International Journal of Management Science
Vol. 2, No. 1, December 1996

# Sojourn Times in a Multiclass Priority Queue with Random Feedback

Sung-Jo Hong* · Tetsuji Hirayama**

## Abstract

We consider a priority-based multiclass queue with probabilistic feedback. There are $J$ service stations. Each customer belongs to one of the several priority classes, and the customers of each class arrive at each station in a Poisson process. A single server serves queued customers on a priority basis with a nonpreemptive scheduling discipline. The customers who complete their services feed back to the system instantaneously and join one of the queues of the stations or depart from the system according to a given probability. In this paper, we propose a new method to simplify the analysis of these queueing systems. By the analysis of busy periods and regenerative processes, we clarify the underlying system structure, and systematically obtain the mean for the sojourn time, i.e.,the time from the arrival to the departure from the system, of a customer at every station. The mean for the number of customers queued in each station at an arbitrary time is also obtained simultaneously.

* Department of Industrial Engineering, Dongguk University, Korea
** Institute of Information Sciences and Electronics, University of Tsukuba, Japan

## 1. Introduction

Single server priority-based multiclass queues with feedback of output customers can model many practical systems. Among them is a computer system in which a central processor simultaneously processes many different characteristics of service requirements for several classes of customers. For example, a central processor may have to do computations, send information to memory, retrieve information from memory, receive input, deliver output, etc. Furthermore, each customer may be served repeatedly for a certain reason. A single application of the system may require more than one task to be processed. For example, somebody may input some data which must be added to something that must be retrieved from memory and then printed. Also in a time-shared computer system, the interval of time during which the customer is permitted to remain in service is referred to as his quantum. The quantum offered may or may not be enough to satisfy his service. If sufficient, the customer depart from the system; if not, he reenters the system of queues and waits within the system until his second quantum starts, and so on. Eventually, after a sufficient number of visits to the processor, the customer will have gained enough service and will depart. In this paper, we consider such a multiclass queue with priority and random feedback. By random feedback we mean that a customer whose service has just been completed immediately joins queues again with assigned probability, or it departs from the system. Figure 1 shows the multiclass queue with random feedback.

Queueing models with feedback have been investigated extensively, and applied to the performance evaluation of computer systems and flexible manufacturing systems, etc. Disney [2] has been concerned with sojourn times in M / G / 1 queues with instantaneous, Bernoulli feedback. Berg, et al. [1] considered the system with deterministic routing, in which each customer requires $N$ services. They derived the set of linear equations for the mean sojourn times per visit to the first come first served discipline. Simon [7] considered the system with $c$ types of customers and $m$ levels of priority. Doshi and Kaufman [3] studied the sojourn time of a tagged customer who has just completed his $m^{th}$ pass in a multiclass M / G / 1 queue with Bernoulli feedback. Epema [4] has investigated a general single server time-sharing model with multiple queues and customer classes, priorities and feedback. The purpose of this paper is to unify above systems and to simplify the analysis.

In Section 2 we define our system as a stochastic process and then define the sojourn times as cost functions of a tagged customer that naturally stem from the analysis of the system. We assume that each customer belongs to one of the $J$ priority classes, 1 through $J$, where class $i$ has priority over class $j$ if $i<j$. We derive equations that are satisfied by the cost functions. The cost functions defined are closely related to busy periods. Thus in Section 3 we analyze busy periods of the system, and then investigate the mean for the number of customers in each station at the completion epoch of some busy periods. In Section 4 we derive expressions of the mean for the initial sojourn times, i.e., the time from the arrival to the completion time of his first service, and the mean for the number of customers in each station at a completion epoch of the initial sojourn time (initial stay) of a tagged customer. Customers are served in the first come first served (FCFS) basis or the last come first served (LCFS) basis for each class. Section 5 is devoted to solve the equations given in Section 2 to obtain the expressions of the mean for the sojourn times, i.e., the times from the arrival to the departure from the system, of a tagged customer for every station. In Section 6 we investigate the steady state value of the sojourn times and the number of customers in each station by using the little's formula and the Poisson arrivals see time averages(PASTA) property. A conservation relation of the system is considered in Section 7. We conclude in Section 8.

## 2. Model description

Let us introduce our model and associated notation specifically. Let there be $J$ classes of customers indexed as $1,2,\ldots,J$. Customers of class $j$ arrive in a Poisson process at station $j$(waiting room of class $j$) with rate $\lambda_j$. We assume that each station has an infinite capacity. Let $\lambda=\sum_{j=1}^{J}\lambda_j$. We assume that the classes of customers are priority classes such that class $i$ has priority over class $j$ if $i<j$. Customers are preferentially served by a single server in the order of priority, and for each class in the first come first served basis or the

last come first served basis. If once the service to a customer is started it is not disrupted until the service is completed even if customers with higher priority arrive (nonpreemptive discipline). After receiving a service, a class $i$ customer either feeds back to the system and proceeds to station $j$ with probability $p_{ij}$, or departs from the system with probability $p_{i0}=1-\sum_{j=1}^{J}p_{ij}$. For notational convenience, $p_{0i}=0$. The feedback probability matrix is given by $\mathbf{P}_m=(p_{ij}: i, j=1,...,m)(m=1,...,J)$. Note that feedback process of every customer can only depend on his current class. Arrival processes, service times and feedback processes are assumed to be independent of each other.

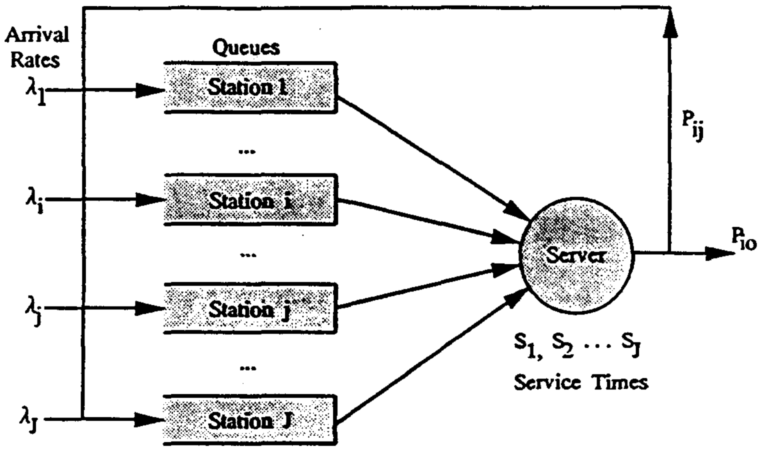


Figure 1. A multiclass priority queue with feedback.

Let $S_j$ be the independent random that denotes the length of the service time for a customer of class $j$. Let $r$ be a remaining service time of a customer found in service. Let $T_{ij}(r)$ be the total amount of service times of a class $i$ customer with his current remaining service time $r$ receives until he departs from the system or leaves for one of the stations $j+1,...,J$ for the first time. We sometimes use a notation $T_{ij}^{(m)}$ to denote a time $T_{ij}$ of an $m^{th}$ class $i$ customer. The expected value of $T_{ij}(r)$ is given by

$$E[T_{ij}(r)] \quad = \quad r+\sum_{k=1}^{j}p_{ik}E[T_{kj}(S_k)], \qquad j=0,1,\ldots,J. \tag{2.1}$$

An empty sum, which often occurs at $j=0$, is defined to be 0 from now on.

Specifically if we let $T_{ij} \equiv T_{ij}(S_i)$, then

$$E[T_{ij}] = E[S_i] + \sum_{k=1}^{j} p_{ik} E[T_{kj}]. \tag{2.2}$$

So we can obtain their solutions in a vector form if $(1-\mathbf{P}_j)^{-1}$ exists. Further, we define

$$\rho_0^+ \equiv 0, \tag{2.3}$$

$$\rho_j^+ \equiv \sum_{i=1}^{j} \lambda_i E[T_{ij}], \qquad j = 1, \ldots, J. \tag{2.4}$$

Let us consider the following assumption:

**Assumption 2.**

1. $\mathbf{P}^n{}_J \to 0$ as $n \to \infty$.
2. $\rho_J^+ < 1$.

The first assumption is a sufficient condition for existence of $(I-\mathbf{P}_j)^{-1}$ for $j=1,\ldots,J$. Since $\rho_J^+$ is the traffic intensity of the system, throughout the paper we assume that the system is not saturated in the steady state. Let $R^n$ be a $n$-tuple of nonnegative real numbers.

The number of class $i$ customers except for a customer found in service is denoted by $n^i$ and its vector is denoted by $\mathbf{n}=(n_1,\ldots,n_J) \in R^J$. These customers who are not currently served are called *waiting customers*. Let $\kappa$ be the class of a customer found in service. The class of a customer found in service at time $t$ is denoted by $\kappa(t)$, and his remaining service time is denoted by $r(t)$. We assume that $\kappa(\tau)=0$ if the system is empty at time $\tau$, or $\tau$ is a service completion time. The number of waiting customers in station $j$ at time $t$ is denoted by $n_j(t)$ and its vector is denoted by $\mathbf{n}(t) = (n_1(t),\ldots,n_J(t))$. Let us assume that customers are numbered in the order of their arrivals from outside the system. Let us consider transition epochs of these processes such as customers's arrival epochs and service completion epochs. Let us consider the $e^{th}$ customer arrives from outside the system at one of the stations at some epoch $\sigma_0^e(e=1,2,\ldots)$. Let $M^e$ be the number of his visits to the stations from his arrival at time $\sigma_0^e$ until his departure from the system. Then let $\sigma_k^e$ be a time epoch just when he would arrive at one of the stations after completing his $k^{th}$ service ($k=1,2,...,M^e$). For convenience, $\sigma_k^e = \sigma_{M^e}^e$ for $k>M^e$. Then let $I(t)$ denote the class (station) of a customer who arrived at the last transition epoch before $t(t \geq 0)$. $I(t)$ is right continuous with left-hand limits and $I(t)=0$ if a customer departs from the system at the last transition epoch before $t$. Then we define a stochastic process $Q \equiv \{\mathbf{Y}(t) \equiv (I(t), \kappa(t), r(t), \mathbf{n}(t) : t \geq 0\}$ that represents an evolution of the system. Possible values of $\mathbf{Y}(t)$ are called *system states* whose generic values are denoted by $\mathbf{Y} \equiv (j,\kappa,r,\mathbf{n})$. The state space is denoted by $\varepsilon$. All of the component processes from $\{\kappa(t) : t \geq 0\}$ through $\{\mathbf{n}(t) : t \geq 0\}$ are left continuous with right-hand limits except for the following cases. First, the process $\{\kappa(t) : t \geq 0\}$ take the value 0 at the moments of the completion epochs of service periods, at which epochs the process has left-hand limits and right-hand limits. Second, the process $\{n_j(t) : t \geq 0\}$ is right continuous with left-hand limits at service completion epochs. We restrict our attention to *work conserving* scheduling algorithms: for any system state $\mathbf{Y} \equiv (j,\kappa,r,\mathbf{n}) \in \varepsilon$, the unfinished work (total amount of remaining service time) of all customers, $T_{\kappa J}(r) + \sum_{(j,m) \in \mathbf{n}} T_{jj}^{(m)}$ (work-in-system: Wolff [10]) is conserved for all scheduling algorithms and where the server is not idle when any customers are waiting. For an $e^{th}$ arrival customer, we would like to derive the cost functions to be defined below. Let

$$G_{W_j}^e(t) \equiv \begin{cases} 1, & \text{if an } e^{th} \text{ customer stays at station } j \text{ at time } t, \\ 0, & \text{if } e^{th} \text{ customer does not stay at station } j \text{ at time } t, \end{cases} \tag{2.5}$$

where $t \geq 0$, $j=1,...,J$ and $e=1,2,....$ Then we define his sojourn time $W_j^e$ at station $j$ as the interval from the arrival to the departure from the system. That is,

$$W_j^e \equiv \int_0^\infty C_{Wj}^e(s)ds. \tag{2.6}$$

If the $e^{th}$ customer has arrived at the system, which the state is $\mathbf{Y}$, at time $\sigma_l^e$. Its expected value $W_j(\mathbf{Y},e,l)$ after time $\sigma_l^e$ conditioned on the system state $\mathbf{Y}$ is defined as follow:

$$W_j(\mathbf{Y},e,l) \equiv E\left[\int_{\sigma_l^e}^{\infty} C_{Wj}^e(t)dt|\mathbf{Y}(\sigma_l^e)=\mathbf{Y}\right], \tag{2.7}$$

for $l$=0,1,.... $W_j(\mathbf{Y},e,l)$ denotes the mean *sojourn time* of an $e^{th}$ customer spent at station $j$ after time $\sigma_l^e$ given that the system is in state $\mathbf{Y}$ at that time. On the other hand, we define his *initial stay* as the interval from the time $\sigma_l^e$ to the completion of his first service just prior to $\sigma_{l+1}^e$. We call his initial stay initial sojourn time. That is,

$$W_j^I(\mathbf{Y},e,l) \equiv E\left[\int_{\sigma_l^e}^{\sigma_{l-1}^e} C_{Wj}^e(t)dt|\mathbf{Y}(\sigma_l^e)=\mathbf{Y}\right], \tag{2.8}$$

where $l$=0,1,... and $e$=1,2,..... $W_j^I(\mathbf{Y},e,l)$ denotes the mean *initial sojourn time* of an $e^{th}$ customer spent at station $j$ during his initial stay after time $\sigma_l^e$ given that the system is in state $\mathbf{Y}$ at that time. Then the mean sojourn time $W_j(\mathbf{Y},e,l)$ is decomposed into two parts: the mean initial sojourn time and the mean sojourn time after his initial stay at station $j$. We mathematically express the fact as follow:

$$W_j(\mathbf{Y},e,l) = W_j^I(\mathbf{Y},e,l)+E[W_j(\mathbf{Y}(\sigma_{l+1}^e),e,l+1)|\mathbf{Y}(\sigma_l^e)=\mathbf{Y}] \tag{2.9}$$

Of course, every scheduling algorithm has its own cost functions. After stating some assumptions, we will explicitly solve the equation (2.9) in section 5.

## 3. Busy period analysis

An important model in the analysis of priority queueing systems is the M/G/1 queue with exceptional first service. An exceptional first service busy period is an interval that begins when an arrival with exceptional service time finds the server is idle, and ends when, for the first time after that, a departure leaves the system empty. One interpretation of this model is that some kind of setup time is needed before an arrival who finds the system empty can begin service. The quantities defined in the last section will be shown to be closely related to some exceptional first service busy periods.

So we define some quantities related to them. Now let us consider the system is in state $\mathbf{Y}=(i,\kappa,r,\mathbf{n})\in\varepsilon$ at some transition epoch. We select a set of customers $C=C(\mathbf{Y})$ who stay in the system at that time. For example, if a customer of the $m^{th}$ queueing position at station $i$ is in $C$, then $(i,m)\in C$. Let $C^c=C^c(\mathbf{Y}:C)$ denote a set of customers who stay in the system at that time and are not in the set $C$. It is assumed that a class $i$ customer found in service is staying at the $0^{th}$ position of his queue.

We define an exceptional first service busy period $B^j(a)$ as the interval that begins when an arrival with exceptional service time $a$ finds the server is idle, and ends when, for the first time after that, the system is cleared of the customers from classes 1 through $j$. Then its expected value is given by [10]

$$E[B^j(a)] = \frac{a}{1-\rho_j^+}. \tag{3.1}$$

For any state $\mathbf{Y}$ and any customer set $C$, let $B^j(\mathbf{Y}:C)$ be a busy period initiated with state $\mathbf{Y}$ until the first time when the system is cleared of the customers in $C$ and the stations 1 through $j$, except for the customers in $C^c$. We call $B^j(\mathbf{Y}:C)$ a *class j busy period* initiated with $(\mathbf{Y}:C)$. For convenience, let $B^0(\mathbf{Y}:C)$ denote a time interval to complete service of a customer found in service and all customers in $C$. Their expected values can be obtained by the usual method [10]. That is,

$$E[B^0(\mathbf{Y};C)] = r + \sum_{(i,m)\in C} E[S_i], \tag{3.2}$$

$$E[B^j(\mathbf{Y};C)] = \frac{E[T_{\kappa j}(r)] + \sum_{(i,m)\in C} E[T_{ij}]}{1-\rho_j^+}. \tag{3.3}$$

Now we consider the number of customers in the system at the completion epoch of a class $j$ busy period. Let $N_{kl}^j(k=1,\ldots,J)$ denote the number of customers in station $l$ at the completion epoch of a class $j$ busy period initiated by a class $k$ customer. Further, if the busy period is initiated by a class $k$ customer with exceptional first service $a$, $N_{kl}^j(a)$ denotes the number of customers in station $l$ at the completion epoch of the class $j$ busy period. Then it can be shown that

$$E[N_{kl}^j(a)] = \lambda_l a + p_{kl} + \sum_{i=1}^{j}\{\lambda_i a + p_{ki}\}E[N_{il}^j], \tag{3.4}$$

where $0 \leq j < l \leq J$. By removing the condition on $r$, we have

$$E[N_{kl}^j] = \lambda_l E[S_k] + p_{kl} + \sum_{i=1}^{j}\{\lambda_i E[S_k] + p_{ki}\}E[N_{il}^j]. \tag{3.5}$$

It can be solved under Assumption 2. Now we define

$$\xi_l^j = \lambda_l + \sum_{i=1}^{j}\lambda_i E[N_{il}^j], \tag{3.6}$$

$$\chi_{kl}^j = \begin{cases} 0, & k=0, \\ p_{kl} + \sum_{i=1}^{j} p_{ki}E[N_{il}^j], & k=1,\ldots,J, \end{cases} \tag{3.7}$$

where $0 \leq j < l \leq J$. Also we let $N_l^j(\mathbf{Y}:C)(0 \leq j < l \leq J)$ be the number of customers in station $l$ at the completion epoch of $B^j(\mathbf{Y}:C)$. Then we obtain the formula:

$$\begin{aligned} E[N_l^j(\mathbf{Y};C)] &= \sum_{m\in C_l^c} 1 + E[N_{\kappa l}^j(r)] + \sum_{(i,m)\in C} E[N_{il}^j] \\ &= \sum_{m\in C_l^c} 1 + r\xi_l^j + \chi_{\kappa l}^j + \sum_{(i,m)\in C}\left\{E[S_i]\xi_l^j + \chi_{il}^j\right\}, \end{aligned} \tag{3.8}$$

where $C_l^c=\{m:(l,m)\notin C\}$. Since we are now considering work conserving priority scheduling algorithms defined in Section 2, these quantities $E[N_l^j(\mathbf{Y}:C)]$ are invariant for all scheduling algorithms under consideration. From these analyses, we can obtain the explicit formulae of the cost functions.

## 4. Initial sojourn times and Queue sizes

In this section we first derive initial sojourn times $W_i^I(\cdot)$ in station $i$ for the $e^{th}$ arriving customer(a tagged customer) at time $\sigma_l^e$ conditioned on the system state $\mathbf{Y}(\sigma_l^e)$ at that time. Also we derive the expected value for the number of customers in each station at the completion epoch of the initial sojourn time. Customers are preferentially served in the order priority, and for each class in the FCFS discipline or the LCFS discipline.

**FCFS discipline**

We first consider the system that has the FCFS discipline for each class. Thus class $i$ customers are served on a first-come-first-served basis if no customers are in the stations 1 through $i-1$ nonpreemptively. Let us consider the tagged customer who arrived at station $i$ with the FCFS discipline at $\sigma_l^e(e=1,2,\dots$ and $l=0,1,\dots)$. Let $\mathbf{Y}(\sigma_l^e)=\mathbf{Y}\equiv(i,\kappa,r,\mathbf{n})\in\varepsilon$ be any state of the system at his arrival epoch. Then the initial sojourn time for the tagged customer consists of his service time and the waiting time, i.e., the time interval from his arrival to the moment at which he first receives service. If we select a customer set $C_i^F$ composed of a customer found in service and all customers from classes 1 through $i$ (except for the tagged customer) who are in the system at time $\sigma_l^e$. Then the waiting time for the tagged customer is a class $i-1$ busy period initiated with $\{\mathbf{Y}:C_i^F\}$, regardless of the disciplines adopted by stations $1,\dots, i-1$. Thus we can write as

$$\begin{aligned} W_i^I(\mathbf{Y},e,l) &= E[B^{i-1}(\mathbf{Y};C_i^F)+S_i] \\ &= \frac{E[T_{\kappa,i-1}(r)]}{1-\rho_{i-1}^+}+\sum_{j=1}^{i} n_j\frac{E[T_{ji-1}]}{1-\rho_{i-1}^+}+E[S_i] \end{aligned} \tag{4.1}$$

Let us now consider the number of customers in each station $k(k=1,\dots,J)$ at the completion epoch of the his initial stay. Since the mean initial sojourn time of the tagged customer is the sum of the class $i-1$ busy period initiated with $\{\mathbf{Y}:C_i^F\}$, and his service time $S_i$, from (3.8) we have

$$\begin{aligned} E[n_k(\sigma_{l+1}^e)|\mathbf{Y}(\sigma_l^e)=\mathbf{Y}] &= E[N_k^{i-1}(\mathbf{Y};C_i^F)]+\lambda_k E[S_i] \\ &= \begin{cases} \lambda_k E[S_i], & k<i, \\ r\xi_i^{i-1}+\chi_{\kappa i}^{i-1}+\sum_{j=1}^{i} n_j\{E[S_j]\xi_i^{i-1}+\chi_{ji}^{i-1}\}+\lambda_i E[S_i], & k=i, \\ n_k+r\xi_k^{i-1}+\chi_{\kappa k}^{i-1}+\sum_{j=1}^{i} n_j\{E[S_j]\xi_k^{i-1}+\chi_{jk}^{i-1}\}+\lambda_k E[S_i], & k>i. \end{cases} \end{aligned} \tag{4.2}$$

**LCFS discipline**

On the other hand if the service discipline for each class is LCFS basis, the class $i$ customers are served on a last-come-first-served basis if no customers are in the stations 1 through $i-1$ nonpreemptively. In this case, if we select a customer set $C_i^L$ composed of a customer found in service and all customers from classes 1 through $i-1$ (except for the tagged customer) who are in the system at time $\sigma_l^e$ Then the initial sojourn time of the tagged customer is composed of a class $i$ busy period initiated with $\{\mathbf{Y};C_i^L\}$, and the tagged customer's service $S_i$, regardless of the disciplines adopted by stations $1,\dots,i-1$. Hence,

$$\begin{aligned} W_i^I(\mathbf{Y},e,l) &= E[B^i(\mathbf{Y};C_i^L)+S_i] \\ &= \frac{E[T_{\kappa,i}(r)]}{1-\rho_i^+}+\sum_{j=1}^{i} n_j\frac{E[T_{ji}]}{1-\rho_i^+}+E[S_i] \end{aligned} \tag{4.3}$$

Also the mean number of customers in each station $k$ at the completion epoch of the initial stay of the tagged customer is obtained as follow:

$$\begin{aligned} & E[n_k(\sigma_{l+1}^e)|\mathbf{Y}(\sigma_l^e)=\mathbf{Y}] = E[N_k^{i-1}(\mathbf{Y};C_i^L)]+\lambda_k E[S_i] \\ &= \begin{cases} \lambda_k E[S_i], & k<i, \\ n_i+\lambda_i E[S_i], & k=i, \\ n_k + r\xi_k^i+\chi_{\kappa k}^i+\sum_{j=1}^{i-1} n_j\{E[S_j]\xi_k^i+\chi_{jk}^i\}+\lambda_k E[S_i], & k>i. \end{cases} \end{aligned} \tag{4.4}$$

Note that the above quantities do not count the feedback of the tagged customer. As we may see from the above expressions, the expected values derived in this section are linear combination of some components of the state. We summarize them in the following lemma.

**Lemma 4.1** Consider the multiclass M / G / 1 system with feedback defined in Section 2. By appropriately choosing a nonnegative vector $\mathbf{w}^i \in R^{J\times 1}$ and nonnegative constants $\varphi^i$, $\psi^{i\kappa}$ and $w^i$, for any $e=1,2,\dots$, we have

$$W_j^I(\mathbf{Y},e,l) = \begin{cases} r\varphi^i+\psi^{i\kappa}+\mathbf{n}\mathbf{w}^i+w^i, & i=j \\ 0, & i\neq j \end{cases} \tag{4.5}$$

for $\mathbf{Y}=(i,\kappa,r,\mathbf{n})\in\varepsilon$ and $l=0,1,\dots$. □

**Lemma 4.2** Consider the multiclass M / G / 1 system with feedback defined in Section 2. By appropriately choosing a nonnegative matrix $\mathbf{u}^i\in R^{J\times J}$ and nonnegative vectors $\mathbf{a}^i\in R^{1\times J}$, $\mathbf{d}^{i\kappa}\in R^{1\times J}$, for any $e=1,2,\dots$, we have

$$E[\mathbf{n}(\sigma_{l+1}^e)|\mathbf{Y}(\sigma_l^e)=\mathbf{Y}] = r\mathbf{a}^i+\mathbf{d}^{i\kappa}+\mathbf{n}\mathbf{u}^i+\mathbf{g}^i \tag{4.6}$$

for $\mathbf{Y}=(i,\kappa,r,\mathbf{n})\in\varepsilon$ and $l=0,1,\dots$. □

The important things to consider about these expressions are first that elements $I(\sigma_l^e)$, $\kappa(\sigma_l^e)$, $r(\sigma_l^e)$, and $\mathbf{n}(\sigma_l^e)$ of state of the system should be sufficient for estimating the expected value of the cost functins and, second that these expected values should be linear functions of components $(r,\mathbf{n})$ for any given class $i$ customer. Of course, every scheduling algorithm has its own coefficients.

## 5. Sojourn times analysis

In this section we derive explicit formulae for the sojourn times $W_j^l(\cdot)$ under some assumptions. As we have defined, $W_j^l(\mathbf{Y},e,l)$ is the mean sojourn time of the $e^{th}$ customer, who arrives at time $\sigma_l^e$ when the system is in state $\mathbf{Y}$, spent at station $j$ until he departs from the system. We make the expressions (4.5) and (4.6) as the following assumption.

**Assumption 5.** For any $e^{th}$ customer ($e$=1,2,...), we assume

$$W_j^I(\mathbf{Y},e,l) = \begin{cases} r\varphi^i + \psi^{i\kappa} + \mathbf{n}\mathbf{w}^i + w^i, & i = j \\ 0, & i \neq j \end{cases}$$

$$E[\mathbf{n}(\sigma_{l+1}^e)|\mathbf{Y}(\sigma_l^e) = \mathbf{Y}] = r\mathbf{a}^i + \mathbf{d}^{i\kappa} + \mathbf{n}\mathbf{u}^i + \mathbf{g}^i,$$

hold for the state $\mathbf{Y}=(i,\kappa,r,\mathbf{n})\in\varepsilon$. □

Of course, the assumption is satisfied by the priority scheduling algorithms considered.

Now we focus on a station $j$ ($j$=1,...,$J$). Let $\mathcal{J}=J^2$ and define the following vectors and matrices:

$$\mathbf{W} = (0,\ldots,0,\mathbf{w}^{j\prime},0,\ldots,0)' \in \mathcal{R}^{\mathcal{J}\times 1},$$

$$\mathbf{U_p} = \begin{pmatrix} \mathbf{u}^1 p_{11} & \mathbf{u}^1 p_{12} & \cdots & \mathbf{u}^1 p_{1J} \\ \mathbf{u}^2 p_{21} & \mathbf{u}^2 p_{22} & \cdots & \mathbf{u}^2 p_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}^J p_{J1} & \mathbf{u}^J p_{J2} & \cdots & \mathbf{u}^J p_{JJ} \end{pmatrix} \in \mathcal{R}^{\mathcal{J}\times\mathcal{J}},$$

where $'$ denotes transposition of a vector. We now suppose $(\mathbf{I}-\mathbf{U_p})^{-1}$ exists, where $\mathbf{I}$ is an identity matrix. Then we can define

$$\begin{pmatrix} \mathbf{w}_{1j} \\ \vdots \\ \mathbf{w}_{Jj} \end{pmatrix} \equiv (\mathbf{I} - \mathbf{U_p})^{-1}\mathbf{W} \in \mathcal{R}^{J\times 1}, \tag{5.1}$$

and define

$$\mathbf{V} = \begin{pmatrix} \mathbf{g}^1 \sum_{k=1}^{J} p_{1k}\mathbf{w}_{kj} \\ \vdots \\ \mathbf{g}^{j-1} \sum_{k=1}^{J} p_{j-1k}\mathbf{w}_{kj} \\ w^j + \mathbf{g}^j \sum_{k=1}^{J} p_{jk}\mathbf{w}_{kj} \\ \mathbf{g}^{j+1} \sum_{k=1}^{J} p_{j+1k}\mathbf{w}_{kj} \\ \vdots \\ \mathbf{g}^J \sum_{k=1}^{J} p_{Jk}\mathbf{w}_{kj} \end{pmatrix} \in \mathcal{R}^{J\times 1},$$

From Assumption 2, $(\mathbf{I}-\mathbf{P}_J)^{-1}$ exists. Then we can define

$$\begin{pmatrix} w_{1j} \\ \vdots \\ w_{Jj} \end{pmatrix} \equiv (\mathbf{I} - \mathbf{P}_J)^{-1}\mathbf{V} \in \mathcal{R}^{J\times 1}, \tag{5.2}$$

$$\varphi_{ij} = \begin{cases} \mathbf{a}^i \sum_{l=1}^{J} p_{il}\mathbf{w}_{lj}, & i \neq j, \\ \varphi^j + \mathbf{a}^j \sum_{l=1}^{J} p_{jl}\mathbf{w}_{lj}, & i = j, \end{cases} \tag{5.3}$$

$$\psi_{ij}^{\kappa} = \begin{cases} 0, & \kappa = 0, \\ \mathbf{d}^{i\kappa} \sum_{l=1}^{J} p_{il}\mathbf{w}_{lj}, & i \neq j \text{ and } \kappa \neq 0, \\ \psi^{j\kappa} + \mathbf{d}^{j\kappa} \sum_{l=1}^{J} p_{jl}\mathbf{w}_{lj}, & i = j \text{ and } \kappa \neq 0. \end{cases} \tag{5.4}$$

As we have stated, for every scheduling algorithm, these vectors and matrices have different values.

Now we can derive the following theorem.

**Theorem 5.** We assume that Assumption 2 and Assumption 5 hold, and that $(\mathbf{I}-\mathbf{U}_\mathrm{p})^{-1}$ exists. Let $\mathbf{Y}=(i,\kappa,r,\mathbf{n})\in\varepsilon$ be a given system state. Then, for any station $j (j=1,\ldots,J)$,

$$\hat{W}_j(\mathbf{Y}, e, l) = \begin{cases} r\varphi_{ij} + \psi_{ij}^{\kappa} + \mathbf{n}\mathbf{w}_{ij} + w_{ij}, & l = 0 \\ \mathbf{n}\mathbf{w}_{ij} + w_{ij}, & l > 0 \end{cases} \tag{5.5}$$

is a solution of equation (2.9). □

*Proof.* We shall show that (5.5) satisfies the equation (2.9). For $l=0$, let $\mathbf{Y}(\sigma_0^e)=\mathbf{Y}\equiv(i,\kappa,r,\mathbf{n})$ be the state of the system at arrival epoch $\sigma_0^e$. For $i=j$,

$$\begin{aligned}
&W_j^I(\mathbf{Y},e,l)+E[W_j(\mathbf{Y}(\sigma_{l+1}^e),e,l+1)|\mathbf{Y}] \\
&= r\varphi^i+\psi^{i\kappa}+\mathbf{n}\mathbf{w}^i+w^i+\sum_{k=1}^{J}p_{ik}E[\mathbf{n}(\sigma_{l+1}^e)\mathbf{w}_{kj}+w_{kj}|\mathbf{Y}] \\
&= r\varphi^i+\psi^{i\kappa}+\mathbf{n}\mathbf{w}^i+w^i+\sum_{k=1}^{J}p_{ik}\left[\left\{r\mathbf{a}^i+\mathbf{d}^{i\kappa}+\mathbf{n}\mathbf{u}^i+\mathbf{g}^i\right\}\mathbf{w}_{kj}+w_{kj}\right] \\
&= r\left\{\varphi^i+\mathbf{a}^i\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}\right\}+\left\{\psi^{i\kappa}+\mathbf{d}^{i\kappa}\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}\right\} \\
&\quad +\mathbf{n}\left\{\mathbf{w}^i+\mathbf{u}^i\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}\right\}+\left\{w^i+\mathbf{g}^i\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}+\sum_{k=1}^{J}p_{ik}w_{kj}\right\} \\
&= \hat{W}_j(\mathbf{Y},e,l).
\end{aligned}$$

The last equation follows from the definition of the constants $\mathbf{w}_{ij}$ *and* $w_{ij}$.

*For* $i\neq j$,

$$\begin{aligned}
&E[W_j(\mathbf{Y}(\sigma_{l+1}^e),e,l+1)|\mathbf{Y}] \\
&= \sum_{k=1}^{J}p_{ik}E[\mathbf{n}(\sigma_{l+1}^e)\mathbf{w}_{kj}+w_{kj}|\mathbf{Y}] \\
&= \sum_{k=1}^{J}p_{ik}\left[\left\{r\mathbf{a}^i+\mathbf{d}^{i\kappa}+\mathbf{n}\mathbf{u}^i+\mathbf{g}^i\right\}\mathbf{w}_{kj}+w_{kj}\right] \\
&= r\left\{\mathbf{a}^i\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}\right\}+\left\{\mathbf{d}^{i\kappa}\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}\right\}+\mathbf{n}\left\{\mathbf{u}^i\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}\right\} \\
&\quad +\left\{\mathbf{g}^i\sum_{k=1}^{J}p_{ik}\mathbf{w}_{kj}+\sum_{k=1}^{J}p_{ik}w_{kj}\right\} \\
&= \hat{W}_j(\mathbf{Y},e,l).
\end{aligned}$$

For $l>0$, let $\mathbf{Y}(\sigma_l^e)=\mathbf{Y}\equiv(i,0,0,\mathbf{n})$ be the state of the system at arrival epoch $\sigma_l^e$.

For $i=j$,

$$
\begin{aligned}
& W_j^I(\mathbf{Y}, e, l) + E[W_j(\mathbf{Y}(\sigma_{l+1}^e), e, l+1)|\mathbf{Y}] \\
&= \mathbf{n}\mathbf{w}^i + w^i + \sum_{k=1}^{J} p_{ik} E[\mathbf{n}(\sigma_{l+1}^e)\mathbf{w}_{kj} + w_{kj}|\mathbf{Y}] \\
&= \mathbf{n}\mathbf{w}^i + w^i + \sum_{k=1}^{J} p_{ik} \left[\left\{\mathbf{n}\mathbf{u}^i + \mathbf{g}^i\right\} \mathbf{w}_{kj} + w_{kj}\right] \\
&= \mathbf{n}\left\{\mathbf{w}^i + \mathbf{u}^i \sum_{k=1}^{J} p_{ik}\mathbf{w}_{kj}\right\} + \left\{w^i + \mathbf{g}^i \sum_{k=1}^{J} p_{ik}\mathbf{w}_{kj} + \sum_{k=1}^{J} p_{ik} w_{kj}\right\} \\
&= \hat{W}_j(\mathbf{Y}, e, l).
\end{aligned}
$$

The last equation follows from the definition of the constants $\mathbf{w}_{ij}$ *and* $w_{ij}$.
*For* $i \neq j$,

$$
\begin{aligned}
& E[W_j(\mathbf{Y}(\sigma_{l+1}^e), e, l+1)|\mathbf{Y}] \\
&= \sum_{k=1}^{J} p_{ik} E[\mathbf{n}(\sigma_{l+1}^e)\mathbf{w}_{kj} + w_{kj}|\mathbf{Y}] \\
&= \sum_{k=1}^{J} p_{ik} \left[\left\{\mathbf{n}\mathbf{u}^i + \mathbf{g}^i\right\} \mathbf{w}_{kj} + w_{kj}\right] \\
&= \mathbf{n}\left\{\mathbf{u}^i \sum_{k=1}^{J} p_{ik}\mathbf{w}_{kj}\right\} + \left\{\mathbf{g}^i \sum_{k=1}^{J} p_{ik}\mathbf{w}_{kj} + \sum_{k=1}^{J} p_{ik} w_{kj}\right\} \\
&= \hat{W}_j(\mathbf{Y}, e, l).
\end{aligned}
$$

Hence, $W_j(\mathbf{Y},e,l)$ satisfies equation (2.9). □

## 6. Steady state mean sojourn times

We have considered the system in an arbitrary state. In this section we evaluate steady state values of the cost functions $W_j(\cdot)$. Let us consider the system operating under some fixed scheduling algorithm defined in Section 2. We define the following time average values:

$$\bar{r}^\kappa \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t r(s) 1\{\kappa(s) = \kappa\} ds, \tag{6.1}$$

$$\bar{q}^\kappa \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t 1\{\kappa(s) = \kappa\} ds, \tag{6.2}$$

where

$$1\{\kappa(s)=\kappa\} = \begin{cases} 1, & \kappa(s)=\kappa \\ 0, & \kappa(s)\neq\kappa \end{cases}$$

Then $\tilde{r}^{\kappa}$ is a cumulative remaining service time when the server serves $\kappa$-customers, and $\tilde{q}^{\kappa}$ is a fraction of time that the server serves $\kappa$-customers. We are willing to assume that

[A1] The stochastic process $Q$ is regenerative [6].

If we let $N_B$ be the number of customers served during a regenerative cycle.

[A2] The system is initially empty and $E[N_B]<\infty$.

The throughput $\theta^j$ of the station $j(j=1,\ldots,J)$ is defined as the following equation:

$$\tilde{\theta}_j = \lambda_j + \sum_{i=1}^{J} p_{ij}\tilde{\theta}_i \tag{6.3}$$

Then it can be shown that

$$\tilde{r}^{\kappa} = \begin{cases} \tilde{\theta}_\kappa \overline{s^2}_\kappa/2, & \kappa=1,\ldots,J, \\ 0, & \kappa=0, \end{cases} \tag{6.4}$$

$$\tilde{q}^{\kappa} = \begin{cases} \tilde{\theta}_\kappa E[S_\kappa], & \kappa=1,\ldots,J, \\ 1-\rho_J^{+}, & \kappa=0. \end{cases} \tag{6.5}$$

If we let $\tilde{r}$ be the remaining service time of a customer found in service at any time,

$$\tilde{r} = \sum_{k=1}^{J} \frac{\tilde{\theta}_k \overline{s^2}_k}{2} \tag{6.6}$$

Since we consider the system with the steady state, we define the following customer average for the sojourn time at station $j$:

$$\bar{W}_j(\cdot) \equiv \lim_{N\to\infty} \frac{1}{N}\sum_{e=1}^{N} W_j^e, \tag{6.7}$$

if the limit exists. Then, from the regenerative assumptions [A1] and [A2], we may find that this customer average value may be represented as:

$$\bar{W}_j(\cdot) = \frac{E[\sum_{e=1}^{N_B} W_j^e]}{E[N_B]}, \tag{6.8}$$

if we may assume that every numerator in the right-hand side of the above expression is finite. Further the customer average values of components of the state are defined by:

$$\bar{n}_j \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} n_j(\sigma_0^e), \tag{6.9}$$

$$\bar{r}^\kappa \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} r(\sigma_0^e) \mathbf{1}\{\kappa(\sigma_0^e) = \kappa\}, \tag{6.10}$$

$$\bar{q}^\kappa \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{e=1}^{N} \mathbf{1}\{\kappa(\sigma_0^e) = \kappa\}, \tag{6.11}$$

if these limits exist, where $k=1,\ldots,J$. Let $\bar{\mathbf{n}} \equiv (\bar{n}_1,\ldots,\bar{n}_J)$. Now we assume that

[A3] $E[\sum_{e=1}^{N_B} n_j(\sigma_0^e) < \infty$ and $E[\sum_{e=1}^{N_B} r(\sigma_0^e) < \infty$ for $j=1,\ldots,J$.

Then we have

$$\bar{n}_j = \frac{E[\sum_{e=1}^{N_B} n_j(\sigma_0^e)]}{E[N_B]}, \tag{6.12}$$

$$\bar{r}^\kappa = \frac{E[\sum_{e=1}^{N_B} r(\sigma_0^e) \mathbf{1}\{\kappa(\sigma_0^e) = \kappa\}]}{E[N_B]}, \tag{6.13}$$

$$\bar{q}^\kappa = \frac{E[\sum_{e=1}^{N_B} \mathbf{1}\{\kappa(\sigma_0^e) = \kappa\}]}{E[N_B]}, \tag{6.14}$$

The time average value for the number of customers $n_j$ is defined by:

$$\tilde{n}_j \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t n_j(s) ds, \tag{6.15}$$

Let $\tilde{\mathbf{n}} \equiv (\tilde{n}_1,\ldots,\tilde{n}_J)$. Now we can see the following lemma concerned with representations of steady state values for the mean sojourn times.

**Lemma 6.** We assume that Assumption 2, Assumption 5 and the steady state assumptions from [A1] through [A3] hold. Then $E[\sum_{e=1}^{N_B} W_j^e] < \infty$ $(j=1,\ldots, J)$. Further we have the following representations:

$$\bar{W}_j(\cdot) = \sum_{i=1}^{J} \frac{\lambda_i}{\lambda} \left\{ \sum_{k=1}^{J} \left( \bar{r}^k \varphi_{ij} + \psi_{ij}^k \bar{q}^k \right) + \bar{\mathbf{n}} \mathbf{w}_{ij} + w_{ij} \right\} \quad \square \tag{6.16}$$

We use the generalized version of Little's formula ($H = \lambda G$) [8] that equates time average values of costs with customer average values of costs to obtain

$$\tilde{n}_j + \tilde{q}^j = \lambda \bar{W}_j(\cdot) \tag{6.17}$$

From the Poisson arrivals see time averages (PASTA) property [9], it can be shown $\bar{r}^\kappa = \tilde{r}^\kappa$ and $\bar{q}^\kappa = \tilde{q}^\kappa$. Then, from the equations (6.6) and (6.16), we have

$$\bar{W}_j(\cdot) = \sum_{i=1}^{J} \frac{\lambda_i}{\lambda} \left\{ \tilde{r} \varphi_{ij} + \sum_{k=1}^{J} \psi_{ij}^k \tilde{q}^k + \tilde{\mathbf{n}} \mathbf{w}_{ij} + w_{ij} \right\} \tag{6.18}$$

Thus we obtain

$$\tilde{n}_j + \tilde{q}^j = \sum_{i=1}^{J} \lambda_i \left\{ \tilde{r} \varphi_{ij} + \sum_{k=1}^{J} \psi_{ij}^k \tilde{q}^k + \tilde{\mathbf{n}} \mathbf{w}_{ij} + w_{ij} \right\} \tag{6.19}$$

If we define the follong vectors

$$\begin{aligned}
\mathbf{S} &\equiv \sum_{i=1}^{J} \lambda_i (\mathbf{w}_{i1}, \ldots, \mathbf{w}_{iJ}) \\
\mathbf{s} &\equiv \sum_{i=1}^{J} \lambda_i (w_{i1}, \ldots, w_{iJ}) \\
\mathbf{s}_0 &\equiv \sum_{i=1}^{J} \lambda_i \left( \tilde{r} \varphi_{i1} + \sum_{k=1}^{J} \psi_{i1}^k \tilde{q}^k, \ldots, \tilde{r} \varphi_{iJ} + \sum_{k=1}^{J} \psi_{iJ}^k \tilde{q}^k \right) - (\tilde{q}^1, \ldots, \tilde{q}^J),
\end{aligned}$$

we obtain the equation that determines the steady state expected value $\tilde{\mathbf{n}}$:

$$\tilde{\mathbf{n}} = \mathbf{s}_0 + \tilde{\mathbf{n}} \mathbf{S} + \mathbf{s}. \tag{6.20}$$

If the inverse matrix $(\mathbf{I} - \mathbf{S})^{-1}$ exists, we have

$$\tilde{\mathbf{n}} = (\mathbf{s}_0 + \mathbf{s})(\mathbf{I} - \mathbf{S})^{-1} \tag{6.21}$$

Finally let $\bar{W}_j(i)$ be the steady state value for the mean sojourn time of a customer spent at station $j$, given that the customer arrives at station $i$ from outside the system. Then we can get the steady state values:

$$\bar{W}_j(i) = \tilde{r}\varphi_{ij} + \sum_{k=1}^{J} \psi_{ij}^k \tilde{q}^k + (\mathbf{s}_0 + \mathbf{s})(\mathbf{I} - \mathbf{S})^{-1}\mathbf{w}_{ij} + w_{ij} \tag{6.22}$$

These results are arranged in the following theorem:

**Theorem 2.** We assume that the multiclass M / G / 1 system with feedback defined in Section 2 satisfies the steady state assumptions from [A1] through [A3]. Let $\tilde{\mathbf{n}} \equiv (\tilde{n}_1,\ldots,\tilde{n}_J)$. be the vector of the steady state mean number of customers defined by the equation (6.15). Further, let $\bar{W}_j(i)$ be the steady state value for the mean sojourn time of a customer, who initially arrives at station $i$ from outside the system, spent at station $j$ until his departure from the system. If we assume that Assumption 2 and Assumption 5 hold, and that the inverse matrix $(\mathbf{I}-\mathbf{S})^{-1}$ exists, for $i,j=1,\ldots,J$,

$$\tilde{\mathbf{n}} = (\mathbf{s}_0 + \mathbf{s})(\mathbf{I} - \mathbf{S})^{-1},$$

$$\bar{W}_j(i) = \tilde{r}\varphi_{ij} + \sum_{k=1}^{J} \psi_{ij}^k \tilde{q}^k + (\mathbf{s}_0 + \mathbf{s})(\mathbf{I} - \mathbf{S})^{-1}\mathbf{w}_{ij} + w_{ij} \quad \square$$

Of course, the total mean sojourn time $\bar{W}.(i)$ of a customer, who arrives at station $i$ from outside the system, spent in the system from the arrival to the departure from the system is given by

$$\begin{aligned} \bar{W}.(i) &= \sum_{j=1}^{J} \bar{W}_j(i) \\ &= \tilde{r}\sum_{j=1}^{J}\varphi_{ij} + \sum_{k=1}^{J}\sum_{j=1}^{J}\psi_{ij}^k\tilde{q}^k + (\mathbf{s}_0 + \mathbf{s})(\mathbf{I} - \mathbf{S})^{-1}\sum_{j=1}^{J}\mathbf{w}_{ij} + \sum_{j=1}^{J} w_{ij}. \end{aligned} \tag{6.23}$$

## 7. Conservation relation

In this section we consider a conservation relation that holds between the mean sojourn times for all customer classes. If a scheduling algorithm improves the sojourn time for some class, the sojourn time for another class will degrade.

Scheduling algorithms that have been considered in this paper are work conserving. Let $\mathbf{Y}=(i,\kappa,r,\mathbf{n})\in\varepsilon$ be any state of the system. In this case, the unfinished work $U(\mathbf{Y})$ at state $\mathbf{Y}$ is defined as the time required to complete the service of all customers present at that time(That is, when the system is in the state $\mathbf{Y}$) until their departure from the system. Then the expected value of the work is

$$E[U(\mathbf{Y})] = E[T_{\kappa J}(r)] + \sum_{k=1}^{J} n_k E[T_{kJ}]. \tag{7.1}$$

The value is conserved for all work conserving scheduling algorithms defined in section 2. Its steady state (time average) value $\tilde{U}$ is given by

$$\tilde{U} = \tilde{r} + \sum_{k=1}^{J} \tilde{q}^k \sum_{l=1}^{J} p_{kl} E[T_{lJ}] + \sum_{k=1}^{J} \tilde{n}_k E[T_{kJ}] \tag{7.2}$$

From equations (6.17) and (6.22), we obtain

$$\tilde{n}_j + \tilde{q}^j = \sum_{k=1}^{J} \lambda_k \bar{W}_j(k), \quad j = 1, \ldots, J. \tag{7.3}$$

Hence, we have

$$\begin{aligned} \tilde{U} &= \tilde{r} + \sum_{k=1}^{J} \tilde{q}^k \sum_{l=1}^{J} p_{kl} E[T_{lJ}] + \sum_{j=1}^{J} \sum_{k=1}^{J} \lambda_k \bar{W}_j(k) E[T_{jJ}] - \sum_{j=1}^{J} \tilde{q}^j E[T_{jJ}] \\ &= \tilde{r} - \sum_{k=1}^{J} \tilde{q}^k E[S_k] + \sum_{j=1}^{J} \sum_{k=1}^{J} \lambda_k \bar{W}_j(k) E[T_{jJ}]. \end{aligned} \tag{7.4}$$

By considering the system with a scheduling algorithm that serves customers nonpreemptively from their external arrival until their final departure in the order of FCFS, it can be shown that the total work $\tilde{U}$ is equivalent to the mean sojourn of customers of such M/G/1 system. Hence the left-hand side of the above expression can be easily calculated by the Pollaczek-Khinchin mean value formula [5].

$$\tilde{U} = \frac{\sum_{j=1}^{J} \lambda_j E[T_{jJ}^2]}{2(1-\rho_J^+)} \tag{7.5}$$

We finally obtain a conservation relation for the multiclass M / G / 1 queues with feedback. Assume that the multiclass M / G / 1 queues with feedback defined in Section 2 is in the steady state defined in Section 6. Further, we assume that the assumptions in Theorem 2 hold. Let $\bar{W}_j(i)$ $(i,j=1,\ldots,J)$ be the mean *sojourn time* of customers, who initially arrive at station $i$ from the outside of the system, spent at station $j$ until their departure from the system. Then the equation

$$C = \sum_{j=1}^{J}\sum_{i=1}^{J} \lambda_i \bar{W}_j(i) E[T_{jJ}], \tag{7.6}$$

holds for all scheduling algorithms defined in Section 2, where

$$C = \tilde{U} - \tilde{r} + \sum_{k=1}^{J} \tilde{q}^k E[S_k], \tag{7.7}$$

is a constant independent of these scheduling algorithms. □

## 8. Conclusions

We have concerned with the multiclass M / G / 1 queues with feedback. Customers are classified into $J$ groups according to their priorities. Non-preemptive scheduling algorithms have been analyzed. That is, if once the service to a customer is started it is not disrupted until the service is completed. Performance measures for every arriving customer such as the expected values of sojourn times are defined as conditional expectations of the system state at their arrival epochs. These values are divided into two parts: the expected values concerned with his initial sojourn times, and the expected values accumulated after his initial sojourn times. We have obtained a set of equations that are satisfied by these values. Then we have obtained these expected values concerned with the initial sojourn times from the analysis of (exceptional first service) busy periods for each service discipline adopted by each station. Further we have considered the steady state of the system and derived the steady state values of these system performance measures by using the generalized little's formula and the PASTA property. The special features of our method are summarized as follows:

1. We have treated the system performance measures such as the mean sojourn times explicitly as the cost functions of the system state.
2. Sufficient conditions that the objective cost functions can be explicitly derived have been given.
3. The mean values for the sojourn times are given in the matrix form. Hence, the algorithm that yields these values is easily constructed.

By appropriately choosing system parameters and combining the service disciplines, we can also construct many of the scheduling algorithms considered up to date.

**Acknowledgements.** The authors are grateful to the referees for their valuable comments and suggestions. We are also grateful to an Associate Editor for his comment on the improvement of the paper.

## References

[1] van den Berg, J.L., Boxma, O.J. and W.P. Groenendijk, "Sojourn times in the M / G / 1 queue with deterministic feedback," *Commun. Statist. -Stochastic Model*, Vol.5, No.1 (1989), pp.115-129.

[2] Disney, R.L., "A note on sojourn times in M / G / 1 queues with instantaneous Bernoulli feedback," *Naval Res. Log. Quarterly*, Vol.28 (1981), pp.679-684.

[3] Doshi, B.T. and J.S. Kaufman, "Sojourn time in an M / G / 1 queue with Bernoulli feedback," in: *Queueing Theory and its Applications (Liber Amicorum for J.W. Cohen)*, eds. O.J. Boxma and R. Sysky, North-Holland, Amsterdam (1988), pp.207-233.

[4] Epema, D.H.J., "Mean waiting times in a general feedback queue with

priorities," *Performance Evaluation*, Vol.13 (1991), pp.45-85.

[5] Kleinrock, L., *Queueing Systems Vol. II: Computer Applications*, Wiley, New York, 1976.

[6] Ross, S.M., *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, 1976.

[7] Simon, B., "Priority queues with feedback," *Journal of ACM*, Vol.31 (1984), pp.134-149.

[8] Whitt, W., "A review of $L=\lambda W$ and extensions," *Queueing Systems*, Vol.9 (1991), pp.235-268.

[9] Wolff, R.W., "Poisson arrivals see time averages," *Operations Research*, Vol.30 (1982), pp.223-231.

[10] Wolff, R.W., *Stochastic Modeling and the Theory of Queues*, Prentice-Hall, New Jersey, 1989.