

Polling Systems with Job Failures and with Station Failures

Merav Shomrony and Uri Yechiali

Department of Statistics and Operations Research
School of Mathematical Sciences, The Raymond and Beverly Sackler
Faculty of Exact Sciences, Tel Aviv University, Tel-Aviv 69978, Israel
(mshomrony@elta.ac.il, uriy@post.tau.ac.il)

ABSTRACT

We consider a polling system with Exhaustive service policy subject to job failures or to station failures. Each of the N stations (nodes) comprising the system is, in isolation, an $M/G/1$ - type queue. Failures at each node follow a Poisson process. A job failure causes the job to be rejected from the system, while a station failure (breakdown) causes the server to switch over to the next node. The case with job failures is transformed into a classical polling model. However, the case with station failures does not yield a solution when analyzed via the station occupancy approach. We therefore bring into use interpolation theory, never employed before in polling systems, and obtain explicit (approximated) solutions of key performance measures.

Key words: Polling; Exhaustive; Failures; Interpolation theory.

1. INTRODUCTION

This paper analyzes Polling systems with job failures and with station failures (breakdowns). Such systems (under various scenarios, described in the sequel) have been partially treated before in the literature (see, for example, Kofman and Yechiali [1996,1997]), but the analysis was incomplete because the Probability Generating Function (*PGF*) and mean of a key variable: X_i^i - the number of customers present in queue i when the server polls this queue - could not be explicitly derived. In this work we overcome this shortcoming by bringing into use interpolation theory, a method never used before for the analysis of polling systems.

In classical polling models, customers (jobs) arrive, independently, to each of several $M/G/1$ - type queues (stations) being attended by a single server that moves from one queue to another, incurring switch-over times. The server visits the stations in a cyclic order.

Consider however, a polling system subject to two types of failures, occurring when the server visits a queue: 1. a job failure, and 2. a station breakdown. Upon occurrence of a failure of the first type, service is interrupted to the currently processed job, which is rejected, while the server proceeds immediately

to serve the next job in the queue (if any). Examples for this type of failure are: an unsuccessful attempt of a job to access an unallocated memory, a transient fault of the communication system, or a short interruption in the power supply to the system, which is recovered immediately.

Upon occurrence of a station breakdown, the service at the queue being attended is stopped *immediately*, the current job is rejected and the server *switches* on to the next station. We assume that by the time the server visits this station again, the station has been recovered from its breakdown state. An example of such a failure is an undetected virus attack on the station.

Service policies control the amount of service given to each queue during a server's visit. We study in this work the Exhaustive service policy. In the classical Exhaustive policy, at each visit the server attends the queue until it becomes empty, and only then the server is allowed to move further. In contrast, in our model here, upon the occurrence of a station breakdown, the server immediately switches on to the next station, even if jobs are still present in the failed station.

Most of the results on polling systems until mid 80s are presented in a manuscript by Takagi [1986]. An update is given in Takagi [1990]. The Globally Gated regime was introduced by Boxma, Levy and Yechiali [1992]. Polling systems with breakdowns are treated in Kofman and Yechiali [1996, 1997] while polling systems involving both breakdowns and repairs were studied in Nakdimon and Yechiali [2003]. Issues of controlling polling systems are discussed in Yechiali [1993].

Our analysis is based on the laws of motion which connect between the occupancy variables X_i^j 's, being the number of customers present in queue j at polling instant of queue i . The classical method of analysis utilizes joint multi-variate probability generating functions of the system's state (a vector of queue sizes) to derive a set of recursive equations for the means (and higher moments) of the variables X_i^j 's.

As indicated, with station failures, the above method does not yield explicit solutions of $\widehat{X}_i^i(z)$, the PGF of X_i^i , and of $E[X_i^j]$ or $E[(X_i^j)^2]$, needed for the complete evaluation of the *PGF* and mean of the stationary queue size at each station, as well as for the calculation of the moments of the waiting times. We therefore bring interpolation theory into use and approximate $\widehat{X}_i^i(z)$. This enables us to obtain explicit (approximated) solutions.

In section 2 we describe the underlying basic model and list the various symbols and notation. We further give some preliminary results, being used throughout the paper. Section 3 treats the Exhaustive policy. For each of the two failure scenarios we derive the multi-dimensional *PGF* of the system's state at polling instants, and calculate the mean queue sizes at those moments. We also calculate the *PGF* of the queue size in each station at an arbitrary moment, and derive the Laplace Stieltjes Transform (*LST*) of the waiting time of a customer in each queue. Numerical examples are presented, showing how the interpolation theory is used to obtain explicit results. We mention that the Gated and the Globally Gated (see Boxma et al [1992] and [1993] policies can be treated in a similar manner (see Shomrony and Yechiali [2005a]).

2. THE BASIC MODEL

2.1 Model description

We consider a polling system comprised of N stations (queues, channels), labeled $1, 2, \dots, N$, where customers (jobs, messages) arrive at station i according to a Poisson process with rate λ_i . However, the system is subject to random failures, independently of the arrival stream, and occurring at the station being served according to a Poisson process with rate γ_i . We consider two failure types: job failures and station failures, described in the introduction.

A single server renders service to the entire network by moving from one station to another, in a cyclic order, i.e. visiting the queues in the order $1, 2, \dots, N-1, N, 1, 2, \dots$. The server resides at a queue for a length of time determined by the service policy and by the failure type, and then moves on to the next station.

The server's switch-over duration to move from station i to the next is a random variable D_i . Each job in queue i ($i = 1, 2, \dots, N$) is characterized by an independent random service requirement B_i with $E[B_i^2] < \infty$. We assume that a failure causes the job in service to leave the system immediately, without completing its service requirement.

2.2 Notation and definitions

$\hat{X}(z) = E[z^X]$ denotes the probability generating function (*PGF*) of a discrete random variable X .

$\tilde{Y}(s) = E[e^{-sY}]$ denotes the Laplace Stieltjes transform (*LST*) of a non negative continuous random variable Y (with mean y and second moment $y^{(2)}$).

$X \sim Y$ denotes two random variables having the same probability distribution function.

X_i^j Number of jobs present in queue j at polling instant of queue i .
 $f_i(j) = E[X_i^j]$

$\theta_{X_i^i}^i = \theta_i$ Sojourn time (busy period) of the server in station i , starting with X_i^i jobs.

Φ_i Length of a busy period in a regular $M/G/1$ queue with arrival rate λ_i and service times M_i .

ϕ_i Length of a busy period in a regular $M/G/1$ queue with arrival rate λ_i and service times B_i .

T_i Inter-arrival time of failures to station i . $T_i \sim \exp(\gamma_i)$.

$A_i(t)$ Number of (Poisson with rate λ_i) arrivals to queue i during a time

interval of length t .

$G_i(\underline{z}) = E \left[\prod_{j=1}^N z_j^{X_i^j} \right]$ The joint probability generating function (*PGF*) of the system state vector \underline{X}_i , where $\underline{z} = (z_1, \dots, z_N)$ and

$$\underline{X}_i = (X_i^1, X_i^2, \dots, X_i^N).$$

B_i Service requirement for a job in channel i , having a density function $f_{B_i}(\cdot)$.

a_i Probability that a service (in channel i) is completed before a failure occurs. $a_i = \widetilde{B}_i(\gamma_i) = P(B_i \leq T_i)$.

D_i Switch-over duration when moving from channel i to the next channel.

$D = \sum_{i=1}^N D_i$ Total switch-over duration during a cycle.

$M_i = \min\{B_i, T_i\}$ Net service time of a customer in queue i .

$\rho_i = \lambda_i m_i$ The traffic load of queue i .

$\rho = \sum_{i=1}^N \rho_i$ The total traffic load of the system.

$K_i = K_i(X_i^i)$ Total number of customers successfully served in channel i during a visit of the server to that channel, while starting with X_i^i customers. That is, K_i is the number of customers whose service was not interrupted by a failure.

$\overline{K}_i = \overline{K}_i(X_i^i)$ Total number of service attempts (including unsuccessful attempts interrupted by a failure) in channel i during a visit of the server to that channel, while starting with X_i^i customers.

N_i Queue size at the end of a busy period.

U_i Duration of a successful (not interrupted) service time of a customer in queue i . That is, $U_i \sim B_i \mid B_i \leq T_i$.

\overline{U}_i The realized service time of an interrupted customer in queue i :
 $\overline{U}_i \sim T_i \mid T_i < B_i$.

L_i Number of customers left behind by an arbitrary departing customer from channel i . L_i also stands for the number of customers at channel i (in a steady-state) at an arbitrary point of time.

W_i Sojourn time of a customer in channel i .

Wq_i Waiting (queueing) time of an arbitrary customer before service at channel i .

C Length of a cycle. That is, the time interval between two consecutive polling instants of a queue.

$$\sigma = \sigma(\underline{z}) = \sum_{j=1}^N \lambda_j (1 - z_j)$$

$$\sigma_i = \sigma_i(z) = \sum_{\substack{j \\ (j \neq i)}} \lambda_j (1 - z_j)$$

$$\delta_i = \gamma_i + \lambda_i (1 - z)$$

2.3 Useful results

We now present a few results that will be useful in the sequel. Let $M_i = \min\{B_i, T_i\}$ be the *actual* net service time of an arbitrary customer in queue i . Then, we have the following:

Result 1: The *LST*, mean and second moment of M_i are given, respectively, by

$$\widetilde{M}_i(s) = \frac{\gamma_i + s \cdot \widetilde{B}_i(\gamma_i + s)}{\gamma_i + s} \quad (2.3-1)$$

$$m_i = E[M_i] = \frac{1 - \widetilde{B}_i(\gamma_i)}{\gamma_i} = \frac{1 - a_i}{\gamma_i} \quad (2.3-2)$$

and

$$m_i^{(2)} = E[M_i^2] = \frac{2}{\gamma_i^2} \cdot \left(\gamma_i \widetilde{B}_i'(\gamma_i) + 1 - \widetilde{B}_i(\gamma_i) \right) \quad (2.3-3)$$

Proof: By conditioning on $P(B_i \leq T_i)$ and $P(B_i > T_i)$, (2.3-1) follows.

By differentiating (2.3-1) once and twice at $s = 0$ we obtain, respectively, m_i and $m_i^{(2)}$.

Let U_i denote the duration of a successful (not interrupted) service time of a customer in queue i . That is, $U_i \sim B_i \mid B_i \leq T_i$. Then,

Result 2: The *LST* of U_i is given by

$$\widetilde{U}_i(s) = \frac{\widetilde{B}_i(\gamma_i + s)}{\widetilde{B}_i(\gamma_i)} \quad (2.3-4)$$

Proof:

$$\begin{aligned} \widetilde{U}_i(s) &= E[e^{-sU_i}] = E[e^{-sB_i} \mid B_i \leq T_i] = \frac{1}{P(B_i \leq T_i)} \cdot \int_{b=0}^{\infty} \int_{t=b}^{\infty} e^{-sb} \cdot \gamma_i e^{-\gamma_i t} \cdot f_{B_i}(b) dt db \\ &= \frac{1}{\widetilde{B}_i(\gamma_i)} \cdot \int_{b=0}^{\infty} e^{-sb} e^{-\gamma_i b} \cdot f_{B_i}(b) db = \frac{\widetilde{B}_i(\gamma_i + s)}{\widetilde{B}_i(\gamma_i)} \end{aligned}$$

Let \overline{U}_i be the **realized** service time of an interrupted customer in queue i . That is, $\overline{U}_i \sim T_i \mid T_i < B_i$. Then,

Result 3: The *LST* of \bar{U}_i is given by

$$\widetilde{\bar{U}}_i(s) = \frac{\gamma_i}{\gamma_i + s} \cdot \frac{(1 - \widetilde{B}_i(\gamma_i + s))}{1 - \widetilde{B}_i(\gamma_i)} \quad (2.3-5)$$

Proof: Similar to the derivation of (2.3-4).

Result 4: For all models considered, the mean cycle time is given by

$$E[C] = \frac{d}{1 - \rho} \quad (2.3-6)$$

Proof Follows directly, as in many polling systems, from the balance equation:

$$E[C] = \sum_{i=1}^N E[D_i] + \sum_{i=1}^N \lambda_i m_i E[C] = d + \left(\sum_{i=1}^N \rho_i \right) \cdot E[C] = d + \rho E[C]$$

3. EXHAUSTIVE SERVICE POLICIES

In the basic Exhaustive service policy model, at every visit the server attends each queue until it becomes empty, and only then is allowed to move on. We consider two Exhaustive-type policies, denoted *E1* and *E2*, as follows.

3.1 Case E1: job failures

3.1.1 Analysis

Under the *E1* model, upon occurrence of a failure, service is interrupted to the currently processed job, which is rejected, while the server proceeds immediately to serve the next job in the queue. That is, during a visit, the server attempts to serve all customers in the queue (including new arrivals), until it becomes empty.

The evolution of the system is given by the following law of motion,

$$X_{i+1}^j = \left\{ \begin{array}{ll} X_i^j + A_j(\theta_{X_i^i}^i + D_i) & i \neq j \\ A_i(D_i) & i = j \end{array} \right\} \quad (3.1-1)$$

where $\theta_{X_i^i}^i = \sum_{k=1}^{X_i^i} \Phi_{ik}$ and Φ_{ik} denote i.i.d busy periods in a regular *M/G/1*

model with service times M_i , and arrival rate λ_i , all distributed like Φ_i .

Equation (3.1-1) is similar to the law of motion of the regular Exhaustive policy, with the only difference that $M_i = \min\{B_i, T_i\}$ replaces B_i . Thus (c.f. Takagi [1986], Yechiali [1993]),

$$G_{i+1}(\underline{z}) = G_i(z_1, \dots, z_{i-1}, \widetilde{\Phi}_i(\sigma_i), z_{i+1}, \dots, z_N) \cdot \widetilde{D}_i(\sigma) \quad (3.1-2)$$

$$\text{where } \sigma_i = \sigma_i(\underline{z}) = \sum_{\substack{j \\ (j \neq i)}} \lambda_j (1 - z_j)$$

The mean number of customers in the various queues at polling instants is given by:

$$f_{i+1}(j) = \begin{cases} f_i(j) + \lambda_j E[\Phi_i] f_i(i) + \lambda_j d_i & i \neq j \\ \lambda_i d_i & i = j \end{cases} \quad (3.1-3)$$

where $E[\Phi_i] = m_i / (1 - \rho_i)$ and $\rho_i = \lambda_i m_i$.

Equation (3.1-3) yields

$$f_i(j) = \begin{cases} \lambda_j \left(\sum_{k=j+1}^{i-1} \rho_k \cdot \frac{d}{1-\rho} + \sum_{k=j}^{i-1} d_k \right) & i \neq j \\ \lambda_i (1 - \rho_i) \cdot \left(\frac{d}{1-\rho} \right) & i = j \end{cases} \quad (3.1-4)$$

where $E[C] = \frac{d}{1-\rho}$.

The expectation of $\overline{K}_i(X_i^i)$ can now be calculated:

$$\overline{K}_i(X_i^i) = X_i^i + A_i \left(\sum_{k=1}^{X_i^i} \Phi_{ik} \right)$$

$$\begin{aligned} E[\overline{K}_i(X_i^i)] &= f_i(i) + \lambda_i \cdot f_i(i) \cdot E[\Phi_i] = f_i(i) + \frac{\lambda_i \cdot f_i(i) \cdot m_i}{1 - \rho_i} \\ &= \frac{f_i(i)}{1 - \rho_i} = \frac{\lambda_i \cdot (1 - \rho_i) \cdot E[C]}{1 - \rho_i} = \lambda_i \cdot E[C] \end{aligned}$$

3.1.2 Waiting times

Waiting times are calculated in section 3.2.2 together with the calculation of waiting times for model E2.

3.2 Case E2: station failures

3.2.1 Analysis

Under the *E2* model, upon occurrence of a failure, the service at the queue being attended is stopped immediately, and the server *switches on to the next station*. In a busy period, the server attempts to serve all customers in the queue until a failure occurs.

The evolution of the system is given by the following law of motion,

$$X_{i+1}^j = \left\{ \begin{array}{ll} X_i^j + A_j(\theta_{X_i^i}^i + D_i) & i \neq j \\ N_i(X_i^i) + A_i(D_i) & i = j \end{array} \right\} \quad (3.2-1)$$

where $N_i(X_i^i)$ is the number of jobs present in queue i at an instant of server's departure from that queue, given that there were X_i^i jobs at the beginning of the busy period.

In order to find the *PGF* $G_{i+1}(z)$, we first define the following joint distribution transform of $(\theta_r^i, N_i(r))$: $\Theta_r^i(w, z) = E \left[e^{-w\theta_r^i} \cdot z^{N_i(r)} \right]$

Conditioning on the first service time $B_{i,1}$ in a busy period in queue i we have:

$$\Theta_r^i(w, z) = E \left[e^{-w\theta_r^i} \cdot z^{N_i(r)} \cdot 1_{B_{i,1} < T_{i,1}} \right] + E \left[e^{-w\theta_r^i} \cdot z^{N_i(r)} \cdot 1_{B_{i,1} \geq T_{i,1}} \right] \quad (3.2-2)$$

and

$$(\theta_r^i, N_i(r)) \stackrel{d}{=} \left\{ \begin{array}{ll} (B_{i,1} + \theta_{r-1+A_i(B_{i,1})}^i, N_i(r-1+A(B_{i,1}))), & B_{i,1} < T_{i,1} \\ (T_{i,1}, r-1+A_i(T_{i,1})), & B_{i,1} \geq T_{i,1} \end{array} \right\}$$

Following a derivation similar to Eliazar and Yechiali [1998] and Shomrony and Yechiali [2001] we get

$$\begin{aligned} E \left[e^{-w\theta_r^i} \cdot z^{N_i(r)} \cdot 1_{B_{i,1} < T_{i,1}} \right] &= \\ &= \sum_{j=0}^{\infty} \left(E \left[\frac{(\lambda_i B_i)^j}{j!} \cdot e^{-(w+\gamma_i+\lambda_i)B_i} \right] \cdot \Theta_{r+j-1}^i(w, z) \right) \end{aligned} \quad (3.2-3)$$

and

$$E \left[e^{-w\theta_r^i} \cdot z^{N_i(r)} \cdot 1_{B_{i,1} \geq T_{i,1}} \right] = \frac{\gamma_i z^{r-1}}{w + \gamma_i + \lambda_i(1-z)} \cdot \left(1 - \widetilde{B}_i(w + \gamma_i + \lambda_i(1-z)) \right) \quad (3.2-4)$$

By substituting (3.2-3) and (3.2-4) in (3.2-2) we obtain

$$\Theta_r^i(w, z) = \begin{cases} \sum_{j=0}^{\infty} (a_j^i(w) \cdot \Theta_{r+j-1}^i(w, z)) + c^i(w, z) \cdot z^r & r \geq 1 \\ 1 & r = 0 \end{cases} \quad (3.2-5)$$

where $a_j^i(w) = E \left[\frac{(\lambda_i B_i)^j}{j!} \cdot e^{-(w+\gamma_i+\lambda_i)B_i} \right]$ and

$$c^i(w, z) = \frac{\gamma_i}{z(w + \gamma_i + \lambda_i(1-z))} \cdot \left(1 - \widetilde{B}_i(w + \gamma_i + \lambda_i(1-z)) \right)$$

The solution for the last set of equations (3.2-5) is

$$\Theta_r^i(w, z) = \varphi^i(w, z) \cdot z^r + (1 - \varphi^i(w, z)) \cdot \left[\widetilde{\phi}_i(w + \gamma_i) \right]^r \quad (3.2-6)$$

where ϕ_i is the length of a busy period in a *regular* $M/G/1$ queue with arrival rate λ_i and service times B_i , and

$$\varphi^i(w, z) = \frac{c^i(w, z) \cdot z}{z - \widetilde{B}_i(w + \gamma_i + \lambda_i(1-z))} = \frac{\gamma_i \cdot \left(1 - \widetilde{B}_i(w + \delta_i) \right)}{(w + \delta_i) \cdot \left(z - \widetilde{B}_i(w + \delta_i) \right)}$$

where $\delta_i = \gamma_i + \lambda_i(1-z)$.

We now return to the derivation of $G_{i+1}(\underline{z})$. By using the law of motion (3.2-1) we get

$$\begin{aligned} G_{i+1}(\underline{z}) &= E \left[\prod_{j=1}^N z_j^{X_{i+1}^j} \right] = E \left[\prod_{\substack{j \\ (j \neq i)}} z_j^{X_i^j + A_j(\theta_{X_i^i}^i + D_i)} \cdot z_i^{N_i(X_i^i) + A_i(D_i)} \right] \\ &= E \left[\prod_{\substack{j \\ (j \neq i)}} z_j^{X_i^j} \cdot \prod_{\substack{j \\ (j \neq i)}} z_j^{A_j(\theta_{X_i^i}^i)} \cdot z_i^{N_i(X_i^i)} \cdot \prod_{j=1}^N z_j^{A_j(D_i)} \right] \end{aligned}$$

$$= E_{\underline{X}_i} \left[\prod_{\substack{j \\ (j \neq i)}} z_j^{X_j^{X_i^j}} \cdot E \left(\prod_{\substack{j \\ (j \neq i)}} z_j^{A_j(\theta_{X_i^i}^i)} \cdot z_i^{N_i(X_i^i)} | \underline{X}_i \right) \right] \cdot \widetilde{D}_i(\sigma) \quad (3.2-7)$$

Now,

$$\begin{aligned} E \left(\prod_{\substack{j \\ (j \neq i)}} z_j^{A_j(\theta_{X_i^i}^i)} \cdot z_i^{N_i(X_i^i)} | \underline{X}_i \right) &= E \left[E \left(\prod_{\substack{j \\ (j \neq i)}} z_j^{A_j(\theta_{X_i^i}^i)} \cdot z_i^{N_i(X_i^i)} | \underline{X}_i, N_i(X_i^i), \theta_{X_i^i}^i \right) \right] \\ &= E \left[\prod_{\substack{j \\ (j \neq i)}} e^{-\lambda_j(1-z_j)\theta_{X_i^i}^i} \cdot z_i^{N_i(X_i^i)} \right] = E \left[e^{-\sum_{\substack{j \\ (j \neq i)}} \lambda_j(1-z_j)\theta_{X_i^i}^i} \cdot z_i^{N_i(X_i^i)} \right] = \Theta_{X_i^i}^i(\sigma_i, z_i) \end{aligned}$$

By substituting this last result and the solution (3.2-6) in (3.2-7) we get

$$\begin{aligned} G_{i+1}(\underline{z}) &= E_{\underline{X}_i} \left[\prod_{\substack{j \\ (j \neq i)}} z_j^{X_j^{X_i^j}} \cdot \Theta_{X_i^i}^i(\sigma_i, z_i) \right] \cdot \widetilde{D}_i(\sigma) \quad (3.2-8) \\ &= E_{\underline{X}_i} \left[\prod_{\substack{j \\ (j \neq i)}} z_j^{X_j^{X_i^j}} \cdot \left(\varphi^i(\sigma_i, z_i) \cdot z_i^{X_i^{X_i^j}} + (1 - \varphi^i(\sigma_i, z_i)) \cdot [\widetilde{\phi}_i(\sigma_i + \gamma_i)]^{X_i^{X_i^j}} \right) \right] \cdot \widetilde{D}_i(\sigma) \\ &= \left[\varphi^i(\sigma_i, z_i) \cdot G_i(\underline{z}) + (1 - \varphi^i(\sigma_i, z_i)) \cdot G_i(z_1, \dots, z_{i-1}, \widetilde{\phi}_i(\sigma_i + \gamma_i), z_{i+1}, \dots, z_N) \right] \cdot \widetilde{D}_i(\sigma) \quad (1) \end{aligned}$$

$$\text{where } \varphi^i(\sigma_i, z_i) = \frac{\gamma_i \cdot (1 - \widetilde{B}_i(\sigma + \gamma_i))}{(\sigma + \gamma_i) \cdot (z_i - \widetilde{B}_i(\sigma + \gamma_i))}$$

By differentiating (3.2-8), the set of first moments $\{f_i(j) = E[X_i^j]\}$ is given by

$$f_{i+1}(j) = \left\{ \begin{array}{ll} f_i(j) + \frac{\lambda_j}{\gamma_i} \cdot \left(1 - \widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i)) \right) + \lambda_j d_i & i \neq j \\ f_i(i) + \left(\frac{\lambda_i}{\gamma_i} - \frac{1}{1 - \widetilde{B}_i(\gamma_i)} \right) \cdot \left(1 - \widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i)) \right) + \lambda_i d_i & i = j \end{array} \right\} \quad (3.2-9)$$

Unfortunately, in contrast to case E1, equations (3.2-9) *do not* lead to an explicit solution for the $f_i(j)$'s.

However, by summing equations (3.2-9) over i , for every $j = 1, \dots, N$, we derive

$$\frac{1 - \widehat{X}_j^j(\widetilde{\phi}_j(\gamma_j))}{1 - \widetilde{B}_j(\gamma_j)} = \lambda_j \cdot d + \lambda_j \cdot \sum_{k=1}^N \frac{1 - \widehat{X}_k^k(\widetilde{\phi}_k(\gamma_k))}{\gamma_k}$$

$$\text{Defining } x_k = \frac{1 - \widehat{X}_k^k(\widetilde{\phi}_k(\gamma_k))}{\gamma_k}$$

$$\text{and } \alpha_j = \frac{\gamma_j}{\lambda_j(1 - \widetilde{B}_j(\gamma_j))} - 1 = \frac{1}{\lambda_j m_j} - 1 = \frac{1 - \rho_j}{\rho_j},$$

the above equations can be written, for every $j = 1, \dots, N$, as

$$d = - \sum_{\substack{i \\ (i \neq j)}} x_i + \frac{\gamma_j}{\lambda_j(1 - \widetilde{B}_j(\gamma_j))} \cdot x_j - x_j = - \sum_{\substack{i \\ (i \neq j)}} x_i + \alpha_j x_j \quad (3.2-10)$$

In a matrix form, equations (3.2-10) can be expressed as $A\underline{x} = \underline{d}$, where A is a square matrix of order N , $\underline{x} = (x_1, x_2, \dots, x_N)^T$ and $\underline{d} = d \cdot \underline{1}$.

That is,

$$\begin{bmatrix} \alpha_1 & -1 & \dots & \dots & -1 & -1 \\ -1 & \alpha_2 & -1 & \dots & -1 & -1 \\ & & & & & \\ & & & & & \\ -1 & -1 & \dots & -1 & \alpha_{N-1} & -1 \\ -1 & -1 & \dots & \dots & -1 & \alpha_N \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_{N-1} \\ x_N \end{bmatrix} = \begin{bmatrix} d \\ d \\ \dots \\ d \\ d \end{bmatrix}$$

By subtracting the bottom row of A from each of the upper rows $i = 1, 2, \dots, N-1$, and then adding each row i ($i = 1, 2, \dots, N-1$) multiplied by $(\alpha_i + 1)^{-1}$ to the bottom row, A is modified to an upper triangular matrix, so that

$$\begin{bmatrix} \alpha_1 + 1 & 0 & \dots & \dots & 0 & -(1 + \alpha_N) \\ 0 & \alpha_2 + 1 & 0 & \dots & 0 & -(1 + \alpha_N) \\ & 0 & & & & \dots \\ & & & & 0 & \\ 0 & 0 & \dots & 0 & \alpha_{N-1} + 1 & -(1 + \alpha_N) \\ 0 & 0 & \dots & \dots & 0 & a_{N,N} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_{N-1} \\ x_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ d \end{bmatrix} \quad (3.2-11)$$

Using $m_i = \frac{1 - a_i}{\gamma_i}$, we have

$$a_{N,N} = \alpha_N + \frac{-(1 + \alpha_N)}{1 + \alpha_1} + \frac{-(1 + \alpha_N)}{1 + \alpha_2} + \dots + \frac{-(1 + \alpha_N)}{1 + \alpha_{N-1}} = \frac{1 - \rho}{\rho_N}$$

Since $\alpha_i + 1 = \frac{1}{\rho_i} \neq 0$, a *unique* solution exists if and only if $a_{N,N} \neq 0$, which is equivalent to $\rho \neq 1$.

The solution of the equations $A\underline{x} = \underline{d}$ is

$$x_N = \frac{d}{a_{N,N}} = \frac{d\rho_N}{1-\rho},$$

$$x_j = \frac{1+\alpha_N}{1+\alpha_j} \cdot x_N = \frac{\rho_j}{\rho_N} \cdot \frac{d\rho_N}{1-\rho} = \frac{d\rho_j}{1-\rho}, \text{ for } j = 1, \dots, N-1$$

That is, for every $j = 1, \dots, N$,

$$x_j = \frac{d\rho_j}{1-\rho}, \quad (3.2-12)$$

leading to

$$\widehat{X}_j^j(\widetilde{\phi}_j(\gamma_j)) = 1 - x_j\gamma_j = 1 - \frac{\lambda_j m_j \gamma_j d}{1-\rho} = 1 - \frac{\lambda_j(1 - \widetilde{B}_j(\gamma_j))d}{1-\rho}.$$

By substituting in equation (3.2-9) we get

$$f_{i+1}(j) = \left\{ \begin{array}{ll} f_i(j) + \frac{\lambda_j \rho_i d}{1-\rho} + \lambda_j d_i & i \neq j \\ f_i(i) + \frac{\lambda_i \rho_i d}{1-\rho} - \frac{\lambda_i d}{1-\rho} + \lambda_i d_i & i = j \end{array} \right\}. \quad (3.2-13)$$

Note that, although the set (3.2-13) looks 'nicer' than (3.2-9), we still **can not** solve it for the $f_i(j)$'s, since it is a non independent set. Thus, although $\widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i))$ is now known, we need the **distribution** of X_i^i , namely $\widehat{X}_i^i(z)$, in order to be able to obtain $f_i(i)$. A procedure for obtaining $\widehat{X}_i^i(z)$, using **interpolation** theory, is developed in section 3.2.3. .

3.2.2 Waiting times

3.2.2.1 Derivation of $\widehat{L}_i(z)$

We first derive the *PGF* of the queue size at an arbitrary moment in queue i . This will enable us to derive the *LST* of the waiting time of an arbitrary customer in that queue. (See also Eliazar and Yechiali [1998] and Shomrony and Yechiali [2001]).

For brevity, in the following derivation and in equations (3.2-14) to (3.2-16), we suppress the index i , indicating the queue number, and use the subscript n to indicate the n -th departing customer from that queue.

Define

τ_n – instant of the n -th departure, starting from time $t = 0$

L_n – queue size at time $\tau_n + 0$

I_n – the event: { at time τ_n the server leaves the queue }

$\overline{X}_n, \overline{N}_n$ – queue sizes at the beginning and the end, respectively, of the busy period containing instant τ_n

X_n^* – queue size at the beginning of the busy period containing instant τ_n as its first departure. i.e. $X_n^* \sim [\bar{X}_n | I_{n-1}]$

$$p_n = P(I_n) ; 1 - p_n = P(I_n^C)$$

By conditioning on the event I_{n-1} , we write,

$$E[z^{L_n}] = E[z^{L_n} | I_{n-1}^C] \cdot (1 - p_{n-1}) + E[z^{L_n} | I_{n-1}] \cdot p_{n-1} \quad (3.2-14)$$

Observe that, if I_{n-1} occurs, then the next departure (the n -th) will be the first in the next busy period, starting with \bar{X}_n customers, so that $L_n = \bar{X}_n - 1 + A(M_n)$.

(Attention: M_n here is the length of the service time of the n -th departing customer in queue i ; $M_n = \min\{B, T\}$ where B is the service time in queue i and T is the inter-arrival time of failures in queue i).

Otherwise, if I_{n-1}^c occurs, then $L_n = L_{n-1} - 1 + A(M_n)$.

Hence, (3.2-14) becomes:

$$\begin{aligned} E[z^{L_n}] &= E[z^{L_{n-1}-1+A(M_n)} | I_{n-1}^C] \cdot (1 - p_{n-1}) + E[z^{\bar{X}_n-1+A(M_n)} | I_{n-1}] \cdot p_{n-1} \\ &= \frac{\widetilde{M}(\lambda(1-z))}{z} \cdot \left[E[z^{L_{n-1}}] - E[z^{\bar{N}_{n-1}}] \cdot p_{n-1} + E[z^{X_n^*}] \cdot p_{n-1} \right] \end{aligned} \quad (3.2-15)$$

We repeat the above computation for L_n instead of z^{L_n} :

$$\begin{aligned} E[L_n] &= [E[A(M_n)] - 1 + E[L_{n-1} | I_{n-1}^C]] \cdot (1 - p_{n-1}) + [E[A(M_n)] - 1 + E[X_n^*]] \cdot p_{n-1} \\ &= \lambda m - 1 + E[L_{n-1}] - E[\bar{N}_{n-1}] \cdot p_{n-1} + E[X_n^*] \cdot p_{n-1} \end{aligned} \quad (3.2-16)$$

Since the Markov chain, embedded at service completion instants, converges in distribution, taking $n \rightarrow \infty$ in (3.2-15) and in (3.2-16) implies, respectively:

$$E[z^L] \cdot (z - \widetilde{M}(\lambda(1-z))) = \widetilde{M}(\lambda(1-z)) \cdot [E[z^{X^*}] - E[z^{\bar{N}}]] \cdot p \quad (3.2-17)$$

$$1 - \lambda m = [E[X^*] - E[\bar{N}]] \cdot p \quad (3.2-18)$$

in which L , X^* , \bar{N} and p are the steady state limits of L_n , X_n^* , \bar{N}_n and p_n , respectively.

From the last two equations we obtain:

$$\widehat{L}(z) = (1 - \rho) \cdot \frac{\widetilde{M}(\lambda(1-z))}{z - \widetilde{M}(\lambda(1-z))} \cdot \frac{E[z^{X^*}] - E[z^{\overline{N}}]}{E[X^*] - E[\overline{N}]} \quad (3.2-19)$$

Since $X^* \sim X|X \geq 1$, $\overline{N} \sim N|X \geq 1$ and $N|\{X=0\} = 0$, we get,
(see definition of N_i in section 3.2.1; X replaces the original X_i^i)

$$\frac{E[z^X] - E[z^N]}{E[X] - E[N]} = \frac{E[z^{X^*}] - E[z^{\overline{N}}]}{E[X^*] - E[\overline{N}]}$$

Substituting the last equation in (3.2-19) yields the *PGF* for L_i at epochs of departure

$$\widehat{L}_i(z) = (1 - \rho_i) \cdot \frac{\widetilde{M}_i(\lambda_i(1-z))}{z - \widetilde{M}_i(\lambda_i(1-z))} \cdot \frac{E[z^{X_i^i}] - E[z^{N_i}]}{E[X_i^i] - E[N_i]} \quad (3.2-20)$$

Note that this decomposition is similar to the Fuhrmann and Cooper decomposition result [1985] for the steady state queue size L , of a nonpreemptive $M/G/1$ queue with general vacation, i.e.,

$$\widehat{L}(z) = \widehat{L}_{M/G/1}(z) \cdot \chi(z)$$

where $\widehat{L}_{M/G/1}(z) = (1 - \rho) \cdot \frac{\widetilde{M}(\lambda(1-z)) \cdot (z-1)}{z - \widetilde{M}(\lambda(1-z))}$ is the steady state queue size of a regular $M/G/1$ queue, and $\chi(z)$ is the *PGF* of the queue size at an arbitrary instant during the vacation. Later, Borst [1995] has shown that $\chi(z) = \frac{1}{z-1} \cdot \frac{E[z^X] - E[z^N]}{E[X] - E[N]}$.

We now obtain the distribution of N_i for the *E2* model by using (3.2-6):

$$\widehat{N}_i(r)(z) = E[z^{N_i(r)}] = \Theta_r^i(0, z) = \varphi^i(0, z) \cdot z^r + (1 - \varphi^i(0, z)) \cdot [\widetilde{\phi}_i(\gamma_i)]^r$$

where,

$$\varphi^i(0, z) = \frac{\gamma_i \cdot (1 - \widetilde{B}_i(\delta_i))}{\delta_i \cdot (z - \widetilde{B}_i(\delta_i))}$$

By putting X_i^i instead of r and taking expectation, we get the *LST* of the number of customers at the **end** of a busy period in queue i :

$$\begin{aligned} \widehat{N}_i(z) &= E[z^{N_i}] = E_{X_i^i} \left[E \left[z^{N_i(X_i^i)} | X_i^i \right] \right] = E_{X_i^i} \left[\varphi^i(0, z) \cdot z^{X_i^i} + (1 - \varphi^i(0, z)) \cdot [\widetilde{\phi}_i(\gamma_i)]^{X_i^i} \right] \\ &= \frac{\widehat{X}_i^i(z) \cdot \gamma_i \cdot (1 - \widetilde{B}_i(\delta_i)) + \widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i)) \cdot [\delta_i z - \gamma_i - \widetilde{B}_i(\delta_i) \cdot \lambda_i(1-z)]}{\delta_i \cdot (z - \widetilde{B}_i(\delta_i))} \end{aligned} \quad (3.2-21)$$

Also, (by comparing (3.2-1) and (3.2-13)):

$$E[N_i] = E[N_i(X_i^i)] = f_i(i) + \frac{\lambda_i \rho_i d}{1 - \rho} - \frac{\lambda_i d}{1 - \rho} = E[X_i^i] - \frac{\lambda_i d}{1 - \rho} \cdot (1 - \rho_i) \quad (3.2-22)$$

Now, from (3.2-21) and (3.2-22) we get:

$$\frac{E[z^{X_i^i}] - E[z^{N_i}]}{E[X_i^i] - E[N_i]} = \frac{\left[\widehat{X}_i^i(z) - \widehat{X}_i^i(\tilde{\phi}_i(\gamma_i)) \right] \cdot \left[\delta_i z - \gamma_i - \widetilde{B}_i(\delta_i) \cdot \lambda_i(1 - z) \right]}{\delta_i \cdot \left(z - \widetilde{B}_i(\delta_i) \right) \cdot \frac{\lambda_i d}{1 - \rho} \cdot (1 - \rho_i)} \quad (3.2-23)$$

Using (2.3-1),

$$\frac{\widetilde{M}_i(\lambda_i(1 - z))}{z - \widetilde{M}_i(\lambda_i(1 - z))} = \frac{\frac{\gamma_i + \lambda_i(1 - z) \cdot \widetilde{B}_i(\delta_i)}{\delta_i}}{z - \frac{\gamma_i + \lambda_i(1 - z) \cdot \widetilde{B}_i(\delta_i)}{\delta_i}} = \frac{\gamma_i + \lambda_i(1 - z) \cdot \widetilde{B}_i(\delta_i)}{\delta_i z - \gamma_i - \lambda_i(1 - z) \cdot \widetilde{B}_i(\delta_i)} \quad (3.2-24)$$

Substituting (3.2-23) and (3.2-24) in (3.2-20) yields,

$$\widehat{L}_i(z) = \frac{\left[\widehat{X}_i^i(z) - \widehat{X}_i^i(\tilde{\phi}_i(\gamma_i)) \right] \cdot \left[\gamma_i + \lambda_i(1 - z) \cdot \widetilde{B}_i(\delta_i) \right]}{\delta_i \cdot \left(z - \widetilde{B}_i(\delta_i) \right) \cdot \frac{\lambda_i d}{1 - \rho}} \quad (3.2-25)$$

Note that, as in many $M/G/1$ -type queues, L_i also stands for the number of customers at channel i (under a steady-state condition) at an arbitrary instant of time.

Calculation of the PGF of L_i for the E1 case

$$\text{From (3.1-4): } E[X_i^i] = \frac{\lambda_i(1 - \rho_i)d}{1 - \rho}, \quad N_i = 0, \quad E[N_i] = 0$$

Hence, from (3.2-20):

$$\widehat{L}_i(z) = \frac{\widetilde{M}_i(\lambda_i(1 - z))}{z - \widetilde{M}_i(\lambda_i(1 - z))} \cdot \frac{\widehat{X}_i^i(z) - 1}{\frac{\lambda_i d}{1 - \rho}} \quad (3.2-26)$$

which is consistent with the result for the regular Exhaustive regime but with service times M_i .

3.2.2.2 Derivation of $\widetilde{W}_i(s)$

Let W_{q_i} be the waiting time of a customer in queue i , and let its sojourn time be $W_i = W_{q_i} + M_i$. Then, by using the standard argument that the number of customers left behind in queue i by a department customer equals the number of arrivals during W_i one gets

$$\widehat{L}_i(z) = \widetilde{W}_i(\lambda_i(1-z)) = \widetilde{W}_{q_i}(\lambda_i(1-z)) \cdot \widetilde{M}_i(\lambda_i(1-z))$$

Thus, based on (3.2-25) we have,

$$\widetilde{W}_i(s) = \widehat{L}_i(1 - \frac{s}{\lambda_i}) = \frac{\left[\widehat{X}_i^i(1 - \frac{s}{\lambda_i}) - \widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i)) \right] \cdot [\gamma_i + s \cdot \widetilde{B}_i(\gamma_i + s)] \cdot (1 - \rho)}{(\gamma_i + s) \cdot \left(1 - \frac{s}{\lambda_i} - \widetilde{B}_i(\gamma_i + s) \right) \cdot \lambda_i d} \quad (3.2-27)$$

and

$$\widetilde{W}_{q_i}(s) = \frac{\widehat{X}_i^i(1 - \frac{s}{\lambda_i}) - \widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i))}{1 - \frac{s}{\lambda_i} - \widetilde{B}_i(\gamma_i + s)} \cdot \frac{1 - \rho}{\lambda_i d} \quad (3.2-28)$$

Finally, one can substitute $\widehat{X}_i^i(\widetilde{\phi}_i(\gamma_i)) = 1 - \frac{\lambda_i(1 - \widetilde{B}_i(\gamma_i))d}{1 - \rho}$ in equations (3.2-25), (3.2-27) and (3.2-28).

3.2.3 Solution for the PGF $\widehat{X}_i^i(z)$ by interpolation

3.2.3.1 Probability distribution relation

As indicated, equations (3.2-13) do not yield an explicit solution for the means $f_i(j) = E[X_i^j]$. Therefore, we have to turn back to finding the PGF $\widehat{X}_i^i(z)$, from which $f_i(i)$ and $f_i(j)$ can be determined.

We will find $\widehat{X}_i^i(z)$ by using an interpolation method. For simplicity we consider a symmetric system.

Equation (3.2-8) relates the PGF $G_i(z_1, \dots, z_N)$ to the PGF $G_{i+1}(z_1, \dots, z_N)$.

We claim that in a symmetrical case, the following relation holds:

$$G_{i+1}(z_1, \dots, z_N) = G_i(z_2, \dots, z_N, z_1) \quad (3.2-29)$$

This follows since $X_i^j \sim X_{i+k}^{j+k}$ for every $i, j = 1, \dots, N$ and $k = 1, 2, 3, \dots$ (where the counting of k is in a cyclic manner - modulo N). In particular, $X_i^j \sim X_{i+1}^{j+1}$. That is, $z^{X_i^j} \sim z^{X_{i+1}^{j+1}}$, so that

$$z_1^{X_{i+1}^1} \cdot z_2^{X_{i+1}^2} \cdot \dots \cdot z_N^{X_{i+1}^N} \sim z_1^{X_i^N} \cdot z_2^{X_i^1} \cdot \dots \cdot z_N^{X_i^{N-1}}$$

Therefore,

$$G_{i+1}(z_1, \dots, z_N) = E \left[z_1^{X_{i+1}^1} \cdot z_2^{X_{i+1}^2} \cdot \dots \cdot z_N^{X_{i+1}^N} \right] = E \left[z_1^{X_i^N} \cdot z_2^{X_i^1} \cdot \dots \cdot z_N^{X_i^{N-1}} \right]$$

$$= G_i(z_2, \dots, z_N, z_1)$$

Using the relation (3.2-29), equation (3.2-8) becomes

$$G_i(z_2, \dots, z_N, z_1) =$$

$$= \left[\varphi^i(\sigma_i, z_i) \cdot G_i(\underline{z}) + (1 - \varphi^i(\sigma_i, z_i)) \cdot G_i \left(z_1, \dots, z_{i-1}, \tilde{\phi}_i(\sigma_i + \gamma_i), z_{i+1}, \dots, z_N \right) \right] \cdot \tilde{D}_i(\sigma)$$

(3.2-30)

where $\varphi^i(\sigma_i, z_i)$ is defined in section 3.2.1 .

3.2.3.2 Approximating $G_i(z_1, z_2, \dots, z_N)$ using polynomial interpolation

According to interpolation theory (see e.g. Atkinson [1989]) a function $h(x)$ with $n + 1$ continuous derivatives can be approximated by a polynomial sum:

$$P(x) = a_0 + a_1x + \dots + a_mx^m \quad (m \leq n) .$$

The coefficients a_0, a_1, \dots, a_m are uniquely evaluated when knowing the values of the function $h(x)$ at some distinct $m + 1$ points: x_0, \dots, x_m .

This is done by extracting the coefficients values from $m + 1$ equations, each equating the polynomial expression to the function value $h(x_i)$, $i = 0, 1, \dots, m$.

The approximation error at a point t is expressed by: $\frac{(t - x_0) \cdot \dots \cdot (t - x_m)}{(m + 1)!}$. $h^{(m+1)}(\xi)$, where ξ is a point lying in the minimal interval connecting the points t, x_0, x_1, \dots, x_m , and $h^{(m+1)}$ is the $(m + 1)$ st derivative of $h(\cdot)$.

The error gets smaller as t approaches one of the points x_0, \dots, x_m and when m becomes larger.

The method is extended to multi-variate functions and polynomials.

Since our goal is to approximate the multi-variate function $G_i(\underline{z})$, we now introduce a modification to the interpolation method in order to solve cases in which, instead of knowing values of the function $G_i(\underline{z})$ at distinct points, we have a relation (equation (3.2-30)), involving three values of the function $G_i(\cdot)$ at three separate points.

According to the method, a polynomial sum is substituted instead of $G_i(\cdot)$ in the above relation. Then, the coefficients a_0, a_1, \dots are calculated in a similar way to the original method by the following steps (which will be illustrated by numerical examples):

1. Insert the polynomial expression (with unknown coefficients) instead of $G_i(\underline{z})$ in the given relation.
2. Substitute a chosen value for \underline{z} in the equation.

3. Repeat step 2, several times (according to the polynomial degree) and get a system of independent equations with unknown parameters (the polynomial coefficients).
4. Solve for the unknown parameters and get the approximation to $G_i(\underline{z})$.

While approximating a generating function, one of the equations must be the normalizing equation: $G_i(\underline{1}) = 1$.

Since the error of the approximation is smaller at the vicinity of any one of the sampled points, we will carry the sampling of those points as close as possible to points in which we want the approximation to hold.

First order approximation

We begin by approximating $G_i(z_1, z_2, \dots, z_N)$ by using a polynomial of first degree with unknown parameters μ_0, \dots, μ_N :

$$G_i(z_1, z_2, \dots, z_N) \cong \mu_0 + \mu_1 z_1 + \dots + \mu_N z_N \quad (3.2-31)$$

From which we have,

$$\widehat{X}_i^i(z) = G_i(1, 1, \dots, z, \dots, 1) \cong \mu_0 + \mu_1 + \dots + \mu_{i-1} + \mu_i z + \mu_{i+1} + \dots + \mu_N \quad (3.2-32)$$

and

$$E[X_i^i] \cong \mu_i \quad (3.2-33)$$

Numerical example 1

We assume exponential service times and exponential switch-over times.

The following values for the system variables are used:

Number of queues: $N = 3$; Poisson rates of customer arrivals and of failures: $\lambda_i = 1$, $\gamma_i = 1$, respectively. Service times: $B_i \sim \exp(4)$, with $E[B_i] = b_i = \frac{1}{4}$

Switch over times: $D_i \sim \exp(10)$, with $E[D_i] = d_i = 0.1$

Starting with queue $i = 2$, the first order approximation for the function $G_2(z_1, z_2, z_3)$ is:

$$G_2(z_1, z_2, z_3) \cong \mu_0 + \mu_1 z_1 + \mu_2 z_2 + \mu_3 z_3 \quad (3.2-34)$$

Substituting (3.2-34) in (3.2-30) yields,

$$\mu_0 + \mu_1 z_2 + \mu_2 z_3 + \mu_3 z_1 =$$

$$= \widetilde{D}_2(\sigma) \cdot \left[\frac{\varphi^2(\sigma_2, z_2) \cdot (\mu_0 + \mu_1 z_1 + \mu_2 z_2 + \mu_3 z_3) + (1 - \varphi^2(\sigma_2, z_2)) \cdot (\mu_0 + \mu_1 z_1 + \mu_2 \cdot \widetilde{\phi}_2(\sigma_2 + \gamma_2) + \mu_3 z_3)}{\sigma + \gamma_2 + \frac{1}{b_2}} \right] \quad (3.2-35)$$

where $\sigma = \sigma(z) = \sum_{j=1}^3 \lambda_j(1-z_j) = 3 - z_1 - z_2 - z_3$ and $\widetilde{D}_2(\sigma) = 10/(10+\sigma)$.

Since $B_i \sim \exp(1/b_i)$, we have,

$$\varphi^2(\sigma_2, z_2) = \frac{\gamma_2 \cdot \left(1 - \frac{\frac{1}{b_2}}{\sigma + \gamma_2 + \frac{1}{b_2}} \right)}{(\sigma + \gamma_2) \cdot \left(z_2 - \frac{\frac{1}{b_2}}{\sigma + \gamma_2 + \frac{1}{b_2}} \right)} = \frac{1 - \frac{4}{\sigma + 5}}{(\sigma + 1)(z_2 - \frac{4}{\sigma + 5})} = \frac{1}{(\sigma + 5)z_2 - 4}$$

In addition, since ϕ_i is the length of a busy period in a regular $M/G/1$ queue, we have: $\widetilde{\phi}_i(s) = \widetilde{B}_i \left(s + \lambda_i(1 - \widetilde{\phi}_i(s)) \right)$.

In particular, for the $M/M/1$ queue, since $\widetilde{B}_i(s) = (1/b_i)/[1/b_i + s]$, we get

$$\widetilde{\phi}_i(s) = \frac{\frac{1}{b_i}}{\frac{1}{b_i} + s + \lambda_i(1 - \widetilde{\phi}_i(s))}.$$

The solution for $\widetilde{\phi}_i(s)$ is

$$\widetilde{\phi}_i(s) = \frac{s + \lambda_i + \frac{1}{b_i} - \sqrt{\left(s + \lambda_i + \frac{1}{b_i} \right)^2 - 4\lambda_i \cdot \frac{1}{b_i}}}{2\lambda_i} \quad i = 1, 2, 3$$

In particular, $\sigma_2 = 2 - z_1 - z_3$

and $\widetilde{\phi}_2(\sigma_2 + \gamma_2) = \frac{1}{2}[(8 - z_1 - z_3) - \sqrt{(8 - z_1 - z_3)^2 - 16}]$.

In order to solve for the unknowns parameters μ_0, μ_1, μ_2 and μ_3 in (3.2-34), we use the normalization equation $G_2(1, 1, 1) = \mu_0 + \mu_1 + \mu_2 + \mu_3 = 1$, and

substitute 3 different values of the point \underline{z} in equation (3.2-35).

Aiming at mean values, we approximate $G(\underline{z})$ at points *close* to $\underline{z} = \underline{1}$ in order to obtain a good approximation. We note that this approximation may be less accurate at other points of \underline{z} , which are *far* from $\underline{z} = \underline{1}$. With $\underline{z} = (1, 1.2, 1)$, $(1, 1, 1.1)$ and $(1.1, 1.1, 1)$ we get a set of 4 linear equations, whose solution is: $\mu_0 = -0.2685$; $\mu_1 = 0.1439$; $\mu_2 = 0.7009$; and $\mu_3 = 0.4237$.

Hence, from (3.2-34),

$$G_2(z_1, z_2, z_3) \cong -0.2685 + 0.1439z_1 + 0.7009z_2 + 0.4237z_3$$

Now,

$$\widehat{X}_2^2(z) = G_2(1, z, 1) \cong 0.2991 + 0.7009z$$

and $E[X_2^2] = f_2(2) \cong 0.7009$ By the symmetry assumption, $E[X_i^i] = E[X_2^2]$ for all i . Finally, all $f_i(j)$'s are determined from equation (3.2-13).

Second order approximation

In order to get more accurate results we now use a polynomial of second degree:

$$G_i(z_1, z_2, \dots, z_N) \cong \mu_0 + \mu_1 z_1 + \dots + \mu_N z_N + \mu_{11} z_1^2 + \dots + \mu_{NN} z_N^2 + \mu_{12} z_1 z_2 + \dots + \mu_{ij} z_i z_j + \dots + \mu_{N-1,N} z_{N-1} z_N \quad (3.2-36)$$

That is, in the general case there are $1 + 2N + \binom{N}{2}$ unknown parameters, implying that this same number of distinct points are used for the parameters evaluation.

Also,

$$\widehat{X}^i(z) = G_i(1, 1, \dots, z, \dots, 1) \cong \mu_0 + \mu_1 + \dots + \mu_i z_i + \dots + \mu_N + \mu_{11} + \dots + \mu_{ii} z^2 + \dots + \mu_{NN} + \mu_{12} + \dots + \mu_{ij} z + \dots + \mu_{N-1,N} \quad (3.2-37)$$

$$E[X_i^i] \cong \mu_i + 2\mu_{ii} + \sum_{j \neq i} \mu_{ij} \quad (3.2-38)$$

Numerical example 2

We use the same values as in numerical example 1. For queue $i = 2$, the polynomial turns out to be

$$G_2(z_1, z_2, z_3) \cong \mu_0 + \mu_1 z_1 + \mu_2 z_2 + \mu_3 z_3 + \mu_{11} z_1^2 + \mu_{22} z_2^2 + \mu_{33} z_3^2 + \mu_{12} z_1 z_2 + \mu_{13} z_1 z_3 + \mu_{23} z_2 z_3 \quad (3.2-39)$$

Substituting (3.2-39) in the left-hand-side of (3.2-30) yields

$$\begin{aligned} G_2(z_2, z_3, z_1) &\cong \mu_0 + \mu_1 z_2 + \mu_2 z_3 + \mu_3 z_1 + \mu_{11} z_2^2 + \mu_{22} z_3^2 + \mu_{33} z_1^2 + \mu_{12} z_2 z_3 + \mu_{13} z_2 z_1 \\ &+ \mu_{23} z_3 z_1 = \widetilde{D}_2(\sigma) \cdot \left[\varphi^2(\sigma_2, z_2) \cdot (\mu_0 + \mu_1 z_1 + \mu_2 z_2 + \mu_3 z_3 + \mu_{11} z_1^2 + \mu_{22} z_2^2 \right. \\ &\left. + \mu_{33} z_3^2 + \mu_{12} z_1 z_2 + \mu_{13} z_1 z_3 + \mu_{23} z_2 z_3) \right] \end{aligned}$$

$$\begin{aligned}
& + (1 - \varphi^2(\sigma_2, z_2)) \cdot \left(\mu_0 + \mu_1 z_1 + \mu_2 \tilde{\phi}_2(\sigma_2 + \gamma_2) + \mu_3 z_3 + \mu_{11} z_1^2 + \mu_{22} [\tilde{\phi}_2(\sigma_2 + \gamma_2)]^2 + \mu_{33} z_3^2 \right. \\
& \left. + \mu_{12} z_1 \tilde{\phi}_2(\sigma_2 + \gamma_2) + \mu_{13} z_1 z_3 + \mu_{23} \tilde{\phi}_2(\sigma_2 + \gamma_2) z_3 \right) \Big] \quad (3.2-40)
\end{aligned}$$

By substituting 9 different points $\underline{z} = (z_1, z_2, z_3)$ in equation (3.2-40) and using the normalizing equation $G_2(\underline{1}) = 1 = \sum_{i=0}^3 \mu_i + \sum_{i,j} \mu_{ij}$ we get a system of 10 linear equations from which we calculate the unknown parameters μ_i 's and μ'_{ij} 's.

The resulting values are

$$\begin{aligned}
\mu_0 &= 1.9254; \mu_1 = -0.3509; \mu_2 = -1.9174; \mu_3 = -1.2515; \mu_{11} = 0.11414; \\
\mu_{22} &= 0.809; \mu_{33} = 0.41114; \mu_{12} = 0.2763; \mu_{13} = 0.16786; \mu_{23} = 0.81576,
\end{aligned}$$

leading to

$$\begin{aligned}
\widehat{X}_2^2(z) &= G_2(1, z, 1) \cong \mu_0 + \mu_1 + \mu_3 + \mu_{11} + \mu_{33} + \mu_{13} + (\mu_2 + \mu_{12} + \mu_{23})z \\
&+ \mu_{22}z^2 = 1.01645 - 0.82534z + 0.809z^2 \quad (3.2-41)
\end{aligned}$$

Hence, $E[X_2^2] \cong \mu_2 + \mu_{12} + \mu_{23} + 2\mu_{22} = 0.79266$

Again, by the symmetry assumption, $E[X_i^i] = E[X_2^2]$, $i = 1, 3$.

Now, using (3.2-41), $E[X_2^2(X_2^2 - 1)] \cong 2\mu_{22} = 1.618$, we get, $E[(X_2^2)^2] \cong 2.41066$.

To calculate the PGF of L_2 , the following are needed:

$$\begin{aligned}
m_i &= \frac{1 - \tilde{B}_i(\gamma_i)}{\gamma_i} = \frac{1}{5}; \quad \rho = \frac{3}{5}; \quad \widehat{X}^i_i(\tilde{\phi}_i(\gamma_i)) = 0.85 \text{ (using (3.2-12));} \\
\delta_i &= 2 - z; \quad \tilde{B}_i(\delta_i) = \frac{4}{6 - z}.
\end{aligned}$$

Now, substituting (3.2-41) in (3.2-25) and using the above, we get

$$\begin{aligned}
\widehat{L}_2(z) &\cong \frac{[(1.01645 - 0.82534z + 0.809z^2) - 0.85][1 + (1 - z)\frac{4}{6 - z}]}{(2 - z)\left(z - \frac{4}{6 - z}\right)0.75} = \frac{5.39z^2 - 5.5z + 1.11}{-z^2 + 6z - 4} \quad (3.2-42)
\end{aligned}$$

Differentiating at $z = 1$, leads to $E[L_2] \cong 1.2815$.

Finally,

$$\widetilde{W}_2(s) = \widetilde{L}_2\left(1 - \frac{s}{\lambda_2}\right) \cong \frac{20}{3} \cdot \frac{0.809s^2 - 0.7927s + 0.15}{-s^2 - 4s + 1}$$

with $E[W_2] \cong 1.2815$ and $E[W_{q_2}] \cong E[W_2] - m_2 = 1.0815$.

Conclusion

Applying the traditional station occupancy approach in studying polling systems subject to job failures or station breakdowns may lead to situations where the analysis can't be completed due to the lack of knowledge of the *PGF* $\widehat{X}_i^i(z)$ of the key variable, X_i^i - number of customers present in queue i when polled. To overcome this difficulty we use interpolation theory, never before utilized in the analysis of polling systems. The new approach enables one to obtain a polynomial approximation whose accuracy increases with the increase of the order of the polynomial function. We use the method to approximate the unknown *PGF* $\widehat{X}_i^i(z)$. Substituting the approximated value of $\widehat{X}_i^i(z)$ in the various derived formulae for the *PGF* of the queue size and the *LST* of the waiting time ($\widehat{L}_i(z)$ and $\widehat{W}_{q_i}(s)$, respectively) leads to an approximated closed-form formulae for those performance measure functions, and result in a complete solution of the models.

In this work we have restricted ourselves to the cases where failures or breakdowns may occur only when a station is visited by the cyclically moving server. In a related work (Shomrony and Yechiali [2005b]) we study the case where failures may occur at all stations, whether being attended by the server or not. This leads us to the notion of 'negative customers' (see e.g. Gelenbe [1991,1993] and Gelenbe and Fourneau [2002]), whose arrival 'kill' jobs queueing in the various stations. The analysis of such systems is complex and requires the use of transient probabilities of the system's state.

References

- Atkinson, K.E., (1978) An introduction to numerical analysis, Wiley
- Borst, S.C., (1995) Polling systems with multiple coupled servers. Queueing systems **20**:369-393.
- Boxma, O.J., Levy, H., Yechiali, U., (1992) Cyclic reservations schemes for efficient operation of multiple queue single server systems. Annals of Operations Research **35**:187-208.
- Boxma, O.J., Weststrate, J.W., Yechiali, U., (1993) A Globally Gated polling system with server interruptions and applications to the repairman problem. Probability in the engineering and informational sciences **7**:187-208.
- Eliazar, I., Yechiali, U., (1998) Randomly timed Gated queueing systems. SIAM journal on applied mathematics **59**:423-441.
- Fuhrmann, S.W., Cooper, R.B., (1985) Stochastic decomposition in the M/G/1 queue with generalized vacations. Operation Research **33**:1117-1129.
- Gelenbe, E., (1991) Queues with negative arrivals. Journal of applied probability **28**:245-250.

- Gelenbe, E., (1993) G-networks with triggered customer movement. *Journal of applied probability* **30**:742-748.
- Gelenbe, E., Fourneau, J.M., (2002) G-networks with resets. *Performance Evaluation* **49**:179-191.
- Kofman, D., Yechiali, U., (1996) Polling systems with station breakdowns. *Performance Evaluation* **27&28**:647-672.
- Kofman, D., Yechiali, U., (1997) Queueing networks with station breakdowns and Globally Gated service regime. In "Teletraffic contribution for the information age" (Ramaswamy, V. and Wirth, P.E., Eds.), North Holland, pp. 285-296.
- Nakdimon, O., Yechiali, U., (2003) Polling systems with breakdowns and repairs. *European journal of Operational Research* **149**:588-613.
- Shomrony, M., Yechiali, U., (2001) Burst arrival queues with server vacations and random timers. *Mathematical Methods of Operations Research* **53**:117-146.
- Shomrony, M., Yechiali, U., (2005a) Polling systems with job failures and with station failures, Technical report, Dept. of Statistics and Operations Research, Tel Aviv University, Israel.
- Shomrony, M., Yechiali, U., (2005b) Polling systems with positive and negative customers. Submitted for publication.
- Takagi, H., (1986) *Analysis of polling systems*, The MIT Press.
- Takagi, H., (1990) Queueing analysis of polling models: An update. In "Stochastic analysis of computer and communications systems" (Takagi, H., Ed.), North Holland, pp. 267-318.
- Yechiali, U., (1993) Analysis and control of polling systems. In "Performance Evaluation of Computer and Communication Systems" (L. Donatiello and R. Nelson, Eds.), Springer-Verlag, pp. 630-650.