A CONTINUOUS-TIME TREATMENT OF CERTAIN QUEUES AND INFINITE DAMS

R. M. LOYNES

(received 9 December 1961)

Summary

The continuous-time behaviour of a model which represents certain queues and infinite dams with correlated inputs is considered. It is shown how the transient behaviour may be investigated, and the asymptotic behaviour is obtained. Finally the methods are illustrated for a queue whose input consists of two superimposed renewal processes.

1. Introduction

In a previous paper (Loynes 1961b), the stationary waiting-time distribution for certain single-server queues with non-independent inter-arrival or service times has been investigated: with the terminology used there it was shown that in the 'finite matrix' case, the derivation of this distribution is straightforward if either the inter-arrival time or the service time has a rational conditional characteristic function. Here we shall give a continuous-time treatment of certain systems, including many of those 'finite matrix' queues whose inter-arrival time distribution is of this type, which will show how we may find the distribution of the waiting-time and of the busy period at any time after starting from arbitrary initial conditions. The particular queue $E_k/G/1$ has in fact been investigated in essentially the same way by Takács (1961), although in order to be able to treat the general case, the detailed methods used are different.

The method used will be that originated by Takács (1955) and studied later by Beneš (1957) (whom we shall follow quite closely), in which the basic variable is what we shall call the potential waiting-time at time t: this is the length of time necessary to complete the service of those in the queue at that time.

The analogy with an infinite dam, due to Smith (1953), allows a re-interpretation of the equations as a description of the behaviour of a dam with a (possibly) correlated input.

https://doi.org/10.1017/S1446788700027488 Published online by Cambridge University Press

2. The Model and its Various Interpretations

We suppose that controlling the system (probabilistically) there is a stochastic process $z(t)$ and events associated with it. The possible values of $z(t)$ are finite in number, and we shall for obvious reasons refer to the value at any time t as the state occupied. The sequence of events, when the state occupied is known, forms a Poisson process whose parameter depends only on that state. When an event occurs, the state of $z(t)$ immediately after is chosen in accordance with the transition probabilities of an irreducible Markov chain (in discrete time). The process $z(t)$ clearly forms a finite homogeneous Markov chain in continuous time, and if the state immediately after an event cannot be the same as that immediately before, these events of which we are talking are merely the moments of transition of the $z(t)$ chain. It is necessary for our purposes to allow the state to be the same immediately before and after an event, but we shall nevertheless refer to such events as transitions.

The variable $W(t)$, equal to the potential waiting-time or the dam content, has non-negative jumps S at the transitions of $z(t)$, the distribution of the magnitude of the jump given the state occupied immediately before and after the transition being independent of past history, and otherwise decreases at a rate $c[z(t)]$, subject to the presence of an impenetrable barrier at $W = 0$.

The pair $(W(t), z(t))$ is itself a Markov process, and we shall use this property to carry out the investigation of the system in the following sections. Here we merely point out various situations which can be represented by giving suitable values to the parameters.

If we make $z(t)$ a cyclic chain (i.e. one in which the successive states occupied form a deterministic sequence), and make S zero except at one transition, then we can clearly build up any inter-arrival time distribution with a characteristic function which is the reciprocal of a polynomial with (negative) real zeroes. By allowing complex rates of transition, giving rise to complex probabilities (cf. Cox (1955)), and non-cyclic chains, it is possible to construct a large class of rational characteristic functions, and successive service times S need only be conditionally independent given the values of $z(t)$ at the transition, rather than completely independent. The same construction can also be used to treat certain queues with bulk service, for we may merely change the state cyclically every time a customer arrives, but not add in the service time until the whole group is present; this idea was applied in the reverse direction by Takács (1961). The example we give in section 7 arises in a slightly different way, but has in common with the previous examples the feature that one may introduce fictitious states (after Erlang) to obtain the structure already postulated.

If the function $c(\cdot)$ is unity for all states, then service continues whenever there is someone in the queue. By allowing $c(\cdot)$ to be zero for certain states, the phenomenon of the server stopping work for a random length of time could be simulated, though with the present model customer arrivals would be inhibited (or at least affected) by this.

Turning now to the infinite dam interpretation, we have considerable freedom of choice. The inputs S can always be considered as 'cloudbursts' of (possibly) correlated amounts, occurring at points which are not necessarily those of a Poisson process. The possibility of $c(\cdot)$ having different values in different states can be regarded in two ways. Either the rate at which water is withdrawn is made to depend on the weather, which is then represented by the chain $z(t)$, or $c(\cdot)$ is supposed to represent the difference between the rate at which water is being withdrawn and the rate at which rain is entering the dam: this continuous entry of rain might well represent a prolonged shower, or even drainage after a short one, quite adequately. With this latter interpretation there is no reason why $c(\cdot)$ should be positive.

3. The Basic Equations

We suppose for convenience that the possible values of $z(t)$ are the integers 1 to k. The pair $\{W(t), z(t)\}$ forms a Markov process, and the forward equations can be written down directly. The result is an integro-differential equation for $p[W(t) \leqq x; z(t) = y]$ similar to that of Takács, but in order to solve this we shall take Laplace-Stieltjes transforms with respect to x; since it is as easy to write down the equation for the transform (and easier to prove the necessary differentiability properties), we shall not concern ourselves with the probabilities, but shall begin with the transforms.

Denote by λ_j the transition rate for $z(t)$ when $z(t) = j$, and by p_{ij} the conditional probability of a transition from state j to state i, given that a transition has taken place. Let the value of $c[z(t)]$ when $z(t) = j$ be c_j, and (when a transition is known to have occurred at time t) let

$$\beta_{ij}(\theta) = E\{e^{-\theta S} \,|z(t-) = j, z(t+) = i\}$$

for $R\theta \geqq 0$. Finally, let

$$\Phi(\theta, y; t) = E\{e^{-\theta W(t)} I[z(t) = y]\} = \int e^{-\theta W(t)} dp[W(t); z(t) = y] \tag{1}$$

for θ with $R\theta \geqq 0$, where $I[A]$ is the characteristic function of the set A.

Let n be the number of transitions in $(t, t+\Delta t)$, and c_i^+ the positive part of c_i. Then for $\Delta t > 0$,

$$\Phi(\theta, y; t+\Delta t) = \sum_{i,n} E\{e^{-\theta W(t+\Delta t)} I[z(t+\Delta t) = y, z(t) = i, n, W(t) > c_i^+ \Delta t]\} + 2 \text{ further terms} \tag{2}$$

The remaining terms are obtained from the first by replacing the conditions on $W(t)$ by $c_i^+\Delta t \geqq W(t) > 0$, and $W(t) = 0$, respectively; the first of these does not occur if $c_i \leqq 0$.

Now the parts into which the right side of (2) is split are straightforward to evaluate in terms of known conditional probabilities. For we know that

$$W(t+\Delta t) = [W(t) - c_i\Delta t]^+ \tag{3}$$

whenever $n = 0$ and $z(t) = i$, and

$$W(t+\Delta t) = W(t) - \alpha c_i \Delta t + S \tag{4}$$

when $n = 1$ and $z(t) = i$, where $0 \leqq \alpha \leqq 1$. The operation of taking the positive part in (3) can be omitted, and α in (4) set equal to unity, when $W(t) \geqq c_i^+\Delta t$. Furthermore, such probabilities as

$$p[n = 0 \,|z(t) = i] = 1 - \lambda_i \Delta t + o(\Delta t) \tag{5}$$

are known.

After evaluating and rearranging the terms, dividing by Δt, and letting Δt tend to zero, we find that $\Phi(\theta, y; t)$ has a right-hand derivative everywhere with respect to t, given by

$$\frac{\partial\Phi(\theta, y; t)}{\partial t} = -\lambda_y\Phi(\theta, y; t) + \theta c_y \Phi(\theta, y; t) - \theta c_y^+ p[W(t) = 0; z(t) = y] + \sum_i \lambda_i p_{yi} \beta_{yi}(\theta)\Phi(\theta, i; t). \tag{6}$$

The left-hand derivative gives more trouble, but we can deal with it in the following way. By applying the same analysis to the interval $(t - \Delta t, t)$ it can be seen that $\Phi(\theta, y; t)$ is continuous on the left, and that its upper and lower left derivatives are bounded. Then according to Theorem 34.5 of McShane (1947) Φ satisfies a Lipschitz condition (in t), and is hence absolutely continuous, which is sufficient for our purposes, for it is then the integral of its derivative, which exists almost everywhere. If we showed that $p[W(t) = 0; z(t) = y]$ is continuous everywhere then we should know directly from (6) (Titchmarsh (1939) § 11.3 Ex (vi)) that Φ is differentiable everywhere.

It is natural and convenient to rewrite (6) in matrix form with the following notation:

https://doi.org/10.1017/S1446788700027488 Published online by Cambridge University Press

$\Phi = \Phi(\theta, t)$ is the column vector formed by $\Phi(\theta, y; t)$; $C = \operatorname{diag}(c_i)$, and $C^+ = \operatorname{diag}(c_i^+)$; $\Lambda = \operatorname{diag}(\lambda_i)$; $Q = Q(\theta) = [q_{ij}]$, where $q_{ij} = p_{ij}\beta_{ij}(\theta)$; $F = F(t)$ is the column vector formed by $p[W(t) = 0;\ z(t) = y]$ and

$$A = A(\theta) = \theta C + (Q(\theta) - I)\Lambda \tag{7}$$

The result is

$$\frac{\partial \Phi}{\partial t} = A\Phi - \theta C^+ F; \tag{8}$$

if we take Laplace transforms with respect to t we have finally

$$(sI - A)\Phi^* = \Phi(\theta, 0) - \theta C^+ F^* \tag{9}$$

where

$$\Phi^* = \int_0^\infty e^{-st} \Phi(\theta, t) dt \tag{10}$$

and

$$F^* = \int_0^\infty e^{-st} F(t) dt, \tag{11}$$

all the transforms existing when $Rs > 0$.

We observe that the number of components of F actually contributing to (8) or (9) is the same as the number of positive c_i.

In order to investigate the busy period, or the wet period in a dam, we introduce the process $W'(t)$, which is formed by 'freezing' $W(t)$ at zero as soon as it reaches that value. If q is the busy period, or the time at which $W(t)$ first reaches zero,

$$\begin{aligned} W'(t) &= W(t) \qquad (t < q) \\ &= 0 \qquad\qquad (t \geqq q) \end{aligned} \tag{12}$$

Similarly, let

$$\begin{aligned} z'(t) &= z(t) \qquad (t < q) \\ &= z(q) \qquad (t \geqq q) \end{aligned} \tag{13}$$

so that $z(t)$ is also 'frozen' as soon as $W(t)$ reaches zero. Then

$$p[W'(t) = 0; z'(t) = y] = p[q \leqq t; z'(q) = y]. \tag{14}$$

For

$$\Psi(\theta, y; t) = E\{e^{-\theta W'(t)} I[z'(t) = y]\} \tag{15}$$

and the corresponding vector $\Psi(\theta, t)$, with Laplace transform Ψ^*, we find in a similar way the equations

$$\frac{\partial \Psi(\theta, y; t)}{\partial t} = \theta c_v \Psi(\theta, y; t) - \lambda_v \Psi(\theta, y; t) + \sum_i \lambda_i p_{vi} \beta_{vi} \Psi(\theta, i; t) \tag{16}$$
$$+ (\lambda_v - \theta c_v) p[W'(t) = 0; z'(t) = y] - \sum_i \lambda_i p_{vi} \beta_{vi} p[W'(t) = 0; z'(t) = y]$$

$$\frac{\partial \Psi}{\partial t} = A\Psi - AG \tag{17}$$

and

$$[sI - A]\Psi^* = \Phi(\theta, 0) - AG^* \tag{18}$$

where G is the vector formed by $p[W'(t) = 0; z'(t) = y]$.

It is obvious that $p[W'(t) = 0;\ z'(t) = y] = 0$ if $c_v \leqq 0$, so that the number of unknowns other than Ψ in (17) or (18) is just the number of positive c_i.

If we do not freeze $z(t)$, then for

$$\Psi_1(\theta, y; t) = E\{e^{-\theta W'(t)} I[z(t) = y]\} \tag{19}$$

we have similarly

$$\frac{\partial \Psi_1(\theta, y; t)}{\partial t} = (\theta c_v - \lambda_v)\Psi_1(\theta, y; t) + \sum_i \lambda_i p_{vi} \beta_{vi} \Psi_1(\theta, i; t) \tag{20}$$
$$-\theta c_v p[W'(t) = 0; z(t) = y]$$
$$+ \sum_i \lambda_i p_{vi}[1 - \beta_{vi}] p[W'(t) = 0; z(t) = i]$$

$$\frac{\partial \Psi_1}{\partial t} = A\Psi_1 - [\theta C + (Q - P)A]G_1 \tag{21}$$

(writing $P = (p_{ij})$), and

$$[sI - A]\Psi_1^* = \Phi(\theta, 0) - [\theta C + (Q - P)A]G_1^* \tag{22}$$

but these are probably of less importance than (16), (17), or (18).

Our restriction that the chain be finite is unnecessary for the derivation of the above equations, but will become necessary for the discussion of their solution in the following sections. For the particular queue $M/G/1$, it is possible to use an infinite chain to investigate the number of customers in a busy period, since we may increase the number of the state by unity each time a customer arrives, and then the quantity $z'(t)$ is exactly the number of customers served before the server first becomes idle. This does not seem possible for other queues.

4. Two Lemmas

The equations derived in the previous section all contain two unknown functions; for instance (8) involves both Φ and F. In order to use these

https://doi.org/10.1017/S1446788700027488 Published online by Cambridge University Press

equations we need some way of finding one of the unknowns, and we shall see below that this is possible in principle under certain conditions.

For this purpose we shall prove two similar Lemmas, in both of which it will be found necessary to assume that all λ_i are real; but it seems quite likely that the results hold without this restriction. It is a consequence of this restriction that any problem in which we wished to introduce complex probabilities would need further investigation, and possibly a different approach entirely.

LEMMA 1. If all λ_i are real, then for θ sufficiently close to the origin on the imaginary axis, the eigenvalues of $A(\theta)$ have a strictly negative real part, unless C is a multiple of I and $|\beta_{ij}| = 1$ for all i, j.

LEMMA 2. If all λ_i are real, and all c_i are positive, then for s with positive real part the matrix $sI - A(\theta)$ is singular for k values of θ inside the right half plane.

The further restriction in Lemma 2, that all c_i should be positive, is unfortunate, but cannot be avoided. For suppose that all $\beta_{ij} = 1$, and that $P = I$. Then if some c_i are not positive, there are certainly not singularities of $sI - A$ for k values of θ in the right half plane. Strictly speaking this counter-example is not valid, as the matrix P is not irreducible, but small perturbations will clearly remedy this without affecting the number of singularities.

It is possible, however, that even if there are $c_i \leqq 0$ there are still enough singularities to determine the unknowns, for we have already noted that the number of unknowns occurring in (8) or (17) is equal to the number of c_i which are positive. This does not appear to be true for (21).

To prove these Lemmas, we first locate the eigenvalues of A in the usual way (as, for instance, in Bartlett (1956) p. 52). If μ is an eigenvalue of A, there is a left eigenvector (t_i) of $A - \mu I$, and if the component of maximum absolute value is t_m, then we have

$$|t_m|\,|\mu - a_{mm}| \leqq \sum_{j \neq m} |t_j|\,|a_{jm}| \leqq \lambda_m\,|t_m|(1 - p_{mm}) \tag{23}$$

whenever $R\theta \geqq 0$, where

$$a_{mm} = \theta c_m + (q_{mm} - 1)\lambda_m. \tag{24}$$

For Lemma 1, we assume θ purely imaginary. It then follows from (23) and (24) that either the real part of μ is negative, or $\mu = \theta c_m$, which can only be true if $|q_{jm}| = p_{jm}$ and $|t_j| = |t_m|$ for j such that $p_{jm} \neq 0$. It is therefore possible to apply (23) for these particular j, and finally for all j, since the chain is irreducible. Thus C must be a multiple of I, and $|\beta_{ij}| = 1$ for all j, i, so that Lemma 1 is proved. It is in fact not difficult to show

https://doi.org/10.1017/S1446788700027488 Published online by Cambridge University Press

that a necessary and sufficient condition for some eigenvalue of A to have zero real part is that C is a multiple of I and

$$\beta_{ij} = t_j/t_i. \tag{25}$$

In this case the analysis of (8) is straightforward.

To prove Lemma 2, we first observe that it follows from (23) and (24) that μ lies inside the circle with centre $\theta c_m - \lambda_m$ and radius λ_m, for some m. The matrix $sI - A$ is singular if and only if

$$s = \mu$$

or

$$s + y - b\theta = \mu(\theta) + y - b\theta \tag{26}$$

where we choose b larger than the maximum of the c_i, and y is positive and to be chosen later. The right side of (26) can be written as $y - \lambda_m - (b - c_m)\theta + \lambda_m v$, where $|v| \leqq 1$, for some m. We can show easily that on any large semi-circle in the right half plane the modulus of the left side of (26) is greater than that of the right side, for one is approximately $b|\theta|$ and the other $(b - c_m)|\theta|$, and we have supposed $0 < b - c_m < b$. Let us now close the semi-circle by a straight line parallel to the imaginary axis on which

$$0 < bR(\theta) < Rs. \tag{27}$$

The difference between the square of the modulus of the left side and the square of that of the right side on such a line will be found to involve $I(\theta)$ quadratically with a positive coefficient, so that it is bounded below as θ varies along the line. The coefficient of y will be found to be positive, so that if y is chosen sufficiently large this difference will be everywhere positive.

An argument of Rouché type, similar to that given in Loynes (1961b), may now be applied to show that (26) has k roots inside the contour for every zero that the left side has there, and Lemma 2 follows.

In the next section we shall require that the matrix X, whose rows are the left eigenvectors of $sI - A$ when θ takes the k values assured by Lemma 2, be non-singular. We can show that except for isolated points s this is so, unless special conditions are satisfied, in the following way. First we observe that $sI - A$ is of rank $k - 1$ when $\theta = \theta_i(s)$, for otherwise A has two equal eigenvalues for this value of θ, and this cannot be true for other than isolated values of θ (and correspondingly isolated values of s) unless the discriminant of the characteristic equation of A vanishes identically, which clearly does not happen in general. It follows that we may use the cofactors of a column of $sI - A(\theta_i)$ as the elements of the eigenvector, and if the determinant of X vanishes, this gives a relationship between the functions $\theta_i(s)$ (and s), which is either true identically, or is only true for isolated values of s. Since it is easy to construct examples for which this does not happen identically, the result is proved for matrices in general. Even if the discriminant of the characteristic equation of A does vanish identically, so that $sI - A$ always has two among the $\theta_i(s)$ coincident, there may still be enough independent vectors to make X non-singular; this situation occurs in the example in section 7.

5. The Solution of the Equations

As remarked above, some method of finding the unknown F occurring in (8), and similarly G and G_1 in (17) and (21) respectively, is needed. If F is known, then (8) can be solved to give

$$\Phi(\theta, t) = e^{At}\Phi(\theta, 0) - \theta \int_0^t e^{A(t-u)}C^+F(u)du. \tag{28}$$

Similarly for (17), though it is usually the unknown G that is of interest, rather than Ψ.

By using Lemma 2, we can in fact show how to determine F^*, in principle, merely by using the analytic properties of Φ, and thus by inverting the transform (11), F itself. To do this we must and therefore shall throughout this section assume that the conditions of Lemma 2 are satisfied; we may then replace C^+ by C.

For given s, there are therefore k values of θ, say $\theta_i(s)$, for which $sI-A$ is singular, and corresponding to these, k left eigenvectors $X'_i(s)$. Putting these values of θ into (9), and multiplying on the left by the appropriate X'_i, we find

$$X'_i\Phi(\theta_i, 0) = \theta_i X'_i CF^* \qquad (1 \leqq i \leqq k). \tag{29}$$

If we write $\Gamma(s)$ for the vector whose components are $X'_i\Phi(\theta_i, 0)/\theta_i$, and X for the matrix whose rows are X'_i, we can solve (29) in the form

$$CF^* = X^{-1}\Gamma, \tag{30}$$

where, as we have remarked at the end of section 4, the inverse exists in general. This can be solved for F^* if required, but it would not normally be necessary, since only CF is needed for use in (28).

We can perform a similar analysis on (18), and using (14) to express $B(s)$, with components

$$B(s, y) = \int_0^\infty e^{-su}d_u p[q \leqq u; z'(q) = y] \tag{31}$$

in terms of G^*, we finally obtain

$$B(s) = sG^* = X^{-1}\Delta \tag{32}$$

where Δ has components $X_i'\Phi(\theta_i, 0)$. $B(s)$ is just the Laplace-Stieltjes transform of the busy-period distribution, conditional on the state of $z(t)$ in which the busy-period ends, multiplied by the probability that it ends in that particular state.

If the matrix X is not invertible, there seems no choice but to invert the matrix $(sI - A)$, leaving the components of F as undetermined constants, and then choose these components to ensure the analyticity of $\Phi^*(\theta)$ in the right half plane. They are certainly determined uniquely in this way in some examples for which the approach leading to (30) does not succeed.

6. The Asymptotic Behaviour

Throughout this section we suppose the conditions of Lemma 1 satisfied; the values of c_i are not, however, restricted as they were in the previous section. We show first that as t tends to ∞ the vector F in (8) converges to a limit which is independent of initial conditions. This is easily done in the way outlined by Smith (1955), which is to map the state $\{W = 0, z = y\}$ onto one state of a two-state Semi-Markov chain, and all other possible values of $\{W, z\}$ onto the other. According to Smith's Theorem 5, $p[W(t) = 0; z(t) = y]$ tends to a limit as t tends to ∞, the fact that a transition may not have occurred at $t = 0$ being easily seen not to affect the result. The expression for the limit given by Smith only depends on initial conditions through the probability that the state $\{W = 0; z = y\}$ ever occurs. If we first consider all sets of initial conditions in which $W(0)$ takes arbitrary values but $z(0)$ takes a single fixed value, then the inequalities

$$0 \leqq W_1(t) - W_0(t) \leqq W_1(0) - W_0(0) = W_1(0),$$

in which the suffix 1 refers to any fixed value of $W(0)$ and the suffix 0 refers to the particular case $W(0) = 0$, are easily obtained. From this it follows that if $W_0(t)$ never reaches zero with $z(t) = y$, neither does $W_1(t)$. Conversely, if $W_0(t)$ does reach zero with $z(t) = y$, with probability one, it does so infinitely often: $W_1(t)$ then reaches the same state provided that the residual lifetime in state y is longer than $W_1(0)/c_y$ (assuming $c_y > 0$), and this clearly happens with probability one at some time. Since the $z(t)$ chain is irreducible, this extends at once to the case when $z(0)$ is also arbitrary, and hence it follows that either the probability is unity for all initial conditions, or the state is transient, and therefore that the limit obtained from Smith's Theorem is also independent of initial conditions, for states with $c_y > 0$. A little thought will show that this argument is in fact also valid when $c_y = 0$, unless no change is possible in the value of

$W(t)$ (i.e. unless $c_i = 0$ for all j and the increments S are zero with probability one), if we consider how it is possible to reach the state $\{W = 0; z = y\}$, and as the limit is clearly zero when $c_y < 0$, the limit is always independent of initial conditions.

Since this is so, we may consider initial conditions determined by probability distributions, and use any of these which may be convenient to investigate the limits in more detail; let us therefore suppose that $W(0) = 0$, and that $z(0)$ has the equilibrium distribution of the $z(t)$ chain, say the vector (M_i). Then we may apply a continuous-time analogue of the argument given in Loynes (1961a) to prove that with these initial conditions a proper limiting distribution of $W(t)$ exists if and only if either

$$\sum_i c_i M_i > \sum_{i,j} \lambda_i M_i p_{ji} \mu_{ji} \tag{33}$$

or

$$c_i = \mu_{ij} = 0 \text{ for all } i, j, \tag{33a}$$

where μ_{ij} is the mean of the distribution corresponding to β_{ij}; the right side of (33) is infinite if any mean does not exist. The left side of (33) is the average rate at which $W(t)$ decreases, and the right side the average rate at which it increases, or in other words the drifts to and from the boundary $W = 0$. The critical situation in which equality holds in (33) needs a slightly more delicate argument than the others: if a proper limiting distribution exists, then it must be such that no overshoot over the boundary occurs (cf. equation (7) of Loynes (1961a)), and this can clearly only be true if each $c_i \leqq 0$; this and the version of (33) with equality together imply (33a).

If we let t tend to ∞ in (28), then it follows by Lemma 1 that for θ close to the origin on the imaginary axis, $\Phi(\theta, t)$ tends to a limit $\Phi(\theta)$ satisfying

$$A\Phi(\theta) = \theta C^+ F, \tag{34}$$

where F is the limiting value of $F(t)$. The limit $\Phi(\theta)$ is a fortiori also independent of initial conditions.

Now suppose that condition (33) is satisfied, and that the process starts from the particular initial conditions used above, so that the distribution of $W(t)$ tends to a proper limit. Then the Laplace-Stieltjes transform of the distribution of $W(t)$ also tends to a limit, which is continuous on the imaginary axis at the origin, and this limiting transform is of course just the sum of the components of $\Phi(\theta)$. The latter sum can be regarded as a weighted sum of conditional characteristic functions, and by considering the real and imaginary parts separately, it is easy to see that each component of $\Phi(\theta)$ is continuous at the origin. It now follows by a result of

Zygmund (1951) that the limit of $\Phi(\theta, t)$ exists and satisfies (34) for all θ with $R\theta \geqq 0$, that $p[W(t) \leqq x;\ z(t) = y]$ tends to a limit, and that these two limits correspond in the sense of equation (1).

If, on the other hand, condition (33) is not satisfied, it is clear that no proper limit can exist for the joint distribution of $W(t)$ and $z(t)$, since this would imply a proper limit for the distribution of $W(t)$.

We have therefore shown that as t tends to ∞ the joint distribution of $W(t)$ and $z(t)$ tends to a proper limit (which is independent of initial conditions) if and only if (33) is satisfied, and that the transform of this limit satisfies (34). By comparison of (34) and (8) it follows that this limit is an equilibrium distribution, and since any other equilibrium distribution would clearly have itself as a limit, it must in fact be the unique equilibrium distribution.

In order to use (34) to find the limiting distribution we need to know the unknown C^+F. It should usually be possible to do this by using the fact that $\Phi(\theta)$ must be analytic in the right half plane, since it seems likely that Lemma 2 remains valid when $s = 0$, and carrying out an analysis similar to that giving rise to (30); it might be necessary to add the further hypothesis that each β_{ij} be analytic at the origin. If this were not so, and yet $C^+F^*(s)$ could be found, C^+F is clearly then determined by

$$C^+F = \lim_{s \to 0} sC^+F^*(s). \tag{35}$$

From (34) we can deduce a formula which may be useful. We multiply on the left by the vector $t' = (1, 1, \ldots, 1)$ (or in other words add the equations together), divide by θ, and let θ tend to zero along the positive real axis; then the continuity of $\Phi(\theta)$, and the fact that $t'\ (Q - I)/\theta$ tends to $t'B$, where B is the matrix $(p_{ij}\mu_{ij})$ imply that

$$\sum_i c_i M_i - \sum_{j,i} p_{ji}\mu_{ji}\lambda_i M_i = \sum c^+{}_i F_i \tag{36}$$

The connection of the left side of this and (33) is clear. If there is only one $c_i > 0$, the right side of (34) is completely determined by this.

If the system is a queue, with $C = I$, then (36) can be rewritten as

$$Pr\,\{\text{Server is unemployed}\} = 1 - \sum_{j,i} p_{ji}\mu_{ji}\lambda_i M_i. \tag{37}$$

7. Examples

Various queues of some interest can be formulated in this way with a suitable choice of $z(t)$ by using Erlang's device in connection with distributions of E_n type.

https://doi.org/10.1017/S1446788700027488 Published online by Cambridge University Press

If a queue whose service-time is of general independent type is put in series with, first, a queue of the type $E_n/E_m/1$ with a finite waiting-room (a condition imposed to ensure a finite upper bound to the number of customers in the first queue), then the triple of numbers giving (for the first queue) the phase of the input, the phase of the service, and the number of customers present, is a suitable $z(t)$ process for the second queue. It is hardly ever possible to find the transient behaviour of the queue because of the difficulty of finding, as explicit functions of s, the values of θ which make $sI - A$ singular, but it is usually true that the stationary state can be found from (34). This seems to be true even in the simpler situation treated by Beneš, and will be illustrated in somewhat greater detail in the example below. If the restriction to a finite waiting-room is dropped, it becomes necessary to use an infinite chain, and our methods of solution fail. For instance, when the first queue is $M/E_2/1$, then if F on the right side of (34) were known, the complete solution could be readily obtained by a generating function procedure; but there seems no way of finding F.

The example we shall treat may perhaps be described as $E_2 + E_2/M/1$; the service time is independent and negative exponential, and the arrival instants are those events belonging to either of two superimposed renewal processes, both of which have an E_2 life-time distribution. This type of arrival pattern arises naturally when two sources of customers feed into the same queue. The equations generalise immediately to include $E_n + E_m/G/1$.

Let us introduce Erlang's fictitious states into the constituent renewal processes. Then in each process we have two states, 1 and 2, such that the transition rate from one to the other is constant, and equal to λ (say), a customer arriving only on the transition from 2 to 1. If we now form the four compound states (i, j) where i $(=1, 2)$ represents the state of one component process, and j $(= 1, 2)$ that of the other, we may analyse the queue using the process $z(t) = (i(t), j(t))$. We take these states in the order $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 2)$, and for simplicity choose the time scale so that the service-time transform is $1/1 + \theta$.

Then the matrices we need are

$$C^+ = C = I, \tag{38}$$

and

$$A = \begin{bmatrix} \theta - 2\lambda & \lambda/1 + \theta & \lambda/1 + \theta & 0 \\ \lambda & \theta - 2\lambda & 0 & \lambda/1 + \theta \\ \lambda & 0 & \theta - 2\lambda & \lambda/1 + \theta \\ 0 & \lambda & \lambda & \theta - 2\lambda \end{bmatrix}. \tag{39}$$

For the analysis given in section 5, the values of θ making the matrix $sI - A$ singular are needed: the determinant is

$$\det(sI - A) = (\theta - 2\lambda - s)^2\{(1+\theta)(\theta - 2\lambda - s)^2 - 4\lambda^2\}/(1+\theta) \tag{40}$$

This has two coincident zeroes at $\theta = 2\lambda + s$, and two linearly independent left eigenvectors corresponding to them, and according to Lemma 2 two of the three zeroes of the other factor must have positive real parts. However, although this implies that in principle the solution may be obtained by the methods of section 5, the explicit expression for the two roots is too complicated to be of use, even ignoring the difficulty of deciding which is which. For this reason we consider the limiting distribution, given by (34).

The condition (33) for valid limits here reduces to the obvious one

$$\lambda < 1, \tag{41}$$

and we shall assume this to be satisfied. The determinant of A has the following zeroes: $2\lambda, 2\lambda, 0, \frac{1}{2}(4\lambda - 1 + \sqrt{(1+8\lambda)}), \frac{1}{2}(4\lambda - 1 - \sqrt{(1+8\lambda)})$, the last of which is negative and therefore irrelevant.

There are three independent eigenvectors corresponding to the positive roots, and a fourth relation between the components of F is obtained from (36). Carrying out the solution in this way we find for the components of F,

$$\begin{aligned} F_1 &= \{1 - 8\lambda - 8\lambda^2 + (4\lambda+1)\alpha\}/8(2\lambda^2 + 2\lambda + 1) \\ F_2 = F_3 &= \{16\lambda^2 + 11\lambda + 3 - (\lambda+1)(4\lambda+1)\alpha\}/8(2\lambda^2+2\lambda+1) \\ F_4 &= (1+2\lambda)F_1, \end{aligned} \tag{42}$$

where

$$\alpha = \sqrt{(1+8\lambda)}. \tag{43}$$

Finally, for the components of $\Phi(\theta)$, we obtain

$$\begin{aligned} \Phi_1(\theta) &= \{8F_1\theta^2 + \beta\theta + 1 + \alpha - 4\lambda\}/4(1+\theta)(2\theta + \alpha + 1 - 4\lambda) \\ \Phi_2(\theta) = \Phi_3(\theta) &= \{8F_2\theta + 1 + \alpha - 4\lambda\}/4(2\theta + \alpha + 1 - 4\lambda) \\ \Phi_4(\theta) &= \{8F_4\theta + 1 + \alpha - 4\lambda\}/4(2\theta + \alpha + 1 - 4\lambda), \end{aligned} \tag{44}$$

where

$$\beta = 2(1-\lambda)\{1 - 2\lambda + (1+2\lambda)\alpha\}/(1 + 2\lambda + 2\lambda^2). \tag{45}$$

8. Acknowledgements

I am indebted to Mr. A. M. Walker for useful discussions and to the Department of Scientific and Industrial Research for a Research Studentship.

References

Bartlett, M. S. (1956), Stochastic Processes. Cambridge University Press.

Beneš, V. (1957), "On Queues with Poisson Arrivals", Ann. Math. Statist., **28**, 670—677.

Cox, D. R. (1955), "A Use of Complex Probabilities in the Theory of Stochastic Processes", Proc. Camb. Phil. Soc., **51**, 313—319.

Loynes, R. M. (1961a), "The Stability of a Queue with Non-Independent Inter-Arrival and Service Times", Proc. Camb. Phil. Soc. (to be published).

Loynes, R. M. (1961b), "Stationary Waiting-Times for Single-Server Queues", Ann. Math. Statist. (to be published).

McShane, E. J. (1947), Integration. Princeton University Press.

Smith, W. L. (1953). "On the Distribution of Queueing Times", Proc. Camb. Phil. Soc., **49**, 449—461.

Smith, W. L. (1955), "Regenerative Stochastic Processes", Proc. Roy. Soc. A, **232**, 6—31.

Takács, L. (1955), "Investigation of Waiting Time Problems by Reduction to Markov Processes", Acta Math. Acad. Sci. Hungaricae, **6**, 101—129.

Takács, L. (1961), "Transient Behaviour of Single-Server Queueing Processes with Erlang Input", Trans. Am. Math. Soc., **100**, 1—28.

Titchmarsh, E. C. (1939), Theory of Functions. Oxford University Press.

Zygmund, A. (1951), "A Remark on Characteristic Functions", Second Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press.

Statistical Laboratory,
University of Cambridge.