Информационные процессы, Том 4, № 3, 2004, стр. 256–260.

A Note on Asymptotics Associated with Limit Theorem for Terminating Renewal Process in Processor Sharing Queue[1]

S.F.Yashkov*, **A.S.Yashkova****

**Institute for Information Transmission Problems, 19, Bolshoi Karetny lane, 101447 Moscow GSP–4, Russia. E-mail: yashkov@iitp.ru*
***Municipal Institute of Zhukovsky, Chair Appl. Informatics in Economics, 15, Mayakovsky Street, 140180 Zhukovsky, Moscow region, Russia.*
Received September 17, 2004

Abstract—We give a simple proof of an asymptotic result associated with terminating renewal process in the M/G/1 queue under egalitarian processor sharing. The proof gives insight into the principles according to which the time–dependent solution is obtained for the queue length at time t. This result was announced without the proof in [6, Theorem 2.17].

1. INTRODUCTION

In this note we explore a useful integral representation for the Laplace transform of the generating function for the number of jobs at time t on the first busy period in the M/G/1 queue with egalitarian processor sharing (EPS), which is one of most celebrated models in queueing theory. The EPS model was introduced by Kleinrock in [1] (see also [2]). Mentioned non–trivial integral representation has been derived in [3], [4], [5], and it is one of the key ingredients of transient analysis of the M/G/1—EPS queue. An additional (but abridged) treatment of this result is given also in memoir [6, Theorem 2.16]. However, this solution is obtained (in terms some function ψ, described in Section 2) in little tractable form, which is very difficult to use in applications.

But it may be carried one important step further to find an explicit solution for the Laplace transform of the corresponding generating function for the queue length at time t under simpler condition that the first busy period is initiated by only one job. The proof of this (asymptotic in some sense) solution (an another key ingredient of transient analysis) is given only in [5], [7] (in Russian) and in [8]. The corresponding theorem is also announced in [6] as Theorem 2.17 without the proof. We give here a relative short and somewhat modified proof which is based on obtaining asymptotics for the components of the decomposition of function ψ by using a limit theorem for terminating renewal processes studied in Feller [9, §11.7]. Although the argument leading to the explicit solution coincides partially with one from [5], [8], we give also new argument associated with using the Esscher transformation of probability measure, which is known from insurance risk theory since 1932 [10], [11, Ch. 5, §2d], [12, p. 31, p. 422].

The remainder of this note is organized as follows. The proof of main result (Theorem 2.2) is given in Section 2. Finally, in Section 3 we state our conclusion.

2. THE QUEUE LENGTH DISTRIBUTION IN TERMS OF DOUBLE TRANSFORMS

We recall first some definitions and known results. We consider the egalitarian processor sharing queue M/G/1 with the intensity λ of Poisson input process and service time (or size of job) distribution $B(x)$

[1] The research of S.F.Y. was partially supported by Grant no. Sci.Sch.–934.2003 (Head R.A.Minlos) and Grant to the Program of Fundamental Researches of Russian Academy of Sciences (the Division of Informatics) “New Physical and Structural Solutions in Infotelecommunications” (Head N.A.Kuznetsov).

$(B(0+) = 0,\ B(\infty) = 1)$ with the mean $\beta_1 < \infty$ and the Laplace–Stieltjes transform $\beta(s)$. When there are $n \geq 1$ jobs in the system then each of them receives service at a rate which is $1/n$ times the rate of service that a solitary job in the system would receive. Let $\rho = \lambda\beta_1$. We do not assume that $\rho < 1$.

The number of jobs at time t in the M/G/1–EPS queue will be denoted as $L(t)$. The (non–Markovian) process $\{L(t)\}$ is considered on the first busy period. Let $\zeta = \inf(t > 0 : L(t) = 0)$ and $\pi(s) = \mathsf{E}[\mathrm{e}^{-s\zeta}]$ is the LST of the busy period distribution, i.e., it is the positive root of the functional equation

$$\pi(s) = \beta(s + \lambda - \lambda\pi(s)) \tag{2.1}$$

with the smallest absolute value.

Definition. The function

$$p_n(z,s) = \int_0^\infty \mathrm{e}^{-st}\mathsf{E}[z^{L(t)}\mathbf{1}_{(\zeta>t)}|L(0)=n]\,dt, \quad \mathrm{Re}s > 0,\ |z| \leq 1 \tag{2.2}$$

defines the Laplace transform w.r.t. t of the generating function for the time–dependent queue–length distribution $\mathrm{P}(L(t) = i|L(0) = n)$, $i = 1, 2, \dots,\ n > 0$ on the first busy period that is initiated with n jobs present at time $t = 0$.

Now we summarize those results from [5, Ch. 2] (see also [3], [4], [6]) which will be need for the proof of theorem 2.2. The following theorem [3], [4], [5], [6] is the starting point to derive the time–dependent distribution of the process $\{L(t)\}$ on the first busy period.

Theorem 2.1. *For the M/G/1—EPS queue, the exact solution for the function $p_n(z,s)$ defined by (2.2) has the following integral representation*

$$p_n(z,s) = \frac{z}{\lambda^n}\int_0^\infty \frac{\partial}{\partial z}\left[-\frac{\partial}{\partial t}\ln\psi(z,s,t)\right]^n dt,\ Re\ s > 0,\ |z| \leq 1, \tag{2.3}$$

where function ψ $(\psi(z,s,0) = 1)$ is given by its Laplace transform w.r.t. t

$$\tilde{\psi}(z,s,q) = \int_0^\infty \mathrm{e}^{-qt}\psi(z,s,t)\,dt = \frac{q+s+\lambda-\lambda z+\lambda z\beta(q+s+\lambda)}{(q+s+\lambda)(q+\lambda\beta(q+s+\lambda))}\ . \tag{2.4}$$

Here $\mathrm{Re}\ s > 0,\ q > -\lambda\pi(s),\ |z| \leq 1$*, and $\pi(s)$ is the unique solution of (2.1).*

We will show how Theorem 2.1 can be used to determine the function $p_n(z,s)$ in explicit form for the case $n = 1$.

Theorem 2.2. *For the M/G/1—EPS queue, the exact solution for the function $p_1(z,s)$ defined by (2.2) is given by*

$$p_1(z,s) = \frac{z(1-\pi(s))}{s+\lambda(1-z)(1-\pi(s))}, \quad Re\ s > 0,\ |z| \leq 1, \tag{2.5}$$

where $\pi(s)$ satisfies to the functional equation (2.1).

Proof. The decomposition of the function $\tilde{\psi}(z,s,q)$ follows from (2.4) after some routine algebra

$$\tilde{\psi}(z,s,q) = \tilde{\kappa}_1(s,q) - z\tilde{\kappa}_2(s,q), \tag{2.6}$$

where $\tilde{\kappa}_1(s,q)$ and $\tilde{\kappa}_2(s,q)$ are the Laplace transforms of some functions $\kappa_1(s,t)$ and $\kappa_2(s,t)$, respectively:

$$\tilde{\kappa}_1(s,q) = [q+\lambda\beta(q+s+\lambda)]^{-1},\ \tilde{\kappa}_2(s,q) = (\lambda-\lambda\beta(q+s+\lambda))\tilde{\kappa}_1(s,q)(q+s+\lambda)^{-1}. \tag{2.7}$$

Using (2.6), the expression (2.3) takes the form for $n = 1$:

$$p_1(z,s) = \frac{z}{\lambda}\int_0^\infty \frac{\partial}{\partial z}\left[-\frac{\partial}{\partial t}\ln(\kappa_1(s,t) - z\kappa_2(s,t))\right] dt,\ Re\ s > 0,\ |z| \le 1. \tag{2.8}$$

Now we can interchange the order of the differentiation w.r.t. z and t in the integrand of (2.8). Then (2.8) reduces after the integration to

$$p_1(z,s) = \frac{z}{\lambda}\lim_{t\to\infty}\frac{\kappa_2(s,t)}{\kappa_1(s,t) - z\kappa_2(s,t)}\ , \tag{2.9}$$

where we have used the fact that $\psi(z,s,0) = 1$.

Now it is possible to obtain the asymptotics of $\kappa_1(s,t)$ and $\kappa_2(s,t)$ as $t \to \infty$. At first we consider $\kappa_1(s,t)$. One way of establishing desired result is to rewrite the first equation in (2.7) in the form

$$\tilde\kappa_1(s,q) = \frac{1}{q+\lambda} + \lambda\frac{1-\beta(q+s+\lambda)}{q+\lambda}\tilde\kappa_1(s,q), \tag{2.10}$$

which is equivalent to the renewal equation

$$\kappa_1(s,t) = \mathrm{e}^{-\lambda t} + h_1 * \kappa_1(s,t), \tag{2.11}$$

where $*$ is the Stieltjes convolution sign on the variable t. Here the function $h_1(s,t)$ is defined by its Laplace transform w.r.t. t $\tilde h_1(s,q) = [\lambda(1-\beta(q+s+\lambda))]/(q+\lambda)$, $\tilde h_1(s,0) < 1$, which shows us that $h_1(s,t)$ is the density of defective (improper, that is, with an atom at $+\infty$) distribution of the time intervals between consecutive renewals (interevent intervals). (Such renewal processes are called transient or terminating ones.) Following Feller [9, Ch. 11, §11.7], we can reduce our defective renewal equation to the standard renewal equation since there exists a constant $C_1(s)$ such that

$$C_1(s) = \int_0^\infty t\mathrm{e}^{\lambda\pi(s)t}h_1(s,t)\,dt = -\tilde h_1'(s,-\lambda\pi(s)) < \infty, \quad s > 0$$

and besides

$$\int_0^\infty \mathrm{e}^{\lambda\pi(s)t}h_1(s,t)\,dt = \tilde h_1(s,-\lambda\pi(s)) = 1, \quad s \ge 0.$$

Then introduce $\kappa_1(s,t)\mathrm{e}^{\lambda\pi(s)t}$ into (2.11) to obtain

$$\kappa_1(s,t)\mathrm{e}^{\lambda\pi(s)t} = \mathrm{e}^{-t(\lambda-\lambda\pi(s))} + \int_0^\infty \kappa_1(s,t-y)\mathrm{e}^{\lambda\pi(s)(t-y)}h_1(s,y)\mathrm{e}^{\lambda\pi(s)y}\,dy,$$

where the function $\mathrm{e}^{-t(\lambda-\lambda\pi(s))}$ is directly Riemann integrable. It allows us to cast (2.11) into the standard renewal set–up. Thus we introduced a new (exponentially tilted) probability measure for the interevent intervals. (In other words, this is so–called Esscher transformation of the probability measure, which is in use in the insurance risk models [10], [11, Ch. 5, §2d].) Now it follows from the Smith's key renewal theorem that

$$\lim_{t\to\infty}\kappa_1(s,t)\mathrm{e}^{\lambda\pi(s)t} = \frac{1}{C_1(s)}\int_0^\infty \mathrm{e}^{-t(\lambda-\lambda\pi(s))}\,dt = \frac{1}{C_1(s)(\lambda-\lambda\pi(s))}\ . \tag{2.12}$$

Now we apply similar arguments to the function $\kappa_2(s,t)$ which is given by its Laplace transform in the second equation in (2.7). It yields the following asymptotics for $\kappa_2(s,t)$ as $t \to \infty$ (the details can be found in [5, §2.8] or in [8])

$$\lim_{t\to\infty}\kappa_2(s,t)\mathrm{e}^{\lambda\pi(s)t} = \frac{1}{C_2(s)(s+\lambda-\lambda\pi(s))}\ , \tag{2.13}$$

where $C_2(s) = C_1(s)$ for $s > 0$.

Using this fact, it follows from (2.12) and (2.13) that

$$\lim_{t\to\infty} \frac{\kappa_2(s,t)}{\kappa_1(s,t)} = \frac{\lambda(1-\pi(s))}{s+\lambda-\pi(s)} \ . \tag{2.14}$$

It allows us to reduce (2.8) to (2.5). □

Remark 21. The Esscher transformation is related to the well–known Girsanov transformation of the probability measure (see, for example, [11, p. 568 in Russian edition]). (One of merits of Girsanov's theorem is that it allows to compute explicitly the law of solutions of stochastic differential equations with constant diffusion coefficients.)

3. CONCLUSION

The proof presented in Section 2 gives an additional insight into the principles according to which the transient solutions have been found for the number of jobs in the M/G/1—EPS queue at time t (on the first busy period) as well as (partially) for the sojourn time of the job with the size u arriving at time t into the EPS queue (see [13]). Of course, these exact solutions are derived in terms of multiple transforms. Although Laplace transforms currently seem to be somewhat in disfavor, but they have proven their worth in very many studies in applied probability. Moreover, these transforms can often be effectively inverted numerically (see, for example, [14]), at least for simpler service disciplines, such as FCFS or LCFS with preemptions.

REFERENCES

1. Kleinrock L. Time-shared systems: a theoretical treatment. *J. Assoc. Comput. Mach.*, 1967, vol. 14, no. 2, pp. 242–251.

2. Kleinrock L. *Queueing Systems*. New-York: Wiley, 1976, vol. 2. Russian edition: Kleinrock L. *Computer Systems with Queues.* Ed. Tsybakov B.S. Moscow: Mir, 1979.

3. Yashkov S.F. The non–stationary distribution of numbers of calls in the M/G/1 processor–sharing queue. In: *Systems Analysis and Simulation 1988. Proc. 3rd Int. Symp.* (Berlin, Sept. 12–16, 1988). Eds. A. Sydow et al. Berlin: Akademie, 1988, vol. 2 (Ser.: Math. Research, Bd. 47), pp. 158–162. Reprinted in: *Advances in Simulation*. Eds. P.A.Lukar and B.Schmidt. Berlin: Springer, 1988, vol. 2, pp. 158–162.

4. Yashkov S.F. On method of analysis of a processor sharing system. In: *Teletraffic Theory in Systems of Informatics.* Moscow: Nauka, 1989, pp. 38–45 (in Russian).

5. Yashkov S.F. *Analysis of Queues in Computers*. Moscow: Radio i Svyaz, 1989 (in Russian).

6. Yashkov S.F. Mathematical problems in the theory of shared-processor systems. In: *Itogi Nauki i Tekhniki. Ser.: Probability Theory*. Moscow: VINITI, 1990, vol. 29, pp. 3–82 (in Russian). Engl. edition: *J. of Soviet Mathematics*, 1992, vol. 58, no. 2 (Jan.), pp. 101–147.

7. Yashkov S.F. On some characteristics of non–classical queueing models with processor sharing. In: *Models and Methods of Information Networks.* Moscow: Nauka, 1990, pp. 14–19 (in Russian).

8. Yashkov S.F. Time–dependent analysis of processsor–sharing queue. In: *Queueing, Performance and Control in ATM. Proc. 13th Int. Teletraffic Congress.* (Copenhagen, June 19–26, 1991). Eds. J.W.Cohen and C.D.Pack. Amsterdam: Elsevier, 1991, pp. 199-204.

9. Feller W. *Introduction to Probability Theory and its Applications.* New York: Wiley, 1966, vol 2. Russian Edition: Feller W. *Introduction to Probability Theory and its Applications.* Ed. Prokhorov Yu.V. Moscow: Mir, 1967, vol. 2.

10. Esscher F. On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift.*, 1932, vol. 15, pp. 175–195.

11. Shiryaev A.N. *Essentials of Stochastic Finance Mathematics.* Moscow: Phasis, 1998, vol. 1: Facts. Models, vol. 2: Theory (in Russian). Engl. edition: Shiryaev A.N. *Essentials of Stochastic Finance: Facts, Models, Theory.* World Scientific, 1999.

12. Embrechts P., Kluppelberg C., Mikosch T. *Modelling Extremal Events for Insurance and Finance.* Heidelberg: Springer, 1997.

13. Yashkova A.S., Yashkov S.F. Distribution of the virtual sojourn time in the M/G/1 processor sharing queue. *Information Processes*, 2003, vol. 3, no. 2, pp. 128–137 (available at http://www.jip.ru/2003/128–137.pdf).

14. Choudhury G.L., Lucantoni D.M., Whitt W. Multidimensional transform inversion with applications to the transient M/G/1 queue. *Annals Appl. Probability*, 1994, vol. 7, pp. 719–740.

This paper was recommended for publication by V.I.Venets, a member of the Editorial Board