The Annals of Applied Probability
1996, Vol. 6, No. 1, 92–115

CONDENSATION IN LARGE CLOSED JACKSON NETWORKS

BY VADIM A. MALYSHEV AND ANDREI V. YAKOVLEV[1]

INRIA and Université d'Orleans

We consider finite closed Jackson networks with N first come, first serve nodes and M customers. In the limit $M \to \infty$, $N \to \infty$, $M/N \to \lambda > 0$, we get conditions when mean queue lengths are uniformly bounded and when there exists a node where the mean queue length tends to ∞ under the above limit (condensation phenomena, traffic jams), in terms of the limit distribution of the relative utilizations of the nodes. In the same terms, we also derive asymptotics of the partition function and of correlation functions.

1. Introduction. We consider sequences of closed Jackson networks J_N, $N = 1, 2, \ldots$. The network J_N consists of N first come, first serve (FCFS) exponential single-server fixed-rate nodes, $M = M(N) \geq 0$ indistinguishable customers and it is assigned a (probability) routing matrix $P_N = \{p_{ij,N}\}_{i,j=1}^N$ defining a finite Markov chain (MC) ($p_{ij,N}$ stands for the probability of the transition of a customer after the service at node i, to the node j; we also denote by $\mu_{i,N}$ the service rate at node i in J_N). The latter is assumed to be irreducible (consisting of one class of essential states, with no inessential states) and aperiodic.

This type of network was introduced by Jackson [11], and it gives a simple example of *product-form* networks that model, for example, many telecommunications and data-processing networks, a large class of which was described by Baskett, Chandy, Muntz and Palacios [1] ("BCMP networks"). (See also [13].)

In the quantitative and qualitative analysis of these networks, the equilibrium distribution has been of main concern. This distribution can be described in the following way. Let $\rho_N = (\rho_{1,N}, \ldots, \rho_{N,N})$ be the positive vector (unique up to a positive factor) satisfying the routing equation

$$\rho_N P_N = \rho_N. \tag{1.1}$$

Then the *relative utilizations* of the nodes of the network will be given by

$$r_{i,N} = C_N \frac{\rho_{i,N}}{\mu_{i,N}}, \qquad i = 1, \ldots, N,$$

Received January 1994; revised August 1995.

[1]This work was partially carried out while the author was receiving a scholarship from the French government.

AMS 1991 *subject classifications*. Primary 60K35, 60K30; secondary 90B15.

Key words and phrases. Closed networks, asymptotic evaluation, condensation.

for any $C_N > 0$. [Note that the right-hand sides of (1.3) and (1.4) below do not depend on C_N.] For convenience, we put

$$C_N = \max_{1\le i\le N} \frac{\mu_{i,N}}{\rho_{i,N}}$$

to obtain

$$\max_{1\le i\le N} r_{i,N} = 1. \tag{1.2}$$

Note that the vector $r_N = (r_{1,N},\dots,r_{N,N})$ is proportional to the vector of the stationary distribution of an MC with the state space $\{1,\dots,N\}$ and with the transition intensities $\{\mu_{i,N}p_{ij,N}\}_{i,j=1}^N$.

Let $\boldsymbol{\xi}_{M,N}(t) = (\xi_{1,M,N}(t), \xi_{2,M,N}(t),\dots,\xi_{N,M,N}(t))$ be the stochastic process of queue lengths J_N. Gordon and Newell [9] have shown that, for the equilibrium probability $\mathbf{P}_{M,N}$ in J_N with M customers, one has the following (product) form:

$$\begin{aligned}\mathbf{P}_{M,N}\{\xi_{1,M,N} = n_1, \xi_{2,M,N} = n_2,\dots,\xi_{N,M,N} = n_N\}\\ = \frac{1}{Z_{M,N}}\prod_{i=1}^N r_{i,N}^{n_i}, \qquad 0\le n_i\le M,\ i=1,\dots,N,\end{aligned} \tag{1.3}$$

where $Z_{M,N}$ is the normalizing constant, or *partition function* (p.f.), of J_N:

$$Z_{M,N} = \sum_{n_1+n_2+\cdots+n_N=M}\prod_{i=1}^N r_{i,N}^{n_i}, \qquad n_i\ge 0,\ i=1,\dots,N.$$

Many important performance characteristics of the network can be derived from the partition function $Z_{M,N}$. For example, to obtain moments of the queue lengths, one can differentiate it: the mean number of customers at node i is

$$m_{i,M,N} := E_{\mathbf{P}_{M,N}}\xi_{i,M,N} = \frac{r_{i,N}}{Z_{M,N}}\frac{\partial Z_{M,N}}{\partial r_{i,N}}. \tag{1.4}$$

It is important to mention that the problem of exact computation of the p.f. has been addressed by several authors. Buzen [5] has proposed a convolution algorithm for the p.f. for closed networks with FCFS constant-rate servers. Harrison [10] has found for it an explicit, closed-form expression. Gordon [8] has generalized his formula to networks with infinite servers (IS's) and with irregular constraint bounds (e.g., clustering processes). Their formulas have the most convenient form if all the relative utilizations of the nodes are different.

However, due to the computational difficulties (large required time, numerical instabilities), asymptotic expansions of the p.f., as the size of the network is large, are of greatest interest.

In the present paper, we develop a new framework for the asymptotic analysis of closed Jackson networks. The network with only FCFS nodes has been chosen mainly for the sake of simplicity of exposition of our ideas and results. For networks with infinite servers, for example, the reader will find a discussion of possible generalizations of our results in Section 6. We also believe that generalizations for networks with state-dependent service rates or multiple classes of clients are possible.

We consider J_N in the following limit:

$$M = M(N) \to \infty, \qquad M/N \to \lambda > 0 \quad \text{as } N \to \infty. \tag{1.5}$$

That is, we let the number of nodes tend to ∞, keeping the *density* $\lambda_N = M(N)/N$ asymptotically constant. Further, we do not specify explicitly the routing matrices P_N, but we require the weak convergence of the relative utilizations. Precisely, we define the following sample measure: for any Borel set A, $A \subset \mathbb{R}$, we put $I_N(A) := (1/N)\#\{i\colon r_{i,N} \in A\}$. We make the following fundamental limit assumption.

ASSUMPTION W. *As* $N \to \infty$ *the measures* I_N *converge weakly to a* (*probability*) *measure* $\mathbf{I}$ *on* $[0, 1]$.

We note that this assumption concerns the properties of solutions of the systems of linear equations (1.1) and can be related to the network topology (see Section 6 for the discussion).

The scaling (1.5) has already been considered by Knessl and Tier. In [14] they analyze closed BCMP networks with a single class of clients, an infinite server and large numbers of customers and single-server queues with state-independent service rates. Starting from the recursion relation of Buzen [5] (see also [6]), they use the ray method of geometrical optics [12] and the method of matched asymptotic expansion [2] to give asymptotic approximations of the p.f. when the servers “can be divided into classes with nearly equal relative utilizations” (see Section 2 for an analytical description of the dependence of $r_{i,N}$ on i). Knessl and Tier [15] use the same methods to study a closed BCMP network with multiple classes of clients. They use the extension of the algorithm of Buzen [5] to BCMP multiclass networks by Kobayashi and Reiser [16] (see also [18]). Mei and Tier [26] extend the results to networks with state-dependent servers and with at least one IS node.

Employing the method of generating functions together with the saddle-point method, Kogan [17] has re-deduced the results of Knessl and Tier [14], with the same assumptions about $r_{i,N}$. Birman and Kogan [3] treat, with the help of the same techniques, a class of closed product-form networks with limited-queue-dependent (LQD) service rates (two cases: one IS + a fixed number of queueing stations + a large number of customers; one IS + several groups of LQD nodes (with “slowly varying parameters” in each group) + a large number of customers). In [4] they apply the same methods to multichain closed product-form networks with several groups of stations, each group consisting of many identical stations.

The hypothesis of Knessl and Tier [14], [15] and Kogan [17] on the relative utilizations of the nodes in the network can be easily accommodated by our model (see Section 2); for networks with one class of clients and without IS nodes, our expansions coincide with those of Knessl and Tier.

In our opinion, the important advantages of our "functional" approach (Assumption W) over the "analytic" one are the following.

First, it allows us to find a criterion for instability of the sequence $\{J_N\}_{N=1}^{\infty}$: we find the critical density λ_{cr}, depending only on the measure **I**, such that for $\lambda < \lambda_{cr}$ the mean queue lengths at all nodes are uniformly bounded (Theorem 2.1). In this case we get the limit distribution for a sequence of closed Jackson networks in equilibrium. It is, of course, the product of geometrical distributions.

If $\lambda \geq \lambda_{cr}$, then, at least at the node with maximal relative utilization (r.u.), the mean queue length grows infinitely as $N \to \infty$ (Theorem 2.2). This is a traffic jam phenomenon. The rate of divergence of queue lengths can be quite arbitrary, since it depends on the way the measures I_N and the values of $r_{i,N}$ converge. In Theorem 2.3 we consider a classical example of a network with a few distinguished nodes, where the mean queue length grows linearly with N.

The knowledge of λ_{cr} is important in evaluating the sensitivity of the network w.r.t. the fluctuations of the traffic and/or the conditions of service.

We also find the asymptotics of Z_N and its first derivatives in terms of the limit measure **I**.

Second, considered as an unordered set of nodes with assigned r.u., a large Jackson network strongly resembles an ideal gas (canonical ensemble) in statistical mechanics. There is, however, a crucial difference between the two: the former can be extremely inhomogeneous in the properties of the particles which build up the system. Under Assumption W and for $\lambda \geq \lambda_{cr}$, the behavior of J_N resembles the Bose condensation phenomenon in quantum statistical physics: particles condense in the minimal energy state. This together with other peculiarities can give rise to features in the mere definition of the thermodynamic limit.

Third, we give an extension of the techniques of the "saddle-point" method [7] which could be useful in further studies of large systems.

We use in our work a representation of the p.f. and its derivatives with the help of Cauchy integrals. We then apply the saddle-point method to estimate these integrals. The necessary theory of functions of the complex variable can be found, for example, in [20].

For completeness, we mention here results on the asymptotic expansions of the p.f. for a class of large closed networks under a scaling different from ours. McKenna, Mitra and Ramakrishan [23] have introduced a method of asymptotic expansion based on the Gamma-function representation of the factorials participating in the product form and on the saddle-point estimation of the resulting integrals for networks with infinite servers. In [23], McKenna and Mitra obtain asymptotic expansions in a large parameter N, which reflects the number of customers in the system, for a network consisting of several IS nodes, each serving its own class of clients, and a common processor-sharing (PS) node. In [24] McKenna and Mitra they generalized their results for networks with multiple queueing nodes under the conditions of "normal usage." In [25] they give analogous expansions for the moments of the queue lengths. McKenna [22] has extended the methods to sojourn time problems.

The outline of the paper is as follows. In Section 2 we define the key functions and we formulate our main results. We prove in Section 3 some auxiliary results concerning the monotonicity of the network and the properties of the functions introduced in Section 1. In Section 4 we prove our results in the regular case (absence of condensation) and in Section 5, in the case of condensation. Finally, we consider some examples and we discuss possible generalizations in Section 6.

2. Main results. Let us put $\varepsilon_N = M/(N\lambda) - 1$. Denote, for $z \in \mathbb{C} \setminus [1, +\infty)$,

$$h(z) := \int_0^1 \frac{r}{1 - zr}\, d\mathbf{I}(r), \tag{2.1}$$

$$S_N(z) := -\lambda(1 + \varepsilon_N)\ln z - \frac{1}{N}\sum_{i=1}^{N} \ln(1 - zr_{i,N}), \tag{2.2}$$

$$S(z) := -\lambda \ln z - \int_0^1 \ln(1 - zr)\, d\mathbf{I}(r). \tag{2.3}$$

It is obvious that $h(z)$ is strictly monotone on $[0, 1)$. We will denote $\lambda_{\mathrm{cr}} := \lim_{z \to 1-} h(z)$.

Remark 2.1. Note that if $d\mathbf{I}(r)$ is absolutely continuous with a smooth density $f(r)$, then $\lambda_{\mathrm{cr}} < \infty$ if, for example, $f(1) = 0$.

Remark 2.2. The case where $\lambda_{\mathrm{cr}} = 0$ is, of course, possible. Then for any $\lambda > 0$ we observe the overflow behavior in the system, but we do not study it in this paper. In the sequel, we will assume that

$$\lambda_{\mathrm{cr}} > 0.$$

The main results of the paper are the following theorems.

Theorem 2.1. *Let* $\lambda < \lambda_{\mathrm{cr}}$. *Then*:

(i)

$$Z_N \sim \frac{1}{\sqrt{2\pi N S''(z_0)}} \frac{1}{z_0} \exp(N S_N(z_{0,N})), \qquad N \to \infty,$$

and asymptotically the free energy satisfies

$$F_N = \frac{1}{N} \ln Z_N \sim S(z_0);$$

(ii) $m_{i,N}$ *are uniformly bounded in* N *and* i; *that is, there exists some constant* Q *such that* $m_{i,N} < Q$, $N \geq 1$, $1 \leq i \leq N$;

(iii) *if there exists* $r_i = \lim_{N\to\infty} r_{i,N}$, $1 \le i \le K$, *for some* K, $K > 0$, *then, for any vector* $(n_1, n_2, \dots, n_K) \ge 0$,

$$\mathbf{P}_N\{\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_K = n_K\} \to \prod_{i=0}^{K} (1 - z_0 r_i)(z_0 r_i)^{n_i}, \qquad N \to \infty;$$

(iv) *if, for some* i, *there exists* $r_i = \lim_{N\to\infty} r_{i,N}$, *then*

$$m_{i,N} \to \frac{z_0 r_i}{1 - z_0 r_i}, \qquad N \to \infty,$$

where $\mathbf{P}_N\{A\}$ *is the equilibrium probability of an event* A *in the* N*th network* J_N, z_0 *is the root of the equation*

$$\frac{\partial S(z)}{\partial z} = 0 \quad \Leftrightarrow \quad h(z) = \frac{\lambda}{z} \tag{2.4}$$

such that $0 < z_0 < 1$ *and* $z_{0,N}$ *is the root of the equation*

$$\frac{\partial S_N(z)}{\partial z} = 0 \tag{2.5}$$

such that $0 < z_{0,N} < 1$.

Note that (iii) is, of course, true for any finite set of nodes, but we choose the "first" ones for the sake of simplicity of the notation.

THEOREM 2.2. *Let* $\lambda \ge \lambda_{\mathrm{cr}}$ *and let* $i(N)$ *be such that* $r_{i(N),N} = 1$. *Then* $m_{i(N),N} \to \infty$ *as* $N \to \infty$.

We remark that one of the ways to obtain approximations for finite networks from our asymptotic formulas is to put $z_0 := z_{0,N}$. The values that we obtain then correspond to the limit, where we fix N relative utilizations and take l nodes for each of these r.u.'s and then we let $l \to \infty$. In this case the condensation is absent and the approximations we obtain coincide with those of [15], formulas 3.46–3.49, for networks with one class of clients. It is shown in [15] that the accuracy of the approximations is very good. However, if there are some nodes with r.u.'s very different from all the other r.u.'s, it may be better to use the approximations for the networks with overflow (as in Theorem 2.3 below).

In [14] and [17] the condition on the r.u. is the following:

$$r_{i,N} = g\left(\frac{i}{[\lambda N]}\right), \qquad i = 1, \dots, N, \tag{2.6}$$

where $g(x)$ is finite-piecewise continuously differentiable on $[0, \lambda^{-1}]$ [assume also that $\max_{x\in[0,\lambda]} g(x) = 1$]; (2.6) gives us the following measure $\mathbf{I}$: $\mathbf{I}(A) = \lambda\nu(g^{-1}(A))$ for any Borel set A, where ν is the Lebesgue measure on $[0, 1]$. This means (but is not equivalent to) that $\mathbf{I}$ has discrete and continuous components and its support is a finite number of continuous sets. It is also easy to see that the above regularity properties of $g(\cdot)$ exclude the condensation in the network. It is clear that our framework is more general.

We also treat a classical example of overflow, traditionally considered in the literature (e.g., [14] and [17]) for the respective networks with one IS node, when there is a finite number of identical distinguished nodes having the maximal relative utilization. The critical value of the density λ for a network of this type was implicity found by Kogan [17] (see the discussion in Section 6).

Let, for some fixed K, $K > 0$, and for all N, $N \geq K$,

$$r_{1,N} = \cdots = r_{K,N} = 1. \tag{2.7}$$

Let there also exist B, $0 < B < 1$, such that

$$r_{l,N} \leq B, \qquad l > K. \tag{2.8}$$

It is clear that under these conditions $\lambda_{\mathrm{cr}} < +\infty$, and $h(z)$ is defined for $z \in \mathbb{C} \setminus [1/B, +\infty)$, and it is strictly monotone on $[0, 1/B)$. Moreover, $S(z) = -\lambda \ln z - \int_0^B \ln(1 - zr)\, d\mathbf{I}(r)$, and it can be continued to a neighborhood of $z = 1$. Note also that $S'(1) < 0$. Suppose also that

$$\lim_{z \to 1/B-} h(z) = +\infty. \tag{2.9}$$

THEOREM 2.3. *Under conditions* (2.12)–(2.14), *if* $\lambda > \lambda_{\mathrm{cr}}$, *then*:

(i)

$$Z_N \sim (-NS'(1))^{K-1} \prod_{j=K+1}^{N} \frac{1}{1 - r_{j,N}}, \qquad N \to \infty,$$

and

$$m_{i,N} \sim -N\frac{S'(1)}{K}, \qquad i = 1, \ldots, K, N \to \infty;$$

(ii) $m_{i,N}$ *are uniformly bounded in* N *and* i, $K < i \leq N$;

(iii) *if, for some* i, $i > K$, *there exists* $r_i = \lim_{N\to\infty} r_{i,N}$, *then*

$$m_{i,N} \to \frac{r_i}{1 - r_i}, \qquad N \to \infty.$$

If $\lambda < \lambda_{\mathrm{cr}}$, *then the asymptotics of Theorem* 2.1 *hold true.*

REMARK 2.3. Condition (2.9) holds, for example, if, in some neighborhood to the left of B, $d\mathbf{I}(r)/dr$ exists and $d\mathbf{I}(r)/dr > L$ for some L, $L > 0$.

In the above example, the maximal mean queue length is of order $O(N)$, and the condition of overflow is similar to that obtained in [17]. Note, however, that the overflow can be of much more general character, and Theorems 2.1 and 2.2 together give its *criterion*. Let us have the same network as in Theorem 2.3 but with $K = K(N) \to \infty$ varying. A natural conjecture would be that, if $K(N)/N \to 0$ as $N \to \infty$ and $\lambda > \lambda_{\mathrm{cr}}$, then $m_{1,N} = O(N/K(N))$. Note that the classical techniques of the saddle-point (s.p.) method will not work in this situation, as the order of the pole coalescing with the s.p. contour grows infinitely with N. The analysis can be again more complicated if, for example, the r.u.'s are different but they are approaching 1 as $N \to \infty$ and so on.

3. Auxiliary results. Recall that $r_i \neq 0$, $i = 1, \dots, N$. We will need the following result on the *monotonicity of the network*.

LEMMA 3.1. *For any* $M_2 \geq M_1 \geq 0$, *any* N, $N \geq 1$, j, $1 \leq j \leq N$, *we have* $m_{i, M_2, N} \geq m_{i, M_1, N}$.

PROOF. We will prove the lemma using the stochastic monotonicity of $\boldsymbol{\xi}_{M,N}(t)$. We note that another quite simple proof can be done using the so-called arrival theorem for closed networks (see, e.g., [19]).

Let us introduce the following partial ordering on $E = Z_+^N$:

$$\{x \leq y\} \equiv \{x_i \leq y_i,\ i = 1, \dots, N\}, \qquad x, y \in E.$$

We call Γ an *increasing set* iff $\Gamma = \{y | y \geq x \text{ for some } x \in \Gamma\}$. The family of increasing sets $\mathscr{T}(E)$ induces the *strong stochastic order* $\leq_{\text{st}}$ (cf. [21]): for two probability measures P, Q on E,

$$\{P \leq_{\text{st}} Q\} \equiv \{P(\Gamma) \leq Q(\Gamma)\ \forall\, \Gamma \in \mathscr{T}(E)\}.$$

Consider now the generator A of the family $\{\boldsymbol{\xi}_{M,N}\}_{M=0}^{\infty}$. It is the sum of the respective generators of $\boldsymbol{\xi}_{M,N}$. Note that, as each network is closed, A is a reducible generator. For A we have

$$A = \sum_{i,j=1,\dots,N} a_{ij}\big(\boldsymbol{\Phi}(\mathbf{f}_{ij}) - \mathbf{I}\big),$$

where

$$a_{ij} = \mu_{i,N} p_{ij,N}, \qquad i, j = 1, \dots, N,$$

$$\mathbf{f}_{ij}\colon E \to E, \qquad \mathbf{f}_{ij}(x) = \begin{cases} (\dots, x_i - 1, \dots, x_j + 1, \dots), & x_i > 0, \\ x, & x_i = 0, \end{cases}$$

$$\mathbf{I}\colon E \to E, \qquad \mathbf{I}(x) = x,$$

$$\boldsymbol{\Phi}(\mathbf{f}) = \mathbf{f}^{-1}, \qquad \mathbf{f}\colon E \to E.$$

It is easy to show that the conditions of Theorem 4.1 of Massey [21] are verified and therefore A is a $\mathscr{T}(E)$-monotone generator.

Let us consider $\boldsymbol{\xi}_{M_1,N}, \boldsymbol{\xi}_{M_2,N}$ with the initial conditions: $\boldsymbol{\xi}_{M_1,N}(0) = (M_1, 0, \dots, 0)$, $\boldsymbol{\xi}_{M_2,N}(0) = (M_2, 0, \dots, 0)$. Then $\boldsymbol{\xi}_{M_1,N}(0) \leq_{\text{st}} \boldsymbol{\xi}_{M_2,N}(0)$ and both processes have the same $\mathscr{T}(E)$-monotone generator. Hence, by Theorem 3.4 of [21], $\xi_{M_1,N}(t) \leq_{\text{st}} \xi_{M_2,N}(t)$ for any $t \geq 0$, and therefore $\mathbf{P}_{M_1,N} \leq_{\text{st}} \mathbf{P}_{M_2,N}$.

Now let $g_i(x) = x_i$. Then

$$\begin{aligned} m_{i,M,N} &= E_{\mathbf{P}_{M,N}} g_i(\boldsymbol{\xi}_{M,N}) \\ &= \sum_{j=1}^{\infty} j \mathbf{P}_{M,N}\{\xi_{i,MN} = j\} = \sum_{j=1}^{\infty} \mathbf{P}_{M,N}\{\xi_{i,MN} \geq j\}. \end{aligned}$$

As the sets $\{x_i \geq j\} \in \mathcal{F}(E)$, we get the desired result. □

To evaluate $Z_{M,N}$, for every fixed N we consider first the grand partition function (the generating function of the sequence of networks)

$$\Xi_N(z) := \sum_{M=0}^{\infty} z^M Z_{M,N} = \prod_{i=1}^{N} \frac{1}{1 - zr_{i,N}} \quad \text{for } |z| < 1. \tag{3.1}$$

In accordance with the Cauchy formula on residues,

$$Z_{M,N} = \frac{1}{2\pi i} \int_\gamma \frac{\Xi_N(z)}{z^{M+1}}\, dz,$$

where γ is a circle of radius less than 1 centered at 0, or, using the notation of (2.2),

$$Z_N = \frac{1}{2\pi i} \int_\gamma \frac{1}{z} \exp(NS_N(z))\, dz. \tag{3.2}$$

Note that the integrand is a meromorphic function with poles $z_{i,N} = 1/r_{i,N} \geq 1$, $i = 1, \dots, N$.

For the means, we have

$$m_{i,N} = \frac{r_{i,N}}{Z_N} \frac{\partial Z_N}{\partial r_{i,N}} = \frac{1}{Z_N} \frac{1}{2\pi i} \int_\gamma \frac{r_{i,N}}{1 - zr_{i,N}} \exp(NS_N(z))\, dz. \tag{3.3}$$

For the joint distributions of the queue lengths, we have

$$\mathbf{P}_N\{\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_K = n_K\} = \frac{1}{Z_N} r_{1,N}^{n_1} r_{2,N}^{n_2} \cdots r_{K,N}^{n_K} Z^K_{M - \Sigma_{i=1}^K n_i, N},$$

where

$$Z^K_{L,N} = \sum_{n_{K+1} + n_{K+2} + \cdots + n_N = L} \prod_{i=K+1}^{N} r_{i,N}^{n_i}.$$

We consider again the generating function of the sequence $\{Z^K_{L,N}\}_{L=0}^\infty$:

$$\Xi_N^K(z) = \sum_{L=0}^{\infty} z^L Z^K_{L,N} = \prod_{i=K+1}^{N} \frac{1}{1 - zr_{i,N}}, \qquad |z| < 1.$$

Using again the Cauchy formula for $Z^K_{L,N}$, we obtain

$$\begin{aligned}
&\mathbf{P}_N\{\xi_1 = n_1, \xi_2 = n_2, \dots, \xi_K = n_K\} \\
&= \frac{r_{1,N}^{n_1} r_{2,N}^{n_2} \cdots r_{K,N}^{n_K}}{Z_N 2\pi i} \int_\gamma z^{-M + \Sigma_{i=1}^K n_i + 1} \prod_{i=K+1}^{N} \frac{1}{1 - zr_{i,N}}\, dz \\
&= \frac{1}{Z_N} \frac{1}{2\pi i} \int_\gamma z^{-M-1} \prod_{i=1}^{K} (1 - zr_{i,N})(zr_{i,N})^{n_i} \prod_{i=K+1}^{N} \frac{1}{1 - zr_{i,N}}\, dz \\
&= \frac{1}{Z_N} \frac{1}{2\pi i} \int_\gamma z^{-1} \prod_{i=1}^{K} (1 - zr_{i,N})(zr_{i,N})^{n_i} \exp(NS_N(z))\, dz.
\end{aligned} \tag{3.4}$$

LEMMA 3.2. *All roots of the equation*

$$\frac{\partial S_N(z)}{\partial z} = 0 \tag{3.5}$$

are real and positive. There is always one and only one root in $(0, 1)$.

PROOF. Rewriting (3.19), we obtain

$$0 = \frac{\partial S_N(z)}{\partial z} = -\frac{\lambda(1+\varepsilon_N)}{z} + \frac{1}{N}\sum_{i=1}^{N} \frac{r_{i,N}}{1 - z r_{i,N}}. \tag{3.6}$$

It is easy to see that, for $\operatorname{Im} z \neq 0$, the imaginary parts of all terms in (3.6) have the same sign. Hence, the roots of (3.6) can only be real. Note also that all terms in (3.6) are positive if $z < 0$, which proves that the roots cannot be negative.

Let us consider the interval $(0, 1)$. All terms of (3.6) are strictly monotone on $(0, 1)$ and their sum has infinite limits of different signs at its ends [at 0—evident, at 1—due to (1.2)]. Therefore, the lemma is proved. □

LEMMA 3.3. *For any fixed* λ, $\lambda > 0$, *there exists* $\lim_{N\to\infty} z_{0,N} =: z_0 = z_0(\lambda) > 0$. *It has the following properties*:

(i) *if* $\lambda < \lambda_{\mathrm{cr}}$, *then* $z_0(\lambda)$ *is the root of the equation*

$$h(z) = \frac{\lambda}{z}, \tag{3.7}$$

$z_0(\lambda)$ *is strictly monotone and* $0 < z_0(\lambda) < 1$, $\lim_{\lambda\to\lambda\mathrm{cr}-} z_0(\lambda) = 1$;

(ii) *if* $\lambda \geq \lambda_{\mathrm{cr}}$, *then* $z_0 = 1$.

PROOF. (i) $\lambda < \lambda_{\mathrm{cr}}$. Note that $zh(z)$ is strictly monotone on $[0, 1)$. Since $\lambda < \lim_{z\to 1-} h(z) = \lim_{z\to 1-} zh(z)$, (3.7) obviously has a unique solution $z_0 = z_0(\lambda) \in (0, 1)$, strictly monotone in λ. Hence, there exists $\lim_{\lambda\to\lambda_0-} z_0(\lambda)$. It is easy to see that it equals 1.

Let us show that $\lim_{N\to\infty} z_{0,N}, = z_0$. For any N, let $g_N(z) := \partial S_N(z)/\partial z$, which is a monotonically increasing [see (3.6)] continuous function on $(0, 1)$; $g(z) := h(z) - \lambda/z$ is also a monotonically increasing function on the same set of z; and, for any z, $z \in (0, 1)$, $g_N(z) \to g(z)$ as $N \to \infty$.

Take any small interval around z_0: $(z_0 - \varepsilon, z_0 + \varepsilon)$. Then $g(z)$ will be different from 0 at the endpoints of that interval, and will have different signs. But, for N large, $g_N(z)$ will have the same signs at the respective points. Hence, the interval will contain $z_{0,N}$.

(ii) $\lambda \geq \lambda_{\mathrm{cr}}$. It is obvious that $\limsup_{N\to\infty} z_{0,N} \leq 1$ (since $z_{0,N} \leq 1/r_{i,N}$, $1 \leq i \leq N$). For arbitrary $\varepsilon > 0$ sufficiently small, $g(1-\varepsilon) < \lim_{z\to 1-} g(z) \leq 0$. For N sufficiently large, $g_N(1-\varepsilon)$ will also be negative for z, $0 < z \leq 1 - \varepsilon$, that is, $z_{0,N} \in (1 - \varepsilon, 1)$. □

LEMMA 3.4. *The quantity* $\operatorname{Re} S_N(r\exp(i\varphi))$ *monotonically decreases when* $\varphi \in (0, \pi)$ *and monotonically increases when* $\varphi \in (\pi, 2\pi)$, $r > 0$.

PROOF. Indeed,

$$\exp(N \operatorname{Re} S_N(z)) = \frac{1}{|z|^{N\lambda(1+\varepsilon_N)}} \prod_{i=1}^{N} \frac{1}{|1 - zr_{i,N}|}. \tag{3.8}$$

The first factor on the right-hand side of (3.8) has constant value on the curve $r\exp(i\varphi)$. Let $d_N(\varphi) = |1 - rr_{i,N}\exp(i\varphi)|^2$. Then $d'_N(\varphi) = 2rr_{i,N}\sin\varphi$, so the denominators in (3.8) increase when $0 < \varphi < \pi$ and decrease when $\pi < \varphi < 2\pi$. □

4. Proof of the main results in the regular case. We will first prove a theorem slightly more general than Theorem 2.1, from which the latter will follow.

In the sequel, we shall denote $U_d(v) := \{z \in \mathbb{C}\colon |z - v| < d\}$, $\gamma := \{z \in \mathbb{C}\colon |z| = z_0(\lambda)\}$.

THEOREM 4.4 *Assume*:

(I) $\lambda < \lambda_{\rm cr}$;

(II) $\{f(\theta, z)\}_{\theta\in\Theta}$ *is a family of functions, holomorphic in the ring* $\{z \in \mathbb{C}\colon z_0(\lambda) - \sigma_0 < |z| < z_0(\lambda) + \sigma_0\}$ *for some* $\sigma_0 > 0$, *and uniformly bounded in that ring, and such that, for a given* ε *sufficiently small*, $\varepsilon > 0$, *there exists* σ_u, $\sigma_u > 0$, *such that* $|f(\theta, z)/f_u - 1| < \varepsilon$, $z \in U_{2\sigma_u}(z_0)$, $\theta \in \Theta$, *for some constant* $f_u \in \mathbb{R} \setminus \{0\}$.

Then there exists N_ε *such that, for any* N, $N > N_\varepsilon$, $\theta \in \Theta$,

$$\begin{aligned} L_N(f) &:= \frac{1}{2\pi i}\int_\gamma f(\theta, z)\exp(NS_N(z))\,dz \\ &= \frac{f_u}{\sqrt{2\pi NS''(z_0)}}\exp(NS_N(z_{0,N}))(1 + \zeta_N), \qquad \zeta_N \in \mathbb{R}, |\zeta_N| < 25\varepsilon. \end{aligned}$$

PROOF. Essentially, we will apply to the integral along γ the usual techniques of the saddle-point method. But the way $S_N(z)$ depends on N makes a generalization of the method necessary.

From the convergence of $z_{0,N}$ to z_0 (see Lemma 3.3), we obviously have the following result.

PROPOSITION 4.1. *For any* σ, $\sigma > 0$, *there exists* N' *such that, for any* N, $N > N'$, $U_\sigma(z_{0,N}) \subset U_{2\sigma}(z_0)$.

LEMMA 4.5. *For any given* ε *sufficiently small*, $\varepsilon > 0$, *and any given* σ_u, $\sigma_u > 0$, *there exist* N' *and* σ_r, $\sigma_u \ge \sigma_r > 0$, *such that the following seven properties hold true for* $N > N'$ *and for* $z \in U_{\sigma_r}(z_{0,N})$, *where applicable*:

(a) $S_N(z)$ *can be expanded into a power series in* $U_{\sigma_r}(z_{0,N})$:

(4.1) $$S_N(z) = S_N(z_{0,N}) + \frac{(z - z_{0,N})^2}{2} S''_N(z_{0,N})(1 + R_{2,N}(z)),$$

where

$$R_{2,N}(z) = \frac{2}{S''_N(z_{0,N})} \sum_{k=3}^{+\infty} \frac{S_N^{(k)}(z_{0,N})}{k!} (z - z_{0,N})^{k-2};$$

(b) $S''_N(z_{0,N}) > F$ *for some constant* F, $F > 0$;

(c) $\left| \sqrt{\frac{S''(z_0)}{S''_N(z)}} - 1 \right| < \varepsilon$;

(d) $|R_{2,N}(z)| < \varepsilon$;

(e) $|(z - z_{0,N}) R'_{2,N}(z)| < \varepsilon$;

(f) $\left| \frac{\operatorname{Im} R_{2,N}(z)}{1 + \operatorname{Re} R_{2,N}(z)} \right| < 1$;

(g) $z_{0,N} \in (r_0 + \sigma_r, z_0 + \sigma_0 - \sigma_r)$.

PROOF. To ensure (a) in $U_{\sigma_a}(z_{0,N})$, is is sufficient to require that $0 < \sigma_a < (1 - z_0)/2$ for $\lambda < \lambda_{\text{cr}}$ implies $z_0 < 1$. Then any $S_N(z)$ can be expanded into the power series when $z \in U_{\sigma_a}(z_{0,N})$:

$$S_N(z) = S_N(z_{0,N}) + \sum_{k=2}^{+\infty} \frac{S_N^{(k)}(z_{0,N})}{k!} (z - z_{0,N})^k,$$

all $S_N^{(n)}(z_{0,N}) \in \mathbb{R}$,

$$S_N^{(n)}(z) = (-1)^n (n-1)!\, \lambda (1 + \varepsilon_N) z^{-n} + (n-1)!\, N^{-1} \sum_{i=1}^{N} \left(r_{i,N}^{-1} - z\right)^{-n}.$$

REMARK 4.4. It is easy to see that there exists N'_a such that $z_{0,N} \in U_{\sigma_a/2}(z_0)$ for $N > N'_a$, and the above expansions are valid in $U_{\alpha_a/2}(z_0)$.

Note that for any n there exists

$$\begin{aligned} \lim_{N \to \infty} S_N^{(n)}(z) &=: S^{(n)}(z) \\ &= (-1)^n (n-1)!\, \lambda z^{-n} \\ &\quad + (n-1)! \int_0^1 \left(r_{i,N}^{-1} - z\right)^{-n} d\mathbf{I}(r), \end{aligned} \tag{4.2}$$

uniformly on every compact set outside $[1, +\infty)$, and hence for $z \in \overline{U}_{\sigma_a/2}(z_0)$. Hence, for any n there exists

$$\lim_{N \to \infty} S_N^{(n)}(z_{0,N}) = S^{(n)}(z_0) \quad \text{and} \quad S''(z_0) > 0. \tag{4.3}$$

This obviously implies (b).

Using the uniform convergence of $S''_N(z)$ in $U_{\sigma_a/2}(z_0)$, $N \to \infty$, and that of (4.3), we can choose σ_c to be small enough, so that

$$\left| \sqrt{\frac{S''(z_0)}{S''_N(z)}} - 1 \right| < \frac{\varepsilon}{2}, \qquad z \in U_{2\sigma_c}(z_0).$$

From Proposition 4.1 it follows that there exists N'_c such that (c) is also true if $N > N'_c$.

From (4.2) it follows that there exist N'_d and M, $M > 0$, such that $|S_N(z)| < M$ for $z = z_{0,N} + \mu \exp(i\varphi)$, $\varphi \in [0, 2\pi)$, for some small μ and $N > N'_d$. Then $|S_N^{(n)}(z_{0,N})/n!| \le M/\mu^n$. Let $|z - z_{0,N}| \le \sigma < \mu$. Then

$$\left| \sum_{k=3}^{+\infty} \frac{S_N^{(k)}(z_{0,N})}{k!} (z - z_{0,N})^{k-2} \right| \le M \frac{\sigma}{\mu^3} \frac{1}{1 - \sigma/\mu}$$

and therefore $|R_{2,N}(z)| \to 0$ if $|z - z_{0,N}| \to 0$, $N > N'_d$, which yields (d).

To ensure (e), we can just evaluate

$$\begin{aligned} |(z - z_{0,N}) R'_{2,N}(z)| &= \frac{2}{|S''_N(z_{0,N})|} \left| \sum_{k=3}^{+\infty} \frac{(k-2) S_N^{(k)}(z_{0,N})}{k!} (z - z_{0,N})^{k-2} \right| \\ &\le \frac{2M}{|S''_N(z_{0,N})| \lambda^2} \sum_{k=3}^{+\infty} (k-2) \left(\frac{\sigma}{\lambda} \right)^{k-2}, \\ &\qquad |z - z_{0,N}| < \sigma \le \mu,\ N > N'_d, \end{aligned}$$

and, by choosing appropriate $N'_e \ge N'_d$ and appropriate small σ_e, we can provide (e).

Property (i) follows from (d), with some σ_f, N'_f.

It is obvious that property (g) can be easily satisfied, too, with some σ_g, N'_g.

Now, by taking $\sigma_r = \min(\sigma_a, \dots, \sigma_g)$ and $N' = \max(N'_a, \dots, N'_g)$, we conclude the proof. □

Now we proceed to the proof of the theorem. Under its conditions, the conditions of Lemma 4.5 are satisfied, and we will refer to the properties proved therein simply by (a)–(g).

First, we deform the integration contour γ into the new integration contour γ_N, depending on N, in the following way. From the two level curves of the function Im $S_N(z)$ passing through $z_{0,N}$: Im $S_N(z) = \text{Im } S_N(z_{0,N}) = 0$, we choose the one orthogonal to the real axis of the z-plane. As the only finite singularity points of $S_N(z)$ are $0, 1/r_{i,N}$, $1 \le i \le N$, this level curve goes out of $U_\sigma(z_{0,N})$, crossing its boundary at two points, symmetrical with respect to the real axis. We denote the intersection of this curve with $U_\sigma(z_{0,N})$ by $\gamma_{\sigma,N}$, and by $z_{1,N}$, $z_{2,N}$ we denote the ends of $\gamma_{\sigma,N}$. Then we connect $z_{1,N}$ and $z_{2,N}$ by an arc centered at $z = 0$, which is located outside of $U_\sigma(z_{0,N})$, thus obtaining the integration curve γ_N since lies inside of the domain of analyticity of the integrands by (g) and by the conditions of the theorem, it would give the same integral values as those counted along γ.

Now we will evaluate the integral along the part of γ_N outside of $U_\sigma(z_{0,N})$. But $z_{j,N} - z_{0,N} = \sigma \exp(i\varphi_j)$, $j = 1,2$, and $\operatorname{Im}((z_j - z_{0,N})^2(1 + R_{2,N}(z_j))) = 0$, $j = 1,2$, from which $\tan(2\varphi_j) = -\operatorname{Im} R_{2,N}(z_j)/(1 + \operatorname{Re} R_{2,N}(z_j))$. From (b) and (4.23) and (f), we can prove that there exists C, $C > 0$, such that, for any N, $N > N'$, $\operatorname{Re} S_N(z_{j,N}) < \operatorname{Re} S_N(z_{0,N}) - C$, $j = 1,2$.

If we take into account the uniform boundedness of $\{f(\theta, z)\}_{\theta\in\Theta}$, we obtain

$$(4.4) \qquad \left| \int_{\gamma_N \setminus \gamma_{\sigma,N}} f(\theta, z) \exp(NS_N(z))\, dz \right| \le G \exp(N(S_N(z_{0,N}) - C))$$

for some $C > 0$, $G > 0$ and any $N > N'$, $\theta \in \Theta$.

Now we proceed to an evaluation of the main part (along $\gamma_{\sigma,N}$) of the integral.

Let us introduce

$$\psi_N(z) = (z - z_{0,N})\sqrt{S_N''(z_{0,N})}\sqrt{1 + R_{2,N}(z)}, \qquad z \in U_{\sigma_r}(z_{0,N}),$$

where the positive branch of the square root is used. It is obvious that $\psi_N'(z_{0,N}) = \sqrt{S_N''(z_{0,N})}$ and that $-i\psi_N(z) \in \mathbb{R}$, $z \in \gamma_{\sigma,N}$.

From (d) it follows that $\psi_N(z)$ in $U_\sigma(z_{0,N})$ has only one 0 at $z_{0,N}$ and that

$$\frac{1}{2\pi i}\int_{\gamma_s} \frac{\psi_N'(z)}{\psi_N(z)}\, dz = 1,$$

where $\gamma_s = \{z\colon z = z_{0,N} + (\sigma_r/2)\exp(i\varphi),\ 0 \le \varphi < 2\pi\}$.

Let $\delta = \inf_{|z - z_{0,N}| = \sigma_r/2,\ N > N'} |\psi_N(z)|$. Note that $\delta > 0$ due to (c). And let $\sigma = \min(\sigma_r/2, \inf_{N > N'} \sup_{\sigma' > 0}\{\sigma'\colon\ \sigma' \exp(i\varphi) \in \psi_N^{-1}(U_\delta(0)),\ \ 0 \le \varphi < 2\pi\})$. Note also that $\sigma > 0$. Then, for any $z' \in U_\sigma(z_{0,N})$,

$$\frac{1}{2\pi i}\int_{\gamma_s} \frac{\psi_N'(z)}{\psi_N(z) - \psi_N(z')}\, dz = 1,$$

which, due to the argument principle, implies that there are no other points $z'' \in U_\sigma(z_{0,N})$: $\psi_N(z'') = \psi_N(z')$. Hence, $\psi_N(z)$ performs the analytical isomorphism of $U_\sigma(z_{0,N})$, $N > N'$, into some neighborhood of 0.

Let us put $z = \psi_N^{-1}(w)$, $w \in \psi_N(U_\sigma(z_{0,N}))$. Then

$$\begin{aligned}
&\frac{1}{2\pi i}\int_{\gamma_{\sigma,N}} f(\theta, z)\exp(NS_N(z))\, dz \\
&\quad = \frac{1}{2\pi i}\exp(NS_N(z_{0,N}))\int_{\psi_N(\gamma_{\sigma,N})} f(\theta, \psi_N^{-1}(w))\exp\left(\frac{Nw^2}{2}\right)\frac{dw}{\psi_N'(\psi_N^{-1}(w))} \\
&\quad = \frac{1}{2\pi}\exp(NS_N(z_{0,N}))I_N,
\end{aligned}$$

$$I_N := \int_{-i\psi_N(\gamma_{\sigma,N})} \exp\left(-\frac{Nu^2}{2}\right)\frac{f(\theta, z_N(u))}{\psi_N'(z_N(u))}\, du,$$

where $w = iu$ and $z_N(u) := \psi_N^{-1}(iu)$. Note again that $u \in \mathbb{R}$. But

$$\frac{1}{\psi_N'(z_N(u))} = \frac{\sqrt{1 + R_{2,N}(z_N(u))}}{\sqrt{S_N''(z_{0,N})}\left[1 + R_{2,N}(z_N(u)) + \frac{1}{2}(z_N(u) - z_{0,N})R_{2,N}'(z_N(u))\right]}.$$

From (c)–(e) together with condition (II) of the theorem, it follows that

$$\frac{f(\theta, z_N(u))}{\psi_N'(z_N(u))} = \frac{f_u}{\sqrt{S''(z_0)}}(1 + \zeta_{1,N}(u)),$$
$$\zeta_{1,N}(u) \in \mathbb{C}, |\zeta_{1,N}(u)| < 10\varepsilon, iu \in \psi_N(U_\sigma(z_{0,N})), \theta \in \Theta$$

(as ε is small). So, for $N > N'$, $\theta \in \Theta$,

$$I_N = \int_{-i\psi_N(\gamma_{\sigma,N})} \exp\left(-\frac{Nu^2}{2}\right) du \frac{f_u}{\sqrt{S''(z_0)}}(1 + \zeta_{2,N}), \tag{4.5}$$
$$\zeta_{2,N} \in \mathbb{C}, |\zeta_{2,N}| < 15\varepsilon.$$

But $-i\psi_N(z_{j,N}) = (z_{j,N} - z_{0,N})\sqrt{S_N''(z_{0,N})}\sqrt{1 + R_{2,N}(z_{j,N})}$, $j = 1, 2$, and due to (c), (d) and (f), for small ε and for $N > N'$,

$$\frac{\sigma}{2}D < |-i\psi_N(z_{j,N})| < 2\sigma D, \qquad j = 1, 2, \; D = \sqrt{S''(z_0)}.$$

So the last integral in (4.5) is bounded by the integrals of $\exp(-Nu^2/2)$ on $[-\sigma D/2, \sigma D/2]$ and $[-2\sigma D, 2\sigma D]$, the latter two being equivalent to $\sqrt{2\pi/N}$ as $N \to \infty$. Hence, we can conclude that there exists N'', $N'' \geq N'$, such that, for any N, $N > N''$,

$$\int_{-\psi_N(\gamma_{\sigma,N})} \exp\left(-\frac{Nu^2}{2}\right) du = \sqrt{\frac{2\pi}{N}}(1 + \zeta_{3,N}),$$
$$\zeta_{3,N} \in \mathbb{R}, |\zeta_{3,N}| < \varepsilon, N > N'',$$

and therefore

$$I_N = \sqrt{\frac{2\pi}{N}} \frac{f_u}{\sqrt{S''(z_0)}}(1 + \zeta_{4,N}), \qquad \zeta_{4,N} \in \mathbb{C}, |\zeta_{4,N}| < 20\varepsilon, N > N'', \theta \in \Theta.$$

From this estimate together with (4.4), it follows that the integrals along $\gamma_N \setminus \gamma_{\sigma,N}$ are exponentially small compared to those along $\gamma_{\sigma,N}$; that is, Theorem 4.4 is proved. □

Now we will prove the statements of Theorem 2.1. Take $f(\theta, z) = f(z) = 1/z$, $f_u = 1/z_0$. Then, for any small ε, $\varepsilon > 0$, we can find small σ_u, $\sigma_u > 0$, such that condition (II) of Theorem 4.4 is satisfied. Therefore, for any ε sufficiently small, $\varepsilon > 0$, $\exists N_\varepsilon: \forall N,\ N > N_\varepsilon$,

$$Z_N = \frac{1}{\sqrt{2\pi N S''(z_0)}} \frac{1}{z_0} \exp(NS_N(z_{0,N}))(1 + \zeta_N), \qquad |\zeta_N| < 25\varepsilon,\ N > N_\varepsilon,$$

which proves statement (i) of Theorem 2.1.

To prove (ii), we fix some small ε, $\varepsilon > 0$, and assume $f(\theta, z) = A/z + \theta/(1 - z\theta)$, $\theta \in \Theta = [0, 1]$, $A > 0$, $f_u = A/z_0$. The $f(\theta, z)/f_u = z_0/z + (z_0\theta/(1 - z\theta))/A$. By choosing $\sigma_u = \varepsilon/8$ and $A = (16z_0/(1 - z_0))/\varepsilon$, we can ensure condition (II) of Theorem 4.4.

Then we obtain that $\exists N_\varepsilon: \forall N,\ N > N_\varepsilon,\ \theta \in \Theta$,

$$\frac{1}{2\pi i}\int_\gamma \left(\frac{A}{z} + \frac{\theta}{1 - z\theta}\right)\exp(NS_N(z))\,dz$$
$$= \frac{1}{\sqrt{2\pi N S''(z_0)}} \frac{A}{z_0} \exp(NS_N(z_{0,N}))(1 + \zeta_N), \qquad \zeta_N \in \mathbb{R},\ |\zeta_N| < 25\varepsilon.$$

After dividing by Z_N and applying the results of (i) to the right-hand side of the resulting equality, we obtain that, $\exists N'_\varepsilon: \forall N,\ N > N'_\varepsilon$,

$$A + \frac{1}{Z_N}\frac{1}{2\pi i}\int_\gamma \frac{\theta}{1 - z\theta}\exp(NS_N(z))\,dz = A(1 + \zeta'_N), \qquad |\zeta'_N| < 30\varepsilon.$$

As $m_{i,N} = Z_N^{-1} L_N(f_{i,N})$ with $f_{i,N}(z) = r_{i,N}/(1 - zr_{i,N})$ and $A/z + f_{i,N}(z) = f(\theta, z)|_{\theta = r_{i,N}} \in \{f(\theta, z)\}_{\theta \in \Theta}$, the above evaluation implies that

$$|m_{i,N}| < 30A\varepsilon, \qquad N > N'_\varepsilon,\ 1 \le i \le N,$$

which proves statement (ii) of the theorem.

To prove (iv), we can take, for some fixed i,

$$f(\theta, z) = f(N, z) = \frac{1}{1 - zr_{i,N}}, \qquad f_u = \frac{1}{1 - z_0 r_i}.$$

Then, for any ε sufficiently small, $\varepsilon > 0$, we can obviously find N_ε^0 and σ_u, $\sigma_u > 0$, such that, for $\Theta = \{N_\varepsilon^0, N_\varepsilon^0 + 1, \dots\}$, condition (II) of Theorem 4.4 is satisfied. Then we obtain

$$\frac{\partial Z_N}{\partial r_{i,N}} = \frac{1}{1 - z_0 r_i} \frac{1}{\sqrt{2\pi N S''(z_0)}} \exp(NS_N(z_{0,N}))(1 + \zeta_N),$$
$$|\zeta_N| < 25\varepsilon,\ N > \max(N_\varepsilon^0, N_\varepsilon).$$

Applying (i), we can see that

$$\frac{1}{Z_N}\frac{\partial Z_N}{\partial r_{i,N}} \to \frac{z_0}{1 - z_0 r_i}, \qquad N \to \infty.$$

And, as $r_{i,N} \to r_i$, $N \to \infty$, we obtain in a straightforward manner the desired result.

Statement (iii) can obviously be proved in exactly the same manner. □

5. Proof of the main results in the condensation case.

PROOF OF THEOREM 2.2. Note that by Lemma 3.3 the value $z_0(\lambda)/(1 - z_0(\lambda))$ is strictly and infinitely increasing when $\lambda \nearrow \lambda_{cr}$. Hence, for any m, $m > 0$, we can choose $\lambda' = \lambda'(m) < \lambda_{cr}$ such that $z_0(\lambda')(1 - z_0(\lambda')) = m + 1$.

Note also that, instead of tracking the varying index $i(N)$ of the node, where $r_{i(N), N} = 1$, we can assume, without loss of generality, that $i(N) = 1$, that is, $r_{1, N} = 1$. Indeed, the value of the expression for $m_{i, N}$ does not depend on i, but only on the value of $r_{i, N}$.

With $M'(N) = [\lambda' N]$ ($[\cdots]$ denotes the integer part of $(\cdots)$), the assumptions of Theorem 2.1 hold true, and therefore there exists N': $\forall N > N'$, $m_{1, M'(N), N} > z_0(\lambda')/(1 - z_0(\lambda')) - 1 = m$. But $M/N \to \lambda \geq \lambda_{cr} > \lambda'$. Hence, $\exists N''$: $\forall N$, $N > N''$, $M(N) = N\lambda(1 + \varepsilon_N) \geq M'(N)$. Therefore, by Lemma 3.1, $\forall N$, $N > \max(N', N'')$, $m_{1, N} = m_{1, M(N), N} \geq m_{1, M', N} > m$. As m is arbitrary, Theorem 2.2 is proved. □

PROOF OF THEOREM 2.3. The case of $\lambda < \lambda_{cr}$ was treated in Theorem 2.1.

Proceeding to the case of $\lambda > \lambda_{cr}$, we note that under the assumptions of the theorem it represents just one of the possibilities for the general case of $z_0 = 1$, which appears to be more difficult to study than the regular one, due to the unknown rate of convergence of the network parameters. But in this particular situation we can evaluate the integral for Z_N by taking into account the residue at $z = 1$, the new integration curve crossing the real axis to the right of that point. Indeed, (3.7) is not satisfied with $z \leq 1$ if $\lambda > \lambda_{cr}$, and 1 is separated from other singularity points.

Let us rewrite

$$Z_N = Z_N(r_1, \dots, r_K) = \frac{1}{2\pi i} \int_\gamma T_N^K(z)\, dz,$$

$$T_N^K(z) := \frac{1}{z} \prod_{j=1}^{K} \frac{1}{1 - zr_j} \exp\bigl(N S_N^K(z)\bigr),$$

where

$$S_N^K(z) = -\lambda(1 + \varepsilon_N)\ln z - \frac{1}{N} \sum_{j=K+1}^{N} \ln\bigl(1 - zr_{j,N}\bigr). \tag{5.1}$$

Let $z_{1, N}$ be the leftmost point, where $\partial S_N^K(z)/\partial z = 0$. Note that, for any n, $n \geq 0$,

$$\frac{\partial^n}{\partial z^n} S_N^K(z) \to \frac{\partial^n}{\partial z^n} S(z), \qquad N \to \infty,$$

uniformly on every compact set outside $[1/B, +\infty)$. It is clear that there exists $\lim_{N\to\infty} z_{1, N} =: z_1 \geq 1$. But z_1 is a root of the equation

$$-\frac{\lambda}{z} + \int_0^B \frac{r}{1 - zr}\, d\mathbf{I}(r) = 0 \quad \Leftrightarrow \quad \frac{\partial S(z)}{\partial z} = 0.$$

Since $S'(z) = -\lambda/z + h(z)$ and $\lambda > \lambda_{cr}$,

$$S'(1) < 0 \quad \text{and} \quad z_1 > 1. \tag{5.2}$$

Note also that $z_1 < 1/B$ because of (2.9).

Now we split the integral for $Z_N(r_1, \dots, r_K)$ into two:

$$\begin{aligned} Z_N(r_1,\dots,r_K) &= \Im_{1,N}(r_1,\dots,r_K) + \Im_{2,N}(r_1,\dots,r_K),\\ \Im_{1,N}(r_1,\dots,r_K) &= -\frac{1}{2\pi i}\int_{\gamma'} T_N^K(z)\,dz, \qquad \Im_{2,N}(r_1,\dots,r_K)\\ &= \frac{1}{2\pi i}\int_{\gamma_N''} T_N^K(z)\,dz, \end{aligned} \tag{5.3}$$

where γ' is a small circle around $z_0 = 1$, and γ_N'' embraces 0 and z_0 and goes through $z_{1,N}$. The quantity $\Im_{2,N}$ can be evaluated in the same way as in Theorem 2.1, and we have

$$\begin{aligned} \Im_{2,N}(1,\dots,1) &= \frac{1}{2\pi i}\int_{\gamma_N''} \frac{1}{z}\frac{1}{(1-z)^K}\exp\bigl(NS_N^K(z)\bigr)\,dz\\ &\sim \frac{\text{const}}{\sqrt{N}}\exp\bigl(NS_N^K(z_{1,N})\bigr), \qquad N \to \infty. \end{aligned} \tag{5.4}$$

Observe that the following equality holds due to the residue formula for r inside γ':

$$\begin{aligned} &-\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{z^K}\frac{1}{1-zr}\exp\bigl(NS_N^K(z)\bigr)\,dz\\ &\quad = r^{K-1}\exp\bigl(NS_N^K(r^{-1})\bigr), \end{aligned} \tag{5.5}$$

which yields that

$$\begin{aligned} &\Im_{1,N}(r,\dots,r)|_{r=1}\\ &\quad = \frac{\partial^{K-1}}{\partial r^{K-1}}\bigl(r^{K-1}\exp\bigl(NS_N^K(r^{-1})\bigr)\bigr)\Big|_{r=1}\\ &\quad \sim (-NS'(1))^{K-1}\exp\bigl(NS_N^K(1)\bigr), \qquad N \to \infty. \end{aligned} \tag{5.6}$$

The asymptotics in (5.6) are easy to establish, since in the expansion of the derivative in (5.6) only one term has the maximal order of growth.

From (5.2) it follows that (5.4) is exponentially small with respect to (5.6), and the asymptotics for Z_N are therefore proved.

Let $i = 1$ (the case of $1 < i \le K$ can be treated in the same way). We will differentiate the representation (5.3), taking into account (5.5),

$$\begin{aligned}\left.\frac{\partial Z_N(r_1,\dots,r_K)}{\partial r_1}\right|_{r_1=\cdots=r_K=1} &= \frac{1}{K}\left.\frac{\partial Z_N(r,\dots,r)}{\partial r}\right|_{r=1}\\ &= \frac{1}{K}\frac{\partial^K}{\partial r^K}\left(r^{K-1}\exp\left(NS_N^K(r^{-1})\right)\right)\Big|_{r=1}\\ &\quad+\frac{1}{K}\frac{\partial}{\partial r}\Im_{2,N}(r,\dots,r)\Big|_{r=1}.\end{aligned} \tag{5.7}$$

The second term of the sum in (5.7) can again be evaluated as in Theorem 2.1, and it is equivalent to const $\times \exp(NS_N^K(z_{1,N}))/\sqrt{N}$. For the first term of (5.7) we have

$$\begin{aligned}&\frac{\partial^K}{\partial r^K}\left(r^{K-1}\exp\left(NS_N^K(r^{-1})\right)\right)\Big|_{r=1}\\ &\qquad\sim (-NS'(1))^K\exp\left(NS_N^K(1)\right),\qquad N\to\infty.\end{aligned} \tag{5.8}$$

The second term is again exponentially small compared to the first one, and therefore

$$m_{1,N} = \left.\frac{r_1(\partial Z_N/\partial r_1)}{Z_N}\right|_{r_1=1} \sim -N\frac{S'(1)}{K},\qquad N\to\infty,$$

which proves statement (i) of the theorem.

For $i > K$,

$$\begin{aligned}&\left.\frac{\partial Z_N}{\partial r_{i,N}}\right|_{r_1=\cdots=r_K=1}\\ &\quad=\left(\frac{\partial \Im_{1,N}(r_1,\dots,r_K)}{\partial r_{i,N}}+\frac{\partial \Im_{2,N}(r_1,\dots,r_K)}{\partial r_{i,N}}\right)\Big|_{r_1=\cdots=r_K=1}.\end{aligned} \tag{5.9}$$

We have

$$\begin{aligned}&\left.\frac{\partial \Im_{1,N}(r_1,\dots,r_K)}{\partial r_{i,N}}\right|_{r_1=\cdots=r_K=1}\\ &\quad=-\frac{1}{2\pi i}\int_{\gamma'}\frac{1}{z}\frac{1}{(1-zr)^K}\frac{z}{1-zr_{i,N}}\exp\left(NS_N^K(z)\right)dz\Big|_{r=1}\\ &\quad=\frac{\partial^{K-1}}{\partial r^{K-1}}\left(\frac{1}{r-r_{i,N}}r^{K-1}\exp\left(NS_N^K(r^{-1})\right)\right)\Big|_{r=1}\\ &\quad\sim\frac{1}{1-r_{i,N}}(-NS'(1))^{K-1}\exp\left(NS_N^K(1)\right),\qquad N\to\infty.\end{aligned} \tag{5.10}$$

By reasoning analogous to that of the proof of Theorem 2.1, it can be shown that the ratio of $\partial \Im_{2,N}/\partial r_{i,N}$ and $\Im_{2,N}$ is uniformly bounded when $N \to \infty$, $0 \le r_{i,N} \le B$. Since $\Im_{2,N}$ is exponentially small w.r.t. $\Im_{1,N}$, we can see, first, that

$$m_{i,N} = \frac{1}{Z_N} r_{i,N} \frac{\partial Z_N}{\partial r_{i,N}}$$

are uniformly bounded in i, $i > K$, and N, and, second, that $m_{i,N} \to r_i/(1 - r_i)$, $N \to \infty$, if $r_{i,N} \to r_i$, $N \to \infty$, $i > K$. Therefore, Theorem 2.3 is completely proved. □

6. Remarks, examples and generalizations. We remark that our approach can also be taken w.r.t. networks including infinite servers. For example, let the system contain one IS node ($i = 0$) with mean service time equal to ν_N^{-1}. Then the relative utilizations of the queueing nodes are given by

$$r_{i,N} = \frac{\nu_N \rho_{i,N}}{\rho_{0,N}\mu_{i,N}}, \qquad i = 1, \dots, N,$$

where $\rho_N = (\rho_{0,N}, \dots, \rho_{N,N})$ are solutions of the routing equation with the extended routing matrix $P_N = \{p_{ij,N}\}_{i,j=0}^N$ (we suppose that $\rho_{0,N} \neq 0$). Then for the stationary distribution on the network we have the product form:

$$\begin{aligned} &\mathbf{P}_{M,N}\{\xi_{1,MN} = n_1, \xi_{2,MN} = n_2, \dots, \xi_{N,MN} = n_N\} \\ &\quad = \frac{1}{Z_{M,N}} \frac{M!}{(M - n_1 - n_2 - \cdots - n_N)!} \prod_{i=1}^N r_{i,N}^{n_i}, \qquad 0 \le \sum_{i=1}^N n_i \le M, \end{aligned} \tag{6.1}$$

and the generating function of the network is

$$\Xi_N(u) = e^{u/r_{\max,N}} \prod_{i=1}^N \frac{1}{1 - us_{i,N}},$$

where $s_{i,N} = r_{i,N}/r_{\max,N}$, $r_{\max,N} = \max_{1 \le i \le N} r_{i,N}$, $u = zr_{\max,N}$.

After making the same assumption as in [14] and [17]:

$$Nr_{\max,N} \to C_0 > 0, \qquad N \to \infty,$$

we can find the critical value of the density $\lambda = \lim_{N \to \infty} M/N$ in terms of the limit distribution $\mathbf{I}(\cdot)$ of $\{s_{i,N}\}$:

$$\lambda_{\mathrm{cr}} = \frac{1}{C_0} + \lim_{u \to 1-} \int_0^1 \frac{s}{1 - us}\, d\mathbf{I}(s),$$

from which we can obtain the critical value for the network without the IS node as the limit case when $\nu_N \to \infty$, since then $1/C_0 \to 0$.

This critical value of λ was implicitly found by Kogan [17], Proposition 2, under assumption (2.6) and when there is one distinguished node with $r_{N,N} = \text{const} > \max_{u \in [0, \lambda^{-1}]} g(u)$. In fact, the density λ is not a parameter in (2.6) it participates in the definition of $g(\cdot)$ and the results of [17] are derived under the assumption of λ and $g(\cdot)$ fixed. One can, however, make λ variable by considering $g(u) := f(\lambda u)$ with f finite-piecewise continuously differentiable on $[0,1]$. Through simple algebra one can see then the equivalence of the results.

We believe that other generalizations should be possible, too.

As far as the limit measures are concerned, the natural question arises: which measures are possible? Let us give some remarks.

REMARK 1. Trivially, for any J_N, P_N we can choose $\mu_{i,N}$ so that multiplication on $\mu_{i,N}$ gives any $\mathbf{I}_N$ we like; so any limiting measure can appear on $[0,1]$.

REMARK 2. Even if $\mu_{i,N}$ are fixed, for example, $\mu_{i,N} \equiv 1$, then for any positive vector ρ_N there exists a stochastic routing matrix P_N such that $\rho_N P_N = \rho_N$ holds. It is sufficient to define

$$p_{ij,N} = \rho_{j,N}\left(\sum_{k=1}^{N} \rho_{k,N}\right)^{-1}, \qquad i,j = 1,\dots,N.$$

REMARK 3. We can get the same results as in Remark 2 even for much more restricted classes of topologies and interactions. An example for the circle topology is as follows.

LEMMA 6.6. *Let N be odd and the vector ρ_N be given such that*

$$0 < \rho_{3,N} \le \rho_{4,N} \le \cdots \le \rho_{2,N} \le 2\rho_{3,N}.$$

Then there exists a stochastic routing matrix P_N satisfying $\rho_N P_N = \rho_N$ and also

$$p_{ij,N} \ne 0 \quad \Leftrightarrow \quad |i-j| \in \{1, N-1\}.$$

In other words, in the circle topology, the interaction radius in this network is equal to 1.

We skip the proof.

REMARK 4. If there exists a symmetry group G_N acting transitively and one-to-one on J_N so that

$$p_{ij} = p_{g(i)g(j)}$$

for all $1 \le i, j \le N$ and $g \in G_N$, then all $\rho_{i,N}$ are equal. If all $\mu_{i,N} \equiv 1$, then I_N is the point measure at 1. Examples include the lattice on the torus with translation invariant probabilities and completely symmetric networks.

REMARK 5. Consider more closely the example of the network on the torus, with the probabilities of jumps invariant w.r.t. the shifts in the state space. If all the service rates are equal to 1, then $\mathbf{I}_N$ is the point measure at 1, and $\lambda_{\mathrm{cr}} = +\infty$; that is, there is no phase transition. If we decrease the service rate only at the origin, $\mu'_{0,N} < 1$, then $\mathbf{I}_N$ will become the point measure at $B = \mu'_{0,N} < 1$. Hence, $h(z) = B/(1 - zB)$ and $\lambda_{\mathrm{cr}} = B/(1-B) < +\infty$—there appears a phase transition.

REMARK 6. Let us look at the maximal mean queue length in a system with given distribution of r.u.'s and with varying density λ. Let $r_{N,N} = 1$. According to our results, the relation $m_{N,N}/N \to 0$ as $N \to \infty$ and $\lambda < \lambda_{\text{cr}}$, and it can tend to some positive increasing function $f(\lambda)$ for $\lambda > \lambda_{\text{cr}}$. We do not have an analytic expression for $f(\lambda)$. Figure 1 depicts a set of graphs of $m_{N,N}/N$ corresponding to the simplest condensation case: a network with one distinguished node (see, e.g., the previous example). The graphs were constructed using the explicit expression for the p.f. A solid boldface line denotes the predicted $f(\lambda)$.

We conjecture that $f(\lambda)$ will have similar properties in the general condensation case. Figure 2 depicts a set of graphs of $m_{N,N}/N$ as per our asymptotic formulas and in the way described in the remark after Theorem 2.2, for a "general" network, with r.u. given by (2.11) with $g(\cdot)$ not satisfying the smoothness conditions of [14] and [17]. One can see that there exists an asymptotic curve (the solid boldface line) resembling that given in Figure 1.

These examples illustrate the fact that knowledge of the critical point λ_{cr} is important in the evaluation of the network's sensitivity w.r.t. traffic and/or service condition fluctuations.

At the end we indicate another problem. It could be natural to define an "infinite" closed Jackson network in the following way. Let us take a countable Markov chain L with state space S and transition probabilites p_{ij}. For any finite $\Lambda \subset S$, define the finite Markov chain S_Λ by truncating the stochastic matrix in some way. For example, put

$$p_{ij}^{\Lambda} = \begin{cases} p_{ij}, & i, j \in \Lambda,\ i \neq j, \\ p_{ii} + \sum_{j \in S \setminus \Lambda} p_{ij}, & i = j. \end{cases}$$

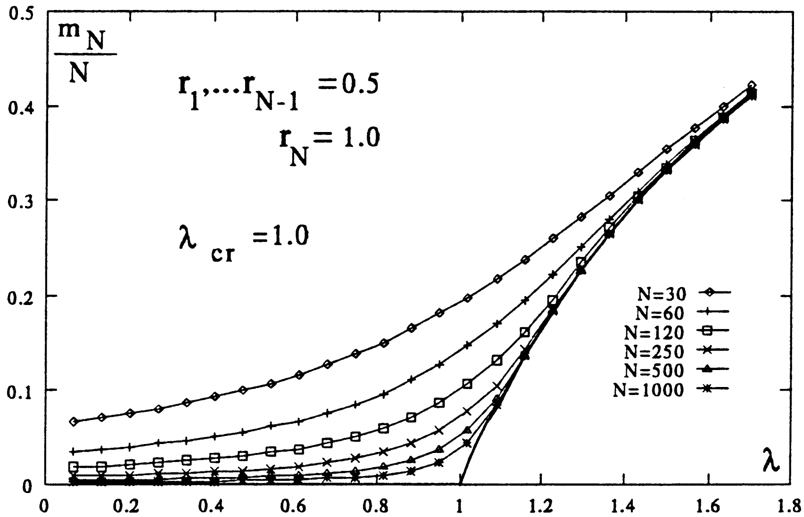

FIG. 1.

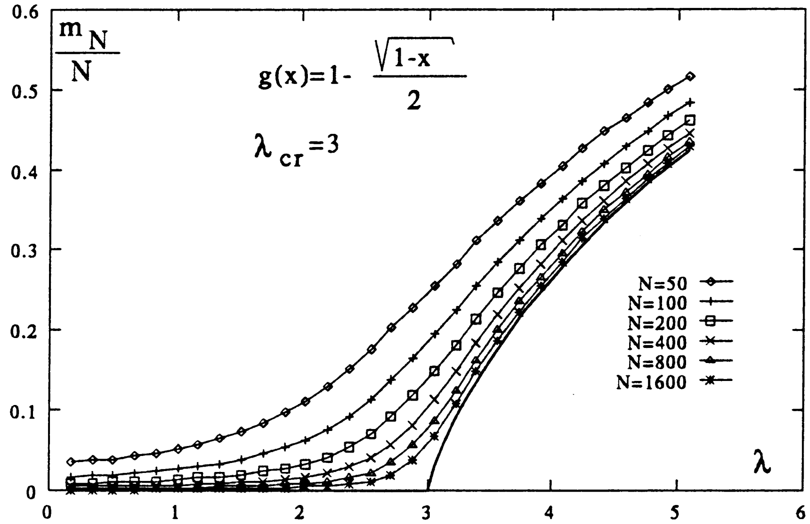

FIG. 2.

Suppose also that for any $i \in S$ the service rate μ_i is given. Then we define a closed Jackson network J_Λ with routing probabilities p_{ij}^{Λ}, rates μ_i and $M_\Lambda = [\lambda|\Lambda|]$.

One of the questions is what limiting measures **I** can appear for various sequences

$$\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_N \subset \cdots$$

such that $|\Lambda_N| = N$, $\bigcup \Lambda_N = S$.

REFERENCES

[1] BASKETT, F., CHANDY, K. M., MUNTZ, R. R. and PALACIOS, F. G. (1975). Open, closed, and mixed networks of queues with different classes of customers. *J. Assoc. Comput. Mach.* **22** 248–260.

[2] BENDER, C. M. and ORSZAG, S. A. (1978). *Advanced Mathematical Methods for Scientists and Engineers*. McGraw-Hill, New York.

[3] BIRMAN, A. and KOGAN, Y. (1991). Asymptotic analysis of closed queueing networks with bottlenecks. In *Proceedings of the International Conference on the Performance of Distributed Systems and Integrated Communication Networks* (T. Hasegawa, H. Tagaki and Y. Takahashi, eds.) 237–252. North-Holland, Amsterdam.

[4] BIRMAN, A. and KOGAN, Y. (1992). Asymptotic evaluation of closed queueing networks with many stations. *Comm. Statist. Stochastic Models* **8** 543–563.

[5] BUZEN, J. P. (1973). Computational algorithms for closed queueing networks with exponential servers. *Comm. ACM* **16** 527–531.

[6] BUZEN, J. P. and DENNING, P. J. (1978). The operational analysis of queueing network models. *Comput. Surveys* **10** 225–261.

[7] COPSON, E. T. (1965). *Asymptotic Expansions*. Cambridge Univ. Press.

[8] GORDON, J. J. (1990). The evaluation of normalizing constants in closed queueing networks. *Oper. Res.* **38** 863–869.

[9] GORDON, W. J. and NEWELL, G. F. (1967). Closed queueing systems with exponential servers. *Oper*. *Res*. **15** 254–265.
[10] HARRISON, P. G. (1985). On normalizing constants in queueing networks. *Oper*. *Res*. **33** 464–468.
[11] JACKSON, J. R. (1963). Jobshop-like queueing systems. *Management Sci*. **10** 131–142.
[12] KELLER, J. B. (1978). Rays, waves, and asymptotics. *Bull*. *Amer*. *Math*. *Soc*. **84** 727–750.
[13] KELLY, F. P. (1980). *Reversibility and Stochastic Networks*. Wiley, New York.
[14] KNESSL, C. and TIER, C. (1990). Asymptotic expansions for large closed queueing networks. *J*. *Assoc*. *Comput*. *Mach*. **37** 144–174.
[15] KNESSL, C. and TIER, C. (1992). Asymptotic expansions for large closed queueing networks with multiple job classes. *IEEE Trans*. *Comput*. **41** 480–488.
[16] KOBAYASHI, H. and REISER, M. (1975). Queueing networks with multiple closed chains: theory and computational algorithms. *IBM J*. *Res*. *Develop*. **19** 283–294.
[17] KOGAN, Y. (1992). Another approach to asymptotic expansions for large closed queueing networks. *Oper*. *Res*. *Lett*. **11** 317–321.
[18] LAVENBERG, S. S., ed. (1983). *Computer Performance Modeling Handbook*. Academic Press, New York.
[19] LAVENBERG, S. S. and REISER, R. (1980). Stationary state probabilities at arrival instants for closed queueing networks with multiple types of customers. *J*. *Appl*. *Probab*. **17** 1048–1061.
[20] MARKUSHEVICH, A. I. (1965). *Theory of Functions of a Complex Variable* **1**, **2**. Prentice-Hall, Englewood Cliffs, NJ.
[21] MASSEY, W. A. (1987). Stochastic orderings for Markov processes on partially ordered spaces. *Math*. *Oper*. *Res*. **12** 350–367.
[22] MCKENNA, J. (1987). Asymptotic expansions of the sojourn time distribution functions of jobs in closed, product-form queueing networks. *J*. *Assoc*. *Comput*. *Mach*. **34** 985–1003.
[23] MCKENNA, J., MITRA, D. and RAMAKRISHAN, K. G. (1981). A class of closed Markovian queueing networks: integral representations, asymptotic expansions, and generalizations. *Bell System Tech*. *J*. **60** 599–641.
[24] MCKENNA, J. and MITRA, D. (1982). Integral representations and asymptotic expansions for closed Markovian queuing networks: normal usage. *Bell System Tech*. *J*. **61** 661–683.
[25] MCKENNA, J. and MITRA, D. (1984). Asymptotic expansions and integral representations of moments of queue lengths in closed Markovian networks. *J*. *Assoc*. *Comput*. *Mach*. **31** 346–360.
[26] MEI, J.-D. and TIER C. (1994). Asymptotic approximations for a queueing network with multiple classes. *SIAM J*. *Appl*. *Math*. **54** 1147–1180.

INRIA
DOMAINE DE VOLUCEAU
ROCQUENCOURT, B.P.105
78153 LE CHESNAY CEDEX
FRANCE
E-mail: Vadim.Malyshev@inria.fr

UNIVERSITÉ D'ORLÉANS
DÉPARTEMENT DE MATHÉMATIQUES
U.F.R. FACULTÉ DES SCIENCES, B.P. 6759
45067 ORLEANS CEDEX 2
FRANCE