LIMITS AND APPROXIMATIONS FOR THE M/G/1 LIFO WAITING-TIME DISTRIBUTION

by

Joseph Abate[1] and Ward Whitt[2]

April 15, 1996
Revision: January 2, 1997

Operations Research Letters 20 (1997) 199–206

[1]900 Hammond Road, Ridgewood, NJ 07450-2908
[2]AT&T Research, Room 2C-178, Murray Hill, NJ 07974-0636 (wow@research.att.com)

Abstract

We provide additional descriptions of the steady-state waiting-time distribution in the M/G/1 queue with the last-in first-out (LIFO) service discipline. We establish heavy-traffic limits for both the cumulative distribution function (cdf) and the moments. We develop an approximation for the cdf that is asymptotically correct both as the traffic intensity $\rho \to 1$ for each time t and as $t \to \infty$ for each ρ. We show that in heavy traffic the LIFO moments are related to the FIFO moments by the Catalan numbers. We also develop a new recursive algorithm for computing the moments.

Key Words: M/G/1 queue; Last-in first-out service discipline; Waiting time; Heavy Traffic; Asymptotics; Asymptotic normal approximation; Moments; Catalan numbers

1. Introduction

The purpose of this paper is to deduce additional properties of the steady-state waiting-time distribution (until beginning service) in the M/G/1 queue with the last-in first-out (LIFO) service discipline. We obtain our results here by applying results in our previous papers [1]–[6]. Our results here complement previous M/G/1 LIFO results by Vaulot [17], Wishart [19], Riordan [13], [14], Takács [16] and Iliadis and Fuhrmann [10].

We start in Section 2 by establishing a heavy-traffic limit theorem for the M/G/1 LIFO steady-state waiting-time distribution. In Section 3 we give a new derivation for the large-time asymptotics of the LIFO steady-state waiting-time distribution, complementing our recent result in [1]. In Section 4 we then develop an approximation for the LIFO waiting-time distribution, called the asymptotic normal approximation, and show that it is asymptotically correct both as time $t \to \infty$ for each value of the traffic intensity ρ and as $\rho \to 1$ for each t. Our approximation generalizes an approximation suggested for the M/M/1 model by Riordan [14], p. 109. Sections 2–4 parallel and draw on our previous limiting results for the M/G/1 busy-period distribution in [5]. In Section 5 we give a numerical example showing that the asymptotic normal approximation for the LIFO steady-state waiting-time cdf performs very well. We show that it is much better than the large-time limit (the limit as $t \to \infty$ for each fixed ρ).

In Section 6 we develop a new recursive algorithm for computing the M/G/1 LIFO waiting-time moments. We believe that this algorithm is an attractive alternative to previous algorithms developed by Takács [16] and Iliadis and Fuhrmann [10]. The first four moments are given explicitly. In Section 7 we show that the moments have a very simple asymptotic form in heavy traffic.

2. Heavy-Traffic Limit for the LIFO CDF

In this section we establish a heavy-traffic limit for the LIFO steady-state waiting-time cdf, denoted by W_L. For this purpose, we establish heavy-traffic limits for the M/G/1 workload process moment cdf's. These workload results extend previous results for the M/M/1 special case in Corollary 5.22 of [3]. The heavy-traffic limits for the workload process are local-limit refinements (versions for densities) to theorems that can be deduced directly from older heavy-traffic limit theorems, e.g., in Whitt [18].

We use the following notation. For any cumulative distribution function (cdf) F, let f be the associated probability density function (pdf) and let $m_k(F)$ be the k^{th} moment. Let $F^c(t) \equiv 1-F(t)$

be the associated complementary cdf (ccdf) and let the associated equilibrium excess ccdf be

$$F_e^c(t) = m_1(F)^{-1} \int_t^\infty F^c(u) \ . \tag{2.1}$$

The pdf of F_e is f_e and its $k^{\rm th}$ moment is $m_k(F_e)$.

For $k \geq 1$, let H_k be the $k^{\rm th}$ moment cdf of canonical (drift 1, diffusion coefficient 1) reflected Brownian motion (RBM) $\{R(t) : t \geq 0\}$, i.e.,

$$H_k(t) = \frac{E[R(t)^k | R(0) = 0]}{E[R(\infty)^k]} \ , \quad t \geq 0 \ . \tag{2.2}$$

By definition, $H_k(t)$ as a function of t is the expectation of a stochastic process, but it also has the structure of a cdf. Throughout this paper we relate the tail behavior of cdf's to large-time asymptotics of stochastic processes. Properties of the RBM moment cdf's are established in [2]. Let h_k be the density of H_k. Let Φ be the standard (mean 0, variance 1) normal cdf and let ϕ be its density. We shall primarily focus on the pdf h_1, which has the formula

$$h_1(t) = 2t^{1/2}\phi(t^{1/2}) - 2[1 - \Phi(t^{1/2})] = 2\gamma(t) - \gamma_e(t) \ , \tag{2.3}$$

where γ is the gamma pdf with mean 1 and shape parameter 1/2, i.e.,

$$\gamma(t) = (2\pi t)^{-1/2} e^{-t/2} \ , \quad t \geq 0 \ , \tag{2.4}$$

and γ_e is the associated equilibrium-excess pdf.

We consider a family of M/G/1 queueing systems indexed by the arrival rate ρ. We assume that the service-time distribution is fixed, having cdf G with pdf g, mean 1 and second moment $m_2(G)$. Let $\{V_\rho(t) : t \geq 0\}$ be the M/G/1 workload (unfinished work or virtual waiting time) process and, for $k \geq 1$, let $H_{\rho k}$ be its associated $k^{\rm th}$ moment cdf, expressed as a function of the arrival rate ρ, i.e.,

$$H_{\rho k}(t) = \frac{E[V_\rho(t)^k | V_\rho(0) = 0]}{E[V_\rho(\infty)^k]} \ , \quad t \geq 0 \ , \tag{2.5}$$

see [4]. Just as with $H_k(t)$ in (2.2), $H_{\rho k}(t)$ as a function of t is simultaneously the expectation of a stochastic process and a cdf for each ρ. Also let $H_{\rho 0}$ be the $0^{\rm th}$-moment cdf or server-occupancy cdf, defined by

$$H_{\rho 0}(t) = (1 - P_{00}(t))/\rho \ , \quad t \geq 0 \ , \tag{2.6}$$

where $P_{00}(t)$ is the probability of emptiness at time t starting empty at time 0 (with arrival rate ρ), as in (23) of [4].

The key formula connecting these expressions to the LIFO cdf W_L is

$$W_L^c(t) = P_{00}(t) - (1-\rho) = \rho H_{\rho 0}^c(t) \ , \tag{2.7}$$

due to Takács [16]; see (36) of [1] and (78) on p. 500 of [16]. The next theorem describes the LIFO steady-state waiting-time cdf W_L in heavy traffic.

Theorem 2.1 *For each $t > 0$,*

$$\begin{aligned} \lim_{\rho\to 1} \quad & 2(1-\rho)^{-1} W_L^c(tm_2(G)(1-\rho)^{-2}) \\ &= \lim_{\rho\to 1} 2(1-\rho)^{-1} H_{\rho 0}^c(tm_2(G)(1-\rho)^{-2}) \\ &= \lim_{\rho\to 1} m_2(G)(1-\rho)^{-2} h_{\rho 1}(tm_2(G)(1-\rho)^{-2}) = h_1(t) \end{aligned}$$

for h_1 in (2.3).

Proof. By (2.7), the first limit is equivalent to the second. By Theorem 2(b) of [4], $h_{\rho 0e}(t) = H_{\rho 1}(t)$, so that $h_{\rho 1}(t) = H_{\rho 0}(t)/h_{\rho 01}$. By Theorem 6(b) of [4], $h_{\rho 01} = \rho^{-1} v_{\rho 1} = m_2(G)/2\rho(1-\rho)$. Hence the third limit is equivalent to the second.

By Theorems 3(a) and 4(b) of [4],

$$\hat{h}_{\rho 1}(s) = \frac{1-\hat{b}_{\rho e}(s)}{s b_{\rho e 1}(1-\rho+\rho \hat{b}_{\rho e}(s))} \ , \tag{2.8}$$

where $\hat{b}_{\rho \epsilon}(s)$ is the transform of the equilibrium excess distribution associated with the busy period and $b_{\rho e1}$ is its mean. For brevity, let $c = (1-\rho)^2/m_2(G)$ and note that $b_{\rho e 1} = (2c)^{-1}$ by Theorem 5(b) of [4]. Then, by (2.8),

$$\hat{h}_{\rho 1}(cs) = \frac{2[1-\hat{b}_{\rho e}(cs)]}{s(1-\rho+\rho \hat{b}_{\rho e}(cs))} \ .$$

Since we have expressions for the transforms, we establish the convergence using them. By Theorem 1 of [5], $\hat{b}_{\rho e}(cs) \to \hat{h}_1(s)$ as $\rho \to 1$. (Equations (2.9) of [5] should read $h_\rho(t) \equiv m_2(1-\rho)^{-2}$ $b_{\rho e}(tm_2(1-\rho)^{-2}) = m_2(1-\rho)^{-1} B_\rho^c(tm_2(1-e)^{-2})$, with m_2 there being $m_2(G)$ here. Incidentally, a space normalization — typically $c(1-\rho)$ for a constant c — is missing before $W_\rho^*(dt(1-\rho)^{-2})$ in Condition C2 on p. 589 in [5] as well.) Hence,

$$\hat{h}_{\rho 1}(cs) \to \frac{2[1-\hat{h}_1(s)]}{s\hat{h}_1(s)} = \frac{\hat{h}_{1e}(s)}{\hat{h}_1(s)} = \hat{h}_1(s) \ ,$$

with the last step following from (2.6) of [5] or p. 568 of [2]. □

The LIFO approximation obtained from Theorem 2.1 is

$$W_L^c(t) \approx \frac{(1-\rho)}{2} h_1(t(1-\rho)^2/m_2(G)) \ . \tag{2.9}$$

The M/M/1 special case of Theorem 2.1 for $h_{\rho 1}(t)$ is given in Corollary 5.2.2 of [3]. In [3] the time scaling is done at the outset; see (2.2) there.

The limit for the density $h_{\rho 1}$ implies a corresponding limit for the associated cdf $H_{\rho 1}$, by Scheffé's theorem, p. 224 of Billingsley [7]. The following result could also be deduced from standard heavy-traffic limits for the workload process [18].

Corollary. *For each $t > 0$,*

$$\lim_{\rho \to 1} H_{\rho 1}(tm_2(G)(1-\rho)^{-2}) = H_1(t) = 1 - 2(1+t)[1 - \Phi(t^{1/2})] + 2t^{1/2}\phi(t^{1/2}) \ .$$

3. Large-Time Asymptotics

In [1] we used (2.7) and an integral representation for $P_{00}(t)$ to establish the asymptotic behavior of $W_L^c(t)$ as $t \to \infty$. Here we give an alternative derivation based on asymptotics for the density $w_L(t)$. Let $f(t) \sim g(t)$ as $t \to \infty$ mean that $f(t)/g(t) \to 1$ as $t \to \infty$.

We start with the elementary observation that the conditional LIFO steady-state waiting time given that it is positive coincides with the busy period generated by the equilibrium excess of a service time. Let $b(t, \theta)$ be the density of a busy period starting with a service time of length θ. Then the LIFO waiting-time distribution has an atom of size $1 - \rho$ at the origin and a density

$$w_L(t) = \rho \int_0^\infty b(t,\theta) g_e(\theta) d\theta \ , \quad t > 0 \ . \tag{3.1}$$

From (3.1), we reason as in Cox and Smith [8] and [1] to obtain the following asymptotic result.

Theorem 3.1 *As $t \to \infty$,*

$$w_L(t) \sim (\omega/\tau)(\pi t^3)^{-1/2} e^{-t/\tau} \tag{3.2}$$

and

$$W_L^c(t) \sim \omega(\pi t^3)^{-1/2} e^{-t/\tau} \ , \tag{3.3}$$

where τ^{-1} is the asymptotic decay rate (reciprocal of the relaxation time), given by

$$\tau^{-1} = \rho + \zeta - \rho \hat{g}(-\zeta) \ , \tag{3.4}$$

ζ is the unique real number u to the left of all singularities of the service-time moment generating function $\hat{g}(-s)$ (assumed to exist) such that

$$\hat{g}'(-u) = -\rho^{-1} \ , \tag{3.5}$$

$$\omega = \rho\alpha/\zeta^2 \ , \tag{3.6}$$

$$\alpha = [2\rho^3 \hat{g}''(-\zeta)]^{-1/2} \ . \tag{3.7}$$

Proof. We apply (19) of [1] with (3.1) to obtain the integral representation

$$w_L(t) = (\rho/t)\mathcal{L}^{-1}(-\hat{g}_e'(s)\exp(-\rho t(1-\hat{g}(s))) \ , \tag{3.8}$$

where $\mathcal{L}^{-1}$ is the Laplace transform inversion operator (Bromwich contour integral)

$$\mathcal{L}^{-1}(\hat{g}(s))(t) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} e^{st}\hat{f}(s)ds \ , \tag{3.9}$$

where the contour $Re(s) = a$ is to the right of all singularities of $\hat{f}(s)$. We then apply Laplace's method as on p. 156 of Cox and Smith [8] and pp. 80, 121 and 127 of Olver [12] to obtain (3.2). We can also obtain the asymptotic result for $w_L(t)$ by relating (3.8) to the integral representation for the busy-period density $b(t)$ in (3) of [1]. From p. 127 of Olver [12], we see that

$$w_L(t) \sim \rho^2(-\hat{g}_e'(-\zeta)b(t) \quad \text{as} \quad t \to \infty \tag{3.10}$$

where

$$-\hat{g}_e'(-\zeta) = 1/\zeta^2\rho\tau \ , \tag{3.11}$$

so that

$$w_L(t) \sim \frac{\rho}{\zeta^2\tau} b(t) \quad \text{as} \quad t \to \infty \ . \tag{3.12}$$

Hence, we can invoke the known asymptotics for $b(t)$ in (5) of [1]. Finally, we integrate (3.2) to obtain (3.3); see p. 17 of Erdélyi [9].

4. The Asymptotic Normal Approximation

In [5] we found that the tail of the M/G/1 busy-period ccdf is well approximated by an asymptotic normal approximation, which is asymptotically correct as $\rho \to 1$ for each t and as $t \to \infty$ for each ρ. We now apply essentially the same reasoning to develop a similar approximation for the LIFO steady-state waiting-time ccdf.

Paralleling Theorem 2 of [5], we can express the asymptotic relation in (3.3) in terms of $h_1(t)$, because

$$h_1(t) \sim 2t^{-1}\gamma(t) \quad \text{as} \quad t \to \infty \ , \tag{4.1}$$

for γ in (2.4). The following is our *asymptotic normal approximation*, obtained by combining (2.4), (3.1) and (4.1). The M/M/1 special case was proposed for the M/M/1 queue by Riordan [14], p. 109.

Theorem 4.1 *If (3.3) holds, then*

$$W_L^c(t) \sim 2\omega\tau^{-3/2} h_1(2t/\tau) \quad as \quad t \to \infty \ .$$

for $h_1(t)$ in (2.3), τ in (3.4) and ω in (3.6).

Theorem 4.1 directly shows that the asymptotic normal approximation is asymptotically correct as $t \to \infty$ for each ρ. Previous asymptotic relations for τ and ω show that it is also asymptotically correct as $\rho \to 1$ for each t. In particular, by (14) and (40) of [1],

$$\tau^{-1} = \frac{(1-\rho)^2}{2m_2(G)}(1 + O(1-\rho)) \quad \text{as} \quad \rho \to 1 \tag{4.2}$$

and

$$\omega = \frac{\tau}{2}\sqrt{\frac{m_2(G)}{2}}(1 + O(1-\rho)) \quad \text{as} \quad \rho \to 1 \ , \tag{4.3}$$

which is consistent with (2.9). For the constant, note that

$$2\omega\tau^{-3/2} \sim \sqrt{m_2(G)/2\tau} \sim \frac{1-\rho}{2} \quad \text{as} \quad \rho \to 1 \ . \tag{4.4}$$

5. A Numerical Example

In this section we compare the approximations for the M/G/1 LIFO steady-state waiting-time ccdf $W_L^c(t)$ to exact values obtained by numerical transform inversion. We consider the gamma service-time distribution with shape parameter 1/2, denoted by $\Gamma_{1/2}$, which has pdf g in (2.4) and Laplace transform

$$\hat{g}(s) = 1/\sqrt{1+2s} \ . \tag{5.1}$$

Its first four moments are $1, 3, 15$ and 105.

For the M/G/1 model there is no explicit result for the transform $\hat{W}_L(s)$. One way to proceed is to determine the transform values $\hat{P}_{00}(s)$ from the busy period via (36) of [4] or directly via the functional equation

$$\hat{P}_{00}(s) = \frac{1}{s + \rho\hat{g}(1/\hat{p}_{00}(s))} \ , \tag{5.2}$$

see (37) of [4], and then use (2.7). However, there is a more direct approach using the contour integral representation

$$W_L^c(t) = t^{-1}\mathcal{L}^{-1}(s^{-2}\exp(-\rho t(1-\hat{g}(s)))) - (1-\rho) \ , \tag{5.3}$$

where $\mathcal{L}^{-1}$ is the inverse Laplace transform operator; see (34) of [1]. Equation (3.8) above is a similar representation for the density.

Table 1 gives the results for $W_L^c(t)$ for the case $\rho = 0.75$. In Table 1 we also display results for three approximations: (i) the standard asymptotic approximation in (3.1), (ii) the heavy-traffic approximation in (2.9) and (iii) the asymptotic normal approximation in Theorem 3.1.

For the standard asymptotic approximation, we need to derive the asymptotic parameters ω and τ^{-1}. For this example, we find from the root equation (3.5) that

$$\zeta = \frac{1}{2}(1 - \rho^{2/3}) \ . \tag{5.4}$$

From (3.4), (3.7) and (3.6),

$$\tau^{-1} = \rho + \frac{1}{2} - \frac{3}{2}\rho^{2/3} \ , \tag{5.5}$$

$$\alpha = \frac{\rho^{-2/3}}{\sqrt{6}} \tag{5.6}$$

and

$$\omega = \frac{4\rho^{1/3}}{\sqrt{6}(1 - \rho^{2/3})^2} \ . \tag{5.7}$$

time t	exact by numerical inversion	asymptotic normal approximation, Theorem 4.1	heavy-traffic approximation, (2.8)	standard asymptotic approximation, (3.3)
10^{-8}	.7500000			
.1	.6893962	1.9	1.6	
.5	.5354077	.80	.66	
1.0	.4252065	.53	.44	
5.0	.1734549	.182	.15	
8.0	.1217696	.126	.11	1.1
15.0	.0700599	.071	.062	.40
30.0	.0328100	.0331	.030	.12
60.0	.0119596	.01203	.011	.029
80.0	.0070322	.00706	.0070	.015
120.0	.0028268	.002835	.0030	.0051
240.0	.0003043	.0003048	.00038	.00044
360.0	.0000444	.0000444	.000066	.000058

Table 1. A comparison of approximations with exact values of the LIFO steady-state waiting-time ccdf $W_L^c(t)$ in the $M/\Gamma_{1/2}/1$ model with $\rho = 0.75$.

Table 1 shows that the asymptotic normal approximation is remarkably accurate for t not too small, e.g., for $t \geq 5$. Moreover, the asymptotic normal approximation is much better than the standard asymptotic approximation based on (3.3). The heavy-traffic approximation is quite good though for t neither too large nor too small, e.g., for $1.0 \leq t \leq 120$. The standard asymptotic approximation is not good until t is very large. From Table 1, note that the standard asymptotic approximation for the tail probability is consistently high, whereas the heavy-traffic approximation (2.8) crosses the exact curve twice.

6. A Recursive Algorithm for the LIFO Moments

In this section we relate the k^{th} moments of the steady-state FIFO and LIFO waiting-time distributions, denote by v_k and w_k, respectively. For previous related work, see Takács [16] and Iliadis and Fuhrmann [10]. Iliadis and Fuhrmann show that the same relationship between LIFO and FIFO holds for a large class of models with Poisson arrivals.

Let h_{0k} be the k^{th} moment of the server-occupancy cdf $H_{\rho 0}$ in (2.6). By (2.7),

$$w_k = \rho h_{0k} \quad \text{for all} \quad k \geq 1 \ . \tag{6.1}$$

Theorem 6 of [4] gives an expression for the moments h_{0k} in terms of the moments b_{ek} of the busy-period equilibrium excess distribution, while Theorem 5 of [4] gives the recursive formula for the busy-period equilibrium-excess distribution moments b_{ek} in terms of the moments v_k. These two results give a recursive algorithm for computing the LIFO moments w_k in terms of the FIFO moments v_k.

First, as in (43) of [4], we can relate the FIFO waiting-time moments to the service-time moments via a recursion. Let the k^{th} service-time moment be denoted by g_k. The recursion is

$$v_k = \frac{\rho}{1-\rho} \sum_{j=1}^{k} \binom{k}{j} \frac{g_{j+1}}{j+1} v_{k-j} \ , \tag{6.2}$$

where $v_0 = 1$, so that the FIF0 moments v_k are readily available given the service-time moments g_k.

By Theorem 3a of [4], the equilibrium-time to emptiness transform $\hat{f}_{\epsilon 0}(s)$ can be expressed as

$$\hat{f}_{\epsilon 0}(s) = 1 - \rho + \rho \hat{b}_e(s) \ , \tag{6.3}$$

so that their moments are related by

$$f_{\epsilon 0 k} = \rho b_{ek} \ , \quad k \geq 1 \ . \tag{6.4}$$

Theorem 5 of [4] shows that these moments can be computed recursively, given the moments v_k. To emphasize the recursive nature, for any Laplace transform

$$\hat{f}(s) = \int_0^\infty e^{-st} f(t) dt \ , \tag{6.5}$$

let $T_n(\hat{f})$ be the truncated power series defined by

$$T_n(\hat{f})(s) = \sum_{k=0}^{n} \frac{f_k}{k!}(-s)^k \ , \tag{6.6}$$

where

$$f_k = \int_0^\infty t^k f(t) dt = \frac{(-1)^k d^k \hat{f}(s)}{ds^k}|_{s=0} \ , \tag{6.7}$$

which is well defined assuming that the integrals in (6.7) are finite. Then Theorem 5(a) of [4] can be restated as

$$f_{\epsilon 0k} = \sum_{j=1}^{k} \binom{k}{j} \frac{v_j}{1-\rho} [T_{k-1}(\hat{f}_{\epsilon 0})]^j_{k-j} \tag{6.8}$$

where $\hat{f}_j$ is understood to the j^{th} moment, which is $j!(-1)^j$ times the j^{th} coefficient of the power series.

Then we can calculate w_k from

$$w_k = \rho b_{ek} - \rho \sum_{j=1}^{k} \binom{k}{j} b_{ej} w_{k-j} \ , \quad k \geq 1 \ , \tag{6.9}$$

where $w_0 = \rho$, which is obtained by combining (6.1) here with Theorem 6(a) of [4].

In summary, if we want to calculate w_k, then we first calculate the first k FIFO waiting-time moments $v_1, \ldots, v_k$ recursively via (6.2). Then we calculate the first k equilibrium-time-to-emptiness moments $f_{\epsilon 01}, \ldots, f_{\epsilon 0k}$ recursively via (6.8). We obtain the associated busy-period equilibrium excess moments $b_{e1}, \ldots, b_{ek}$ via (6.4). Finally, we obtain the LIFO waiting-time moments $w_1, \ldots, w_k$ recursively via (6.9). For example, the first four are:

$$\begin{aligned} w_1 &= v_1 \ , \quad w_2 = \frac{v_2}{1-\rho} \\ w_3 &= \frac{v_3 + 3v_2v_1}{(1-\rho)^2} \ , \\ w_4 &= \frac{v_4 + 8v_3v_1 + 12v_2v_1^2 + 6v_2^2}{(1-\rho)^3} \ , \end{aligned} \tag{6.10}$$

which is consistent with Theorem 6 of [4]. In the case (e) of Theorem 6 in [4] there is a typographical error in h_{04}; all four terms there should be divided by $\rho(1-\rho)^3$.

7. Heavy-Traffic Limits for the LIFO Moments

We now show that the LIFO waiting-time moments have a simple asymptotic form in heavy traffic. As in Section 6, let g_k and w_k be the k^{th} moments of the service time and LIFO waiting time, respectively.

Theorem 7.1 *Assume that* $g_{n+2} < \infty$. *Then* $w_{n+1} < \infty$ *and*

$$w_n \sim \frac{(2n-2)!}{(n-1)!} \frac{w_1^n}{(1-\rho)^{n-1}} \quad \text{as} \quad \rho \to 1 \ . \tag{7.1}$$

To put Theorem 7.1 in perspective, it should be contrasted with the known heavy traffic results for first-in first-out (FIFO) and random order of service (ROS). For FIFO, the waiting-time distribution is asymptotically exponentially distributed as $\rho \to 1$, so that

$$v_n \sim n! v_1^n \quad \text{as} \quad \rho \to 1 \ . \tag{7.2}$$

For ROS, with W_R as the cdf, Kingman [11] showed that

$$1 - W_R(t) \approx 2\sqrt{t/w_{R1}} K_1(2\sqrt{t/w_{R1}})$$

where K_1 is the Bessel function, so that

$$w_{Rn} \sim (n!)^2 w_{R1}^n \quad \text{as} \quad \rho \to 1 \ . \tag{7.3}$$

Of course, $w_1 = w_{R1} = v_1$. It is interesting that, as $\rho \to 1$, the n^{th} moments v_n, w_{RN} and w_n in (7.1)–(7.3) are v_1^n times $n!$, $(n!)^2$ and $(2n-2)/(n-1)!(1-\rho)^{n-1}$, respectively. Hence, v_n and w_{RN} are $O((1-\rho)^{-n})$ as $\rho \to 1$, while w_n is $O((1-\rho)^{-(2n-1)})$.

We give another expression using the Catalan numbers. Let C_n be the n^{th} *Catalan number*, i.e.,

$$C_n = \frac{1}{n+1}\binom{2n}{n} \ ; \tag{7.4}$$

e.g., $1, 1, 2, 5, 14, \ldots$ see Riordan [15]. The Catalan numbers have the self-convolution property

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i} \ , \quad n \geq 1 \tag{7.5}$$

The Catalan numbers are also associated with the distribution of the number of customers served in an M/M/1 busy period; see p. 65 of Riordan [14]. We combine Theorem 7.1 and (7.2) to obtain the following representation.

Corollary. *Under the conditions of Theorem 7.1,*

$$w_n \sim C_{n-1} \frac{v_n}{(1-\rho)^{n-1}} \quad as \quad \rho \to 1 \ .$$

The Corollary should provide a better approximation than Theorem 7.1; e.g., it is exact for $n = 2$.

Proof of Theorem 7.1. Theorem 2.1 can be interpreted as convergence of ccdf's, because $h_1(t) = h_{0e}(t) = 2H_0^c(t)$. Hence we can obtain convergence of the associated moments under the regularity condition of uniform integrability; see p. 32 of [7]. The direct moment calculation from Theorem 2.1 is

$$w_n \sim nh_{1,n-1} 2^n w_1^n / (1-\rho)^{n-1} \quad \text{as} \quad \rho \to 1 \ , \tag{7.6}$$

where h_{1n} is the n^{th} moments of $h_1(t)$ in (2.3). By (10.15) of [6],

$$h_{1n} = \frac{(2n)!}{(n+1)! 2^n} \ . \tag{7.7}$$

Combining (7.6) and (7.7) gives the formula.

The extra moment condition that g_{n+2} be finite is used to establish the required uniform integrability. From (6.2), (6.4), (6.8) and (6.9) we see that $v_k, f_{\epsilon 0k}, b_{ek}$ and w_k are all finite for $k \leq n$ when $g_{k+1} < \infty$. To establish upper bounds implying uniform integrability, note that by (6.2),

$$v_k \leq \frac{g_{k+1} 2^k}{(1-\rho)^k} \ . \tag{7.8}$$

Using (7.8) and (6.8),

$$f_{\epsilon 0k} \leq \frac{K_k}{(1-\rho)^{2k}} \tag{7.9}$$

where K_k is α constant independent of ρ. Finally, by (6.9) and (7.9),

$$w_k \leq \frac{M_k}{(1-\rho)^{2k-1}} \ , \tag{7.10}$$

where M_k is a constant independent of ρ. The bounds in (7.8)–(7.10) provide the uniform integrability to go from convergence of distributions to associated convergence of moments.

Alternatively, a proof can be based on the recursion (6.9) and asymptotics for the moments b_{en}. By (7.2) and the recursion (6.8),

$$b_{en} \sim \frac{C_n n! v_1^n}{(1-\rho)^n} \quad \text{as} \quad \rho \to 1 \ . \tag{7.11}$$

From (6.9), (7.5) and (7.11), we get the conclusion.

References

[1] J. Abate, G. L. Choudhury and W. Whitt, "Calculating the M/G/1 busy-period density and LIFO waiting-time distribution by direct numerical transform inversion," *Opns. Res. Letters* 18, 113–119 (1995).

[2] J. Abate and W. Whitt, "Transient behavior of regulated Brownian motion, I: starting at the origin," *Adv. Appl. Prob.* 19, 560–598 (1987).

[3] J. Abate and W. Whitt, "Transient behavior of the M/M/1 queue via Laplace transforms," *Adv. Appl. Prob.* 20, 145–178 (1988).

[4] J. Abate and W. Whitt, "Transient behavior of the M/G/1 workload process," *Opns. Res.* 42, 750–764 (1994).

[5] J. Abate and W. Whitt, "Limits and approximations for the busy-period distribution in single-server queues," *Prob. Engr. Inf. Sci.* 9, 581–602 (1995).

[6] J. Abate and W. Whitt, "An operational calculus for probability distributions via Laplace transforms," *Adv. Appl. Prob.*, 28, 75–113 (1996).

[7] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[8] D. R. Cox and W. L. Smith, *Queues*, Methuen, London, 1961.

[9] A. Erdélyi, *Asymptotic Expansions*, Dover, New York, 1956.

[10] I. Iliadis and S. W. Fuhrmann, "Moment relationships for queues with Poisson input," *Queueing Systems* 12, 243–256 (1992).

[11] J. F. C. Kingman, "Queue disciplines in heavy traffic," *Math. Opns. Res.* 7, 262–271 (1982).

[12] F. W. J. Olver, *Asymptotics and Special Functions*, Academic Press, New York, 1974.

[13] J. Riordan, "Delays for last-come first-served service and the busy period," *Bell System Tech. J.* 40, 785–793 (1961).

[14] J. Riordan, *Stochastic Service Systems*, Wiley, New York, 1962.

[15] J. Riordan, *Combinatorial Identities*, Wiley, New York, 1968.

[16] L. Takács, "Delay distributions for one line with Poisson input, general holding times, and various orders of service," *Bell System Tech. J.* 42, 487–503 (1973).

[17] E. Vaulot, "Delais d'attente des appels telephoniques dans l'ordre inverse de leur arrive," *C. R. Acad. Sci. Paris* 238, 1188–1189 (1954).

[18] W. Whitt, Weak convergence theorems for priority queues: preemptive-resume discipline, *J. Appl. Prob.* 8 (1971) 74–94.

[19] D. M. Wishart, "Queueing systems in which the discipline is last-come first-served," *Oper. Res.* 8, 591–599 (1960).