Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

J.L. van den Berg, O.J. Boxma, W.P. Groenendijk

Sojourn times in the M/G/1 queue
with deterministic feedback

Department of Operations Research and System Theory Report OS-R8708 March

Sojourn Times in the M/G/1 Queue with Deterministic Feedback

J.L. van den Berg, O.J. Boxma, W.P. Groenendijk
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

In this paper we consider an M/G/1 queueing model, in which each customer is fed back a fixed number of times. For the case of negative exponentially distributed service times at each visit, we determine the joint distribution of the sojourn times of the consecutive visits. As a by-result, we obtain the total sojourn time distribution; it can be related to the sojourn time distribution in the M/D/1 queue with processor sharing. For the case of generally distributed service times at each visit, a set of linear equations is derived, from which the mean sojourn times per visit can be calculated.

1980 Mathematics Subject Classification: 60K25, 68M20.
Key Words & Phrases: M/G/1 queue, feedback, sojourn times.
Note: This report will be submitted for publication elsewhere.

1. Introduction

In this paper we study an M/G/1 queueing system in which each customer is fed back a fixed number of times. Feedback systems occur in many practical situations; for instance, in computer systems tasks that are scheduled for resources may have to come back several times for additional service. In the literature much attention has been paid to feedback queues. However, most studies concerned so-called Bernoulli feedback: when a customer completes his service he departs from the system with probability 1-p and is fed back with probability p; see Takács [10], Disney and König [2], Disney et al. [3] and Doshi and Kaufman [4]. Fontana and Diaz Berzosa [5,6] extend some results obtained for the M/G/1 model with Bernoulli feedback to a more general feedback model with priorities.

Simon [9] studies a somewhat more general model than the one presented in this paper. He allows different types of customers and priority levels, that may change after a service completion. The main result of his paper is the derivation of a set of linear equations for the mean sojourn time of each visit.

In the present model, to be described in Section 2, the priority mechanism is omitted. In Section 3 we derive, for the case of negative exponentially distributed service times, the joint distribution of the successive sojourn times of a customer. As a by-result we obtain an explicit expression for the distribution of the total sojourn time of a customer. This expression can be used to obtain the sojourn time distribution in the M/D/1 queueing system with processor sharing, a result previously found by Ott [8]. In Section 4 we show that, for the case of generally distributed service times at each visit, the set of linear equations for the mean sojourn times per visit can be explicitly solved, yielding some interesting results. Finally these results are extended to a model with a more general feedback mechanism.

Report OS-R8708
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

2. Model description

We consider a single server queueing system with infinite waiting room. Customers arrive at the system according to a Poisson process with intensity $\lambda>0$. Each customer requires N services: a customer who enters the queue will return to the queue (feedback) after service $N-1$ times before leaving. Fed back customers return instantaneously, joining the end of the queue. The service discipline is First Come First Served (FCFS). The N service times of a customer are mutually independent random variables having distribution functions $B_i(\cdot)$, with mean β_i and second moment $\beta_i^{(2)}$, $i=1,...,N$. These service times are also independent of the service times of other customers.

Obviously the stability condition is that $\lambda\sum_{i=1}^{N}\beta_i < 1$.

We define

- type-i customer: customer who is visiting the queue for the i-th time, $i=1,...,N$.

- X_i: number of type-i customers in the system at an arbitrary epoch, $i=1,...,N$.

- S_i: time between i-th arrival and i-th service completion of a customer, $i=1,...,N$.

- $\mathrm{S} := \sum_{i=1}^{N}\mathrm{S}_i$ (total sojourn time).

3. The negative exponential case

For the case that the service times are identically, negative exponentially, distributed,

$$B_i(t) = 1-e^{-t/\beta}, \quad i=1,...,N,$$

we derive an expression for the Laplace-Stieltjes transform $E\{e^{-(\omega_1\mathrm{S}_1+...+\omega_N\mathrm{S}_N)}\}$ of the joint distribution of the successive sojourn times S_i, $i=1,...,N$, of a customer.

First note that the system described above can be considered as a queueing network consisting of one queue with N types of customers. Type-N customers leave the system with probability 1 after service. Type-i customers return to the queue with probability 1 after service, and change into type-$(i+1)$ customers, $i=1,...,N-1$.

Because the service times are assumed to be exponentially distributed, the results obtained by Baskett et al. [1] can be applied to find the joint distribution of the number of type-i customers in the system. It is found that, for $x_1,...,x_N = 0,1,2,...,$

$$P(x_1,\ldots,x_N) := Pr\{\mathrm{X}_1=x_1,\ldots,\mathrm{X}_N=x_N\} = (1-N\lambda\beta)(\lambda\beta)^{(x_1+...+x_N)}(x_1+...+x_N)!\prod_{i=1}^{N}\frac{1}{x_i!}. \quad (3.1)$$

We follow a customer, say K, from the moment he arrives as a type-1 customer until he leaves the system as a type-N customer. For the successive sojourn times $\mathrm{S}_1,\ldots,\mathrm{S}_N$ of K it holds that

$$E\{e^{-(\omega_1\mathrm{S}_1+...+\omega_N\mathrm{S}_N)}\} = \sum_{x_1=0}^{\infty}\cdots\sum_{x_N=0}^{\infty}P(x_1,\ldots,x_N)E\{e^{-(\omega_1\mathrm{S}_1+...+\omega_N\mathrm{S}_N)}\,|\mathrm{X}_1=x_1,\ldots,\mathrm{X}_N=x_N\}. \quad (3.2)$$

Note that we have used the PASTA property ([11]).

Let $\omega:=(\omega_1,\ldots,\omega_N)$, and

$$A_1(\omega) := \frac{1}{1+\beta\omega_N},$$

$$A_i(\omega) := \frac{1}{1+\beta(\omega_{N-i+1}+\lambda-\lambda\prod_{j=1}^{i-1}A_j(\omega))}, \qquad i=2,...,N.$$

We now prove the following theorem.

Theorem

In the M/M/1 queue with deterministic feedback,

$$E\{e^{-(\omega_1 \mathbf{S}_1+...+\omega_N \mathbf{S}_N)}\} = \frac{(1-N\lambda\beta)\prod_{i=1}^{N}A_i(\omega)}{1-\lambda\beta(\prod_{i=1}^{N}A_i(\omega)+\prod_{i=2}^{N}A_i(\omega)+...+A_N(\omega))}, \qquad \mathrm{Re}\ \omega_i \geqslant 0,\ i=1,...,N. \tag{3.3}$$

Proof: Conditioning on the number of external arrivals, n_i, during the i-th sojourn time, $i=1,...,N-1$, it is easily seen that

$$E\{e^{-(\omega_1 \mathbf{S}_1+...+\omega_N \mathbf{S}_N)}\ |\mathbf{X}_1=x_1,\ldots,\mathbf{X}_N=x_N\} =$$

$$\int_{t_1=0}^{\infty} e^{-\omega_1 t_1} \int_{t_2=0}^{\infty} e^{-\omega_2 t_2} \cdots \int_{t_{N-1}=0}^{\infty} e^{-\omega_{N-1} t_{N-1}} \int_{t_N=0}^{\infty} e^{-\omega_N t_N} \sum_{n_1=0}^{\infty} e^{-\lambda t_1}\frac{(\lambda t_1)^{n_1}}{n_1!} \sum_{n_2=0}^{\infty} e^{-\lambda t_2}\frac{(\lambda t_2)^{n_2}}{n_2!} \cdots \sum_{n_{N-1}=0}^{\infty} e^{-\lambda t_{N-1}}\frac{(\lambda t_{N-1})^{n_{N-1}}}{n_{N-1}!}$$

$$dB(t_N)^{(x_1+n_1+...+n_{N-1}+1)^*}\ dB(t_{N-1})^{(x_1+x_2+n_1+...+n_{N-2}+1)^*} \cdots dB(t_2)^{(x_1+...+x_{N-1}+n_1+1)^*}\ dB(t_1)^{(x_1+...+x_N+1)^*} =$$

$$\int_{t_1=0}^{\infty} e^{-\omega_1 t_1} \int_{t_2=0}^{\infty} e^{-\omega_2 t_2} \cdots \int_{t_{N-1}=0}^{\infty} e^{-\omega_{N-1} t_{N-1}}$$

$$\sum_{n_1=0}^{\infty} e^{-\lambda t_1}\frac{(\lambda t_1)^{n_1}}{n_1!} \sum_{n_2=0}^{\infty} e^{-\lambda t_2}\frac{(\lambda t_2)^{n_2}}{n_2!} \cdots \sum_{n_{N-1}=0}^{\infty} e^{-\lambda t_{N-1}}\frac{(\lambda t_{N-1})^{n_{N-1}}}{n_{N-1}!} \left[\frac{1}{1+\beta\omega_N}\right]^{x_1+n_1+...+n_{N-1}+1}$$

$$dB(t_{N-1})^{(x_1+x_2+n_1+...+n_{N-2}+1)^*} \cdots dB(t_2)^{(x_1+...+x_{N-1}+n_1+1)^*}\ dB(t_1)^{(x_1+...+x_N+1)^*} =$$

$$\left[\frac{1}{1+\beta\omega_N}\right]^{x_1+1} \int_{t_1=0}^{\infty} e^{-\omega_1 t_1} \int_{t_2=0}^{\infty} e^{-\omega_2 t_2} \cdots \sum_{n_1=0}^{\infty} e^{-\lambda t_1}\frac{(\lambda t_1)^{n_1}}{n_1!} \sum_{n_2=0}^{\infty} e^{-\lambda t_2}\frac{(\lambda t_2)^{n_2}}{n_2!} \cdots$$

$$\left[\frac{1}{1+\beta\omega_N}\right]^{n_1+...+n_{N-2}} \int_{t_{N-1}=0}^{\infty} e^{-\omega_{N-1} t_{N-1}} e^{-\lambda t_{N-1}} \sum_{n_{N-1}=0}^{\infty} \left[\frac{\lambda t_{N-1}}{1+\beta\omega_N}\right]^{n_{N-1}} \frac{1}{n_{N-1}!}$$

$$dB(t_{N-1})^{(x_1+x_2+n_1+...+n_{N-2}+1)^*} \cdots dB(t_2)^{(x_1+...+x_{N-1}+n_1+1)^*}\ dB(t_1)^{(x_1+...+x_N+1)^*} =$$

$$\left[\frac{1}{1+\beta\omega_N}\right]^{x_1+1}\int_{t_1=0}^{\infty}e^{-\omega_1 t_1}\int_{t_2=0}^{\infty}e^{-\omega_2 t_2}\cdots\sum_{n_1=0}^{\infty}e^{-\lambda t_1}\frac{(\lambda t_1)^{n_1}}{n_1!}\sum_{n_2=0}^{\infty}e^{-\lambda t_2}\frac{(\lambda t_2)^{n_2}}{n_2!}\cdots$$

$$\left[\frac{1}{1+\beta\omega_N}\right]^{n_1+\ldots+n_{N-2}}\left[\frac{1}{1+\beta(\omega_{N-1}+\lambda-\frac{\lambda}{1+\beta\omega_N})}\right]^{x_1+x_2+n_1+\ldots+n_{N-2}+1}$$

$$dB(t_{N-2})^{(x_1+x_2+x_3+n_1+\ldots+n_{N-3}+1)^*}\cdots dB(t_2)^{(x_1+\ldots+x_{N-1}+n_1+1)^*}\,dB(t_1)^{(x_1+\ldots+x_N+1)^*}.$$

Proceeding in this way we find

$$E\{e^{-(\omega_1 S_1+\ldots+\omega_N S_N)}\,|X_1=x_1,\ldots,X_N=x_N\}=\prod_{i=1}^{N}A_i^{1+\sum_{j=1}^{i}x_j}(\omega). \tag{3.4}$$

Substituting (3.1) and (3.4) in (3.2), and observing that

$$(1-N\lambda\beta)\prod_{i=1}^{N}A_i(\omega)\sum_{x_1=0}^{\infty}\cdots\sum_{x_N=0}^{\infty}(\lambda\beta)^{x_1+\ldots+x_N}\frac{(x_1+\ldots+x_N)!}{x_1!\cdots x_N!}\prod_{i=1}^{N}A_i^{x_1+\ldots+x_i}(\omega)=$$

$$(1-N\lambda\beta)\prod_{i=0}^{N}A_i(\omega)\sum_{m=0}^{\infty}\sum_{x_1}\cdots\sum_{\substack{x_N\\x_1+\ldots+x_N=m}}\frac{m!}{x_1!\cdots x_N!}\left[\lambda\beta\prod_{i=1}^{N}A_i(\omega)\right]^{x_1}\left[\lambda\beta\prod_{i=2}^{N}A_i(\omega)\right]^{x_2}\cdots\left[\lambda\beta A_N(\omega)\right]^{x_N}=$$

$$(1-N\lambda\beta)\prod_{i=1}^{N}A_i(\omega)\sum_{m=0}^{\infty}\left[\lambda\beta\prod_{i=1}^{N}A_i(\omega)+\lambda\beta\prod_{i=2}^{N}A_i(\omega)+\ldots+\lambda\beta A_N(\omega)\right]^m,$$

we obtain the required result.

Remarks

i) Consider a type-i customer, present when the tagged customer K arrives. During his $(i+y)$-th service, $y=0,1,\ldots,N-i$, he influences K's sojourn times $S_{y+1},\ldots,S_N$ in two ways. His service time contributes to S_{y+1}, and customers arriving during this service time influence $S_{y+2},\ldots,S_N$. These contributions are collected in the term $A_{N-y}(\omega)$; the total contribution of all x_i type-i customers to the expression in the right-hand side of (3.4) is $\{\prod_{y=0}^{N-i}A_{N-y}(\omega)\}^{x_i}=\{\prod_{j=i}^{N}A_j(\omega)\}^{x_i}$.

ii) The (marginal) distribution of the i-th sojourn time, S_i, $i=1,\ldots,N$, can be obtained from (3.3) by taking $\omega_j=0$, $j=1,\ldots,N$, $j\neq i$. It is found that

$$E\{e^{-\omega_i S_i}\}=\frac{1-N\lambda\beta}{1-N\lambda\beta+\beta\omega_i}. \tag{3.5}$$

Hence, the sojourn times S_i, $i=1,\ldots,N$, are identically, negative exponentially, distributed with mean $\beta/(1-N\lambda\beta)$. This coincides with the sojourn time distribution in an ordinary M/M/1 queue with mean service time β and arrival rate $N\lambda$.

iii) In order to investigate the dependence between the i-th and j-th sojourn time we have computed the Laplace-Stieltjes transform of the joint distribution of $\mathbf{S}_i$ and $\mathbf{S}_j$, $1 \leqslant i < j \leqslant N$. It is found from (3.3) that

$$E\{e^{-(\omega_i \mathbf{S}_i + \omega_j \mathbf{S}_j)}\} = \frac{1 - N\lambda\beta}{1 - N\lambda\beta + \beta\omega_i + \beta\omega_j + \beta^2(1+\lambda\beta)^{j-i-1}\omega_i\omega_j}. \tag{3.6}$$

From (3.6) the correlation coefficient, $corr(\mathbf{S}_i,\mathbf{S}_j)$, can easily be obtained:

$$corr(\mathbf{S}_i,\mathbf{S}_j) = 1 - (1 - N\lambda\beta)(1+\lambda\beta)^{j-i-1}, \quad 1 \leqslant i < j \leqslant N. \tag{3.7}$$

It follows that $corr(\mathbf{S}_i,\mathbf{S}_j)$ as a function of i and j only depends on $j-i$, and that it decreases if $j-i$ grows. It is also seen that $corr(\mathbf{S}_i,\mathbf{S}_j)$ is an increasing function of the total offered load $N\lambda\beta$ for fixed i and j. These intuitively appealing properties are illustrated in Fig. 1.

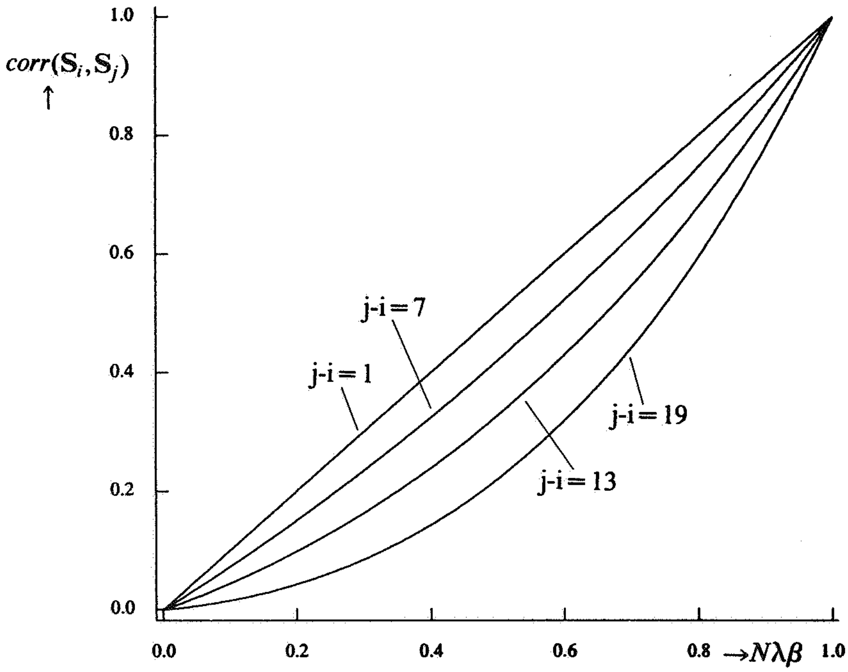

Fig. 1 $corr(\mathbf{S}_i,\mathbf{S}_j)$ as a function of offered load $N\lambda\beta$, with $N=20$

iv) From (3.3), substituting $\omega_i = \omega_0$, $i=1,...,N$, it is found that the Laplace-Stieltjes transform of the total sojourn time distribution is given by

$$E\{e^{-\omega_0 \mathbf{S}}\} = \frac{(1-N\lambda\beta)(\lambda+\omega_0)^2}{\omega_0^2(1+\lambda\beta+\beta\omega_0)^N + \lambda(\lambda+\omega_0)(1-N\lambda\beta) + \lambda\omega_0}, \quad \text{Re } \omega_0 \geqslant 0. \tag{3.8}$$

Formula (3.8) can easily be inverted using Jagerman's inversion technique [7].
From (3.8) we obtain the variance of the total sojourn time:

$$Var(\mathbf{S}) = \left[\frac{\beta}{1-N\lambda\beta}\right]^2 \left[\frac{2}{(\lambda\beta)^2} - N^2 - 2(1-N\lambda\beta)\frac{(1+\lambda\beta)^N}{(\lambda\beta)^2}\right]. \tag{3.9}$$

Formula (3.9) could also have been derived from the above results for $Var(\mathbf{S}_i)$ and $corr(\mathbf{S}_i,\mathbf{S}_j)$.

v) If we let $N\uparrow\infty$ and $\beta\downarrow 0$ in such a way that $\tilde{\beta}:=N\beta$ remains constant, then the distribution of the total service time received by each customer approaches the deterministic distribution fixed at $\tilde{\beta}$:

$$\lim_{\substack{N\uparrow\infty \\ \beta\downarrow 0}} \left[\frac{1}{1+\beta\omega_0}\right]^N = \lim_{N\uparrow\infty} \left[\frac{1}{1+\frac{\tilde{\beta}}{N}\omega_0}\right]^N = e^{-\tilde{\beta}\omega_0}, \quad \mathrm{Re}\ \omega_0 \geqslant 0.$$

This limiting procedure apparently reduces the deterministic feedback model to the M/D/1 queueing model with processor sharing. Indeed, in the limit the distribution of the total sojourn time equals the sojourn time distribution in the M/D/1 system with processor sharing:

$$\lim_{\substack{N\uparrow\infty \\ \beta\downarrow 0}} E\{e^{-\omega_0 S}\} = \frac{(\lambda+\omega_0)^2(1-\lambda\tilde{\beta})e^{-\tilde{\beta}(\lambda+\omega_0)}}{\omega_0^2 + \lambda(\lambda+2\omega_0-\lambda\tilde{\beta}(\lambda+\omega_0))e^{-\tilde{\beta}(\lambda+\omega_0)}}, \quad \mathrm{Re}\ \omega_0 \geqslant 0, \tag{3.10}$$

a result previously obtained by Ott [8].

4. The general case

In this section we first consider the case where the service time distribution of a customer depends on the number of times he has been fed back. As in Simon [9], we derive a set of linear equations from which the mean sojourn time per visit can be calculated. Next, we show that for the special case that all service time distributions are equal (but not necessarily negative exponential), this set of linear equations can be easily solved explicitly. It appears that from the second visit on, all mean sojourn times are equal. Finally we introduce a generalized feedback model, which includes both the deterministic feedback model and the Bernoulli feedback model. Again, a set of linear equations for the mean sojourn times is derived. This set is solved for the case of Bernoulli feedback.

4.1 *Derivation of a set of linear equations*

We consider the case that the service time distribution of a customer who has been fed back $i-1$ times is given by $B_i(\cdot)$, $i=1,2,...,N$. Denote by $\rho_i:=\lambda\beta_i$ the offered traffic due to type-i customers. We start by obtaining a relation for $E\mathbf{S}_1$. Note that a newly arriving customer is a Poisson arrival and hence PASTA ([11]) applies. Consider the mean amount of work that has to be handled before this newly arriving customer receives his first service. This quantity consists of two components:

1. the mean amount of waiting work found upon his arrival that is handled before his first service, given by: $\sum_{i=1}^{N}\beta_i E\mathbf{X}_i^w$;
2. the mean amount of work currently in service: $\sum_{i=1}^{N}\rho_i \frac{\beta_i^{(2)}}{2\beta_i}$;

where $\mathbf{X}_i^w$ denotes the number of *waiting* type-i customers. It may now be seen that,

$$E\mathbf{S}_1 = \sum_{i=1}^{N}\beta_i E\mathbf{X}_i^w + \sum_{i=1}^{N}\rho_i\frac{\beta_i^{(2)}}{2\beta_i} + \beta_1. \tag{4.1.1}$$

With $E\mathbf{X}_i^w = E\mathbf{X}_i - \lambda\beta_i$ we obtain:

$$E\mathbf{S}_1 = \sum_{i=1}^{N}\beta_i E\mathbf{X}_i - \lambda\sum_{i=1}^{N}\beta_i^2 + \frac{\lambda}{2}\sum_{i=1}^{N}\beta_i^{(2)} + \beta_1. \tag{4.1.2}$$

$E\mathbf{S}_{i+1}$ is composed of mean service times of "old" customers and of customers who have arrived during the first i sojourn times:

$$ES_{i+1} = \sum_{j=1}^{N-i}{}' EX_j\beta_j + \sum_{j=1}^{i}\beta_{i+1-j}\lambda ES_j + \beta_{i+1}, \quad i=1,...,N-1. \tag{4.1.3}$$

Now apply Little's formula to (4.1.2) and (4.1.3):

$$ES_1 = \sum_{i=1}^{N}\rho_i ES_i + \frac{\lambda}{2}\sum_{i=1}^{N}(\beta_i^{(2)} - 2\beta_i^2) + \beta_1, \tag{4.1.4}$$

$$ES_{i+1} = \sum_{j=1}^{N-i}{}' \rho_j ES_j + \sum_{j=1}^{i}\rho_{i+1-j} ES_j + \beta_{i+1}, \quad i=1,...,N-1. \tag{4.1.5}$$

Formulas (4.1.4) and (4.1.5) represent a set of N linear equations in N unknowns. In the next subsection this set of equations will be solved in a special case.

4.2 *Special case:* $B_i(\cdot) \equiv B(\cdot),\ i=1,...,N$

In this subsection we assume that all service time distributions are the same. In fact, for our purposes it suffices to assume that $\beta_i \equiv \beta$ and $\beta_i^{(2)} \equiv \beta^{(2)}$ for all i. The equations (4.1.4) and (4.1.5) now become:

$$ES_1 = \lambda\beta ES + \frac{\lambda}{2}N(\beta^{(2)} - 2\beta^2) + \beta, \tag{4.2.1}$$

$$ES_{i+1} = \lambda\beta\sum_{j=1}^{N-i}{}' ES_j + \lambda\beta\sum_{j=1}^{i} ES_j + \beta, \quad i=1,...,N-1. \tag{4.2.2}$$

Due to the symmetry in (4.2.2) we have that

$$ES_{i+1} = ES_{N-i+1}, \quad i=1,...,N-1.$$

Subtracting ES_i from ES_{i+1}, we obtain

$$ES_{i+1} - ES_i = -\lambda\beta ES_{N-i+1} + \lambda\beta ES_i = -\lambda\beta(ES_{i+1} - ES_i), \quad i=2,...,N-1.$$

Hence, $ES_i = ES_{i+1}$ and we obtain

$$ES_2 = ES_3 = \cdots = ES_N. \tag{4.2.3}$$

Now from (4.2.1) and (4.2.3):

$$ES_1 = \lambda\beta ES_1 + (N-1)\lambda\beta ES_2 + \frac{\lambda}{2}N(\beta^{(2)} - 2\beta^2) + \beta. \tag{4.2.4}$$

And from (4.2.2) and (4.2.3):

$$ES_2 = 2\lambda\beta ES_1 + (N-2)\lambda\beta ES_2 + \beta. \tag{4.2.5}$$

Solving equations (4.2.4) and (4.2.5) yields:

$$ES_1 = \frac{\beta}{1-N\lambda\beta} + \frac{(1-(N-2)\lambda\beta)\frac{\lambda}{2}N(\beta^{(2)} - 2\beta^2)}{(1+\lambda\beta)(1-N\lambda\beta)}, \tag{4.2.6}$$

$$ES_2 = ES_3 = \cdots = ES_N = \frac{\beta}{1-N\lambda\beta} + \frac{\lambda^2\beta N(\beta^{(2)} - 2\beta^2)}{(1+\lambda\beta)(1-N\lambda\beta)}. \tag{4.2.7}$$

Hence

$$ES = \frac{N\beta}{1-N\lambda\beta} + \frac{1+N\lambda\beta}{1-N\lambda\beta}\frac{\lambda}{2}\frac{N(\beta^{(2)} - 2\beta^2)}{1+\lambda\beta}. \tag{4.2.8}$$

For $N = 1$ this gives:

8

$$ES = \frac{\lambda\beta^{(2)}}{2(1-\lambda\beta)} + \beta, \tag{4.2.9}$$

as could be expected.
Finally observe from (4.2.6) and (4.2.7) that $ES_1 = ES_2$ if the service times are negative exponentially distributed.

4.3 A generalized M/G/1 Bernoulli feedback model

An obvious generalization of our model and of the Bernoulli feedback model is the model in which a customer who just had his j-th service departs from the system with probability $1-p(j)$ and is fed back with probability $p(j)$, $j=1,2,\ldots$. By definition, $p(0)=1$. Let

$$q_i := \prod_{j=0}^{i-1} p(j), \quad i=1,2,\ldots .$$

The definitions of type-i customers and their characteristic quantities, as given in Section 2, are extended in an obvious way. Note that the stability condition for this system is that $\lambda\sum_{i=1}^{\infty} q_i\beta_i < 1$. We assume in the following that $B_i(\cdot) \equiv B(\cdot)$. In exactly the same way as in Subsection 4.2 we derive:

$$ES_1 = \beta\sum_{i=1}^{\infty} EX_i + \frac{\lambda}{2}(\beta^{(2)}-2\beta^2)\sum_{i=0}^{\infty} q_{i+1} + \beta, \tag{4.3.1}$$

$$ES_{k+1} = \lambda\beta\sum_{i=0}^{k-1} q_{i+1}ES_{k-i} + \beta\sum_{i=1}^{\infty}\frac{q_{k+i}}{q_i}EX_i + \beta, \quad k=1,2,\ldots . \tag{4.3.2}$$

Rewriting (4.3.1) and (4.3.2) and using Little's formula yields:

$$ES_1 = \lambda\beta\sum_{i=1}^{\infty} q_i ES_i + \frac{\lambda}{2}(\beta^{(2)}-2\beta^2)\sum_{i=1}^{\infty} q_i + \beta, \tag{4.3.3}$$

$$ES_{k+1} = \lambda\beta\sum_{i=0}^{k-1} q_{i+1}ES_{k-i} + \lambda\beta\sum_{i=1}^{\infty} q_{k+i}ES_i + \beta, \quad k=1,2,\ldots . \tag{4.3.4}$$

For some special cases (in particular cases with $p(N)=0$ for some finite N) this set of equations can be easily solved. Below we present the solution for the case of Bernoulli feedback.
Introducing

$$M := \frac{\lambda}{2}(\beta^{(2)}-2\beta^2)\frac{1}{1-p},$$

$$M_i := \frac{1}{M}\left[ES_i - \frac{\beta}{1-\lambda\beta/(1-p)}\right], \quad i=1,2,\ldots,$$

we can rewrite (4.3.3) and (4.3.4) into

$$M_1 = \lambda\beta\sum_{i=1}^{\infty} p^{i-1}M_i + 1, \tag{4.3.5}$$

$$M_{k+1} = \lambda\beta\sum_{i=0}^{k-1} p^i M_{k-i} + \lambda\beta\sum_{i=1}^{\infty} p^{k+i-1}M_i, \quad k=1,2,\ldots . \tag{4.3.6}$$

From (4.3.6),

$$M_{k+2} = (\lambda\beta+p)M_{k+1} = \cdots = (\lambda\beta+p)^k M_2, \quad k=1,2,\ldots . \tag{4.3.7}$$

Substitution of (4.3.7) into (4.3.5) and (4.3.6) leads to a set of two linear equations with two unknowns M_1 and M_2; finally

$$M_1 = \frac{1-p-\lambda\beta p}{1-p-\lambda\beta},$$
$$M_2 = \lambda\beta\frac{1-p(\lambda\beta+p)}{1-p-\lambda\beta},$$

so

$$ES_1 = \frac{\beta}{1-\lambda\beta/(1-p)} + \frac{\lambda}{2}(\beta^{(2)}-2\beta^2)\frac{1}{1-p}\frac{1-p-\lambda\beta p}{1-p-\lambda\beta}, \tag{4.3.8}$$

$$ES_k = \frac{\beta}{1-\lambda\beta/(1-p)} + \frac{\lambda}{2}(\beta^{(2)}-2\beta^2)\frac{1}{1-p}\lambda\beta\frac{1-p(\lambda\beta+p)}{1-p-\lambda\beta}(\lambda\beta+p)^{k-2}, \quad k=2,3,\ldots . \tag{4.3.9}$$

Note that $ES_k \rightarrow \frac{\beta}{1-\lambda\beta/(1-p)}$ for $k\rightarrow\infty$, which is the mean sojourn time per visit in the case of a negative exponential service time distribution. Also note that (cf. Takács [10])

$$ES = \sum_{i=1}^{\infty} p^{i-1}ES_i = \frac{\beta}{1-p-\lambda\beta} + \frac{\lambda}{2}(\beta^{(2)}-2\beta^2)\frac{1}{1-p-\lambda\beta}. \tag{4.3.10}$$

In a future paper it will be shown that the results obtained in Section 3 for the deterministic feedback model can be extended to the generalized Bernoulli feedback model with exponential service times.

REFERENCES

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

[2] DISNEY, R.L., KÖNIG, D. (1985) Queueing networks: a survey of their random processes. *SIAM Review* **27**, 335-403.

[3] DISNEY, R.L., KÖNIG, D., SCHMIDT, V. (1984) Stationary queue-length and waiting-time distributions in single-server feedback queues. *Adv. Appl. Prob.* **16**, 437-446.

[4] DOSHI, B.T., KAUFMAN, J.S. (1987) Unpublished manuscript.

[5] FONTANA, B., DIAZ BERZOSA, C. (1984) Stationary queue-length distributions in an M/G/1 queue with two non-preemptive priorities and general feedback. In: *Performance of Computer-communication Systems,* eds. H. Rudin and W. Bux. North-Holland Publ. Cy., Amsterdam, 333-347.

[6] FONTANA, B., DIAZ BERZOSA, C. (1985) M/G/1 queue with N-priorities and feedback: joint queue-length distributions and response time distribution for any particular sequence. In: *Teletraffic Issues in an Advanced Information Society,* ITC-11, ed. M. Akiyama. North-Holland Publ. Cy., Amsterdam, 452-458.

[7] JAGERMAN, D.L. (1978) An inversion technique for the Laplace transform with application to approximation. *Bell System Tech. J.* **57**, 669-710.

[8] OTT, T.J. (1984) The sojourn-time distribution in the M/G/1 queue with processor sharing. *J. Appl. Prob.* **21**, 360-378.

[9] SIMON, B. (1984) Priority queues with feedback. *J. ACM.* **31**, 134-149.

[10] TAKÁCS, L. (1963) A single-server queue with feedback. *Bell System Tech. J.* **42**, 505-519.

[11] WOLFF, R.W. (1982) Poisson arrivals see time averages. *Oper. Res.* **30**, 223-231.