UNIVERSITÄT HAMBURG

Weak convergence limits for sojourn times in cyclic queues under heavy traffic conditions

H. Daduna, C. Malchin , R. Szekli

Preprint–No. 2007-04 July 2007

FACHBEREICH MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

Weak convergence limits for sojourn times in cyclic queues under heavy traffic conditions

Hans Daduna *
Hamburg University
Department of Mathematics
Bundesstrasse 55
20146 Hamburg, Germany

Christian Malchin †
Hamburg University
Department of Mathematics
Bundesstrasse 55
20146 Hamburg, Germany

Ryszard Szekli ‡
Wrocław University
Mathematical Institute
pl. Grunwaldzki 2/4
50-384 Wrocław, Poland

July 6, 2007

Abstract

We consider sequences of closed cycles of exponential single server nodes with a single bottleneck. We study the cycle time and the successive sojourn times of a customer when the population sizes go to infinity. Starting from old results on the mean cycle times under heavy traffic conditions we prove a central limit theorem for the cycle time distribution. This result is then utilized to prove a weak convergence characteristic of the vector of a customer's successive sojourn times during a cycle for a sequence of networks with population sizes going to infinity. The limiting picture is a composition of a central limit theorem for the bottleneck node and an exponential limit for the unscaled sequences of sojourn times for the non-bottleneck nodes.

AMS (2000) subject classification: 60K25, 60F05

Keywords: joint sojourn time distribution, cycle times, customer stationary behaviour, bottleneck nodes, central limit theorem

1 Introduction

We study the behaviour of cyclic networks of exponential single server queues when a fixed number of nodes is filled with an ever increasing population. Such studies date back to the general model construction of Gordon and Newell [GN67]. There are two essentially different pictures.
(i) All servers have the same load: Then the total population in system is shared equally by all nodes up to random fluctuations.
(ii) Differently loaded servers exist: Then bottlenecks occur, which in the simplest case with exactly one slowest server means that almost the whole population is queued up at this slowest server.

*Work supported by Deutscher Akademischer Austauschdienst and Deutsche Forschungsgemeinschaft
†Work supported by Deutscher Akademischer Austauschdienst and Deutsche Forschungsgemeinschaft
‡Work supported by the by KBN Grant 2P03A02023

In the cyclic case the load carried by the different nodes is proportional to their mean service times. This means that in (i) (the case of balanced machines) all servers have the same service rate, and (ii) if there is exactly one server with smallest service rate then this node evolves as a unique bottleneck.

Our main interest is in the detailed travel time behaviour of individual customers in case (ii): The starting point is the steady state distribution of a customer's vector of successive sojourn times at the different nodes during a cycle. We further are interested in the customer's cycle time distribution. These distributions are known, given in the transform domain by their respective Laplace-Stieltjes-transforms ([BKK84], [SD83a]). We transform these formulae in a way that allows to prove weak convergence results for the customer's travel time behaviour when the bottleneck dominates the travel times. This was studied as the *influence of the slowest server* by Boxma [Box88].
The *usual interpretation* of the results obtained by Gordon and Newell [GN67] is that with increasing number of customers the bottleneck node approaches asymptotically a Poissonian source for the network, while all the other nodes eventually form an open ergodic tandem system, the behaviour of which is well understood: Local geometrical queue length distribution and independence over the nodes in steady state. While this is generally understood as a statement about the queue length description of the cycle, it seems to be rather obvious that a similar property should hold for cycle times and their asymptotic behaviour, respectively for the joint sojourn times of a customer during a cycle.

The *influence of the slowest server* [Box88] and Chow's [Cho80] observation about the asymptotic distributional behaviour of cycle times in a two-stage cycle suggest that there should be a standard normal limit of the scaled cycle times; this is proved in Theorem 4.1. On the other hand the *usual interpretation* suggests that even the unscaled sojourn times at the non-bottleneck nodes should converge in some sense to exponential distributions. This will come out from our main Theorem 5.1 as an immediate corollary and supports anew the *usual interpretation*.
Theorem 5.1 aggregates all sojourn times of a customer during a cycle, in the bottleck and the non-bottleneck nodes. Due to what is said before the limiting procedure should not be performed using a general scaling over all nodes. It turns out that for the bottleneck node a central limit scaling is appropriate, while for the non-bottleneck nodes no scaling is necessary.
This is different to the case of balanced machines ((i) above), where Kelly [Kel84] proved that an overall law of large numbers scaling is the appropriate one to obtain distributional limits for sojourn times, which nevertheless are not normal.

The paper is organized as follows: In Section 2 we describe the model and collect the necessary results on sojourn time distributions. In Section 3 we recall Boxma's result on the influence of the slowest server on the total cycle time and add on a somewhat curious observation about the behaviour of the joint vector of a customers sojourn times given his total cycle is known. This and the asymptotic moments of the cycle times and the local sojourn times (presented here) will be utilized in the proof of the main Theorem 5.1. In Section 4 we prove the central limit theorem for the overall cycle time when the population size grows unboundedly. The main problem in performing the proof is to suppress the strong dependencies in the summands which constitute the cycle times. The asymptotic behaviour of the joint vector of sojourn times is derived in Section 5.

2 Closed cycle with M stations

We consider a closed cyclic queueing network with M nodes (stations), where station $Q[i]$ is a single-server with infinite waiting room under first-come-first-serve (FCFS) queueing discipline, $i = 1, \ldots, M$. A fixed number $N \geq 1$ of identical customers circulate in the network: If the service of a customer at node $Q[i]$ terminates, he moves instantaneously to node $Q[i+1]$ (resp. to $Q[1]$ if a service ends at node $Q[M]$). The customers at station $Q[i]$ (if there are any) move one position forward and the service of the customer at the head of the queue begins immediately. If there are further customers present at queue $Q[i]$ at the arrival instant of the moving customer, he joins the tail of the queue; if there aren't any further customers his service begins immediately. Jumps and reorganisation of the queues take zero time. Service times of all customers at station $Q[i]$ are $\exp(\mu_i)$-distributed and are an independent family and independent of anything else. The joint queue length process of the cycle is a strong Markov process, which is irreducible and positive recurrent with steady-state distribution ([GN67])

$$\pi^{(M,N)}(n_1, \ldots, n_M) = \frac{1}{G(M,N)} \prod_{i=1}^{M} \left(\frac{1}{\mu_i}\right)^{n_i} , \tag{1}$$

with $(n_1, \ldots, n_M) \in Z(M,N) = \{(n_1, \ldots, n_M) \in \mathbb{N} : n_1 + \cdots + n_M = N\}$ and normalising constant

$$G(M,N) := \sum_{n \in Z(M,N)} \prod_{i=1}^{M} \left(\frac{1}{\mu_i}\right)^{n_i} . \tag{2}$$

We fix some test customer TC and evaluate the asymptotic and stationary characteristics of TC's sojourn times and cycle times. The starting instants of TC's cycles are his successive entrance times into station $Q[1]$. The limiting distribution of the vector of TC's successive sojourn times during his cycles is [BKK84] (Theorem 1, p. 130) determined by the Laplace-Stieltjes transform (LST)

$$\phi^{(M,N)}(\theta_1, \ldots, \theta_M) \quad = \quad \sum_{n \in Z(M,N-1)} \pi^{(M,N-1)}(n) \prod_{j=1}^{M} \left(\frac{\mu_j}{\mu_j + \theta_j}\right)^{n_j+1} , \quad 0 \leq \theta_j < 1 . \tag{3}$$

Here $\pi^{(M,N-1)}(n_1, \ldots, n_M)$ (prescribed by (1))is the steady-state distribution, that in arrival-instants of TC at $Q[1]$ there are n_1 further customers present at node $Q[1]$ (without counting TC himself), and n_j customers at the nodes $Q[j]$ for $j \in \{2, \ldots, M\}$ respectively.
The LST of the limiting distribution of TC's cycle time is (see [SD83b])

$$\psi^{(M,N)}(\theta) = \sum_{n \in Z(M,N-1)} \pi^{(M,N-1)}(n) \prod_{j=1}^{M} \left(\frac{\mu_j}{\mu_j + \theta}\right)^{n_j+1} . \tag{4}$$

The sojourn and cycle time distributions (3) and (4) are by definition limiting distributions. They can be considered as stationary distributions as well under a customer stationary regime [DS02]. Our results can be interpreted in both situations. We fix henceforth the following notation for a cycle with M nodes and N customers, including TC:
An M-dimensional vector with non negative real coordinates

$$(S_1^{(N)}, S_2^{(N)}, \ldots, S_M^{(N)}) \sim \phi^{(M,N)}(\theta_1, \ldots, \theta_M),$$

i.e., having distribution with LST $\phi^{(M,N)}(\theta_1, \dots, \theta_M)$ is considered as TC's successive sojourn times during a cycle under customer stationary conditions, and a non negative real variable

$$S(N) \sim \psi^{(M,N)}(\theta),$$

i.e., having distribution with LST $\psi^{(M,N)}(\theta)$ is considered as TC's cycle time under customer stationary conditions.

It is well known that in case that not all service rates μ_j are the same in networks with many customers cycling there will occur bottlenecks, i.e., nodes where almost all customers will queue up in the long run. The bottleneck nodes are the stations $Q[i]$ with $\mu_i = \min\{\mu_j : j = 1, \dots, M\}$.
We shall assume throughout (unless otherwise stated) that $Q[1]$ is the only bottleneck station, and, for technical reasons, that all service rates are distinct: $\mu_1 < \mu_2 < \cdots < \mu_M$. From the product form steady state and sojourn time distributions this ordering assumption does not reduce generality.

3 The Influence of the slowest server

If station $Q[1]$ is the only bottleneck of the cycle it is not surprising to TC that in case of large population almost all other customers are waiting before him at $Q[1]$ when his cycle commences. Then in particular it follows ([Box88]) that even for $\mu_1 < \mu_2 \leq \cdots \leq \mu_M$

$$E\big(S(N)\big) = N\mu_1^{-1}\big\{1 + O\big([\frac{\mu_1}{\mu_2}]^N\big)\big\}, \quad \mathrm{Var}\big(S(N)\big) = N\mu_1^{-2}\big\{1 + O\big([\frac{\mu_1}{\mu_2}]^N\big)\big\}, \quad N \to \infty \ . \tag{5}$$

From (5) obviously in heavy traffic the slowest queue generates the main fraction of the cycle time of TC. This clearly reflects the bottleneck behaviour with respect to the number of customers. So it is reasonable to approximate the distribution of the cycle time for great values of N by the sum of N consecutive service times at the slowest queue. This tempting conjecture is supported by Chow's observation that in a two-stage cycle a result in parallel to (5) holds for the LST of the cycle times as well [Cho80]. Moreover this suggests that there should hold some form of a central limit theorem for the rescaled cycle time, when the number of customers tends to infinity, while the number of stations remains fixed. This will indeed be proved in Section 4 below.
Although the quantitative results in (5) are completely in line with intuition, there is still something behind the first and second order picture expressed there. This can be expressed as a conditional invariance property of the sojourn times, for a discussion see [MD05]. We state the result here without further comments because we shall need it for the derivations later on.

Theorem 3.1 (An invariance property for the conditional sojourn time distribution)**.** *For cycles with general service rates (with or without bottlenecks) the conditional distribution of TC's successive sojourn times, given his cycle time, is independent of the number N of customers.*

For measurable and bounded $h\colon \mathbb{R}_+^M \to \mathbb{R}$ we have the conditional expectation

$$E\left(h \circ \left(S_1^{(N)},\dots,S_M^{(N)}\right) \mid S_1^{(N)}+\dots+S_M^{(N)}=t\right)$$

$$= \frac{\displaystyle\int_{\Delta_{M-1}^t} h(t-\sum_{j=2}^M x_j, x_2,..,x_M)\mu_1 e^{-\mu_1 t}\prod_{j=2}^M \mu_j e^{-x_j(\mu_j-\mu_1)}\, d(x_2,..,x_M)}{\displaystyle\int_{\Delta_{M-1}^t} \mu_1 e^{-\mu_1 t}\prod_{j=2}^M \mu_j e^{-x_j(\mu_j-\mu_1)}\, d(x_2,\dots,x_M)},$$

where for $M \in \mathbb{N}_+, t \geq 0$, we define $\Delta_M^t = \{(x_1,\dots,x_M) \in \mathbb{R}_+^M : x_1+\dots+x_M \leq t\}$.

As preparation (which is of independent interest) for to perform the analysis we state the asymptotics of moments' behaviour for cycle times (which follows directly from Boxma's results) and for local sojourn times (which is done by direct but tedious computations from (3)).

Lemma 3.1. *For pairwise distinct service rates we have the following local asymptotics of moments:*

(i) Mean cycle time and cycle time variance.

$$\lim_{N\to\infty}\frac{1}{N}E(S(N)) = \frac{1}{\mu_1}\ , \qquad \lim_{N\to\infty}\frac{1}{N}\mathrm{Var}\left(S(N)\right) = \frac{1}{\mu_1^2}\ .$$

(ii) Bottleneck mean and variance of sojourn time.

$$\lim_{N\to\infty}\frac{1}{N}E(S_1^{(N)}) = \frac{1}{\mu_1}\ , \qquad \lim_{N\to\infty}\frac{1}{N}\mathrm{Var}(S_1^{(N)}) = \frac{1}{\mu_1^2}\ .$$

(iii) Non-bottleneck mean and variance of sojourn time.

$$\lim_{N\to\infty} E(S_k^{(N)}) = \frac{1}{\mu_k-\mu_1}\ , \qquad \lim_{N\to\infty}\mathrm{Var}(S_k^{(N)}) = \frac{1}{(\mu_1-\mu_k)^2}\ .$$

(iv) Covariance of sojourn times at bottleneck and some other station.

$$\lim_{N\to\infty}\mathrm{Cov}\left(S_1^{(N)},S_l^{(N)}\right) = -\frac{1}{(\mu_1-\mu_l)^2}\ .$$

(v) Covariance of sojourn times at two non-bottleneck stations.

$$\lim_{N\to\infty}\mathrm{Cov}\left(S_k^{(N)},S_l^{(N)}\right) = 0\ , \quad k,l \neq 1\ .$$

A short comment may be in order here: From (*iv*) in Lemma 3.1 we see that the covariances between a customer's sojourn times at the bottleneck and the other non-bottleneck nodes do not vanish in the limit. In contrast to this the *usual interpretation* as described on page 2 states that in the limiting open tandem system the Poissonian source is independent from the service mechanism at the stations. By direct computations it can be shown that results similar to (*iv*) and (*v*) in Lemma 3.1 hold for the joint queue length vector as well.

4 A central limit theorem for the cycle time

As suggested by the influence of the slowest server we shall prove in this section a central limit theorem for TC's cycle time. The running index which tends to infinity in this statement will be the number of customers in the cycle.
The main problem which we have to overcome are the strong dependencies of TC's successive sojourn times during his cycle. Although the result of Boxma suggests that cycle times can be approximated by sums of independent $\exp(\mu_1)$ variables, the dependent sojourn times at the non-bottleneck nodes will come into the play. To suppress these dependencies eventually needs lengthy computations.

Theorem 4.1. *The sequence $(T(N))_{N\in\mathbb{N}}$ of the normalized and centered sojourn times*

$$T(N) := \frac{S(N) - E(S(N))}{\sqrt{\mathrm{Var}(S(N))}} \tag{6}$$

converges with increasing number of customers weakly to the Standard Normal Distribution $N(0,1)$.

Before proving the theorem we prepare some preliminary steps. For pairwise distinct service rates $\mu_i \neq \mu_j$ for $i, j \in \underline{M}$, $i \neq j$, we can write the normalizing constant in a closed form ([Har85]):

$$G(M,N) = \sum_{i=1}^{M} \left(\frac{1}{\mu_i}\right)^N \prod_{\substack{j=1\\ j\neq i}}^{M} \left(\frac{\mu_j}{\mu_j - \mu_i}\right), \tag{7}$$

We define coefficients $C_{i,N+1}$, $i \in \underline{M}$, $N \in \mathbb{N}_0$ related to $G(M,N)$ by

$$C_{i,N+1} := \frac{\left(\frac{1}{\mu_i}\right)^N \prod\limits_{\substack{j=1\\ j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}}{\sum\limits_{i=1}^{M} \left(\frac{1}{\mu_i}\right)^N \prod\limits_{\substack{j=1\\ j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}} \quad . \tag{8}$$

Lemma 4.1. *For $\mu_1 < \mu_2 < \cdots < \mu_M$ the coefficients $C_{i,N}$ have the following properties:*

$$(i) \quad \sum_{i=1}^{M} C_{i,N} = 1 \ , \qquad (ii) \quad \lim_{N\to\infty} C_{1,N} = 1 \ , \qquad (iii) \quad \lim_{N\to\infty} N^r C_{k,N+1} = 0, \quad k \neq 1, \quad r \in \mathbb{R}_+$$

$$(iv) \quad C_{k,N} > 0 \textit{ for } k \textit{ odd and } C_{k,N} < 0 \textit{ for } k \textit{ even} \ .$$

For the proof of the lemma, see the Appendix. From Harrison's formla ([Har85]) we obtain

Theorem 4.2. *The Laplace-transform $\psi^{(M,N)}$ of the cycle time $S(N)$ under steady-state conditions can in case of pairwise distinct service rates be written as*

$$\psi^{(M,N)}(\theta) = \sum_{i=1}^{M} C_{i,N} \left(\frac{\mu_i}{\mu_i + \theta}\right)^N \quad . \tag{9}$$

Proof. To apply Harrison's formula we have to ensure $\mu_k+\theta \neq \mu_l+\theta$ for $k \neq l$ and $\theta \in [0,1)$. Therefore

$$\psi^{(M,N)}(\theta) = \frac{1}{G(M,N-1)} \sum_{i=1}^{M} \left(\prod_{\substack{j=1 \\ j\neq i}}^{M} \frac{\mu_j}{\mu_j-\mu_i} \right) \left(\frac{1}{\mu_i}\right)^{N-1} \left(\frac{\mu_i}{\mu_i+\theta}\right)^{N} = \sum_{i=1}^{M} C_{i,N} \left(\frac{\mu_i}{\mu_i+\theta}\right)^{N} .$$

□

$\psi^{(M,N)}$ in (9) is not a mixture distribution because some coefficients are negative but nevertheless the n-th moment of the cycletime can be expressed as

$$E\left(S(N)^n\right) = \left(\prod_{j=1}^{n}(N+j-1)\right) \sum_{i=1}^{M} C_{i,N} \left(\frac{1}{\mu_i}\right)^{n}.$$

Hence the cycle time for N customers circulating fulfills

$$E\left(S(N)\right) = \sum_{i=1}^{M} C_{i,N} N \left(\frac{1}{\mu_i}\right), \quad \text{and} \tag{10}$$

$$\mathrm{Var}\left(S(N)\right) = N \sum_{i=1}^{M} C_{i,N} \left(\frac{1}{\mu_i}\right)^2 + N^2 \sum_{(k<l)} C_{k,N} C_{l,N} \left(\frac{1}{\mu_k} - \frac{1}{\mu_l}\right)^2. \tag{11}$$

We have now prepared the ground for the

Proof. (**Theorem 4.1**). The Fourier Transform (FT) $\mathcal{F}_{S(N)}$ of $S(N)$ is

$$\mathcal{F}_{S(N)}(x) := \int_{\mathbb{R}} e^{ixy}\, dP^{S(N)} = \sum_{j=1}^{M} C_{j,N} \left(\frac{\mu_j}{\mu_j - ix}\right)^{N}, \quad x \in \mathbb{R}.$$

Therefore the FT $\mathcal{F}_{T(N)}$ of the normalized and centered sojourn time $T(N)$ is

$$\begin{aligned} \mathcal{F}_{T(N)}(x) &= \exp\left[-\frac{ixE(S(N))}{\sqrt{\mathrm{Var}(S(N))}}\right] \cdot \mathcal{F}_{S(N)}\left(\frac{x}{\sqrt{\mathrm{Var}(S(N))}}\right) \\ &= \exp\left[-\frac{ixE(S(N))}{\sqrt{\mathrm{Var}(S(N))}}\right] \cdot \sum_{j=1}^{M} C_{j,N} \left(\frac{\mu_j}{\mu_j - \frac{ix}{\sqrt{\mathrm{Var}(S(N))}}}\right)^{N}, \quad x \in \mathbb{R}. \end{aligned}$$

Inserting $E(S(N))$ and $\mathrm{Var}(S(N))$ from (10) and (11) we obtain

$$\exp\left[-\frac{ixE(S(N))}{\sqrt{\mathrm{Var}(S(N))}}\right] = \left(\exp\left[-\frac{ix}{\sqrt{N}} \frac{C_{1,N} + \sum_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)}{\sqrt{C_{1,N} + \sum_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)^2 + N \sum_{(k<l)} C_{k,N}C_{l,N}\mu_1^2\left(\frac{1}{\mu_k}-\frac{1}{\mu_l}\right)^2}}\right]\right)^{N}$$

and

$$\frac{\mu_1}{\mu_1 - \frac{ix}{\sqrt{\mathrm{Var}(S_{(N)})}}} = \frac{1}{1 - \frac{ix}{\sqrt{N}} \frac{1}{\sqrt{C_{1,N} + \sum\limits_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)^2 + N \sum\limits_{(k<l)} C_{k,N}C_{l,N}\mu_1^2\left(\frac{1}{\mu_k} - \frac{1}{\mu_l}\right)^2}}}.$$

We show (i)

$$\lim_{N\to\infty} C_{1,N}\left(\frac{\exp\left[-\frac{ix}{\sqrt{N}}\frac{C_{1,N}+\sum\limits_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)}{\sqrt{C_{1,N}+\sum\limits_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)^2+N\sum\limits_{(k<l)} C_{k,N}C_{l,N}\mu_1^2\left(\frac{1}{\mu_k}-\frac{1}{\mu_l}\right)^2}}\right]}{1 - \frac{ix}{\sqrt{N}}\frac{1}{\sqrt{C_{1,N} + \sum\limits_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)^2 + N\sum\limits_{(k<l)} C_{k,N}C_{l,N}\mu_1^2\left(\frac{1}{\mu_k}-\frac{1}{\mu_l}\right)^2}}}\right)^N = e^{-\frac{x^2}{2}}$$

and (ii)

$$\lim_{N\to\infty} \exp\left[-\frac{ixE(S_{(N)})}{\sqrt{\mathrm{Var}(S_{(N)})}}\right] \cdot \sum_{j=2}^{M} C_{j,N}\left(\frac{\mu_j}{\mu_j - \frac{ix}{\sqrt{\mathrm{Var}(S_{(N)})}}}\right)^N = 0 \ .$$

We denote

$$f(N) := C_{1,N} + \sum_{j=2}^{M} C_{j,N}\left(\frac{\mu_1}{\mu_j}\right)^2 + \sum_{(k<l)} N C_{k,N}C_{l,N}\mu_1^2\left(\frac{1}{\mu_k} - \frac{1}{\mu_l}\right)^2 \ .$$

To prove (i), it suffices to show

$$\left|\left(\frac{\exp\left[-\frac{ix}{\sqrt{N}}\frac{C_{1,N}}{\sqrt{f(N)}}\right]}{1 - \frac{ix}{\sqrt{N}}\frac{1}{\sqrt{f(N)}}}\right)^N - \left(1 - \frac{x^2}{2N}\right)^N\right| \underset{N\to\infty}{\longrightarrow} 0 \ .$$

Recall (Lemma 4.1) that $\lim\limits_{N\to\infty} C_{1,N} = 1$, and $\lim\limits_{N\to\infty} N^r C_{j,N} = 0,\ j = 2, \ldots, M,\ r \geq 0,$ hence

$$\lim_{N\to\infty} f(N) = 1, \qquad \lim_{N\to\infty} N^r \cdot f(N) = \infty,\ r > 0.$$

Note that for every $x \in \mathbb{R}$ and N sufficiently large

$$\left|\frac{\exp\left[-\frac{ix}{\sqrt{N}}\frac{C_{1,N}}{\sqrt{f(N)}}\right]}{1 - \frac{ix}{\sqrt{N}}\frac{1}{\sqrt{f(N)}}}\right| = \frac{1}{\sqrt{1 + \frac{x^2}{N}\frac{1}{f(N)}}} \leq 1, \quad \text{and} \quad \left|1 - \frac{x^2}{2N}\right| \leq 1.$$

From $|u^n - v^n| \le n\,|u-v|$ for every $u, v \in \mathbb{C}$ with $|u| \le 1$, $v \le 1$ and $n \in \mathbb{N}$ we have

$$\left| \left(\frac{\exp\left[-\frac{ix}{\sqrt{N}} \frac{C_{1,N}}{\sqrt{f(N)}} \right]}{1 - \frac{ix}{\sqrt{N}} \frac{1}{\sqrt{f(N)}}} \right)^N - \left(1 - \frac{x^2}{2N}\right)^N \right| \le N \left| \left(\frac{\exp\left[-\frac{ix}{\sqrt{N}} \frac{C_{1,N}}{\sqrt{f(N)}} \right]}{1 - \frac{ix}{\sqrt{N}} \frac{1}{\sqrt{f(N)}}} \right) - \left(1 - \frac{x^2}{2N}\right) \right|$$

$$\le \quad N \left| \left(\frac{1 - \frac{ix}{\sqrt{N}} \frac{C_{1,N}}{\sqrt{f(N)}} - \frac{x^2 C_{1,N}^2}{2N f(N)} + \sum_{k=3}^{\infty} \left(\frac{-ix}{\sqrt{N}} \frac{C_{1,N}}{\sqrt{f(N)}} \right)^k \frac{1}{k!}}{1 - \frac{ix}{\sqrt{N}} \frac{1}{\sqrt{f(N)}}} \right) - \left(1 - \frac{x^2}{2N}\right) \right|$$

$$= \left| -\frac{ixN(1 - C_{1,N})}{ix - \sqrt{N f(N)}} - \frac{\frac{x^2 C_{1,N}^2}{2 f(N)}}{1 - \frac{ix}{\sqrt{N f(N)}}} + \frac{x^2}{2} + \frac{\sum_{k=3}^{\infty} \left(\frac{-ix C_{1,N}}{\sqrt{N^{1-\frac{2}{k}} f(N)}} \right)^k \frac{1}{k!}}{1 - \frac{ix}{\sqrt{N f(N)}}} \right| \underset{N\to\infty}{\longrightarrow} 0.$$

To prove (ii), note that

$$\left| \exp\left[-\frac{ixE(S_{(N)})}{\sqrt{\mathrm{Var}(S_{(N)})}} \right] \cdot \sum_{j=2}^{M} C_{j,N} \left(\frac{\mu_j}{\mu_j - \frac{ix}{\sqrt{\mathrm{Var}(S_{(N)})}}} \right)^N \right| = \left| \sum_{j=2}^{M} C_{j,N} \left(\frac{\mu_j}{\mu_j - \frac{ix}{\sqrt{\mathrm{Var}(S_{(N)})}}} \right)^N \right|$$

$$\le \sum_{j=2}^{M} |C_{j,N}| \left| \frac{\mu_j}{\mu_j - \frac{ix}{\sqrt{\mathrm{Var}(S_{(N)})}}} \right|^N = \sum_{j=2}^{M} |C_{j,N}| \left| \frac{1}{\sqrt{1 + \frac{x^2}{\mu_j^2 \mathrm{Var}(S_{(N)})}}} \right|^N \le \sum_{j=2}^{M} |C_{j,N}| \underset{N\to\infty}{\longrightarrow} 0 \ .$$

$\square$

5 Weak convergence limits for sojourn times

Theorem 5.1. *Consider a cycle of M nodes with N customers and with rates $\mu_1 < \cdots < \mu_M$. For TC's sojourn times $(S_1^{(N)}, S_2^{(N)}, \ldots, S_M^{(N)})$, consider the partly rescaled sequence*

$$\widetilde{S(N)} = \left(\frac{S_1^{(N)} - ES_1^{(N)}}{\sqrt{\mathrm{Var}S_1^{(N)}}}, S_2^{(N)}, \ldots, S_M^{(N)}\right), \quad N \to \infty \ . \tag{12}$$

The sequence $\widetilde{S(N)}$ converges for $N \to \infty$ to a random vector which is distributed

$$\mathcal{N}(0,1) \otimes \exp(\mu_2 - \mu_1) \otimes \cdots \otimes \exp(\mu_M - \mu_1) \ . \tag{13}$$

Proof. For the cycle with M nodes fixed and N customers we abbreviate for the cycle time the expectation and square root of the variance by

$$\nu_N := E(S(N)) \quad \text{and} \quad \sigma_N := \sqrt{\mathrm{Var} S(N)}\ . \tag{14}$$

We shall compute for $\theta_j \in \mathbb{C}, \quad j = 1, \ldots, M$

$$\lim_{N\to\infty} E\left[\exp(i\theta_1 \frac{S_1^{(N)} - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}) \cdot \prod_{j=2}^{M} \exp\left(i\theta_j S_j^{(N)}\right)\right] = e^{-\theta_1^2/2} \cdot \prod_{j=2}^{J} \frac{\mu_j - \mu_1}{\mu_j - \mu_1 - i\theta_j}\ . \tag{15}$$

Applying Theorem 3.1 we obtain (recall $\Delta_k^t := \{(x_1, \ldots, x_k) \in \mathbb{R}_+^k \mid x_1 + \cdots + x_k \le t\}$)

$$\begin{aligned}
& E\left[\exp\left(i\theta_1 \frac{S_1^{(N)} - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}\right) \cdot \prod_{j=2}^{M} \exp\left(i\theta_j S_j^{(N)}\right)\right] \\
& = \int_{\mathbb{R}} P^{T(N)}(dt) E\left[\exp\left(i\theta_1 \frac{S_1^{(N)} - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}\right) \cdot \prod_{j=2}^{M} \exp\left(i\theta_j S_j^{(N)}\right) | T(N) = t\right]
\end{aligned}$$

$$\begin{aligned}
= & \int_{\mathbb{R}} P^{T(N)}(dt) E\left[\exp\left(i\theta_1 \frac{S_1^{(N)} - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}\right) \cdot \prod_{j=2}^{M} \exp\left(i\theta_j S_j^{(N)}\right) | S(N) = t\sigma_N + \nu_N\right] \\
= & \int_{\mathbb{R}} P^{T(N)}(dt) \int_{\Delta_{M-1}^{t\sigma_N + \nu_N}} \exp\left(i\theta_1 \frac{t\sigma_N + \nu_N - \sum_{j=2}^{M} x_j - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}\right) \cdot \prod_{j=2}^{M} \exp\left(i\theta_j x_j\right) \\
& \cdot \mu_1 e^{-\mu_1(t\sigma_N + \nu_N)} \cdot \prod_{j=2}^{M} \mu_j e^{(-x_j(\mu_j - \mu_1))}(dx_2, \ldots, x_M) \\
& \cdot \left(\int_{\Delta_{M-1}^{t\sigma_N + \nu_N}} \mu_1 e^{-\mu_1(t\sigma_N + \nu_N)} \cdot \prod_{j=2}^{M} \mu_j e^{(-x_j(\mu_j - \mu_1))}(dx_2, \ldots, x_M)\right)^{-1} \\
= & \int_{\mathbb{R}} P^{T(N)}(dt) \int_{\Delta_{M-1}^{t\sigma_N + \nu_N}} \exp\left(i\theta_1 \frac{t\sigma_N + \nu_N - \sum_{j=2}^{M} x_j - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}\right) \cdot \prod_{j=2}^{M} \exp\left(i\theta_j x_j\right) \\
& \cdot \prod_{j=2}^{M} (\mu_j - \mu_1) e^{(-x_j(\mu_j - \mu_1))}(dx_2, \ldots, x_M) \\
& \cdot \left(\int_{\Delta_{M-1}^{t\sigma_N + \nu_N}} \prod_{j=2}^{M} (\mu_j - \mu_1) e^{(-x_j(\mu_j - \mu_1))}(dx_2, \ldots, x_M)\right)^{-1}
\end{aligned}$$

Let $(Y_2,\dots,Y_M)$ denote a vector with non negative coordinates which are independent with $Y_j \sim \exp(\mu_j-\mu_1)$ distributed. Then the last expression can be written as

$$\int_{\mathbb{R}} P^{T(N)}(dt)\int_{\mathbb{R}_+^{M-1}} P^{(Y_2,\dots,Y_M)}(dx_2,\dots,x_M)1_{\Delta_{M-1}^{t\sigma_N+\nu_N}}(x_2,\dots,x_M)\cdot\left(P((Y_2,\dots,Y_M)\in\Delta_{M-1}^{t\sigma_N+\nu_N})\right)^{-1}$$
$$\cdot\exp\left(i\theta_1\frac{t\sigma_N+\nu_N-\sum_{j=2}^M x_j - ES_1^{(N)}}{\sqrt{\mathrm{Var}S_1^{(N)}}}\right)\cdot\prod_{j=2}^M\exp(i\theta_j x_j).$$

For fixed $\theta_j\in\mathbb{C},\ \ j=1,\dots,M$ we define random variables

$$H(N)=H(N)(\theta_j : j=1,\dots,M):\mathbb{R}\to\mathbb{C},$$

by

$$H(N)(t)=\exp\left(\frac{i\theta_1 t\sigma_N}{\sqrt{\mathrm{Var}S_1^{(N)}}}\right)\exp\left(\frac{i\theta_1(\nu_N-ES_1^{(N)})}{\sqrt{\mathrm{Var}S_1^{(N)}}}\right)$$
$$\cdot\int_{\mathbb{R}_+^{M-1}} P^{(Y_2,\dots,Y_M)}(dx_2,\dots,x_M)1_{\Delta_{M-1}^{t\sigma_N+\nu_N}}(x_2,\dots,x_M)$$
$$\cdot\left(P((Y_2,\dots,Y_M)\in\Delta_{M-1}^{t\sigma_N+\nu_N})\right)^{-1}\cdot\exp\left(\frac{i\theta_1\sum_{j=2}^M x_j}{\sqrt{\mathrm{Var}S_1^{(N)}}}\right)\cdot\prod_{j=2}^M\exp(i\theta_j x_j),$$

and

$$H=H(\theta_j : j=1,\dots,M):\mathbb{R}\to\mathbb{C},$$

by

$$H(t)=\exp(i\theta_1 t)\cdot\int_{\mathbb{R}_+^{M-1}} P^{(Y_2,\dots,Y_M)}(dx_2,\dots,x_M)\cdot\prod_{j=2}^M\exp(i\theta_j x_j)\ .$$

By direct inspection we see that the set of points $t\in\mathbb{R}$ such that there exists a sequence $(t_N\in\mathbb{R} : N\in\mathbb{N})$ with $(t_N\to t)$ but $\neg(H(N)(t_N)\to H(t))$, has under $N(0,1)$ probability 0.
Here we used $\nu_N-ES_1^{(N)}\to 0$ and $\mathrm{Var}S_1^{(N)}\to\infty$ according to the influence of the slowest server.
Applying Theorem 5.5 in [Bil68] we conclude from weak convergence of the cycle time distributions $P^{T(N)}$ to $N(0,1)$ weak convergence of the sequence of image measures $P^{T(N)}H(N)^{-1}$ to $N(0,1)H^{-1}$.
Finally, we define the bounded continuous function

$$f:\mathbb{C}\to\mathbb{C}\ ,\quad x\to\begin{cases}x & \text{if } |x|\le 1\\ \frac{x}{|x|} & \text{if } |x|>1\end{cases}$$

and conclude

$$\int_{\mathbb{C}} P^{T(N)}H(N)^{-1}(dx)f(x)\to\int_{\mathbb{C}} N(0,1)H^{-1}(dx)f(x),\quad\text{for }\ N\to\infty.$$

But

$$\int_{\mathbb{C}} P^{T(N)} H(N)^{-1}(dx) f(x) = \int_{\mathbb{R}} P^{T(N)}(dt) f(H(N)(t))$$

$$= \int_{\mathbb{R}} P^{T(N)}(dt) \int_{\mathbb{R}_+^{M-1}} P^{(Y_2,\dots,Y_M)}(dx_2,\dots,x_M) 1_{\Delta_{M-1}^{t\sigma_N+\nu_N}}(x_2,\dots,x_M)$$

$$\cdot \left(P((Y_2,\dots,Y_M) \in \Delta_{M-1}^{t\sigma_N+\nu_N})\right)^{-1} \cdot \exp\left(i\theta_1 \frac{t\sigma_N + \nu_N - \sum_{j=2}^{M} x_j - ES_1^{(N)}}{\sqrt{\mathrm{Var} S_1^{(N)}}}\right) \cdot \prod_{j=2}^{M} \exp(i\theta_j x_j),$$

and

$$\int_{\mathbb{C}} N(0,1) H^{-1}(dx) f(x) = \int_{\mathbb{R}} N(0,1)(dt) f(H(t))$$

$$= \int_{\mathbb{R}} N(0,1)(dt) \exp(i\theta_1 t) \cdot \int_{\mathbb{R}_+^{M-1}} P^{(Y_2,\dots,Y_M)}(dx_2,\dots,x_M) \cdot \prod_{j=2}^{M} \exp(i\theta_j x_j)$$

$$= e^{-\theta_1^2/2} \cdot \prod_{j=2}^{J} \frac{\mu_j - \mu_1}{\mu_j - \mu_1 - i\theta_j}.$$

The proof is complete. □

6 Appendix

Proof. (Lemma 4.1) From $1 > \frac{\mu_1}{\mu_2} > \frac{\mu_1}{\mu_3} > \cdots > \frac{\mu_1}{\mu_M} > 0$ we conclude directly (ii):

$$\lim_{N\to\infty} C_{1,N+1} = \lim_{N\to\infty} \frac{\prod\limits_{\substack{j=1\\j\neq 1}}^{M} \frac{\mu_j}{\mu_j - \mu_1}}{\prod\limits_{j=2}^{M} \frac{\mu_j}{\mu_j - \mu_1} + \sum\limits_{i=2}^{M} \left(\frac{\mu_1}{\mu_i}\right)^N \prod\limits_{\substack{j=1\\j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}} = 1 \ .$$

In order to prove (iii) we take into account $\mu_1/\mu_k < 1$ for every $i = 2, \dots, M$, and write

$$N^r C_{k,N+1} = \frac{N^r \left(\frac{1}{\mu_k}\right)^N \prod\limits_{\substack{j=1\\j\neq k}}^{M} \frac{\mu_j}{\mu_j - \mu_k}}{\sum\limits_{i=1}^{M} \left(\frac{1}{\mu_i}\right)^N \prod\limits_{\substack{j=1\\j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}} = \frac{N^r \left(\frac{\mu_1}{\mu_k}\right)^N \prod\limits_{\substack{j=1\\j\neq k}}^{M} \frac{\mu_j}{\mu_j - \mu_k}}{\prod\limits_{j=2}^{M} \frac{\mu_j}{\mu_j - \mu_1} + \sum\limits_{i=2}^{M} \left(\frac{\mu_1}{\mu_i}\right)^N \prod\limits_{\substack{j=1\\j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}} \ .$$

Finally, the alternating signs of $\left(\frac{1}{\mu_i}\right)^N \prod\limits_{\substack{j=1\\j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}$ prove property (iv) because by (7) the denominator of the rhs of (8) is the normalizing constant G(M,N) and therefore positive. □

References

[Bil68] P. Billingsley. *Convergence of Probability Measures.* John Wiley and Sons, Inc., New York, 1968.

[BKK84] O. J. Boxma, F. P. Kelly, and A. G. Konheim. The product form for sojourn time distributions in cyclic exponential queues. *Journal of the Association for Computing Machinery*, 31:128–133, 1984.

[Box88] O. J. Boxma. Sojourn times in cyclic queues - the influence of the slowest server. In O.J. Boxma, P.J. Courtois, and G. Iazeolla, editors, *Computer Performance and Reliability*, pages 13–24. North - Holland, Amsterdam, 1988.

[Cho80] W. M. Chow. The cycle time distribution of exponential cyclic queues. *Journal of the Association for Computing Machinery*, 27:281–286, 1980.

[DS02] H. Daduna and R. Szekli. Conditional job observer property for multitype closed queueing networks. *Journal of Applied Probability*, 39:865–881, 2002.

[GN67] W.J. Gordon and G.F. Newell. Closed queueing networks with exponential servers. *Operations Research*, 15:254–265, 1967.

[Har85] P. G. Harrison. On normalizing constants in queueing networks. *Operations Research*, 33:464–468, 1985.

[Kel84] F.P. Kelly. The dependence of sojourn times in closed queueing networks. In G. Iazeolla, P.J. Courtois, and A. Hordijk, editors, *Mathematical Computer Performance and Reliability*, pages 111–121. North–Holland, Amsterdam, 1984. Proceedings of the Innternational Workshop, Pisa, Italy, 1983.

[MD05] C. Malchin and H. Daduna. An invariance property of sojourn times in cyclic networks. *Operations Research Letters*, 33:1–8, 2005.

[SD83a] R. Schassberger and H. Daduna. A discrete–time technique for solving closed queueing network models of computer systems. In Paul J. Kühn and K.M. Schulz, editors, *Messung,Modellierung und Bewertung von Rechensystemen*, pages 122–134, Berlin, 1983. Informatik–Fachberichte 61, Springer.

[SD83b] R. Schassberger and H. Daduna. The time for a roundtrip in a cycle of exponential queues. *Journal of the Association for Computing Machinery*, 30:146–150, 1983.