



maximum of the total number of jobs that can be present in the system at any instant of time (i.e. one job being served and  $K-1$  others waiting for service). Since the queue is of finite capacity, jobs arriving when the system is full (i.e.  $K$  jobs in the system) are lost and have to leave the system without getting any service. Such jobs are also referred to as being blocked. The probability of jobs being lost (also referred to as the blocking probability,  $P_B$ ) is as important a performance parameter for the finite capacity system as its delay and throughput.

Following our usual practice, we define the system state at time  $t$  to be the number in the system at that instant. As done for the M/G/1 queue of Section 3.2, we consider the imbedded Markov Chain of the system states just after the departure instants of the jobs that leave the queue after obtaining service. Note that the jobs, which get lost because of blocking, do not actually enter the queue. Their departures (without getting service) do not contribute to the imbedded time instants considered here. Let  $\lambda$  be the average arrival rate of jobs (from the Poisson arrival process) to the queue. Note that of these arrivals, only a fraction  $(1-P_B)$  will actually be able to enter the queue.

Consider the imbedded Markov Chain of system states at these time instants  $t_i$   $i=1,2,3,\dots$  when the  $i^{\text{th}}$  job departs from the system after obtaining service. At a time instant  $t_i$ , the system state  $n_i$  will be the number of jobs left behind in the system when the  $i^{\text{th}}$  job departs. Note that  $n_i$  will range between 0 and  $K-1$  since the departure of the job cannot leave the system completely full, i.e. with system state  $K$ . Let  $a_i$  be the number of arrivals (from the Poisson arrival process) in the  $i^{\text{th}}$  service time. The equations for the corresponding Markov Chain can then be written as

$$\begin{aligned} n_{i+1} &= \min\{a_{i+1}, K-1\} & \text{for } n_i = 0 \\ &= \min\{n_i - 1 + a_{i+1}, K-1\} & \text{for } n_i = 1, \dots, (K-1) \end{aligned} \quad (1)$$

Note that the approach followed in Section 3.2 was to obtain the generating function of the system state at equilibrium using directly the expressions for the imbedded Markov Chain. For the M/G/1/K queue, it would be easier to directly compute the equilibrium state probabilities  $p_{d,k}$   $k=0,1,\dots,(K-1)$  at the departure instants of jobs from the queue. For this, we will need transition probabilities of the imbedded Markov Chain at equilibrium. At equilibrium, these are defined to be

$$p_{d,jk} = P\{n_{i+1} = k \mid n_i = j\} \quad 0 \leq j, k \leq K-1 \quad (2)$$

Let  $\mathbf{a}_k$  be the probability of  $k$  job arrivals to the queue during a service time, where the pdf of the service time is given as  $b(t)$  (with cdf  $B(t)$  and Laplace Transform  $L_B(s)$ ). Using this, we can find  $\mathbf{a}_k$  as

$$\mathbf{a}_k = \int_{t=0}^{\infty} \frac{(\mathbf{I}t)^k}{k!} e^{-\mathbf{I}t} b(t) dt \quad (3)$$

It may be noted that  $\mathbf{a}_k$  can also be found as the coefficient of  $z^k$  in the expansion of  $L_B(\mathbf{I}-\mathbf{I}z)$ . This may be proved by using  $\mathbf{a}_k$  of (3) to show that the generating function of the number of arrivals in a service duration will be  $L_B(\mathbf{I}-\mathbf{I}z)$ . This also follows from the comments made at the end of Section 2.6.2 and the approach used for obtaining equations like (2.36) or (3.13).

The transition probability  $p_{d,jk}$  for the two cases  $j=0$  and  $j=1, \dots, K-1$  will be found separately using the values of  $\mathbf{a}_k$  found in (3). The expressions for these are given in (4) and (5), respectively, based on the observation that the final state  $k$  cannot exceed  $K-1$ .

$$p_{d,0k} = \begin{cases} \mathbf{a}_k & 0 \leq k \leq K-2 \\ \sum_{m=K-1}^{\infty} \mathbf{a}_m & k = K-1 \end{cases} \quad j=0 \quad (4)$$

$$p_{d,jk} = \begin{cases} \mathbf{a}_{k-j+1} & j-1 \leq k \leq K-2 \\ \sum_{m=K-j}^{\infty} \mathbf{a}_m & k = K-1 \end{cases} \quad j=1, \dots, K-1 \quad (5)$$

Using the transition probabilities of (4) and (5), the equilibrium state probabilities  $p_{d,k}$   $k=0, 1, \dots, K-1$  at the departure instants may be calculated in the usual fashion by solving the  $K-1$  balance equations along with the normalisation condition. These equations will be as follows.

$$p_{d,k} = \sum_{j=0}^{K-1} p_{d,j} p_{d,jk} \quad k = 0, 1, \dots, K-1 \quad (5)$$

$$\sum_{k=0}^{K-1} p_{d,k} = 1 \quad (\text{Normalisation Condition}) \quad (6)$$

The transition probabilities  $p_{d,jk}$  of (4) and (5) may now be substituted in (5) and (6). This gives a set of linear equations that may be solved to get the corresponding state probabilities. Note that only  $K$  independent equations are needed, as there are only  $K$  unknowns (i.e.  $p_{d,k}$   $k=0,1,\dots,K-1$ ) to be found. This implies that apart from the normalisation condition of (6), only  $K-2$  equations are needed from the  $K-1$  equations of (5). This set of  $K-1$  equations is summarised in (7).

$$p_{d,k} = p_{d,0} \mathbf{a}_k + \sum_{j=1}^{k+1} p_{d,j} \mathbf{a}_{k-j+1} \quad k = 0, 1, \dots, K-2 \quad (7)$$

$$\sum_{k=0}^{K-1} p_{d,k} = 1$$

Alternatively, one can solve first for the normalised variables ( $p_{d,k}/p_{d,0}$ ) using

$$\frac{p_{d,k+1}}{p_{d,0}} = \frac{1}{\mathbf{a}_0} \left[ \frac{p_{d,k}}{p_{d,0}} - \sum_{j=1}^k \frac{p_{d,j}}{p_{d,0}} \mathbf{a}_{k-j+1} - \mathbf{a}_k \right] \quad k = 0, 1, \dots, K-2 \quad (8)$$

and then solve for  $p_{d,0}$  using the normalisation condition to get

$$p_{d,0} = \frac{1}{\sum_{k=0}^{K-1} \frac{p_{d,k}}{p_{d,0}}} \quad (9)$$

Using this and the values obtained earlier for ( $p_{d,k}/p_{d,0}$ ), one can then obtain the actual state probabilities  $p_{d,k}$   $k=1,\dots,K-1$  at the job departure instants.

Considering a system at equilibrium, let  $p_{a,k}$   $k=0,1,\dots,K$  be the probability that a newly arriving job, irrespective of whether it finally joins the queue or not, finds  $k$  jobs waiting in the queue. For this system, let  $p_k$   $k=0,1,\dots,K$  be the probability that the queue has  $k$  jobs in it at an arbitrarily chosen instant of time. Using the PASTA property of Section 2.5.1, we can then claim that

$$p_k = p_{a,k} \quad k = 0, 1, \dots, K \quad (10)$$

We can also define  $p_{ac,k}$   $k=0,1,\dots,K-1$  as the equilibrium probability of the system state  $k$  as seen by an arrival which does actually enter the queue. Based on the fact, that the state of the queue can change by at most  $\pm 1$  because of these arrivals and the departures from it, we can claim that

$$p_{d,k} = p_{ac,k} \quad k = 0, 1, \dots, K-1 \quad (11)$$

Using  $P_B$  as the equilibrium probability that an arrival is blocked (because the queue is full, i.e. in state  $K$ ,  $P_B = p_K$ ), we can see that

$$p_k = p_{a,k} = (1 - P_B) p_{ac,k} = (1 - P_B) p_{d,k} \quad k = 0, 1, \dots, K-1 \quad (12)$$

Note that this may also be confirmed by observing that

$$\sum_{k=0}^{K-1} p_{a,k} = 1 - P_B = \sum_{k=0}^{K-1} (1 - P_B) p_{ac,k}$$

$$\text{since } \sum_{k=0}^K p_{a,k} = 1 \text{ and } \sum_{k=0}^{K-1} p_{ac,k} = 1$$

Let  $\bar{X}$  be the mean service time of a job in the queue. The traffic load  $\mathbf{r}$  offered to the queue will then be given by  $\mathbf{r} = \mathbf{I}\bar{X}$ . Since the average arrival rate of jobs actually entering the queue (also the average departure rate of jobs leaving the queue) is  $\mathbf{I}_c = \mathbf{I}(1 - P_B)$ , the actual traffic throughput of the queue will be  $\mathbf{r}_c = \mathbf{r}(1 - P_B)$ .

Note that this implies that the probability  $p_0$  of finding the queue empty at an arbitrary time will be

$$p_0 = 1 - \mathbf{r}_c$$

Using (12) for the case  $k=0$ , we can then write

$$1 - \mathbf{r}(1 - P_B) = (1 - P_B) p_{d,0} \quad (13)$$

Since  $p_{d,0}$  has been found earlier using (7)-(9), we can use (13) to find the blocking probability  $P_B$  (or  $p_K$ ) as

$$P_B = 1 - \frac{1}{p_{d,0} + \mathbf{r}} \quad (14)$$

Using the values of  $p_{d,k}$  obtained using (7)-(9) and the results of (12) and (14), the equilibrium state distribution  $p_k$ ,  $k=0,1,\dots,(K-1)$  of the queue at arbitrary time instants may then be shown to be

$$p_k = \frac{1}{p_{d,0} + \mathbf{r}} p_{d,k} \quad k=0,1,\dots,K-1 \quad (15)$$

The equilibrium state distribution may now be used in the usual fashion to find the mean number  $N$  in the system as

$$N = \sum_{k=0}^K k p_k = \frac{1}{(p_{d,0} + \mathbf{r})} \sum_{k=0}^{K-1} k p_{d,k} + K \left( 1 - \frac{1}{(p_{d,0} + \mathbf{r})} \right) \quad (16)$$

Note that the effective arrival rate  $I_c$  to the queue will be given by

$$I_c = I(1 - P_B) = \frac{I}{(p_{d,0} + \mathbf{r})} \quad (17)$$

Using this and Little's result, the mean total time spent in system by a job actually entering the queue will be

$$W = \frac{N}{I_c} = \frac{\sum_{k=0}^{K-1} k p_{d,k} + K(p_{d,0} + \mathbf{r} - 1)}{I} \quad (18)$$

This may be used to get the mean time spent waiting in the queue  $W_q$  as

$$W_q = W - \bar{X} = \frac{1}{I} \sum_{k=0}^{K-1} k p_{d,k} + \frac{K}{I} (p_{d,0} + \mathbf{r} - 1) - \bar{X}$$

where  $\bar{X}$  is the mean service time. The second moment of the time spent waiting in queue has also been obtained in [Takagi2] and is given by

$$\begin{aligned} \overline{w_q^2} = (K-1) & \left[ (K-2)(\bar{X})^2 + \overline{X^2} - \frac{K\bar{X}p_{d,0}}{I} - \frac{2\bar{X}}{I} \sum_{k=1}^{K-1} k p_{d,K-k} \right] \\ & + \frac{1}{I^2} \sum_{k=1}^{K-1} k(k+1) p_{d,K-k} \end{aligned}$$

where  $\overline{X^2}$  is the second moment of the service time.

## Proportionality Relationship between the M/G/1 and the M/G/1/K Queues

Consider the way (7) would be written for a M/G/1 (i.e. M/G/1/ $\infty$ ) queue. Denoting the corresponding state probabilities  $p_{\mu,k}$  this equation for the M/G/1 queue would be

$$p_{\infty,k} = p_{\infty,0} \mathbf{a}_k + \sum_{j=1}^{k+1} p_{\infty,j} \mathbf{a}_{k-j+1} \quad k = 0, 1, \dots, \infty \quad (19)$$

$$\sum_{k=0}^{\infty} p_{\infty,k} = 1$$

Note that for the M/G/1 queue, the state probabilities at the departure instants, arrival instants and at an arbitrary time instant would be the same. (This has been discussed earlier in Section 3.2.) Comparing the form of (7) and (19), we conclude that the state probabilities  $p_{d,k}$  of the M/G/1/K queue and the state probabilities  $p_{\mu,k}$  of the M/G/1 queue will be proportional to each other for  $k=0, 1, \dots, K-1$ . Using the normalisation condition, the proportional relation between them may be shown to be

$$p_{d,k} = \frac{P_{\infty,k}}{\sum_{j=0}^{K-1} P_{\infty,j}} \quad k = 0, 1, \dots, K-1 \quad (20)$$

Note that (20) implies that the equilibrium state probabilities at the departure instants of the M/G/1/K queue may be obtained by a simple truncation and scaling of the equilibrium state probabilities of the corresponding M/G/1 queue. This may be used to find  $p_{d,k}$  for states  $k=0, 1, \dots, (K-1)$  and then (12) may be used to find the equilibrium state probabilities  $p_k$  for  $k=0, 1, \dots, K-1$ . Finally, the normalisation result may be used to find the blocking probability  $P_B=p_k$ . In order to find the equilibrium state probabilities of the M/G/1 queue, we can either directly use (19) or invert the generating function  $P(z)$  obtained in (3.14) of Section 3.2.2.

## Pushout Operation of the M/G/1/K Queue

In our earlier description for the M/G/1/K queue, we have used the approach that a newly arriving job which sees the queue full, leaves without service. Note that the order in which jobs are served can be FCFS, LCFS or SIRO. As usual, we may note that the sequence, in which the jobs are served once they are in the queue, will not affect the mean performance parameters, i.e.  $N$ ,  $N_q$ ,  $W$  and  $W_q$ .

One can consider an alternate method for handling the jobs that arrive when the system is full. In this, a newly arriving job is always accepted. If the queue is full when the job arrives, it discards the one that has waited in the queue for the longest time. (Note that a job in service is never discarded and is allowed to continue its service until completion.) This strategy is referred to as the *pushout* strategy. Note that even with the pushout strategy, one can still operate the queue following FCFS, LCFS or SIRO service disciplines.

It may be noted that the pushout strategy is a reasonable one to follow in systems where a later job/message/packet arrival makes an earlier one redundant in some way. For example, this may happen in a system handling voice or video packets where one would prefer to discard the oldest packet waiting in the queue rather than the more recent arrival.

The mean number (i.e.  $N$  or  $N_q$ ) in a M/G/1/K queue following a pushout strategy would be the same as for one where such a strategy is not being followed. The average departure rate of jobs  $I_c$  and the actual throughput of the queue  $r_c$  will also be the same in the two cases.

A detailed derivation of the queue's performance under the pushout strategy may be found in [Takagi2]. This derivation and the associated results are somewhat difficult. However, one can easily comment on the relative mean performance of the M/G/1/K queue operated with and without the pushout strategy.

Applying Little's result to the M/G/1/K queue with and without pushout strategy, we get that

$$I_c W_q = I(1 - P_B)W_q = I W_{q,P} \quad (21)$$

where  $W_q$  is the mean waiting time in the M/G/1/K queue without pushout and  $W_{q,P}$  is the mean time spent in the queue prior to service by a job in the M/G/1/K queue with pushout. Note that in the latter case, this waiting time will include both the jobs which eventually get served and ones which get pushed out and, hence, leave without service.

Note that since  $I_c = I(1 - P_B)$ , (21) implies that



$$W_{q,P} = (1 - P_B)W_q \quad (22)$$

and that, therefore

$$W_{q,P} \leq W_q \quad (23)$$

We may also note that for the M/G/1/K queue with pushout, the delay parameter  $W_{q,P}$  actually consists of two components. One component,  $W_{q,PS}$  is the mean waiting time in the queue as seen by jobs which eventually do get served. The other component,  $W_{q,PNS}$  corresponds to the time spent waiting in the queue by jobs which get pushed out (after spending some time waiting for service) and leave without service. It may also be noted that  $P_B$  and  $(1 - P_B)$  are the respective probabilities that a job is not eventually served and that a job does get service. Using this, we get that

$$W_{q,P} = (1 - P_B)W_{q,PS} + P_B W_{q,PNS} \quad (24)$$

Substituting (22) in (24) gives

$$(1 - P_B)(W_q - W_{q,PS}) = P_B W_{q,PNS} \geq 0 \quad (25)$$

and that, therefore

$$W_{q,PS} \leq W_q \quad (26)$$

Note that (26) leads to the following important conclusion. For a M/G/1/K queue with pushout, the queueing delay seen by the jobs which eventually get served will be less than what one would observe for a queue without pushout. Since the system throughput will be the same in both cases, the queue with pushout provides a way of giving improved service (lower delays) to the jobs that actually do get service. The reader is referred to [Takagi2] for more detailed analysis of the M/G/1/K queue operated with the pushout strategy.

### **An Alternate Derivation for the State Probabilities at an Arbitrary Instant in a M/G/1/K Queue**

It is possible to provide a more direct approach to finding the state probabilities of the M/G/1/K queue at an arbitrary instant of time for a queue

in equilibrium. For this, we first note that the mean time interval between successive imbedded points (at the job departure instants) will be

$$\frac{1}{I} + \bar{X} \quad \text{if the queue is empty at the earlier imbedded point}$$

$$\bar{X} \quad \text{if the queue is not empty at the earlier imbedded point}$$

Using the above, we obtain the probability  $p_0$  that the queue is empty at an arbitrarily chosen time to be

$$p_0 = \frac{\left(\frac{1}{I}\right) p_{d,0}}{\left(\frac{1}{I} + \bar{X}\right) p_{d,0} + \bar{X}(1 - p_{d,0})} = \frac{p_{d,0}}{p_{d,0} + \mathbf{r}} \quad (27)$$

which agrees with the expression  $p_0 = 1 - \mathbf{r}_c$  obtained earlier or as given in (12) and (15) for  $k=0$ .

Now consider the situation where the arbitrarily chosen time instant falls within a service duration where  $x$  is the amount of service already provided. We consider the case where there are  $k$  jobs in the system for  $k=1, \dots, (K-1)$ .

The pdf of  $x$  may be found from residual life arguments to be  $\frac{1 - B(x)}{\bar{X}}$  where

$\mathbf{r}_c$  is the probability that the arbitrarily chosen time instant will fall within a service time. Considering separately the two cases where the previous departure left the queue empty or left the queue with  $j, j=1, \dots, k$  jobs in it, we get

$$p_k = \mathbf{r}_c \left[ p_{d,0} \int_0^{\infty} \frac{(Ix)^{k-1}}{(k-1)!} e^{-Ix} \frac{1 - B(x)}{\bar{X}} dx \right] + \mathbf{r}_c \sum_{j=1}^k p_{d,j} \int_0^{\infty} \frac{(Ix)^{k-j}}{(k-j)!} e^{-Ix} \frac{1 - B(x)}{\bar{X}} dx \quad k=1, \dots, (K-1) \quad (28)$$

Let  $A_k$  be the probability that there are  $k$  or more arrivals during a service time where

$$\begin{aligned}
 A_k &= \sum_{j=k}^{\infty} \mathbf{a}_j = \sum_{j=k}^{\infty} \int_0^{\infty} \frac{(\mathbf{I}x)^j}{j!} e^{-Ix} b(x) dx \\
 &= \int_0^{\infty} \frac{(\mathbf{I}x)^{k-1}}{(k-1)!} e^{-Ix} [1 - B(x)] \mathbf{I} dx
 \end{aligned}
 \tag{29}$$

$k=1,2,\dots,\mathbf{K}$

and  $\sum_{k=1}^{\infty} A_k = \mathbf{I}\bar{X} = \mathbf{r}$  (30)

(The derivation of these results is given in the Appendix.)

Substitution of (29) in (28) gives

$$p_k = \frac{\mathbf{r}_c}{\mathbf{r}} \left[ p_{d,0} A_k + \sum_{j=1}^k p_{d,j} A_{k-j+1} \right] \quad k=1,\dots,(K-1) \tag{31}$$

To simplify this further, we need the following result which may be proved by recursion.

$$p_{d,k} = p_{d,0} A_k + \sum_{j=1}^k p_{d,j} A_{k-j+1} \tag{32}$$

Substituting this in (31), gives  $p_k$ ,  $k=1,\dots,(K-1)$  with the same expression as obtained earlier in (12) and (15).

We can use a similar approach to find the probability  $p_K$ , i.e.  $p_k$  for  $k=K$ , when the queue is found to be full at an arbitrary time instant. In this case, we do need to take into account the situation that arrivals coming when the system is already full will be denied entry into the queue and will be lost. Following the same arguments as given earlier, we get

$$\begin{aligned}
 p_K &= \mathbf{r}_c \left[ p_{d,0} \sum_{k=K-1}^{\infty} \int_0^{\infty} \frac{(\mathbf{I}x)^k}{k!} e^{-Ix} \frac{1-B(x)}{\bar{X}} dx \right] \\
 &+ \mathbf{r}_c \sum_{j=1}^{K-1} p_{d,j} \sum_{k=K-j}^{\infty} \int_0^{\infty} \frac{(\mathbf{I}x)^k}{k!} e^{-Ix} \frac{1-B(x)}{\bar{X}} dx
 \end{aligned}
 \tag{33}$$

Using (29), we can rewrite (33) as

$$p_K = \frac{\mathbf{r}_c}{\mathbf{r}} \left[ p_{d,0} \sum_{k=K}^{\infty} A_k + \sum_{j=1}^{K-1} p_{d,j} \sum_{k=K-j+1}^{\infty} A_k \right] \quad (34)$$

To simplify (34) further, we need the result that

$$\left[ p_{d,0} \sum_{k=K}^{\infty} A_k + \sum_{j=1}^{K-1} p_{d,j} \sum_{k=K-j+1}^{\infty} A_k \right] = \mathbf{r} + p_{d,0} - 1 \quad (35)$$

Note that this follows from summing  $p_{d,k}$  over  $k=1, \dots, (K-1)$  and using (30) to get

$$\begin{aligned} 1 - p_{d,0} &= \sum_{k=1}^{K-1} p_{d,k} = p_{d,0} \sum_{k=1}^{K-1} A_k + \sum_{k=1}^{K-1} \sum_{j=1}^k p_{d,j} A_{k-j+1} \\ &= \mathbf{r} - p_{d,0} \sum_{k=K}^{\infty} A_k - \sum_{j=1}^{K-1} p_{d,j} \sum_{k=K-j+1}^{\infty} A_k \end{aligned} \quad (36)$$

Substituting (35) in (34), we get the same expression for  $p_K$  as given for the blocking probability  $P_B$  (which is the same quantity) in (14).

## Appendix:

Consider (29) for  $k=1$ , i.e. for  $A_1$ . For this, we need to prove that

$$\begin{aligned} \sum_{j=1}^{\infty} \int_0^{\infty} \frac{(\mathbf{1}x)^j}{j!} e^{-\mathbf{1}x} b(x) dx &= 1 - \int_0^{\infty} e^{-\mathbf{1}x} b(x) dx \\ &= \mathbf{1} \int_0^{\infty} e^{-\mathbf{1}x} [1 - B(x)] dx \end{aligned} \quad (A.1)$$

Using integration by parts, we can show that

$$\int e^{-\mathbf{1}x} b(x) dx = e^{-\mathbf{1}x} B(x) + \mathbf{1} \int e^{-\mathbf{1}x} B(x) dx \quad (A.2)$$

and that

$$\int_0^{\infty} e^{-Ix} b(x) dx = I \int_0^{\infty} e^{-Ix} B(x) dx \quad (\text{A.3})$$

$$\int_0^{\infty} e^{-Ix} dx = \frac{1}{I}$$

Substituting (A.3) in (A.1), we can show that (29) holds for  $k=1$ . We can also show from (A.2) that

$$\int e^{-Ix} [I - IB(x) + b(x)] dx = -e^{-Ix} [1 - B(x)] \quad (\text{A.4})$$

and that

$$\begin{aligned} & \int_0^{\infty} \frac{(Ix)^k}{k!} e^{-Ix} [I[1 - B(x)] + b(x)] dx \\ &= - \left( \frac{(Ix)^k}{k!} e^{-Ix} [1 - B(x)] \right) \Big|_{x=0}^{\infty} + I \int_0^{\infty} \frac{(Ix)^{k-1}}{(k-1)!} e^{-Ix} [1 - B(x)] dx \quad (\text{A.5}) \\ &= I \int_0^{\infty} \frac{(Ix)^{k-1}}{(k-1)!} e^{-Ix} [1 - B(x)] dx \end{aligned}$$

Since we have shown that (29) holds for  $k=1$ , we can now use induction to show the general result. Note that from the definition of  $A_k$ , we can also write that

$$A_{k+1} = A_k - \mathbf{a}_k = A_k - \int_0^{\infty} \frac{(Ix)^k}{k!} e^{-Ix} b(x) dx \quad (\text{A.6})$$

Assuming that (29) holds for  $k$ , we can substitute that result for  $A_k$  in the *RHS* of (A.5) and use (A.5) to get

$$\begin{aligned}
RHS &= A_k - \int_0^{\infty} \frac{(Ix)^k}{k!} e^{-Ix} b(x) dx \\
&= I \int_0^{\infty} \frac{(Ix)^{k-1}}{(k-1)!} e^{-Ix} [1 - B(x)] dx - \int_0^{\infty} \frac{(Ix)^k}{k!} e^{-Ix} b(x) dx \\
&= I \int_0^{\infty} \frac{(Ix)^k}{k!} e^{-Ix} [1 - B(x)] dx \\
&= LHS
\end{aligned}$$

We can therefore claim that if (29) holds for  $A_k$ , then it holds for  $A_{k+1}$ . Since we have shown that it holds for  $A_1$ , we can claim that it holds for all  $k, k=1,2,\dots,\infty$ .

Actually, one can give a physical reasoning to justify (29), without going through the details of the proof given above. For this, consider arrivals to a simple M/G/1 queue. The justification for  $A_k$  as given by (29), then follows by considering an arbitrary time instant within a service time. Let this time instant be the time at which the  $k^{\text{th}}$  arrival in the on-going service time enters the system. For this, let  $x$  be the time that has elapsed from the beginning of the currently on-going service to this instant. Note that  $[1-B(x)]$  will be the probability that the service time will be greater than or equal to  $x$ . Combining these, we get that

$$A_k = \int_{x=0}^{\infty} \frac{(Ix)^{k-1}}{(k-1)!} e^{-Ix} [1 - B(x)] (I dx)$$

from which (29) follows.

Once we have shown that (29) holds, (30) follows directly using

$$\begin{aligned}
\sum_{k=1}^{\infty} A_k &= \sum_{k=1}^{\infty} \int_{x=0}^{\infty} \frac{(Ix)^{k-1}}{(k-1)!} e^{-Ix} [1 - B(x)] (I dx) \\
&= \int_0^{\infty} [1 - B(x)] (I dx) \\
&= (I [1 - B(x)] x) \Big|_{x=0}^{\infty} + I \int_{x=0}^{\infty} x b(x) dx \\
&= I \bar{X}
\end{aligned}$$