PERFORMANCE '81
F.J. Kylstra (editor)

A STATIC PRIORITY QUEUE WITH PIECEWISE MARKOVIAN DISTRIBUTED SERVICE TIME AND WAITING-TIME OPTIMAL DISCIPLINE

Bernhard Walke

AEG-TELEFUNKEN, Research Institute Ulm,
Sedanstraße 10, D-7900 Ulm, Germany

An infinite queue single-server model with preemptive resume priorities is considered. Customers with priority $i (1 \leq i \leq N)$ arrive according to a Poisson process, join the i-th priority level and demand service according to a general distribution function. In each level the service discipline may be either FCFS of MLFCFS (multi-level FCFS). In the MLFCFS discipline customers are assigned to different important sublevels within a priority level, dependent on the amount of attained service. The customer of a more important sublevel has preemptive priority over customers of all less important sublevels of the same priority level. The PM (piecewise Markovian) d.f. is introduced as a possible approximation for general d.f.'s . Using this approximation for service time with a large coefficient of variance the MLFCFS discipline gains the minimum mean waiting time.

Computation of mean waiting time, both with and without condition on a customers service requirement, is demonstrated by means of examples. The appendix gives a sketch of the proof for the optimality of the used discipline.

1. INTRODUCTION

In real-time computer systems sometimes concurrent customers of different importance have to be serviced. This is, usually, taken into account by assigning static priority numbers to customers and serving the highest priority (lowest priority number) customers first.

In this paper we consider service disciplines in terms of a queueing model. Customers with priority i are completely described by their interarrival and service times. From theory of scheduling [7, 8] it is know that customers with a large squared coefficient of variance C^2 (variance over squared mean) of their service-time distribution, namely $C^2 > 1$, must be serviced by means of an appropriate feedback (FB) discipline to reach a minimum mean waiting time. We observe this law by applying to customers of each priority level such a multi-level FB discipline that their mean waiting time is minimized.

Multi-level queueing models with static priorities and FB disciplines are rarely to be found in the literature [1, 2, 3]. Typical for these papers is that service quanta are used which are not interruptable. Moreover the external priority of a customer is not consequently observed there: a customer with lower external priority may get service before another with higher external priority and large service time requirement, due to the FB disciplines applied. These models may be called multi-level feedback models with arrivals not only at the first but also subsequent levels.

Our model differs from these models in that 1) the external priority of a customer is used to guarantee preferential service over all lower priority customers, 2) the FB discipline with finite or infinite quantum size is used only within each priority level, 3) preemption (of the resume type) of service quanta within a priority level and between different priority levels is allowed. The service discipline within each of the feedback queues is FCFS. Using service quanta of infinite lenght within each priority level, our model degenerates to the well-known model with preemptive resume priorities, first studied by White and Christie [4] and Miller [5].

In sections 2 and 3 we introduce the model and define the mean waiting time in the MLFCFS discipline, respectively. In section 4 we present the application to customers with piecewise Markovian (PM) service-time distribution.

2. THE MODEL

We consider an infinite queue model of a computer system, cf. Fig. 1, with N static-priority levels of the preemptive resume type.

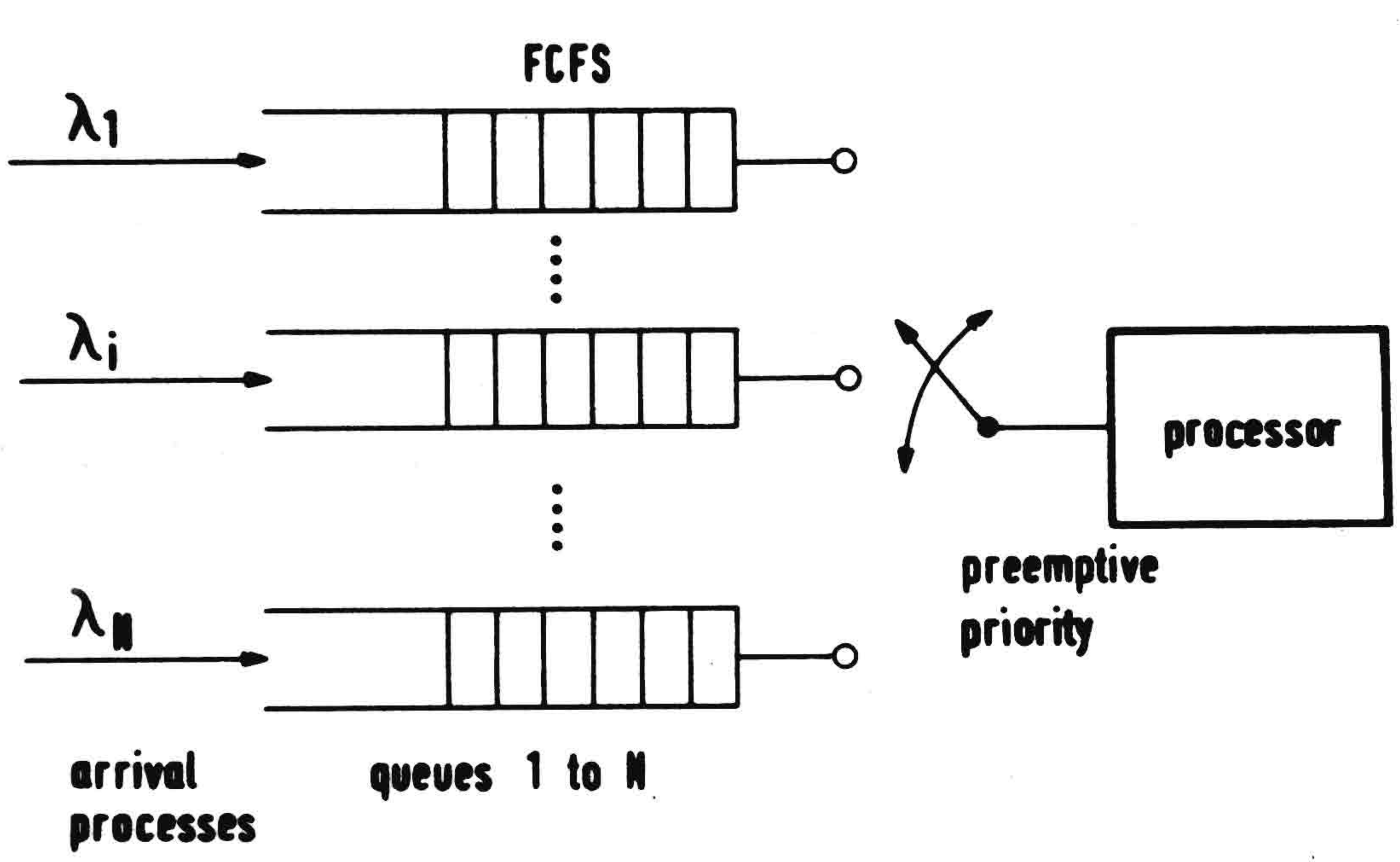

Fig. 1: Model with N queues one for each external preemtive-resume priority level

Customers with, say, priority number i ($1 \leq i \leq N$) arrive from a Poisson arrival process with parameter λ_i and demand service. Their service times are assumed to be independent and indentically distributed with general distribution function (d.f.) and finite second moment. The related arrival processes are assumed to be independent, too. Apparently, we consider a multi-entrance M/G/1 model. Small priority numbers denote high priorities. Within priority level i we use the multi-level first come first serve (MLFCFS) diciplіne which uses interruptable service quanta. The functioning of this discipline can be explained by means of the model in Fig. 2 and by Table I.

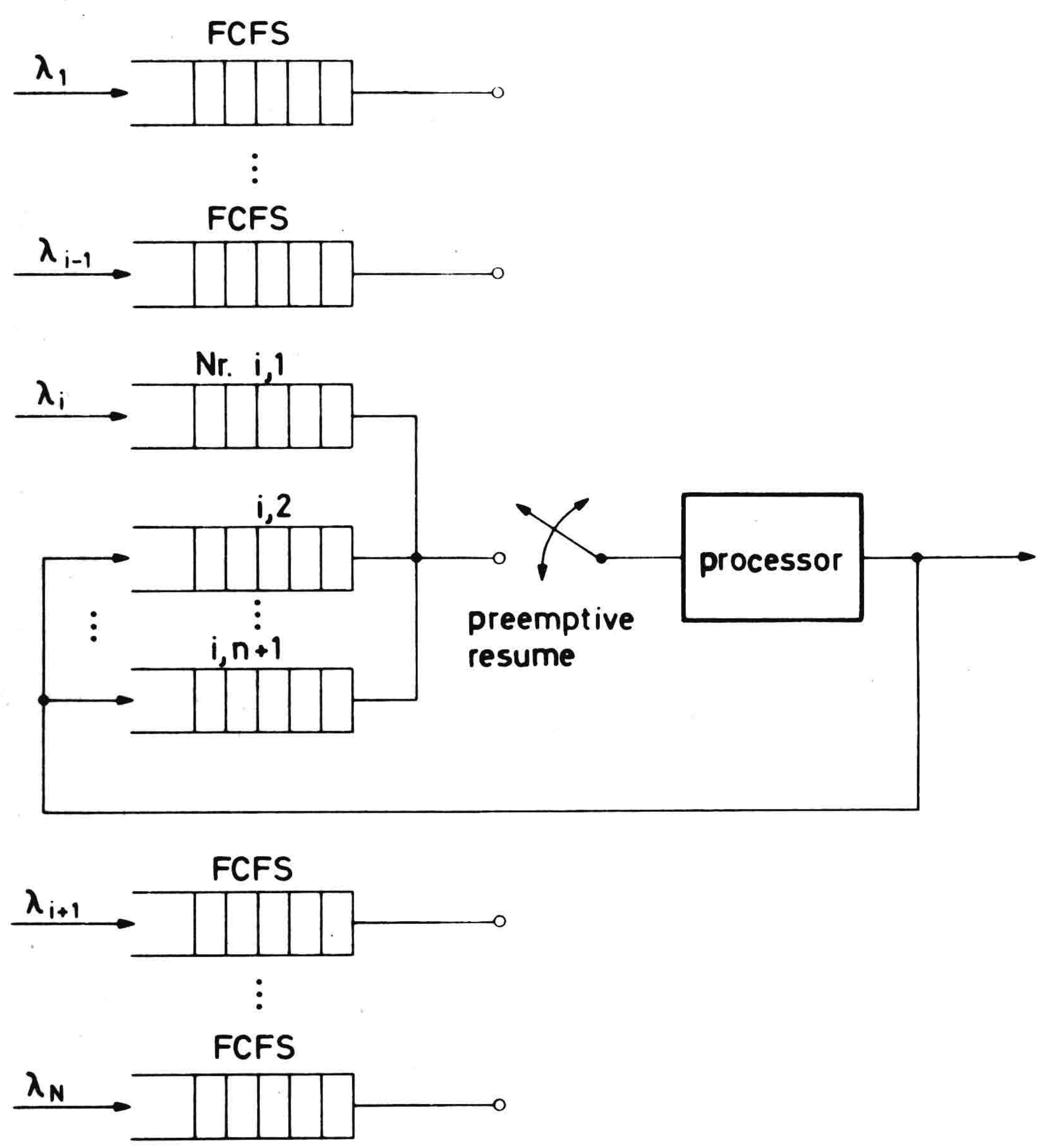

Fig. 2: Model with N external preemtive-resume priority levels and the MLFCFS and FCFS disciplines applied to levels i and $j \neq i$, respectively ($i,j \in N$).

In the MLFCFS discipline a new arriving customer of priority i joins the queue number i,1, where he remains until either his service requirement is fulfilled during the first quantum of size $t_{gi,1}$ and he leaves the model or he has consumed the total quantum. Customers which need more than the first quantum are fed back, when the quantum has expired, to queue number i,2 where the quantum size is $(t_{gi,2}-t_{gi,1})$. Customers which need still more service are fed back to queue number i,3, where the quantum size is $(t_{gi,3}-t_{gi,2})$, and so on. In queue number i, n+1 the quantum size is infinite, cf. Table I.

level of the queueing system	i,1	i,2	...	i, n+1
attained service time t is in the interval	$0 \leq t < t_{gi,1}$	$t_{gi,1} \leq t < t_{gi,2}$	...	$t_{gi,n} \leq t < t_{gi,n+1} = \infty$

Table I: Attained service time in queue number i,k of the model in Fig. 2

Upon its arrival a customer with lower priority index $j<i$ interrupts immediately any other customer of priority i. Moreover service of queues number (i,k), $(2 \leq k \leq n+1)$, is interrupted in favour of a new arriving customer in queue number (i,1). Preemted service is resumed later on at the same sublevel. Inside each queue the service discipline is FCFS. It can easily be seen that, for an infinite quantum assigned to customers of the first queue (i,1) of priority i, the other queues inside this level are obsolete and a simple FCFS queue remains there.

3. PREVIOUSLY KNOWN RESULTS

Computation of the first two moments of the mean waiting time of the M/G/1 model in Fig. 1, presuming Poisson arrival processes, arbitrary service-time d.f.'s and

FCFS scheduling in each priority level, is known [4,5]. The feedback model inserted in level i in Fig. 2 has been analysed separately in the MLFCFS discipline and the first moment of the mean waiting time was found in [6] for an arbitrary number of levels. The complete model in Fig. 2 has been analysed in [10]. The main results are shortly repeated in section 3.1.

3.1 Results of the static preemtive-resume priority model with MLFCFS in priority level i

First note that the waiting time of customers in level i does not depend on the service disciplines applied to levels $j \neq i$. This follows from the assumption of preemptive-resume priority for levels j, (j=1,2,...N), in the model in Fig. 2. Therefore it is sufficient to know the results for only one level $i>1$.
The different queues at priority level i are indexed (i,k) with $(1 \leq k \leq n+1)$ and are called sublevels.
The following abbreviations are used thoughout:

$\lambda_{\leq i} = \sum_{j=1}^{i} \lambda_j$	arrival rate up to and including level i,
$\rho_i = \lambda_i / \beta_i$	traffic offered to priority level i,
$\rho = \sum_{i=1}^{N} \rho_i$	offered traffic to priority level 1 through N, total offered traffic,
$w_i^{(r)}, \beta_i^{(r)}$	r^{th} moment of waiting time and service time d.f.'s, respectively
$w_i = w_i^{(1)}, \beta_i = \beta_i^{(1)}$	means (= first moments)
$B_i(t) = P(b_i \leq t)$	probability of a service-time b_i not to exceed t; service-time d.f. of customers of level i
$B_{\leq i,k}(t)$	weighted service-time d.f. of all customers with priorities 1 through (i-1) and those of priority i, which leave the system from sublevel (i,k) or lower.

$$B_{\leq i,k}(t) = \lambda_{\leq i}^{-1} \{ \sum_{j=1}^{i-1} \lambda_j B_j(t) + \lambda_i P_i(t) \} \tag{3.1}$$

with
$$P_i(t) = \begin{cases} B_i(t) & \text{for } 0 \leq t \leq t_{gi,k} \\ 1 & \text{otherwise} \end{cases}$$

$t_{\leq i,k}^{(r)}$	r^{th} moment of $B_{\leq i,k}(t)$.

$$t^{(r)}_{\leq i,k} = \lambda^{-1}_{\leq i} \{ \sum_{j=1}^{i-1} \lambda_j \beta_j^{(r)} + \lambda_i \, t^{(r)}_{i,k} \} \tag{3.2}$$

$t^{(r)}_{i,k}$ r^{th} moment of $B_i(t)$ truncated at $t = t_{gi,k}$

$$t^{(r)}_{i,k} = \int_0^{t_{gi,k}} t^r \, dB_i(t) + t^r_{gi,k} \{1-B_i(t_{gi,k})\}$$

$$\gamma^{(r)}_{\leq i,k} = t^{(r)}_{\leq i,k} \lambda_{\leq i} .$$

Using these expressions the mean waiting time $W_i(t)$ of customers depending on their service requirement t and the mean waiting time W_i are

$$W_i(t) = (W_{i,k} + t\,\gamma_{\leq i,k-1})/(1-\gamma_{\leq i,k-1}) \tag{3.3}$$

$$\text{for } t_{gi,k-1} < t \leq t_{gi,k}$$

with $$W_{i,k} = \gamma^{(2)}_{\leq i,k} / \{2(1-\rho_{<i} - \lambda_i t_{i,k})\}$$

$$W_i = \sum_{k=1}^{n+1} \frac{P_{i,k}(W_{i,k} + E_{i,k})}{1 - \gamma_{\leq i,k-1}} - \beta_i \, , \tag{3.4}$$

with $$P_{i,k} = B(t_{gi,k}) - B(t_{gi,k-1}) \tag{3.5}$$

$$E_{i,k} = 1/P_{i,k} \int_{t_{gi,k-1}}^{t_{gi,k}} t \, dB_i(t) \; . \tag{3.6}$$

4. MINIMUM WAITING TIME OF CUSTOMERS WITH PM d.f.

First we recall that for customers with a large coefficient of variance, C>1, an appropriate feedback discipline gains a smaller mean waiting time than the FCFS discipline. If C≤1 than FCFS is an optimal discipline.
It is common practice to approximate service-time d.f.'s with C>1 by members of the hyperexponential d.f. family which is defined by

$$P(\leq t) = 1 - \sum_{i-1}^{n} w_i \, e^{-\mu_i t}; \quad \mu_i > 0, \; w_i > 0, \; \sum_{i=1}^{n} w_i = 1. \tag{4.1}$$

Such approximations imply that the FB_∞ diciplime with (in some situations) infinite small service quanta must be used to yield the minimum mean waiting time [7,8], cf. the appendix. Such small quanta are unacceptable for practical applications due to the vast amount of switching overhead associated with the FB_∞ discipline. Therefore in practice suboptimal feedback disciplines with finite sized quanta, are used. For example problems arise in sizing the quanta to be applied. In the next section we use another family of d.f.'s to approximate service-time d.f.'s with arbitrary C. This family is useful to approximate d.f.'s with C>1 in that the MLFCFS discipline (with finite sized quanta) then minimizes the mean waiting time, cf. the appendix. Moreover, it will be seen later on that the approximation selected defines the number of feedback levels and the sizes of the service quanta.

4.1 Piecewise Markovian d.f.'s

During this section we drop the static priority index i, i.e. we consider B(t) instead of $B_i(t)$. A piecewise exponential d.f. (PM-d.f.) firstly was used in [9] to demonstrate the optimality of a two level MLFCFS discipline. The family of the PM d.f.'s is introduced in [11] where the mean waiting time for a two and three-level MLFCFS discipline is computed. The optimal choice of the sizes of quanta to service customers with PM service-time d.f. was studied and solved in [9] and [11].

Compared to other efforts to approximate service-time d.f.'s by means of piecewise functions, e.g. [12], our approach has as a main advantage that the markovian property is sustained during the intervals of the service time t, ($t_{gk-1} < t \leq t_{gk}$), which form the pieces. In contrary to the M d.f.

$$P(t) = 1 - \exp(-\mu t)$$

where the departure rate μ of a customer being serviced in a server remains constant, independent of the attained service-time t, this rate is μ_1 for customers with attained service t in the interval $\{0 = t_{go} < t \leq t_{g1}\}$, μ_2 for customers with attained service in the interval $\{t_{g1} < t \leq t_{g2}\}$, and so on until finally the tail of the PM d.f. is reached, which has a rate μ_{n+1} and simply is a M d.f. The PM d.f. of order (n+1) is defined by (cf. [11])

$$P(t) = 1 - \exp\left\{-\sum_{i=1}^{j} t_{gi}(\mu_i - \mu_{i+1}) - \mu_{j+1} t\right\} \tag{4.3}$$

$$\text{for } t_{gj} \leq t \leq t_{gj+1}$$

$$j = 0,1,\dots,n;\; 0 = t_{go} < t_{g1} \dots < t_{gn+1} = \infty .$$

A PM d.f. of first order is a M d.f. By choosing appropriate parameter values t_{gi} and μ_i and a sufficient number of pieces almost general d.f.'s can be approximated, see Fig. 3. By including zero and infinite service rates also step functions are possible.

We aim here at demonstrating the usefulness of PM d.f.'s to approximate service-time d.f.'s with C>1 and decide to consider only the subfamily defined by the relation

$$\mu_1 < \mu_2 < \dots \mu_i \dots < \mu_{n+1} . \tag{4.4}$$

Just this subfamily is comparable to the hyperexponential d.f.'s.

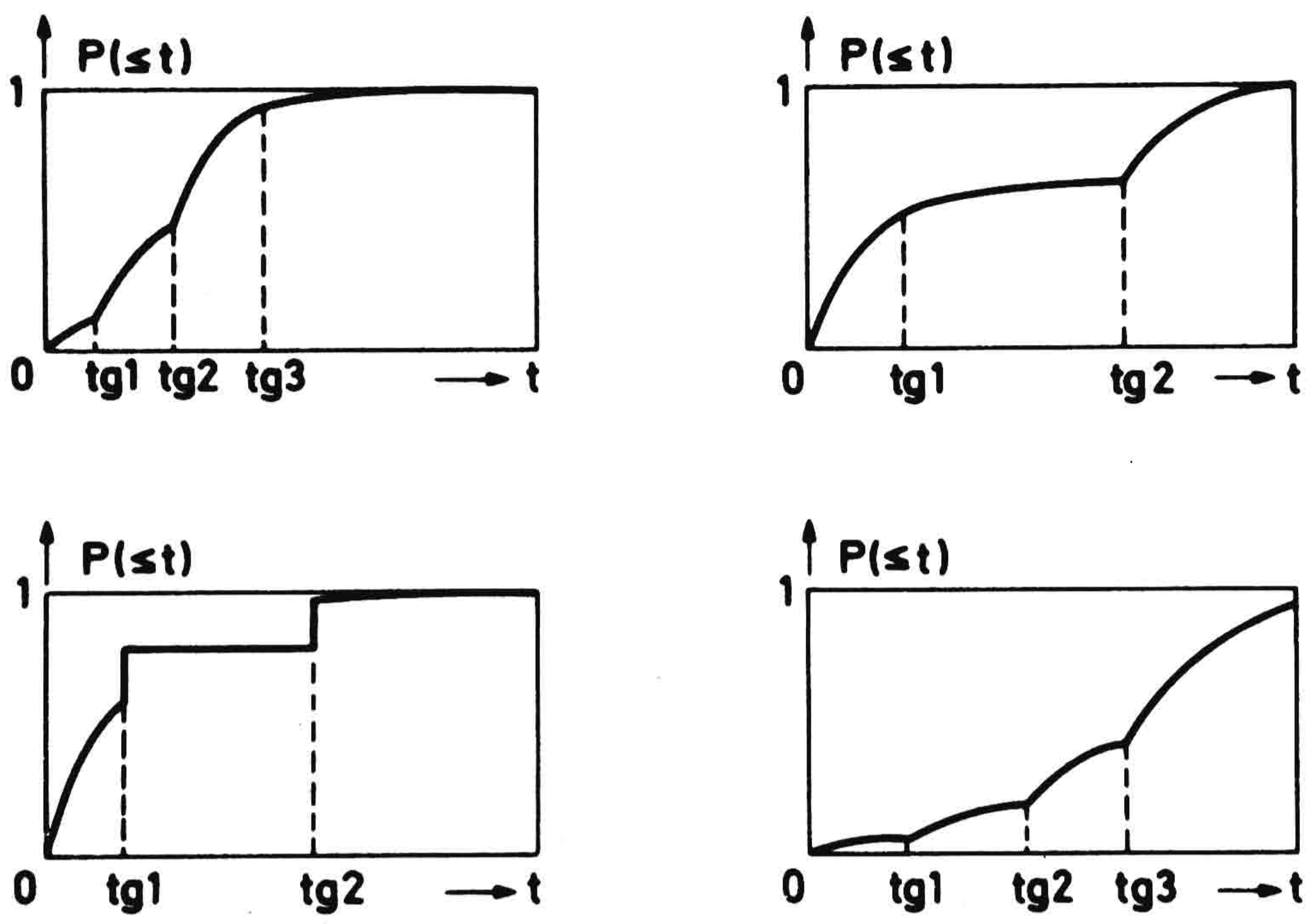

Fig. 3: Examples of piecewise markovian (PM) d.f.'s

The parameters t_{gj} in Eq (4.3) form the interval limits for the MLFCFS discipline, cf. Table I. Using results of [7], cf. the appendix, it is shown in [9] that a combination of the FCFS discipline, applied to customers with a service requirement belonging to the same interval, and the preemtive resume priority discipline applied to customers belonging to different intervals, minimizes the mean waiting time. Thereby customers with attained service time in the interval j have the preemtive priority j as long as they are member of level j. Discussing the MLFCFS discipline we use the term level instead of service interval to point out that an individual priority number is associated with each interval. The related model is that of level i in Fig. 2 with an input process only to the first queue of the queueing system, as is shown there. The term level instead of interval will be used throughout section 4.1, where we assume the model in Fig. 2 with only one external priority, namely the isolated level i. In section 4.2 we consider the case of more than one external priority, i.e. $N>1$. Then the term level as used hitherto becomes sublevel and level is used to define an external priority level.

The reason for the MLFCFS discipline being optimal is 1) that the remaining service time t of a customer belonging to level j follows a M d.f. with rate μ_j, as long as t does not exceed t_{gj}. From [7] it can be derived that 2) FCFS is an optimal discipline to service customers belonging to the same level j, cf. the appendix. A well-known special case is that the total service time of all customers follows a M d.f. which may be looked at as a degenerated PM d.f. with only one (the first) level. For this special case undoubtedly FCFS is an optimal discipline. From 1) and 2) it follows that the MLFCFS discipline yields the minimum mean waiting time for customers with PM service-time d.f. if the levels are chosen according to Table I. Deviations between the level structure prescribed by a given PM d.f. and/or the assignment of different service quanta can result in a disastrous increase of the mean waiting time [11] .

We find it appropriate to facilitate the application of the PM d.f. family by giving here expressions which apply to the total subfamily defined by Eq.(4.4). These expressions are needed to make use of the results of section 3.1. The first three moments of the service-time d.f. truncated at t_{gk} are

$$t_k = \sum_{j=1}^{k} P_j/\mu_j, \tag{4.5}$$

$$t_k^{(2)} = \sum_{j=1}^{k} 2P_j/\mu_j^2 + \sum_{j=1}^{k-1} 2t_{gj}\{1-B(t_{gj})\}(\mu_{j+1}^{-1}-\mu_j^{-1}) - 2t_{gk}/\mu_k\{1-B(t_{gk})\}, \tag{4.6}$$

$$t_k^{(3)} = \sum_{j=1}^{k} 6P_j/\mu_j^{-3} - 3\sum_{j=1}^{k-1} t_{gj}\{1-B(t_{gj})\}(2/\mu_j^2-2/\mu_{j+1}^2+t_{gj}/\mu_j-t_{gj}/\mu_{j+1})-$$

$$-\,3t_{gk}/\mu_k\{1-B(t_{gk})\}\,(2/\mu_k+t_{gk}). \tag{4.7}$$

with P_k, as difined by Eq.(3.5):

$$P_k = \exp\{-\sum_{j=1}^{k-1} t_{gj}(\mu_j-\mu_{j+1})\}\{\exp(-t_{gk-1}\mu_k)-\exp(t_{gk}\mu_k)\} \tag{4.8}$$

and

$B(t_{gj}) = P(b_i \leq t_{gj})$ with $t = t_{gj}$ (cf. abbreviations in section 3.1).

The first and second moments of customers with service requirements $t, (t_{gk-1} < t \leq t_{gk})$ are, cf. Eq.(3.6)

$$E_k = 1/\mu_k + \{t_{gk-1}[1-B(t_{gk-1})] + t_{gk}[1-B(t_{gk})]\}/P_k, \tag{4.9}$$

$$E_k^{(2)} = \{[1-B(t_{gk}-1)]\,(t_{gk-1}^2+2t_{gk-1}/\mu_k+2/\mu_k^2) - [1-B(t_{gk})]\,(t_{gk}^2+2t_{gk}/\mu_k+2/\mu_k^2)\}/P_k. \tag{4.10}$$

Eqs.(4.8), (4.9) and (4.10) are applicable for k=1,2,...n+1.
Eqs.(4.5), (4.6) and (4.7) are applicable for k=0,1,2...n+1, where $\mu_0 \neq 0$ can be chosen arbitrarily. The last summands of Eqs.(4.5) to (4.7) equal zero if k=n+1. Just in this case the moments $t_{n+1}^{(r)}$ equal the moments $\beta^{(r)}$ of the related PM d.f. As an example we given here the first three moments of a PM-2 d.f. (with two exponential pieces, i.e. n=1):

$$\beta^{(1)} = \mu_1^{-1} + (\mu_2^{-1}-\mu_1^{-1})\, e^{\mu_1 t_g}$$

$$\beta^{(2)} = 2\{\mu_1^{-2}+(\mu_2^{-1}-\mu_1^{-1})\, e^{-\mu_1 t_g}(t_g+\mu_1^{-1}+\mu_2^{-1})\}$$

$$\beta^{(3)} = 6\{\mu_1^{-3}+e^{-\mu_1 t_g}[t_g^2/2(\mu_2^{-1}-\mu_1^{-1})+t_g(\mu_2^{-2}-\mu_1^{-2})+(\mu_2^{-3}-\mu_1^{-3})]\}\,. \tag{4.11}$$

Finally, it should be mentioned that d.f.'s with C≤1, too, can be approximated by the PM d.f., but this is only possible without observing the relation given in Eq.(4.4).

4.2 Examples

Throughout this section we assume the model in Fig. 2 having N=3 static priority levels and equally to the levels distributed total offered traffic, namely $\rho_i = \rho/N$. Moreover, we assume the same PM service-time d.f. for customers of all three priority levels.

We consider an

Example 1 with a PM-2 service-time d.f., defined by

$$B_i(t) = B(t) = \begin{cases} 1-e^{-\mu_1 t} & 0 \leq t \leq t_g \\ 1-e^{-(\mu_1-\mu_2)t_g-\mu_2 t} & t_g < t \end{cases}$$

with parameters $\mu_1 = 1.24\ s^{-1}$, $\mu_2 = 0.0485\ s^{-1}$, $t_g = 2s$. For this parameter values one computes $C = 3.46$, $t_1 = 0.739$, $t_2 = \beta = 2.466$, $t_2^{(2)} = 79.03$, $t_1^{(3)} = 1.42$, $t_2^{(3)} = 4854$, $P_1 = 0.916$, $E_1 = 0.625$, $E_2 = 22.62$, $E_1^{(2)} = 0.64$, $E_1^{(2)} = 936.72$.

Using the expressions from section 3.1 the mean waiting times for both disciplines FCFS and MLFSFC, applied to customers of priority level i ($1 \leq i \leq N = 3$), can be computed. The corresponding results are shown in Fig. 4 where the mean waiting time W_i normalized by the mean service time β_i at the ordinate is plotted over the total offered traffic ρ. The lines and broken lines show the results of the FCFS and MLFCFS disciplines, respectively. The case of saturation, $\rho > 1$, is included.

A comparison of the mean waiting times of the FCFS and MLFCFS disciplines, applied to customers of the same priority level, reveals that for a fixed ρ the difference is the larger the lower the priority number is. For a total offered traffic $\rho = 0.5$ the mean waiting time of the MLFCFS discipline is only a percentage x of that of FCFS according to Tabel II.

For a larger traffic ρ than in Table II, x becomes substantially smaller.

priority level i	1	2	3
percentage $x = \frac{W_{iMLFCFS}}{W_{iFCFS}} * 100\ \%$	12.8	52.3	61.8

Table II: Mean waiting time in the MLFCFS discipline in percent of the mean waiting time of FCFS for an offered traffic $\rho = 0.5$.

Next we consider an Example 2 with customer's service times of all three priority levels following the same PM-1 d.f., namely a M d.f. with parameter $\mu = \mu_i = 1/\beta_i$.

In this case the disciplines FCFS and MLFCFS both applied to customers of level i yield the same results which are shown in Fig. 5 for i=1,2,3. A comparison with the results of Example 1, cf. Fig. 4, reveals that for a fixed priority level i the broken lines in Fig. 4 always are below the lines in Fig. 5. This and other similar examples, we studied, show that the waiting-time optimal discipline yields a smaller mean waiting time, the greater the coefficient of varianve, $C \geq 1$, of the service-time d.f. is.

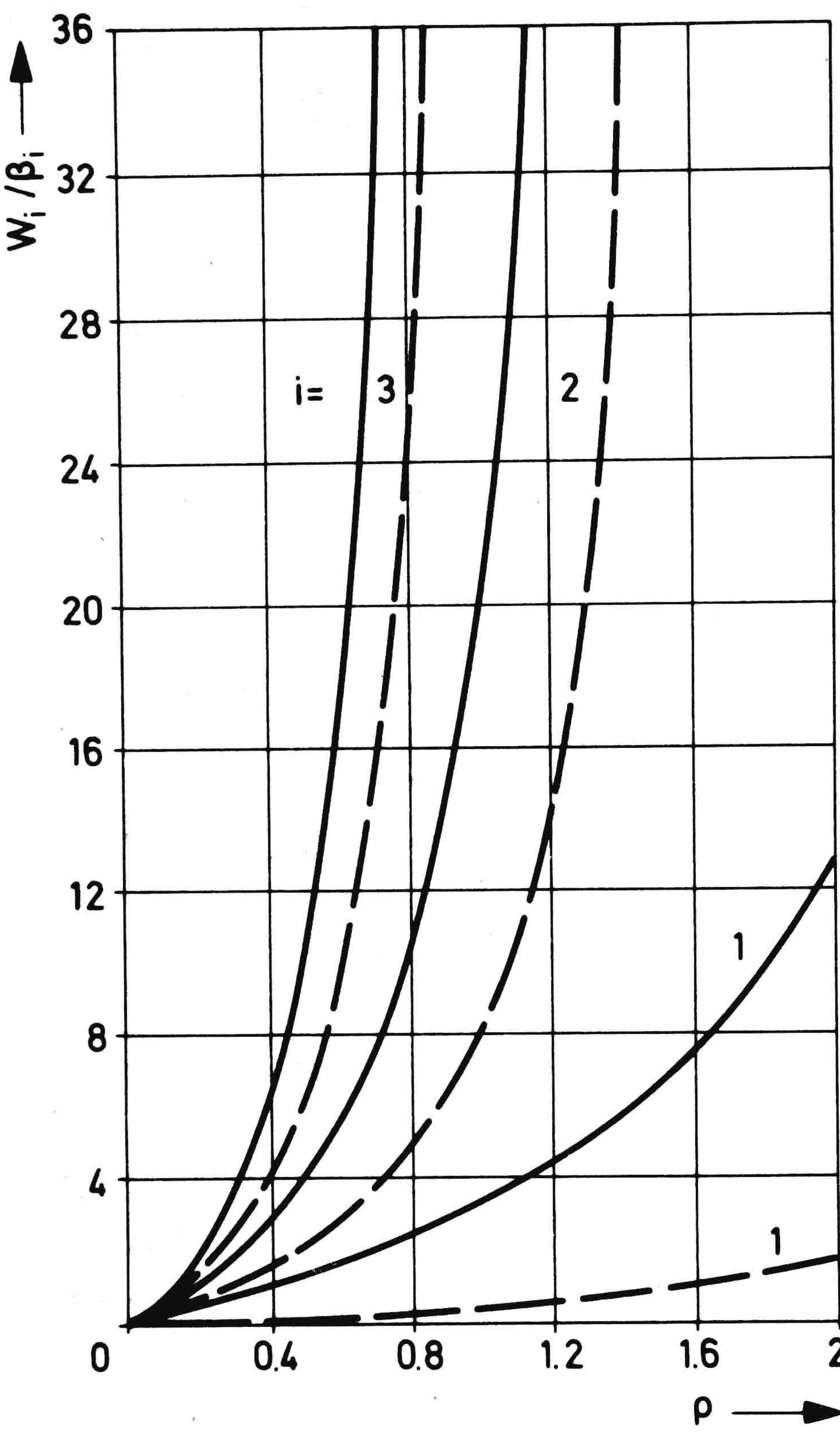

Fig. 4: Normalized mean waiting time over total offered traffic for the model in Fig. 2 and the assumptions described in Example 1, section 4.2. The curves show results of the FCFS (lines) and MLFCFS (broken lines) disciplines applied to priority level i.

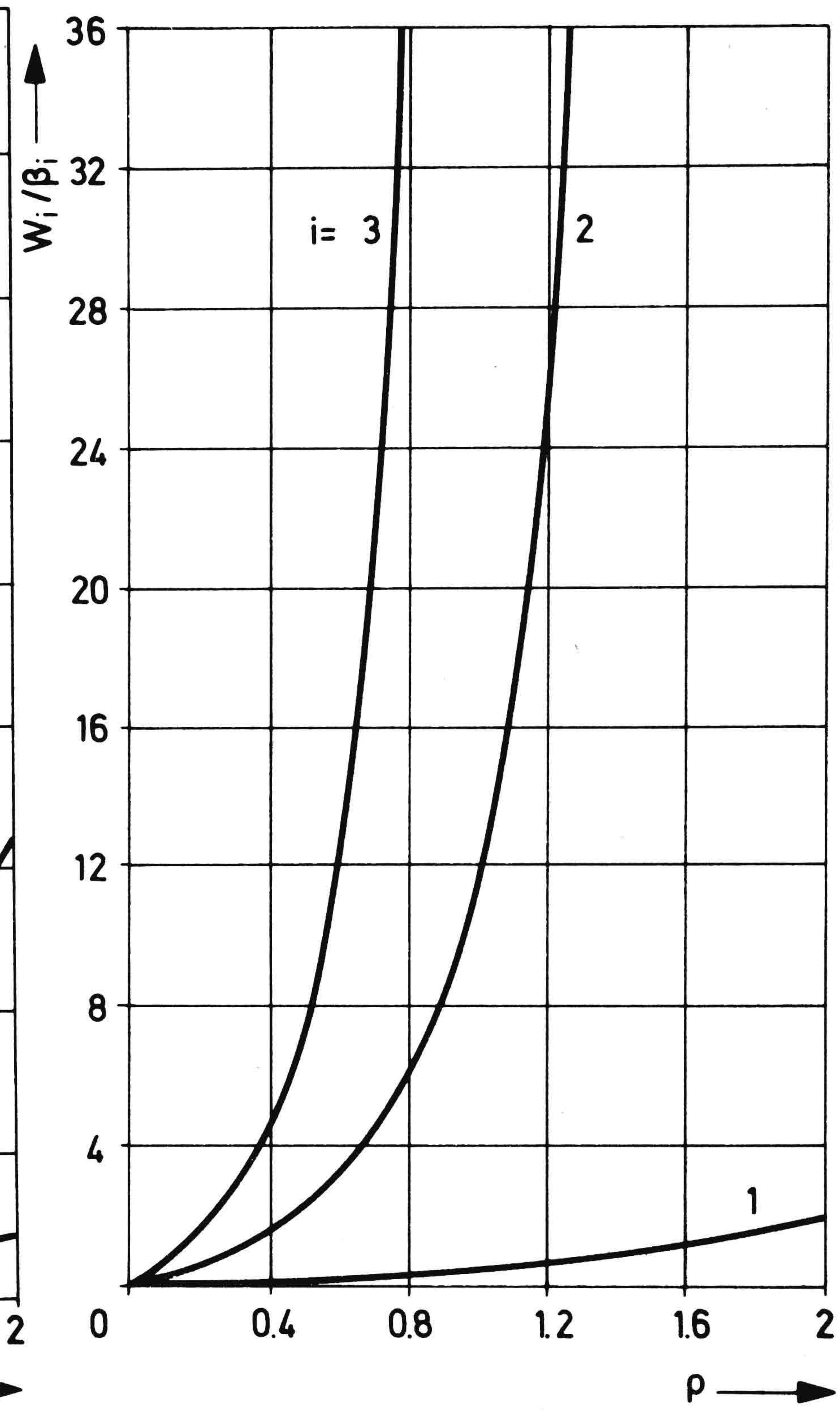

Fig. 5: Normalized mean waiting time over total offered traffic for the model in Fig. 2 and the assumptions described in Example 2, section 4.2.

An expression for the mean waiting time $W_i(t)$ conditioned on the service requirement t of a customer of priority level i has also been presented in section 3.1. Fig. 6 shows results for the assumptions of example 1 : $W_2(t)$ is shown over t for three different total offered traffics $\rho = 0.3, 0.6, 0.9$. The jump at $t=t_g=2s$ is a consequence of the two-level queueing system needed for the MLFCFS discipline with the first level having preemptive priority over the second.

For some applications the size of the jumps at $t = t_g$ may be undesirable large. Instead of this a more continous increase of $W_i(t)$ dependent on the service time t might be appropriate.

That this is possible is shown by an

Example 3: The service-time d.f.'s of customers of all three priority levels are presumed to be taken from the same PM-3 d.f. with parameters $\mu_1=3$, $\mu_2=0.4$, $\mu_3=0.04$, $t_{g1}=0.25$, $t_{g2}=5.8$. The first and second moments are $\beta=2.5$ and $\beta^{(2)}=83.4$, respectively, and nearly the same as in Example 1, cf. t_2 and $t_2^{(2)}$. In other words we have approximated the same service-time d.f. by two different PM d.f.'s namely one of second order (Example 1) and one of third order (Example 3).

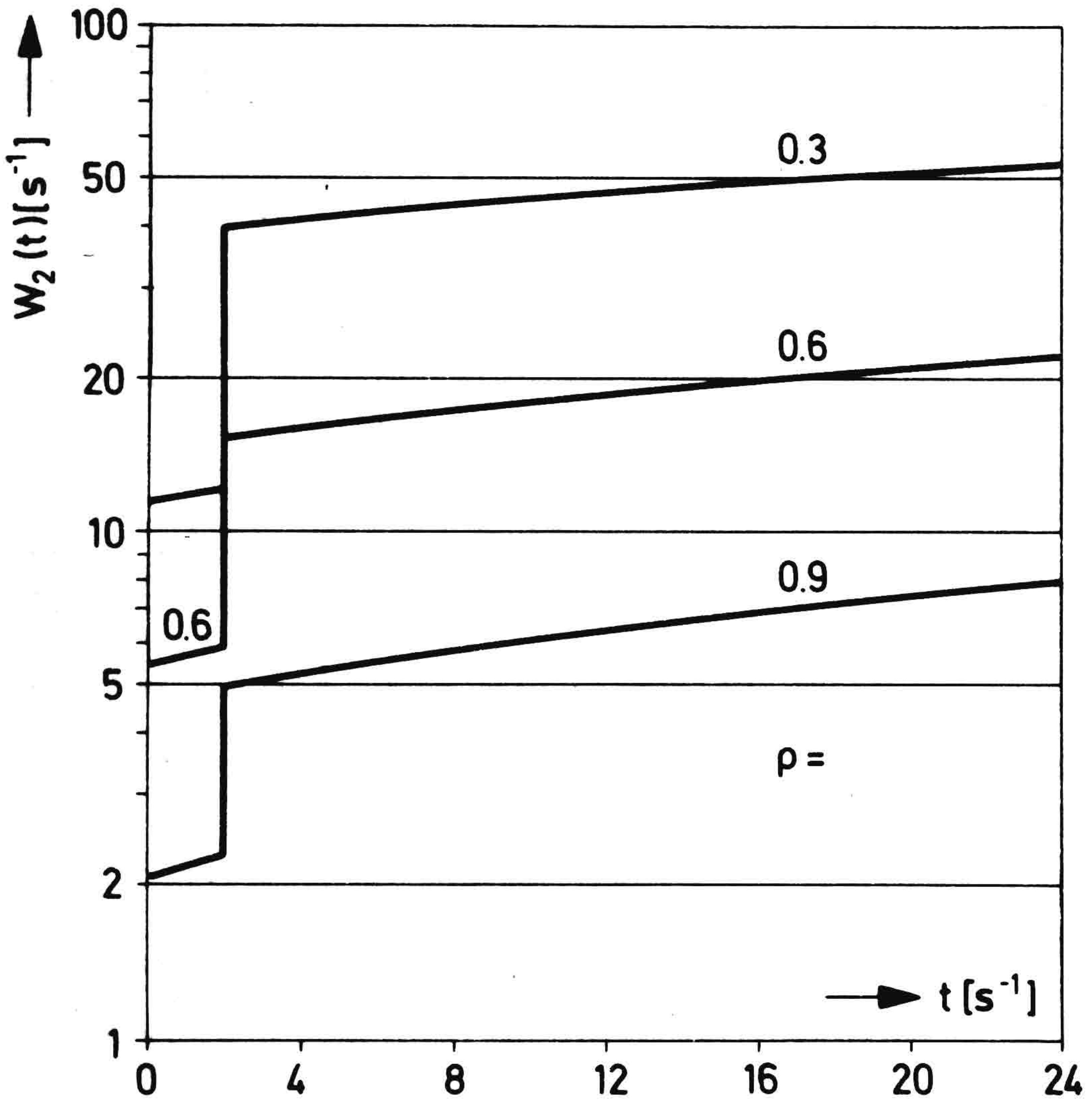

Fig. 6: Conditioned mean waiting time over service requirement t of customers of priority level 2 of the model in Fig. 2 in the MLFCFS discipline.
Assumptions are those of Example 1, section 4.2. Parameter is the total offered traffic.

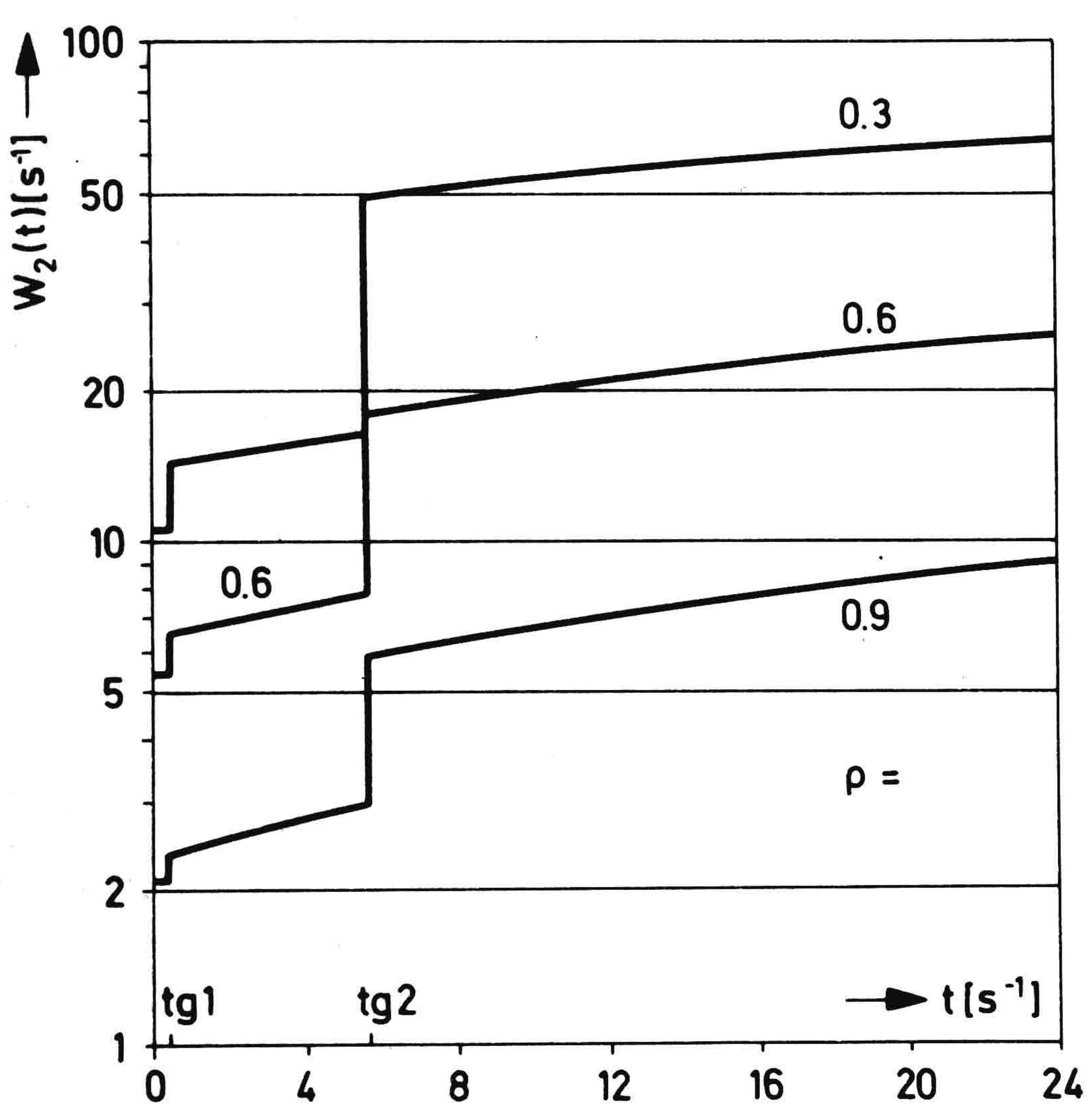

Fig. 7: Conditioned mean waiting time over service requirement t of customers of priority level 2 of the model in Fig. 2 in the MLFCFS discipline. Assumptions are those of Example 3, section 4.2. Parameter is the total offered traffic.

From Fig. 7 it can be seen that the conditioned waiting time now has two jumps at $t=t_{g1}$ and $t=t_{g2}$, each of smaller size than at t_g in Fig. 6, whose sizes summarized are greater that at t_g. Two such jumps better reach the goal to attain a more continous increase of $W_i(t)$ dependent on t (without large sized jumps), than is the case with only one jump.

Apparently, by using PM-(n+1) d.f.'s one implicitely chooses the number of jumps for $W_i(t)$ to be n and thereby directly influences the continuity of the increase of $W_i(t)$ dependent on t. Obviously for n approaching infinity a completely continous function $W_i(t)$ would arise as would be the case for the FB discipline, cf. [6] Note that the number of exponential pieces (n+1) directly defines the number of interruptable sublevels in level i, cf. Fig. 2, and thereby the amount of switching overhead. The mean waiting time W_i for both examples 1 and 3 nearly remains the same.

It should be mentioned that, by slightly modifying the MLFCFS discipline, the sizes of the jumps at $t=t_{gi}$ can be lowered distinctly without increasing the mean waiting time. Instead of consequently applying the discipline FCFS within level (i,k), cf. Fig. 2, exeptions are permitted whenever a customer, being the only one in level (i,k-1), is switched over to level (i,k) where other concurrent customers are waiting. In this situation the switched over customer is not fed back to the end of the queue in level (i,k) but remains in service. The other customers in level (i,k)are serviced after this customer. Remember that the service-time d.f. of all customers within a distinct level (i,k) is the same M d.f., presupposed a PM d.f. is used for customers of priority level i. Therefore every discipline applied to customers of such a level (i,k) yields the same mean waiting time. We are free to prescribe instead of FCFS inside a level (i,k), as is done by the MLFCFS discipline, other disciplines which take into account our demands concerning $W_i(t)$. To accomplish demands for small sized jumps, the last come first serve (LCFS) discipline inside a level (i,k), without preemptions, is appropriate. Together with the feedback discipline between levels a MLLCFS discipline for priority level i arises which minimizes switching overhead for any given number of feedback levels. While the mean waiting time W_i remains unaffected, it is an open problem to solve for $W_i(t)$ in the MLLCFS discipline. Another paper recently has appeared [13] which, too, uses the PM d.f. and discusses the related advantages compared to other modelling approaches.

Acknowledgement

I have to thank the referees for their proposal to outline the proof [7] of the optimality of MLFCFS vs. FCFS for different C as an appendix.

Appendix

This appendix contains a sketch of the voluminous proof in [7] and contains the

- criterion which easily can be applied to construct waiting-time optimal disciplines. Assumptions and outline of the proof:

  - Model A/G/1, A = homogenious arrival process.
  - Analysis of a single busy period is sufficient.
  - Cost function $g_i(t)$ with $0 \leq g_i(t) \leq C > 0$, dependent on the consumed service time t of request i (i,t) is used to characterize such a request.
  - Discrete service-time (s.t.) d.f. $B_i(b_i \leq Q) = 1$, Q large, finite.
  - Slotted time axis, slot size Δ. This assumption is dropped later on.
  - For each slot the request, to be served among those waiting, is selected according to the request's rank r(i,t).
  - A rank scale is an ordered sequence of falling rank numbers, one for each waiting request. A rank number gives a request's importance for being serviced during the next slot.
  - At any time during the running busy period the rank scale is to be constructed such that the expected cost K of all, say N, requests together to be served, is minimized, $K = E\{\sum_{j=1}^{N} g_j(t)\, w_j\}$ with w_j being the actual waiting time of request j.
  - The existence of an optimal rank scale is proofed by deriving an upper bound for K.
  - The rest of the proof demonstrates that it is sufficient to consider the differences in cost of rank scales, in which two requests are interchanged pairwise.
  - Define at the beginning of each time slot two random variables $G(i,t,n) = \begin{cases} 0 & \text{request not finished in n future slots} \\ g_i(t) & \text{request is finished in n future slots} \end{cases}$

    i.e. the cost of a macro interval of length n slots (not necessarily uninterrupted) for a request (i,t).

    T(i,t,n) service time during this macro interval (the request might be finished before the end of that interval)
  - Define expection $E\{..\}$ and a request's merit function

    $$L(i,t) = \max_n \frac{E\{G(i,t,n)\}}{E\{T(i,t,n)\}} \quad \text{(discrete case)}$$
  - The proof [7] shows that for any two requests (i,t_i), (j,t_j), not necessarily $i \neq j$

$$r(i,t_i) \left\{ \begin{matrix} > \\ < \\ = \end{matrix} \right\} r(j,t_j) \quad \text{iff} \quad L(i,t_i) \left\{ \begin{matrix} > \\ < \\ = \end{matrix} \right\} L(j,t_j)$$

i.e. the rank for a request is computable from it's merit function. A linear (not combinatoric) pairwise comparison of merit values of request is sufficient to decide, which request is to be served during the next slot.

- In the limit $\Delta \to 0$, the merit function is the quotient of
  - the expected cost of a service interval of length δ

$$E\{k_i\} = \int_t^{t+\delta} k_i(x,t)dx \quad \text{with} \quad k_i(x,t) = g_i(x)\, b_i(x)/(1-B_i(t)), \quad (x \geq t)$$

  i.e. the weighted finishing rate of request i with
  $b_i(x) = d\, B_i(x)/dx$ density of the s.t.d.f. of request i
  - the expected consumed portion of the interval of length δ

$$E\{\delta\} = \left[\int_t^{t+\delta} (x-t)b_i(x)dx + \delta \int_{t+\delta}^{\infty} b_i(x)dx\right]/(1-B_i(x)) .$$

$$\boxed{L(i,t) = \max_{\delta \geq 0} \frac{E\{k_i\}}{E\{\delta\}}} \qquad \text{(continuous case)}$$

Consequences

The permanent application of the merit function as decision criterion to decide, which request is to be served, results in a waiting-time optimal discipline. Sevcic has derived the same result for a much simpler model, in which all jobs to be executed are assumed to be in the system at time zero (i.e. no further arrivals) [14]. Apparently, the above given results are much more general.

It is very instructive to consider some special cases of general queueing models:

If the weighted finishing rate $k_i(x,t)$ is monotonic increasing in t, then the merit function can be simplified to

$$L(i,t) = \frac{\int_t^{\infty} g_i(x)b_i(x)dx}{\int_t^{\infty}(x-t)b_i(x)dx}$$

due to the fact that $\delta = \infty$ gains the maximum of L(i,t).

If $k_i(x,t)$ is monotonic decreasing in t, then L(i,t) equals $k_i(t,t)$.

Such monotonicity properties can be observed for all s.t.d.f.'s usually applied in queueing models. E.G. for $g_i(t) = g_i$ the finishing rates are monotonic for M, H_k, E_k, n-point and special cases of PM and GE (general Erlang) d.f.'s.

The overhead consumed to make decisions based on the merit functions of waiting requests depends on the monotonicity of $k_i(x,t)$:
If all request have <u>monotonic increasing or nondecreasing</u> finishing rates (e.g. M,

E_k s.t. d.f.'s) then a decision based on L(i,t) is needed only just after the arrival of a new request and just after having finished service of a request. At this situations the next request is to be selected and will be served until min(next arrival, next departure). The overhead needed for decision in this special case can be very small.

If some (or all) request have monotonic decreasing finishing rates, due to the assumption of a H_k s.t.d.f. then periods may arise, during which the value of the merit function of a request in service after a very small time period $\delta \to 0$ is inferior to the merit-function value of one or more other waiting requests. The resulting discipline is comparable to processor sharing among those requests and the decision and switching overhead needed exceeds any limit. Just in such situations it is adviseable to work with PM instead of H_k s.t.d.f.'s, which results in piecewise constant finishing rates (no need for preemptions during exponential pieces without arrivals) and as a consequence results in time periods $\delta \gg 0$ and thereby a very small overhead.

In [15] examples are given which demonstrate, how to serve requests with different s.t.d.f.'s. The resulting waiting-time optimal disciplines differ from those usually applied in time-sharing systems. Their analysis remains for further research. Results of such analysis may serve to derive a basis useful to evaluate service disciplines, currently in use.

REFERENCES:

[1] J.M. Babad: A generalized multi-entrance time-sharing priority queue. Journ. ACM, Vol. 22, No. 2, April 1975, pp. 231-247.

[2] M. Fujiki, E. Gambe: Conditional waiting time in a multilevel feedback priority queueing model. Electron. and Comm. in Japan, Vol. 52-C, No. 1, 1969, pp. 129-136.

[3] Z. Rosberg, I. Adiri: Multi-level queues with external priorities. Journ. ACM, Vol. 23, No. 4, Oct. 1976, pp. 680-690.

[4] H. White, L.S. Christie: Queueing with preemptive priorities or with breakdown. Oper.Res. (USA), Vol. 6, No. 1, 1958, pp. 75-95.

[5] R.G. Miller: Priority queues. Annals Math. Stat. 31, 1960, pp. 86-103.

[6] L.Kleinrock, R.R. Muntz: Multi-level processor-sharing queueing models for time-shared systems. Proc. 6th Int. Teletraff. Congr., 1970, Congressbook, pp. 341/1-8.

[7] G. Olivier: Cost-minimum priorities in queueing systems of type M/G/1. Elektron. Rechenanlagen 14, (1972), No. 6, pp. 262-271 (In German).

[8] M. Sakata, S. Noguchi, J. Oizumi: Optimal scheduling in single-processor systems. Proc. 3rd Hawaii Int. Conf. Syst. Science. 1970, Univ. of Hawaii, pp. 708-711.

[9] G. Marte: Optimal time scheduling for time-shared computer systems with piecewise exponential computing-time distribution. Proc. Intern. Comp. Sympos. 1970, Published by the German Chapter of the ACM, 1973, pp. 750-774.

[10] B. Walke: A static-priority queue with time slicing. Proc. 9th Internat. Teletraff. Congr., 1979, Congressbook, session 51,and Elektron.Rechenanlagen 22 (1980), H. 3, pp. 107-114, (English and German version,respectively).

[11] B. Walke, H.J. Küspert: Multi-access systems: mean response time for optimal policies. Elektron. Rechenanlagen 13,(1971),No. 5, pp. 193-199 (In German).

[12] V.A. Popov, M.L. Litvinov: Solution of queueing problems using piecewise distribution functions. Cybernetics (USA), 9, (1973), May-Juni No. 3, pp. 451-455.

[13] U. Herzog: Performance characteristics for hierarchically organized multiprocessor computer systems with generally distributed processing time. Proc. 9th Int. Teletraffic Congr., 1979, Torremolinos-Spain, Congressbook, session 44.

[14] K.C. Sevcik: A proof of the optimality of "smallest rank" scheduling. Journ. ACM, 21, 1974, pp. 66-75.

[15] B. Walke: Zur Anwendung optimaler Bedienstrategien in Rechnern (How to construct optimal service disciplines). Informatik Fachberichte 19, Springer Verlag, Berlin 1979, pp. 352-367.