ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIV, 1 (1974)

H.-J. ROSSBERG and G. SIEGEL (Leipzig)

THE $GI/G/1$ MODEL WITH WARMING-UP TIME

1. Introduction. The object of the present paper is to survey the most important results about the waiting time distributions in single server queues with specialities at the beginning of a busy period.

We start by considering the basic model $GI/G/1$. The customers arrive in a recurrent stream; they are enumerated by $n = 0, 1, \ldots$ The corresponding service times $\beta_0, \beta_1, \ldots$ are independent and subject to the common distribution function (d. f.) $B(x)$. Analogously, the interarrival times $a_1, a_2, \ldots$ are independent with d. f. $A(x)$. We are particularly interested in the waiting times $w_0, w_1, \ldots$; they are defined recursively by

$$(1.1)\qquad w_{n+1} = \begin{cases} w_n + X_{n+1}, & w_n + X_{n+1} > 0 \\ 0, & w_n + X_{n+1} \leqslant 0 \end{cases} \qquad (n = 0, 1, \ldots),$$

provided that w_0 is given, and $X_n = \beta_{n-1} - a_n$.

This sequence of waiting times is intimately connected with the random walk $S_0 = 0$, $S_n = X_1 + \ldots + X_n$ (see fig.). As is easily seen from (1.1) we have $w_n = S_n$ as long as $S_n \geqslant 0$ if we assume $w_0 = 0$. In the figure, $S_3 < 0$; so the third customer is a "happy" one whose service starts immediately; the following waiting times can be obtained from the random walk S_n by a parallel translation (dotted line) as long as no other "happy" customer arrives. Then another translation is necessary and so on. As is well known (see, e. g., [3]) the condition

$$(1.2)\qquad \mathrm{P}\{S_n \to -\infty\} = 1$$

is necessary and sufficient for the existence of a stationary waiting time distribution $W(x)$, i. e. for the weak convergence

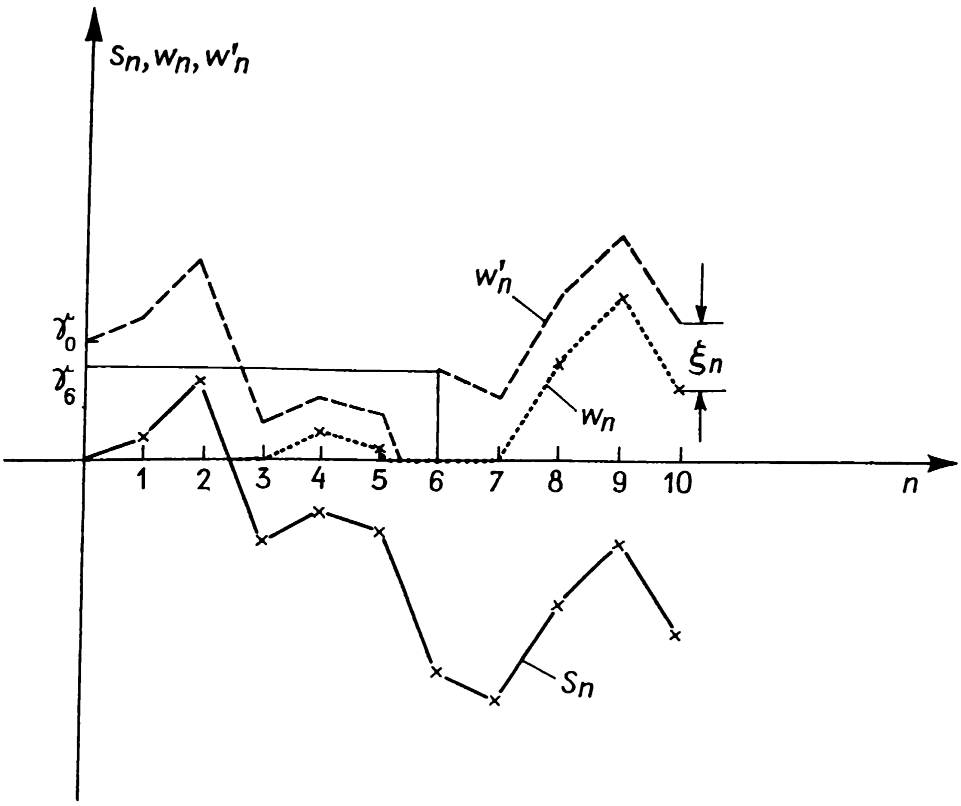
$$\mathrm{P}\{w_n < x\} \Rightarrow W(x) \qquad (n \to \infty).$$

This limit distribution is the only d. f. which is solution of the Wiener--Hopf equation

$$(1.3)\qquad W(x) = K * W(x) \qquad (x > 0).$$

In this report we consider two generalizations of this model which are characterized by peculiarities at the beginning of every busy period. We collect together the most important results from [11] and [15]-[18]. Some special cases were obtained by other authors as will be indicated. Only the general solution of (1.3) given by Klimova in [8] cannot be considered a special case of our results; it was obtained by algebraic methods, but it seems to be hard to pick out the special cases where rational characteristic functions (c. f.) are involved.

As a matter of surprise the study of models with warming-up time leads to new results concerning the basic model $GI/G/1$ (cf. theorems 3.2 and 3.3).



2. Finch's model. Once in a while it happens to all of us that we enter a service station, that there is no queue, but service does not start immediately. Sometimes the server has another occupation or, if the server is a machine, a random warming-up time is necessary. This experience was put in model by Finch in [4]. Further results where published in [2].

For every customer there exists a non-negative random variable (r. v.) γ_n $(n = 0, 1, \ldots)$; they are independent and subject to a common d. f. $C(x)$. In the model of the service process only a random subsequence of $\{\gamma_n\}$ is actually used; γ_n plays a role if and only if the n-th customer finds the server idle. So, for the sequence w_n' of waiting times, we have the recursive formula

$$w'_{n+1} = \begin{cases} w'_n + X_{n+1}, & w'_n + X_{n+1} > 0 \\ \gamma_n, & w'_n + X_{n+1} \leqslant 0 \end{cases} \qquad (n = 0, 1, \ldots). \tag{2.1}$$

A realization is given in the figure (dashed line). For $n = 0, \ldots, 5$ it can be obtained from $\{S_n\}$ by a parallel translation. We have $w'_n = w_n + \gamma_0$ $(n = 0, 1, 2)$ but $w'_n = w_n + \xi$ $(n = 3, 4, 5)$ with $0 < \xi < \gamma_0$. Then γ_6 must be used. Generally, we have $w'_n = w_n + \xi_n$ $(n = 0, 1, \ldots)$; the r. v. ξ_n are positive and depend on w_n. For example, it depends on the process $\{w_n\}$ for how long we have $w'_n = w_n + \gamma_0$. We tackelled this model with different methods.

a. *Results obtained by means of the random walk and renewal theory.* We first state a generalization of Finch's fundamental theorem [4]; it was proved in [16].

THEOREM 2.1. *Let us assume that the X_i are non-lattice and $K(x) < 1$ for $x < 0$ or $C(x) > 0$ for $x > 0$. If, moreover,* $\mathrm{P}\{S_n \to -\infty\} = 1$ *and* $m_C = \mathrm{E}\,\gamma_n < \infty$, *then there exists a stationary waiting time distribution W' such that*

$$\mathrm{P}\{w'_n < x\} \Rightarrow W'(x) \qquad (n \to \infty).$$

W' is the only d. f. satisfying

$$W'(x) = K * W'(x) - \beta \bar{C}(x) \qquad (x > 0), \tag{2.2}$$

where $\beta = K * W'(+0)$ and $\bar{C}(x) = 1 - C(x)$.

For $\mathrm{P}\{S'_n \to -\infty\} < 1$ *it follows* $\mathrm{P}\{w'_n < x\} \to 0$ $(n \to \infty)$.

Remark. Equation (2.2) is an inhomogeneous counterpart of (1.3); note that we obtain (1.3) putting $C(x) \equiv \varepsilon_0(x)$ (degenerated d. f. concentrated in the origin). This equation plays a considerable role in the theory, but it is not the only integral equation of interest.

THEOREM 2.2 (cf. [16]). *If $W'(x)$ exists, then also*

$$\lim_{n \to \infty} \mathrm{P}\{\xi_n < x\} = G(x)$$

exists; G is the unique d. f. satisfying

$$G(x) = U_- * G(x) - \delta \bar{C}(x) \qquad (x > 0), \tag{2.3}$$

where $\delta = U * G(+0)$, and $U_-(x) = 1 - U(-x)$ is the d. f. of the idle time in the basic model $GI/G/1$.

Putting $\bar{G}(x) = 1 - G(x)$, we have the representation

$$\bar{G}(x) = \delta \Big[\bar{C}(x) + \int_0^\infty M(y)\, dC(y + x) \Big], \tag{2.4}$$

where

$$M(y) = \sum_{i=1}^{\infty} U_{-}^{i*}(x)$$

is the renewal function corresponding to $U_{-}(x)$.

Remarks. 1. Equations (2.2) and (2.3) resemble each other on the very first glance, but there is an essential difference concerning the kernels: $0 < K(0) < 1$, but $U(+0) = 1$.

2. $M(y)$ is a complicated function even if $U_{-}(x)$ is known explicitely, but often we do not know much about $U_{-}(x)$. But there exist estimations below and above for the renewal function, and each of them yields an estimation of (2.4); some of them can be found in [18].

3. In case of the basic model $M/G/1$ (with parameter λ of the Poisson process) we have $M(y) = \lambda y$, and so we obtain from (2.4) the simple formula

$$G(x) = \delta\Big[C(x)+\lambda\int_{0}^{x}\bar{C}(y)\,dy\Big], \qquad \delta = (1+\lambda m_c)^{-1} \qquad (x > 0). \tag{2.5}$$

So far the essential role of the function $G(x)$ has not become sufficiently evident, but it is made clear by the following

THEOREM 2.3 (cf. [16]). *If* $W'(x)$ *exists, then* w_n *and* ξ_n $(n = 1, 2, \ldots)$ *are asymptotically independent and* [1]

$$W' = W*G. \tag{2.6}$$

Remark. The first part of the theorem is the deeper result. Representation (2.6) is well known from the general theory of the inhomogeneous Wiener-Hopf equations. Interesting is that in this case G is a d. f. with an intuitive meaning.

On the other hand, W can be considered as known from the vast literature on the basic model. Therefore, our interest focuses to the relations between $C(x)$ and $G(x)$. We give the following examples (see [18]):

(i) The estimations of $M(y)$ in (2.4) yield that for

$$\int_{0}^{\infty} x\,dA(x) < \infty$$

it holds, for every $a > 0$,

$$\int_{1}^{\infty} x^{a}\,dG(x) < \infty \text{ if and only if } \int_{1}^{\infty} x^{a+1}\,dC(x) < \infty.$$

[1] The first special case of (2.6) was established in [11].

(ii) If $\bar{C}(x+y) \leqslant \bar{C}(x)\bar{C}(y)$, then $\bar{G}(x) \leqslant \bar{C}(x)$. Further, in this assertion $\leqslant$ can be replaced by $\geqslant$ (see [18]). Note that this is a generalization of one part of theorem 2.6.

(iii) Now let us suppose that

$$C(x) \leqslant \frac{1}{m_c} \int_0^x \bar{C}(y)\,dy \qquad (x \geqslant 0).$$

(For instance, this is true if the assumption of (ii) holds.) Then we obtain

$$\int_0^\infty x^n\, dG(x) \leqslant m_c^n n! < \infty \qquad (n = 1, 2, \ldots).$$

If the moments of G exist, then we can replace the signs $\leqslant$ by $\geqslant$ in both relations.

In theorem 2.4 we shall give another estimation by comparing two different d. f. C_1 and C_2. Setting

$$S_c^0(x) = \bar{C}(x) \quad \text{and} \quad S_c^{n+1}(x) = \int_x^\infty S_c^n(y)\,dy \qquad (n = 0, 1, \ldots),$$

it was proved in [18].

THEOREM 2.4. *Let, for some* $n \geqslant 0$,

$$\int_0^\infty x^{n+1}\, dC(x) < \infty \quad \textit{and} \quad S_{C_1}^n(x) \leqslant S_{C_2}^n(x) \qquad (x \geqslant 0).$$

Then

$$\frac{\beta_2}{\beta_1} S_{W_1'}^n(x) \leqslant S_{W_2'}^n(x) \qquad (x \geqslant 0).$$

For $n = 0$ *it follows* $\beta_2 \leqslant \beta_1$. *But for* $n > 1$ *this assertion is not true generally.*

Analogously, we compare different d. f. K_1 and K_2 for fixed C. For instance, from $K_1 \geqslant K_2$ it follows

$$\frac{\beta_2}{\beta_1} \bar{W}_1' \leqslant \bar{W}_2' \quad \text{and} \quad \beta_2 \leqslant \beta_1.$$

It is important to choose A, B or C such that $m_{W'}$ attains a minimum. As well known about $GI/G/1$ (see [10]), for fixed A, $\min\limits_{\forall B,\, m_B=\text{const}} m_W$ is attained for $B \equiv \varepsilon_{m_B}$ (degenerated d. f. concentrated in m_B) and, for fixed B, $\min\limits_{\forall A,\, m_A=\text{const}} m_W$ is attained for $A \equiv \varepsilon_{m_A}$. For $GI/G/1$ with warming-up time, ε_{m_B} and ε_{m_A} are only good solutions but generally not the best ones. Analogously, for fixed A, B and $m_C = \text{const}$ the special function $C \equiv \varepsilon_m$ is a good, but generally not the best solution (see [18]).

Now let us consider the model $M/G/1$ with warming-up time, where the parameter λ = const. In this case, as a matter of surprise,

$$B \equiv \varepsilon_{m_B} \quad \text{for } m_B = \text{const and fixed } C$$

or

$$C \equiv \varepsilon_{m_C} \quad \text{for } m_C = \text{const and fixed } B$$

are the best solutions. We have another pleasant case, if C is a negative exponential distribution with parameter λ, then Rogozin's results mentioned above hold true also in the model with warming-up time.

b. *Results obtained by analytical methods.* If we want to treat (2.2) and (2.3) with characteristic functions, we must add a term such that we have equations not only for $x > 0$ but on the whole line. It is an interesting feature of our special case that these additional terms have an intuitive meaning while they are not of interest themselves in the general theory of integral equations on the half-line. For (2.3), e. g., we obtain [15]

$$\delta U' + G = U * G + \delta C,$$

where $1 - U'(-x)$ is the d. f. of the idle time in Finch's model. Denoting the c. f. by asterisks, we now get

$$G^*(t)\big(1 - U^*(t)\big) + \big(U'^*(t) - C^*(t)\big) = 0. \tag{2.7}$$

As is well known, the general solution of (1.3) and (2.2) is intimately connected with the Wiener-Hopf factorization. But special cases where rational c. f. are involved can be treated in a much more elementary way; we pointed it out in [11], see also [6], chapter 5. This method can also be applied to (2.7) (see [1]) and yields, for instance,

THEOREM 2.5. *G^* rational if and only if C^* rational, and both functions have the same poles.*

As the poles of G^* and C^* coincide we obtain immediately the following

CONSEQUENCE. *G is a finite mixture of exponential d. f. if and only if so is C.*

Analytical considerations yield also [16]

THEOREM 2.6. *G or C are exponential d. f. if and only if $G \equiv C$.*

Note that one part of the theorem follows immediately from the above-mentioned consequence. The other rests on a unity theorem for an integral equation proved in [12], namely, putting $C \equiv G$ in (2.3) we obtain, after some rearrangements $(C_-(x) = 1 - C(-x+0))$,

$$(1-\delta) C_-(x) = C_- * U_-(x) \qquad (x \leqslant 0).$$

A representation for $G^*(z)$ in $\{z = t + iy;\ y > 0\}$ can also be obtained by the following idea (cf. [15]), which makes use of the fact $U(+0) = 1$ (cf. remark 1 to theorem 2.2). In this case the general theory of inhomogeneous Wiener-Hopf equations ([2]) proves superfluous. From the very beginning we know, by the inversion formula of Laplace-transforms,

$$G^*(z)-G^*(+0) = \frac{1}{2\pi i}\,\mathrm{VP}\int_{-\infty}^{\infty} \frac{G^*(t)-G(+0)}{v-z}\,dv$$

(VP denotes "valeur principale"). Substituting $G^*(t)$ by means of (2.7), we obtain after some rearrangements the following

THEOREM 2.7. *Let be*

$$\int_{\infty}^{0} K(x)\,dx < \infty \quad \textit{and} \quad \overline{\lim_{t\to\infty}}|K^*(t)| < 1.$$

Then we have, for the solution of (2.3),

$$W(+0)[G^*(z)-G(+0)] = \frac{\delta}{2\pi i}\,\mathrm{VP}\int_{-\infty}^{\infty}\left[\frac{C^*(v)-1}{1-U^*(v)}+1-C(+0)\right]\frac{dv}{v-z}.$$

Remark. This formula can be inverted in several special cases, e. g. we obtain (2.5) under more general assumptions. Furthermore, we find a new proof of the general formula (2.4).

3. Yeo's model. We next consider a model due to Yeo [21] which is intimately connected with Finch's model. But there is no warming-up time. Welch [20] considers this model independent of Yeo. This model with further specialities was considered in [19].

So far we had exactly one random service time $\beta_0, \beta_1, \ldots$ for each of the successive customers. In Yeo's model we have two service times for each customer. In other words, there is a second sequence $\hat{\beta}_0, \hat{\beta}_1, \ldots$ of independent positive random variables subject to the common d. f. $\hat{B}(x)$. For each customer his actual service time is determined by the following rule: If the n-th customer finds the server busy, then β_n; if he finds the server idle, then $\hat{\beta}_n$. So this model may be considered a very simple case of a system in which the service time depends on the queue length. The counterpart to formulas (1.1) and (2.1) now reads as

$$(3.1)\qquad \hat{w}_{n+1} = \begin{cases} \hat{w}_n + X_{n+1}, & \hat{w}_n + X_{n+1} > 0,\ \hat{w}_n > 0, \\ \hat{X}_{n+1}, & \hat{X}_{n+1} > 0,\ \hat{w}_n = 0, \\ 0, & \hat{w}_n = 0,\ \hat{X}_{n+1} \leqslant 0 \ \text{or}\ \hat{w}_n + X_{n+1} < 0,\ \hat{w}_n = 0, \end{cases}$$

where $X_n = \beta_{n-1} - a_n$, $\hat{X}_n = \hat{\beta}_{n-1} - a_n$ $(n = 1, 2, \ldots)$. Obviously, there are two interesting special cases:

a. $B \equiv \hat{B}$ (basic model);

b. $\tilde{\beta}_n = \beta_n + \gamma_n$, where $\mathrm{P}\{\gamma_n \geqslant 0\} = 1$.

([2]) It was applied to (2.2) in [9] where more restrictions had to be made.

In the latter case we have Finch's model with a minor alteration: In Yeo's model every waiting time at the begin of a busy period is zero, but in Finch's model we have $w'_n > 0$ $(n = 0, 1, \ldots)$ provided that $P\{\gamma_n > 0\} = 1$. Namely, in Finch's case, $\gamma_{n_\nu} = w'_{n_\nu}$ holds for a random subsequence n_ν of the natural numbers, but in Yeo's model γ_n is part of a service time. This difference is unessential as the sojourn times of all customers coincide.

So, in a certain sense, Yeo's model is a generalization of Finch's. But in spite of that there is a trick to reduce the study of Yeo's model to that of Finch's. Let us consider the sequence of waiting times defined by (3.1). Then, we pick out a random subsequence by cancelling all those $\hat{w}_n$ which are equal to zero. As can be shown, this procedure yields a sequence w'_ν for which (2.1) holds; we call it the *imbedded model*; the corresponding d. f. $C(x)$, obviously, depends on K, B, $\hat{B}$. This is the basic idea on which the proof of theorem 3.1 rests [17].

THEOREM 3.1. *If* $\lim_{n\to\infty} P\{\hat{w}_n < x\} = \hat{W}(x)$ *exists, then*

$$\hat{W}(x) = a\varepsilon_0(x) + (1-a)W * G(x) \qquad (x > 0), \tag{3.2}$$

where $a = \hat{W}(+0)$, W *can be considered as known from the basic model and* $G = G(K, B, \hat{B})$ *is determined by the imbedded model.*

A special case of (3.2) was found in [7] and suggested the supposition that (3.2) might hold generally; in this paper the analytical method mentioned in section 2b was applied.

Now, we return to the basic model $GI/G/1$ putting $\hat{B} \equiv B$. Therefore, $\hat{W} \equiv W$ and (3.2) turns into a renewal equation for W, where $G = G(K, B, \hat{B})$ is determined only by the basic model. It is well known (see [3]) that W is not only the solution of (1.3) but also the solution of the renewal equation

$$W(x) = 1 - H(\infty) + H(\infty) W * H(x) \qquad (x > 0),$$

where $H(x)$ is the distribution of the "ascending ladder heights" of the random walk $\{S_n\}$. As we assume (1.2), the renewal process of these ladder heights must terminate and, therefore, $H(\infty) < 1$. Generally, it is difficult to find explicite formulas for $H(x)$. So, the following result (see [17]) might be of some interest:

THEOREM 3.2. $H(\infty) G(x) \equiv H(x)$ *if and only if*

$$C(x) \equiv \frac{K(x) - K(0)}{1 - K(0)} \qquad (x > 0).$$

The roundabout way *via* the imbedded model yields also [18]

THEOREM 3.3. *Let be $a > 0$ and $m_A < \infty$. In the basic model GI/G/1, $\int_0^\infty x^a dW(x)$ exists if and only if $\int_0^\infty x^{1+a} dB(x)$ exists.*

Only very few special cases of this result seem to be known.

References

[1] M. Dewess, *Erhaltungssätze im Modell GI/G/1 mit verzögertem Bedienungsbeginn*, Operationsforschung u. Math. Statistik, submitted.

[2] W. J. Ewens and P. D. Finch, *A generalized single-server queue with Erlang input*, Biometrika 49 (1962), p. 242-245.

[3] W. Feller, *An introduction to probability theory and its applications*, Vol. II, J. Wiley, New York 1966.

[4] P. D. Finch, *A probability limit theorem with application to a generalization of queuing theory*, Acta Math. Acad. Sci. Hung. 10 (1959), p. 317-325.

[5] I. Gherghei (И. Гергей), *Однолинейная система обслуживания с „разогревом"*, in *Вычислительные методы и программирование*, Москва 1967, p. 99-107.

[6] B. W. Gnedenko and I. N. Kovalenko, *Einführung in die Bedienungstheorie*, Berlin 1971.

[7] H.-J. Haeske, *Eine verallgemeinerte Behandlung der Integralgleichung von G. F. Yeo nach einer Methode von H.-J. Rossberg*, Operationsforschung und Math. Statistik 2 (1970), p. 31-38.

[8] E. S. Klimova (E. C. Климова), *Исследование однолинейной системы обслуживания с „разогревом"*, Техническая кибернетика 1 (1968), p. 91-97.

[9] P. G. Petrov (П. Г. Петров), *Распределение время ожидания в стационарном режиме для одной системы обслуживания с „разогревом"*, in *Вычислительные методы и программирование*, Москва 1967, p. 117-130.

[10] B. A. Rogozin (Б. А. Рогозин), *Некоторые экстремальные задачи теории массового обслуживания*, Теор. вероятн. и примен. 11 (1966), p. 161-169.

[11] H.-J. Rossberg, *Eine neue Methode zur Behandlung der Lindleyschen Integralgleichung und ihrer Verallgemeinerung durch Finch*, Elektr. Informationsverarbeitung u. Kybernetik 3 (1967), p. 215-238.

[12] — *Characterization of the exponential and the Pareto distributions by means of some properties of the distributions which the differences and quotients of order statistics are subject to*, Operationsforschung u. Math. Statistik 3 (1972), p. 207-211.

[13] — and G. Siegel, *Die Bedeutung von Kingmans Integralungleichungen bei der Approximation der stationären Wartezeitverteilung im Modell GI/G/1 mit und ohne Verzögerung beim Beginn einer Beschäftigungsperiode*, ibidem, submitted.

[14] — *On Kingman's integral inequalities for approximations of the waiting time distribution in the queuing model GI/G/1 with and without warming-up time*, this fascicle, p. 27-30.

[15] — *Новый метод к изучению распределения времени ожидания в стационарном режиме для системы массового обслуживания GI/G/1 с „разогревом"*, in preparation.

[16] G. Siegel, *Verallgemeinerung einer Irrfahrt im R^1 und deren Bedeutung für Bedienungssysteme mit verzögertem Bedienungsbeginn*, Operationsforschung u. Math. Statistik, submitted.

[17] — *The stationary waiting time and other variables in single-server queues with specialities at the beginning of a busy period*, Zastosow. Matem. 13 (1973), p. 465-479.

[18] — *Abschätzungen für die Wartezeitverteilung und ihre Momente beim Wartemodell GI/G/1 mit Erwärmung,* Zeitschr. angew. Math. u. Mechanik, submitted.

[19] S. Takomatsu, *Queueing processes with accumulated service,* Ann. Inst. Statist. Math. 22 (1970), p. 349-379.

[20] P. D. Welch, *On a generalized M/G/1 queuing process in which the first customer of each busy period receives exceptional service,* Operat. Res. 12 (1964), p. 736--752.

[21] G. F. Yeo, *Single server queues with modified service mechanisms,* J. Australian Math. Soc. 2 (1962), p. 499-507.

SEKTION MATHEMATIK
KARL-MARX-UNIVERSITÄT
701 LEIPZIG, DDR

Received on 26. 2. 1973

H.-J. ROSSBERG i G. SIEGEL (Leipzig)

SYSTEM *GI/G/1* Z ROZGRZEWANIEM

STRESZCZENIE

Praca zawiera streszczenie ważniejszych wyników dotyczących stacjonarnych rozkładów czekania w dwóch systemach obsługi masowej z wyróżnionym zachowaniem się na początku okresu zajętości (bez dowodów). Systemy te, będące prostymi uogólnieniami systemu *GI/G/1*, były po raz pierwszy badane przez Fincha [4] i Yeo [21]. Wyniki otrzymano przy użyciu metod błądzenia przypadkowego, teorii odnowy i metod analitycznych.