# Introduction to Queueing Theory

**Prof. N. Selvaraju**
**Department of Mathematics**
**Indian Institute of Technology Guwahati, India**

## Lecture - 47
## M/G/1 Queues with Vacations

Hi and hello, everyone. What we have been seeing was that this queuing systems or queuing models with vacations. We have just introduced certain what is a vacation queuing model, and we have seen the $M/M/1$ vacation model with multiple vacations and how we can modify it to a single vacation model. What we will see next is the little bit more generic one, which is basically $M/G/1$ queues with vacation. The reason is we are not looking at in much generality $G/G/1$ because then the decomposition is what we want to see exactly. We did not see it in $M/M/1$ because we are going to see it in here, like in the case of $M/G/1$, because with Poisson input, the decomposition holds both for queue length as well as waiting time. In the general $G/G/1$ model, it holds only for waiting time; that is the idea that we need to keep in mind. So, since we are taking $M/G/1$, which is the Poisson input model, the queue length distribution level itself, we will find the decomposition. So, we will now discuss this $M/G/1$ model with single and multiple vacations under exhaustive service discipline.

As usual, we will assume that the vacation sequence $\{v_n\}$ is stationary and the system is steady-state.
Now, let us introduce a random variable $N^*$ be the number of customers present at the start of a busy period meaning the normal period following a vacation or vacation period, a vacation or a vacation period because if it goes from multiple vacation and it will be after multiple vacation period. So, whenever the busy period starts, that means when the system becomes empty, the server will go for vacation. After returning from vacation, it will either start serving the customer if there is at least one customer or wait in case of a single vacation period. During multiple vacation period, it will go for another vacation. So, whenever he starts the service, the number of customers present at the start of that busy period is what we call it $N^*$. So, $N^*$ has to be at least 1, I mean $N^* \geq 1$, and $N^*$ can be deterministic or a random variable; both is possible.

- First, consider that $N^*$ is an RV having PGF

$$R(z) = \sum_{n=1}^{\infty} \mathrm{P}\left\{N^* = n\right\} z^n, \quad |z| < 1.$$

- Let $P(z)$ be the PGF of the number in the system at a departure epoch of a usual $M/G/1$ queue without vacation. It is a normal $M/G/1$ queue without vacation is what the PGF of the departure epoch, but we know in such a queue the distribution of the number in the system at the random epoch or at the arrival epoch, or at the departure epochs, they are all one and the same because of the Poisson input queue. So, this $P(z)$ is given by the PK transformed formula or Pollaczek Khinchin transform formula. We know all those things; that is what we take it from $M/G/1$ case. So, this $P(z)$ is basically the number in the system at the departure epoch.

Now, $Q(z)$ is the PGF of the number in the system at a departure epoch of a $M/G/1$ queue with vacation. $P(z)$ is without vacation, $Q(z)$ is with vacation, and $V(z)$ is another one which is PGF of the number in the system at a random point in time when the server is on vacation.
This is at the random point, but when the server is on vacation. So, that number, whatever be its distribution, and this V of z is its PGF. Now, for Poisson input queue, the basic decomposition result is $Q(z) = P(z)V(z)$. What is that?

This is, in a way, queue length, distribution of the departure epoch is given by this product of these two generating functions. So, if you think in terms of the random variable, it is the sum of two independent random variables; one is the ordinary M G 1 queue, the other is a random variable connected with the vacation duration random variable.

So, this basic decomposition result shows that the number of customers at a departure epoch of a Poisson input queue with vacation is the sum of two random variables.
▶ One is the number of customers at a departure epoch of the corresponding Poisson input queue without vacation, usual $M/G/1$, and
▶ The other is the number of customers at a random point of time given that the server is on vacation. So, while variable $(i)$ is vacation independent, the $(ii)$ is basically is a vacation-related quantity.

That is the decomposition result that we set holds in the case of vacation queuing models. And that is what is these two random variables listed out. So, the main result the way decomposition result is what is given in this theorem which is what it says is

**Theorem.** *[Decomposition Result]*
*For an $M/G/1$ queue with server vacations,*

$$Q(z) = P(z)\frac{1 - R(z)}{(1 - z)E(N^*)} \tag{1}$$

So, this is what is one can obtain. But the proof we are not going to do. One can look elsewhere. For example, one can look at the basic paper of this, which is in 1984 by Foreman, 1984, like that is one can have a look at that. But what we look at it some special cases like what it happens, this generally general result which one can prove; what will happen in some special cases.

**Some special cases:**

(A) Let us take N* is deterministic first, which means that this result we are going to take we are going to deduce this result to something simpler form depending upon what is our system is actually.

(A)-(i) $\mathrm{P}\{\mathrm{N}^* = 1\} = 1$, we get the usual queue with the vacation period corresponding to the idle period of the system. Then $Q(z) = P(z)$.

(A)-(ii) $N$ is a fixed number, say $N$–that is, $\mathrm{P}\{N^* = N\} = 1$. This corresponds to the case when the server is on vacation (or remains busy with other work or secondary customers) until the (primary) queue size builds up to a preassigned fixed number $N$, known as $N$-policy. This was considered first by Heyman (1968), who shows that a system with such a policy possesses some optimal properties. From (1), we get

$$Q(z) = P(z)\frac{1 - z^N}{(1 - z)N}.$$

- In the preceding two cases under $(A)$, the length of the server vacation depends on the arrival process during but not after the vacation. Under $(\boldsymbol{B}\ )$ (given below), the length of server vacation is independent of the arrival process.

(B) The second case that we consider is $N^*$ is a random variable

(B)-(a) Now, in this, the first specific case here is that this is an $M/G/1-V_m$; in this case, what happens.

- Let $A_v$ be the number of arrivals during a typical vacation period $v$. Then the PGF $\alpha(z)$ of $A_v$ is given by

$$\alpha(z)=\sum_{n=0}^{\infty}\mathrm{P}\{A_v=n\}z^n=F_v^*[\lambda(1-z)] \tag{2}$$

We have

$$\mathrm{P}\left\{A_v=0\right\}=F_v^*(\lambda) \tag{3}$$

$$\text{so that}\qquad \mathrm{P}\left\{A_v\geq 1\right\}=1-F_v^*(\lambda). \tag{4}$$

- Now the event $N^*=n$ is the event that the number of arrivals during the last vacation period equals $n$, given that this number is at least $1$. That is,

$$\mathrm{P}\{N^*=n\}=\mathrm{P}\left\{A_v=n\mid A_v\geq 1\right\}=\frac{\mathrm{P}\left\{A_v=n\right\}}{1-F_v^*(\lambda)},\quad n=1,2,\ldots \tag{5}$$

$$\text{Thus, } R(z)=\sum_{n=1}^{\infty}\mathrm{P}\left\{N^*=n\right\}z^n=\frac{F_v^*[\lambda(1-z)]-F_v^*(\lambda)}{1-F_v^*(\lambda)} \tag{6}$$

$$\Longrightarrow\ E\left(N^*\right)=R'(1)=\frac{-\lambda F_v^{*\prime}(0)}{1-F_v^*(\lambda)}=\frac{\lambda E(v)}{1-F_v^*(\lambda)}. \tag{7}$$

- Substituting in (1), we get

$$Q(z)=P(z)\frac{1-F_v^*[\lambda(1-z)]}{\lambda E(v)(1-z)}. \tag{8}$$

That is what you can obtain as this expression here. So, again, this is the decomposition thing. $P(z)$ is the usual PGF that you have for the number the system in the usual $M/G/1$ without vacation. And $\frac{1-F_v^*[\lambda(1-z)]}{\lambda E(v)(1-z)}$ is the additional term; the second term is the additional quantity that you would multiply $P(z)$ to get the corresponding quantity of the $M/G/1$ with vacation model.

Now, let us see some remarks with; before we go to the another model, we let us see some remarks with respect to the second quantity first and then overall also as well.

Some remarks on (B)-(a).

**Remark 1** The second factor in $Q(z)$ has an interesting interpretation. Let $Z(t)$ be the residual lifetime of the vacation random variable $v$.

At any given point of time, suppose if I get the residual vacation period, residual vacation time. If it is exponential, you know that it is same as the original one, but if it is nonexponential, then really, this is a different one. Then the limiting distribution $Z$ of $Z(t)$ as $t \to \infty$ is given by

$$F_z(x) = \mathrm{P}\{Z \le x\} = \frac{\int_0^x \left[1 - F_v(y)\right] dy}{E(v)},$$

where $F_v(y) = \mathrm{P}\{v \le y\}$. Let $b_n$ be the probability that $n$ arrivals occur during $Z$ and let $\beta(z) = \sum_{n=0}^{\infty} b_n z^n$ be the PGF of the number of arrivals during $Z$. Then

$$\begin{aligned}\beta(z) &= \sum_{n=0}^{\infty} z^n \int_0^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} dF_z(t) = \int_0^{\infty} \left\{ \sum \frac{e^{-\lambda t}(\lambda t z)^n}{n!} \right\} \frac{1 - F_v(t)}{E(v)} dt \\ &= \int_0^{\infty} \frac{e^{-\lambda t(1-z)} \left(1 - F_v(t)\right)}{E(v)} dt = \frac{1 - F_v^*[\lambda(1-z)]}{\lambda E(v)(1-z)}\end{aligned}$$

which is equal to the second factor on the RHS of (8).

Thus, while the first factor is the PGF of the number at departure epoch in the standard $M/G/1$ queue without vacation, the second factor is the PGF of the number of arrivals during the limiting residual vacation period.

Basically, it is not just some expression that we are obtaining it; it has an interpretation that is what is important. That is relevant; you want to look at that particular scenario. That is what is one remark.

**Remark 2** We have $\alpha'(1) = E\left(A_v\right) = \lambda E(v)$, so that the second factor on the RHS of (8) can be written as

$$\frac{1 - \alpha(z)}{(1-z)\alpha'(1)}.$$

Thus, for $M/G/1 - V_m$, (1) can be put as

$$Q(z) = P(z) \frac{1 - \alpha(z)}{(1-z)\alpha'(1)}, \tag{9}$$

where $\alpha(z)$ is the PGF of the number of arrivals during the vacation.

Note: The factor $\dfrac{1 - \alpha(z)}{(1-z)\alpha'(1)}$ is the PGF of $\mathrm{P}\left\{A_v > k\right\}/E(A_v), \quad k = 0, 1, 2, \ldots$.
This is also the PGF of the number of units that arrive during an interval from the commencement of a vacation period to a random point in the vacation period.

**Remark 3** In an $M/G/1 - V_m$ queue, the (server) idle period $I$ has mean

$$E(I) = E(v) / \left[1 - F_v^*(\lambda)\right].$$

**Remark 4** For $M^X/G/1 - V_m$ system, $Q(z)$ will become

$$Q(z) = V(z) \frac{1 - F_v^*[1 - \lambda A(z)]}{\lambda E(v)[1 - A(z)]}$$

where $A(z)$ is the PGF of $X$, and

$$V(z) = \frac{(1-\rho)(1-z)B^*(\lambda - \lambda A(z))}{B^*(\lambda - \lambda A(z))}.$$

This is for $M/G/1$ when $N^*$ is random and $M/G/1$ with multiple vacation.

(B)-(b) $M/G/1 - V_s$ model

- Here there is only one vacation, and there may be no arrivals or one arrival or more than one arrival during the server vacation period.
- If there is no arrival, the server waits for an arrival to occur, and then $N^* = 1$. If there is an arrival during the vacation, then $N^*$ is equal to the number of arrivals during the vacation. Thus,

$$\begin{aligned} \mathrm{P}\left\{N^* = 1\right\} &= \mathrm{P}\left\{A_v = 0\right\} + \mathrm{P}\left\{A_v = 1\right\} \\ \mathrm{P}\left\{N^* = n\right\} &= \mathrm{P}\left\{A_v = n\right\}, \quad n = 2, 3, \ldots. \end{aligned}$$

- Thus, using (2) and (3), we get

$$\begin{aligned} R(z) &= \sum_{n=1}^{\infty} \mathrm{P}\left\{N^* = n\right\} z^n = \mathrm{P}\left\{A_v = 0\right\} z + \sum_{n=1}^{\infty} \mathrm{P}\left\{A_v = n\right\} z^n \\ &= zF_v^*(\lambda) + F_v^*[\lambda(1-z)] - F_v^*(\lambda) = F_v^*[\lambda(1-z)] - (1-z)F_v^*(\lambda) \\ \text{and} \quad E\left(N^*\right) &= R'(1) = -\lambda F_v^{*\prime}(0) + F_v^*(\lambda) = \lambda E(v) + F_v^*(\lambda). \end{aligned}$$

- Substitution in (1) gives

$$Q(z) = P(z)\, \frac{1 - F_v^*[\lambda(1-z)] + (1-z)F_v^*(\lambda)}{(1-z)\left[\lambda E(v) + F_v^*(\lambda)\right]}. \tag{10}$$

So, this is the total queue length distribution at the departure epoch as far as your $M/G/1$ queues with vacation is concerned. You see that we have given the decomposition result on these two specific cases; we have obtained what is the PGF of the $M/G/1$ queue with vacation model.

We will also quickly see the waiting time distribution in this particular case. Basically, we are looking at the waiting time, the sojourn time, or waiting time in the system. We assume that the queue discipline is FIFO, and the waiting time of a customer is independent of the input process that occurs after the epoch of arrival of the customer considered, which is not satisfied for an N-policy queue. So, whatever we are saying, it is not applicable for $N$-policy queues because you are considering an arbitrary customer in a vacation system and the waiting time of an arbitrary customer in a vacation system with Poisson input. Now, if the waiting time of a customer is dependent on the input process that occurs after the epoch of arrival of the customer, if you consider the customer and if you have an $N$-policy, that means that if this customer number is less than $N$ and greater than $N$ matters. So, if so, this assumption is not hold for $N$-policy. What we are doing without that, we are assuming under this assumption only we are deriving this result. Now, let us call $W_1(.)$ as the distribution function or the CDF of the waiting time of a customer in a standard $M/G/1$ queue which is without vacation, and $W_1^*$ its Laplace-Stieltjes transform. We are following this $G/G/1$ notation, only using this W itself directly for the distribution as well as its mean. So, you have to be careful when you are interpreting this. And let $W(.)$ and $W^*(.)$ be the corresponding functions of an $M/G/1$ queue with vacation under the assumptions we have stated here. $W_1(.)$ is without vacation $W(.)$ is with vacation, then the basic decomposition result now is given by this

**Theorem.**
*For an $M/G/1$ queue with vacations,*

$$W^*(s) = W_1^*(s) V\left(1 - \frac{s}{\lambda}\right). \tag{11}$$

Proof Under FIFO discipline, the customers left behind by an (arbitrary) departing customer are precisely those customers that arrived during the waiting time of the departing customer. It follows that

$$Q(z) = \int_0^\infty e^{-\lambda(1-z)t} dW(t) = W^*[\lambda(1-z)]$$

so that, putting $\lambda(1-z) = s$, we get

$$W^*(s) = Q\left(1 - \frac{s}{\lambda}\right) \tag{12}$$

for an $M/G/1$ queue with vacation.

Similarly, for the queue without vacation

$$W_1^*(s) = P\left(1 - \frac{s}{\lambda}\right) \tag{13}$$

Using the basic decomposition result (**??**), one gets

$$W^*(s) = W_1^*(s)\, V\left(1 - \frac{s}{\lambda}\right).$$

Case (i) $M/G/1 - V_m$
Putting $s = \lambda(1-z)$ in $(8)$, we get

$$Q\left(1 - \frac{s}{\lambda}\right) = P\left(1 - \frac{s}{\lambda}\right) \frac{1 - F_v^*(s)}{sE(v)}$$

so that for such a system

$$W^*(s) = W_1^*(s) \frac{1 - F_v^*(s)}{sE(v)}, \tag{14}$$

where $W_1^*(s)$ is given by the Pollaczek-Khinchin formula.

- Note that

$$V\left(1 - \frac{s}{\lambda}\right) = \frac{1 - F_v^*(s)}{sE(v)}$$

is the LST of the forward-recurrence time of a vacation.

- In this vacation model, the waiting-time distribution decomposes into two independent components.
- One is the waiting time distribution in the corresponding model without vacation. The other is the forward-recurrence time of the vacation.
- When the vacation distribution is given, the decomposition result reduces the problem to a convolution problem.

- The mean queueing time $W_q$ in a vacation model is given by

$$W_q = \frac{\lambda E(S^2)}{2(1-\rho)} + \frac{E(v^2)}{2E(v)}$$
$$= \text{ mean waiting time in the standard } M/G/1 \text{ queue}$$
$$+ \text{ mean residual vacation time}$$
$$i.e. \quad W_q\left(M/G/1 - V_m\right) = W_q(M/G/1) + \text{ mean residual vacation time.}$$

Case (ii) $M/G/1 - V_s$ model

Putting $s = \lambda(1-z)$ in (10), we get (using (12) and (13)

$$W^*(s) = W_1^*(s)\frac{F_v^*(\lambda) + \left(\frac{\lambda}{s}\right)\left[1 - F_v^*(s)\right]}{\lambda E(v) + F_v^*(\lambda)}. \tag{15}$$

So, this is what we have with respect to the case of queues with vacation. Of course, there are plenty of things that one can do for such queues, but we are just giving the basic idea and how one can analyze what the difference that you would say is. The decomposition property is the main aspect when dealing with a vacation queue, so you will not look at it in totality; you will basically break it into two pieces. One is connected with the without vacation model, and then what comes additionally is what then you need to be concerned about it. So, this is you can explore if you want further the vacation queueing models. So, with that, we end our discussion of these queues with the vacation. So, in summary, what we have seen in this course is that we have seen a wide variety of queuing models starting from birth-death models to general Markovian models. General Markovian models arise because of various features that you try to incorporate like retrial, impatience, vacation, or any other feature. And which can be applied to a large number of systems or networks depending upon where you find this application. Mainly, we have concentrated on majority of the course of the Markovian type, this being the introductory course. And one can do a complete analysis in that particular case; that is what was the advantage. And these ideas were either extended or is extendable to semi Markovian models very easily because you just carry over, you try to still extract a Markov process out of a non-Markov process and try to do the analysis and also to general queuing network models, again there you are trying to get the Markovian structure. Now, where these models can be applied is up to you wherever you find. Our idea was that we get exposure to the different kinds of queuing models and how one can analyze that queuing models, and what one looks for in a typical one queuing system, and how one can utilize. So, we have there is whole lot beyond this point. We have not talked about how this can be used to optimize the system in some sense. So, that is a complete domain that is an optimal optimization problem you have to frame based upon whatever quantities that we estimate or derive based upon our modelling framework and then try to give answers to that. That is always is there. Like there are further directions one can take from here. And till date, and even now like, you would find N number of queueing models, but you will encounter a situation where none of these models may be applicable, so that should lead you to develop a new queueing model. When you do that, you should by now at least have some familiarity of how what you will do with the model, how you can analyze, what are the thing that you are looking for from such models in order to make some sense out of your the whole process that you might be employing here. So, this theory and application, in any applied field, is always the case that the theory is applicable in more situations, and then newer situations demand a new development in the case of the theory and models and analysis and so on. So, this two-way connection is always what leads to a healthy growth of the subject as well as its application. Hope, you have had exposure, at least to some extent, to the basic queuing models in this course. And hope it will be useful to you when you want to really apply to a some

scenario that you encounter in practice in your field and as well as if you want to explore further in the area. So, with wish you all the very best in all your endeavours, whether you are using queuing theory or otherwise.
Thank you. Bye.