Another look at transient versions of Little's law, and $M/G/1$ preemptive Last-Come-First-Served queues

Brian H. Fralix
EURANDOM
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Germán Riaño
Strategic Operations Research Team
Kimberly-Clark, Latin American Operations
Bogotá, Colombia

November 20, 2008

Abstract

We take a new look at transient, or time-dependent Little laws for queueing systems. Through the use of Palm measures, we will show that previous laws (see [6]) can be generalized; furthermore, within this framework a new law can be derived as well, which gives higher-moment expressions for very general types of queueing systems; in particular, the laws hold for systems that allow customers to overtake one another. What's especially novel about our approach is the use of Palm measures that are induced by nonstationary point processes, as these measures are not commonly found in the queueing literature. This new higher-moment law is then used to provide closed-form expressions of all moments of the number of customers in the system in an $M/G/1$ preemptive-LCFS queue at a time $t > 0$, for any initial condition and for any of the more famous preemptive disciplines (i.e. preemptive-resume, and preemptive-repeat with and without resampling). The phrase "closed-form" is used here to stress that the moments can be expressed in terms of probabilities that consist of convolutions of busy periods and residual busy periods, and so moment-matching methods can be used to generate very simple approximations of these quantities, as in [3]. It is also worth noting that these results appear to be new for the $M/M/1$ queue as well (see [3], [4]), and so we use them to derive a nice structural form for all of the time-dependent moments of a regulated Brownian motion (see [1], [2]).

1 Introduction

Little's law is one of the most fundamental laws of queueing theory. For a queueing system in steady-state, it relates L, the expected number of customers present in the system, to λ, the rate at which customers arrive to the system, and W, the expected waiting time of a customer that arrives during steady state; more precisely, $L = \lambda W$. Numerous papers devoted to this law have appeared in the literature over the past forty years: a nice overview of what was discovered through 1991 can be found in Whitt [23].

In this paper we will focus on transient, or time-dependent versions of Little's law, and their applications. To the best of our knowledge, the first paper that specifically focused on these sort of laws for queueing systems was Bertsimas and Mourtzinou [6]. Their method of proof involved the use of sample-path arguments to compute the first moment of the number of customers present in a queueing system at time t (denoted $Q(t)$). They also use the same type of argument to establish a distributional relationship between $Q(t)$ and the waiting times of all customers that arrive in the system during the interval $(0, t]$, so long as customers depart from the system in the same order at which they arrived. They allow for arbitrary initial conditions, and they also consider multiclass queueing systems as well. What's important to notice is that the proofs of these results require that the necessary waiting time distributions can be expressed in terms of a limit (this is needed in order to "condition" on having an arrival at a fixed time), and that the mean measure of the arrival process is absolutely continuous with respect to Lebesgue measure. In particular, if N represents the point process of arrivals (made up of points $\{T_n\}_{n\geq 1}$), they assume that there exists a function h and a random process $\{W(t) : t \geq 0\}$ such that for each $t \geq 0$,

$$h(t) = \lim_{\delta t \to 0} \frac{E[N(t)] - E[N(t - \delta t)]}{\delta t}$$

and for each $t, \tau > 0$,

$$h(t)dtP(W(t) > \tau) := \sum_{n=1}^{\infty} P(t - dt < T_n \leq t)P(W_n > t | T_n = t)$$

where W_n represents the sojourn time of the n^{th} arrival to the system in $(0, \infty)$. From this definition, $P(W(t) > \tau)$ can intuitively be interpreted as the probability that a customer that arrives to the system at time t stays in the system for at least τ units of time.

We will begin by showing how Palm theory can be used to generate the laws given in [6] under less restrictive conditions, in that this approach no longer requires that the limits mentioned above exist, nor do we have to assume that the mean measure is absolutely continuous. Putting these laws into a Palm framework is nice, because it gives us a natural analogue of the known Palm interpretation for the classical versions of Little's law. Furthermore, our approach also leads to a new law that allows us to relate *any* moment of $Q(t)$ to the sojourn times of customers that interact with the system in $[0, t]$. What's especially interesting about this law is that it is very general: it even holds for queueing systems that allow customers to overtake one another.

Typically the type of Palm measure that is found in the queueing literature is the one that is induced by a stationary point process. These measures are often used to analyze systems from the perspective of an arriving or departing customer that interacts with the system while it is in equilibrium. Introductions to the theory can be found in many places: see, for instance, Baccelli and Brémaud [5] and Serfozo [22]. We will instead use a family of Palm measures that are induced by point processes that do not necessarily have to be stationary, and the reader will see that their use will allow us to "condition" on an arrival occurring at a fixed time t in the appropriate way. These were first introduced in Ryll-Nardzewski [21], and a nice discussion on these measures can be found in the book of Kallenberg [14]. Usage of these measures is rare in the queueing literature, which adds to the novelty of the approach that's featured in this paper.

The application we present is as follows: in the papers of Abate and Whitt [3, 4] they were interested in how the moments of $Q(t)$ behave as a function of t, where $\{Q(t); t \geq 0\}$ corresponds to an $M/M/1$ queue with arrival rate λ and service rate μ. One of the main results of [3] (Theorem 3.2) showed that

$$E[(Q(t))_n | Q(0) = 0] = n!(\lambda E[\tau])^n P\left(\sum_{j=1}^{n} R_{\tau,j} \leq t\right) \tag{1}$$

where $\{R_{\tau,k}\}_{k\geq 1}$ is an i.i.d. sequence of residual busy periods, and for a given $x \in \mathbb{R}$ and an integer $n \geq 1$, $(x)_n = x(x-1)\cdots(x-n+1)$. The proof of this result involved using the fact that, when $Q(0) = 0$, $Q(t) \stackrel{d}{=} M(t)$ (here $\stackrel{d}{=}$ denotes equality in distribution), where $M(t)$ represents the maximum value over $[0, t]$ of a birth-death process that moves along the integers, with a birth rate λ and a death rate μ. Later, in [4] transform techniques were used to derive a decomposition of the queue-length at an exponential time, and this is then used to study the behavior of $E[Q(t)^n | Q(0) = k]$ for any $n, k \geq 1$: however, even though they are able to generate equations that give insight into how these moments behave (see Theorems 8.4 and 8.5 of [4]), they do not give an expression for this quantity that's as clear as, or is analogous to, identity (1). This approach does give a nice expression when $n = 1$, for arbitrary k, but it is not clear how it can immediately be used to compute the second and higher moments, for such k. We should also point out that the time-dependent moments of regulated Brownian motion were also studied in another series of papers by Abate and Whitt (see [1], [2]), which do not rely on the results found for the $M/M/1$ queue. Similarly, in these papers it is shown that the moments are much more difficult to compute when the process does not start at the origin, and only the first two moments are given for any initial condition; moreover, their derived form of the second moment does not immediately allow one to guess what the higher moments should look like.

We will show how to use our new transient versions of Little's law to very quickly derive the time-dependent moments of an $M/GI/1$ preemptive-LCFS (Last Come First Served) queue, for any initial condition. Our use of the term preemptive-LCFS queue will refer to systems that operate under either the preemptive-resume or preemptive-repeat disciplines. It is also worth observing that our results will also hold for queueing systems that "mix" both the preemptive-resume and repeat disciplines, i.e. it also holds, for instance, if at each time a given customer begins to receive service, it either continues where it left off (preemptive-resume), restarts from where it began the last time (preemptive-repeat without resampling), or restarts with a new amount of work (preemptive-repeat with resampling), where the choice of preemption used is governed by another random element. The expressions we find are as pleasing as (1), in that they are in terms of probabilities that can be approximated with moment-matching techniques, in the same way as (1) was approximated within section 4 of [3].

It is interesting to note that for $Q(0) = 0$, the factorial moments of $Q(t)$ were implicity computed for the $M/G/1$ preemptive-resume LCFS queue as well in the work of Kella et al. [17], who were interested in various time-dependent properties of symmetric $M/G/1$ queues (see Kelly [15] for a definition). There they were able to compute the distribution of $Q(\tau(q))$, where $\tau(q)$ is an exponential random variable with rate q, by making use of the fact that when $Q(0) = 0$, the distribution of $Q(t)$ can be expressed in terms of a Lévy process. In particular,

$$Q(t) \stackrel{d}{=} \sharp \left\{ s \in [0,t] : X(s-) = \inf_{r \in [s,t]} X(r) \right\}$$

where $\{X(t); t \geq 0\}$ represents the "net-input" process, i.e.

$$X(t) = \sum_{k=1}^{N(t)} S_k - t$$

and $\sharp\{s : S(s)\}$ denotes the number of s values for which the statement $S(s)$ is true. Formula (1) then quickly follows from inverting the transform of $Q(\tau(q))$. In a sequel to this paper [12], we will show how a transient analogue of results found in the classical ASTA (Arrivals See Time Averages) literature can be used to derive the distributions of $Q(\tau(q))$ for queueing models that are more general than the ones considered here.

We begin in Section 2 by setting up the mathematical framework in which we will work throughout this study. The derivation of our transient Little laws will be given in Section 3. In Section 4 we will show how these results can be used to gain additional insight into the time-dependent behavior of the moments of the $M/G/1$ preemptive-LCFS queue, and we will conclude in Section 5 by demonstrating how the time-dependent moments of the $M/M/1$ queue can be used to derive all of the corresponding moments for a regulated Brownian motion.

2 Palm Measures

Suppose $N := \{N(t); t \geq 0\}$ is a point process on $(0, \infty)$, whose points consist of the arrival times of customers to a given queueing system. We identify these points with the sequence $\{T_n; n \geq 1\}$, where T_k denotes the arrival time of the k^{th} customer to enter the system in the interval $(0,\infty)$. Associated with the k^{th} arrival is its waiting time W_k, and these waiting times generate a real-valued stochastic process $\{W(s); s \in \mathbb{R}_+\}$, where $W(s)$ represents the waiting time of the last customer to arrive at or before time s: we assume $W(s) = 0$ if no customers have arrived in $(0, s]$. Finally, let μ denote the mean measure of N, i.e. $\mu(A) = E[N(A)]$, which we will assume is σ-finite, in that $\mu(K) < \infty$ for all compact sets K. It is further assumed that throughout this paper, all of our processes reside on the space $(\Omega, \mathcal{F}, P)$, where Ω is a complete separable metric space, $\mathcal{F}$ are its associated Borel sets, and P is an arbitrary probability measure that determines the laws of all processes on the space. Such assumptions should not be considered to be too restrictive, due to the fact that many interesting processes associated with queueing networks reside on the space $D[0,\infty)$ that consists of right-continuous functions with left-hand-limits, and it is well-known that this space can be equipped with a metric that satisfies such properties.

To show how some of our transient laws simplify to their well-known stationary variants, we will also consider stationary versions of the processes given above, which will actually be defined on the entire real line. We will refrain from giving explicit definitions of all our stationary processes, as their proper definition can easily be inferred from our current setting. Rather, to signify that a process is stationary we will merely place a tilde over each random element, e.g. $\tilde{N}$, $\tilde{Q}(t)$, $\tilde{W}(s)$, etc..

Under these assumptions, we know that there exists a μ-a.e. unique collection of Palm measures $\{P_s\}_{s\in\mathbb{R}_+}$ induced by N, that satisfy the following: for any Borel set $B \subset \mathbb{R}$ and $A \in \mathcal{F}$,

$$E[N(B)\mathbf{1}_A] = \int_B P_t(A)\mu(dt)$$

where $\mathbf{1}_A$ is shorthand for the indicator function $\mathbf{1}_A(\omega)$ with $\omega \in \Omega$, which is 1 if $\omega \in A$, and 0 otherwise.

A major consequence of this definition is what is known as the Campbell-Mecke formula, which relates the Palm distributions to expectations of stochastic integrals with respect to a point process.

Theorem 2.1 *For a given stochastic process $\{X(t); t \geq 0\}$, the following equality holds:*

$$E\int_{\mathbb{R}} X(s)N(ds) = \int_{\mathbb{R}} E_s[X(s)]\mu(ds).$$

The proof of this theorem follows by applying an extension argument to the local definition of the Palm kernel that we have provided here, and is well-known in the literature.

The reader should note that the more classical (from a queueing perspective) definition of a Palm probability follows from this definition, when we further assume the existence of a measurable flow $\{\theta_t\}_{t\in\mathbb{R}}$ on our underlying space such that all processes of interest are adapted to the flow, and that P is θ_t-invariant, i.e. $P(\theta_t A) = P(A)$ for any $A \in \mathcal{F}$. By measurable flow, we mean that for each $\theta_t : \Omega \to \Omega$ (i) θ_0 is the identity mapping, (ii) θ_t is a (jointly measurable) bijection, and (iii) $\theta_{s+t} = \theta_s\theta_t$ for any $s, t \in \mathbb{R}$. Readers that are uncomfortable with the notion of such a flow can think of it as a stationary process that contains all the information about a process we may be interested in, so that all stationary processes of interest (i.e. queue-length and workload processes) are nice (i.e. shift-invariant) functionals of the flow. As a matter of fact, when the underlying probability space is $D[0,\infty)$, the standard shift operator on this space plays the role of the measurable flow.

Under these assumptions, the locally defined Palm measures are related to the classical Palm measure in the following way: $P_t(\theta_t A) = P_0(A)$. This will also be used at various points throughout the paper.

3 Little Laws

Throughout this section we will assume that $Q(0) = 0$, but it is not difficult to extend the formulas given below to the case when $Q(0) = n$, for any $n \geq 1$. Indeed, we will do this when we compute the time-dependent moments of the $M/G/1$ preemptive-LCFS queue in Section 4.

Our first result is a generalization of Theorem 1 in Section 3 of [6]. It is used in the Ph.D. thesis of Riaño [19], and can also be found for the case of Poisson arrivals in Rolski [20].

Theorem 3.1 *The first moment of $Q(t)$ satisfies the following equality:*

$$E[Q(t)] = \int_0^t P_s(W(s) > t - s)\mu(ds)$$

Proof As is true in [20], the proof of this result immediately follows from applying the Campbell-Mecke formula:

$$\begin{aligned} E[Q(t)] &= E\left[\int_0^t \mathbf{1}(W(s) > t - s)N(ds)\right] \\ &= \int_0^t P_s(W(s) > t - s)\mu(ds). \end{aligned}$$

This completes the proof. $\diamondsuit$

Remark It is interesting to notice that the well-known version of Little's law immediately follows from this result. If we assume that $\tilde{Q} := \{\tilde{Q}(t); t \in \mathbb{R}\}$ is stationary, then

$$\begin{aligned} E[\tilde{Q}(0)] &= \int_{-\infty}^0 \tilde{P}_s(\tilde{W}(s) > -s)\lambda ds \\ &= \lambda \int_{-\infty}^0 \tilde{P}_0(\tilde{W}(0) > -s) ds \\ &= \lambda \tilde{E}_0[\tilde{W}(0)]. \end{aligned}$$

$\square$

For queueing systems that satisfy the following assumptions, it has been shown that even stronger relationships hold between the steady-state number of customers in the system and the steady-state sojourn time. These assumptions are also given in Theorem 1 of [7].

Assumptions 3.1 *All arriving customers enter the system one-at-a-time, and remain in the system until their service requirements are satisfied.*

Assumptions 3.2 *The customers leave the system in the same order in which they arrived, i.e. the system is overtake-free.*

Assumptions 3.3 *The sojourn time of a customer is independent of the arrival times and service requirements of all customers that arrive during its sojourn time.*

Our next result is a generalization of Theorem 6 in [6].

Theorem 3.2 *Suppose that a queueing system satisfies assumptions (3.1), (3.2), and (3.3). Then the generating function of $Q(t)$ is as follows:*

$$E[z^{Q(t)}] = 1 + (z - 1)\int_0^t P_s(W(s) > t - s) E_s[z^{N(s,t]}]\mu(ds)$$

Proof Proving this involves applying the Campbell-Mecke formula to

$$\mathbf{1}(Q(t) \geq n) = \int_0^t \mathbf{1}(W(s) > t-s, N(s,t] = n-1)N(ds).$$

In other words,

$$\begin{aligned} P(Q(t) \geq n) &= \int_0^t P_s(W(s) > t-s, N(s,t] = n-1)\mu(ds) \\ &= \int_0^t P_s(W(s) > t-s)P_s(N(s,t] = n-1)\mu(ds). \end{aligned}$$

The generating function of $Q(t)$ can now be obtained after some simple algebra has been performed. $\diamondsuit$

Remark Again, it is easy to see that this result can be related to the steady-state distributional version of Little's law. Notice that if the arrival process is a renewal process, then

$$\begin{aligned} E[z^{\tilde{Q}(0)}] &= 1 + (z-1)\int_{-\infty}^0 \tilde{P}_s(\tilde{W}(s) > -s)\tilde{E}_s[z^{\tilde{N}(s,0]}]\lambda ds \\ &= 1 + (z-1)\int_0^\infty \tilde{P}_0(\tilde{W}(0) > s)\tilde{E}_0[z^{\tilde{N}(0,s]}]\lambda ds. \end{aligned}$$

The form of this result is different from the standard $\tilde{Q}(0) \stackrel{d}{=} \tilde{N}_e(0, \tilde{W})$ representation of this law (see [7] and [13]: also, see [16] for the Poisson arrival case), where $\tilde{W}$ is the stationary waiting time and $\tilde{N}_e$ is the equilibrium version of the renewal process. However, it is equivalent, and moreover, the appearance of the $z-1$ term allows for very simple calculations of all factorial moments $E[(\tilde{Q}(0))_n]$, where $(x)_n = x(x-1)\cdots(x-n+1)$:

$$E[(\tilde{Q}(0))_n] = n\lambda \int_0^\infty \tilde{P}_0(\tilde{W}(0) > s)]\tilde{E}_0[(\tilde{N}(0,s])_{n-1}]ds.$$

$\square$

It is theoretically interesting that the following alternative transient distributional law can also be derived, when P_s^* is the Palm measure induced by the departure process of our overtake-free system (we still assume that we have Poisson arrivals). Let $\{V(t); t \geq 0\}$ denote a stochastic process, where $V(t)$ represents the sojourn time of the first customer to depart at or after time t. This result is a transient analogue of the main result of Keilson and Servi [16].

Theorem 3.3 *For $0 < z < 1$, we find that*

$$E_t^*[z^{X(t)}] = E_t^*[z^{N(t-V(t),t]}]. \tag{2}$$

Furthermore,

$$E_t^*[X(t)] = E_t^*[N(t-V(t),t]] = \lambda E_t^*[V(t)]. \tag{3}$$

Proof The proof of this statement is simple: it merely follows from the fact that since the system is overtake-free, $X(t) = N(t-V(t),t]$ if we have a departure at time t (which occurs w.p. 1 under P_t^*). The rest then follows from differentiation. $\diamondsuit$

Notice that equation (3) is in the classical form of Little's law, even though the expected values are with respect to time-dependent Palm probabilities. If we generalize our setting by assuming that N is stationary, and that the number of arrivals observed by a customer that departs at time t is independent under P_t^* of the sojourn time of that customer, we end up with this form as well. Eventually, for large t, $E_t^*[X(t)]$ is approximately $E[X(t)]$, which gives the classical version of Little's law.

At this point we will begin to discuss a result that not only does not appear to be previously known in any sense, but also does not have any type of stationary interpretation. In particular, we will show that, regardless of the service discipline invoked by the queueing system, there is still a relationship between all moments of the number of customers in the system, and their waiting times. To do this we will have to briefly introduce a collection of multi-indexed Palm measures. These are discussed in [14], and a queueing application can be found in [8, 9], where they use higher-order (reduced) Palm measures to derive approximations for various performance measures associated with queues that are experiencing light-traffic.

We will now give a very rough sketch as to how these measures are derived; the details can be found in [14]. Based on our assumptions we know that there exists a μ_2-a.e. unique probability kernel $\{P_{s_1,s_2}\}_{s_1,s_2\in\mathbb{R}}$ such that, for $A \in \mathcal{F}$ and B_1, $B_2 \in \mathcal{B}(\mathbb{R})$,

$$E[N(B_1)N(B_2)\mathbf{1}_A] = \int_{B_1}\int_{B_2} P_{s_1,s_2}(A)\mu_{s_1}(ds_2)\mu(ds_1).$$

where $\mu(B_1 \times B_2) = E[N(B_1)N(B_2)] = \int_{B_1} \mu_{s_1}(B_2)\mu(ds_1)$. Such a construction can also be found in [14]; furthermore, it is known that we can also derive measures $P_{s_1,s_2,\ldots,s_n}$ for any $n \geq 1$, and we can still interpret $P_{s_1,s_2,\ldots,s_n}(A)$ as the probability of A, given N has points at $s_1, s_2, \ldots, s_n$. Moreover, from Lemma 11.2 in Kallenberg (see also Proposition 2.4 in [8] for the reduced case), we also know that these Palm measures are consistent under iterations, in that the Palm measure P_{s_2} induced by N, with respect to the probability measure P_{s_1} is the same as P_{s_1,s_2}. This result can be used to prove the following:

Theorem 3.4 *The factorial moments of $Q(t)$ satisfy the following relationship: for each $n \geq 1$,*

$$\begin{aligned} E[(Q(t))_n] \quad &= \quad \int_0^t\int_0^t \cdots \int_0^t P_{s_1,s_2,\ldots,s_n}(W(s_1) > t - s_1, W(s_2) > t - s_2, \ldots, W(s_n) > t - s_n) \\ &\quad \mu_{s_1,s_2,\ldots,s_{n-1}}(ds_n) \cdots \mu_{s_1}(ds_2)\mu_{ds_1} \end{aligned} \tag{4}$$

where $\mu_{s_1,\ldots,s_k}(A) = E_{s_1,\ldots,s_k}[N(A - \{s_1, \ldots, s_k\})]$.

Proof We will provide the details for the proof when $n = 2$; it will be obvious to the reader that the same argument follows for any arbitrary n. Notice that

$$\begin{aligned} &E[Q(t)(Q(t) - 1)] = E[\int_0^t (Q(t) - 1)\mathbf{1}(W(s) > t - s)N(ds)] \\ = \quad &\int_0^t E_s[\mathbf{1}(W(s) > t - s)(Q(t) - 1)]\mu(ds) \\ = \quad &\int_0^t E_s\left[\mathbf{1}(W(s) > t - s)\left[\int_0^t \mathbf{1}(W(u) > t - u)N(du) - \mathbf{1}(W(s) > t - s) - \mathbf{1}(W(s) \leq t - s)\right]\right]\mu(ds) \\ = \quad &\int_0^t E_s\left[\mathbf{1}(W(s) > t - s)\left[\int_{0,u\neq s}^t \mathbf{1}(W(u) > t - u) - \mathbf{1}(W(s) \leq t - s)\right]\right]\mu(ds) \\ = \quad &\int_0^t \int_{(0,s)\cup(s,t]} P_{u,s}(W(u) > t - u, W(s) > t - s)\mu_s(du)\mu(ds). \end{aligned}$$

$\diamondsuit$

Therefore, we see that the n^{th} factorial moment of $Q(t)$ can be expressed in terms of the joint distribution of the waiting times of n customers that arrive at times $s_1, s_2, \ldots, s_n \in (0, t]$. Unfortunately, applying this result is typically a very difficult task, mainly because the mean factorial measures found in the integral typically do not have a nice form.

For queues with Poisson arrivals, however, it is actually possible to simplify this relationship. If N is a stationary Poisson process with rate $\lambda > 0$, then it's known that

$$\mu_{s_1,\ldots,s_{n-1}}(ds_n)\cdots\mu(ds_1) = \lambda^n ds_n \cdots ds_1 \tag{5}$$

which gives us the following corollary.

Corollary 3.1 *If N is a stationary Poisson process with rate $\lambda > 0$, then*

$$\begin{aligned} E[(Q(t))_n] &= n!\lambda^n \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} P_{s_1,s_2,\ldots,s_n}(W(s_1) > t - s_1, W(s_2) > t - s_2, \ldots, W(s_n) > t - s_n) \\ &\times \; ds_n \ldots ds_2 ds_1. \end{aligned} \tag{6}$$

Proof The proof of this result involves applying (5) to (4), along with the fact that $P_{s_1,\ldots,s_n}(W(s_1) > t - s_1, \ldots, W(s_n) > t - s_n)$ is symmetric with respect to $(s_1, \ldots, s_n)$. $\diamondsuit$

Remark Clearly, if we also assume that our queueing system is stationary and satisfies Assumptions (3.1), (3.2), and (3.3), we find that

$$\begin{aligned} E[(\tilde{Q}(0))_n] &= n!\lambda^n \int_{-\infty}^0 \cdots \int_{-\infty}^{s_{n-1}} \tilde{P}_{s_1,s_2,\ldots,s_n}(\tilde{W}(s_n) > -s_n) ds_n \ldots ds_1 \\ &= n!\lambda^n \int_{-\infty}^0 \cdots \int_{-\infty}^{s_{n-1}} \tilde{P}_{s_n}(\tilde{W}(s_n) > t - s_n) ds_n \ldots ds_1 \\ &= n!\lambda^n \int_{-\infty}^t \cdots \int_{-\infty}^{s_{n-1}} \tilde{P}_0(\tilde{W}(0) > -s_n) ds_n \ldots ds_1 \\ &= n!\lambda^n \tilde{E}_0[\tilde{W}(0)] \int_{-\infty}^0 \cdots \int_{-\infty}^{s_{n-2}} \tilde{P}_0(R_{\tilde{W}(0)} > -s_{n-1}) ds_{n-1} \ldots ds_1 \\ &= \ldots = n!\lambda^n \prod_{k=0}^{n-1} E[R_{\tilde{W}(0),k}] \end{aligned}$$

where $R_{0,\tilde{W}(0)} = \tilde{W}(0)$, $R_{1,\tilde{W}(0)}$ is the residual version of $\tilde{W}(0)$, i.e. for any $t \geq 0$,

$$P(R_{1,\tilde{W}(0)} \leq t) = \frac{1}{E[\tilde{W}(0)]} \int_0^t P(\tilde{W}(0) > s) ds$$

and for any $n \geq 1$ $R_{n+1,\tilde{W}(0)} = R_{1,R_{n,\tilde{W}(0)}}$. However, it is well known that

$$\prod_{k=0}^{n-1} E[R_{k,\tilde{W}(0)}] = \frac{E[\tilde{W}(0)^n]}{n!}$$

and so we conclude that $E[(\tilde{Q}(0))_n] = \lambda^n \tilde{E}_0[\tilde{W}(0)^n]$, which is of course known, and can also be computed from the distributional Little's law. $\square$

Remark Now suppose we consider the case when $Q(0) = 0$, and the sojourn time of each customer that enters the system is its service time. If all services are independent of one another, then from the proof of equation (6) (without making use of symmetric properties of the integrand) we see that

$$\begin{aligned} E[(Q(t))_n] &= \lambda^n \int_0^t \int_0^t \cdots \int_0^t P_{s_1,s_2,\ldots,s_n}(W(s_1) > t - s_1, W(s_2) > t - s_2, \ldots, W(s_n) > t - s_n) ds_n \ldots ds_2 ds_1 \\ &= (\lambda E[B] P(R_{1,B} \leq t))^n. \end{aligned}$$

This of course agrees with the well-known, elementary fact that for an $M/G/\infty$ queue, the distribution of $Q(t)$ is Poisson with mean $\lambda E[B] P(R_{1,B} \leq t)$. $\square$

4 Time-dependent moments of a preemptive-LCFS queue

Consider an $M/GI/1$ preemptive LCFS queue, where we assume that the arrival rate is λ, and that each customer brings with it a generally distributed amount of work S, with mean $E[S]$. We will assume throughout that the first moment of the busy period is finite for each type of preemptive model considered (recall the discussion of our use of the term "preemptive LCFS queue" in the introduction). With that being said, the reader should realize that the transient Little laws themselves are valid for any choice of parameter values associated with our arrival and service times. The main reason why this assumption is being made is because it will make the form of our final results more pleasing, from an interpretation standpoint.

We begin by first computing $E[(Q(t))_n]$, while assuming that $Q(0) = 0$. Then from our previous observations,

$$E[(Q(t))_n] = n!\lambda^n \int_0^t \cdots \int_0^{s_2} P_{s_1,\ldots,s_n}(W(s_1) > t - s_1, \ldots, W(s_n) > t - s_n) ds_1 \ldots ds_n. \tag{7}$$

However, a little thought shows that under this queue discipline, if $s_1 < s_2 < \ldots < s_n$, then

$$\begin{aligned} & P_{s_1,s_2,\ldots,s_n}(W(s_1) > t - s_1, W(s_2) > t - s_2, \ldots, W(s_n) > t - s_n) \\ = \; & P_{s_1,s_2,\ldots,s_n}(W(s_1) > s_2 - s_1, W(s_2) > s_3 - s_2, \ldots, W(s_n) > t - s_n) \\ = \; & P(\tau > s_2 - s_1) P(\tau > s_3 - s_2) \cdots P(\tau > t - s_n) \end{aligned}$$

where τ is a random variable that represents the busy period. Here the first equality comes from the fact that, on the set where $\{W(s_n) > t - s_n\}$, $W(s_{n-1}) > t - s_{n-1}$ if and only if $W(s_{n-1}) > s_n - s_{n-1}$. The second equality follows from the fact that each event $\{W(s_i) > s_{i+1} - s_i\}$ can be expressed in terms of the work S_i that the customer arriving at time s_i brings to the system, along with the points of N in (s_i, s_{i+1}) and their respective marks. These facts, along with N being Poisson shows that the events are independent. Plugging this into (7) gives

$$\begin{aligned} E[(Q(t))_n] &= n!\lambda^n \int_0^t \cdots \int_0^{s_2} P(\tau > s_2 - s_1) P(\tau > s_3 - s_2) \ldots P(\tau > t - s_n) ds_1 ds_2 \cdots ds_n \\ &= n!\lambda^n E[\tau] \int_0^t \cdots \int_0^{s_3} P(\tau > s_3 - s_2) \ldots P(\tau > t - s_n) P(R_\tau \leq s_2) ds_2 \ldots ds_n \\ &= n!\lambda^n E[\tau] \int_0^t \cdots \int_0^{s_3} P(\tau > s_2) \ldots P(\tau > t - s_n) P(R_\tau \leq s_3 - s_2) ds_2 \ldots ds_n \\ &= \ldots = n!(\lambda E[\tau])^n P(\sum_{k=1}^n R_{\tau,k} \leq t) \end{aligned}$$

where $\{R_{\tau,k}\}_{k \geq 1}$ represents an i.i.d. sequence of residual busy periods. This of course agrees with the double transform computed in [17], and it is also in agreement with what's found in [3], [4].

Similarly, if there is one customer in the system at time zero with an amount of service that is equal in distribution to S, we see that if W_1 represents its sojourn time,

$$\begin{aligned} E[\mathbf{1}(W_1 > t)(Q(t))_n] &= n!\lambda^n \int_0^t \cdots \int_0^{s_2} P_{s_1,\ldots,s_n}(W_1 > t, W(s_1) > t - s_1, \ldots, W(s_n) > t - s_n) ds_1 \ldots ds_n \\ &= n!\lambda^n \int_0^t \cdots \int_0^{s_2} P(\tau > s_1)P(\tau > s_2 - s_1)P(\tau > s_3 - s_2)\ldots P(\tau > t - s_n) ds_1 \ldots ds_n \\ &= n!(\lambda E[\tau])^n \left(P\left(\sum_{j=1}^n R_{\tau,j} \le t \right) - P\left(\tau + \sum_{j=1}^n R_{\tau,j} \le t \right) \right) \\ &= n!(\lambda E[\tau])^n \left(P\left(\tau + \sum_{j=1}^n R_{\tau,j} > t \right) - P\left(\sum_{j=1}^n R_{\tau,j} > t \right) \right). \end{aligned}$$

Notice that if there are two customers in the system at time zero instead of one, the τ found in the above expression would have to be replaced with a convolution of two busy periods, and similarly for any other number of customers that happen to be present at time zero.

With these calculations in mind, we are now ready to derive an expression for the time-dependent moments.

Theorem 4.1 *Suppose $Q(0) = n_0 \ge 0$, where the customers are labeled with the numbers $1, 2, ..., n_0$, with service times that are independent and equal in distribution to S, and are served in this order (but still under preemptive LCFS). Then*

$$\begin{aligned} E[Q(t)^n] &= \sum_{l=0}^n S(n,l) l! (\lambda E[\tau])^l P(\sum_{k=1}^l R_{\tau,k} \le t) \\ &+ \sum_{m=1}^n \binom{n}{m} \sum_{r=1}^{n_0} [(n_0 - r + 1)^m - (n_0 - r)^m] \sum_{k=0}^{n-m} S(n-m,k) k! (\lambda E[\tau])^k \\ &\times \left[P\left(\sum_{j=1}^r \tau_j + \sum_{l=1}^k R_{\tau,l} > t \right) - P\left(\sum_{l=1}^k R_{\tau,l} > t \right) \right]. \end{aligned}$$

Here the sequences $\{\tau_k\}_{k\ge1}$ and $\{R_{\tau,k}\}_{k\ge1}$ represent two independent, i.i.d. sequences of busy periods and residual busy periods, respectively. Moreover, the doubly indexed sequence of integers $\{S(n,k)\}_{n\ge1,1\le k\le n}$ represent the Stirling numbers of the second kind.

Proof Our new Little's result will allow us to write down a closed-form expression for the n^{th} moment of $Q(t)$. Notice that we can write

$$Q(t) = \sum_{k=1}^n \mathbf{1}(W_k > t) + Q_0(t)$$

where $Q_0(t) = \int_0^t \mathbf{1}(W(s) > t - s) N(ds)$. Furthermore, it is also clear that

$$\begin{aligned} Q(t)^n &= \sum_{m=0}^n \binom{n}{m} \sum_{1 \le r_1,\ldots,r_m \le n_0} \mathbf{1}(W_{\min(r_i, 1\le i \le m)} > t) Q_0(t)^{n-m} \\ &= \sum_{m=0}^n \binom{n}{m} \sum_{1 \le r_1,\ldots,r_m \le n_0} \mathbf{1}(W_{\min(r_i, 1\le i \le m)} > t) \sum_{k=0}^{n-m} S(n-m,k)(Q_0(t))_k. \end{aligned} \quad (8)$$

We can further simplify this expression by noticing that for a fixed nonnegative integer $m \leq n$ and a fixed positive integer $r \leq m$, the number of times r appears as the index of W in (8) is just

$$1 + \sum_{k=1}^{m-1} \binom{m}{k} (n_0 - r)^{m-k} = (n_0 - r + 1)^m - (n_0 - r)^m$$

which follows from a simple counting argument. Thus, after making use of this fact and taking expectations in (8), we find that

$$E[Q(t)^n] = \sum_{m=0}^{n} \binom{n}{m} \sum_{r=1}^{m} [(n_0 - r + 1)^m - (n_0 - r)^m] \sum_{k=0}^{n-m} S(n-m,k) E[\mathbf{1}(W_r > t)(Q_0(t))_k]$$

so it suffices to compute the expectations found within the sum. But by applying the Campbell-Mecke formula in the same way as before, it is easy to see that if $m = 0$,

$$E[(Q_0(t))_k] = k!(\lambda E[\tau])^k P\left(\sum_{j=1}^{k} R_{\tau,j} \leq t\right).$$

Otherwise, if $m \geq 1$, we see from our previous calculations that for a positive integer r,

$$E[\mathbf{1}(W_r > t)(Q_0(t))_k] = k!(\lambda E[\tau])^k \left[P\left(\sum_{j=1}^{r} \tau_j + \sum_{l=1}^{k} R_{\tau,l} > t\right) - P\left(\sum_{l=1}^{k} R_{\tau,l} > t\right)\right].$$

This completes the proof. $\diamondsuit$

Remark It would be very interesting to see if there is still a connection between the marginal distributions of the $M/G/1$ queues under LCFS-PR and processor sharing, for an arbitrary initial condition. In general, it appears very difficult to make use of the work of Kitaev [18] in the hopes of establishing that they are indeed the same. Indeed, it is easy to see that the argument given in the proof of Theorem 2.2 in Denisov and Sapozhnikov [11] cannot be used to prove this statement, if we assume that each customer present in the system at time zero has a remaining amount of work that is equal in distribution to the service time of all customers that arrive after time zero. $\square$

5 Moments of a regulated Brownian motion

As is mentioned in [3], it is possible to rescale time and space in such a way so that the sample path of the $M/M/1$ queue converges in distribution (under the Skorohod metric) to a regulated Brownian motion (RBM) $\{R(t); t \geq 0\}$, with drift parameter $\mu = -1$ and diffusion coefficient $\sigma^2 = 1$, as $\rho \to 1$. In particular, as $\rho \to 1$, it is known that since the sample paths of RBM are continuous with probability one, we see that for each $t \geq 0$,

$$\frac{(1-\rho)}{2} Q\left(\frac{2t}{(1-\rho)^2}\right) \Rightarrow R(t)$$

where $(\Rightarrow)$ is used to denote weak convergence. Therefore, we can use our time-dependent moments of the $M/M/1$ queue-length to derive the time-dependent moments of a regulated Brownian motion for any initial condition $x \geq 0$, which complements and extends the results given in [1], [2].

Theorem 5.1 *For each $n \geq 1$, $x \geq 0$, we find that*

$$E[R(t)^n | R(0) = x] = \frac{n!}{2^n}\left[P\left(\sum_{k=1}^{n} I_k \leq t\right) + \sum_{m=1}^{n}\int_0^x \left[P\left(\sum_{l=1}^{n-m} I_l \leq t\right) - P\left(T_{x-u,0} + \sum_{l=1}^{n-m} I_l \leq t\right)\right]\frac{u^{m-1}}{(m-1)!}du\right]$$

where $T_{x,0} := \inf\{t > 0 : R(t) = 0\}$ (assuming $R(0) = x$), and $\{I_k\}_{k\geq 1}$ is an i.i.d. sequence of random variables such that

$$P(I_1 \leq t) = \int_0^\infty P(T_{x,0} \leq t)2e^{-2x}dx.$$

Proof To prove the result, we will define a sequence of $M/M/1$ queues Q_{n_0} that begin in state $Q_{n_0}(0) = 2^{n_0+1}$, with a service rate of 1 and a traffic intensity $\rho_{n_0} = 1 - x/2^{n_0}$. From Theorem 4.1, we find that

$$\begin{aligned}
& E\left[\left(\frac{1-\rho_{n_0}}{2}\right)^n Q_{n_0}\left(\frac{2t}{(1-\rho_{n_0})^2}\right)^n\right] \\
= & \left(\frac{1-\rho_{n_0}}{2}\right)^n \sum_{l=0}^{n} S(n,l)l!\left(\frac{\rho_{n_0}}{1-\rho_{n_0}}\right)^l P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{k=1}^{l} R_{\tau,k} \leq t\right) \\
+ & \left(\frac{1-\rho_{n_0}}{2}\right)^n \sum_{m=1}^{n}\binom{n}{m}\sum_{r=1}^{2^{n_0+1}}\left[(2^{n_0+1}-r+1)^m - (2^{n_0+1}-r)^m\right]\sum_{k=0}^{n-m} S(n-m,k)k!\left(\frac{\rho_{n_0}}{1-\rho_{n_0}}\right)^k \\
\times & \left[P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{l=1}^{k} R_{\tau,l} \leq t\right) - P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{j=1}^{r}\tau_j + \frac{(1-\rho_{n_0})^2}{2}\sum_{l=1}^{k} R_{\tau,l} \leq t\right)\right] \\
= & \frac{1}{2^n}\sum_{l=0}^{n} S(n,l)l!\rho_{n_0}^l(1-\rho_{n_0})^{n-l} P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{k=1}^{l} R_{\tau,k} \leq t\right) \\
+ & \frac{1}{2^n}\sum_{m=1}^{n}\binom{n}{m}\sum_{r=1}^{2^{n_0+1}}\left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m\right]\sum_{k=0}^{n-m} S(n-m,k)k!\rho_{n_0}^k(1-\rho_{n_0})^{n-m-k} \\
\times & \left[P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{l=1}^{k} R_{\tau,l} \leq t\right) - P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{j=1}^{r}\tau_j + \frac{(1-\rho_{n_0})^2}{2}\sum_{l=1}^{k} R_{\tau,l} \leq t\right)\right].
\end{aligned}$$

Notice that as $n_0 \to \infty$, $\rho_{n_0} \to 1$, and so all terms that are multiplied by the constant $1 - \rho_{n_0}$ disappear. Hence, as $n_0 \to \infty$, the limit of the scaled n^{th} moment is the same as the limit of

$$\begin{aligned}
& \frac{n!}{2^n}\rho_{n_0}^n P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{k=1}^{n} R_{\tau,k} \leq t\right) \\
+ & \frac{n!}{2^n}\sum_{m=1}^{n}\frac{1}{m!}\rho_{n_0}^{n-m}\sum_{r=1}^{2^{n_0+1}}\left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m\right] \\
\times & \left[P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{l=1}^{n-m} R_{\tau,l} \leq t\right) - P\left(\frac{(1-\rho_{n_0})^2}{2}\sum_{j=1}^{r}\tau_j + \frac{(1-\rho_{n_0})^2}{2}\sum_{l=1}^{n-m} R_{\tau,l} \leq t\right)\right].
\end{aligned}$$

It has already been established in Corollary 5.2.2(a) of [4] that, as $n_0 \to \infty$,

$$\frac{(1-\rho_{n_0})^2}{2} R_{\tau,1} \Rightarrow I_1 \tag{9}$$

where the distribution of I_1 is as given in the theorem. Hence, we see that for each $t \geq 0$,

$$\lim_{n_0\to\infty} \frac{n!}{2^n} \rho_{n_0}^n P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{k=1}^{n} R_{\tau,k} \leq t\right) = \frac{n!}{2^n} P\left(\sum_{k=1}^{n} I_k \leq t\right).$$

Furthermore, it's also clear that for each $t \geq 0$,

$$\begin{aligned} &\lim_{n_0\to\infty} \frac{n!}{2^n} \sum_{m=1}^{n} \frac{1}{m!} \rho_{n_0}^{n-m} \sum_{r=1}^{2^{n_0+1}} \left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m\right] P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{l=1}^{n-m} R_{\tau,l} \leq t\right) \\ = &\frac{n!}{2^n} \sum_{m=1}^{n} \frac{1}{m!} x^m P\left(\sum_{l=1}^{n-m} I_l \leq t\right). \end{aligned}$$

To complete the proof it will suffice to compute, for each $t \geq 0$,

$$\lim_{n_0\to\infty} \frac{n!}{2^n} \sum_{m=1}^{n} \frac{1}{m!} \rho_{n_0}^{n-m} \sum_{r=1}^{2^{n_0+1}} \left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m\right] P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{j=1}^{r} \tau_j + \frac{(1-\rho_{n_0})^2}{2} \sum_{l=1}^{n-m} R_{\tau,l} \leq t\right).$$

We claim that, uniformly in $t \geq 0$,

$$\begin{aligned} &\lim_{n_0\to\infty} \sum_{r=1}^{2^{n_0+1}} \left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m\right] P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{j=1}^{r} \tau_j \leq t\right) \\ = &\int_0^x P(T_{x-u,0} \leq t) m u^{m-1} du. \end{aligned} \tag{10}$$

To see this, first define a sequence of functions f_{n_0}, where $f_{n_0}(z) = P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{j=1}^{r} \tau_j \leq t\right)$ for $z \in [r/2^{n_0+1}, (r+1)/2^{n_0+1})$. Notice that $P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{j=1}^{r} \tau_j \leq t\right)$ can be interpreted as the probability that the properly time and space-scaled $M/M/1$ queue that starts in state $r/2^{n_0+1}$ reaches level zero before time t. Thus, since the sample paths of regulated Brownian motion are continuous and leave zero immediately after reaching it, we can apply the continuous mapping theorem to conclude that

$$\lim_{n_0\to\infty} f_{n_0}(z) = f(z)$$

where $f(z) = P(T_{z,0} \leq t)$. Establishing that (10) holds uniformly in t follows from applying the Dominated Convergence Theorem, along with Lemma 4.1 of Dai [10]. Finally, by combining (9) with the uniform convergence of (10), we see that

$$\begin{aligned} &\lim_{n_0\to\infty} \frac{n!}{2^n} \sum_{m=1}^{n} \frac{1}{m!} \rho_{n_0}^{n-m} \sum_{r=1}^{2^{n_0+1}} \left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m\right] \\ &\times P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{j=1}^{r} \tau_j + \frac{(1-\rho_{n_0})^2}{2} \sum_{l=1}^{n-m} R_{\tau,l} \leq t\right) \end{aligned}$$

$$
\begin{aligned}
&= \lim_{n_0\to\infty} \int_0^t \frac{n!}{2^n} \sum_{m=1}^{n} \frac{1}{m!} \rho_{n_0}^{n-m} \sum_{r=1}^{2^{n_0+1}} \left[(x(1-\frac{r-1}{2^{n_0+1}}))^m - (x(1-\frac{r}{2^{n_0+1}}))^m \right] \\
&\quad \times P\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{j=1}^{r} \tau_j \le t-y \right) dP\left(\frac{(1-\rho_{n_0})^2}{2} \sum_{l=1}^{n-m} R_{\tau,l} \le y \right) \\
&= \frac{n!}{2^n} \sum_{m=1}^{n} \frac{1}{m!} \int_0^x P\left(T_{x-u,0} + \sum_{l=1}^{n-m} R_{\tau,l} \le t \right) m u^{m-1} du
\end{aligned}
$$

which completes the proof. ♢

Acknowledgments The authors would like to thank Yoni Nazarathy for stimulating discussions, which eventually led to the creation of this article. We would also like to thank both Onno Boxma and Richard Serfozo for their comments and suggestions.

References

[1] J. Abate and W. Whitt (1987). Transient behavior of regulated Brownian motion, I: starting at the origin. *Advances in Applied Probability* **19**, 560-598.

[2] J. Abate and W. Whitt (1987). Transient behavior of regulated Brownian motion, II: non-zero initial conditions. *Advances in Applied Probability* **19**, 599-631.

[3] J. Abate and W. Whitt (1987). Transient behavior of the $M/M/1$ queue: starting at the origin. *Queueing Systems* **2**, 41-65.

[4] J. Abate and W. Whitt (1988). Transient behavior of the $M/M/1$ queue via Laplace transforms. *Advances in Applied Probability* **20**, 145-178.

[5] F. Baccelli and P. Brémaud (2003). *Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences.* Springer, New York.

[6] D. Bertsimas and G. Mourtzinou (1997). Transient laws of non-stationary queueing systems and their applications. *Queueing Systems* **25**, 115-155.

[7] D. Bertsimas and D. Nakazato (1995). The distributional Little's law and its applications. *Operations Research* **43**, 298-310.

[8] B. Blaszczyszyn (1995). Factorial moment expansion for stochastic systems. *Stochastic Processes and their Applications* **56**, 321-335.

[9] B. Blaszczyszyn, T. Rolski and V. Schmidt (1995). Light-traffic approximation in queues and related stochastic models. *Advances in Queueing: Theory, Methods, and Open Problems*, CRC Press, 379-406

[10] J.G. Dai (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability*, **5**, 49-77.

[11] D. Denisov and A. Sapozhnikov (2006). On the distribution of the number of customers in the symmetric $M/G/1$ queue. *Queueing Systems* **54**, 237-241.

[12] B. H. Fralix (2008). A time-dependent view of the ASTA problem, with applications to preemptive queues and birth-death processes. To be submitted.

[13] R. Haji and G. Newell (1971). A relation between stationary queue and waiting time distributions. *Journal of Applied Probability* **8**, 617-620.

[14] O. Kallenberg (1983). *Random Measures.* Akademie-Verlag, Berlin.

[15] F. Kelly (1979). *Reversibility and Stochastic Networks.* John Wiley & Sons, Chichester.

[16] J. Kelson and L. Servi (1988). A distributional form of Little's law. *Operations Research Letters* **7**, 223-227.

[17] O. Kella, B. Zwart and O. Boxma (2005). Some time-dependent properties of symmetric $M/G/1$ queues. *Journal of Applied Probability* **42**, 223-234.

[18] M. Yu. Kitaev (1993). The $M/G/1$ processor-sharing model: transient behavior. *Queueing Systems* **14**, 239-273.

[19] G. Riaño (2002). *Transient Behavior of Stochastic Networks: Application to Production Planning with Load-Dependent Lead Times.* Ph. D. thesis, Georgia Institute of Technology.

[20] T. Rolski (1989). Relationships between characteristics in periodic Poisson queues. *Queueing Systems* **4**, 17-26.

[21] Ryll-Nardzewski, C. (1961). Remarks on process of call. *Proc. 4th Berkeley Symposium, vol 2.* University of California Press, Berkeley, 455–465.

[22] R.F. Serfozo (1999). *Introduction to Stochastic Networks.* Springer, New York

[23] W. Whitt (1991). A review of $L = \lambda W$ and extensions. *Queueing Systems* **9**, 235-268.