots & a_1^M(t) \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_M^M(t) \end{bmatrix} + \sum_{n=1}^{\infty} z^n \begin{pmatrix} a_1(t) \\ \vdots \\ a_M(t) \end{pmatrix} * a_1^{(n-1)}(t) * \begin{pmatrix} a_1^1(t) & \cdots & a_1^M(t) \end{pmatrix}. \tag{10}$$

By interpreting this expression directly, it is clear that this is the phase renewal (generating) function of the arrival process.

We apply the uniformization technique (see Keilson[12]) to the phase renewal function:

$$e^{-(A_0+zA_1)t} = e^{-\nu t} e^{\nu t(I - \frac{1}{\nu}A_0 - \frac{z}{\nu}A_1)}$$

where we choose $\nu \geq \max_{i=1,\ldots,M} \lambda_i$.

Let $P_0 = I - \frac{1}{\nu}A_0$ and $P_1 = \frac{-1}{\nu}A_1$. Then

$$e^{-(A_0+zA_1)t} = \sum_{n=0}^{\infty} e^{-\nu t} \frac{(\nu t)^n}{n!} [P_0 + zP_1]^n.$$

The interpretation of this formula is that a transition occurs in a Poisson manner at rate ν with ATC phase transition probability P_0 and with the *effective* arrival probability P_1. Note that a transition is either an arrival or a shift to the next exponential stage. We will use this interpretation in the next two subsections.

3.3 The relation between the pre-transition and post-departure probabilities

Our goal in this subsection is to find a relation between the probabilities at pre-transition Poisson epochs of the arrival process and post-departure probabilities as a generalization of Burke's [5] theorem.

In the previous subsection we have introduced the matrices: $P_0 = I - \frac{1}{\nu}A_0$ with elements $P_{0,i,j}$ $\{i,j = 1,\ldots,M\}$ and $P_1 = \frac{-1}{\nu}A_1$ with elements $P_{1,i,j}$ $\{i,j = 1,\ldots,M\}$.

We now introduce the notation we will use:

L^+ = The number of customers in the system (or queue) immediately after a departure epoch.

L^- = The number of customers in the system (or queue) just before a *transition* epoch of the arrival process. A transition includes both arrivals in the system and shifts to the next exponential stage of the ATC. We emphasize that L^- in this and the following subsection is <u>not</u> the number of customers before an *arrival* epoch. The motivation for considering L^- is that, because of the uniformization, the epochs of transition of the arrival process (n,i), are Poisson distributed with rate ν. Therefore, by Wolff's [19] PASTA result, the pre-transition probabilities are the same with the time average probabilities. Therefore, if we are able to find the distribution of L^-, we will immediately find L.

R^+ = The ATC stage immediately after a departure epoch.

R^- = The ATC stage just before a transition epoch of the arrival process.

$\vec{P_n^+} = \{\Pr[L^+ = n \cap R^+ = i]\}_{i=1}^{i=M}$ and $\vec{P_n^-} = \{\Pr[L^- = n \cap R^- = i]\}_{i=1}^{i=M}$.

$\vec{P^+}(z) = \sum_{n=0}^{\infty} z^n \vec{P_n^+}$ and $\vec{P^-}(z) = \sum_{n=0}^{\infty} z^n \vec{P_n^-}$.

We observe the system from $t = 0$ to $t = T$. We also define

$u(n,i)$ = The number of upwards jumps during the period $(0,T)$, such that $L^- = n$ and $R^- = i$.

$u^0(n,i)$ = The number of shifts from the ATC stage i to $i+1$ during the period

$(0, T)$ such that $L^- = n$ and $R^- = i$.

$d(n, i)$ = The number of downwards jumps during the period $(0, T)$ such that $L^+ = n$ and $R^+ = i$.

U = The total number of transitions (upward jumps and shifts) of the arrival process during the period $(0, T)$.

D = The total number of departures (downward jumps) during the period $(0, T)$.

We now prove the following theorem on the relation of the pre-transition and post-departure probabilities.

Theorem 2 *Let $\vec{P^-}(z), \vec{P^+}(z)$ be the generating functions for the pre-transition and the post-departure probabilities as defined above. Then*

$$\vec{P^-}(z)[I - P_0 - zP_1] = \frac{\lambda(1-z)}{\nu}\vec{P^+}(z),$$

which is the same with

$$\vec{P^-}(z) = \lambda(1-z)\vec{P^+}(z)(A_0 + zA_1)^{-1}.$$

Proof

We follow a method used by Papaconstantinou and Bertsimas [18] and Hebuterne [10] to establish the relation between pre-arrival and post-departure probabilities in stochastic processes with random upward and downward jumps.

We first write down the flow balance equations; the left hand sides correspond to flow out and the right sides correspond to flow in.

$$\begin{aligned}
d(n-1,i) + u^0(n,i) + u(n,i) &= d(n,i) + u^0(n,i-1) & \{i > 1, n > 0\} \\
d(n-1,1) + u^0(n,1) + u(n,1) &= d(n,1) + \textstyle\sum_{i=1}^{M} u(n-1,i) & \{n > 0\} \\
u^0(0,i) + u(0,i) &= d(0,i) + u^0(0,i-1) & \{i > 1\} \\
u^0(0,1) + u(0,1) &= d(0,1). &
\end{aligned} \tag{11}$$

We divide all equations by U and we then take the limit as $T \to \infty$. Note that

$$\frac{D}{U} \to \frac{\lambda}{\nu}$$

$$\begin{aligned}
\frac{1}{D}d(n,i) &\rightarrow \Pr[L^+ = n \cap R^+ = i] \\
\frac{1}{U}u^0(n,i) &\rightarrow \Pr[L^- = n \cap R^- = i]P_{0,i,i+1} \\
\frac{1}{U}u(n,i) &\rightarrow \Pr[L^- = n \cap R^- = i]P_{1,i,1} \\
\frac{1}{U}(u^0(n,i) + u(n,i)) &\rightarrow \Pr[L^- = n \cap R^- = i](1 - P_{0,i,i}).
\end{aligned}$$

Then (11) becomes in matrix form

$$\begin{aligned}
\vec{P_n^-}(I - P_0) - \vec{P_{n-1}^-}P_1 &= \tfrac{\lambda}{\nu}\vec{P_n^+} - \tfrac{\lambda}{\nu}\vec{P_{n-1}^+} \quad \{n > 0\} \\
\vec{P_0^-}(I - P_0) &= \tfrac{\lambda}{\nu}\vec{P_0^+}
\end{aligned}$$

Computing the generating functions $\vec{P^-}(z), \vec{P^+}(z)$, we obtain

$$\vec{P^-}(z)[I - P_0 - zP_1] = \frac{\lambda(1-z)}{\nu}\vec{P^+}(z).\square$$

A critical observation, which actually motivated the uniformization technique, is that the epochs of transition of the arrival process (n, i), are Poisson distributed with rate ν. Therefore, by Wolff's [19] PASTA result, the pre-transition probabilities are the same with the time average probabilities. Therefore, we formally state

Proposition 1 *The time average number in the system L is equal to the pre-transition number in the system L^-*

$$L \stackrel{d}{=} L^-$$

Therefore, the time average generating function $G_L(z)$ is equal to

$$G_L(z) = \vec{P^-}(z)\vec{1}.$$

3.4 The relation of the waiting time and the post-departure probabilities

Let $\phi_W(s)$ be the transform pdf of the waiting time of a class C customer and $F_W(t)$ be the cdf of the waiting time. In this subsection we relate the waiting time and the post-departure probabilities.

Proposition 2 *For systems satisfying the assumptions of theorem 1 the post-departure probability generating function* $\vec{P^+}(z)$ *is represented as*

$$\vec{P^+}(z) = \vec{e_1}'\Phi_W(A_0 + zA_1),$$

where for any matrix D we symbolically define:

$$\Phi_W(D) \triangleq \int_0^\infty e^{-Dt} dF_W(t).$$

Proof

Because of assumptions 2, 3, 4 in theorem 1 the number of customers left behind by a customer departing from the system (or queue), is precisely the same with the number of customers that arrived during this customer's waiting time in the system (or queue). Therefore,

$$\Pr[L^+ = n \cap R^+ = i] = \int_0^\infty a^{(n)}(t) * a_1^i(t) dF_W(t).$$

Taking generating functions and using (10) we find

$$\vec{P^+}(z) = \mathrm{E}[\vec{e_1}'e^{-(A_0+zA_1)W}],$$

and thus the result holds. □

3.5 A matrix view of the distributional law

We now have all the necessary ingredients to give the second proof of the distributional law.

Theorem 3 *For systems satisfying the assumptions 2, 3 and 4 of theorem 1 and for Coxian interarrival times characterized by the matrices* A_0, A_1, *the generating function* $G_L(z)$ *and the cdf* $F_W(t)$ *are related by:*

$$G_L(z) = \int_0^\infty K(z,t)\, dF_W(t)$$

where

$$K(z,t) = \lambda(1-z)\vec{e_1}'e^{-(A_0+zA_1)t}(A_0 + zA_1)^{-1}\vec{1},$$

which leads to

$$G_L(z) = \lambda(1-z)\vec{e_1}'\Phi_W(A_0 + zA_1)(A_0 + zA_1)^{-1}\vec{1}. \quad (12)$$

The Laplace transform of $K(z,t)$ is

$$K^*(z,s) = \frac{1}{s} - \lambda\frac{(1-z)(1-\alpha(s))}{s^2(1-z\alpha(s))},$$

which from a well-known result from renewal theory, leads to

$$K(z,t) = \sum_{n=0}^{\infty} z^n \Pr[N_a^*(t) = n].$$

Proof

By proposition 1, $G_L(z) = \vec{P^-}(z)\vec{1}$ and by theorem 2,

$$G_L(z) = \lambda(1-z)\vec{P^+}(z)(A_0 + zA_1)^{-1}\vec{1}.$$

Then by proposition 2,

$$G_L(z) = \lambda(1-z)\vec{e_1}'\Phi_W(A_0 + zA_1)(A_0 + zA_1)^{-1}\vec{1}.$$

Note that this is a matrix geometric (see Neuts [17]) generalization of Keilson and Servi's [13] result $G_L(z) = \phi_W(\lambda - \lambda z)$.

Therefore,

$$G_L(z) = \int_0^{\infty} \lambda(1-z)\vec{e_1}'e^{-(A_0+zA_1)t}(A_0 + zA_1)^{-1}\vec{1}\,dF_W(t),$$

i.e.,

$$K(z,t) = \lambda(1-z)\vec{e_1}'e^{-(A_0+zA_1)t}(A_0 + zA_1)^{-1}\vec{1},$$

The transform of $K(z,t)$ is thus given by

$$\begin{aligned} K^*(z,s) &= \lambda(1-z)\vec{e_1}'(Is + A_0 + zA_1)^{-1}(A_0 + zA_1)^{-1}\vec{1} \\ &= \frac{\lambda(1-z)}{s}\vec{e_1}'\{(A_0 + zA_1)^{-1} - (Is + A_0 + zA_1)^{-1}\}\vec{1}. \end{aligned}$$

But

$$\begin{aligned}
\vec{e}_1{}'(Is + A_0 + zA_1)^{-1}\vec{1} &= \vec{e}_1{}'\{(Is + A_0)^{-1} - \frac{z}{1 - z\alpha(s)}(Is + A_0)^{-1}A_1(Is + A_0)^{-1}\}\vec{1} \\
&= \{1 + \frac{z}{1 - z\alpha(s)}\alpha(s)\}\vec{e}_1{}'(Is + A_0)^{-1}\vec{1} \\
&= \frac{1}{1 - z\alpha(s)}\frac{1 - \alpha(s)}{s} \\
&= \frac{1 - \alpha(s)}{s(1 - z\alpha(s))},
\end{aligned}$$

and similarly by taking $s \to 0$,

$$\vec{e}_1{}'(A_0 + zA_1)^{-1}\vec{1} = \frac{1}{\lambda(1 - z)}. \tag{13}$$

Therefore

$$K^*(z, s) = \frac{\lambda(1 - z)}{s}\{\frac{1}{\lambda(1 - z)} - \frac{1 - \alpha(s)}{s(1 - z\alpha(s))}\} = \frac{1}{s} - \lambda\frac{(1 - z)(1 - \alpha(s))}{s^2(1 - z\alpha(s))}.$$

Cox [7] contains an interpretation of this formula as the transform renewal generating function of the equilibrium renewal process. □

Theorem 3 is a special case of theorem 1 for the case of Coxian arrival process. Since Coxian distributions are dense in the space of distributions, one can prove the distributional law for arbitrary arrival processes by simply taking limits of Coxian distributions. The advantage of the proof technique that led to theorem 3 is that it also provides a closed form expression for the Kernel $K(z, t)$. This has some interesting applications as we show next.

4 Applications of the distributional law

The distributional Little's law is primarily a structural result relating the distributions of L and W. In this section we investigate some important structural and algorithmic consequences of the distributional law.

4.1 Relations among the second moments

A useful application of the distributional law is a relation of the first two moments of the queue length and the waiting time:

Theorem 4 *Under the conditions of theorem 1, Little's law for the first and second moments is*

$$\mathrm{E}[L] = \lambda\, \mathrm{E}[W]. \tag{14}$$

$$\mathrm{E}[L^2] = \lambda \left(\mathrm{E}[W] + 2\,\mathrm{E}\left[\int_0^W \mathrm{E}[N_a(\tau)]\, d\tau \right] \right), \tag{15}$$

where $\mathrm{E}[N_a(t)]$ *is the renewal function whose Laplace transform is given by*

$$\int_0^\infty e^{-st}\, \mathrm{E}[N_a(t)]\, dt = \frac{\alpha(s)}{s(1-\alpha(s))}.$$

Asymptotically,

$$\mathrm{E}[L^2] = \lambda^2 E[W^2] + \lambda c_a^2 E[W] - \frac{\lambda^3 E[A^3]}{3} + \frac{(c_a^2+1)^2}{2} + o(1), \tag{16}$$

$$E[L^+] = \lambda E[W] + \frac{c_a^2 - 1}{2} + o(1), \tag{17}$$

where c_a^2 *is the square coefficient of variation of the interarrival distribution and* $E[A^3]$ *is the third moment of the interarrival time.*

Proof

We first expand $K(z,t)$ as a Taylor series in terms of $\log(z)$:

$$K(z,t) = 1 + \lambda t \log(z) + \lambda \left(t + 2\int_0^t \mathrm{E}[N_a(\tau)]\, d\tau \right) \frac{(\log(z))^2}{2} + o((\log(z))^2).$$

To see this, we substitute $z = e^u$ in (4) first, expand the expression in terms of u and then perform the inverse Laplace transform term by term.

Now if we compare it with

$$G_L(z) = \sum_{r=0}^{\infty} \mathrm{E}[L^r] \frac{(\log(z))^r}{r!},$$

we obtain (14) and (15).

Let A be the interarrival time. Since

$$\frac{\alpha(s)}{s(1-\alpha(s))} = \frac{\lambda}{s^2} + \frac{c_a^2 - 1}{2s} - \frac{\lambda^2 E[A^3]}{6} + \frac{\lambda^3 E[A^2]^2}{4} + o(1),$$

inverting term by term, we obtain that

$$E[N_a(t)] = \lambda t + \frac{c_a^2 - 1}{2} + [\frac{\lambda^3 E[A^2]^2}{4} - \frac{\lambda^2 E[A^3]}{6}]\delta(t) + o(\delta(t)),$$

where $\delta(t)$ is Dirac's function. Substituting into (15) and performing the integration term by term we obtain (16). Starting with (7) and using the same technique as before we obtain (17). □

For the Poisson case ($c_a^2 = 1$) the asymptotic expression (16) is exact. Moreover, in this case the distributional law leads to an easy expression between the factorial moments. Since

$$E[z^L] = E[e^{-\lambda(1-z)W}]$$

successive differentiation leads to:

$$E[L(L-1)\ldots(L-r+1)] = \lambda^r E[W^r], \; r = 1, 2, \ldots.$$

Moreover, since $E[L]$, $E[L^+]$ and $E[W]$ are independent of the queue discipline, (17) leads to

$$E[L^+] = E[L] + \frac{c_a^2 - 1}{2} + o(1),$$

independent of the queue discipline.

The above expressions not only offer structural insight linking together fundamental properties in queues, but they lead, as we see next, to a closed form formulae for the expected waiting time for systems that have a distributional law for both the number in the system and in the queue.

4.2 Closed form approximations for systems with no overtaking

Consider a queueing system in which the distributional law holds for both the number in the system and the number in the queue. Examples in this category include the GI/G/1, GI/D/s queues with priorities, as well as the GI/G/1 with vacations. We will show in this subsection that the formulae we gave in the previous subsection lead to a closed form formula for the expected waiting time.

The GI/G/1 queue

Let L, Q is the number in the system and queue respectively and S and W is the time spent in the system and queue. Let $1/\lambda$, $E[X]$, c_a^2, c_x^2 be the means and the square coefficients of variation for the interarrival and service time distribution. Let $E[A^3]$ be the third moment of the interarrival distribution. We want to develop a formula for the expected waiting time as a function of just these parameters. We will develop a formula which is asymptotically correct for all ρ.

Theorem 5 *For the GI/G/1 queue the expected waiting time is asymptotically given*

$$E[W_{GI/G/1}] = \frac{\rho^2 c_x^2 + \rho^2 + \rho c_a^2 - \rho + o(1)}{2\lambda(1-\rho)}. \tag{18}$$

Proof

For the GI/G/1 queue the distributional law holds for both the number in the system and the number in the queue. Applying (14) and (16) we obtain

$$E[L] = \lambda E[S] \tag{19}$$

$$E[Q] = \lambda E[W] \tag{20}$$

$$E[L^2] = \lambda^2 E[S^2] + \lambda c_a^2 E[S] - \frac{\lambda^3 E[A^3]}{3} + \frac{(c_a^2+1)^2}{2} + o(1) \tag{21}$$

$$E[Q^2] = \lambda^2 E[W^2] + \lambda c_a^2 E[W] - \frac{\lambda^3 E[A^3]}{3} + \frac{(c_a^2+1)^2}{2} + o(1) \tag{22}$$

But $S = W + X$, where X is the service time and W, X are independent. Thus

$$E[S] = E[W] + E[X] \tag{23}$$

$$E[S^2] = E[W^2] + E[X^2] + 2E[W]E[X]. \tag{24}$$

Finally it straightforward to verify that the transforms of $G_L(z) = E[z^L]$, and $G_Q(z) = E[z^Q]$, are related as follows:

$$G_L(z) = (1-z)(1-\rho) + zG_Q(z),$$

which leads by successive differentiation to

$$E[L] = E[Q] + \rho \tag{25}$$

$$E[L^2] = 2E[Q] + E[Q^2] + \rho \tag{26}$$

Substituting (23), (24) and (26) to (21) we obtain

$$E[Q^2] + 2E[Q] + \rho = \lambda^2(E[W^2] + E[X^2] + 2E[W]E[X]) +$$
$$\lambda c_a^2(E[W] + E[X]) - \frac{\lambda^3 E[A^3]}{3} + \frac{(c_a^2+1)^2}{2} + o(1).$$

Substituting (20), (22) to the above equation and solving for $E[W]$ we obtain

$$E[W] = \frac{\lambda^2 E[X^2] + \rho c_a^2 - \rho + o(1)}{2\lambda(1-\rho)}$$

which leads to (18), since $\lambda^2 E[X^2] = \rho^2(c_x^2 + 1)$. □

Note that (18) is exactly the diffusion approximation for the expected waiting time in a GI/G/1 queue (see for example Heyman and Sobel [11], p. 483). We have just shown that just the distributional law leads to an asymptotically exact formula for the expected waiting time. Note that the formula is exact for $c_a^2 = 1$ (Pollaczek-Kintchine formula) and it also agrees with the heavy traffic ($\rho \to 1$) limit. Note that the formula addresses the dependence of the expected waiting time on just the first two moments of the interarrival and service time distributions. The $o(1)$ terms, which we have neglected, include the dependence of the expected waiting time on higher order momements.

The G/D/s queue

Another system which has a distributional law for both the number in the system and the number in the queue is the G/D/s.

Theorem 6 *For the GI/D/s queue the expected waiting time is asymptotically given*

$$E[W_{G/D/s}] = \frac{\rho c_a^2 + s(\rho^2 - \rho) + o(1)}{2\lambda(1-\rho)}, \tag{27}$$

where $\rho = \lambda E[X]/s$.

Proof

Since the service times are deterministic, every s customers are served by the same server. Therefore, each customer sees a $G^{(s)}/D/1$ queue, where $G^{(s)}$ is the s convolution of the interarrival distribution. As a result, the waiting time in the G/D/s queue is the same as in the $G^{(s)}/D/1$ queue. Moreover, (18) is valid (note that we need to use λ/s for the arrival rate and c_a^2/s for the coefficient of variation of the interarrival distribution). Therefore, the expected waiting time for the G/D/s queue is given in (27). □

4.3 Stochastic decomposition in vacation queues

In this subsection we will illustrate that the distributional law leads to the well known decomposition result for the expected waiting time in vacation queues. Although the decomposition result is more general we will only demonstrate it for the expected waiting time. The server of a GI/G/1 queue works as follows. When the system becomes empty, the server becomes inactive ("on vacation") for a duration V having Laplace transform $v(s)$. At the end of the vacation period another vacation period begins if the system is empty. Otherwise the system is again served exhaustively. It is assumed that V is independent of the arrival process. We offer a new simple proof of the decomposition result for the expected waiting time in vacation queues based on the distributional law.

Theorem 7 *(Doshi [8]) For the GI/G/1 with vacations V the expected waiting time is the sum of the expected waiting time of a GI/G/1 and the forward recurrence time of the vacation V.*

Proof

Let L_v, Q_v, W_v and S_v be the number of customers in the system, the number of customers in the queue and the time spent in the queue and in the system *with vacations* respectively. Let B be the number of customers in queue given that the server is on vacation. Let A be the number of customers in queue given that the server is not in vacation. Let V^* be the forward recurrence time of the vacation period.

Applying the distributional law to this system we will get (19)-(24) but for L_v, Q_V, W_v and S_v.

By conditioning on whether the server is on vacation we obtain

$$G_{Q_v}(z) = \rho G_A(z) + (1-\rho)G_B(z)$$

and

$$G_{L_v}(z) = \rho z G_A(z) + (1-\rho)G_B(z)$$

which lead to

$$G_{L_v}(z) = zG_{Q_v}(z) + (1-\rho)(1-z)G_B(z).$$

Differentiating twice we obtain

$$E[L_v] = E[Q_v] + \rho \tag{28}$$

$$E[L^2] = 2E[Q_v] + E[Q_v^2] + \rho - 2(1-\rho)E[B]. \tag{29}$$

Moreover, applying the Little's law for the number B and V^* we obtain

$$E[B] = \lambda E[V^*] \tag{30}$$

Solving the system of equations (19)-(24) and (28)-(30) we obtain

$$E[W_v] = E[V^*] + \frac{\rho^2 c_x^2 + \rho^2 + \rho c_a^2 - \rho + o(1)}{2\lambda(1-\rho)},$$

i.e. from (18)

$$E[W_v] = E[V^*] + E[W_{GI/G/1}]. \ \square$$

We close this subsection by emphasizing that the same method leads to similar expressions for every system that has a distributional law for both the number in the system and the number in the queue. For example a GI/G/1 queue with priorities can also be analyzed using the same techniques.

4.4 Structural implications

In this subsection we present another interesting consequence of the distributional law, which offers structural insight into the class of distributions that can arise in queueing systems.

Theorem 8 *If the waiting time distribution $F_W(t)$ is a mixture of exponential distributions, i.e.*

$$F_W(t) = 1 - \sum_u A_u e^{-x_u t} \tag{31}$$

(x_u can be a complex number), the queue length distribution is a mixture of geometric terms. Namely

$$\begin{aligned}
\Pr[L = n] &= \lambda \sum_u A_u \frac{(1-w_u)^2}{x_u w_u} w_u^n, & (32)\\
G_L(z) &= 1 - \lambda \sum_u A_u \frac{(1-z)(1-w_u)}{x_u(1-zw_u)} & (33)\\
G_{L^+}(z) &= \sum_u A_u \frac{1-w_u}{(1-zw_u)}. & (34)\\
w_u &= \alpha(x_u).
\end{aligned}$$

Proof

From (2) we obtain

$$\begin{aligned}
G_L(z) &= \int_0^\infty K(z,t) \left\{ \sum_u A_u x_u e^{-x_u t} \right\} dt \\
&= \sum_u A_u \int_0^\infty x_u e^{-x_u t} K(z,t)\, dt \\
&= \sum_u A_u x_u K^*(z, x_u).
\end{aligned}$$

Using (4) and then McLaurin expanding $G_L(z)$ in terms of z, we obtain (32). In a similar way we obtain (34). □

The previous theorem is applicable in a wide variety of queueing systems. In Bertsimas and Nakazato [2] we show that a general GI/G/s system with heterogeneous servers satisfies the exponentiality assumption of the theorem 8. Moreover, if the service time distributions have rational Laplace transform then the number of exponential terms in (31) is finite.

As an illustration of the usefullness of theorem 8, we apply it to find a closed form expression for the queue length distribution of the G/R/1 queue. Let $\beta(s) = \frac{\beta_N(s)}{\beta_D(s)}$ be the Laplace transform of the service time distribution, where $\beta_D(s)$, $\beta_N(s)$ are polynomials of degree m and less than m respectively. Let $\alpha(s)$ be the Laplace transform of the interarrival distribution. Using the Hilbert factorization method, one can derive the waiting time distribution for the GI/R/1 queue (see for example Bertsimas et. al. [3]) as follows:

$$\phi_W(s) = \frac{\beta_D(s)}{\beta_D(0)} \prod_{r=1}^{m} \frac{x_r}{x_r + s}$$

where x_r, $r = 1, \dots, m$ are the m roots of the equation

$$\alpha(z)\beta(-z) = 1, \quad Re(z) > 0.$$

Expanding $\phi_W(s)$ in partial fractions and inverting we find that

$$F_W(t) = 1 - \sum_{r=1}^{m} \frac{\beta_D(-x_r)}{\beta_D(0)} \prod_{i \neq r} \frac{x_i}{x_i - x_r} e^{-x_r t}.$$

Applying theorem 8 we find that the queue length distribution is given by

$$Pr\{Q = n\} = \lambda \sum_{r=1}^{m} \frac{\beta_D(-x_r)}{\beta_D(0)} \prod_{i \neq r} \frac{x_i}{x_i - x_r} \frac{(1 - w_r)^2}{x_r w_r} w_r^n, \tag{35}$$

where $w_r = \alpha(x_r)$.

4.5 The inverse problem

Equation (2) gives the generating function of the number in the system (or in the queue) as an integral transformation of the distribution of the time in the system (or in the queue). Therefore, once the waiting time is known we can easily find through (2) the queue length distribution. It is interesting, however, to find an inverse of this linear transformation in order to express W in terms of L. Our goal is to find a kernel $\bar{K}(z,t)$ so that

$$F_W(t) = \frac{1}{2\pi\sqrt{-1}} \oint \bar{K}(z,t) G_L(z)\, dz$$

where the contour contains all the singularities of $G_L(z)$ but none of $\bar{K}(z,t)$.

Theorem 9 *When the waiting time distribution $F_W(t)$ is a mixture of exponential distributions,*

$$F_W(t) = \frac{1}{2\pi\sqrt{-1}} \oint \bar{K}(z,t) G_L(z)\, dz$$

where

$$\bar{K}(z,t) = \frac{\alpha^{-1}(\frac{1}{z})}{\lambda(1-z)^2}(e^{-\alpha^{-1}(\frac{1}{z})t} - 1).$$

Proof

Assuming that the waiting time distribution is a mixture of exponential distributions, we obtain from (33)

$$\begin{aligned} G_L(z) &= 1 - \lambda \sum_u A_u \frac{(1-z)(1-\alpha(x_u))}{x_u(1-z\alpha(x_u))} \\ &= 1 - \lambda \sum_u A_u \frac{(1-\alpha(x_u))}{x_u\alpha(x_u)} + \lambda \sum_u A_u \frac{(1-\alpha(x_u))^2}{x_u\alpha(x_u)(1-z\alpha(x_u))}. \end{aligned}$$

Since both the lhs and the rhs must have the same singularity structure, $G_L(z)$ must be singular at $z = \frac{1}{\alpha(x_u)}$. Therefore from the last term of the rhs, we obtain

$$\underset{z=\frac{1}{\alpha(x_u)}}{\text{Residual}}\, G_L(z) = -\frac{\lambda(1-\alpha(x_u))^2}{x_u\alpha(x_u)^2} A_u.$$

Let z_0 be a singular point of $G_L(z)$, i.e. $x_u = \alpha^{-1}(\frac{1}{z_0})$ (assuming that there exist a unique x_u such that $\Re x_u > 0$ for each given singular point $|z_0| > 1$). From (31), we have

$$F_W(t) = -\sum_{z_0} \frac{\alpha^{-1}(\frac{1}{z_0})}{\lambda(1-z_0)^2}(1 - e^{-\alpha^{-1}(\frac{1}{z_0})t}) \left\{ \underset{z=z_0}{\text{Residual}}\, G_L(z) \right\}.$$

Expressing the last expression in terms of a Cauchy integral, we obtain

$$F_W(t) = \frac{1}{2\pi\sqrt{-1}} \oint -\frac{\alpha^{-1}(\frac{1}{z})}{\lambda(1-z)^2}(1 - e^{-\alpha^{-1}(\frac{1}{z})t}) G_L(z)\, dz.$$

Therefore,

$$\bar{K}(z,t) = -\frac{\alpha^{-1}(\frac{1}{z})}{\lambda(1-z)^2}(1 - e^{-\alpha^{-1}(\frac{1}{z})t}).\square$$

4.6 Algorithmic applications

In this section we use the distributional law to derive the distribution for the number in the system.

The R/G/1 and G/R/1 queues

In (35) we have derived the queue length distribution of the G/R/1 queue using the Laplace transform of the waiting time. We will now use the distributional law to find the queue length distribution of the R/G/1 queue.

Let $\alpha(s) = \frac{\alpha_N(s)}{\alpha_D(s)}$ be the Laplace transform of the interarrival distribution, where $\alpha_D(s)$, $\alpha_N(s)$ are polynomials of degree m and less than m respectively. Let $\beta(s)$ be the Laplace transform of the service time distribution. Using the Hilbert factorization method, one can derive the waiting time distribution for the R/G/1 queue (see Bertsimas et. al. [3]) as follows:

$$\phi_W(s) = \frac{\alpha_D(0)}{\alpha_D(-s)} \frac{(1-\rho)s}{\lambda(1-\alpha(-s)\beta(s))} \prod_{r=1}^{k-1} \frac{x_r + s}{x_r}$$

where x_r, $r = 1, \ldots, k-1$ are the $k-1$ roots of the equation

$$\alpha(z)\beta(-z) = 1, \quad Re(z) < 0.$$

Note that for $k = 1$ the product $\prod_{r=1}^{k-1} \frac{x_r+s}{x_r}$ is defined to be 1. In addition for $k = 1$ the formula reduces to the well known Pollaczek-Kintchine formula for the M/G/1 queue. Applying the distributional law (12) and diagonalizing the matrix $\Phi_W(A_0 + zA_1)$ we can find the queue length distribution as follows:

$$G_Q(z) = \lambda(1-z)\vec{e_1}'S(z) \begin{bmatrix} \frac{\phi_W(\theta_1(z))}{\theta_1(z)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\phi_W(\theta_k(z))}{\theta_k(z)} \end{bmatrix} S^{-1}(z)\vec{1}, \qquad (36)$$

where $S(z)$ is a matrix with columns the eigenvectors of $A_0 + zA_1$ and $\theta_i(z)$, $i = 1, \ldots, k$ are the eigenvalues of $A_0 + zA_1$, which are calculated from the equation:

$$z\alpha(\theta_i(z)) = 1.$$

The G/D/s queue

As we observed in subsection 4.2 the waiting time for the G/D/s is exactly the same as in a $G^{(s)}/D/1$ queue, where $G^{(s)}$ is the s convolution of the interarrival distribution. Therefore we can solve the $R/D/s$ queue using the results of the previous paragraph for the $R/D/1$, since the class R is closed under convolutions.

The GI/D/∞ queue

In this case L is the number in the system and W is the time spent in the system. Because of the deterministic service with mean $\frac{1}{\mu}$ it is clearly a system with no overtaking. Moreover, because of the presence of infinite number of servers there is no waiting and thus $f_W(t) = \delta(t - \frac{1}{\mu})$. From the distributional law therefore

$$G_L(z) = K(z, \frac{1}{\mu}).$$

If in addition the arrival time is Poisson, i.e. $K(z,t) = e^{-\lambda(1-z)t}$, then we obtain the well known result that the number in the system has a Poisson distribution with parameter λ/μ.

Systems with no overtaking

In subsection 4.2 we used the distributional lawa to find closed form approximations for the expected waiting time in systems for which the distributional law holds for both the number in the system and the number in the queue. We want to argue that for such systems the distributional law completely characterizes all the distributions of interest, i.e. just the knowledge of the distributional law has all the probabilistic information needed to solve for these distributions. Although the actual solution might need arguments from complex analysis, the distributional laws fully characterize such systems. This important idea was observed by Keilson and Servi [14] for systems with Poisson arrival process. We generalize it here for systems with arbitrary distributions.

Let L, Q be the number in the system and queue respectively and S and W is the time spent in the system and queue. From theorem 1

$$G_L(z) = \frac{1}{2\pi\sqrt{-1}} \oint K^*(z,s)\phi_S(-s)\,ds \tag{37}$$

$$G_Q(z) = \frac{1}{2\pi\sqrt{-1}} \oint K^*(z,s)\phi_W(-s)\,ds. \tag{38}$$

But if $\beta(s)$ is the transform of the service time distribution then

$$\phi_S(s) = \phi_W(s)\beta(s). \tag{39}$$

Finally, depending on the system being solved $G_L(z), G_Q(z)$ are related. For example for the $GI/G/1$

$$G_L(z) = (1-\rho)(1-z) + zG_Q(z). \tag{40}$$

Solving the system of equations (37), (38), (39) and (40) we can find an integral equation for the transform of the waiting time pdf

$$\frac{1}{2\pi\sqrt{-1}} \oint K^*(z,s)\phi_W(-s)(\beta(-s)-z)\,ds = (1-\rho)(1-z). \tag{41}$$

For the special case of the M/G/1 queue one can derive easily the Pollaczek-Khintchine formula. In order to solve (41) we need to use the calculus of residuals and regularity arguments from complex analysis. What is important is that just the knowledge of the distributional Little's law for systems with no overtaking in both the number in the queue and in the system is sufficient to fully characterize the queueing system.

Acknowledgements

We would like to thank Ward Whitt for his comments on an earlier version of the paper and for pointing out [9]. The research of the authors was partially supported by grants from the Leaders for Manufacturing program at MIT and from Draper Laboratory.

References

[1] Bertsimas D. and Nakazato D. (1990). "The departure process from a G/G/1 queue and its applications to the analysis of tandem queues", Operations Research Center technical report, MIT.

[2] Bertsimas D. and Nakazato D. (1990). "On the characteristic equation of general queueing systems", Operations Research Center technical report, MIT.

[3] Bertsimas D., Keilson J., Nakazato D. and Zhang H. (1990), "Transient and busy period analysis of the GI/G/1 queue as a Hilbert factorization problem", *Jour. Appl. Prob.*, to appear.

[4] Borovkov, A.A. (1976). *Stochastic Processes in Queueing Theory*, Springer-Verlag, New York.

[5] Burke, P.J. (1956). "The output of queueing systems", *Operations Research*, 6, 699-704.

[6] Cox, D.R. (1955). "A use of complex probabilities in the theory of stochastic processes", *Proc. Camb. Phil. Soc.*, 51, 313-319.

[7] Cox, D.R. (1962). *Renewal Theory*, Chapman and Hall, New York.

[8] Doshi B. (1985). "A note on Stochastic decomposition in a Gi/G/1 quue with vacations or set-up times", *Jour. Appl. Prob.*, 22, 419-428.

[9] Haji R. and Newell G. (1971). "A relation between stationary queue and waiting time distributions", *Jour. Appl. Prob.*, 8, 617-620.

[10] Hebuterne, G. (1988). "Relation between states observed by arriving and departing customers in bulk systems", *Stochastic Process and their Applications*, 27, 279-289.

[11] Heyman D. and Sobel M. (1982). *Stochastic models in Operations Research*, Vol. 1, New York.

[12] Keilson, J. (1965). *Green's function method in probability theory*, Hafner, New York.

[13] Keilson, J. and Servi, L. (1988). "A distributional form of Little's law", *Operations Research Letters*, Vol.7, No.5, 223-227.

[14] Keilson, J. and Servi, L. (1990). "The distributional form of Little's law and the Fuhrmann-Cooper decomposition", *Operations Research Letters*, Vol.9, No.4, 239-247.

[15] Little, J. (1961). "A proof of the theorem $L = \lambda W$", *Operations Research*, 9, 383-387.

[16] Melamed B. and Whitt W. (1990). "On arrivals that see time averages", *Operations Research*, Vol. 38, 1, 156-172.

[17] Neuts, M. (1981). *Matrix-geometric solutions in stochastic models; an algorithmic approach*, The John Hopkins University Press, Baltimore.

[18] Papaconstantinou, X. and Bertsimas, D. (1990) "Relation between the prearrival and postdeparture state probabilities and the FCFS waiting time distribution in the $E_k/G/s$ Queue", *Naval Research Logistics*, 37, 135-149.

[19] Wolff, R. (1982) "Poisson arrivals see time averages", *Operations Research*, 30, 223-231.