PHYSICAL AND MATHEMATICAL QUEUES IN THE APPLIED QUEUING THEORY

Alexander Kirpichnikov¹, Anton Titovtsev² §
¹,²Kazan National Research Technological University
68, K. Marksa, Kazan, 420015, RUSSIA

Abstract: The problem of various types of queues arising in queuing systems is considered in the paper. The concept of a physical queue is introduced; universal mathematical formulas of the first and second moments of main discrete and continuous random variables characterizing behavior of physical queues for various models of mixed type queuing systems are obtained.

AMS Subject Classification: 60K25
Key Words: physical queue, queue, quality of service (QoS), queuing system (QS)

1. Main Definitions

In practice, when different technical facilities working according to the principle of queuing systems (QS) are employed, very often one should deal with the habitual and well-studied mathematical queue and with other types of queues as well. In particular, it is possible to consider another numerical characteristic of QS which, for example, in the GPSS standard report of simulation modeling system has the name ”a queue without zero inputs”.

In this case, zero input is understood as such arrival of the claim to the system at which there is, at least, one free server in the multichannel facility; and in this case the claim newly arrived in the system is served immediately. Thus, unlike a usual mathematical queue, ”the queue without zero inputs” means such a situation when at the time of a new claim arrival in the system all servers can be engaged in the service facility, but at the same time the queue, per se, can
be absent. In the latter case, an arrived claim does not meet other claims, and is just in front of the service facility in which all service channels are engaged.

The "queue without zero inputs" defined as such has to be calculated considering only those claims which really expected the beginning of service and without claims which did not have to wait as at the time of their arrival in the system at least one service channel was free. The mean length of the queue without zero inputs has apparently to be more than the mean length of a well-known and more habitual mathematical queue. At the same time the minimum mean length of such queue is zero as well as general mathematical queue is, i.e. on average any number of claims can be found both in such a queue and in a mathematical queue as well.

Thus, if the general mathematical queue is calculated as mean on all claims that visited the system, then the queue without zero inputs has to be calculated as an average quantity minus those demands which were served immediately as they got into the system while, at least, one of service channel was free.

According to the nomenclature accepted in GPSS simulation modeling system we will call a zero input as such an arrival of the claim into the system when there is, at least, one free service channels, and the claim is served immediately in this case.

Let's call the queue calculated without zero inputs a physical queue. In case when all service channels are employed the claim newly arrived into the system will have to wait for the service in the memory, at the same time the minimum quantity of claims in the memory is equal to zero.

As well as the habitual mathematical queue calculated with zero inputs taken into account, a physical queue is also characterized by servicing-waiting moments of delayed demands, and, accordingly, by servicing-waiting moments of a newly arrived claim in the corresponding queue. The idea of these moments calculation is that in this case when queue parameters are calculated it is necessary to redefine the normalization condition considering only those QS states at which zero inputs are absent. In other words, only those states of the system which have a physical queue should be regarded upon calculation.

Thus, the calculation algorithm of physical queue parameters is completely identical to the calculation algorithm of a mathematical queue for an exception of probabilities of possible QS states which have to be calculated on the basis of the new normalization condition excluding states corresponding to zero inputs.

The paper suggests universal mathematical expressions for first and second moments of demands number expecting service in a physical queue within the most common queuing system model with the arbitrary quantity of sources and size-limited queue [5]. Besides, corresponding distribution functions are
calculated, and as a result, the first and second moments of waiting time to start claim service in a queue of the specified type are also calculated.

2. Probabilistic Characteristics and Moments of Demands Number in a Physical Queue

The work of authors [5] presented the most general mathematical model of an open multi-channel QS having \( m \) service units of identical efficiency with exponentially distributed service time. The input demand stream in this case is a superposition of the arbitrary number of \( h \) components each of which represents a Poisson stream of claims served in the queue. For each type of demands arriving in the system from \( j \)-th source limitation for a queue length \( \varepsilon_j \) acts, at the same time \( \varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_h \).

We will also use designations accepted in work [5] in the present paper:

\[
\varepsilon_0 = E_0 = 0; \quad \varepsilon_1 = E_1; \quad \varepsilon_2 = E_1 + E_2; \quad \cdots \varepsilon_j = \sum_{i=0}^{j} E_i = \sum_{i=1}^{j} E_i; \quad \text{a size-limited queue (memory volume) for claims of the } j\text{-th component;}
\]

\[
\Lambda_0 = \sum_{j=0}^{h} \lambda_j; \quad \Lambda_1 = \sum_{j=1}^{h} \lambda_j; \quad \Lambda_2 = \sum_{j=2}^{h} \lambda_j; \quad \cdots \Lambda_h = \lambda_h; \quad \text{where } \lambda_j \text{ claim stream intensity of the } j\text{-th component; } \mu \text{ claims service intensity by one service facility.}
\]

\[
R_0 = \sum_{j=0}^{h} \rho_j; \quad R_1 = \sum_{j=1}^{h} \rho_j; \quad R_2 = \sum_{j=2}^{h} \rho_j; \quad \cdots \quad R_h = \rho_h; \quad R_i = \frac{\Lambda_i}{\mu}, \quad \text{where } \rho_j \text{ is the given claim stream intensity of the } j\text{-th component.}
\]

To calculate parameters of a physical queue on the basis of the developed and presented in work [5] universal mathematical model it is necessary to rewrite a normalization condition in such a way as upon totaling in this condition to consider only those states of the system which correspond to physical queue. As a result, we have

\[
\sum_{i=m}^{m+\varepsilon_h} P_i = 1 \quad (1)
\]

– normalization condition to calculate a physical queue.

Substituting corresponding expressions for probabilities of steady states \( P_i \) in the normalization condition presented above, and solving the obtained equation in regard to \( P_0 \), we will obtain the following expression of \( \tilde{P}_0 \) quantity used for a physical queue calculation:
$\tilde{P}_0 = \left[ \frac{R_0^m}{m!} + \frac{R_0^m}{m!} \sum_{g=1}^{h} \prod_{j=0}^{g-1} \left( \frac{R_j}{m} \right)^{E_j} \times \right. \\
\times \left\{ \frac{R_g}{m-R_g} \left( 1 - \frac{R_g}{m} \right)^{E_g}, \quad R_g \neq m \right\}^{-1}$ ;

The formulas for other probabilistic characteristics intended for physical queue parameters calculation will look the same as when characteristics of the usual mathematical queue with replacement $P_0$ for $\tilde{P}_0$ are calculated. We have

\[ \tilde{P}_i = \left( \frac{R_{j+1}}{m} \right)^{i-m-\varepsilon_j} \prod_{g=0}^{i} \left( \frac{R_g}{m} \right)^{E_g} \frac{R_0^m}{m!} \tilde{P}_0, \]

\[ m + \varepsilon_j \leq i \leq m + \varepsilon_{j+1}, \quad 0 \leq j \leq h - 1 \quad (2) \]

– the probabilities of various QS states in a steady state mode;

\[ \tilde{P}_{Bi} = \prod_{g=0}^{i-1} \left( \frac{R_g}{m} \right)^{E_g} \frac{R_0^m}{m!} \tilde{P}_0 \left\{ \frac{m}{m-R_i} \left( 1 - \frac{R_i}{m} \right)^{E_i}, \quad R_i \neq m \right\} \quad (3) \]

– the basic probability for QS characteristics calculation;

\[ \tilde{P}_{m+\varepsilon_h} = \prod_{g=1}^{h} \left( \frac{R_g}{m} \right)^{E_g} \frac{R_0^m}{m!} \tilde{P}_0 \quad (4) \]

– congestion probability of the system;

\[ \tilde{P}_W = \frac{1}{R_0} \sum_{i=1}^{h} R_i \tilde{P}_{Bi} \]

– probability of a newly arrived claim service expectation in the physical queue.

In this case the relative throughout capacity of the system is $\tilde{q} = \tilde{P}_W$, absolute throughout capacity is apparently $\tilde{A} = \Lambda_0 \tilde{q}$.

Let’s emphasize that in this case $\tilde{P}_0$ does not represent the probability of a zero state of the system $P_0$ any more, i.e. the probability of total claims absence in the system, which is not necessary in this calculation, but some coefficient the value of which can be even larger than unity. At the same time they have
a clear physical sense; only those probabilities of the system stationary states which correspond to states \( i \geq m \) are necessary for calculations. The values of these probabilities calculated according to formulas (2) – (5) are, of course, quantities which are smaller than unity, anyway.

The moments of claims number expecting service in a physical queue are calculated according to the same scheme as for a mathematical queue; however, the probabilities with a tilde obtained with the corresponding normalization condition in view are inserted into corresponding formulas. A detailed conclusion of formulas for the moments of queue length is given in work [5]. As a result, we have

- average number of claims in a physical queue

\[
\bar{l} = \sum_{i=1}^{h} \left\{ \frac{R_i}{m-R_i} \left[ \tilde{P}_{Bi} - E_i \tilde{P}_{m+i_1} \right], \quad R_i \neq m \right\} + 
\end{equation}

\[
+ \sum_{i=2}^{h} \varepsilon_i \frac{R_i}{m} \tilde{P}_{Bi};
\]

- the second initial moment of demands number in a physical queue

\[
\bar{l}^2 = \sum_{i=1}^{h} \left[ \varepsilon_i \tilde{P}_{Bi} + \left\{ \frac{R_i}{m-R_i} \left( \frac{m+R_i}{m-R_i} + 2\varepsilon_i \right) \tilde{P}_{Bi} - \frac{mE_i}{m-R_i} \left( E_i + \frac{2R_i}{m-R_i} + 2\varepsilon_i \right) \tilde{P}_{m+i_1}, \quad R_i \neq m \right\} + \varepsilon_i^2 \tilde{P}_{m+i_2} \right]\]

Let’s also give options of these dependences for several specific cases of particular practical interest.

For the combined model of queuing presented in work [4] we have

\[
\tilde{P}_0 = \left[ \frac{R_0^m}{m!} \left( 1 + \left\{ \frac{R_0^m}{m!} \left[ 1 - \frac{(R_0^m)}{E_1} E_1 \right], \quad R_1 \neq m \right\} + \right\} + \right]\]

\[
\left. + \left( \frac{R_1}{m} \frac{E_1}{m-\rho} \right) \right]^{-1};
\]

at the same time, the probabilities of steady states are

\[
\tilde{P}_i = \left\{ \begin{array}{ll}
\frac{(R_i)}{m}^{i-m} \frac{R_0^m}{m!} \tilde{P}_0, & m \leq i \leq m + E_1 \\
\frac{(\rho)}{m}^{i-m-E_1} \frac{(R_i)}{m} \frac{E_1}{m} \frac{R_0^m}{m!} \tilde{P}_0, & i \geq m + E_1
\end{array} \right.
\]
In this model the following system of designations is accepted:

\[ \rho_0 = \frac{\lambda_0}{\mu}; \quad \rho_1 = \frac{\lambda_1}{\mu}; \quad \rho = \frac{\lambda}{\mu}; \]

\[ R_0 = \frac{\Lambda_0}{\mu} = \rho_0 + \rho_1 + \rho; \quad R_1 = \frac{\Lambda_1}{\mu} = \rho_1 + \rho. \]

In this case the first and second moments of claims number expecting service in a physical queue have the form

\[ \bar{I} = \left\{ \frac{R_1}{m-R_1} \left( \tilde{P}_{B1} - E_1 \tilde{P}_{m+E_1} \right), R_1 \neq m \right\} + \]

\[ + \rho \left( \frac{E_1}{m} + \frac{1}{m-\rho} \right) \tilde{P}_{MFL}; \]

\[ \bar{I}^2 = \left\{ \frac{R_1}{m-R_1} \left[ \frac{m+R_1}{m-R_1} \tilde{P}_{B1} - E_1 \left( E_1 + \frac{2m}{m-R_1} \right) \tilde{P}_{m+E_1} \right), R_1 \neq m \right\} + \]

\[ + \rho \left[ \frac{E_1^2}{m} + \frac{1}{m-\rho} \left( 2E_1 + \frac{m+\rho}{m-\rho} \right) \right] \tilde{P}_{MFL}, \]

Where \( \tilde{P}_{MFL} = \sum_{i=m+E_1}^{\infty} \tilde{P}_i \) probability of the memory full loading, \( \rho \) the given claim stream intensity which are not influenced by size-limited queue.

For a specific QS model with full set of memories proposed in work [6], the probabilities of steady states of system upon physical queue calculation take the form

\[ \tilde{P}_0 = \left[ \frac{R_0^m}{m!} + \frac{R_0^m}{m!} \sum_{g=1}^{h} \prod_{j=1}^{g} \frac{R_j}{m} \right]^{-1}; \]

\[ \tilde{P}_i = \prod_{g=1}^{i-m} \frac{R_g R_0^m}{m! \tilde{P}_0}, \quad m + 1 \leq i \leq m + h. \]

and then, moments of claims number expecting service in a physical queue are

\[ \bar{I}^{(K)} = \sum_{i=m+1}^{m+h} (i-m)^K \prod_{g=1}^{i-m} \frac{R_g R_0^m}{m! \tilde{P}_0}, \quad K = 1, 2. \]
For H. Takagi model [7] steady states of system probabilities for calculation of the physical queue have the form

\[
P_0 = \left( \frac{R_0}{m!} + \frac{R_m}{m!} \left\{ \frac{R_1}{m-R_1} \left[ 1 - \frac{(R_1)}{m} \right]^E_1, \quad R_1 \neq m \right\} \right)^{-1};
\]

\[
P_i = \left( \frac{R_1}{m} \right)^{i-m} \frac{R_0^m}{m!} \tilde{P}_0, \quad m \leq i \leq m + E_1
\]

and moments of claims number expecting service in a physical queue

\[
\tilde{l} = \left\{ \frac{R_1}{m-R_1} \frac{\tilde{P}_{B1}^m - \tilde{E}_1 \tilde{P}_{m+E1}}{E_1+1}, \quad R_1 \neq m \right\};
\]

\[
\tilde{l}^2 = \left\{ \frac{R_1}{m-R_1} \frac{m+R_1}{(E_1+1)(m+R_1)} \tilde{P}_{B1}^m - \tilde{E}_1 \left( \frac{1}{E_1+1} \frac{m+R_1}{m-R_1} \tilde{P}_{m+E1}^m \right), R_1 \neq m \right\}.
\]

Just the same system of formulas will be, apparently, actual for a one-component QS model with the queue of finite length (according to J. Kendall’s classification – M/M/m/E model). In this case in the above-stated formulas it is necessary to put only \( R_0 = R_1 = \rho_1 \) hence according to [2], [3] we have

\[
\tilde{l} = \left\{ \frac{\rho_1}{m-\rho_1} \frac{1 - \tilde{E}_1 \tilde{P}_{L}}{\tilde{P}_W}, \quad \rho_1 \neq m \right\};
\]

\[
\tilde{l}^2 = \left\{ \frac{\rho_1}{(m-\rho_1)^2} - \frac{\rho_1}{E_1(\rho_1)} \left( \frac{1}{E_1+1} \frac{m}{m-\rho_1} \right) \frac{\tilde{P}_{L}}{\tilde{P}_W}, \quad \rho_1 \neq m \right\}.
\]

In these formulas

\[
\tilde{P}_W = \left\{ \frac{\tilde{P}_0}{(m-\rho_1)!} \left[ 1 - \rho_1 \right]^E_1, \quad \rho_1 \neq m \right\}
\]

\[
\tilde{P}_L = \left\{ \frac{\rho_1}{m!} \frac{1}{m-E_1} \tilde{P}_0, \quad \rho_1 \neq m \right\}
\]

- the probability of claim service expectation, i.e. the probability that an arriving demand will find all servers engaged.

\[
\tilde{P}_L = \left\{ \frac{\rho_1}{m!} \tilde{P}_0, \quad \rho_1 = m \right\}
\]

- the probability of refusal which the claim receives if it comes at the moment when all \( m \) service channels and all \( E_1 \) places in the queue are engaged.
Finally, for a specific model studied in J. Cohen’s work [1] for the first time, we have the following steady states probabilities of the system for physical queue parameters calculation:

\[
\tilde{P}_0 = \left[ \frac{R_0^m}{m!} \frac{m}{m - \rho} \right]^{-1};
\]

(6)

In this case the first and second moments of claims number expecting service in a physical queue are

\[
\tilde{l} = \frac{\rho}{m - \rho} \tilde{P}_{MFL}; \quad \tilde{l}^2 = \frac{\rho (m + \rho)}{(m - \rho)^2} \tilde{P}_{MFL}.
\]

For a simpler one-component model with an unlimited queue (M/M/m model) in formulas (6) \( R_0 = \rho \) should be accepted, and in this case we have

\[
\tilde{l} = \frac{\rho}{m - \rho}; \quad \tilde{l}^2 = \frac{\rho (m + \rho)}{(m - \rho)^2}.
\]

3. The Claim Servicing Waiting in the Physical Queue

The algorithm to calculate servicing-waiting moments of the claim in the physical queue is quite identical to the algorithm to calculate servicing-waiting moments of the claim in a general mathematical queue presented in work [5]. Corresponding probabilities with a tilde obtained with a new normalization condition in view have to be inserted instead of common probability characteristics (1). Besides, taking into account this circumstance it is necessary to redefine the relative throughout capacity of \( \tilde{q} \) system. In this case, we have the following finite formulas for the claim servicing-waiting moments in the physical queue for the QS universal model presented in work [5]:

\[
\overline{t_W} = \frac{1}{\Lambda_0 \tilde{q}} \sum_{i=1}^{h} \left\{ \frac{R_i}{m - R_i} \left[ \tilde{P}_{Bi} - E_i \tilde{P}_{m + \varepsilon_i} \right] + \frac{R_i}{m} \varepsilon_{i-1} \tilde{P}_{Bi} \right\} = \frac{\tilde{l}}{A} ;
\]

\[
\overline{t_W^2} = \frac{1}{\Lambda_0 \tilde{q}} \sum_{i=1}^{h} R_i \left\{ \frac{2(\tilde{P}_{Bi} - E_i \tilde{P}_{m + \varepsilon_i})}{\mu (m - R_i)^2} \left[ 1 + \frac{\varepsilon_{i-1}}{m} (m - R_i) \right] + \frac{\tilde{P}_{m + \varepsilon_i}}{3m^2 \mu} \prod_{g=0}^{i-1} \left( \frac{R_g}{R_i} \right)^{E_g} \times \right\}
\]

\[ + \varepsilon_{i-1}(\varepsilon_{i-1} + 1) \frac{\hat{P}_{Bi}}{m^2 \mu} - \frac{E_i(E_{i+1}) \hat{P}_{m+B} + \varepsilon_i}{m\mu(m-R_i)}, \quad R_i \neq m \}
\times \left[ \varepsilon_i(\varepsilon_i + 1)(\varepsilon_i + 2) - \varepsilon_{i-1}(\varepsilon_{i-1} + 1)(\varepsilon_{i-1} + 2) \right], \quad R_i = m
\]

For the combined model of queuing given in work [4] we have

\[ \overline{t_W} = \frac{1}{A} \left[ \frac{R_1}{m-R_1} \left( \hat{P}_{B1} - E_1 \hat{P}_{m+E_1} \right) + \frac{\rho}{m} \hat{P}_{MFL} \left( E_1 + \frac{m}{m-\rho} \right) \right] \]

\[ = \frac{\tilde{l}}{A}; \]

\[ \overline{t^2_W} = \frac{1}{A \tilde{q}} \left\{ \frac{1}{\mu(m-R_1)} \left[ 2 \frac{R_1}{m-R_1} \left( \hat{P}_{B1} - E_1 \hat{P}_{m+E_1} \right) - E_1 (E_1 + 1) \hat{P}_{m+E_1} \right], \quad R_1 \neq m \right\} + \]

\[ + \frac{1}{\mu(m-\rho)} \left[ 2 \frac{\rho}{m} \hat{P}_{MFL} \left( E_1 + \frac{m}{m-\rho} \right) + E_1 (E_1 + 1) \hat{P}_{m+E_1} \right] \] \]

For the specific QS model with a full set of memories [6], we have

\[ \overline{t_W} = \frac{\tilde{l}}{A}; \quad \overline{t^2_W} = \frac{1}{A \mu m} \sum_{i=m+1}^{m+h} (i-m)(i-m+1) \hat{P}_i. \]

For a Takagi model [7]:

\[ \overline{t_W} = \frac{1}{\Lambda_0 \tilde{q}} \frac{R_1}{m-R_1} \left( \hat{P}_{B1} - E_1 \hat{P}_{m+E_1} \right) = \frac{\tilde{l}}{A}; \]

\[ \overline{t^2_W} = \frac{1}{A} \left\{ \frac{1}{\mu(m-R_1)} \left[ 2 \frac{R_1}{m-R_1} \left( \hat{P}_{B1} - E_1 \hat{P}_{m+E_1} \right) - E_1 (E_1 + 1) \hat{P}_{m+E_1} \right], \quad R_1 \neq m \right\} + \]

\[ + \frac{1}{\mu \mu m} E_1 (E_1 + 1) \hat{P}_{m+E_1}. \]

The same formulas will be fair for the classical M/M/m/E model as well if we put \( R_0 = R_1 = \rho_1 \) in them, in this case we have

\[ \overline{t_W} = \frac{\tilde{l}}{A} = \left\{ \frac{\rho_1 (\hat{P}_W - E_1 \hat{P}_L)}{E_{1+1}^{m-\rho_1}}, \quad \rho_1 \neq m \right\}; \]

\[ \frac{E_{1+1}^{m-\rho_1}}{2\lambda}, \quad \rho_1 = m \]
- variance of a claim waiting time in a physical queue

\[ \tilde{\sigma}_{W}^{2} = \begin{cases} \frac{(E_1 + 1)(E_1 + 2)}{3 \lambda_1 A} \tilde{P}_W - \bar{t}_W^2, & \rho_1 \neq m \\ \frac{(E_1 + 1)(E_1 + 5)}{12 \lambda_1^2}, & \rho_1 = m \end{cases} \]

And finally, for J. Cohen’s model [1]:

\[ \bar{t}_W = \frac{1}{\Lambda_0 \tilde{q}} \frac{\rho}{m - \rho} \tilde{P}_{MFL} = \frac{\bar{l}}{\tilde{A}}; \]

\[ \bar{t}_W^2 = \frac{2 \rho \tilde{P}_{MFL}}{\tilde{A} \mu (m - \rho)^2}. \]

For the simplest one-component model with an unlimited queue (M/M/m model) provided \( R_0 = \rho \) it apparently follows

\[ \bar{t}_W = \frac{\rho}{\lambda (m - \rho)} = \frac{\bar{l}}{\lambda}; \]

\[ \tilde{\sigma}_{W}^{2} = \frac{2}{\mu^2 (m - \rho)^2}. \]

It should be noted that results obtained according to this algorithm completely coincide with results of QS simulation modeling of the above-stated types in the system of GPSS World simulation modeling.

References


