A fluid queue driven by a Markovian queue

Bruno Sericola and Bruno Tuffin

IRISA-INRIA, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France
E-mail: {Bruno.Sericola;Bruno.Tuffin}@irisa.fr

Received 23 January 1998; revised 9 October 1998

We consider an infinite buffer fluid queue receiving its input from the output of a Markovian queue with finite or infinite waiting room. The input is characterized by a Markov modulated rate process. We derive a new approach for the computation of the stationary buffer content. This approach leads to a numerically stable algorithm for which the precision of the result can be given in advance.

Keywords: fluid queue, Markovian queue, Markov process

1. Introduction

In performance evaluation of telecommunication and computer systems, fluid queues models with Markov modulated input rates have been widely used in many papers, see among others [1,3,6,8,9]. The traffic arriving to a network queue has already traversed parts of the network and has been modified along its traversal. In such cases, it is the output from a queue which forms the input to the next network element.

In the most important part of the literature on this subject, see, for instance, [3,6] and the references therein, the state space of the Markov process that modulates the input rate in the fluid queue is supposed to be finite. The case where this state space is infinite has been analysed in [1,9] for the $M/M/1$ queue and in [8] for a birth and death process.

In this paper, we generalize the problem to an infinite buffer fluid queue driven by a Markovian queue. The only requirement needed on the Markov process that modulates the input and output rates is that it has a single state such that the input rate is smaller than the output rate of the fluid queue and that it has a uniform infinitesimal generator, that is, the suppremum of the output rates of the states is bounded. These Markov processes include not only the well-known $M/M/1/L$, $M/M/K/L$, $M/PH/1/L$ and $M/PH/K/L$ queues with finite ($L < \infty$) or infinite ($L = \infty$) waiting room but also the superposition of on–off sources with exponential off periods and phase-type on periods. Nevertheless, our method cannot be used if the Markov process that modulates the input and output rates in the fluid queue has more than one state with a negative effective input rate. This is the case if the Markovian queue is, for instance, the $PH/PH/K$ queue with non Poisson arrivals. In this queue the number of states corresponding to 0 customer in the queue is equal to the number of phases of the arrival process and this number is at least equal to 2 if we suppose that the arrival process in not a Poisson process.

The method used here to obtain the distribution of the stationary buffer content is neither based on spectral analysis nor on the use of Bessel functions as done in [1, 8,9], but a direct approach is used which leads to simple recursions. This method is particularly interesting due to the fact that it uses only additions and multiplications of positive numbers bounded by one. Thus we obtain a stable algorithm which, moreover, gives the result with a precision that can be specified in advance.

The rest of the paper is organized as follows. In the next section, we present the model and we obtain the solution in terms of recurrence relations whose behavior is studied. In section 3 we present the algorithm and numerical illustrations are given in section 4.

2. Model and solution

We describe in this section a fluid model with an infinite buffer for which the input and output rates are controlled by a homogeneous Markov process $\{X_t,\ t \geqslant 0\}$ on the state space S with infinitesimal generator denoted by A and stationary probability distribution denoted by π.

Let r_i be the input rate and c_i be the output rate when the Markov process $\{X_t\}$ is in state i. We denote by θ_i the effective input rate of state i, that is, $\theta_i = r_i - c_i$. We suppose that for every $i \in S$ we have $\theta_i \neq 0$. It is shown in the appendix that the case where $\theta_i = 0$ for some i can be reduced to this one.

We assume in this paper that the state space S contains only one state with negative effective input rate. This state is denoted by 0 and thus we have $S = \{0\} \cup S^+$ with $\theta_0 < 0$ and $\theta_i > 0$ for $i \in S^+$. It is, moreover, assumed that $\inf\{\theta_i \mid \theta_i > 0\} > 0$.

We suppose that the stability condition is satisfied, that is

$$\rho = \frac{\sum_{i \in S} r_i \pi_i}{\sum_{i \in S} c_i \pi_i} < 1,$$

where ρ is the traffic intensity, so that the limiting behavior exists. We denote by X the stationary state of the Markov process $\{X_t\}$ and by Q the stationary amount of fluid in the buffer.

Let $F_j(x) = \Pr\{X = j, Q \leqslant x\}$. We then have the following differential equations, see, for instance, [3], for all $j \in S$

$$\theta_j \frac{\mathrm{d}F_j(x)}{\mathrm{d}x} = \sum_{i \in S} F_i(x) A(i,j), \tag{1}$$

with the initial condition given by $F_j(0) = 0$ for every $j \in S^+$. We then have $F_0(0) = \Pr\{Q = 0\}$. We assume that $\sup\{-A(i,i)\colon\ i \in S\}$ is finite and we denote by

P the transition probability matrix of the uniformized Markov chain [5] with respect to the uniformization rate λ which verifies $\lambda \geqslant \sup\{-A(i,i),\ i \in S\}$. The matrix P is then related to A by $P = I + A/\lambda$, where I denotes the identity matrix. The following result, giving the stationary probability that the fluid queue is empty, will be used in the sequel.

Lemma 1.

$$F_0(0) = \frac{\sum_{j\in S} \theta_j \pi_j}{\theta_0} = (1-\rho)\frac{\sum_{j\in S} c_j \pi_j}{-\theta_0}.$$

Proof. Consider equation (1). By integrating from 0 to ∞ and summing over index j, we get $\sum_{j\in S} \theta_j (F_j(\infty) - F_j(0)) = 0$. Now, since $F_j(\infty) = \pi_j$ and $F_j(0) = 0$ for $j \in S^+$, we obtain the first equality. The second equality follows immediately from the definition of ρ. □

The main result of this paper, which is the distribution of the pair (X, Q), is given by the following theorem.

Theorem 2. For every $j \in S$, we have

$$F_j(x) = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} b_j(n), \tag{2}$$

where $\theta = \inf\{\theta_i \mid \theta_i > 0\}$ and the coefficients $b_j(n)$ are given by the following recursive expression:

$$b_0(0) = F_0(0) \quad \text{and} \quad b_j(0) = 0 \quad \text{for } j \in S^+,$$

and for $n \geqslant 1$ and $j \in S$,

$$b_j(n) = \left(1 - \frac{\theta}{\theta_j}\right) b_j(n-1) + \frac{\theta}{\theta_j} \sum_{i\in S} b_i(n-1) P(i,j). \tag{3}$$

Proof. We replace $F_j(x)$ by expression (2) in equation (1). Thus,

$$\begin{aligned}
&\theta_j \mathrm{e}^{-\lambda x/\theta} \frac{\lambda}{\theta} \left[\sum_{n=1}^{\infty} \frac{(\lambda x/\theta)^{n-1}}{(n-1)!} b_j(n) - \sum_{n=0}^{\infty} \frac{(\lambda x/\theta)^n}{n!} b_j(n) \right] \\
&\quad = \mathrm{e}^{-\lambda x/\theta} \sum_{n=0}^{\infty} \frac{(\lambda x/\theta)^n}{n!} \sum_{i\in S} b_i(n) A(i,j),
\end{aligned}$$

which can be reduced to

$$\theta_j \frac{\lambda}{\theta} \sum_{n=0}^{\infty} \frac{(\lambda x/\theta)^n}{n!} \big(b_j(n+1) - b_j(n)\big) = \sum_{n=0}^{\infty} \frac{(\lambda x/\theta)^n}{n!} \sum_{i\in S} b_i(n) A(i,j).$$

We then have for every $n \geqslant 0$

$$\theta_j \frac{\lambda}{\theta}\big(b_j(n+1) - b_j(n)\big) = \sum_{i \in S} b_i(n) A(i,j).$$

Using $A = \lambda(P - I)$, we obtain relation (3).

For $x = 0$, we have $F_j(0) = b_j(0)$ for every $j \in S$ from equation (2), which completes the proof. □

We give now some properties of the numbers $b_j(n)$ which will be used in the next section in order to develop a precise and stable algorithm to compute the distribution of the buffer content.

Proposition 3. For every $n \geqslant 0$, we have

$$b_0(n) = F_0(0) + \frac{\sum_{j \in S^+} \theta_j b_j(n)}{-\theta_0}. \tag{4}$$

Proof. Consider relation (3). By multiplying both sides by θ_j and by summing over index j, we obtain for $n \geqslant 1$

$$\sum_{j \in S} \theta_j b_j(n) = \sum_{j \in S} \theta_j b_j(n-1).$$

It follows that for every $n \geqslant 0$ we have

$$\sum_{j \in S} \theta_j b_j(n) = \sum_{j \in S} \theta_j b_j(0) = \theta_0 F_0(0),$$

which is equivalent to relation (4). □

Proposition 4. For every $j \in S$ and $n \geqslant 0$, we have $0 \leqslant b_j(n) \leqslant \pi_j$.

Proof. We proceed by induction. By definition of $F_j(x)$, we have $0 \leqslant F_j(x) \leqslant \pi_j$ for every $x \geqslant 0$ and $j \in S$. Since $F_j(0) = b_j(0)$ for every $j \in S$, we have $0 \leqslant b_j(0) \leqslant \pi_j$. Suppose now that we have $0 \leqslant b_j(n-1) \leqslant \pi_j$.

For $j \in S^+$, we have $\theta/\theta_j \in (0,1)$, so we easily obtain from relation (3), by using the relation $\pi P = \pi$, that $0 \leqslant b_j(n) \leqslant \pi_j$.

For $j = 0$, since $\theta_0 < 0$, $\theta_j > 0$ and $b_j(n) \geqslant 0$ for $j \in S^+$, we obtain from relation (4) that $b_0(n) \geqslant 0$ and

$$\begin{aligned} b_0(n) &= F_0(0) + \frac{\sum_{j \in S^+} \theta_j b_j(n)}{-\theta_0} \leqslant F_0(0) + \frac{\sum_{j \in S^+} \theta_j \pi_j}{-\theta_0} \\ &= F_0(0) + \frac{\sum_{j \in S} \theta_j \pi_j}{-\theta_0} + \pi_0 = \pi_0 \quad \text{(from lemma 1)}, \end{aligned}$$

and the result follows. □

Proposition 5. For every $j \in S$, the sequence $b_j(n)$ increases and converges to π_j.

Proof. We have, for $j \in S^+$, $b_j(1) \geqslant 0 = b_j(0)$ and

$$b_0(1) = F_0(0) + \frac{\sum_{j \in S^+} \theta_j b_j(1)}{-\theta_0} \geqslant F_0(0) = b_0(0).$$

Moreover, from relations (3) and (4), we have for $j \in S^+$

$$\begin{aligned} &b_j(n+1) - b_j(n) \\ &\quad = \left(1 - \frac{\theta}{\theta_j}\right)\bigl(b_j(n) - b_j(n-1)\bigr) + \frac{\theta}{\theta_j} \sum_{i \in S} \bigl(b_i(n) - b_i(n-1)\bigr) P(i,j) \end{aligned}$$

and

$$b_0(n+1) - b_0(n) = \frac{\sum_{j \in S^+} \theta_j (b_j(n+1) - b_j(n))}{-\theta_0}.$$

Since, for $j \in S^+$, we have $\theta/\theta_j \in (0,1)$ and $\theta_0 < 0$ we deduce by induction that for every $j \in S$ the sequence $b_j(n)$ is increasing. It then converges by using proposition 4.

For every $j \in S$, we denote by l_j the limit of $b_j(n)$ when n goes to infinity. We then have

$$F_j(x) = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} b_j(n) \to l_j \quad \text{when } x \to \infty.$$

Thus, since $F_j(x) = \Pr\{X = j, Q \leqslant x\}$ tends to π_j when x tends to ∞, we have $l_j = \pi_j$. $\square$

3. Algorithmical aspects

We suppose in this section that the infinitesimal generator of the process X has the following block tridiagonal structure:

$$A = \begin{pmatrix} A_{0,0} & A_{0,1} & & & & & & \\ A_{1,0} & A_{1,1} & A_{1,2} & & & & & \\ & A_{2,1} & A_{2,2} & A_{2,3} & & & & \\ & & A_{3,2} & \cdot & \cdot & & & \\ & & & \cdot & \cdot & A_{k-1,k} & & \\ & & & & A_{k,k-1} & A_{k,k} & A_{k,k+1} & \\ & & & & & \cdot & \cdot & \cdot \end{pmatrix}$$

where $A_{0,0}$ is the output rate from state 0. Such a structure leads to the infinitesimal generators of Markovian queues such as the $M/M/1/L$, the $M/M/K/L$, the $M/PH/1/L$ and the $M/PH/K/L$ queues with finite ($L < \infty$) or infinite ($L = \infty$) waiting room [4].

To compute the probability distribution $\Pr\{Q \leqslant x\}$ of the buffer content we use relations (2), (3) and (4) together with propositions 4 and 5. Relation (3) is used only for $j \in S^+$, and for $j = 0$ we use relation (4). These relations are particularly interesting from a computational point of view. Indeed, the fact that only additions and multiplications of positive and bounded numbers are used in their recurrences is a very important property for what concerns the numerical stability of the computation. Propositions 4 and 5 will be used as a criterion to stop the computation in the case where the sequence of the $b_j(n)$ is close to its limit π_j.

We denote by n_i the dimension of the square matrix $A_{i,i}$. Note that $n_0 = 1$. The transition probability matrix of the uniformized Markov chain has the same block tridiagonal structure as the matrix A. The blocks of matrix P are denoted by $P_{i,j}$ and we have, since $P = I + A/\lambda$, $P_{i,i} = I + A_{i,i}/\lambda$ and $P_{i,j} = A_{i,j}/\lambda$ for $i \neq j$, where I is in this case the identity matrix of dimension n_i.

We also consider the infinite row vector containing the $b_j(n)$ for $j \in S$. This infinite row vector can be rearranged according to the structure of matrix P to be written as

$$\left(b^{[0]}(n), b^{[1]}(n), b^{[2]}(n), \ldots\right),$$

where $b^{[0]}(n)$ is the scalar $b_0(n)$ and for $j \geqslant 1$, $b^{[j]}(n)$ is a row vector of dimension n_j. This consists in rearranging the state space S as $S = \{0\} \cup S_1 \cup S_2 \cup \cdots$, where, for $j \geqslant 1$, S_j contains n_j states with the same effective input rate equal to θ_j. With this notation, relation (3) can be written, for $j \geqslant 1$ and $n \geqslant 1$ as

$$\begin{aligned} b^{[j]}(n) = {} & \left(1 - \frac{\theta}{\theta_j}\right) b^{[j]}(n-1) \\ & + \frac{\theta}{\theta_j}\left(b^{[j-1]}(n-1)P_{j-1,j} + b^{[j]}(n-1)P_{j,j} + b^{[j+1]}(n-1)P_{j+1,j}\right). \end{aligned} \tag{5}$$

Using this recursion, it can be easily checked that, since $b^{[j]}(0) = 0$ for $j \geqslant 1$, we have

$$b^{[j]}(n) = 0 \quad \text{for } n \geqslant 0 \text{ and } j \geqslant n+1.$$

Relation (4) can then be written as

$$b^{[0]}(n) = F_0(0) + \frac{\sum_{j=1}^{n} \theta_j b^{[j]}(n)\mathbf{1}}{-\theta_0}, \tag{6}$$

where $\mathbf{1}$ is a column vector with all the entries equal to 1, its dimension being given by the context. Denoting by $F(x)$ the probability distribution function of the buffer content Q, that is, $F(x) = \Pr\{Q \leqslant x\}$, we finally get

$$F(x) = \sum_{j \in S} F_j(x) = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} b(n), \tag{7}$$

where $b(n) = \sum_{j \in S} b_j(n) = \sum_{j=0}^{n} b^{[j]}(n)\mathbf{1}$.

From proposition 5 and from the dominated convergence theorem, we obtain that the sequence $b(n)$ is an increasing sequence that converges to 1 when n goes to infinity.

The computation of $F(x)$ can then be done as follows. For a given error tolerance ε, we define integer N as

$$N = \min\left\{ n \in \mathbb{N} \,\middle|\, \sum_{i=0}^{n} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^i}{i!} \geqslant 1 - \varepsilon \right\} \tag{8}$$

and we denote by $F(N, x)$ the sum of the $N + 1$ first terms of relation (7), that is,

$$F(N, x) = \sum_{n=0}^{N} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} b(n).$$

We then have

$$F(x) = F(N, x) + e(N),$$

where the rest $e(N)$ of the series satisfies

$$\begin{aligned} e(N) &= \sum_{n=N+1}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} b(n) \leqslant \sum_{n=N+1}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} \\ &= 1 - \sum_{n=0}^{N} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} \leqslant \varepsilon. \end{aligned}$$

We also consider integer N' defined by

$$N' = \min\bigl\{ n \in \mathbb{N} \mid b(n) \geqslant 1 - \varepsilon \bigr\}.$$

Since the sequence $b(n)$ is increasing and converges to 1, we have $b(n) \geqslant 1 - \varepsilon$ for every $n \geqslant N'$. So we get

$$\begin{aligned} F(x) &= F\bigl(N', x\bigr) + \sum_{n=N'+1}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} b(n) \\ &= F\bigl(N', x\bigr) + 1 - \sum_{n=0}^{N'} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} - e'\bigl(N'\bigr), \end{aligned}$$

where the rest $e'(N')$ satisfies

$$e'\bigl(N'\bigr) = \sum_{n=N'+1}^{\infty} \mathrm{e}^{-\lambda x/\theta} \frac{(\lambda x/\theta)^n}{n!} \bigl(1 - b(n)\bigr) \leqslant \varepsilon.$$

The integer N' is not known a priori so we will first compute the integer N and start the computation of $F(N, x)$. This computation will be then stopped in the case where $N' < N$. Note also that the integer N, defined in (8), is an increasing function of x, say $N(x)$. So if the function $F(x)$ has to be evaluated at M points, say $x_1 < \cdots < x_M$, we only need to evaluate the values of $b(n)$ for $n = 0, 1, \ldots, N(x_M)$ since these values are independent of the values of $x_1, \ldots, x_M$.

The pseudocode of the algorithm is given below.

Input: $x_1 < \cdots < x_M$, ε
Output: $\Pr\{Q \leqslant x_1\}, \ldots, \Pr\{Q \leqslant x_M\}$
Compute N from relation (8) with $x = x_M$
$N' = N$
$b^{[0]}(0) = F_0(0)$ computed using lemma 1
$n = 0$
while $[n < N']$ **do**
  $n = n + 1$
  for $j = 1$ **to** n **do** Compute $b^{[j]}(n)$ from relation (5) **endfor**
  Compute $b^{[0]}(n)$ from relation (6)
  $b(n) = \sum_{j=0}^{n} b^{[j]}(n)\mathbf{1}$
  if $(b(n) \geqslant 1 - \varepsilon)$ **then**
    $N' = n$
  endif
endwhile
if $(N' = N)$ **then**
  for $i = 1$ **to** M **do** Compute $F(N, x_i)$ **endfor**
else
  for $i = 1$ **to** M **do**
    Compute $1 - \displaystyle\sum_{n=0}^{N'} \mathrm{e}^{-\lambda x_i/\theta} \frac{(\lambda x_i/\theta)^n}{n!} + F\big(N', x_i\big)$
  endfor
endif

The method that we have developed leads to a simple algorithm which gives very accurate results with a high precision that can be specified in advance. It can be applied to a wide class of Markovian queues; the only requirement being that only one state of the Markov process, which modulates the input and output rates of the fluid queue, must have a negative effective input rate. These results generalize those obtained in [1,9] for the $M/M/1$ queue where the solution is obtained by means of an integral representation. In [8], the authors consider a fluid queue driven by a birth and death process. Their results are based on the study of polynomials and, as said by the authors, the main problem in concrete examples is to find a signed measure with respect to which these polynomials are orthogonal. Our method can also be applied to a birth and death process if it has only one state with a negative effective input rate and if it has a uniform infinitesimal generator.

4. Numerical results

We have shown that the algorithm described in the previous section applies to a large class of block tridiagonal infinitesimal generators A with a single state having a negative effective input rate. Such a structure for the infinitesimal generator includes the following Markovian systems:

- The $M/M/1$ queue with arrival rate β and service rate γ. Take $A_{0,0} = -\beta,\ A_{0,1} = \beta$ and for $i \geqslant 1,\ A_{i,i+1} = \beta,\ A_{i,i-1} = \gamma$ and so $A_{i,i} = -(\beta + \gamma)$.
- The $M/M/K$ queue with arrival rate β and service rate per server γ. Take $A_{0,0} = -\beta,\ A_{0,1} = \beta$ and for $i \geqslant 1,\ A_{i,i+1} = \beta,\ A_{i,i-1} = \min(i,K)\gamma$ and so $A_{i,i} = -(\beta + \min(i,K)\gamma)$.
- The $M/PH/1$ queue with arrival rate β and (α, T) as phase-type representation of the service time distribution [4]. In this case, we must take $A_{0,0} = -\beta,\ A_{0,1} = \alpha\beta,\ A_{1,0} = -T\mathbf{1}$, and for $i \geqslant 1,\ A_{i,i+1} = \beta I,\ A_{i,i} = T - \beta I,\ A_{i+1,i} = -\alpha T\mathbf{1}$.
- The $M/PH/K$ queue with arrival rate β and (α, T) as phase-type representation of the service time distribution per server. The blocks $A_{i,j}$ of its infinitesimal generator can be obtained using tensor algebra as done in [7].
- All these Markovian queues can also be considered when their waiting room is finite since in this case the infinitesimal generator A is a finite block tridiagonal matrix.
- The superposition of a finite number of independent on–off sources where the off periods are exponentially distributed and the on periods have a phase-type distribution.

In order to illustrate our algorithm, we consider the $M/M/K$ queue with arrival rate β and service rate γ per server. The input rate in the fluid queue when the $M/M/K$ queue is in state i is then given by $r_i = \min(i,K)r$ for every $i \geqslant 0$, where r is the input rate per server in the fluid queue. We suppose that the output rate of the fluid queue is constant being equal to c, that is, $c_i = c$ for every $i \geqslant 0$ and such that $r > c$. We then obtain that the effective input rate in the fluid queue is given by $\theta_i = \min(i,K)r - c$. We suppose that $\beta < K\gamma$ so that the π_i exists and that $\rho = \beta r/(\gamma c) < 1$, which implies that the limiting behavior of the buffer contents exists.

Figure 1 shows the complementary cumulative distribution function of the buffer content of a fluid queue driven by an $M/M/K$ for $K = 1$ and $K = 10$. In both cases, the arrival rate is $\beta = 0.8$, the service rate per server is $\gamma = 1$, the input rate per server in the fluid queue is $r = 1.2$ and the output rate is constant $c = 1$. In this example we have taken $\varepsilon = 10^{-5}$.

The same function, but for larger values of x, is shown in figure 2 for $\beta = 0.4,\ \gamma = 1,\ r = 2$ and $c = 1$. In this figure, the vertical axis is in logarithmic scale and we have taken $\varepsilon = 10^{-10}$.

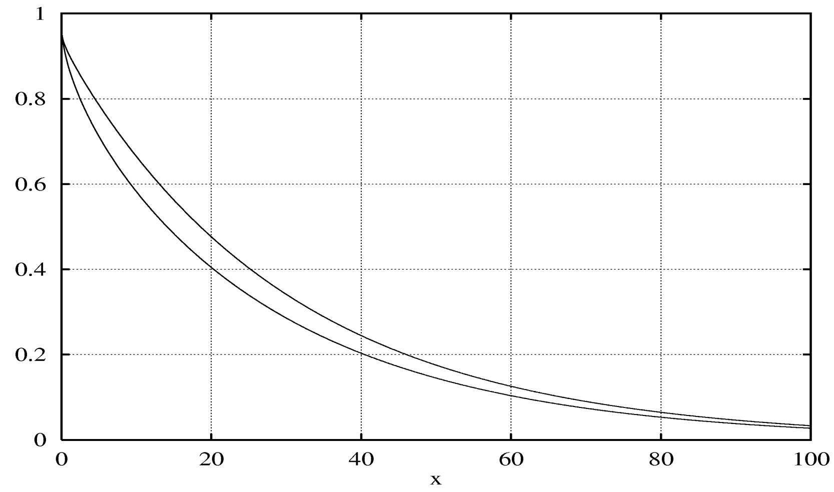

Figure 1. From top to the bottom: $\Pr\{Q > x\}$ versus x for the $M/M/10$ and the $M/M/1$ queues as input queues with arrival rate $\beta = 0.8$, service rate $\gamma = 1$ per server, input rate $r = 1.2$ per server and constant output rate $c = 1$, which gives $\rho = 0.96$.

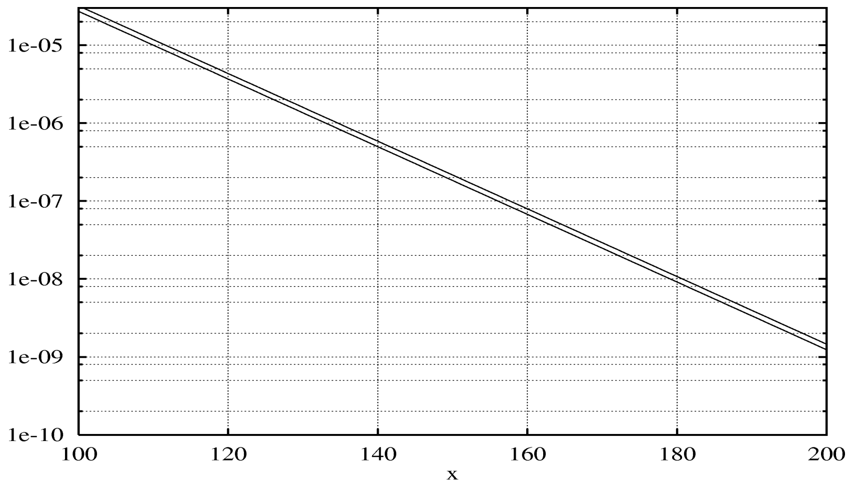

Figure 2. From top to the bottom: $\Pr\{Q > x\}$ versus x for the $M/M/10$ and the $M/M/1$ queues as input queues with arrival rate $\beta = 0.4$, service rate $\gamma = 1$ per server, input rate $r = 2$ per server and constant output rate $c = 1$, which gives $\rho = 0.8$.

Appendix

We consider here that the state space S^* of process $\{X_t\}$, which we suppose irreducible, contains a finite number of zero effective input rates.

We write $S^* = S \cup S^0$, where S (respectively S^0) is the set of states with nonzero (respectively zero) effective input rates.

The infinitesimal generator A^* of the process $\{X_t\}$ and the diagonal matrix D^* of the effective input rates can then be written in the obvious notation as

$$A^* = \begin{pmatrix} A_{SS} & A_{SS^0} \\ A_{S^0S} & A_{S^0S^0} \end{pmatrix} \quad \text{and} \quad D^* = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}.$$

In the same way, we denote by $F_S(x)$ and $F_{S^0}(x)$ the row vectors containing the $F_j(x)$ for $j \in S$ and $j \in S^0$, respectively.

The differential equations (1) can then be written as

$$\begin{aligned} \frac{\mathrm{d}F_S(x)}{\mathrm{d}x} D &= F_S(x)A_{SS} + F_{S^0}(x)A_{S^0S}, \\ 0 &= F_S(x)A_{SS^0} + F_{S^0}(x)A_{S^0S^0}. \end{aligned} \tag{9}$$

As A^* is irreducible, $-A_{S^0S^0}$ is a nonsingular M-matrix [2], so $A_{S^0S^0}$ is invertible. Let $(\pi_i^*)_{i \in S^*}$ be the stationary distribution of $\{X_t\}$. We have:

Proposition.

$$\begin{aligned} F_{S^0}(x) &= -F_S(x)A_{SS^0}A_{S^0S^0}^{-1}, \\ \frac{\mathrm{d}F_S(x)}{\mathrm{d}x} D &= F_S(x)A, \end{aligned} \tag{10}$$

where

$$A = A_{SS} - A_{SS^0}A_{S^0S^0}^{-1}A_{S^0S}.$$

The results given by theorem 2 can then be used to obtain the solution in the following way: from the solution $G(x)$ of section 2 for $(\mathrm{d}G/\mathrm{d}x)(x)D = G(x)A$ we obtain $F_S(x) = (\sum_{i \in S} \pi_i^*)G(x)$ and then $F_{S^0}(x)$ is given from (10).

Proof. Equations (10) follow immediately from (9). It is well known that A is an infinitesimal generator and that the stationary probability measure $\pi_S = (\pi_i)_{i \in S}$ of the Markov process with infinitesimal generator A is given for every $i \in S$ by

$$\pi_i = \frac{\pi_i^*}{\sum_{j \in S} \pi_j^*}.$$

Section 2 then gives for equation

$$\frac{\mathrm{d}G}{\mathrm{d}x}(x)D = G(x)A$$

a solution $G(x)$ which tends to π_S as $x \to +\infty$. Given that the solution of this equation is unique up to a multiplicative constant, and given that $F_S(x)$ tends to $\pi_S^* = (\pi_i^*)_{i \in S}$ as $x \to +\infty$, we obtain

$$F_S(x) = G(x) \sum_{j \in S} \pi_j^*. \qquad \square$$

References

[1] I. Adan and J. Resing, Simple analysis of a fluid queue driven by an $M/M/1$ queue, Queueing Systems 22 (1996) 171–174.
[2] A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences* (Academic Press, New York, 1979).
[3] D. Mitra, Stochastic theory of a fluid model of producers and consumers coupled by a buffer, Adv. in Appl. Probab. 20 (1988) 646–676.
[4] M.F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach* (Johns Hopkins Univ. Press, Baltimore/London, 1981).
[5] S.M. Ross, *Stochastic Processes* (Wiley, New York, 1983).
[6] T.E. Stern and A.I. Elwalid, Analysis of separable Markov-modulated rate models for information-handling systems, Adv. in Appl. Probab. 23 (1991) 105–139.
[7] Y. Takahashi, Asymptotic exponentiality of the tail of the waiting-time distribution in a $PH/PH/c$ queue, Adv. in Appl. Probab. 13 (1981) 619–630.
[8] E.A. van Doorn and W.R.W. Scheinhardt, A fluid queue driven by an infinite state birth-death process, in: *Proc. ITC '15*, Washington (June 1997).
[9] J. Virtamo and I. Norros, Fluid queue driven by an $M/M/1$ queue, Queueing Systems 16 (1994) 373–386.