The Joint Distribution of the Virtual Waitingtime and the Residual Busy Period for the M|G|1 Queue*

by

Marcel F. Neuts

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 116

August, 1967

*This research was partly supported by the Office of Naval Research Contract NONR 1100(26).

ABSTRACT

In the $M|G|1$ queue, the joint distribution of the virtual waitingtime and the residual duration of the busy period at any time t, is given by a formula analogous to the one for the virtual waitingtime by itself. In the limit, as $t \to \infty$, the limiting joint distribution is given by a formula of the Pollaczek-Khintchine type.

The Joint Distribution of the Virtual Waitingtime and the Residual Busy Period for the M|G|1 Queue*

by

Marcel F. Neuts

Purdue University

Introduction

In this paper, we study the joint probability distribution of the virtual waitingtime $\eta(t)$ and the residual busy period $\theta(t)$ for a queue with Poisson input and general service times.

We may phrase the problem considered as follows. We consider a queue with Poisson input of rate λ and general service time distribution $H(\cdot)$ in which the customers are served first come, first served. If at some time t, the input to the queue is shut off it will take an additional length of time $\eta(t)$ before the server becomes free. This is the <u>virtual waitingtime</u>. If the input is not shut off, we denote the additional length of time until the server becomes free by $\theta(t)$ and call it the <u>residual busy period</u>. We make the convention that $\theta(t) = 0$ if $\eta(t) = 0$.

We define the joint probability distribution of $\eta(t)$ and $\theta(t)$ by

$$(1) \qquad \Phi_t(x,y) = P\{\theta(t) \leq x;\ \eta(t) \leq y \mid \xi_0 = i\}$$

*This research was partly supported by the Office of Naval Research Contract NONR 1100(26). Reproduction in whole or in part is permitted for any purpose of the United States Government.

The initial condition $\xi_0 = i$, which we will postulate throughout our discussion, says that at $t = 0$, there are $i > 0$, customers present and that one of them is beginning service at $t = 0$. It is easy to modify our results for different initial conditions. Before stating the main theorems, we will recall some known results from the theory of $M|G|1$ queues. If we denote the distribution function of the length of the busy period by $G(\cdot)$ and its Laplace-Stieltjes transform by $\gamma(s)$, then for every s with $\text{Re } s > 0$ we have:

$$(2) \qquad \gamma(s) = h[s + \lambda - \lambda\gamma(s)], \qquad h(s) = \int_0^\infty e^{-sx} d\, H(x) ,$$

and $\gamma(s)$ is the unique root in the unit disk of the equation:

$$(3) \qquad z = h[s + \lambda - \lambda z], \qquad |z| \leq 1 .$$

Moreover, a busy period with i customers initially has a length, whose distribution is the i-fold convolution $G^{(i)}(\cdot)$ of $G(\cdot)$.

The beginnings of busy periods form a (general) renewal process, whose renewal function we denote by $M(t)$. Its Laplace-Stieltjes transform $m(s)$ is given by:

$$(4) \qquad m(s) = \frac{\lambda}{\lambda + s} \gamma^i(s) \left[1 - \frac{\lambda}{\lambda + s} \gamma(s)\right], \qquad \text{Re } s > 0 .$$

Furthermore, in one of the proofs below, we will make use of the following quantities. Let $G_{ij}^{(n)}(x)$ be the probability that a busy period with i customers initially $(i \geq 1)$ involves at least n services, that the n-th service is completed before x and that at the end of the n-th service j customers are waiting. We define $G_{ij}^{(o)}(x) = \delta_{ij}\, U(x)$, where $U(\cdot)$ is the distribution degenerate at zero. The quantities $G_{ij}^{(n)}$ were studied earlier by Takács [1]. They satisfy the recurrence:

$$(5) \qquad G_{ij}^{(n+1)}(x) = \sum_{a=1}^{j+1} \int_o^x \int_o^{x-u} e^{-\lambda v} \frac{(\lambda v)^{j-a+1}}{(j-a+1)!}\, d\, H(v)\, d\, G_{i\alpha}^{(n)}(u)$$

for $n \geq 0$.

Upon taking Laplace-Stieltjes transforms of the $G_{ij}^{(n)}(x)$ we readily verify that the generating function of these transforms is given by:

$$(6) \qquad \sum_{n=o}^{\infty} \sum_{j=1}^{\infty} g_{ij}^{(n)}(s)\, Z^j = Z\left[Z^i - \gamma^i(s)\right]\left[Z - h(s+\lambda-\lambda Z)\right]^{-1}$$

The two theorems proved below are:

Theorem 1

For $\mathrm{Re}\,\xi > 0$, $\mathrm{Re}\, s_1 \geq 0$, $\mathrm{Re}\, s_2 \geq 0$, we have that:

$$(8) \qquad \int_o^\infty e^{-\xi t} dt \int_o^\infty \int_o^\infty e^{-s_1 x - s_2 y} d_{x,y}\, P\{\theta(t) < x, \eta(t) \leq y \mid \xi_o = i\}$$

$$= \gamma^i(\xi)\, [\xi + \lambda - \lambda\gamma(\xi)]^{-1}$$

$$+ \left[s_1 + s_2 - \xi + \lambda\, h[s_1+s_2+\lambda-\lambda\gamma(s_1) - \lambda\gamma(s_1)\right]^{-1}$$

$$\left\{\gamma^i(\xi)\, \frac{\xi+\lambda-\lambda h[s_1+s_2+\lambda-\lambda\gamma(s_1)]}{\xi + \lambda - \lambda\gamma(\xi)} - h^i[s_1 + s_2 + \lambda - \lambda\gamma(s_1)]\right\}$$

$$= \int_o^\infty e^{-\xi t} dt\, E\left\{\exp\{-\eta(t)\, [s_1 + s_2 + \lambda - \lambda\gamma(s_1)]\}\right\} .$$

Theorem 2:

If and only if $1 - \lambda\alpha > 0$, does the limiting distribution $\Phi_\infty(x,y)$ exist. (α is the mean service time)

Its Laplace-Stieltjes transform is then given by:

$$(9) \qquad \int_0^\infty \int_0^\infty e^{-s_1 x - s_2 y} \, d\, \Phi_\infty(x,y) =$$

$$\frac{\left[s_1 + s_2 + \lambda - \lambda\gamma(s_1)\right](1 - \lambda\alpha)}{s_1 + s_2 + \lambda\, h\left[s_1+s_2+\lambda-\lambda\gamma(s_1)\right] - \lambda\gamma(s_1)} ,$$

The proof of Thm. 2 is immediate from formula (8), and the proof of the existence of the limiting distribution for the virtual waiting-time. We notice that (9) is analogous to the famous Pollaczek-Khinchine formula to which it reduces for $s_1 = 0$.

Proof of Theorem 1

We will present two proofs of theorem 1. The first proof assumes the distribution of $\eta(t)$ known and proceeds via a conditional expectation argument. The second method is from first principles and though much longer, gives a novel proof of the formula for the distribution of the virtual waitingtime. It also yields a number of expressions of independent interest, which we will use in a forthcoming paper on queues with alternating priorities.

The proof by a branching process argument.

Let $\eta(t) = y \geq 0$, then the probability that ν customers arrive during the time interval $(t, t+y]$ is given by $c^{-\lambda y}(\lambda y)^\nu/\nu!$, $\nu \geq 0$.

The instant $t+y$ is clearly the time of a service completion in the queue. It may also be considered as the beginning of a long busy period with ν customers initially. Therefore:

$$(10) \qquad P\{\theta(t) \leq x \mid \eta(t) = y, \quad \nu \text{ new arrivals}\}$$

$$= G^{(\nu)}(x-y), \qquad \nu \geq 0 ,$$

and:

$$(11) \qquad P\{\theta(t) \leq x \mid \eta(t) = y\} = \sum_{\nu=0}^{\infty} G^{(\nu)}(x-y)\, e^{-\lambda y} \frac{(\lambda y)^{\nu}}{\nu!} \quad .$$

Integrating, we obtain:

$$(12) \qquad E\left\{e^{-s_2\theta(t)} \mid \eta(t) = y\right\} =$$

$$\sum_{\nu=0}^{\infty} e^{-\lambda y} \frac{(\lambda y)^{\nu}}{\nu!} \int_y^{\infty} e^{-s_2 x}\, d\, G^{(\nu)}(x-y) = e^{-y[s_2+\lambda-\lambda\gamma(s_2)]}$$

The last part of formula (8) now follows by integration over y and t .

If the queuelength at $t = 0$ is $i > 0$, then we have

$$(13) \qquad P\{\eta(0) \leq y\} = H^{(i)}(y) \quad .$$

The result in the first part of (8) is then obtained by substitution of (13) and the second part of (8) is the known formula for the Laplace-Stieltjes transform of the distribution of $\eta(t)$.

The proof from first principles.

The crucial part in this proof is the evaluation of the following probability, which is of independent interest.

Suppose we have i customers at $t = 0$, one of who just enters service. We define $\Psi_t(x,y)$ as the probability of the event that at time t the queue has not yet emptied out, but that it does so before time $t + x$ and that the virtual waitingtime at t does not exceed y. We note that the definition of $\Psi_t(x,y)$ requires in addition to that of $\Phi_t(x,y)$, that the server has not yet become idle in $[o,t]$.

In order to realize the event in the definition of $\Psi_t(x,y)$ none of the service completions before t can result in an empty queue. Let the last service completion before t occur at time u $(u \leq t)$ and be the n-th $(n \geq 0)$ and let there be j customers at time u. Let there be ν $(\nu \geq 0)$ arrivals of customers in $(u,t]$ and let the first service completion after t occur at time τ $(\tau \geq 0)$. Let there be r_1 customers in the queue at time τ, so that $r_1 - j - \nu + 1 \geq 0$ arrivals have occurred in $(t,\tau]$.

Furthermore we denote by τ_1 the time at which the $j + \nu - 1$ customers who were in the queue before t complete service. The total service time of these $j + \nu - 1$ customers is then $\tau_1 - \tau \geq 0$.

Finally let r be the number of customers in the queue at time τ_1, then $r - r_1 + j + \nu - 1 \geq 0$ new customers have arrived in the interval $(t,\tau_1]$. The busy period with r customers initially at τ_1 comes to an end at time τ_2. $(\tau_2 \geq \tau_1)$

The event in the definition of $\Psi_t(x,y)$ occurs if and only if in addition:

(14) $$\tau_1 \le t + y \ , \qquad \tau_2 \le t + x$$

If we express the complex event described above in terms of known probabilities and sum over all allowable choices of

$n, u, j, \tau, r_1, \tau_1, t, \tau_2$ we obtain:

(15) $$\Psi_t(x,y) = \sum_{n=o}^{\infty} \sum_{j=1}^{\infty} \sum_{\nu=o}^{\infty} \sum_{r_1=j+\nu-1}^{\infty} \sum_{r=r_1-j-\nu+1}^{\infty}$$

$$\underset{(u)}{\int_o^t} \underset{(\tau)}{\int_t^{t+y}} \underset{(\tau_1)}{\int_\tau^{t+y}} \underset{(\tau_2)}{\int_{\tau_1}^{t+x}} d\ G_{ij}^{(n)}(u) e^{-\lambda(t-u)} \frac{\lambda^\nu (t-u)^\nu}{\nu!} \ e^{-\lambda(\tau-t)}$$

$$\frac{\lambda^{r_1-j-\nu+1} (\tau-t)^{r_1-j-\nu+1}}{(r_1-j-\nu+1)!} \ d\ H(\tau-u)\ e^{-\lambda(\tau_1-\tau)} \ \frac{\left[\lambda(\tau_1-\tau)\right]^{r-r_1+j+\nu-1}}{(r-r_1+j+\nu-1)!}$$

$$d\ H^{(j+\nu-1)}(\tau_1-\tau) d\ G^{(r)}(\tau_2-\tau_1) \quad .$$

Next, we evaluate the triple Laplace-(Stieltjes) transform

(16) $$\int_o^\infty e^{-\xi t} dt \int_o^\infty e^{-s_2 y} \int_o^\infty e^{-s_1 y} d_{x,y}\ \Psi_t(x,y) =$$

$$\sum_{n=o}^{\infty} \sum_{j=1}^{\infty} \sum_{\nu=o}^{\infty} \sum_{r_1=j+\nu-1}^{\infty} \sum_{r=r_1-j-\nu+1}^{\infty}$$

$$\underset{(t)}{\int_o^\infty} e^{-\xi t} \underset{(y)}{\int_o^\infty} e^{-s_2 y} \underset{(x)}{\int_y^\infty} s^{-s_1 x} \underset{(u)}{\int_o^t} \underset{(\tau)}{\int_t^{t+y}} d\ G_{ij}^{(n)}(u) d\ H(\tau-u)$$

$$d\ H^{(j+\nu-1)}(t+y-\tau)\ d\ G^{(r)}\ e^{-\lambda(y-u)} \ \frac{[\lambda(t-u)]^\nu}{\nu!} \ e^{-\lambda(\tau-t)}$$

$$\frac{[\lambda(\tau-t)]^{r_1-j-\nu+1}}{(r_1-j-\nu+1)!} \quad e^{-\lambda(t+y-\tau)} \ \frac{[\lambda(t+y-\tau)]^{r-r_1+j+\nu-1}}{(r-r_1+j+\nu-1)!} \quad .$$

After routine, but lengthy calculations involving in particular formulae (2), (6) and (7), we obtain:

$$(17) \qquad \int_0^\infty e^{-\xi t}dt \int_0^\infty \int_0^\infty e^{-s_1x-s_2y} d_{x,y} \Psi_t(x,y) =$$

$$\frac{\gamma^i(\xi) - h\left[s_1+s_2+\lambda-\lambda\gamma(s_1)\right]}{s_1+s_2-\xi+\lambda h\left[s_1+s_2+\lambda-\lambda\gamma(s_1)\right]-\lambda\gamma(s_1)}$$

The relationship between $\Phi_t(x,y)$ and $\Psi_t(x,y)$ is simple. We distinguish in $\Phi_t(x,y)$ between the cases in which the queue has never emptied out in $(o,t]$ and the one in which one or more beginnings of busy periods have occurred. We get in this way that:

$$(18) \qquad \Phi_t(x,y) - \Phi_t(0,0) =$$

$$\Psi_t^{(i)}(x,y) + \int_0^t \Psi_{t-u}^{(1)}(x,y)d\, M(u) \quad .$$

The superscript refers to the number of customers at the beginning of the timeperiod to which the Ψ-function relates.

Furthermore:

$$(19) \qquad \int_0^\infty e^{-\xi t}\, \Phi_t(0,0)dt = \gamma^i(\xi)[\xi+\lambda-\lambda\gamma(\xi)]^{-1}$$

by consideration of the renewal process of the ends of busy periods.

Finally, if we take the desired transforms in formula (18) and substitute from (4), (17) and (19), we obtain the expression given in (8). A direct way of proving (9) consists in multiplying by ξ in (8) and taking the limit as ξ tends to zero. One must then give an extra argument to show that the limit so obtained is the

Laplace-Stieltjes transform of a bivariate probability distribution. The easiest way of doing this is to note the second relation is (8) and to appeal to the known theorem concerning the limit law for the virtual waitingtime.

Moments of the joint limit law

We consider the moments of the limiting joint distribution of $\theta(t)$ and $\eta(t)$. These may be obtained by differentiations in formula (9). Since the calculations are lengthy, we only record the results found. We denote by M_θ and M_η the means of the distribution corresponding to (9). Provided the second moment α_2 of the service time distribution exists, we find:

$$M_\theta = \frac{\lambda\ \alpha_2}{2}\ (1 - \lambda\alpha)^{-2} \tag{20}$$

$$M_\eta = \frac{\lambda\ \alpha_2}{2}\ (1 - \lambda\alpha)^{-3}\ . \tag{21}$$

Likewise if the third moment α_3 of $H(\cdot)$ is finite, we find for the variances and covariance of the distribution (9):

$$\mathrm{var}_\theta = \frac{1}{12}\ \lambda(1 - \lambda\alpha)^{-3}\left[4\alpha_3 - 3\lambda(1 - \lambda\alpha)\alpha_2^2\right], \tag{22}$$

$$\mathrm{var}_\eta = \frac{1}{12}\ \lambda(1 - \lambda\alpha)^{-2}\left[4\alpha_3 - 3\lambda\left[(1-\lambda\alpha)^{-4}-2\right]\alpha_2^2\right], \tag{23}$$

and

$$\mathrm{cov}(\theta,\eta) = \frac{1}{12}\ \lambda(1-\lambda\alpha)^{-2}\left[4\alpha_3-3\lambda(1-\lambda\alpha)^{-1}\left[(1-\lambda\alpha)^{-2}-2\right]\alpha_2^2\right]. \tag{24}$$

We stress the fact that $\theta(t)$ was defined to be zero whenever $\eta(t) = 0$. Since moreover $\theta(t) = 0$ also implies that $\eta(t) = 0$ a.s. it is easy to find the joint distribution of the random variables $\theta(t)$ and $\eta(t)$ conditional on the fact that t is not during an idle period.

Acknowledgement

This problem originated during stimulating discussions with Dr. M. Yadin, to whom the author expresses thanks for his interest.

Bibliography

[1] Takács, Lajos (1961). The transient behavior of a single server queueing process with Poisson Input. Proc. Fourth Berkeley Symp., University of California press. 535-567.

[2] Takács, Lajos (1962). Introduction to the Theory of Queues. Oxford University press, New York.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Purdue University	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

The Joint Distribution of the Virtual Waitingtime and the Residual Busy Period for the M|G|1 Queue

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report

5. AUTHOR(S) *(Last name, first name, initial)*

Neuts, Marcel F.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August, 1967	11	2

8a. CONTRACT OR GRANT NO. NONR 1100(26) b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Mimeograph Series # 116 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research Washington, D. C.

13. ABSTRACT

In the M|G|1 queue, the joint distribution of the virtual waitingtime and the residual duration of the busy period at any time t, is given by a formula analogous to the one for the virtual waitingtime by itself. In the limit, as $t \to \infty$, the limiting joint distribution is given by a formula of the Pollaczek-Khintchine type.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
virtual waitingtime						
busy period						
M|G|1 queue						
Pollaczek-Khintchine formula						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

DD FORM 1 JAN 64 1473 (BACK)

D 23551

Unclassified

Security Classification