Queueing with Balking

Author(s): Frank A. Haight

Source: *Biometrika*, Dec., 1957, Vol. 44, No. 3/4 (Dec., 1957), pp. 360-369

Published by: Oxford University Press on behalf of Biometrika Trust

Stable URL: https://www.jstor.org/stable/2332868

QUEUEING WITH BALKING

By FRANK A. HAIGHT
Auckland University College, New Zealand

1. Introduction

In dealing with problems of queueing, several writers (Kolmogoroff, 1932; Erlang*; Kendall, 1951, 1953; Lindley, 1952; Takács, 1955) have discussed the situation where queue stability is obtained by assuming that the demand for service does not overload the service mechanism. Thus λ, the average number of arrivals per unit time, is assumed to be less than μ, the average number of departures per unit time, so that their ratio, ρ, is less than unity. Kawata (1955), on the other hand, has shown that queue stability can also be obtained by assuming that, although arrivals occur more frequently than departures, some arrivals choose not to join the queue. In theory this case can be included in the original one, simply by supposing λ to be computed only from the values provided by those who actually join the queue. However, if the decision to join or not depends on some random variables, it is sensible to inquire into the relationship between these variables and those which characterize the queue, such as queue length and waiting time.

The factors which influence the decision of a person to join a queue or not may be considered under two general headings: (*a*) those relating to the importance of being served, and (*b*) those relating to the obstacle which the queue presents, namely the waiting time which he must experience. It is obvious that the waiting time cannot be found without a knowledge of the service times of all those in the queue; we shall, therefore, make the simplifying assumption that the individual measures the obstacle presented by the queue by its length when he arrives, which will be denoted by $k(t)$.

The factors included in (*a*) may be much more complicated, and may produce an opinion ranging from absolute urgency, so that a queue of arbitrary length will be joined, to absolute indifference, so that no non-zero queue will be joined. It will be assumed that these factors have been weighed by the individual before he arrives, and have produced in his mind an integer K, which is the greatest queue length that he will tolerate. If then he observes $k(t) \leqslant K$ on arrival, he joins the queue; but if $k > K$, he goes away and does not return. We assume that the values K chosen by the various individual arrivals may be regarded as being random samples from a certain distribution—the balking distribution. We shall sometimes allow K to be infinite with non-zero probability, so that a finite proportion of the arrivals will join any queue. In this paper we assume that if a person joins a queue, he must remain for service. In another paper we will permit him to test sequentially whether to stay or go, while he waits.

While the principal interest here is $\rho \geqslant 1$, it should be noted that the classic queueing cases are included, and that the appropriate results can be obtained by letting $K \to \infty$ when $\rho < 1$.

These results are (implicitly) special cases of a general set-up considered by Kendall & Reuter (1957). We derive them by assuming equilibrium, and investigate them numerically for various balking distributions.

* See Brockmeyer *et al.* (1948).

2. DIFFERENTIAL EQUATIONS

Suppose both arrivals and departures occur as events of homogeneous Poisson processes, of density λ, μ, respectively. Let

$$P(x,t) = \Pr(k \leqslant x \text{ at time } t), \quad F(x) = \Pr(K \leqslant x),$$
$$Q(x,t) = 1-P(x,t), \quad G(x) = 1-F(x),$$
$$p(x,t) = P(x,t)-P(x-1,t), \quad f(x) = F(x)-F(x-1).$$

Note that in general $p(0,t) = P(0,t) \neq 0$; this is the probability that no queue exists at time t. Also $f(0) = F(0) \neq 0$; this is the probability that an individual is absolutely queue-resistant. The distribution of K will be called the balking distribution.

Given a queue of length x, then the probability that an arrival joins it

$$= \Pr(\text{his balking value} \geqslant x)$$
$$= G(x-1).$$

The build-up of differential equations, following Feller, is

$$p(x,t+\Delta t) = [1-(\lambda+\mu)\Delta t]\,p(x,t)+\mu p(x+1,t)\Delta t + \lambda p(x-1,t)\,G(x-2)\Delta t+\lambda p(x,t)\,F(x-1)\Delta t+O(\Delta t)^2,$$

where the four contributions on the right-hand side are from

(a) no arrivals or leavers in Δt,
(b) one leaver in Δt,
(c) one arrival in Δt who joins,
(d) one arrival in Δt who balks,

respectively.

Take $p(x,t)$ from both sides, divide by Δt, and take the limit

$$(\partial/\partial t)\,p(x,t) = -(\lambda+\mu)\,p(x,t)+\mu p(x+1,t)+\lambda p(x-1,t)\,G(x-2)+\lambda p(x,t)\,F(x-1),$$

where, if $x = 0$, we delete μ in the first term; $F(-1) = 0 = F(-2)$. Add over x, giving

$$(\partial/\partial t)\,P(x,t) = \mu p(x+1,t)-\lambda p(x,t)\,G(x-1). \tag{1}$$

Writing $\lambda G(x) = \lambda_x$, (1) can be reduced to a form similar to the equation for a birth-and-death process given by Feller (1950), but differing both with respect to one subscript and with respect to the initial equation, namely

$$(\partial/\partial t)\,P(x,t) = \lambda_{x-1}P(x-1,t)-(\mu+\lambda_{x-1})\,P(x,t)+\mu P(x+1,t).$$

If, however, each person must join the queue, so that $F(x) = 0$ for all x, then (1) becomes a special case of the birth-and-death equation.

Using a method suggested by Koopman (1953) we can give a method for computing solutions of (1) in cases where they exist. Let $\phi(x,s)$ be the Laplace transformation of $p(x,t)$. Transforming, (1) becomes

$$-\lambda G(x-2)\,\phi(x-1,s)+(x+\lambda G(x-1)+\mu)\,\phi(x,s)-\mu\phi(x+1,s) = \delta_x,$$

where $\phi(-2,s) = \phi(-1,s) = 0$, $\delta_x = 1$ for $x = 0$ and 0 otherwise, thus assuming the queue empty at $t = 0$. Writing

$$\psi(x,s) = \phi(x,s)/\phi(x-1,s)$$

and using the equations corresponding to $x = 1, 2, \ldots$, it is seen that $\psi(1,s)$ may be written as a continued fraction

$$\psi(1,s) = \frac{\lambda}{B_1 -} \frac{\lambda\mu G(0)}{B_2 \quad -} \frac{\lambda\mu G(1)}{B_3 \quad -} \cdots.$$

Substituting this value back into the equation for $x = 0$, we find $\phi(0,s)$, and hence each $\phi(x,s)$. Values of $p(x,t)$ may then be found by the numerical inversion of the Laplace transform.

3. Equilibrium distributions

If equilibrium distributions of queue length exist as $t \to \infty$, they will be denoted by suppressing the letter t, and can be found from (1) by setting the left side equal to zero

$$p(1) = \rho p(0), \quad p(x+1) = \rho G(x-1)\, p(x) \quad (x = 1, 2, \ldots). \tag{2}$$

Defining
$$c_0 = c_1 = 1, \quad c_x = \prod_0^{x-2} G(i) \quad (x = 2, 3, \ldots)$$

this can be written
$$p(x) = \rho^x c_x p(0). \tag{3}$$

Summing,
$$1 = p(0) \sum_0^\infty \rho^x c_x.$$

Kawata (1955) and Kendall & Reuter (1957) have shown that the convergence of the series $\sum_0^\infty \rho^x c_x$ is necessary and sufficient for equilibrium to be attained, replacing the classical condition $\rho < 1$ (which is contained); the condition is that $p(0) > 0$.

Thus, the tails of the balking distribution are proportional to the ratio of ordinates of the queue length distribution, and when ρ is known, either may be computed from the other. Some examples will be given in § 5.

From (2), we have (if $p(x), p(x+1) \neq 0$)

$$f(x) = G(x-1) - G(x) = \frac{p(0)}{p(1)} \left(\frac{p(x+1)}{p(x)} - \frac{p(x+2)}{p(x+1)} \right), \tag{4}$$

so we must have
$$p^2(x+1) \geqslant p(x)\, p(x+2). \tag{5}$$

If $p(n) \neq 0$ but $p(n+1) = 0$, then from (2) we have $G(n-1) = 0$ and so $p(x) = 0$ for all $x > n$. This is the case where there is complete balking for queues of length n or more. Any finite distribution of queue-length satisfying (5) is attainable in this way; the condition for equilibrium will be satisfied for all ρ.

Any infinite distribution of queue-length (satisfying (5)) will correspond to a balking distribution (from (4)); there will be a positive probability that K is infinite (i.e. $G(\infty) \neq 0$) unless

$$\lim_{x \to \infty} \frac{p(x+1)}{p(x)} = 0. \tag{6}$$

If (6) is not satisfied, there is an upper limit to ρ for equilibrium to be attained. We have $c_x \sim A(G(\infty))^x$, so $\Sigma \rho^x c_x$ converges only if

$$\rho < \frac{1}{G(\infty)} = \lim_{x \to \infty} \frac{p(x)}{p(x+1)}.$$

The reverse problem, of finding $p(x)$ when $f(x)$ is given can be solved, at least numerically, in every case. However, there are some algebraic difficulties in the three steps of an analytical solution: (*a*) expressing $\rho G(x)$ in a convenient form, (*b*) evaluation of c_x, and (*c*) summing the series $\Sigma \rho^x c_x$.

4. GENERATING FUNCTIONS

Let
$$\eta(s,t) = \sum_0^\infty s^x p(x,t) = (1-s)\sum_0^\infty s^x P(x,t),$$
$$\xi(s,t) = \sum_0^\infty s^x p(x+1,t)\,F(x).$$

Summing (1) after multiplication by s^x, we obtain
$$\frac{\partial}{\partial t}\left(\frac{\eta(s,t)}{1-s}\right) = -\lambda p(0,t) + \left(\frac{\mu}{s}-\lambda\right)\{\eta(s,t)-p(0,t)\} + \lambda s\xi(s,t).$$

Letting $t\to\infty$ with the usual change of notation, we have (assuming equilibrium)
$$\eta(s) - \rho s\eta(s) - p(0) + \rho s^2\xi(s) = 0. \tag{7}$$

With $s = 1$,
$$\eta(0) = p(0) = 1-\rho+\rho\xi(1)$$

and since $0 < p(0) \leqslant 1$,
$$\frac{\rho-1}{\rho} < \xi(1) \leqslant 1.$$

It will be noted that $\xi(1)$ is just the (asymptotic) probability that an individual balks. Recalling that ρ was calculated for all arrivals, whether they joined or not, it can now be seen that an effective value of the traffic intensity, computed only from those who join the queue, say ρ', can be written $\rho' = \rho - \rho\xi(1)$.

Writing now $p_\rho(x)$, $\eta_\rho(x)$, m_ρ (mean queue length), etc., to denote the dependence on ρ, we have from (3)
$$\eta_\rho(s) = p_\rho(0)\sum_0^\infty s^x\rho^x c_x = \sum_0^\infty (\rho s)^x c_x \Big/ \sum_0^\infty \rho^x c_x = p_\rho(0)/p_{\rho s}(0), \tag{8}$$

and using (7)
$$\xi_\rho(s) = \frac{p_\rho(0)}{\rho s^2}\left\{1 - \frac{1-\rho s}{p_{\rho s}(0)}\right\}. \tag{9}$$

Thus the queue length and balking distributions are uniquely determined once $p_\rho(0)$ is known as a function of ρ.

We have from (3)
$$\rho\frac{\partial}{\partial\rho}\log p_\rho(x) = \rho\frac{\partial}{\partial\rho}\left\{\log c_x + x\log\rho - \log\sum_0^\infty \rho^x c_x\right\} = x - m_\rho \tag{10}$$

which essentially determines $p_\rho(x)$ (and in particular $p_\rho(0)$) when m_ρ is known (as a function of ρ). Also, the variance of the queue length,
$$v_\rho = \sum_0^\infty p_\rho(x)\,(x-m_\rho)^2 = \sum_0^\infty (x-m_\rho)\,\rho\frac{\partial}{\partial\rho}p_\rho(x) = \rho\frac{\partial}{\partial\rho}m_\rho \tag{11}$$

(since $\Sigma p_\rho(x) = 1$).

$\eta_\rho(s)$ and $\xi_\rho(s)$ each satisfy a simple partial differential equation; from (8) we have
$$\log\eta_\rho(s) = \log\eta_\rho(0) + g_1(\rho s)$$

whence
$$\left(\rho\frac{\partial}{\partial\rho} - s\frac{\partial}{\partial s}\right)\log\eta_\rho(s) = \rho\frac{\partial}{\partial\rho}\log\eta_\rho(0) = -m_\rho.$$

Also from (9)
$$\log\xi_\rho(s) = \log\eta_\rho(0) - \log\rho s^2 + g_2(\rho s),$$
$$\left(\rho\frac{\partial}{\partial\rho} - s\frac{\partial}{\partial s}\right)\log\xi_\rho(s) = -m_\rho - 1 + 2 = 1 - m_\rho,$$

where $g_1(\ )$ and $g_2(\ )$ are used to denote 'a function of' and 'another function of' in the above equations.

Differentiating (7) and setting $s = 1$ gives

$$m_\rho = \frac{\rho}{1-\rho}\{1-\xi_\rho'(1)-2\xi_\rho(1)\}. \tag{12}$$

Differentiating (7) twice and setting $s = 1$ gives

$$v_\rho = \frac{\rho}{(1-\rho)^2}\{1-\rho[\xi_\rho'^2(1)+4\xi_\rho^2(1)+4\xi_\rho(1)\,\xi_\rho'(1)]-(1-\rho)[\xi_\rho''(1)+5\xi_\rho'(1)+4\xi_\rho(1)]\}. \tag{13}$$

5. Some examples

(i) *Binomial queue*

Condition (5) is satisfied by the finite distribution

$$p(x) = {}^nC_x\pi^x(1-\pi)^{n-x} \quad (x = 0, 1, \ldots, n),$$

where $0 < \pi < 1$. We find

$$\rho = \frac{n\pi}{1-\pi}, \quad 0<\rho<\infty; \qquad m_\rho = \frac{n\rho}{n+\rho} \to n \quad \text{as} \quad \rho\to\infty.$$

The corresponding balking distribution is given by

$$G(x) = \begin{cases} \dfrac{n-x-1}{n(x+2)} & 0 \leqslant x \leqslant n-1, \\ 0 & n \leqslant x, \end{cases} \qquad f(x) = \begin{cases} \dfrac{n+1}{n}\dfrac{1}{(x+1)(x+2)} & 0 \leqslant x \leqslant n-1 \\ 0 & n \leqslant x \end{cases}$$

and is independent of π. As $\rho\to\infty$, $\pi\to 1$ and the queue is almost always of length just n. The relations (8) to (11) may be verified immediately:

$$p_\rho(0) = \left(\frac{n}{n+\rho}\right)^n, \quad \eta_\rho(s) = \left(\frac{n+\rho s}{n+\rho}\right)^n, \quad v_\rho = \rho\left(\frac{n}{n+\rho}\right)^2.$$

(ii) *Negative binomial queue*

An example of an infinite queue-length distribution is

$$p(x) = {}^{N+x-1}C_{N-1}\chi^x(1+\chi)^{-N-x} \quad (x = 0, 1, 2, \ldots),$$

where $\chi > 0$ and $N > 1$. We have

$$\lim_{x\to\infty}\frac{p(x+1)}{p(x)} = G(\infty) = \frac{1}{N},$$

$$\rho = \frac{N\chi}{1+\chi}, \quad 0<\rho<N; \qquad m_\rho = \frac{N\rho}{N-\rho}\to\infty \quad \text{as} \quad \rho\to N.$$

The balking distribution is

$$G(x) = \frac{N+x+1}{N(x+2)}, \quad f(x) = \frac{N-1}{N}\frac{1}{(x+1)(x+2)} \quad (x = 0, 1, 2, \ldots),$$

with a probability $1/N$ at $x = \infty$. The balking distribution is independent of χ:

$$p_\rho(0) = \left(\frac{N-\rho}{N}\right)^N, \quad \eta_\rho(s) = \left(\frac{N-\rho}{N-\rho s}\right)^N, \quad v_\rho = \rho\left(\frac{N}{N-\rho}\right)^2.$$

(iii) *Poisson queue*

Each of the queue distributions in (i) and (ii) approaches the Poisson form as n, $N \to \infty$:

$$p(x) = \frac{\rho^x e^{-\rho}}{x!} \quad (x = 0, 1, 2, \ldots).$$

We now have $G(\infty) = 0$, so $0 < \rho < \infty$; $m_\rho = \rho = v_\rho$. The balking distribution becomes

$$G(x) = \frac{1}{x+2}; \quad f(x) = \frac{1}{(x+1)(x+2)} \quad (x = 0, 1, 2, \ldots)$$

which may be regarded as a discrete analogue of the Cauchy distribution

$$p_\rho(0) = e^{-\rho}, \quad \eta_\rho(s) = e^{-\rho+\rho s}.$$

(iv) *Type III ordinates*

Another possible infinite queue-length distribution is

$$p(x) = A(x+a)^\nu e^{-\lambda x} \quad (x = 0, 1, 2, \ldots),$$

where a, ν, $\lambda > 0$; A is a complicated function of a, ν and λ. Here

$$G(\infty) = \left(\frac{a}{a+1}\right)^\nu, \quad \rho = \left(\frac{a+1}{a}\right)^\nu e^{-\lambda}, \quad \text{so} \quad 0 < \rho < \left(\frac{a+1}{a}\right)^\nu.$$

The balking distribution is

$$G(x) = \left[\frac{a(a+x+2)}{(a+1)(a+x+1)}\right]^\nu \quad (x = 0, 1, 2, \ldots)$$

which is independent of λ. For ρ near $(a+1)^\nu/a^\nu$, we have approximately

$$m_\rho \simeq \frac{\nu+1}{\log\left[\frac{1}{\rho}\left(\frac{a+1}{a}\right)^\nu\right]} - a, \quad v_\rho \simeq \frac{(m_\rho + a)^2}{\nu+1}.$$

If we let $\nu \to 0$ we obtain the classic case where every arrival must join the queue

$$p(x) = \frac{e^{-\lambda x}}{1-e^{-\lambda}}, \quad G(\infty) = 1, \quad \rho = e^{-\lambda}, \quad 0 < \rho < 1.$$

(v) *Normal ordinates*

A very simple result is obtained when we assume that the queue-length distribution is

$$p(x) = A \exp - \frac{(x-m)^2}{2v} \quad (x = 0, 1, 2, \ldots),$$

where m and v are very nearly the mean and variance of the distribution if v and m^2/v are 'large', say both > 9. We find

$$G(\infty) = 0, \quad \rho = \exp(m - \tfrac{1}{2})/v, \quad 0 < \rho < \infty.$$

The balking distribution is Pascal (geometric)

$$G(x) = \exp - (x+1)/v, \quad f(x) = (1-\lambda)\lambda^x \quad (x = 0, 1, 2, \ldots), \tag{14}$$

where $\lambda = \exp v^{-1}$, and is independent of m. Denoting the mean of this distribution by $M = \lambda/(1-\lambda)$, we have

$$m_\rho \simeq m = \tfrac{1}{2} + v\log\rho, \quad v_\rho \simeq v = \left[\log\left(\frac{M+1}{M}\right)\right]^{-1}. \tag{15}$$

The relation $m_\rho = \text{const.} + v_\rho \log\rho$ cannot hold exactly for any distribution whatever, as from (11) it implies first $v_\rho = c = \text{const.}$ and then $m_\rho < 0$ for $\rho < \exp -(c/v)$.

(vi) *Deterministic balking*

One case in which these calculations are easy is the following: each person possesses exactly the same degree of queue resistance, i.e. the balking distribution is deterministic, $f(x) = 1$ for $x = K$ and zero otherwise. Although rather trivial, this case will be important in an application to be mentioned subsequently. We have

$$G(x) = \begin{cases} 1 & 0 \leqslant x \leqslant K-1, \\ 0 & K \leqslant x, \end{cases} \qquad c_x = \begin{cases} 1 & 0 \leqslant x \leqslant K+1, \\ 0 & K+2 \leqslant x, \end{cases}$$

so

$$p_\rho(0) = \frac{1-\rho}{1-\rho^{K+2}}; \qquad p(x) = \begin{cases} \dfrac{1-\rho}{1-\rho^{K+2}}\rho^x & 0 \leqslant x \leqslant K+1, \\ 0 & K+2 \leqslant x. \end{cases}$$

From (8) and (9), or direct from the definitions, we have

$$\xi_\rho(s) = \frac{1-\rho}{1-\rho^{K+2}}\rho^{K+1}s^K, \quad \eta_\rho(s) = \frac{1-\rho}{1-\rho^{K+2}}\,\frac{1-(\rho s)^{K+2}}{1-\rho s}.$$

Hence, from (12) and (13), or from (10) and (11), we have

$$m_\rho = \frac{\rho}{1-\rho} - (K+2)\frac{\rho^{K+2}}{1-\rho^{K+2}},$$

$$v_\rho = \frac{\rho}{(1-\rho)^2} - (K+2)^2\frac{\rho^{K+2}}{(1-\rho^{K+2})^2}.$$

In the trivial situation where each individual insists on immediate service, $K = 0$, and

$$m_\rho = \frac{\rho}{1+\rho}, \quad v_\rho = \frac{\rho}{(1+\rho)^2}.$$

In the classic situation, in which each individual must join the queue, $K = \infty$, and

$$m_\rho = \frac{\rho}{1-\rho}, \quad v_\rho = \frac{\rho}{(1-\rho)^2}.$$

Each of (i) to (v) above has led to a balking distribution which is unimodal with mode at $x = 0$. It seems very difficult to obtain explicit results for a balking distribution with mode not zero; numerical calculations were carried out for $\rho = 2, 3, 4, 5, 6, 7$, for each of the following balking distributions:

(vii) Poisson with mean 10.

(viii) Ordinates of a lognormal distribution $f(y)$ where $5(\log y - 1)$ is a unit normal variable.

(ix) Ordinates of a χ^2 distribution with 10 degrees of freedom.

(x) The Pascal distribution (14) with $\lambda = 10/11$.

Table 1 shows the means and variances of the queue-length distributions obtained, together with those of the 'input' balking distributions. Also, in the last column, are given the approximate values for the Pascal distribution (x) obtained from (15). It will be seen that the approximation is very close except when $\rho = 2$; here m^2/v is only 5·5.

Table 1. *Mean and variance of queue length for various balking distributions*

	Poisson	Lognormal	χ^2	Pascal	Formulae (15)
Input					
Mean	10	10·57	9·498	10	10
Variance	10	19·5384	19·23	110	110
$\rho=2$					
Mean	10·320206	10·212608	9·475843	7·823407	7·772564
Variance	4·908937	6·506500	6·110196	9·726074	10·49205
$\rho=3$					
Mean	11·941052	12·578738	11·633753	12·026812	12·026641
Variance	3·351061	5·382238	4·725268	10·467543	10·49205
$\rho=4$					
Mean	12·826379	14·078101	12·914084	15·041116	15·045024
Variance	2·844219	5·074961	4·248282	10·490947	10·49205
$\rho=5$					
Mean	13·430370	15·192921	13·830718	17·363218	17·386325
Variance	2·582158	4·929776	4·032884	10·463681	10·49205
$\rho=6$					
Mean	13·885518	16·086535	14·547050	19·284075	19·299236
Variance	2·416736	4·858583	3·839912	10·533106	10·49205
$\rho=7$					
Mean	14·248864	16·831880	15·130198	20·919959	20·916585
Variance	2·301017	4·813526	3·728680	10·485757	10·49205

6. SOME GENERALIZATIONS

An interesting application of queueing with balking is furnished by the problem of a sequence of transporting mechanisms which move discrete units of cargo. In the terminology developed by the Department of Engineering of the University of California at Los Angeles (1953), each transporting agency constitutes a 'link', and the place where two such mechanisms transfer their loads is a 'node'. If there is room at a node for the storage of S items, the number in storage at time t may be regarded as a queue of length $0, 1, \ldots, S$. If the link setting down items carries A_1 units at a time, and the link picking up items carries A_2 at a time, then we must assume $S \geqslant A_1 + A_2 - 1$, to prevent the process stopping altogether. The time required for the journeys of the links (in which we absorb the time required for pick up or set down) constitute the inter-arrival and inter-departure times of the queue at the node between them. If the storage $k(t)$ at the node is such that the arriving link cannot set down its cargo, i.e. if $S - k(t) < A_1$, the link goes back and arrives again after another inter-arrival time. If the storage when removing link arrives satisfies $k(t) < A_2$, so that a full load is not available, it also goes back and forth until the inequality is reversed.

Thus, considering only two links separated by one node, the following generalizations of the probabilistic queueing model discussed in our introduction are suggested: (a) bulk arrivals and departures, (*b*) balking distributions associated with both arrivals and departures, (*c*) a finite number of states possible. Also, the cargo-handling case requires deterministic balking; we shall begin with stochastic balking, and then specialize for cargo-handling.

Let K_1 and K_2 be two integers such that when A_1 items arrive, they are added to the queue when $k(t) \leqslant K_1$ and not otherwise, and when the removal mechanism arrives, it takes away A_2 items if $k(t) \geqslant K_2$ and not otherwise. Let $F_j(x) = \Pr(K_j \leqslant x) = 1 - G_j(x)$.

Using the same argument that was employed in § 2,

$$\begin{aligned} p(x, t+\Delta t) = {} & (1-\lambda\Delta t - \mu\Delta t)\, p(x,t) \\ & + \lambda\Delta t(1-\mu\Delta t)\,[p(x,t)\, F_1(x-1) + p(x-A_1, t)\, G_1(x-A_1-1)] \\ & + \mu\Delta t(1-\lambda\Delta t)\,[p(x,t)\, G_2(x) + p(x+A_2, t)\, F_2(x+A_2)] + o(\Delta t), \end{aligned}$$

where departures as well as arrivals are now partitioned into two cases. Also, we have special cases not only for $x = 0$, but for x less than A_1 or A_2. However, this can be dealt with by conventions regarding negative arguments. We put

$$\begin{aligned} p(x) &= 0 \quad \text{for} \quad x<0 \quad \text{or} \quad x>S, \\ F_1(x) &= 0 \quad \text{for} \quad x<0, \\ F_2(x) &= 0 \quad \text{for} \quad x \leqslant 0. \end{aligned}$$

Passing to the limit, and letting $t \to \infty$ (assuming equilibrium) we obtain

$$0 = \rho G_1(x-A_1-1)\, p(x-A_1) + \{\rho G_1(x-1) + F_2(x)\}\, p(x) - F_2(x+A_2)\, p(x+A_2) \qquad (x = 0, 1, \ldots, S).$$

These $S+1$ equations are linearly dependent; there is the further relation $\Sigma p(x) = 1$. The system of equations may be written $\mathbf{Bp} = \mathbf{a}$, where

$$\mathbf{p} = \{p(0), p(1), \ldots, p(S)\}, \quad \mathbf{a} = \{0, 0, \ldots, 1\}$$

and $\mathbf{B}$ is an $(S+1)\times(S+1)$ matrix having zeros everywhere except in the last row (where all the elements are unity), the principal diagonal, the A_2th super-diagonal, and the A_1th subdiagonal.

If $A_1 = A_2 = 1$, these equations can be solved exactly as in § 3; only a redefinition of c_x is needed

$$c_x = \prod_0^{x-2} G_1(j) \Big/ \prod_1^{x} G_2(j).$$

In case of arbitrary A_1 and A_2, but with deterministic balking defined by

$$K_1 = S - A_1, \quad K_2 = A_2$$

(which are the values suggested by the cargo-handling problem), the matrix $\mathbf{B}$ simplifies as follows:

(*a*) the elements in the A_1th subdiagonal become $-\rho$,
(*b*) the elements in the A_2th super-diagonal become -1,
(*c*) the principal diagonal consists of $-A_2$ elements of value ρ, followed by

$$S - A_1 - A_2 + 1 \geqslant 0$$

elements of value $1+\rho$ and, finally, A_1 elements of value 1.

I have evaluated this determinant only in special cases, but offer the following conjectures about the polynomial $B(\rho; A_1, A_2, S)$:

(i) it vanishes unless A_1 and A_2 are relatively prime;

(ii) if A_1 and A_2 are relatively prime, it consists of $S-A_1-A_2+3$ terms, beginning with a term of degree $S-A_1+1$ and ending with a term of degree A_2-1;

(iii) $B(\rho; m, n, S) = \rho^S B(\rho^{-1}; n, m, S)$;

(iv) the first and last coefficients are A_1 and A_2, respectively;

(v) $B(\rho; 1, 1, S) = \sum_{j=0}^{S} \rho^j$.

I wish to thank Drs F. N. David, D. G. Kendall, C. L. Mallows and R. Bellman for their helpful suggestions.

REFERENCES

BROCKMEYER, E., HALSTRØM, H. L. & JENSEN, A. (1948). *The Life and Works of A. K. Erlang.* Copenhagen Telephone Company.

FELLER, W. (1950). *An Introduction to Probability Theory and its Applications*, Vol. I. New York: John Wiley and Sons.

KAWATA, T. (1955). A problem in the theory of queues. *Rep. Statist. Appl. Res. Un. Jap. Sci. Engrs*, **3**, 122–9.

KENDALL, D. G. (1951). Some problems in the theory of queues. *J. R. Statist. Soc.* B, **13**, 151–73.

KENDALL, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *Ann. Math. Statist.* **24**, 338–54.

KENDALL, D. G. & REUTER, G. E. H. (1957). The calculation of the ergodic projection for Markov chains and processes with a countable infinity of states. *Acta Mathematica*, **97**, 103–44.

KOLMOGOROFF, A. (1932). Sur la problèm d'attente. *Math. Sbornik.* **38**, 101.

KOOPMAN, B. O. (1953). New mathematical methods in operations research. *J. Opns. Res. Soc. of America*, **1**, 3–9.

LINDLEY, D. V. (1952). The theory of queues with single server. *Proc. Camb. Phil. Soc.* **48**, 277–89.

TAKÁCS, L. (1955). Investigation of waiting time problems by reduction to Markov processes. *Acta Math. Acad. Sci. Hungary*, **6**, 101–29.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES, DEPT. OF ENGINEERING (1953). *An Engineering Analysis of Cargo Handling.* Report 53–21.