SOME PROBLEMS IN QUEUEING, STORAGE

AND TRAFFIC THEORY

by

G.F. YEO

A thesis submitted to the Australian National University for the degree of Doctor of Philosophy in the Department of Statistics

Canberra

1963

CONTENTS

PREFACE

The work for this thesis was carried out while I was a research student at the Australian National University from January, 1961, to September, 1963. Not all the work done during this period is reported here, as some results were not relevant to the main research described in the thesis; a joint paper with J. Gani (Gani and Yeo (1962)) arising from a virus model is appended in support of the thesis.

Several of the problems described in the thesis have been discussed to some extent by other writers; however, most of the results presented here are new and previous work in the field is acknowledged in the appropriate part of the text. Part or whole of Chapters 1, 2, 3, 4, 6 and 7 have been published or submitted for publication under my own name or jointly with B. Weesakul.

In Chapter 1 the work on the infinite dam is primarily a review, except for some minor new results for the discrete time dam. The new results for the finite dam are part of a paper written jointly with B. Weesakul (Weesakul and Yeo (1963)) which is to appear in the Zeitschrift für Wahrscheinlichkeitstheorie. Chapter 2 is substantially the same as my article (Yeo (1961c)) in the Research Report of the First Summer Research Institute of the Australian Mathematical Society and is entirely new. The major results of Chapter 3 have been published (Yeo (1962)) in the Journal of the Australian Mathematical Society and part of Chapter 4 is to appear shortly (Yeo (1963a)) in the same journal. Since the major part of the work for this last chapter was carried out, Gaver (1962) has published a paper which independently obtains some similar results, and the sections on the preemptive priority repeat different policy were added afterwards for completeness.

The problems of Chapters 5 and 6 arose out of discussions during 1961 with B. Weesakul, and the original model of independent major road traffic in the discrete case and the basic model of one-way major road traffic in the continuous case are due jointly to Weesakul and myself. The remaining extensions and results of both chapters are my own work. Part of Chapter 6 has been submitted to Applied Probability for publication as a joint paper (Yeo and Weesakul (1964)). The results of Chapter 7 are new, and the first two chapter sections are to appear in Biometrika (Yeo (1963b)).

My deepest thanks are due to Dr J. Gani for his understanding supervision throughout my course and for many stimulating discussions and suggestions. I wish to express my appreciation to Professor P.A.P. Moran for his continual encouragement and assistance, particularly during Dr Gani's absence on leave; to Dr B. Weesakul, with whom I now have a substantial and interesting correspondence; to Mr C.C. Heyde for some helpful discussions; to the Australian National University and General Motors Holdens' for their financial assistance; and last but certainly not least to Mrs Beryl Cranston who has made such an excellent job of typing the manuscript.

G.F. Yeo

G.F. Yeo

Department of Statistics,
The Australian National University,
CANBERRA, 1963.

SUMMARY

PART 1

DAMS

The first part of this thesis is concerned with some problems in the theory of dams in both discrete and continuous time.

Chapter 1 contains a brief review of the existing work in the theory of dams and some extensions of this theory. In §2 several discret models are described. For an infinite dam with unit release, the general time-dependent solution is discussed for non-homogeneous inputs; an example of a discrete time queueing model is given, and the known stationary distribution found by taking the limit as time tends to infinity. When the maximum release may be greater than unity a method is given for determining the stationary content distribution. For the finite dam with inputs of a modified geometric type the time-dependent solution is found by determinantal methods. In §3 a review is given of time-dependent results for an infinite dam with possibly non-homogeneous inputs.

In §4 a limiting procedure is described for passing from the discret to the continuous models, so that there is an analogy between the two cases. This enables us to obtain some results for a finite dam in

continuous time with homogeneous Poisson inputs and a negative exponential jump distribution, which cannot be obtained directly. The method has also been applied in §2.5 and §5.5 to some other problems.

<u>Chapter 2</u> describes an infinite dam in discrete time fed by ordered inputs, such that in successive unit time intervals these are independent but their distribution in the time interval $(2t, 2t+1)$ may differ from that in $(2t+1, 2t+2)$. The content is considered separately at odd and even points of time, and the time-dependent distribution of the content at these points is found by the method of generating functions, which may be inverted to give the result in terms of the probabilities of emptiness. These are obtained by combinatorial methods from the probabilities of first emptiness, for which a recurrence relation is given. The result is extended to the case where there may be $K(> 2)$ independent, additive types of input occurring cyclically. A method of solution for the stationary distribution is described, and a numerical example of two geometric input distributions is given. Under the limiting procedure described in §1.4 the continuous analogue is of a form rather different from that of the discrete problem, as the inputs no longer occur cyclically, but independently.

PART 2

SINGLE SERVER QUEUES

In Part 2 two models in the theory of queues with a single server are considered; the first of these has a number of applications later in the thesis. In contrast to Part 1 where we were concerned primarily with time-dependent solutions the stationary properties of the processes now considered in greater detail.

Chapter 3 considers an extension of the M/G/1 queueing system to the case where customers joining an empty queue may have a service time distribution different from those joining a non-empty queue. For a general independent input distribution a necessary and sufficient condition for the existence of a proper stationary distribution is given in §3; subject to a minor additional restriction this is the same as for the G/G/1 system.

A partial solution is obtained in §4 for the time-dependent problem but the probability of emptiness has not been found explicitly. In §5 two alternative methods for obtaining the stationary waiting time distribution are described. The first considers only the points of arrival of customers, which are regeneration points of the process, while the second considers arbitrary points; the results in the two cases are identical. The former is used to extend the result to an Erlangian inter-arrival time distribution. The stationary queue size distribution is found by a similar regeneration point method. In §7 the joint distribution of the length of a busy period and of the number of customers served during the busy period is obtained. A problem of server absenteeism, which may be reduced to a special case of the main problem, is discussed in §8, and the case of bunch arrivals is considered in §10, while in §9 a comparison is made with the results of a similar problem of Finch (1959).

Chapter 4 describes some problems in priority queueing, where interruptions to the service of customers by higher priority customers is allowed. Customers in each class arrive independently in a Poisson process, and independently of other classes of customers, and have a general distribution of service time requirement. Each of the following four priority policies are considered: preemptive priority repeat identical, preemptive priority repeat different, preemptive priority resume and head-of-the-line priority. In §2 and §3 the case of two classes of customers is considered. The distribution of time which customers spend at the counter, and also from the arrival of a non-priority customer at a queue free of other non-priority customers to their reaching the counter for the first time are obtained. These are used to find the stationary waiting time and queue size distributions for each class of customer, and the busy period distributions for the system and for individual classes. The results are extended in §4 to any number $K(\geq 2)$ of classes of customers.

PART 3

ROAD TRAFFIC THEORY

This part is devoted to the application of queueing theory to problems in road traffic theory. Chapters 5 and 6 are concerned with delay at intersections in discrete and continuous time respectively, while delay on a long two-way road is discussed in Chapter 7.

Chapter 5 considers a discrete time model for the delay to vehicles in a one-way minor road at an intersection with a two-way major road, in which traffic has absolute right of way. The minor road vehicles arrive at the intersection with a geometric inter-arrival time distributi while in the major road there is a Markovian relation between successive vehicles. After considering two special cases the service time distribution for minor road vehicles obtained in §3, and used in §4 to find the stationary waiting time and delay distributions. The continuous analogue which results from allowing the units of measurement to tend to zero is considered in §5, and some further problems are discussed in the remaining two sections.

Chapter 6 describes a general continuous time model for the delay to traffic in a one-way minor road at an intersection with a one-way major road. A general description of the major road traffic is given in terms of alternate bunches and gaps, while the minor road traffic is described in a more restricted manner. There are variable gap acceptance times for minor road vehicles waiting to enter the intersection, and several alternate manoeuvres to allow for different conditions are given. In §2 the distribution of the time minor road vehicles spend at the head of the minor road queue is obtained; in §3 this is applied to finding the stationary waiting time, delay, queue size and busy period distributions by using formulae derived in Chapter 3. There is a discussion in §4 of the problem of drivers who accept shorter gaps if they have been waiting a long time. A partial solution is given in §5 to the case of a two-way major road, and some other problems are considered in §6. In §7 the results of some numerical work are given and a comparison made between the delay to minor road vehicles for variable and constant gap acceptance times.

Chapter 7 considers a long, uninterrupted two-way road with one lane in each direction. The description of traffic in each direction in terms of alternate bunches and gaps is of a form similar to that in the previous chapter. However, there is an additional vehicle travelling in one direction at a speed faster than the constant speed of the other vehicles moving in the same direction. Its average speed over a long journey is obtained by considering the distribution of the periods of its restricted and unrestricted travelling. In the final section a flow of fast vehicles is considered; this is a much more difficult problem, and some approximate results are obtained only after making several further simplifying assumptions.

PART 1

DAMS

CHAPTER I

GENERAL PROBLEMS IN THE THEORY OF DAMS

1.1 Introduction

Since the formulation of the theory of dams by Moran (1954) a considerable literature has accumulated on both discrete and continuous dam processes. Descriptive details and bibliographies have been given by Gani (1957), Moran (1959) and Yeo (1961b). In this chapter we shall describe some of the basic models in both discrete and continuous time, extend some of the known results, and discuss some of the analogies between the discrete and continuous time problems. In Chapter 2 we shall consider some problems of a dam with ordered inputs.

We are concerned only with dams, subject to a random input, built to fulfil some deterministic demand of water, for irrigation, say; other uses, such as providing a head of water for power, or storage in case of exceptional rainfall, or as one of a set of dams in series, are not considered.

In the theory of dams some of the important problems are (i) the study of the storage process when it is in statistical equilibrium, (ii) the distribution of the time taken for the dam to dry up for the first time, and (iii) the time-dependent distribution of the dam content. In the next two chapters each of these problems is dealt with, although most emphasis is attached to the last of them. In contrast the problems concerned with stationary properties are dominant in Part 2, where some problems in queueing theory are considered.

Among recent contributions to the theory of dams are works on first emptiness in an infinite dam by Kendall (1957) and Gani (1958), and in a finite dam by Prabhu (1958), Ghosal (1960) and Weesakul (1961a). Time-dependent solutions for the content of an infinite dam have been found by Gani and Prabhu (1959a,b), Yeo (1960, 1961a) and Gani (1962a), while Weesakul (1961b) has found the temporal solution for a finite dam fed by geometric inputs.

1.2 Discrete time models

Let us consider a dam of finite capacity k (a positive integer) in discrete time, with content $Z_t = 0,1,2,\ldots.$ at time $t = 0,1,2,\ldots.$, fed by inputs $X_t = 0,1,2,\ldots.$ in the time interval $(t, t+1)$. If $Z_t + X_t > k$ then an amount $Z_t + X_t - k$ overflows and is lost; there is an integral release of size m $(0 < m < k)$ just before the end of each unit time interval, unless the dam is empty. The content satisfies the relation

$$Z_{t+1} = \min (Z_t + X_t, k) - \min (Z_t + X_t, m). \tag{2.1}$$

We suppose that the input in any unit time interval is independent of that in any other unit time interval, and is a random variable having the probabilities $\Pr\{X_t = i\} = p_i^{t+1}$; Z_t is then a Markov chain. The assumption of the independence of (yearly) inputs seems reasonable in many cases; for example Bhat and Gani (1959) in an unpublished technical report have shown that the serial correlation of annual inputs is not significant for several Australian rivers.

We define $p_i^t(j) = \Pr\left\{\sum_{r=t-j}^{t-1} X_r = i\right\}$ as the probability of having i inputs in the interval (t-j, t), with the probability generating function (p.g.f.) $A^t(s,j) = \sum_{i=0}^{\infty} p_i^t(j)\, s^i$ $(|s| \leq 1)$; we write $q_i^t(j) = \sum_{r=0}^{i} p_r^t(j)$. When the process is homogeneous in time, i.e. $p_i^t(j) \equiv p_i(j)$, then $A^t(s,j) \equiv A(s,j) = \{A(s,1)\}^j = \{A(s)\}^j$, and the $p_i(j)$ are the coefficients of s^i in the expansion of $\{A(s)\}^j$.

Let us first consider the case m = 1. The cumulative probabilities $Q_i(v,t|u,k) \equiv Q_i(v,t) = \Pr\{Z_t \leq i | Z_v = u > 0;\ k\}$ $(i = 0,1,2,\ldots)$ of the content of the dam with capacity k and integral content u at time $v(\geq 0)$ are readily shown to satisfy the difference equations

$$Q_i(v,t+1) = \begin{cases} \sum_{j=0}^{i+1} p_j^{t+1+v} Q_{i+1-j}(v,t) & 0 \leq i \leq k-2 \\ 1 & i \geq k-1. \end{cases}$$

We define $g(v,u,t|k) = \Pr\{Z_t = 0 | Z_v = u > 0;\ Z_r > 0,\ v < r < t; k\}$ as the probability of first emptiness at time t $(\geq v)$ of the dam with content u at time v. For an infinite dam the probabilities $g(v,u,t)$ of first emptiness satisfy the relations

$$g(v,u,t+v) = \begin{cases} p_0^{u+v}(u) & t = u \\ \sum_{j=1}^{t-u} p_j^{u+v}(u)\, g(u+v,j,t+v) & t > u. \end{cases} \tag{2.3}$$

This has not been solved explicitly in the general case, although a solution may be obtained iteratively for any given values of u, v and t from the recurrence relations (2.3). Defining the transform $G(v,u,s)$ $= \sum_{t=u+v}^{\infty} g(v,u,t)\ s^t$ $(|s| \leqq 1)$ we obtain from (2.3) that

$$G(v,u,s) = \sum_{j=o}^{\infty} p_j^{u+v}(u) \quad G(u+v,j,s).$$

In the time-homogeneous case Gani (1958) has shown that

$$g(0,u,t) = u\ t^{-1}\ p_{t-u}(t) \qquad\qquad t \geqq u. \tag{2.4}$$

The probability of emptiness, not necessarily for the first time, is [Gani (1962a)]

$$Q_o(v,t+v) = \left\{p^{t+1+v}\right\}^{-1} \sum_{j=u+1}^{t+1} g(v,j,t+1+v). \tag{2.5}$$

We define $P(v,s,t) = Q_o(v,t) + \sum_{i=1}^{\infty} \left\{Q_i(v,t) - Q_{i-1}(v,t)\right\} s^i$ $(|s| \leqq 1)$ as the p.g.f. of the dam content distribution. From (2.2) we obtain for $0 < s \leqq 1$

$$P(0,s,t) = s^{u-t}\ A^t(s,t) - (1-s^{-1}) \sum_{j=1}^{t-u} Q_o(0,t-j)\ p_o^{t+1-j}\ A^t(s,j-1), \tag{2.6}$$

where $Q_o(0,t)$ is determined from (2.5). This may be inverted as for time-homogeneous inputs to give

$$Q_i(0,t) = q_{i+t-u}^t(t) - \sum_{j=1}^{t-u} p_o^{t+1-j}\ p_{i+j}^t(j-1)\ Q_o(0,t-j) \qquad i = 1,2,\ldots \tag{2.7}$$

If $\lim_{t\to\infty} p_i^t = p_i$ exists, where $\sum_{i=o}^{\infty} p_i = 1$, and if $\rho = \sum_{i=1}^{\infty} ip_i < 1$

then the content distribution tends with time to a proper stationary distribution for which the stationary p.g.f. $P(s)$ of the content has been found by Moran (1956) as

$$P(s) = \frac{(1-\rho)\ (1-s)}{A(s) - s}. \tag{2.8}$$

As an example of the above theory we consider a discrete time queueing model which is a generalisation of that due to Meisling (1958). Customers are assumed to arrive independently at a counter with a single server, who serves customers in the order of their arrival. Customers arrive just before the time points $t = 0,1,2,\ldots$, i.e. at $t-0$, such that there is a probability $a(t)$ $(0 \leqq a(t) \leqq 1)$ of one customer arriving and a probability $b(t) = 1 - a(t)$ of no customers arriving. The service times of customers are independently and identically distributed random variables with probabilities c_i $(i = 1,2,\ldots)$ and p.g.f. $\gamma(s) = \sum_{i=1}^{\infty} c_i\, s^i$ $(|s| \leqq 1)$. This is a discrete analogue of the M/G/1 queueing process; the waiting time is equivalent to the content of a dam, and the time-dependent solution is obtained from (2.3), (2.5), (2.6) and (2.7) by substituting $b(t) + a(t)\ \gamma(s)$ for $A^{t-1}(s)$.

We now turn to the dam with finite capacity k and with $m = 1$. When there are independent geometric inputs Weesakul (1961a,b) has obtained the probability of first emptiness, with and without overflow being allowed, and hence the time-dependent solution. This has been extended by Weesakul and Yeo (1963) to the case where the input distribution has the p.g.f. $A(s) = b + a\ \alpha(s)$, where $0 < a \leqq 1$,

$b = 1 - a$ and where $\alpha(s) = p(1 - qs)^{-1}$ $(0 < p < 1,\ q = 1 - p)$ is the p.g.f. of a geometric distribution. This has led to results for an analogous continuous time process, which we describe in §1.4.

For this model it is known that the content Z_t forms a Markov chain with k states $0,1,\ldots, k-1$, whose transition matrix is

$$\underline{S} = \begin{bmatrix} b+ap & apq & apq^2 & . & . & . & apq^{k-2} & aq^{k-1} \\ b & ap & apq & . & . & . & apq^{k-3} & aq^{k-2} \\ o & b & ap & . & . & . & apq^{k-4} & aq^{k-3} \\ . & . & . & . & . & . & . & . \\ & & & & & & & \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & b & ap & aq \\ . & . & . & . & . & o & b & a \end{bmatrix} . \tag{2.9}$$

Using determinantal methods it may be shown that the generating function $\psi(i,s|u,k) = \sum_{t=u}^{\infty} Q_i(0,t|u,k)\, s^t$ $(0 \leq s < 1)$ with respect to time of the cumulative content probabilities at time t given a dam capacity k and initial content u is given by

$$\psi(i,s|u,k) = \frac{as^2 \sum_{j=1}^{i+1} C_j(1-q^{i+1-j}) + bs^2 \sum_{j=1}^{i+2} C_j}{|\underline{I} - s\underline{S}|} , \tag{2.10}$$

where $\underline{I}$ is the identity matrix, C_j is a determinant similar to $|\underline{I} - s\underline{S}|$ except that the jth row is replaced by $(0,\ldots,0,b,ap,\ldots,apq^{k-u-2}, aq^{k-u-1})$ with b in the uth position. The determinants on the right hand side of (2.10) have the values

$$|\underline{I} - s\underline{S}| = (1-s)(\lambda_1-\lambda_2)^{-1}\{(1-\lambda_2)\lambda_1^{\,k} - (1-\lambda_1)\lambda_2^{\,k}\}$$

$$C_n = \begin{cases} s^{u-n}\, b^{u-n+1}\, E_{k-u}\, G_{n-1} & n = 1,2,\ldots,u \\ q^{n-u-1} F_u\{(ap-bq)E_{k-n+1} + bqE_{k-n}\} & n = u+1,\ldots,k \end{cases}$$

$$E_n = (\lambda_1-\lambda_2)^{-1}\{(1-sa-\lambda_2)\lambda_1^{\,n} - (1-sa-\lambda_1)\lambda_2^{\,n}\} \qquad n \geqq 1$$

$$F_n = (\lambda_1-\lambda_2)^{-1}\{(\lambda_1-sb)\lambda_1^{\,n} - (\lambda_2-sb)\lambda_2^{\,n}\} \qquad n \geqq 1$$

$$G_n = (\lambda_1-\lambda_2)^{-1}\{(1-\lambda_2)(\lambda_1-sb)\lambda_1^{\,n-1} - (1-\lambda_1)(\lambda_2-sb)\lambda_2^{\,n-1}\} \qquad n \geqq 1$$

$$2\lambda_{1,2} = \gamma \pm \{\gamma^2 - 4sbq\}^{\frac{1}{2}}$$

$$\gamma = 1 - s(ap-bq),$$

where $E_0 = E_{-1} = F_0 = G_0 = 1$. Similar expressions may be obtained for the probabilities of first emptiness and emptiness, with and without emptiness being allowed. The stationary distribution may be obtained by the method of Prabhu (1958), or as the limit of the time-dependent solution.

All the results can readily be extended to the case in which units of time and content are Δ (> 0), where Δ is not necessarily unity. We use this in §1.4 and §2.5, where we then let $\Delta \to 0$ in a suitable manner in order to obtain results for an analogous continuous time process.

We now briefly consider an infinite dam in discrete time with time-homogeneous inputs where the maximum release m may be greater than unity. A stationary content distribution exists independently of the initial distribution if and only if $\rho < m$, where $\rho = p'(1)$ is the mean of the input distribution. The stationary probabilities $P_i = \lim_{t\to\infty} \Pr\{Z_t = i | Z_o = u\}$ satisfy the difference equations

$$P_o = \sum_{j=o}^{m} P_j (p_o + p_1 + \ldots + p_{m-j})$$

$$P_n = \sum_{j=o}^{m+n} P_j \, p_{m+n-j} \qquad n = 1, 2, \ldots \qquad (2.11)$$

From (2.11) we find that the p.g.f. $P(z) = \sum_{i=o}^{\infty} P_i \, z^i \; (|z| \leq 1)$ is given by

$$P(z) = \frac{\sum_{j=o}^{m-1} P_j \sum_{i=o}^{m-1-j} (z^m - z^{i+j}) \, p_i}{z^m - p(z)} \qquad (2.12)$$

where $P_o, P_1, \ldots, P_{m-1}$ have still to be determined [c.f. Bailey (1954)]. If $p(e^{i\theta})$ is analytic for $\mathcal{J}(\theta) > -\delta$ for some real $\delta > 0$ and if $\rho < m$ then by an argument similar to that for Lemma 1 of Ewens and Finch (1962) [see §3.5] it may be shown that $z^m - p(z) = 0$ has m-1 roots in $|z| < 1$, and one root at $z = 1$. Denote these roots by $z_1, z_2, \ldots, z_{m-1}$, $z_m = 1$, and suppose they are distinct; if there are multiple roots further differentiation enables us to obtain a total of m equations in (2.13). At the zeros in $|z| \leq 1$ of the denominator of (2.12) there must be zeros of the numerator so that we have

$$\sum_{j=o}^{m-1} P_j \sum_{i=o}^{m-1-j} (m-i-j) \, p_i = m - \rho$$

$$\sum_{j=o}^{m-1} P_j \sum_{i=o}^{m-1-j} (z_n^{\,m} - z_n^{\,i+j})\; p_i = 0 \qquad n = 1,2,\ldots,m-1. \quad (2.13)$$

We thus have m equations in m unknowns so that we can find $P_o, P_1, \ldots, P_{m-1}$, although this is very tedious if m is at all large, as in the applications to traffic delays where m may be 20 or more.

The mean content $P'(1)$ is given by

$$P'(1) = \frac{p''(1) - m(m-1) + \sum_{j=o}^{m-1} P_j \sum_{i=o}^{m-1-j} \{m(m-1) - (i+j)(i+j-1)\} p_i}{2(m - p'(1))} \qquad (2.14)$$

1.3 <u>Continuous time models</u>

In a dam of finite capacity $K(> 0)$ in continuous time, we suppose that there is an input $X(T,T+\tau)$ in the time interval $(T,T+\tau)$ whose distribution is independent in non-overlapping time intervals. There is a continuous release at unit rate per unit time unless the dam is empty, and any input which would increase the content beyond K overflows and is lost. The content $Z(\tau)$ $(\geqq 0)$ of the dam at time $\tau(\geqq 0)$ satisfies the Markovian relation

$$Z(\tau+\delta\tau) = \min(Z(\tau)+\delta X(0,\tau),K) - \min\,(\delta\tau,(1-\eta)\delta\tau), \qquad (3.1)$$

where $\eta\delta\tau$ is the time during the small interval $(\tau,\tau+\delta\tau)$ in which the dam is empty.

We suppose that inputs occur in a Poisson process with parameter $\lambda(\tau)$ at time τ such that the size of each input is a random variable with d.f. $F(x)$ $(0 \leqq x < \infty)$ and Laplace-Stieltjes transform (LST) $\alpha(\theta) = \int_{0-}^{\infty} e^{-\theta x} dF(x)$ $(\theta \geqq 0)$. The content process satisfies the forward Kolmogorov equation [Gani and Prabhu (1959a)]

$$\frac{\partial W(z,\tau)}{\partial \tau} - \frac{\partial W(z,\tau)}{\partial z} = -\lambda(\tau)W(z,\tau) + \lambda(\tau)\int_{v=0}^{z} W(z-v,\tau)dF(v) \qquad 0 \leqq z \leqq K$$

$$W(z,\tau) = 1 \qquad z \geqq K, \tag{3.2}$$

where $W(z,\tau) \equiv W(z,\tau|U,K) = \Pr\{Z(\tau) \leqq z|Z(o) = U;K\}$ is the d.f. of the content $Z(\tau)$ at time τ given an initial content U and dam capacity K; $W(z,\tau) = 0$ for all $z < 0$. $W(z,\tau)$ is continuous for all z in $\tau \geqq U$, $0 < z \leqq K$, but has a discontinuity at $z = 0$, the concentration $W(0,\tau)$ being the probability that the dam is empty at time τ.

When $K = \infty$ this process is equivalent to the M/G/1 single server queueing system with a negative exponential inter-arrival time distribution and a general service time distribution. The integro-differential equation (3.2) when $K = \infty$ has been obtained by Takács (1955), who has shown that the LST $\mathcal{H}(\theta,\tau) = \int_{0-}^{\infty} e^{-\theta z} dW(z,\tau)$ $(\theta \geqq 0)$ is given by

$$\mathcal{H}(\theta,\tau) = e^{(\tau-U)\theta-\lambda\tau(1-\alpha(\theta))} - \theta\int_{v=0}^{\tau} W(0,\tau-v)e^{v(\theta-\lambda(1-\alpha(\theta)))}dv, \tag{3.3}$$

where $W(0,\tau)$ requires to be determined. This may be inverted [c.f. Gani (1962a)]; when $F(x)$ is an absolutely continuous function

$$W(z,t) = K(z+t-U,0,\tau) - \int_{v=0}^{\tau} W(0,\tau-v)k(z+v,\tau-v,\tau)d\tau, \tag{3.4}$$

where

$$K(x,v,\tau) = \sum_{n=0}^{\infty} e^{-(\rho(\tau)-\rho(v))}(\rho(\tau)-\rho(v))^n F^{n*}(x)(n!)^{-1},$$

$F^{n*}(x)$ being the n-fold iterated convolution of $F(x)$, $F^{0*}(x) = 1$ for $x \geqq 0$, $k(x,\tau) = \frac{d}{dx} K(x,\tau)$ and $\rho(\tau) = \int_{v=0}^{\tau} \lambda(v)dv$. Gani (1962a) has also shown that

$$W(0,\tau) = e^{-\rho(\tau)} + \int_{v=U+}^{\tau} g(0,v,\tau)dv, \tag{3.5}$$

where the probability is $dG(V,U,\tau)$ that first emptiness occurs in the interval $(\tau,\tau+d\tau)$ $(\tau \geqq V)$ given content U at time V;

$$dG(V,U,\tau+V) = \begin{cases} e^{-(\rho(V+U) - \rho(V))} & \tau = U \\ g(V,U,\tau+V)d\tau & \tau > U, \end{cases}$$

with a continuous probability for $\tau > U$, but a discrete concentration at $\tau = U$. We have the relation

$$dG(V,U,\tau+V) = \int_{v=0-}^{\tau-U} dG(V+U,v,\tau+V)dK(v,V,V+U). \tag{3.6}$$

This has not been solved explicitly in the general case; however, for time-homogeneous inputs Kendall (1957) has found that

$$dG(0,U,\tau) = \begin{cases} 0 & \tau < U \\ u\tau^{-1}dK(\tau-U,\tau) & \tau \geqq U \end{cases} \tag{3.7}$$

where $dK(0,\tau) = e^{-\lambda\tau}$ and

$$K(x,0,\tau) \equiv K(x,\tau) = \sum_{n=0}^{\infty} e^{-\lambda\tau}(\lambda\tau)^n F^{n*}(x)(n!)^{-1}. \tag{3.8}$$

The stationary d.f. $W(z)$ $(0 \leqq z < \infty)$ of the dam content has been found as the limit of the time-dependent solution $W(z,\tau)$ as $\tau \to \infty$ by Takács (1955) and directly by Lindley (1952) and others. If $\lim_{\tau\to\infty} \lambda(\tau) = \lambda$ exists, then the stationary distribution exists as a proper distribution if and only if $\lambda\rho < 1$, where $\rho = -\alpha'(0)$; the LST $\mathcal{H}(\theta) = \int_{0-}^{\infty} e^{-\theta z}\, dW(z)$ is

$$\mathcal{H}(\theta) = \frac{(1 - \lambda\rho)}{1 - \lambda\{1 - \alpha(\theta)\}/\theta} . \tag{3.9}$$

1.4 <u>Limiting processes</u>

In §1.2 we considered dams discrete in both time and content, and in §1.3 dams continuous in time and content. We can pass from the former to the latter by letting the units of measurement for time and content tend to zero in an appropriate manner. For the continuous time model we consider only inputs that occur in a homogeneous Poisson process; other types of input may be considered, but are not discussed here as the only new solutions we obtain are for the special case of homogeneous Poisson inputs.

The discrete time process may be defined in units of size Δ (> 0) instead of unity, so that time, content, input and release are all measured in multiples of Δ. We shall use the substitution

$$\text{(a)}\quad p_{i\Delta}(\Delta) = \begin{cases} 1 - \lambda\Delta & i = 0 \\ \lambda\Delta\{A(i\Delta) - A((i-1)\Delta)\} & i = 1,2,\ldots \end{cases}$$

$$\text{(b)}\quad i = x\Delta^{-1},\ t = \tau\Delta^{-1},\ k = K\Delta^{-1},\ u = U\Delta^{-1}, \tag{4.1}$$

where $A(i\Delta)$ $(i = 1,2,...)$ are cumulative probabilities with $\lim_{i\to\infty} A(i\Delta) = 1$, and such that in the limit as $i \to \infty$, $\Delta \to 0$ and $i\Delta \to x$ we have

$$\lim_{\Delta\to 0} A(i\Delta) = F(x) \qquad\qquad 0 \leqq x < \infty,$$

where $F(x)$ is a d.f. We suppose that x,τ,U and K are real and non-negative. When $k = \infty$ the capacity of K of the dam in continuous time is also infinite. As $\Delta \to 0$ the release becomes continuous at unit rate per unit time unless the dam is empty, and inputs, whose d.f. is $F(x)$, occur in a Poisson process with parameter λ.

Under the substitution (4.1) it is readily shown that equation (2. 2) tends in the limit as $\Delta \to 0$ to (3.2). Reich (1958,1959) has shown for $K = \infty$ that (3.2) has a unique d.f. solution, so that if we obtain a solution of (3.2) with $K = \infty$ by a limiting procedure from the analogous discrete process, then we shall have a unique solution. When $K < \infty$ it does not appear to be known if there is a unique d.f. which is the solution of (3.2), although it seems reasonable to expect that this is so; in any case the result obtained from the analogous discrete process is a solution of (3.2).

Writing the n-fold convolution of $A(i\Delta)$ as $A^{n*}(i\Delta)$ we have under the substitution (4.1) that

$$\lim_{\Delta\to 0} q_{i\Delta}(t\Delta) = \lim_{\Delta\to 0} \sum_{j=o}^{i} p_{j\Delta}(t\Delta)$$

$$= \lim_{\Delta \to 0} \sum_{n=0}^{t} \binom{t}{n} (1-\lambda\Delta)^{t-n} (\lambda\Delta)^n A^{n*}(i\Delta)$$

$$= \sum_{n=0}^{\infty} e^{-\lambda\tau} (\lambda\tau)^n F^{n*}(x) (n!)^{-1}$$

$$= K(x,\tau). \tag{4.2}$$

For an infinite dam the results (2.3), (2.5) and (2.7) tend under the substitution (4.1) and the limit $\Delta \to 0$ to (3.6), (3.5) and (3.4) respectively; with $K = \infty$ this gives the unique d.f. solution of (3.2), as has been found directly.

Let us now consider the finite dam, which has been discussed by Weesakul and Yeo (1963). We suppose that

$$A(i\Delta) = 1 - (1 - \mu\Delta)^i \qquad i = 1,2,\ldots \tag{4.3}$$

is a cumulative geometric distribution, so that $F(x) = 1-e^{-\mu x} (0 \leqq x < \infty)$ is the d.f. of a negative exponential distribution. We wish to find the time transform

$$\varphi(z,\theta|U,K) = \int_{\tau=0}^{\infty} e^{-\theta\tau} W(z,\tau|U,K)\,d\tau$$

$$= \lim_{\Delta \to 0} \Delta \;\psi(z\Delta^{-1}, e^{-\theta\Delta}|U\Delta^{-1}, K\Delta^{-1})$$

of the d.f. $W(z,\tau|U,K)$ of the content distribution for a dam with capacity K and initial content U by using the substitutions (4.1) and (4.3). In (2.10) the terms $as^2 C_j(1-q^{i+1-j})(|\underline{I}-s\underline{S}|)^{-1}$ $(1 \leqq j \leqq i + 1)$ all tend to zero as $\Delta \to 0$, while the remaining terms yield

$$\varphi(z,\theta|U,K) = \begin{cases} e^{-U(\theta+\lambda)}\{(\eta_1-\lambda)e^{-(K-U)\eta_2}-(\eta_2-\lambda)e^{-(K-U)\eta_1}\} \times \\ \left\{ \dfrac{\eta_1 e^{-z(\eta_2-\lambda-\theta)}-\eta_2 e^{-z(\eta_1-\lambda-\theta)}}{\nu\theta(\eta_1 e^{-K\eta_2}-\eta_2 e^{-K\eta_1})} \right\} & 0 \leqq z \leqq U \\ \varphi(U,\theta|U,K) + \{(\theta+\lambda-\eta_2)e^{-U\eta_2}-(\theta+\lambda-\eta_1)e^{-U\eta_1}\} \times \\ \left\{ \dfrac{\eta_2(\eta_1-\lambda)e^{-K\eta_2}}{\theta+\lambda-\eta_1}(e^{z(\eta_2-\mu)+\mu U}-e^{U\eta_2}) - \dfrac{\eta_1(\eta_2-\lambda)e^{-K\eta_1}}{\theta+\lambda-\eta_2} \times \right. \\ \left. (e^{z(\eta_1-\mu)+\mu U}-e^{U\eta_1}) \right\} \left\{ \nu\theta(\eta_1 e^{-K\eta_2}-\eta_2 e^{-K\eta_1}) \right\}^{-1} & U \leqq z \leqq K, \end{cases} \quad (4.5)$$

where $\nu = \{(\lambda+\mu+\theta)^2 - \lambda\mu\}^{\frac{1}{2}}$ and $2\eta_{1,2} = \theta+\lambda+\mu\pm\nu$. This may be inverted to give the time-dependent content distribution. The stationary d.f. $W(z|K)$ may be obtained by using an extension of Abel's theorem [Widder (1946), Chapter 5] as $F(z|K) = \lim_{\theta\to 0}\theta\varphi(z,\theta|U,K)$. The result found in this way agrees with that determined directly, and is

$$F(z|K) = \frac{\lambda e^{z(\lambda-\mu)}-\mu}{\lambda e^{K(\lambda-\mu)}-\mu}\,. \quad (4.6)$$

Probabilities of first emptiness and emptiness, with and without overflow being allowed, have been obtained in a similar manner by Weesakul and Yeo (1963). The probability of first overflow before emptiness has been found, and this has been applied to a problem in insurance risk [Cramér (1955), Bartlett (1955)].

CHAPTER 2

A DAM WITH ORDERED INPUTS

2.1 Introduction

In the discrete time models discussed in Chapter I the inputs were of the same form for every unit time interval. In this chapter we are concerned with a dam of infinite capacity in which inputs may vary seasonally so that their distribution in the wet season, say, varies from that in the dry season.

A continuous time model in which two types of ordered input of fixed (different) size occur alternately in a Poisson process has been considered by Gani (1960); this model is rather more restricted than the one we wish to consider, and needs to be solved step by step from recurrence relations, while we can employ generating functions to obtain the time-dependent solution.

We have that equation (1.2.1) still holds (with $k=\infty$). We let the inputs in successive time intervals be independent, but such that the distribution of the input X_{2t} in $(2t,2t+1)$ differs from the distribution of the input X_{2t+1} in $(2t+1, 2t+2)$. However, both types of input are additive and their distribution in unit time intervals is denoted, respectively, by

$$\begin{aligned} p_i(1) &= p_i = \Pr\{X_{2t} = i\} \\ q_i(1) &= q_i = \Pr\{X_{2t+1} = i\} \end{aligned} \qquad i,t = 0,1,2,\ldots. \tag{1.1}$$

We shall consider only inputs homogeneous in time, although the results may be extended to the more general case as in Chapter I.

The process Z_t is no longer Markovian; however, $\{Z_t,t\}$ jointly define a time-homogeneous Markov chain.

We let $P_i(t) = \Pr\{Z_t = i|Z_o = u\}$ $(i = 0,1,2,\ldots)$ be the probability distribution of the dam content at time t, given that there is an integral-valued initial dam content $u > 0$. The form of the content distribution is different at odd and at even times; we can construct the following set of difference equations for these distributions:

$$\begin{aligned} P_o(2t+1) &= (p_o+p_1)P_o(2t)+p_oP_1(2t) \\ P_i(2t+1) &= \sum_{j=o}^{i+1} p_jP_{i+1-j}(2t) \qquad i = 1,2,\ldots, \end{aligned} \tag{1.2}$$

$$\begin{aligned} P_o(2t+2) &= (q_o+q_1)P_o(2t+1)+q_oP_1(2t+1) \\ P_i(2t+2) &= \sum_{j=o}^{i+1} q_jP_{i+1-j}(2t+1) \qquad i = 1,2,\ldots, \end{aligned} \tag{1.3}$$

subject to $\sum_{i=o}^{\infty} P_i(2t+1) = \sum_{i=o}^{\infty} P_i(2t) = 1$.

We define the p.g.f.'s

$$Q(s,t) = \sum_{i=o}^{\infty} P_i(t)s^i,$$

$$A(s) = \sum_{i=o}^{\infty} p_is^i, \quad B(s) = \sum_{i=o}^{\infty} q_is^i. \qquad |s| \leq 1$$

If we write $A(s,t) = \sum_{i=0}^{\infty} p_i(t)s^i$ ($|s| \leqq 1$) as the p.g.f. of the distribution $p_i(t)$ of inputs in any t unit time intervals of the form (2r, 2r+1), where r is an integer, we have (as in Chapter I) that $A(s,t) = \{A(s)\}^t$, so that $p_i(t)$ ($i = 0,1,2,\ldots$) is the coefficient of s^i in the expansion of $\{A(s)\}^t$. Similarly $B(s,t) = \sum_{i=0}^{\infty} q_i(t)s^i = \{B(s)\}^t$.

By multiplying the ith relation in (1.2), and (1.3), by s^{i+1} and summing we obtain the difference equations

$$sQ(s,2t+1) = A(s)Q(s,2t) - (1-s)p_o\ P_o(2t)$$

$$sQ(s,2t+2) = B(s)Q(s,2t+1) - (1-s)\ q_o\ P_o(2t+1).$$

By a simple combination of equations of this type we obtain

$$s^2Q(s,2t+2)-A(s)B(s)Q(s,2t) = -p_o(1-s)B(s)P_o(2t)-q_o(1-s)sP_o(2t+1) \quad (1.4)$$

$$s^2Q(s,2t+3)-A(s)B(s)Q(s,2t+1) = -q_o(1-s)A(s)P_o(2t+1)-p_o(1-s)sP_o(2t+2). \quad (1.5)$$

It can be seen that we must evaluate the probabilities $P_o(2t)$ and $P_o(2t+1)$ of emptiness before we can find explicit solutions to (1.4) and (1.5). To do this we shall first discuss the probabilities of first emptiness, and shall proceed from these to the probabilities of emptiness.

2.2 Probabilities of emptiness

We define the probabilities

$$g(v,u,t) = \Pr\{Z_t = 0 | Z_v = u > 0;\ Z_n > 0, v < u < t\}$$

that the dam with content u at time v is empty for the first time at t (as in §1.2). Clearly emptiness cannot occur prior to the time $u + v$, but it may happen at any instant $t = u+v+r$ $(r = 0,1,2,\ldots)$; for this there must be exactly r inputs in such a way that there is no emptiness before $t = u+v+r$. Until the dam becomes empty for the first time $Z_t = Z_v+S_t-S_v-t+v$, where $S_t = X_o+X_1+\ldots+X_{t-1}$ $(t \geq 1)$ is the total input up to time t; hence for first emptiness beyond $u+v$ we have $Z_{u+v} = S_{u+v}-S_v > 0$.

We write $r_i(t,T) = \sum_{j=o}^{i} p_j(t)q_{i-j}(T)$ as the probability of having i inputs distributed over t unit time intervals of the type $(2r,2r+1)$ and T of the type $(2r+1,2r+2)$, for integral r.

We can now write out the following set of difference equations for the probabilities of first emptiness:

$$g(0,2u,2t+1) = \sum_{j=1}^{t-u} r_{2j}(u,u)g(0,2j,2t+1-2u) + \sum_{j=o}^{t-u} r_{2j+1}(u,u)g(0,2j+1,2t+1-2u)$$

$$g(0,2u+1,2t+1) = \sum_{j=1}^{t-u} r_{2j}(u+1,u)g(1,2j,2t+1-2u) + \sum_{j=o}^{t-u-1} r_{2j+1}(u+1,u)g(1,2j+1,2t+1-2u)$$

$$g(0,2u,2t) = \sum_{j=1}^{t-u} r_{2j}(u,u)g(0,2j,2t-2u) + \sum_{j=o}^{t-u-1} r_{2j+1}(u,u)g(0,2j+1,2t-2u)$$

$$\begin{aligned}
g(0,2u+1,2t+2) &= \sum_{j=1}^{t-u} r_{2j}(u+1,u)g(1,2j,2t+2-2u) \\
&\quad + \sum_{j=o}^{t-u} r_{2j+1}(u+1,u)g(1,2j+1,2t+2-2u) \\
g(1,2u,2t+1) &= \sum_{j=1}^{t-u} r_{2j}(u,u)g(1,2j,2t+1-2u) \\
&\quad + \sum_{j=o}^{t-u-1} r_{2j+1}(u,u)g(1,2j+1,2t+1-2u) \\
g(1,2u\ 1,2t+1) &= \sum_{j=1}^{t-u} r_{2j}(u,u+1)g(0,2j,2t-1-2u) \\
&\quad + \sum_{j=o}^{t-u-1} r_{2j+1}(u,u+1)g(0,2j+1,2t-1-2u) \\
g(1,2u,2t) &= \sum_{j=1}^{t-u-1} r_{2j}(u,u)g(1,2j,2t-2u) \\
&\quad + \sum_{j=o}^{t-u-1} r_{2j+1}(u,u)g(1,2j+1,2t-2u) \\
g(1,2u\ 1,2t+2) &= \sum_{j=1}^{t-u} r_{2j}(u,u+1)g(0,2j,2t-2u) \\
&\quad + \sum_{j=o}^{t-u-1} r_{2j+1}(u,u+1)g(0,2j+1,2t-2u) \qquad (2.1)
\end{aligned}$$

with

$$g(0,2u,2u) = r_o(u,u) \qquad g(1,2u,2u+1) = r_o(u,u)$$
$$g(0,2u+1,2u+1) = r_o(u+1,u) \qquad g(1,2u+1,2u+2) = r_o(u,u+1). \tag{2.2}$$

As we have homogeneous inputs we can write, as we have done in equation (2.1), that

$$g(2v,u,t) = g(0,u,t-2v) \qquad\qquad t \geqq 2v + u.$$
$$g(2v+1,u,t) = g(1,u,t-2v).$$

We are actually interested only in the probabilities of emptiness of the form $g\ (0,u,t)$, but we need those of the form $g(1,u,t)$ in order to solve the equations in (2.1). The probabilities in (2.1) do not have a readily obtainable explicit solution, but they may be computed iteratively for any particular values of u and t. Rather than obtaining them in this way, it may be easier to formulate an occupancy problem of which the probabilities are the solution, and use the method of truncated polynomials as developed by Gani (1958,1961).

We can formulate the problem of first emptiness as an occupancy problem in the following way. Let us consider the probability $g(0,u,2t)$; there must be exactly $r = 2t-u$ inputs such that there are

$$\begin{array}{l}
x_o \text{ inputs in } (0,2u), \text{ where } x_o \geqq 1 \\
x_1 \text{ inputs in } (2u,2u+1), \text{ where } x_o+x_1 \geqq 2 \\
\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
x_{r-1} \text{ inputs in } (u+r-2,u+r-1), \text{ where } x_o+..+x_{r-1} = r \\
x_r \text{ inputs in } (u+r-1,u+r), \text{ where } x_r = 0.
\end{array} \tag{2.3}$$

We can obtain similar relations when t is odd; we shall use these combinatorial relations to obtain the probabilities of emptiness, not necessarily for the first time.

Let us now define the p.g.f.'s of the probabilities of first emptiness as

$$\varphi(v,u,\theta) = \sum_{t=u+v}^{\infty} g(v,u,t)\theta^{t} \qquad |\theta| \leq 1.$$

We can find from (2.1) and (2.2) a set of difference equations for the φ's as

$$\varphi(0,2u,\theta) = \theta^{2u}\sum_{j=o}^{\infty}\{\varphi(0,2j,\theta)r_{2j}(u,u)+\varphi(0,2j+1,\theta)r_{2j+1}(u,u)\}$$

$$\varphi(0,2u+1,\theta) = \theta^{2u+1}\sum_{j=o}^{\infty}\{\varphi(1,2j,\theta)r_{2j}(u+1,u)+\varphi(1,2j+1,\theta)r_{2j+1}(u+1,u)\}$$

$$\varphi(1,2u,\theta) = \theta^{2u+1}\sum_{j=o}^{\infty}\{\varphi(1,2j,\theta)r_{2j}(u,u)+\varphi(1,2j+1,\theta)r_{2j+1}(u,u)\}$$

$$\varphi(1,2u+1,\theta) = \theta^{2u+2}\sum_{j=o}^{\infty}\{\varphi(0,2j,\theta)r_{2j}(u,u+1)+\varphi(0,2j+1,\theta)r_{2j+1}(u,u+1)\} \qquad (2.4)$$

When the two input distributions are identical the problem reduces to that considered in Chapter I and has the solution given by (1.2.4).

We now consider the probabilities of emptiness, not necessarily for the first time, of the dam. Let us first consider emptiness at time $2t = u+r$ given content u at time zero. Since the dam may dry up before time

$u+r$, emptiness at $u+r$ may result with less than r inputs in $(0,u+r)$, but there cannot be more than r inputs in this interval.

As has been done by Yeo (1961a) for simple additive inputs we include the time interval $(u+r,u+r+1)$ as an artificial measure in order to formulate the problem in the same manner as that used for first emptiness. In this interval we specify zero input; this has probability p_o, as it is in the time $(2t,2t+1)$, and since what happens beyond $u+r$ is independent of $P_o(u+r)$ we must divide our final result by p_o.

We consider now the case where there must be exactly $j(0 \leqq j \leqq r)$ inputs in $(0,u+r)$; then for $Z_{u+r} = 0$, and subject to zero input in $(u+r,u+r+1)$, there must be

$$\begin{aligned}
&x_o \text{ inputs in } (0,u+r-j+1), \text{ where } x_o \geqq 1\\
&x_1 \text{ inputs in } (u+r-j+1,u+r-j+2), \text{ where } x_o+x_1 \geqq 2\\
&\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\
&x_{j-1} \text{ inputs in } (u+r-1,u+r), \text{ where } x_o+x_1+..+x_{j-1} = j\\
&x_j \text{ inputs in } (u+r,u+r+1), \text{ where } x_j = 0.
\end{aligned} \tag{2.5}$$

This is precisely the same formulation as for the first emptiness problem (2.3), with $u+r-j+1$ substituted for u and $u+r+1$ for $u+r$; now, however, the number j of inputs may no longer be sufficient to prevent emptiness from occurring prior to $u+r$ as well as at $u+r$. We conclude that the probability of emptiness at time $2t = u+r$, when there are $j(\leqq r)$ inputs into the dam, subject to zero input in $(u+r,u+r+1)$, is

$$\Pr\{Z_{2t}=0|Z_o=u;\ 2t=u+r;\ S_{u+r-1}=j;\ X_{u+r}=0\}=g(0,u+r+1-j,u+r+1) \tag{2.6}$$

If we remove the restrictive condition that there be zero input in $(u+r,u+r+1)$, and consider that the dam may become empty at time $2t$ when there are any number of inputs $j = 0,1,\ldots,2t-u$, it follows that

$$P_o(2t) = \begin{cases} p_o^{-1} \sum_{j=u+1}^{2t+1} g(0,j,2t+1) & 2t \geqq u \\ 0 & 2t < u \end{cases} \qquad (2.7)$$

Similarly we have

$$P_o(2t+1) = \begin{cases} q_o^{-1} \sum_{j=u+1}^{2t+1} g(0,j,2t+2) & 2t+1 \geqq u \\ 0 & 2t+1 < u. \end{cases} \qquad (2.8)$$

2.3 The time-dependent solution

To solve equations (1.4) and (1.5) we employ the method of transforms. We let

$$\begin{aligned} \alpha_1(s,\theta) &= \sum_{t=o}^{\infty} Q(s,2t+1)\theta^{2t+1}, \ \alpha_2(s,\theta) = \sum_{t=o}^{\infty} Q(s,2t)\theta^{2t} \\ \psi_1(\theta) &= \sum_{t=o}^{\infty} P_o(2t+1)\theta^{2t+1}, \ \psi_2(\theta) = \sum_{t=o}^{\infty} P_o(2t)\theta^{2t} \end{aligned} \qquad |\theta| < 1$$

be the transforms with respect to time of the Q's and the probabilities of emptiness. Taking transforms in (1.4) and (1.5) we readily obtain

$$\alpha_1(s,\theta) = \{s^{u-1}A(s) - p_o(1-s)s^{-1}\theta\psi_2(\theta) - q_o(1-s)s^{-2}A(s)\theta^2\psi_1(\theta)\} \times \left\{1 - \frac{A(s)B(s)}{s^2}\theta^2\right\}^{-1} .$$

$$\alpha_2(s,\theta) = \{s^u - p_o(1-s)s^{-2}B(s)\theta^2\psi_2(\theta) - q_o(1-s)s^{-1}\theta\psi_1(\theta)\} \times$$
$$\left\{1 - \frac{A(s)B(s)}{s^2}\theta^2\right\}^{-1}.$$

These may be expanded, for suitably small values of θ, in powers of θ to give the p.g.f.'s of the content distribution as

$$\begin{aligned} Q(s,2t+1) = {} & s^{u-2t-1}A(s,t+1)B(s,t) - \\ & - p_o(1-s)\sum_{j=o}^{t} P_o(2j)A(s,t-j)B(s,t-j)s^{-2t+2j-1} \\ & - q_o(1-s)\sum_{j=o}^{t-1} P_o(2j+1)A(s,t-j)B(s,t-1-j)s^{-2t+2j} \end{aligned} \qquad (3.1)$$

$$\begin{aligned} Q(s,2t) = {} & s^{u-2t}A(s,t)B(s,t) \\ & -p_o(1-s)\sum_{j=o}^{t-1} P_o(2j)A(s,t-j)B(s,t-j)s^{-2t+2j-1} \\ & -q_o(1-s)\sum_{j=o}^{t-1} P_o(2j+1)A(s,t)1)j)B(s,t-1-j)s^{-2t+2j+1} \end{aligned} \qquad (3.2)$$

These may be expanded for $0 < s < 1$ in powers of s to determine

$$\begin{aligned} P_i(2t+1) = {} & r_{2t+1-u+i}(t+1,t) - p_o\sum_{j=o}^{t-1} P_o(2j)s_{2t-2j+i+1}(t-j,t-j) \\ & -q_o\sum_{j=o}^{t-1} P_o(2j+1)s_{2t-2j+i}(t-j,t-j) \qquad i \geqq u-2t-1 \end{aligned} \qquad (3.3)$$

$$P_i(2t) = r_{2t-u+i}(t,t) - p_o \sum_{j=o}^{t-1} P_o(2j)s_{2t-2j+i}(t-1-j,t-j)$$

$$-q_o \sum_{j=o}^{t-2} P_o(2j+1)s_{2t-2j+i-1}(t-1-j,t-1-j) \qquad i \geqq u-2t, \tag{3.4}$$

where $P_o(2t+1)$ and $P_o(2t)$ are given by (2.7) and (2.8) respectively, and where for simplicity we have put

$$s_o(t,T) = r_o(t,T),\ s_i(t,T) = r_i(t,T) - r_{i-1}(t,T) \qquad (i \geqq 1).$$

The validity of this last inversion rests on proving that the right hand sides of (3.3) and (3.4) vanish for $i < \max(0,u-2t-1)$ and $i < \max(0,u-2t)$ respectively. Let us consider (3.4); the inversion is obviously valid for $2t \leqq u$, and for $2t > u$ we require to show for $-i > 2t-u$ that

$$r_{2t-u-i}(t,t) = p_o \sum_{j=[\frac{u+1}{2}]}^{t-1-[\frac{u}{2}]} P_o(2j)s_{2t-2j-i}(t-1-j,t-j)$$

$$+ q_o \sum_{j=[\frac{u}{2}]}^{t-1-[\frac{u+1}{2}]} P_o(2j+1)s_{2t-2j-i-1}(t-1-j,t-1-j)$$

$$+ \delta_{i,2m}p_o r_o(\tfrac{i}{2} - 1, \tfrac{i}{2})P_o(2t-i)$$

$$+ \delta_{i,2m+1}q_o r_o(\tfrac{i-1}{2}, \tfrac{i-1}{2})\ P_o(2t-i), \qquad i = 1,2,..,2t-u, \tag{3.5}$$

where m is a non-negative integer, δ_{ij} is the Kronecher delta, and $[x]$ is the greatest integer not exceeding x. We consider the process $\bar{Z}_t = Z_o+S_t-t-1$; $\{\bar{Z}_t,t\}$ is a time homogeneous Markov chain. If we use the same type of procedure as Yeo (1961a) we find that with $Z_o = \zeta$

$$s_{2t+1-i-\zeta}(t,t) = \sum_{j=[\frac{i+1}{2}]}^{t-[\frac{\zeta+1}{2}]} s_{2t+1-i}(j,j)g(0,\zeta,2t-2j)$$

$$+ \sum_{j=[\frac{i}{2}]+1}^{t-[\zeta/2]} s_{2j-i}(j-1,j)g(1,\zeta,2t-2j)$$

$$+ \delta_{i,2m} r_o(\tfrac{i}{2} - 1, \tfrac{i}{2})g(1,\zeta,2t_i)$$

$$+ \delta_{i,2m+1} r_o(\tfrac{i-1}{2}, \tfrac{i-1}{2})g(0,\zeta,2t-i+1).$$

Now by summing ζ over the range $(u+1, 2t-i)$ we obtain, after a few easy steps, (3.5). Similarly we can show that $P_i(2t+1) = 0$ for $i < \max(0,u-2t-1)$. We now have a valid expression for the content probabilities in the form

$$P_i(2t+1) = r_{2t+1-u+i}(t+1,t)-p_o\sum_{j=o}^{t-1} P_o(2j)s_{2t-2j+i+1}(t-j,t-j)$$

$$-q_o\sum_{j=o}^{t-1} P_o(2j+1)\, s_{2t-2j+i}(t-j,t-j) \qquad i > \max(0,u-2t-2)$$

(3.6)

$$P_i(2t) = r_{2t-u+i}(t,t) - p_o \sum_{j=o}^{t-1} P_o(2j)s_{2t-2j+i}(t-1-j,t-j)$$

$$-q_o \sum_{j=o}^{t-2} P_o(2j+1)s_{2t-2j+i-1}(t-1-j,t-1-j) \qquad i > \max\ (0,u-2t-1). \tag{3.7}$$

The probability distribution $P_i(t)$ $(i,t = 0,1,2,\ldots)$ of the content of the dam is given by (3.6) or (3.7) depending on whether t is odd or even. When the two input distributions are identical this reduces to (1.2.7).

The method used to obtain these results is equally applicable to the case where there are $K > 2$ independent additive types of input occurring cyclically, although the results become rather unwieldy. In this case we obtain

$$P_i(Kt+n) = \Pr\{Z_{Kt+n} = i|Z_o = u > 0\}$$

$$= r_{Kt+n-u+i}(t_1,\ldots,t_K)$$

$$-\sum_{v=n}^{K-1}{}' \sum_{j=o}^{t-1} p_{K+n-1-v,o}P_{K+n-1-v}(0,Kj+K+n-1-v) \times$$

$$s_{Kt-Kj-K+1+v+i}(t-j_1,\ldots,t-j_K)$$

$$-\sum_{v=o}^{n-1}{}'' \sum_{j=o}^{t} p_{n-1-v,o}P_{n-1-v}(0,Kj+n-1-v) \times$$

$$s_{Kt-Kj+1+v+i}(t-j_1+1,\ldots,t-j_K+1)$$

$$+\sum_{v=o}^{n-1} p_{n-1-v,o}P_{n-1-v}(0,n-1-v)s_{Kt+1+v+i}(t_1,\ldots,t_K)$$

$$i > 0, t > 0, n = 0,1,\ldots,K-1, \tag{3.8}$$

where $\{p_{o,i}\},\ldots,\{p_{K-1,i}\}$ are the K input distributions occurring in the time intervals $(Kt,Kt+1),\ldots,(Kt+K-1,Kt+K)$; $r_i(t_1,\ldots,t_K)$ and $s_i(t_1,\ldots,t_K)$ are the natural generalisations of $r_i(t_1,t_2)$ and $s_i(t_1,t_2)$; $j \leqq j_i \leqq j+1$, $t \leqq t_i \leqq t+1$ $(1 \leqq i \leqq K)$, and the values of the j_i and t_i can readily be determined from the difference equations for the p.g.f.'s; the first term needs to be slightly modified if $0 < u < n$; $\sum'$ and $\sum''$ indicate that when $v = n = 0$, and $v = 0$, the j sums range over $(0,t-2)$ and $(0,t-1)$ respectively; and the last term of (3.8) is non-zero only if $u \leqq n-1$. The probabilities of emptiness are given by

$$P_o(Kt+n) = p_{n,o}^{-1} \sum_{v=n+1}^{Kt+n+1} g(0,j,Kt+n+1) \qquad 0 \leqq n \leqq K-1, \tag{3.9}$$

and we can find recurrence relations for the probabilities $g(0,u,Kt+n)$ of first emptiness of the dam.

2.4 Stationary distributions

The stationary probabilities $P_1(i) = \lim_{t\to\infty} P_i(2t+1)$ and $P_2(i) = \lim_{t\to\infty} P_i(2t)$ exist, as the limits of (3.6) and (3.7); these are proper distributions if and only if the mean input is less than unity i.e. $\frac{1}{2}(\rho_1+\rho_2) < 1$, where $\rho_1 = A'(1)$ and $\rho_2 = B'(1)$. The first terms of (3.6) and (3.7) converge to zero as $t \to \infty$, and the remaining terms give

$$P_1(i) = -p_o P_2(o) \sum_{j=1}^{\infty} s_{2j+1+i}(j,j) - q_o P_1(0) \sum_{j=1}^{\infty} s_{2j+i}(j,j-1).$$

$$i > 0 \tag{4.1}$$

$$P_2(i) = -p_o P_2(0) \sum_{j=1}^{\infty} s_{2j+i}(j,j) - q_o P_1(0) \sum_{j=1}^{\infty} s_{2j-1+i}(j,j-1)$$

$$i > 0, \qquad (4.2)$$

subject to $\sum_{i=o}^{\infty} P_1(i) = \sum_{i=o}^{\infty} P_2(i) = 1$. Determining the sums of the infinite series of (4.1) and (4.2) is likely to be very difficult, unless the series happen to converge quite rapidly when good approximations may be made; we shall concentrate on a more direct method.

We can obtain the p.g.f.'s

$$Q_1(s) = \sum_{i=o}^{\infty} P_1(i)s^i, \quad Q_2(s) = \sum_{i=o}^{\infty} P_2(i)s^i \qquad |s| \leq 1$$

of the stationary content distributions from a consideration of the initial difference equations which are of the same form as (1.2) and (1.3); after a few easy steps we obtain

$$Q_1(s) = \frac{p_o(1-s)sP_2(0)+q_o(1-s)A(s)P_1(0)}{A(s)B(s)-s^2} \qquad (4.3)$$

$$Q_2(s) = \frac{p_o(1-s)B(s)P_2(0)+q_o(1-s)sP_1(0)}{A(s)B(s)-s^2}, \qquad (4.4)$$

where for the present $P_1(0)$ and $P_2(0)$ are unknown. However, we can find them from (4.3) and (4. 4) (c.f. Bailey (1954) and §1.2). As $s \to 1 - 0$ we find that

$$p_o P_2(0) + q_o P_1(0) = 2 - \rho_1 - \rho_2. \qquad (4.5)$$

We note that when the input distributions are identical (4.5) reduces to the relation $P(0) = (1-\rho)p_o^{-1}$ obtained by Moran (1956). The denominator $C(s) = A(s)B(s)-s^2$ of (4.3) and (4.4) has one and only one zero at $s = 1$, since $\rho_1+\rho_2 < 1$. We have $|A(-s)| \leqq A(s)$ and $|B(-s)| \leqq B(s)$ for $s \geqq 0$. Suppose $A(s)$ and $B(s)$ exist and are finite for $|s| < 1+\delta$ for some $\delta > 0$. For sufficiently small $\delta > 0$ we have $|s|^2 > |A(s)B(s)|$ on $|s| = 1+\delta$, as $\rho_1+\rho_2 < 1$. Hence by Rouché's theorem $C(s) = 0$ has exactly two roots in $|s| \leqq 1$, namely $s = 1$ and $s = s_1$ ($|s_1| < 1$) which are distinct. For a proper p.g.f. $Q_1(s)$ to exist, $Q_1(s)$ must be finite for $|s| \leqq 1$, so that there must be a zero at $s = s_1$ of the numerator of (4.3) as well as a zero of the denominator, and so

$$p_o s_1 P_2(0)+q_o A(s_1)P_1(0) = 0. \tag{4.6}$$

The equations (4.5), (4.6) are consistent as the determinant

$$\begin{vmatrix} p_o & q_o \\ p_o s_1 & q_o A(s_1) \end{vmatrix} = p_o q_o (A(s_1)-s_1) \neq 0;$$

if $A(s_1) = s_1$ then $B(s_1) = s_1$, which is not possible; for suppose $\rho_1 < 1$ (if not then certainly $\rho_2 < 1$ and we consider $B(s)-s$), then by Rouché's theorem $A(s)-s$ has only the one root $s = 1$ in the region $|s| \leqq 1$. Solving (4.5), (4.6) we obtain

$$P_1(0) = \frac{-s_1(2-\rho_1-\rho_2)}{q_o(A(s_1)-s_1)}$$

$$P_2(0) = \frac{A(s_1)(2-\rho_1-\rho_2)}{p_o(A(s_1)-s_1)} \tag{4.7}$$

We now have explicit expressions for the p.g.f.'s $Q_1(s)$ and $Q_2(s)$ which we wish to expand to give the stationary probabilities. In general this may be done by determining the remaining roots of $C(s)$, breaking up $Q_1(s)$ and $Q_2(s)$ into partial fractions and expanding (c.f. Prabhu (1958)).

As an example we consider two geometric input distributions $A(s) = a(1-bs)^{-1}$, $B(s) = \alpha(1-\beta s)^{-1}$, $0 < a < 1$, $0 < \alpha < 1$, $b = 1-a$, $\beta = 1-\alpha$, $\rho_1 = ba^{-1}$, $\rho_2 = \beta\alpha^{-1}$; then

$$Q_1(s) = \frac{(s-1)(1-\beta s)\{a(1-bs)sP_2(0)+a\alpha P_1(0)\}}{b\beta s^4-(b+\beta)s^3+s^2-a\alpha} .$$

$$Q_2(s) = \frac{(s-1)(1-bs)\{a\alpha P_2(0)+\alpha(1-\beta s)P_1(0)\}}{b\beta s^4-(b+\beta)s^3+s^2-a\alpha} \tag{4.8}$$

The denominator of $Q_1(s)$ and $Q_2(s)$ has four roots $s = 1$, s_1, s_2, s_3 with only $|s_1| < 1$, say. From (4.7) we have

$$P_1(0) = \frac{-s_1(1-bs_1)(2-ba^{-1}-\beta\alpha^{-1})}{\alpha\{a-s_1(1-bs_1)\}}$$

$$P_2(0) = \frac{2-b\,a^{-1}-\beta\alpha^{-1}}{a-s_1(1-bs_1)} . \tag{4.9}$$

We break (4.8) up into partial fractions of the form

$$Q_1(s) = aP_2(0)-N_1(s_2)(s-s_2)^{-1}+N_1(s_3)(s-s_3)^{-1}$$

$$Q_2(s) = \alpha P_1(0)-N_2(s_2)(s-s_2)^{-1}+N_2(s_3)(s-s_3)^{-1}, \tag{4.10}$$

where

$$N_1(s) = [\{b\beta(s_2+s_3)-(b+\beta-b\beta s_1)\}aP_2(0)s-a\alpha\beta P_1(0)$$

$$+\{(1-bs_1)(1-\beta s_1)-b\beta s_2 s_3)\}aP_2(0)]\{b\beta(s_3-s_2)\}^{-1}$$

$$N_2(s) = [\{b\beta(s_2+s_3)-(b+\beta-b\beta s_1)\}\alpha P_1(0)s-a\alpha bP_2(0)$$

$$+\{(1-bs_1)(1-\beta s_1)-b\beta s_2 s_3\}\alpha P_1(0)]\{b\beta(s_3-s_2)\}^{-1}.$$

We can now readily expand (4.10) to give

$$\left.\begin{aligned} P_1(i) &= N_1(s_2)s_2^{-i-1} - N_1(s_3)s_3^{-i-1} \\ P_2(i) &= N_2(s_2)s_2^{-i-1} - N_2(s_3)s_3^{-i-1}. \end{aligned}\right\} \quad i > 0$$

As examples we put (i) $a = 0.5$, $\alpha = 0.6$ (ii) $a = 0.4$, $\alpha = 0.8$. Then (i) $Q_1'(1) = 4.3332$, $Q_2'(1) = 4.0257$, (ii) $Q_1'(1) = 8.1528$, $Q_2'(1) = 7.4540$ and

	$P_1(i)$		$P_2(i)$	
i	(i)	(ii)	(i)	(ii)
0	0.2930	0.1891	0.3150	0.2467
1	0.1138	0.0815	0.1136	0.0797
2	0.0965	0.0732	0.0940	0.0678
3	0.0809	0.0658	0.0782	0.0611
4	0.0679	0.0592	0.0653	0.0549
5	0.0567	0.0532	0.0546	0.0484
$\sum_{i=6}^{\infty}$	0.2902	0.4780	0.2793	0.4404

It would be useful if we could find some simple approximations when, as is likely in practice, there are quite different wet and dry seasons. We could consider an input p.g.f. $A(s)B(s)$ with possible releases of size two occurring only at the end of dry seasons i.e. 2t-0; however, this is not any simpler to deal with than the model we have used. A simpler model is one where there is an input p.g.f. $\{A(s)B(s)\}^{\frac{1}{2}}$ with a possible unit release at the end of each interval. Although this has been solved by Moran (1956), Yeo (1961a) it is not very useful as the approximations have been found in the cases considered to be poor; in the examples considered above the stationary probabilities of emptiness are (i) 0.3043 and (ii) 0.2210, which are not very useful.

2.5 A continuous analogue

We now consider a continuous analogue of our problem by allowing the units of measurement to tend to zero in a similar manner to that of Chapter 1. We suppose for the discrete problem that there are K types of input occurring cyclically. There is a unit of $\Delta(> 0)$ so that content, time, input and release are measured in multiples of Δ. We put

$$\text{(a)} \quad p_{j,i\Delta}(\Delta) = \begin{cases} 1-\lambda_j\Delta & i = o \\ \lambda_j\Delta\{A_j(i\Delta)-A_j((i-1)\Delta)\} & i = 1,2,\ldots \end{cases}$$

$$\text{(b)} \quad i = x\Delta^{-1}, \; t = \tau\Delta^{-1}, \; u = U\Delta^{-1}, \qquad (5.1)$$

where x, τ, U are real and non-negative, $A_j(i\ )$ $(1 \leq j \leq K)$ are cumulative probabilities with $\lim_{i\to\infty} A(i\Delta) = 1$, and such that in the limit as $i \to \infty$, $\Delta \to 0$, $i\Delta \to x$ we have $\lim\limits_{\Delta \to 0} A_j(i\Delta) = C_j(x)$ $(0 \leq x < \infty)$, where $C_j(x)$ is a d.f.

Under the condition (5.1) the set of K equations of the type (1.2) and (1.3) that we obtain for the discrete process tend as $\Delta \to 0$ to the integro-differential equation

$$\frac{\partial}{\partial \tau} W(x,\tau) - \frac{\partial}{\partial x} W(x,\tau) = - \sum_{j=1}^{K} \lambda_j \left\{ W(x,\tau) - \int_{v=0}^{x} W(x-v,\tau) dC_j(x) \right\}, \tag{5.2}$$

where $W(x,\tau)$ $(0 \leq x, \tau < \infty)$ is the d.f. of the content for the continuous time process. We now have a process which is equivalent to the single server queueing system with arrivals occurring in K independent Poisson processes with service time distributions having a general form depending on the class of the customer. The inputs no longer occur cyclically, but independently; this is thus different from the problem of Gani and Pyke (1962) where inputs in a continuous time process do occur cyclically. We have reduced a set of K equations for the discrete problem to a single equation for the continuous analogue.

A solution of the equation (5.2) may be obtained from (3.8) by using the substitution (5.1) and taking the limit $\Delta \to 0$, but it is simpler to solve (5.2) directly. This is of the form (1.3.2) with capacity $K = \infty$ and $\lambda(\tau)F(x) = \sum_{i=1}^{K} \lambda_i C_i(x)$, and can be solved by the same method as used in Chapter 1. The time-dependent solution is given by (1.3.3) to (1.3.7) with $\lambda(\tau)F(x) = \sum_{i=1}^{K} \lambda_i C_i(x)$ and $K(x,\tau)$ is such that its IST is

$$\int_{x=0}^{\infty} e^{-\theta x}\, dK(x,\tau) = e^{-\tau\xi(\theta)} ,$$

where

$$\xi(\theta) = \sum_{i=1}^{K} \lambda_i(1-\xi_i(\theta)),\ \xi_i(\theta) = \int_0^{\infty} e^{- x} dC_i(x) \qquad \theta \geq 0.$$

When K = 1 $K(x,\tau)$ is given by (1.3.8) and when K = 2

$$K(x,\tau) = \sum_{n=0}^{\infty} \sum_{j=0}^{n} e^{-(\lambda_1+\lambda_2)\tau} \binom{n}{j}\lambda_1^j\lambda_2^{n-j} \frac{t^n}{n!} \int_{y=0}^{x} C_1^{j*}(x-y)\,dC_2^{(n-j)*}(y). \tag{5.3}$$

The LST $\mathcal{H}_K(\theta) = \int_0^{\infty} e^{-\theta} dW_K(x)$ (Rl $\theta \geq 0$) of the stationary content (or waiting time) distribution, which exists for $\sum_{i=1}^{K} \lambda_i\rho_i < 1$, where $\rho_i = -\alpha_i'(0)$, is from (1.3.9)

$$\mathcal{H}_K(\theta) = \left\{1 - \sum_{i=1}^{K} \lambda_i\rho_i\right\} / \left\{1 - \theta^{-1}\xi(\theta)\right\}. \tag{5.4}$$

The mean waiting time $\bar{h}_K = -\mathcal{H}_K'(0)$ is

$$\bar{h}_K = \frac{\sum_{i=1}^{K}\lambda_i\xi_i''(0)}{2(1 - \sum_{i=1}^{K}\lambda_i\rho_i)} . \tag{5.5}$$

In the context of queueing theory discussed in Part 2, we can also consider the distribution of the number of customers in the queue, including the one being served. The p.g.f. $r_K(z) = \sum_{i=0}^{\infty} r_i z^i$ ($|z| \leq 1$) of the stationary queue size distribution may be obtained as

$$r_K(z) = \frac{(1 - \sum_{i=1}^{K}\lambda_i\rho_i)(1-z)\xi(\nu_K-\nu_K z)}{\xi(\nu_K-\nu_K z)} , \tag{5.6}$$

where $\nu_K = \sum_{i=1}^{K}\lambda_i$. The LST $\Lambda_K(\theta) = \int_0^{\infty} e^{-\theta x}\, dG_K(x)$ (Rl $\theta \geq 0$)

of the distribution of the length of a busy period may be found as for $K = 1$ [Takács (1955)] as the unique solution in real $\theta > 0$ with $\lim_{\theta\to\infty} \Lambda_K(\theta) = 0$ of

$$\Lambda_K(\theta) = \nu_K^{-1}\ \xi(\nu_K+\theta-\nu_K\Lambda_K(\theta)). \tag{5.7}$$

PART 2

SINGLE SERVER QUEUES

CHAPTER 3

QUEUES WITH MODIFIED SERVICE MECHANISMS

3.1 General remarks

The problems in the theory of dams described in Part I may be considered as particular examples of queueing theory; however, it has been found convenient to develop dam theory along slightly different lines, particularly for dams with a finite capacity. In this Part some problems of single server queueing systems are discussed, which have some similarity to the infinite dam problems of Part I, but which will be considered separately.

Work in queueing theory, which has been applied to many branches of science in the last half century, has been concerned mainly with (i) the properties of the process when it has settled down to statistical equilibrium, (ii) the busy period distribution and (iii) the time-dependent properties of waiting time and queue size. We are concerned in the next two chapters primarily with the first two problems, in contrast to Part I where the time-dependent problem for dams was studied in the most detail.

Queueing problems may be specified by the following three properties: (i) the inter-arrival time distribution, (ii) the service mechanism and (iii) the queue discipline. We shall consider only single server queues; in this chapter the queue discipline is first-come first-served, i.e. all customers are served strictly in the order of their arrival, while in Chapter 4 we shall discuss problems where some types of customer have precedence to service over other types of customer.

Throughout we suppose that the time intervals between the arrival of successive customers (or bunches of customers) are random variables which are independently and identically distributed. This distribution is in most cases taken to be negative exponential, so that customers arrive independently in a Poisson process.

We suppose that the service time of a customer is a random variable, whose distribution in the most general case may be dependent on (a) the waiting time in the queue when the customer arrives or (b) the queue size at this point. Although it is possible to write down equations for the waiting time and the joint waiting time-queue size distributions, it has not been found possible to obtain important results for the general case; however, it is possible to solve some special cases, one of which we are concerned with in this chapter. Here the service time of a customer may be different if he arrives at an empty or a non-empty queue.

3.2 Introduction

Let us consider the following generalisation of the queueing system GI/G/1. In a single server system customers arrive at a counter at the instants $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ such that $t_n = \tau_n - \tau_{n-1}$ $(n \geq 1)$, $\tau_o = 0$, are independently distributed random variables with common d.f.'s $A(x)$, finite means $a = E(t_n) = \int_0^\infty x dA(x)$ and LST's $a^*(\theta) = \int_{0-}^\infty e^{-\theta x} dA(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$. If the n-th customer joins a non-empty queue, let his service time be s_n, while if he joins an empty queue let it be r_n; $\{s_n\}$ and $\{r_n\}$ $(n \geqq 1)$ are sequences of independently and identically distributed random variables, which are independent of t_n, with common d.f.'s $B(x)$ and $D(x)$ $(0 \leqq x < \infty)$ respectively, finite but non-zero means $b = E(s_n) = \int_0^\infty x dB(x)$ and $d = E(r_n) = \int_0^\infty x dD(x)$, and LST's $\psi(\theta) = \int_{0-}^\infty e^{-\theta x} dB(x)$ and $\zeta(\theta) = \int_{0-}^\infty e^{- \ x} dD(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$. Let w_n be the time the n-th arrival waits before commencing service; then

$$w_{n+1} = \begin{cases} w_n + u_n & w_n + u_n > 0,\ w_n > 0 \\ c_n & c_n > 0,\ w_n = 0 \\ 0 & \text{otherwise}, \end{cases} \tag{2.1}$$

where $u_n = s_n - t_{n+1}$ and $c_n = r_n - t_{n+1}$; $\{u_n\}$ and $\{c_n\}$ are independently distributed random variables with the d.f.'s

$$U(x) = \int_{0-}^\infty B(x+y) dA(y),\ C(x) = \int_{0-}^\infty D(x+y) dA(y), \quad -\infty < x < \infty, \tag{2.2}$$

with means $E(u_n) = b-a$, $E(c_n) = d-a$ respectively $(E(|u_n|) < \infty$, $E(|c_n|) < \infty)$.

This queueing system may arise in several ways, such as when a machine is shut down when no items remain to be served, and the service mechanism is different for the first item of a group. It will be used in the priority queueing in Chapter 4, and in the traffic problems of Part 3.

Finch (1959) has considered a process which differs from ours only when the n-th arrival joins an empty queue; the customer then waits a time v_n before commencing a service of length s_n, instead of receiving an immediate service of r_n. If we put $D(x) = \int_0^x B(x-y)dV(y)$, where $V(x)$ is the d.f. of the v_n, then we can compare the results obtained for the two systems, as we do in §3.9.

3.3 The stationarity condition

For the GI/G/1 queueing system it has been shown by Lindley (1952) that a unique proper stationary waiting time distribution exists if and only if $E(u) = b-a < 0$. Finch (1959) has shown that this is the stationarity condition for his process, provided $E(v) < \infty$, and we shall show that it is also a necessary and sufficient condition, provided only $0 < d < \infty$, for the existence of a proper stationary waiting time distribution for our process. The condition is independent of the service time distribution for a customer joining a non-empty queue; we may intuitively expect this, for regardless of the (non-zero) finite size of the service time for a customer finding the queue empty on arrival, the waiting time must eventually reduce to zero again with probability unity if $E(u) < 0$, and may build up indefinitely if $E(u) \geqq 0$. We shall prove

Theorem 3.1. A sequence of non-negative random variables $\{w_n\}$ is defined by (2.1) for $n = 1,2,\ldots$, and w_0 is a given non-negative random variable with d.f. $W_0(x)$ $(0 \leqq x < \infty)$, where $\{u_n\}$ and $\{c_n\}$ are defined above in §2. Write $W_n(x) = \Pr\{w_n \leqq x\}$ $(n = 1,2,\ldots)$, then $W(x) = \lim_{n\to\infty} W_n(x)$ exists. If $E(u) \geqq 0$ and $d > 0$ then $W(x) = 0$ for all $x \geqq 0$; if $E(u) < 0$ and $0 < d < \infty$ then $W(x)$ is the d.f. of a non-negative random variable, $W(x)$ is independent of W_o, and is the unique d.f. solution of the integral equation

$$W(x) = \begin{cases} 0 & x < 0 \\ \int_{-\infty}^{x} W(x-y)dU(y) - W(0)\{U(x)-C(x)\} & x \geqq 0. \end{cases} \tag{3.1}$$

Proof. Since $w_n \geqq 0$ we have $\Pr\{w_n \leqq x\} = 0$ for $x < 0$. For $x \geqq 0$

$$\begin{aligned} \Pr\{w_{n+1} \leqq x\} &= \Pr\{w_n+u_n \leqq x,\ w_n > 0\} + \Pr\{w_n = 0,\ c_n \leqq x\} \\ &= \Pr\{w_n+u_n \leqq x,\ u_n \leqq x\} + \Pr\{w_n = 0,\ c_n \leqq x\} \\ &\quad - \Pr\{w_n = 0,\ u_n \leqq x\}. \end{aligned} \tag{3.2}$$

Thus

$$W_{n+1}(x) = \begin{cases} 0 & x < 0 \\ \int_{-\infty}^{x} W_n(x-y)\, dU(y) - W_n(0)\, \{U(x) - C(x)\} & x \geqq 0. \end{cases} \tag{3.3}$$

From (3.2) by iteration we have

$$W_{n+1}(x) = A_n(x) - L_n(x) \tag{3.4}$$

where $L_n(x) = H_n(x) - G_n(x)$ and

$$A_n(x) = \Pr\{w_o+u_o+u_1+\ldots+u_n \leqq x, u_1+\ldots+u_n \leqq x, \ldots, u_{n-1}+u_n \leqq x, u_n \leqq x\}$$

$$\begin{aligned} H_n(x) &= \sum_{i=1}^{n} \Pr\{w_i = 0, u_i+u_{i+1}+\ldots+u_n \leqq x, u_{i+1}+\ldots+u_n \leqq x, \ldots, u_{n-1}+u_n \leqq x, u_n \leqq x\} \\ &= \sum_{i=1}^{n} \Pr\{w_i = 0\} \Pr\{u_i+u_{i+1}+\ldots+u_n \leqq x, u_{i+1}+\ldots+u_n \leqq x, \ldots \\ &\qquad \ldots, u_{n-1}+u_n \leqq x, u_n \leqq x\} \end{aligned}$$

$$= \sum_{i=1}^{n} \Pr\{w_{n+1-i} = 0\}\Pr\{u_1+u_2+\dots+u_{n+1-i} \leqq x, \dots, u_1+u_2 \leqq x, u_1 \leqq x\}$$

$$G_n(x) = \sum_{i=1}^{n} \Pr\{w_i = 0, c_i+u_{i+1}+\dots+u_n \leqq x, u_{i+1}+\dots+u_n \leqq x, \dots$$
$$\dots, u_{n-1}+u_n \leqq x, u_n \leqq x\}$$

$$= \sum_{i=1}^{n} \Pr\{w_i = 0\}\Pr\{c_i+u_{i+1}+\dots+u_n \leqq x, u_{i+1}+\dots+u_n \leqq x, \dots$$
$$\dots, u_{n-1}+u_n \leqq x, u_n \leqq x\}$$

$$= \sum_{i=1}^{n} \Pr\{w_{n+1-i} = 0\}\Pr\{u_1+\dots+u_{n-i}+c_{n+1-i} \leqq x, \dots$$
$$\dots, u_1+u_2 \leqq x, u_1 \leqq x\}; \tag{3.5}$$

here we have used the fact that $\Pr\{w_i = 0\}$ is independent of $\{u_i, u_{i+1}, \dots, u_n\}$ and $\{c_i, u_{i+1}, \dots, u_n\}$.

By defining a sequence of random variables $\{w_n^*\}$ by

$$w_{n+1}^* = \begin{cases} w_n^* + u_n & w_n^* + u_n > 0 \\ 0 & w_n^* + u_n \leqq 0 \end{cases} \qquad n = 1, 2, \dots \tag{3.6}$$

with $w_o^* = w_o$, we obtain $\Pr\{w_{n+1}^* \leqq x\} = A_n(x)$. Following Lindley (1952), we see that $A(x) = \lim_{n\to\infty} A_n(x)$ exists, is independent of $W_o(x)$, and if $E(u) \geqq 0$ then $A(x) = 0$ and if $E(u) < 0$ then $A(x)$ is a d.f.

Let us now consider $E(u) \geqq 0$. This includes both $E(u) > 0$ and $E(u) = 0$, which may be taken together if we make use of the results of Lindley (1952), who employs the strong law of large numbers for $E(u) > 0$ and a result of Chung and Fuchs (1951) for $E(u) = 0$.

We distinguish two cases: (i) $C(0) - U(0) \leqq 0$ and (ii) $C(0) - U(0) > 0$. In case (i) $L_n(0) \geqq 0$ and therefore $W_n(0) \leqq A_n(0)$;

it follows from Lindley (1952) that $\lim_{n\to\infty} W_n(0) = 0$, and as a consequence $\lim_{n\to\infty} W_n(x) = 0$ for all x $(0 \leqq x < \infty)$. For case (ii) we have $G_n(0) > H_n(0)$. We write

$$W_{n+1}(0) = A_n(0) + \sum_{i=1}^{n_o} + \sum_{i=n_o+1}^{n} \Pr\{w_i = 0\}[\Pr\{u_i+\ldots+u_{n+1-i} \leqq 0, \ldots$$

$$\ldots, u_1+u_2 \leqq 0, u_1 \leqq 0\} - \Pr\{u_i+\ldots+u_n+c_{n+1-i} \leqq 0, \ldots$$

$$\ldots, u_1+u_2 \leqq 0, u_1 \leqq 0\}] \qquad (3.7)$$

for $1 < n_o < n$. By a suitable choice of n_o for n sufficiently large we can show that $\lim_{n\to\infty} W_n(0) = 0$ and hence $\lim_{n\to\infty} W_n(x) = 0 \quad (0 \leqq x < \infty)$.

The remainder of this section is concerned with the case $E(u) < 0$ and $0 < d < \infty$. Let $\mathcal{E}$ be the event $\{w_n = 0\}$; $\mathcal{E}$ is an aperiodic certain recurrent event [Lindley (1952), Feller (1957)]. The $\lim_{n\to\infty} \Pr\{w_n = 0\}$ exists, and is greater than zero (and $\leqq 1$), i.e. the mean recurrence step number is finite. The mean recurrence step number of the event $\mathcal{E}$ is not greater than the expectation of the number of steps to the first zero of the sequence of random variables $\{w_n^*\}$ given by (3.6) with w_o being given in the following way: if $r_1 < s_1$ choose $w_o{}^* = 0$, and if $r_1 \geqq s_1$ choose $w_o{}^* = r_1 - s_1$; these define $W_o(x)$. Since $E(u) < 0$ and $E(r) < \infty$ the required result follows from the theorem of Lindley (1952). We also have $\min_{1\leqq i<n} \Pr\{w_i = 0\} > 0$ for any finite value of n, so that $\min_{1\leqq i<\infty} \Pr\{w_i = 0\} > 0$. It follows from the expression for $L_n(x)$ that $L(x) = \lim_{n\to\infty} L_n(x)$ exists, is independent of w_o and is given by

$$L(x) = [\lim_{n\to\infty} \Pr\{w_n = 0\}] \sum_{i=1}^{\infty} \{g_i(x) - h_i(x)\},$$

where

$$h_i(x) = \Pr\{u_1+u_2+\ldots+u_i > x, u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\}$$

$$g_i(x) = \Pr\{c_1+u_2+\ldots+u_i > x, u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\}.$$

This last follows as

$$\Pr\{u_1+u_2+\ldots+u_i \leq x, u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\}$$

$$= \Pr\{u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\} - \Pr\{u_1+u_2+\ldots+u_i > x, u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\}$$

$$\Pr\{c_1+u_2+\ldots+u_i \leq x, u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\}$$

$$= \Pr\{u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\} - \Pr\{c_1+u_2+\ldots+u_i > x, u_2+\ldots+u_i \leq x, \ldots, u_i \leq x\}.$$

From (3.4) we have

$$-1 \leq L_n(x) = \sum_{i=1}^{n} \Pr\{w_{n+1-i} = 0\} \{g_i(x) - h_i(x)\} \leq 1.$$

Now it may be shown [c.f. Finch (1959), p.319 with $v_s = 0$] that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \Pr\{w_{n+1-i} = 0\} h_i(x) \leq \sum_{i=1}^{\infty} h_i(x) < \infty,$$

and as a result we can show that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \Pr\{w_{n+1-i} = 0\} g_i(x) \leq \sum_{i=1}^{\infty} g_i(x) < \infty.$$

We have shown that $W(x)$ exists independently of w_o. The existence of $W(x)$ could also be deduced from the result of Finch (1959).

For whatever the value of $\{r_n\}$ there exists a non-negative random variable v_n with d.f. $V(x)$ and mean $v < \infty$ such that $D(x) < \int_0^x B(x-y)dV(y)$ thus our process is bounded above by that of Finch, for which $W(x)$ exists, so $W(x)$ exists for our process as well.

Let us for the moment consider the problem in terms of a random walk with an impenetrable barrier in which a particle starts at the origin, so that $w_1 = 0$. Following Lindley (1952) let the next jump which meets the origin be the n_1th, giving $w_{n_1+1} = 0$, and put $v_{n_1+1} = w_{n_1} + u_{n_1}$, so that v_{n_1} is the amount that the particle would have gone over the barrier if it had not been there. Let the next jump meeting the origin be the n_2th, define v_{n_2+1} similarly, and so on. We have

$$v_{n_1+1} = c_1 + \sum_{i=2}^{n_1} u_i, \ \ldots, \ v_{n_r+1} = c_r + \sum_{i=n_{r-1}+1}^{n_r} u_i,$$

so that

$$\frac{v_{n_1+1} + v_{n_2+1} + \ldots + v_{n_r+1}}{r} - \frac{c_1+c_2+\ldots+c_r}{r} + \frac{U_r}{r} = \frac{U_{n_r}}{n_r} \cdot \frac{n_r}{r}, \tag{3.8}$$

where $U_n = \sum_{i=1}^n u_i$. Now $|v_{n_r+1}| \leq |u_{n_r}|$, and hence v_{n_r}, which is a random variable, has a finite mean $E(v)$. The left hand side of (3.8) tends, by the strong law of large numbers, to $E(v) - E(c) + E(u)$; U_{n_r}/n_r tends to $E(u)$ (< 0), and thus $n_r/r \to \{E(v)-E(c)+E(u)\}/E(u)$. By the converse of the strong law of large numbers $(n_r - n_{r-1})$ is a random variable with finite mean equal to $\{E(v) - E(c) + E(u)\}/E(u)$; this is the mean recurrence time. The theory of Feller (1957) shows that the mean recurrence time is $\{W(0)\}^{-1}$, and thus $W(0) > 0$.

We have shown that $W(x)$ exists independently of w_o; it remains to show that $W(x)$ is a d.f. Since $W(x) = \lim_{n\to\infty} W_n(x)$ and the $W_n(x)$ are non-decreasing functions of x, it is easily shown that $W(x)$ is a non-decreasing function of x. We require to show that $\lim_{x\to\infty} W(x) = 1$. The joint occurrence of $w_n+u_n \leqq x$ and $c_n \leqq x$ together imply $w_{n+1} \leqq x$, so that

$$\Pr\{w_{n+1} \leqq x\} \geqq \Pr\{w_n+u_n \leqq x, c_n \leqq x\};$$

without loss of generality let us take $w_o = 0$, then by iteration

$$\Pr\{w_{n+1} \leqq x\} \geqq \Pr\{w_o+u_o+u_1+\ldots+u_n \leqq x, c_1+u_2+\ldots+u_n \leqq x, \ldots c_{n-1}+u_n \leqq x \; c_n \leqq x\}$$

$$= \Pr\{u_1+\ldots+u_{n-1}+c_n \leqq x, \ldots, u_1+c_2 \leqq x, c_1 \leqq x\}$$

$$= \Pr\{V_i \leqq x \text{ all } i \text{ in } 1 \leqq i \leqq n\}$$

where $V_i = u_1+u_2+\ldots+u_{i-1}+c_i \quad (i = 1,2,\ldots)$.

The events $M_n(x) = [V_i \leqq x \text{ all } i \text{ in } 1 \leqq i \leqq n]$ are monotonic non-increasing functions of n and hence by the property of probability measure $\lim_{n\to\infty} \Pr\{M_n(x)\} = \Pr\{M(x)\}$, where $M(x)$ is the limit event $[V_i \leqq x \text{ all } i \geqq 1]$. Thus

$$W(x) \geqq \Pr\{M(x)\}.$$

In order to prove that $\lim_{x\to\infty} W(x) = 1$, it is sufficient to show that $\lim_{n\to\infty} \Pr\{M_n(x)\} = 1$. This follows from the law of large numbers, since

$E(u) < 0$ and $-\infty < E(c) < \infty$, by the same type of argument used by Lindley (1952) in the corresponding part of his theorem. Thus $W(x)$ is a d.f.

The fact that $W(x)$ is a solution to the integral equation (3.2) follows from (3.4) by Lebesgue's dominated convergence theorem. If $W^*(x)$ is another solution of (3.2) which is a d.f., take $W_o(x) = W^*(x)$, then from (3.4) we have that $W_n(x) = W^*(x)$ for each n and hence $W^*(x) = W(x)$. This completes the proof.

3.4 The time-dependent problem

We now briefly consider the time-dependent queueing problem, where the inter-arrival time distribution is assumed to be negative exponential, i.e. customers arrive in a homogeneous Poisson process so that in a small interval of time $\delta t > 0$ the probability of (i) no arrivals is $1-\lambda\delta t+o(\delta t)$, (ii) one arrival is $\lambda\delta t+o(\delta t)$ and (iii) more than one arrival is $o(\delta t)$. We define the d.f. $W(u,x,t) \equiv W(x,t)$ as the probability that at time t the waiting time in the queue is not greater than x, given that it was u at time zero. For $x \geqq 0$ we have

$$W(x,t+\delta t) = (1-\lambda\delta t)W(x+\delta t,t)+\lambda\delta t\int_{y=o}^{x} W(x-y,t)dB(y)$$
$$+\lambda\delta t W(o,t)\{D(x)-B(x)\}+o(\delta t). \tag{4.1}$$

We have $W(x,t) = 0$ for all $x < 0$; $W(x,t)$ is continuous for all $x > 0$, but has a discontinuity at $x = 0$, the concentration $W(0,t)$ being the probability that the server is idle at time t. Let $\frac{\partial W(x,t)}{\partial x}$ be a right derivative of $W(x,t)$. From (4.1) by letting $t \to 0$ we obtain

$$\frac{\partial W(x,t)}{\partial t} - \frac{\partial W(x,t)}{\partial x} = -\lambda W(x,t) + \lambda \int_{y=o}^{x} W(x-y,t)\,dB(y)$$

$$+\lambda W(o,t)\{D(x)-B(x)\} \qquad x \geqq 0. \qquad (4.2)$$

If we assume more generally that the service time d.f. $B(x|y)$ is a function of the waiting time y at the point when a customer arrives then we would obtain

$$\frac{\partial}{\partial t} W(x,t) - \frac{\partial}{\partial x} W(x,t) = -\lambda W(x,t) + \lambda \int_{y=o}^{x} W(x-y,t)\,dB(y|x-y), \qquad (4.3)$$

which has (4.2) as a special case. However, it has not been found possible to solve (4.3) explicitly, and we shall concentrate on the solution to (4.2). If $B(x) = D(x)$ then (4.2) reduces to the integro-differential equation of Takács (1955) and has the solution (1.3.4).

Taking transforms in (4.2) we obtain a differential equation for the IST $\Omega(\theta,t) = \int_o^\infty e^{-\theta x} dW(x,t)$ (Rl $\theta \geqq 0$) of the waiting time d.f. $W(x,t)$ as

$$\frac{\partial}{\partial t} \Omega(\theta,t) = \Omega(\theta,t)\{\theta - \lambda(1-\psi(\theta))\} - W(o,t)\{\theta + \lambda(\psi(\theta) - \zeta(\theta))\}, \qquad (4.4)$$

which has the solution

$$\Omega(\theta,t) = e^{-\theta u + \theta t - \lambda t(1-\psi(\theta))} - \int_o^t W(0,\tau)\{\theta + \lambda(\psi(\theta) - \zeta(\theta))\} \times$$

$$e^{(t-\tau)\{\theta - \lambda(1-\psi(\theta))\}} d\tau, \qquad (4.5)$$

where for the moment $W(0,t)$ is unknown. This may be inverted in a similar manner to that used by Gani and Prabhu (1959a). We write

$$K(x,t) = \sum_{n=0}^{\infty} e^{-\lambda t}(\lambda t)^n B^{n}*(x)(n!)^{-1} \qquad x \geqq 0$$

where $B^n*(x)$ is the n-fold convolution of $B(x)$ $(B^0*(x) = 1, x \geqq 0; = 0, x < 0)$. When the derivative $k(x,t) = \frac{d}{dx} K(x,t)$ exists we obtain

$$W(x,t) = K(x+t-u,t) - \int_{\tau=0}^{t} W(0,t-\tau)\{k(x+\tau,\tau)d\tau$$
$$+\lambda \int_{y=-\tau}^{x} \{dB(x-y)-dD(x-y)\}k(x+\tau,\tau)d\tau\} \qquad x \geqq u-t. \tag{4.6}$$

To complete the inversion of (4.5) it is necessary to show that the right hand side of (4.6) vanishes for $x < \max(0,u-t)$; this I have been unable to do, although from physical considerations this should be so.

Let $\Omega^*(\theta,s) = \int_0^\infty e^{-st}\Omega(\theta,t)dt$ (Rl $s > 0$) and $W^*(s) = \int_0^\infty e^{-st} W(0,t)dt$ (Rl $s > 0$) be the transforms of $\Omega(\theta,t)$ and $W(0,t)$ respectively with respect to time. Then from (4.4) by taking transforms we obtain

$$\Omega^*(\theta,s) = \frac{\Omega(\theta,0)-W^*(s)\{\theta+\lambda\psi(\theta)-\lambda\zeta(\theta)\}}{s-\theta+\lambda-\lambda\psi(\theta)} . \tag{4.7}$$

To find the unknown function $W^*(s)$ we argue as Beneš (1957) has done for the M/G/1 system. There is a unique solution $\eta(s)$ of

$$\eta(s) = s+\lambda-\lambda\psi(\eta(s)) \tag{4.8}$$

in Rl $(\theta) > 0$ for Rl $(s) > 0$. In this region $\Omega^*(\theta,s)$ must converge so that the zeros of the denominator of (4.7) must coincide with zeros of the numerator of (4.7). Thus

$$W^*(s) = \frac{\Omega(\eta\rho)}{\eta+\lambda\psi(\eta)-\lambda\zeta(\eta)} = \frac{e^{-\eta u}}{\eta+\lambda\psi(\eta)-\lambda\zeta(\eta)}, \tag{4.9}$$

where u is the waiting time at time zero. When $\psi(\theta) = \zeta(\theta)$ this has been inverted, and the probability of an empty queue at time t is given by (1.3.5).

3.5 The stationary waiting time

We now return to the stationary properties of our process and consider the stationary waiting time distribution; we assume throughout the remainder of this chapter that the stationarity condition $E(u) < 0$ and $0 < d < \infty$ holds. We prove the following

Theorem 3.2. If $A(x) = 1-e^{-\lambda x}(0 \leqq x < \infty)$, then the LST $\Omega(\theta)$ of the stationary waiting time distribution is given by

$$\Omega(\theta) = \frac{W(0)\{\theta+\lambda\psi(\theta)-\lambda\zeta(\theta)\}}{\theta-\lambda+\lambda\psi(\theta)} \tag{5.1}$$

where the probability $W(0)$ that a customer arrives to find the server idle is

$$W(0) = (1-\lambda b)(1-\lambda b+\lambda d)^{-1}. \tag{5.2}$$

Proof (a). We shall prove this theorem by two slightly different methods; the first is similar to Theorem 2 of Finch (1959). For this it is simpler to deal with characteristic functions (c.f.'s) rather than LST's. We put $s = -i\theta$, and thus wish to show

$$\Omega(s) = \frac{W(0)\{is-\lambda\psi(s)+\lambda\zeta(s)\}}{is+\lambda-\lambda\psi(s)} \tag{5.3}$$

where $W(0)$ is given by (5.2).

We require to find a solution to (3.1). Following Lindley (1952) and Finch (1959) we define a function $W^*(x)$ by

$$W^*(x) = \int_{-\infty}^{x} W(x-y)dU(y) \qquad -\infty < x < \infty. \tag{5.4}$$

When $x < 0$ we have

$$W^*(x) = \int_{y=-\infty}^{x} W(x-y) \int_{z=0}^{\infty} dB(y+z)\lambda e^{-\lambda z}dz$$

$$= \int_{y=0}^{\infty}\int_{z=0}^{\infty} W(y)\lambda e^{-\lambda(z+y-x)}dydB(z)$$

$$= C\ e^{\lambda x}$$

where

$$C = W^*(0) = \int_{-\infty}^{0} W(-y)dU(y).$$

For $x \geqq 0$ we have from (3.1) and (5.4) that

$$W^*(x) = W(x) + W(0)\{U(x) - C(x)\}, \qquad x \geqq 0,$$

so that

$$W^*(0) = W(0)\{1+U(0) - C(0)\} = W(0)\{1+\psi(i\lambda) - \zeta(i\lambda)\}$$

From the definitions (2.2) of $U(x)$ and $C(x)$ we have

$$\int_{-\infty}^{\infty} e^{isx}dU(x) = \frac{\lambda\psi(s)}{\lambda+is},\ \int_{-\infty}^{\infty} e^{isx}dC(x) = \frac{\lambda\zeta(s)}{\lambda+is}.$$

Using these and taking Fourier transforms in (3.1) and (5.4), and using (5.5), we obtain

$$\int_{-\infty}^{\infty} e^{isx}dW^*(x) = \int_{x=-\infty}^{\infty} e^{isx}\int_{y=-\infty}^{x} dW(x-y)\ dU(y)$$

$$= \lambda(\lambda+is)^{-1}\psi(s)\Omega(s)$$

$$= \int_{-\infty}^{0} e^{isx} dW^*(x) + \int_{0-}^{\infty} e^{isx} dW(x) - W(0)$$

$$+ W(0)\int_{0-}^{\infty} e^{isx}\, dU(x) - W(0)\int_{0-}^{\infty} e^{isx} dC(x)$$

$$= \lambda(\lambda+is)^{-1}W^*(0) + \Omega(s) - W(0)$$

$$+ \lambda(\lambda+is)^{-1}W(0)\{\psi(s)-\zeta(s)-\psi(i\lambda)+\zeta(i\lambda)\};$$

equation (5.3) follows from these relations and (5.2) is obtained from (5.3) by the limiting process $s \to 0$.

Proof (b). We can prove (5.1) by several different ways, each of which is related to the method due to Takács (1955). As $t \to \infty$ $\frac{\partial}{\partial t} W(x,t) \to 0$ and from (4.2) we have

$$\frac{dW(x)}{dx} = \lambda W(x) - \lambda\int_{0}^{x} W(x-y)dB(y) + \lambda W(0)\{B(x)-D(x)\} \qquad x \geqq 0 \qquad (5.6)$$

Taking LST's in (5.6) readily yields (5.1) and hence (5.2). Allowing t to tend to infinity in (4.5) also gives the result, as does the $\lim_{s\to 0} s\Omega^*(\theta,s)$ by using an extension of Abel's theorem [Widder (1946) Chapter 5].

The moments of the stationary waiting time distribution may be found from (5.1) by differentiation; the mean is

$$E(w) = -\Omega'(0) = \frac{\lambda E(r^2) - \lambda^2 bE(r^2) + \lambda^2 dE(s^2)}{2(1-\lambda b)(1-\lambda b+\lambda d)}\,. \qquad (5.7)$$

The IST $d(\theta)$ of the stationary distribution of the delay, i.e. the total time a customer spends waiting or being served, is the convolution of the distributions of waiting time and service time, so that

$$d(\theta) = \Omega(\theta)\{W(0)\zeta(\theta)+(1-W(0)\psi(\theta)\}. \tag{5.8}$$

The mean delay $\bar{d}$ is

$$\bar{d} = E(w) + W(0)d + (1-W(0)b.$$

$$= \frac{(1-\lambda b)\{\lambda E(r^2)+2d\}+\lambda^2 dE(s^2)}{2(1-\lambda b)(1-\lambda b+\lambda d)}. \tag{5.9}$$

We may observe that proof (b) is simpler than proof (a) as it is not necessary to introduce an artificial function such as $W^*(x)$. However, each method is useful as we can use each procedure to find some further results; we have seen in §4 that the second method is used for the time-dependent case, while the first may be employed for an Erlangian inter-arrival time distribution. In this case we have

Theorem 3.3. When $A(x) = 1 - \sum_{r=0}^{k-1} \lambda^r x^r e^{-\lambda x}(r!)^{-1}$ $(k = 1,2,\ldots;$ $0 \leqq x < \infty)$ the c.f. $\Omega(s)=\int_{0-}^{\infty} e^{isx}dW(x)$ of the stationary waiting time distribution is given by

$$\Omega(s) = W(0)\left[\sum_{n=0}^{k-2}\left\{isL_n\left(\prod_{m\neq n}(is-is_m)\right) - \left((\lambda+is)^{k-1}-(\lambda+is)^n\right)\beta_n\right\}\right.$$
$$\left. + is(\lambda+is)^{k-1}-\lambda^k\psi(s)+\lambda^k\zeta(s)\right]\left[(\lambda+is)^k-\lambda^k\psi(s)\right]^{-1}, \tag{5.10}$$

where

$$W(0) = \lambda^{k-1}(k-\lambda b)\left[\sum_{n=o}^{k-2} L_n \prod_{m\neq n}(-is_m)+\lambda^{k-1}-(k-1)\gamma\lambda^{k-2}-\lambda^k(b-d)\right.$$
$$\left. +\sum_{n=1}^{k-1} n\beta_n\lambda^{n-1}\right]^{-1}, \qquad (5.11)$$

and

$$\chi_n = -\lambda^k(\psi(s_n)-\zeta(s_n))-\sum_{m=o}^{k-2}\{(\lambda+is_n)^{k-1}-(\lambda+is_n)^m\}\beta_m+is_n(\lambda+is_n)^{k-1}$$

$$\beta_n = \frac{(-1)^n\lambda^k}{n!}\left(\frac{d}{d\lambda}\right)^n\{\psi(i\lambda)-\zeta(i\lambda)\}$$

$$\gamma = \frac{(-1)^{k-1}\lambda^{k+1}}{k-1!}\left(\frac{d}{d\lambda}\right)^{k-1}\left\{\frac{\psi(i\lambda)-\zeta(i\lambda)}{\lambda}\right\} = \sum_{n=o}^{k-1}\beta_n$$

$$L_n = (is)^{-1}\chi_n \prod_{m\neq n}(is_n-is_m)^{-1},$$

s_n $(0 \leqq n \leqq k-1)$ being the k roots of $(\lambda+is)^k-\lambda^k\psi(s) = 0$ in the upper half plane $I(s) \geqq 0$. For this theorem we also suppose $\psi(s)$ and $\zeta(s)$ are in the form of c.f.'s.

Proof. As in Theorem 3.2 we write

$$W^*(x) = \int_{-\infty}^{x} W(x-y)dU(y) \qquad -\infty < x < \infty. \qquad (5.12)$$

From (5.12) and the definition of $U(x)$ we have for $x < 0$ that

$$W^*(x) = \frac{\lambda^k e^{\lambda x}}{k-1!}\int_{y=o}^{\infty} W(y)e^{-\lambda y}\int_{z=o}^{\infty} e^{-\lambda z}(z-x+y)^{k-1}\,dB(z)$$

$$= e^{\lambda x}\sum_{r=o}^{k-1} c_r(-x)^r,$$

where

$$C_r = \frac{k}{k-1!} \binom{k-1}{r} \int_{y=0}^{\infty} \int_{z=0}^{\infty} W(y) e^{-\lambda y - \lambda z} (z+y)^{k-1-r} dB(z)$$

In particular

$$C_0 = W^*(0) = \int_{-\infty}^{0} W(-y) dU(y).$$

For $x \geqq 0$

$$W^*(x) = W(x) + W(0)\{U(x) - C(x)\}, \tag{5.13}$$

so that

$$W^*(0) = W(0)\{1+U(0) - C(0)\} = W(0)\{1+\lambda^{-1}\gamma\}, \tag{5.14}$$

where

$$U(0) = \int_0^{\infty} \frac{\lambda^k}{k-1!} e^{-\lambda y} y^{k-1} B(y) dy$$

$$= \frac{\lambda^k (-1)^{k-1}}{k-1!} \left(\frac{d}{d\lambda}\right)^{k-1} \left\{\frac{\psi(i\lambda)}{\lambda}\right\}$$

$$C(0) = \frac{\lambda^k (-1)^{k-1}}{k-1!} \left(\frac{d}{d\lambda}\right)^{k-1} \left\{\frac{\zeta(i\lambda)}{\lambda}\right\},$$

$$\gamma = \lambda\{U(0) - C(0)\}.$$

We also have

$$\int_0^{\infty} e^{isx}\{dU(x) - dC(x)\} = \left(\frac{\lambda}{\lambda+is}\right)^k \{\psi(s) - \zeta(s)\} - \sum_{r=0}^{k-1} (\lambda+is)^{r-k} \beta_r,$$

where

$$\beta_r = \frac{(-1)^r \lambda^r}{r!} \left(\frac{d}{d\lambda}\right)^r \{\psi(i\lambda) - \zeta(i\lambda)\}, \quad \sum_{r=0}^{k-1} \beta_r = \gamma.$$

Taking Fourier transforms in (5.12) yields

$$\int_{-\infty}^{\infty} e^{isx}\, dW^*(x) = \int_{-\infty}^{\infty} e^{isx} \int_{-\infty}^{x} dW(x-y)dU(y)$$

$$= \left(\frac{\lambda}{\lambda+is}\right)^{k} \psi(s)\Omega(s)$$

$$= \int_{-\infty}^{0} e^{isx} dW^*(x) + \int_{0-}^{\infty} e^{isx} dW(x) - W(0)$$

$$+ W(0) \int_{0}^{\infty} e^{isx}\{dU(x) - dC(x)\}$$

$$= \frac{\lambda C_o}{\lambda+is} - is \sum_{r=1}^{k-1} \frac{r!C_r}{(\lambda+is)^{r+1}} + \Omega(s) - W(0)$$

$$+ W(0) \left[\frac{\lambda^k}{(\lambda+is)^k}\{\psi(s)-\zeta(s)\} - \sum_{r=o}^{k-1} (\lambda+is)^{r-k}\beta_r\right].$$

From these relations we obtain

$$\Omega(s) = [is \sum_{r=1}^{k-1} r!\; C_r(\lambda+is)^{k-r} + W(0)\{is(\lambda+is)^{k-1} - \lambda^k\psi(s) + \lambda^k\zeta(s)$$
$$- \sum_{r=o}^{k-1} \{(\lambda+is)^{k-1} - (\lambda+is)^r\}\beta_r\}][(\lambda+is)^k - \lambda^k\psi(s)]^{-1}. \tag{5.15}$$

To eliminate the constants $C_r (r = 0,1,\ldots,k-1)$ we make use of Lemma 1 of Ewens and Finch (1962) which states that if there exists a $\delta > 0$ such that the c.f. $\psi(s)$ is analytic for $I(s) > -\delta$ and if $\lambda b < k$ then the equation $(\lambda+is)^k = \lambda^k\psi(s)$ has exactly k-1 roots in the upper half plane $I(s) > 0$; it also has a root $s = 0$ as $\psi(0) = 1$.

By writing $B_r = (k-1-r)!\ C_{k-1-r}$ $(r = 0,1,\ldots,k-1)$ equation (5.15) becomes

$$\Omega(s) = [is\sum_{r=o}^{k-2} B_r(\lambda+is)^r+W(0)\{is(\lambda+is)^{k-1}-\lambda^k\psi(s)+\lambda^k\zeta(s)$$

$$-\sum_{r=o}^{k-1}\{(\lambda+is)^{k-1}-(\lambda+is)^r\}\beta_r\}][(\lambda+is)^k-\lambda^k\psi(s)]^{-1}. \tag{5.16}$$

Let the k-1 roots of $(\lambda+is)^k-\lambda^k\psi(\theta) = 0$ in $I(s) > 0$ be $s_n(n = 0,1,\ldots,k-2)$ and put $z_j = \lambda+is_j$. Since $|\Omega(s)| \leq 1$ for $I(s) > 0$ the s_n must also be roots of the numerator of (5.16). Then

$$(z_j-\lambda)\sum_{r=o}^{k-2} B_r z_j^r = W(0)\{(z_j-\lambda)z_j^{k-1}-\lambda^k(\psi_j-\zeta_j)-\sum_{r=o}^{k-2}(z_j^{k-1}-z_j^r)\beta_r\}$$

$$= W(0)\chi_j,$$

where $\psi_j = \psi(s_j)$, $\zeta_j = \zeta(s_j)$. For any z we have

$$\sum_{r=o}^{k-2}\beta_r z^r = \sum_{j=o}^{k-2} H_j \prod_{m\neq j}(z-z_m)$$

where

$$H_j = \sum_{m=o}^{k-2} B_m z_j^m \prod_{n\neq j}(z_j-z_n)^{-1}$$

$$= \frac{W(0)\chi_j}{(z_j-\lambda)}\prod_{n\neq j}(z_j-z_n)^{-1}$$

Thus

$$\sum_{r=o}^{k-2} B_r z^r = W(0)\sum_{m=j}^{k-2}\left\{\prod_{m\neq j}(z-z_m)\right\}(z_j-\lambda)^{-1}\chi_j\left\{\prod_{n\neq j}(z_j-z_n)^{-1}\right\}. \tag{5.17}$$

As $\Omega(0) = 1$ we also have

$$\lambda^{k-1}(k-\lambda b) = W(0)\{\lambda^{k-1}-\lambda^k(b-d)-(k-1)\gamma\lambda^{k-2}+\sum_{r=1}^{k-1} r\beta_r\lambda^{r-1}\sum_{r=o}^{k-2} B_r\lambda^r\}. \tag{5.18}$$

By substituting (i) $z = \lambda+is$ and (ii) $z = \lambda$ in (5.17) we readily obtain (5.10) and (5.11), which completes the proof.

3.6 The stationary queue size

Let $Q_m(n)$ $(m = 0,1,\ldots)$ be the probability that the n-th arrival finds m customers in the queue, and let $R_m(n)$ $(m = 0,1,2,\ldots)$ be the probability that when he departs there are m customers in the queue. We have

$$Q_o(n) = \Pr\{w_{n-1}+s_{n-1}\leqq t_n, w_{n-1}>0\}+\Pr\{r_{n-1}\leqq t_n, w_n = 0\}$$

$$\sum_{m=j}^{\infty} Q_m(n) = \Pr\{w_{n-j}+s_{n-j}>t_{n-j+1}+t_{n-j+2}+\ldots+t_n, w_{n-j}>0\}$$

$$+\Pr\{r_{n-j}>t_{n-j+1}+t_{n-j+2}+\ldots+t_n, w_{n-j} = 0\} \qquad j \geqq 1, \tag{6.1}$$

and

$$R_o(n) = \Pr\{w_n+s_n\leqq t_{n+1}, w_n>0\}+\Pr\{r_n\leqq t_{n+1}, w_n = 0\}$$

$$\sum_{m=j}^{\infty} R_m(n) = \Pr\{w_n+s_n>t_{n+1}+t_{n+2}+\ldots+t_{n+j}, w_n>0\}$$

$$+ \Pr\{r_n>t_{n+1}+\ldots+t_{n+j}, w_n = 0\}. \qquad j \geqq 1 \tag{6.2}$$

If $w_{n-j}+s_{n-j}>t_{n-j+1}+\ldots+t_n$ and $w_{n-j}>0$ or $r_{n-j}>t_{n-j+1}+\ldots+t_n$ and $w_{n-j}=0$, then the (n-j)th customer has not departed when the n-th customer arrives and the number of customers present on the n-th arrival is not less than j; conversely, if this is true then either $w_{n-j}>0$ and $w_{n-j}+s_{n-j}>t_{n-j+1}+\ldots+t_n$ or $w_{n-j}=0$ and $r_{n-j}>t_{n-j+1}+\ldots$ $\ldots+t_n$. From the existence of a limiting distribution for w_n we have the existence of a limiting distribution $Q_m = \lim_{n\to\infty} Q_m(n)$, and similarly $R_m = \lim_{m\to\infty} R_m(n)$ exists. From (6.1) and (6.2) we see that

$$R_o(n) = Q_o(n+1)$$

$$\sum_{m=j}^{\infty} R_m(n) = \sum_{m=j}^{\infty} Q_m(n+j),$$

so that $Q_m = R_m \quad (m = 0,1,2,\ldots)$.

Let us now suppose that the inter-arrival time distribution is negative exponential so that $A(x) = 1-e^{-\lambda x}$; write $k_n = \int_o^\infty e^{-\lambda x}\frac{(\lambda x)^n}{n!}dB(x)$, $k_n{}^* = \int_o^\infty e^{-\lambda x}\frac{(\lambda x)^n}{n!}dD(x)$, $k(z) = \sum_{n=o}^{\infty} k_n z^n = \psi(\lambda-\lambda z)$, $k^*(z) = \sum_{n=o}^{\infty} k_n{}^* z^n = \zeta(\lambda-\lambda z) \quad (|z| \leqq 1)$.

It is readily seen that the limiting probabilities R_n satisfy the difference equations

$$R_n = R_{n+1}k_o+R_nk_1+\ldots+R_1k_n+R_ok_n{}^* \qquad n = 0,1,2,\ldots \tag{6.3}$$

We multiply (6.3) by z^n, sum and obtain the p.g.f. $r(z) = \sum_{n=o}^{\infty} R_n z^n$ $(|z| \leqq 1)$ of the stationary queue size distribution as

$$r(z) = \frac{(1-\lambda b)\{\psi(\lambda-\lambda z)-z\zeta(\lambda-\lambda z)\}}{(1-\lambda b+\lambda d)\{\psi(\lambda-\lambda z)-z\}} \quad . \tag{6.4}$$

The mean number of customers in the stationary queue may be obtained from (6.4) by differentiation as

$$r'(1) = \frac{\lambda(1-\lambda b)\{\lambda E(r^2)+2d\}+\lambda^3 dE(s^2)}{2(1-\lambda b)(1-\lambda b+\lambda d)} \tag{6.5}$$

$$= \lambda\bar{d},$$

where $\bar{d}$ is the mean delay given by (5.9).

3.7 The busy period

A busy period is the time from the arrival of a customer at an empty queue until the next point in time that the server again becomes free. The queueing process consists of alternate busy periods and idle periods when the server is unoccupied. We are interested in the distribution of the length of busy periods and of the number of customers served in these periods. We consider in this section only homogeneous Poisson arrivals.

If there is a waiting time of $x > 0$ in the queue at time zero then the d.f. $F(x,t)$ to the time that the queue empties for the first time has the LST $\Phi(x,\theta)$ found by Kendall (1957) as

$$\Phi(x,\theta) = e^{-x\eta(\theta)}$$

$$= \int_0^\infty e^{-\theta x} d_x F(x,t) \tag{7.1}$$

where $\eta(\theta)$ is the unique solution with $\eta(0) = 0$ of the functional equation [c.f. (4.8)]

$$\eta(\theta) = \lambda+\theta-\lambda\psi(\eta(\theta)). \tag{7.2}$$

As a customer joining an empty queue has a service time with d.f. $D(x)$ we find that the LST $\gamma(\theta)$ of the d.f. $G(x)$ of the length of a busy period is

$$\gamma(\theta) = \int_{x=0}^{\infty} \Phi(x,\theta)dD(x) = \zeta(\eta(\theta)). \tag{7.3}$$

This result (7.3) may also be obtained by a generalisation of the method of Takács (1955); by a similar generalisation we can show that the p.g.f. $F(z) = \sum_{i=0}^{\infty} f_i z^i (|z| \leq 1)$ of the probability f_j that j customers are served in a busy period is

$$F(z) = z\zeta(\xi(z)) \tag{7.4}$$

where $\xi(z)$ is the unique solution with $\xi(1) = 0$ of the functional equation

$$\xi(z) = \lambda-\lambda z\psi(\xi(z)). \tag{7.5}$$

When $\zeta(\theta) = \psi(\theta)$ the results (7.3) and (7.4) reduce to (47) and (67) of Takács (1955). The moments of $\gamma(\theta)$ and $F(z)$ may be obtained by differentiation, e.g.

$$-\gamma'(0) = d(1-\lambda b)^{-1}$$

$$\gamma''(0) = \frac{E(r^2)-\lambda bE(r^2)+\lambda dE(s^2)}{(1-\lambda b)^3}. \tag{7.6}$$

Let us now turn to the joint distribution of the length of a busy period and of the number of customers served in this period. In order to find this we first consider the related problem of zero avoiding transition probabilities [c.f. Gaver (1959)]. Let $T_o = 0$, T_n $(n = 1,2,\ldots)$ be the sequence of departure times of customers from the system, and $R(t)$ the number of customers in the system at time t, including the one being served; $R(T_n+0) \equiv R(T_n)$ is the number of customers left in the system immediately after the departure of the n-th customer. We define

$$P_j^{(n)}(t) = \Pr\{R(T_n) = j;\ T_n \leq t;\ R(\tau)>0,\ 0\leq\tau<T_n \mid R(0) = 1\}$$

as the probability that the number of customers in the system passes from one at time zero to $j > 0$ after the n-th departure, which occurs prior to time t, without having been through the state zero. As initial conditions we have

$$P_j^{(n)}(0-) = 0(n>0),\ P_j^{(0)}(t) = \delta_{1j}\, H^*(t), \tag{7.7}$$

where $H^*(x)$ is the unit step function.

We have that $R(T_{n+1}) = j > 0$ if $R(T_n) = j-i(> 0)$ and exactly $i+1$ customers arrive in (T_n,T_{n+1}), so that by enumeration

$$P_j^{(1)}(t) = k_j^*(t)$$

$$P_j^{(n+1)}(t) = \sum_{i=o}^{j} \int_{\tau=o}^{t} P_{j+1-i}^{(n)}(t-\tau)dk_i(\tau) \qquad n = 1,2,\ldots, \tag{7.8}$$

where

$$dk_n(t) = e^{-\lambda t}\,\frac{(\lambda t)^n}{n!}\,dB(t),\ dk_n^*(t) = e^{-\lambda t}\,\frac{(\lambda t)^n}{n!}\,dD(t).$$

We define the transforms

$$\alpha_j(\theta,z) = \sum_{n=0}^{\infty} \int_{t=0}^{\infty} z^n e^{-\theta t} dP_j^{(n)}(t) \qquad \mathrm{Rl}(\theta) \geqq 0, \quad |z| < 1,$$

$$p(\theta,z,x) = \sum_{j=1}^{\infty} \alpha_j(\theta,z)x^j \qquad |x| < 1$$

$$\Xi_n(\theta) = \int_0^{\infty} e^{-\theta t} dk_n(t), \quad \Xi_n^*(\theta) = \int_0^{\infty} e^{-\theta t}\, dk_n^*(t) \qquad \mathrm{Rl}\ (\theta) \geqq 0$$

$$\kappa(\theta,x) = \sum_{n=0}^{\infty} \Xi_n(\theta)x^n = \psi(\lambda+\theta-\lambda x) \qquad |x| < 1$$

$$\kappa^*(\theta,x) = \sum_{n=0}^{\infty} \Xi_n^*(\theta)x^n = \zeta(\lambda+\theta-\lambda x) \qquad |x| < 1.$$

Taking transforms in (7.8) we find

$$\alpha_j(\theta,z)-\delta_{1j} = z\sum_{i=0}^{j} \alpha_{j-i+1}(\theta,z)\, \Xi_i(\theta) + z\Xi_j^*(\theta) - z\Xi_j(\theta), \tag{7.9}$$

from which we obtain

$$p(\theta,z,x) = \frac{x+z\{\kappa^*(\theta,x)-\kappa(\theta,x)\}-z\alpha_1(\theta,z)\,\Xi_0(\theta)}{1-z\kappa(\theta,x)/x}. \tag{7.10}$$

We require to determine $\alpha_1(\theta,z)$; it is known [Gaver (1959)] that for $\mathrm{Rl}\ (\theta) > 0$ and $0 < z \leqq 1$ that there exists a unique root of the equation

$$x = z\psi(\lambda+\theta-\lambda x). \tag{7.11}$$

Now $p(\theta,z,x)$ must be bounded for all x in $0 \leqq x < 1$; thus the zero of the denominator must coincide with a zero of the numerator, so that

$$z\alpha_1(\theta,z)\,\Xi_0(\theta) = x+z\kappa^*(\theta,x)-z\kappa(\theta,x).$$

We now return to the problem of the busy period; if τ_m is the length of a busy period in which exactly m customers are served and

$$F_m(t) = \Pr\{\tau_m \leq t\}$$

then

$$F_m(t) = \int_{\tau=0}^{t} P_1^{(m-1)}(t-\tau)\,dk_0(\tau) \qquad m = 1,2,\ldots \tag{7.13}$$

Taking transforms in (7.13) yields

$$\begin{aligned} \gamma(\theta,z) &= \sum_{m=1}^{\infty} z^m \int_{t=0}^{\infty} e^{-\theta t}\, dF_m(t) \\ &= z\alpha_1(\theta,z)\,\Xi_0(\theta) \\ &= x + z\kappa^*(\theta,z) - z\kappa(\theta,z) \\ &= z\zeta(\lambda+\theta-\lambda x), \end{aligned} \tag{7.14}$$

where x is defined by (7.11). When $\theta = 0$ (7.14) reduces to (7.4) and when $z = 1$ it reduces to (7.3).

By extending the direct argument of Prabhu (1960) we can show that if $G_n(t)$ is the d.f. of the length of a busy period in which n customers are served then

$$dG_n(t) = e^{-\lambda t}\, \frac{\lambda^{n-1} t^{n-2}}{n-1!} \int_{x=0}^{t} x\,dB_{n-1}(t-x)\,dD(x) \tag{7.15}$$

When a busy period commences with n customers present the transform $\gamma_n(\theta,z)$ of the busy period distribution, counting the number of customers served, is not equal to $(\gamma(\theta,z))^n$, as only the first customer to be served has the service time d.f. $D(x)$, which is different from the other n-1 initial customers.

3.8 Cases of server absenteeism

We now consider another generalisation of the M/G/1 queueing system which for stationary distributions may be classed as a special case of our previous problem. Whenever a departing customer leaves the queue empty the server departs from the counter for a time which is a random variable with d.f. $G_1(x)$ $(0 \leqq x < \infty)$, mean $g_1 < \infty$ and LST $\xi_1(\theta)$. When the server returns to find at least one customer waiting he commences serving immediately; when he returns to find the counter still free he departs again for a time which is a random variable with d.f. $G_2(x)$, mean $g_2 < \infty$ and LST $\xi_2(\theta)$, and he continues to come and go until he finds at least one customer waiting. The server's n-th successive absentee period is a random variable with d.f. $G_n(x)$, mean $g_n < \infty$ and LST $\xi_n(\theta)$. From the time the server is first occupied customers have a service time d.f. $B(x)$ with mean $b < \infty$ and LST $\psi(\theta)$. If we were to consider service beginning only when the server actually commences to serve a customer then we would make use of the result of Finch (1959) [see Chapter 4]. However, we suppose a customer commences 'service' immediately he reaches the head of the queue. This is similar to the queueing problem used in Part 3 for road traffic theory.

We require to find the LST $\chi(\theta)$ of the d.f. $L(x)$ of the time from the arrival of a customer at an empty queue to the return of the server. When the process is in statistical equilibrium we can integrate over the possible times of arrival of a customer after the departure of the last customer of the previous busy period. Considering only the first server absentee period we obtain the LST of the (improper) distribution of the time to the return of the server as

$$\chi_1(\theta) = \int_{x=0}^{\infty}\int_{y=0}^{\infty} \lambda e^{-\lambda y} e^{-\theta(x-y)} dG_1(x)dy$$

$$= \frac{\lambda}{\lambda-\theta}\{\xi_1(\theta)-\xi_1(\lambda)\} \tag{8.1}$$

which has $\chi_1(0) = 1-\xi_1(\lambda) = 1-g_1^*$, g_1^* being the probability that no customers arrive while the server is absent. For following server absentee periods the problem is similar and we obtain

$$\chi(\theta) = \sum_{n=1}^{\infty}\left(\prod_{i=1}^{n-1} g_i^*\right)\chi_n(\theta) \tag{8.2}$$

$$\chi_n(\theta) = \frac{\lambda}{\lambda-\theta}\{\xi_n(\theta)-\xi_n(\lambda)\},$$

where $g_i^* = \xi_i(\lambda)$ is the probability that no customers arrive during the server's i-th successive absentee period. As special cases we suppose (i) $G_n(x) = G(x)$ $(n = 1,2,\ldots)$ and (ii) $G_1(x) = G(x)$, $G_n(x) = 1$ $(x \geq 0,\ n = 2,3\ldots)$. In the first case

$$\chi(\theta) = \sum_{n=1}^{\infty}(g_i^*)^{n-1}\chi_1(\theta) = \frac{\lambda\{\xi_1(\theta)-\xi_1(\lambda)\}}{(\lambda-\theta)(1-\xi_1(\lambda))}, \tag{8.3}$$

while in the second

$$\chi(\theta) = \chi_1(\theta)+g_1^* = \frac{\lambda\xi_1(\theta)-\theta\xi_1(\lambda)}{(\lambda-\theta)} \tag{8.4}$$

Once the server actually begins serving the time to completion has the same distribution for all customers, although this may be generalised as is done for some traffic models in Part 3; thus

$$\zeta(\theta) = \psi(\theta)\chi(\theta) \tag{8.5}$$

may be taken as the service time LST for customers arriving at an empty queue. The stationary waiting time and queue size distributions may now be obtained from the results of the previous sections.

3.9 Comparison with Finch (1959)

As a special case of our process we put $D(x) = \int_0^x B(x-y)dV(y)$ where $V(x)$ is the d.f. of a non-negative random variable v. The distribution of R_n (Q_n) is identical to that in Finch (1959), but the waiting time distribution differs, as Finch considers v as part of the waiting time while we have taken it as part of a service time. However, the delay caused, i.e. waiting time plus service time, is the same in both cases.

The LST $\Phi(\theta) = \int_0^\infty e^{-\theta x}dW(x)$ $(\mathrm{Rl}\ \theta \geq 0)$ of the stationary waiting time distribution, which exists for $\lambda b < 1$ and $\bar{v} < \infty$, is given by

$$\Phi(\theta) = \frac{W(0)\{(\theta-\lambda)\vartheta(\theta)+\lambda\}}{\theta-\lambda+\lambda\alpha(\theta)} \tag{9.1}$$

where the probability $W(0)$ that the server is idle is

$$W(0) = (1-\lambda b)(1+\lambda\bar{v})^{-1}, \tag{9.2}$$

and $\vartheta(\theta) = \int_0^\infty e^{-\theta x} dV(x)$ with $\bar{v} = -\vartheta'(0)$. The stationary queue size and busy period distributions are given by (6.4) and (7.3) with $\zeta(\theta) = \psi(\theta)\vartheta(\theta)$.

As an example suppose $A(x) = 1-e^{-\lambda x}$, $B(x) = 1-e^{-\mu x}$, $V(x) = 1-e^{-\nu(x)}$ so that

$$D(x) = 1-(\mu-\nu)^{-1} \ (\mu e^{-\mu x}-\nu e^{-\nu x}) \qquad x \geqq 0.$$

We obtain by inverting (5.1) that

$$W(x) = 1 - \frac{\lambda\nu}{\mu(\lambda+\nu-\mu)} e^{-(\mu-\lambda)x} + \frac{\lambda(\mu-\lambda)}{(\nu+\lambda)(\nu+\lambda-\mu)} e^{-\nu x} \qquad x \geqq 0,$$

with $W(0) = \nu\mu^{-1}(\mu-\lambda)(\nu+\lambda)^{-1}$, while Finch obtains

$$W(x) = 1 - \frac{\lambda\nu}{\mu(\lambda+\nu-\mu)} e^{-(\mu-\lambda)x} - \frac{(\mu-\lambda)(\nu-\mu)}{\mu(\lambda+\nu-\mu)} e^{-\nu x} \qquad x \geqq 0.$$

3.10 Batch arrivals

In the traffic model of Chapter 6 we shall be concerned with vehicles in a minor road arriving at an intersection in bunches. Batch arrivals for the M/G/1 process have been discussed by Gaver (1959) and Foster (1961), and we shall extend these results to our generalised problem.

At instants τ_n $(n = 1,2,\ldots)$ a bunch of I customers arrives at a counter with a single server; I is a random variable with probabilities $\Pr\{I = i\} = b_i$ $(i = 1,2,\ldots)$, which has p.g.f. $b^*(z) = \sum_{i=1}^\infty b_i z^i$ $(|z| \leqq 1)$ and mean $g = \sum_{i=1}^\infty ib_i < \infty$. The times

$t_n = \tau_n - \tau_{n-1}$ $(n \geqq 1)$, $\tau_o = 0$ between the arrival of bunches of customers are identically distributed random variables with common d.f.'s $A(x) = 1-e^{-\lambda x}$. If a bunch of customers arrives at a counter when the server is busy then the service times of the customers are identically distributed random variables with d.f.'s $B(x)$ $(0 \leqq x < \infty)$ (as in §3.2). If the bunch of customers arrives to find the server idle then the first customer of this bunch (the customers may be ordered for service) to be served has a service time which is a random variable with d.f. $D(x)$ $(0 \leqq x < \infty)$, while the other customers of the bunch have service times which are identically distributed random variables with the d.f. $B(x)$.

The arguments used for the particular case of all bunches being of size one carry through for our more general case. The stationarity condition is $\lambda bg < 1$ provided $0 < d < \infty$. The equation (4.2) becomes

$$\frac{\partial W(x,t)}{\partial t} - \frac{\partial W(x,t)}{\partial x} = -\lambda W(x,t) + \lambda \sum_{i=1}^{\infty} b_i \left[\int_{y=o}^{x} W(x-y,t)\, dB^{n*}(y) + W(0,t) \left\{ \int_{y=o}^{x} B^{(n-1)*}(x-y)\, dD(y) - B^{n*}(x) \right\} \right] . \qquad (10.1)$$

The LST $\Omega(\theta) = \int_o^\infty e^{-\theta x} dW(x)$ $(Rl\ \theta \geqq 0)$ of the stationary waiting time distribution may be found using the methods of §3.4 as

$$\Omega(\theta) = \frac{W(0)\{\theta\psi(\theta) + \lambda(\psi(\theta) - \zeta(\theta))b^*(\psi(\theta))\}}{\psi(\theta)\{\theta - \lambda + \lambda b^*(\psi(\theta))\}} \qquad (10.2)$$

$$W(0) = (1-\lambda bg)(1-\lambda b+\lambda d)^{-1}; \tag{10.3}$$

its mean $E(w)$ is

$$E(w) = \lambda[\psi''(0)(g-1+\lambda dg)+\zeta''(0)(1-\lambda bg)-2b(b-d)(g-1)(1-\lambda bg)$$

$$-b^2 b^{*''}(1)(1-\lambda b+\lambda d)][2(1-\lambda bg)(1-\lambda b+\lambda d)]^{-1}, \tag{10.4}$$

and the mean delay is

$$\bar{d} = [\lambda\psi''(0)(g-1+\lambda dg)+(1-\lambda bg)(\lambda\zeta''(0)+2d)$$

$$\frac{+b(1-\lambda b+\lambda d)\{2(1-\lambda bg)(g-1)-\lambda b^2 b^{*''}(1)]}{2(1-\lambda bg)(1-\lambda b+\lambda d)} . \tag{10.5}$$

The stationary p.g.f. $r(z)$ of the queue size distribution is given by

$$r(z) = \frac{W(0)\{\psi(\lambda(1-z))-z\zeta(\lambda(1-z))\}b^*(\psi(\lambda(1-z)))}{\psi(\lambda(1-z))\{b^*(\psi(\lambda(1-z)))-z\}} . \tag{10.6}$$

The LST $\gamma(\theta)$ of the length of a busy period is given by

$$\gamma(\theta) = \zeta(\eta(\theta))b^*(\psi(\eta(\theta)))/\psi(\eta(\theta)) \tag{10.7}$$

where $\eta(\theta)$ is the solution with $\eta(0) = 0$ of the functional equation

$$\eta(\theta) = \lambda+\theta-\lambda b^*(\psi(\eta(\theta))). \tag{10.8}$$

CHAPTER 4

PRIORITY QUEUES

4.1 Introduction

We now turn to some queueing problems where certain classes of customers have priority to service over other classes of customers. We consider a priority queueing system where there are K classes $1,2,\dots,K$ of customers arriving at a counter with a single server. Within each class there is a first-come first-served queue discipline, but between classes there is a relative priority of service such that a class i customer is always served in preference to a class j customer whenever $i < j$.

When a class j customer is being served and a class $i(< j)$ customer arrives there are four possibilities we may consider. The first three of these are preemptive, i.e. the service of the class j customer ceases immediately in favour of the class i customer, and it resumes only when the queueing system is next cleared of all customers of all classes $1,2,\dots,j-1$. When a preempted customer returns to actual service there are three different cases we may discuss. A preemptive priority resume policy allows the customer to re-enter service at the point at which it was preempted; thus the service may be carried out in sections. In a preemptive priority repeat identical policy the class j customer has to begin service again with the same service time as that from which it was preempted; with a preemptive priority repeat different policy the customer commences service again from the beginning but with a new service time which, however, has the same distribution as that from which it was preempted (some of our results may be extended to the case where the new service time requirement may have a distribution different from that of the preempted one). In these two cases the service is completed when a service time requirement is finished without interruption. The fourth case is the head-of-the-line priority policy; here the class j customer completes service without interruption, but the class i customer is placed ahead of all customers of lower classes. This case will not be discussed in detail, although some of the similarities with the preemptive priority resume policy will be discussed, since these two policies yield several very similar results.

We assume that class i ($1 \leqq i \leqq K$) customers arrive independently in a Poisson process with parameter λ_i, and independently of customers in other classes. Class i customers have a service time requirement (total time for a head-of-the-line or preemptive resume policy and uninterrupted time for a preemptive repeat policy), which is a random variable with d.f. $F_i(x)$ ($0 \leqq x \leqq \infty$), LST $\alpha_i(\theta) = \int_0^\infty e^{-\theta x} dF_i(x)$ ($\mathrm{Rl}\ \theta \geqq 0$) and finite mean $\mu_i = -\alpha_i'(0) < \infty$; as a special case we take $F_i(x) = 1$ for $x \geqq b_i$ ($\geqq 0$) and $F_i(x) = 0$ for $x < b_i$. We write $\rho_i = \lambda_i \mu_i (1 \leqq i \leqq K)$.

If customers may arrive in batches rather than singly the results of this chapter can be extended to this more general case as in §3.10.

When there are only two classes of customers, i.e. $K = 2$, then class 1 customers are called priority and class 2 non-priority.

For the head-of-the-line priority policy, results have been obtained by Cobham (1954), Holley (1954), Miller (1960), Jaiswal (1962) and others. For $K = 2$ the time-dependent joint distribution of the number of customers in the queue has been found by Jaiswal (1962) and as a consequence the stationary distribution, which is different from that obtained by Miller (1960) from an imbedded Markov chain. Thus the stationary queue length probabilities for complex queues obtained from the imbedded Markov chain method may not be the same as those obtained from the limit of the time-dependent solution. Miller (1960) has found the stationary waiting time distribution for the K-th priority class and the busy period distribution for the system.

Preemptive priority queueing systems have been discussed by Barry (1956), White and Christie (1958), Stephan (1958), Heathcote (1959, 1960), Miller (1960), Gaver (1962, 1963) and several others. For the non-exponential queue these authors are primarily concerned with a preemptive priority resume policy. The early papers consider the equilibrium behaviour of the exponential queue for $K = 2$, while Heathcote has determined the temporal queue size distribution. Miller (1960) has obtained the waiting time distribution for class j ($j = 1,2,\ldots,K$) customers and the busy period distribution for the system, while Gaver

(1962) has found the stationary queue size distribution for non-priority customers when $K = 2$ for each of the four priority policies discussed in this chapter. Before Gaver's work (1962) little was known about preemptive repeat policies; most of the work of this chapter was done at about the same time but independently of Gaver's.

We are primarily concerned with the stationary waiting time, queue size, and busy period distributions for the various classes of customers. We shall first consider the case $K = 2$ and then extend the results to $K \geqq 2$. To do this we make use of several of the formulae derived in the previous chapter.

4.2 Waiting time distributions for $K = 2$

We are here concerned with the waiting time distributions for priority and for non-priority customers. For priority customers waiting time is defined as the time from arrival in the queue to reaching the counter for the first time. For non-priority customers we distinguish two possibilities: waiting time is the length of time a non-priority customer spends from the time of arrival to the time (a) it reaches the counter for the first time and (b) it reaches the head of the non-priority queue. Once a customer has commenced service for the first time it is no longer waiting, even if it has been preempted, but is regarded as being 'in service'. A service time, which we call completion time to distinguish it from service times in isolation, is the time a customer spends from first reaching the head of the queue to its final departure from the system. This is slightly modified in one case for a head-of-the-line priority policy.

For a preemptive policy the behaviour of priority customers is independent of the non-priority customers, and the priority system is fully described by the M/G/1 process, which has been discussed in §1.3 and §2.5. The time-dependent waiting time distribution is given by (1.3.4) with $\lambda(\tau) = \lambda_1$, $F(x) = F_1(x)$. The transforms $\mathcal{H}_1(\theta)$, $R_1(z)$ and $\Lambda(\theta) = \varphi_1(\theta)$ of the stationary waiting time, stationary queue size and the busy period distributions are for values $\rho_1 = \lambda_1\mu_1 < 1$, assumed throughout this chapter,

$$\mathcal{H}_1(\theta) = (1-\rho_1)\{1-\theta^{-1}\lambda_1(1-\alpha_1(\theta))\}^{-1} \tag{2.1}$$

$$R_1(z) = \frac{(1-\rho_1)(1-z)\alpha_1(\lambda_1-\lambda_1 z)}{\alpha_1(\lambda_1-\lambda_1 z)-z} \tag{2.2}$$

$$\varphi_1(\theta) = \alpha_1(\lambda_1+\theta-\lambda_1\varphi_1(\theta)), \tag{2.3}$$

where $\lim_{\theta\to\infty} \varphi_1(\theta) = 0$. The moments of these distributions may be obtained by differentiation; for example

$$\bar{h}_1 = -\mathcal{H}_1'(0) = \frac{\lambda_1\alpha_1''(0)}{2(1-\rho_1)}$$

$$R_1'(1) = \frac{\lambda_1^2\alpha_1''(0)+2\rho_1(1-\rho_1)}{2(1-\rho_1)}$$

$$\bar{\varphi}_1 = -\varphi_1'(0) = \mu_1(1-\rho_1)^{-1}$$

$$\varphi_1''(0) = \alpha_1''(0)(1-\rho_1)^{-3}.$$

Before proceeding to the completion times for the various priority policies, we obtain the following result which is common to all cases. A non-priority customer may arrive to find the queue free of other non-priority customers, but not of priority customers, so that it is kept away from the server until the priority stream has cleared. This delay is zero if there are no priority customers in the queue, and lasts until the end of the priority busy period if there is at least one priority customer in the queue when the non-priority customer arrives. This depends on the length of time since the departure of the last non-priority customer (or end of completion time), and is similar to the problem of server absenteeism discussed in §3.8, its solution being obtained by the same method. From the point of view of non-priority customers the server appears to be alternately present, for a length of time which has a negative exponential distribution with mean λ_1^{-1}, and absent, for a length of time which has the busy period d.f. $G_1(x)$ of priority customers. We suppose that $\xi(\theta)$ of (3.8.3) is the busy period LST of the priority customers so that $\xi(\theta) = \varphi_1(\theta)$; the probability that a non-priority customer arrives while the server is 'present' is $\lambda_2(\lambda_1+\lambda_2)^{-1}$. The LST $\vartheta_2(\theta)$ of the delay for a non-priority customer to reach the counter for the first time, when it arrives at a queue free of other non-priority customers, is using (3.8.3)

$$\vartheta_2(\theta) = \left\{\frac{\lambda_2}{\nu_2} + \frac{\lambda_1}{\nu_2} \cdot \frac{\lambda_2(\varphi_1(\theta)-\varphi_1(\lambda_2))}{(\lambda_2-\theta)}\right\} \Big/ \left\{1 - \frac{\lambda_1}{\nu_2}\varphi_1(\lambda_2)\right\}$$

$$= \frac{\lambda_2\{\lambda_2-\theta+\lambda_1\varphi_1(\theta)-\lambda_1\varphi_1(\lambda_2)\}}{(\lambda_2-\theta)(\nu_2-\lambda_1\varphi_1(\lambda_2))}\ , \tag{2.4}$$

where $\nu_2 = \lambda_1+\lambda_2$. The probability is $g_1 = \lambda_1\nu_2^{-1}\varphi_1(\lambda_2)$ that a non-priority customer does not arrive during a server 'present-absent' period. The first two moments of this distribution are

$$\bar{v}_2 = -\vartheta_2'(0) = \frac{(1+\lambda_1\bar{\varphi}_1)}{\nu_2(1-g_1)} - \frac{1}{\lambda_2} \tag{2.5}$$

$$\vartheta_2''(0) = \frac{\lambda_1\varphi_1''(0)}{\nu_2(1-g_1)} - \frac{2(1+\lambda_1\bar{\varphi}_1)}{\lambda_2\nu_2(1-g_1)} + \frac{2}{\lambda_2^2} \tag{2.6}$$

We shall now obtain the completion time distributions for each of the four priority policies in turn; we begin with the preemptive repeat policies as less has previously been discovered for these than for the other two under consideration.

(i) The preemptive priority repeat identical policy

We suppose initially that a non-priority customer requires uninterrupted occupation of the server for a fixed time b (> 0) before being able to depart from the system; we later generalise this service time requirement.

Suppose a priority customer arrives at time zero at a counter free of other priority customers, but not necessarily free of non-priority customers. The LST $X_1(\theta,b) = \int_{x=0}^{\infty} e^{-\theta x}\, dL_1(x,b)$ (Rl $\theta \geq 0$) of the d.f. $L_1(x,b)$ $(0 \leq x < \infty)$ of the time until the first gap of at least b, including b, appears in the priority stream is the continuous analogue of a generalisation of the success run problem of Feller (1957, p.299); it may be obtained by renewal theory methods or as follows:

$$\chi_1(\theta,b) = \int_{z=o}^{\infty} dG_1(z)\Big\{e^{-(\lambda_1+\theta)b-\theta z} + \int_{y_1=o}^{\infty} \lambda_1 e^{-\lambda_1 y} dy_1 \int_{z_1=y_1}^{\infty} dG_1(z_1-y_1)$$

$$\Big\{e^{-(\lambda_1+\theta)b-\theta(z+z_1)} + \int_{y_2=o}^{\infty} \lambda_1 e^{-\lambda_1 y_2} dy_2 \int_{z_2=y_2}^{\infty} dG_2(z_2-y_2)$$

$$\Big\{e^{-(\lambda_1+\theta)b-\theta(z+z_1+z_2)} + \ldots \Big\}\Big\}\Big\}$$

$$= \frac{(\lambda_1+\theta)\varphi_1(\theta)e^{-(\lambda_1+\theta)b}}{\lambda_1+\theta - \lambda_1\varphi_1(\theta)+\lambda_1\varphi_1(\theta)e^{-(\lambda_1+\theta)b}} . \tag{2.7}$$

When a non-priority customer reaches the counter for the first time it has a completion time of length b if no priority customers arrive in this period; otherwise it is delayed until the server is first free of priority customers for a least time b. The LST $\psi_2(\theta,b) = \int_o^{\infty} e^{-\theta x} dB_2(x,b)$ (Rl $\theta \geqq 0$) of the d.f. B_2 (x,b) of the time a non-priority customer spends from arrival at the counter to departure from the system is

$$\psi_2(\theta,b) = e^{-(\lambda_1+\theta)b} + \int_{y=o}^{b} \lambda_1 e^{-\lambda_1} dy \int_{z=y}^{\infty} e^{-\theta z} dL_1(z-y,b)$$

$$= e^{-(\lambda_1+\theta)b} + \frac{\lambda_1\chi_1(\theta,b)\{1-e^{-(\lambda_1+\theta)b}\}}{\lambda_1+\theta}$$

$$= \frac{(\lambda_1+\theta)e^{-(\lambda_1+\theta)b}}{\lambda_1+\theta-\lambda_1\varphi_1(\theta)+\lambda_1\varphi_1(\theta)e^{-(\lambda_1+\theta)b}} \; . \tag{2.8}$$

The LST of the distribution of the time a non-priority customer actually occupies the server is given by (2.8) with $\varphi_1(\theta) = 1$, i.e. the time priority customers are being served is neglected.

For a non-priority customer joining an empty non-priority queue we have for definition (b) of the waiting time that the LST $\zeta_2(\theta,b) = \int_{x=0}^{\infty} e^{-\theta x} dD_2(x,b)$ (Rl $\theta \geqq 0$) of the d.f. $D_2(x,b)$ $(0 \leqq x < \infty)$ of the completion time is

$$\zeta_2(\theta,b) = \psi_2(\theta,b)\vartheta_2(\theta), \tag{2.9}$$

while for definition (a) of the waiting time the LST of the completion time distribution is $\psi_2(\theta,b)$ and the customer has a wait with the LST $\vartheta_2(\theta)$.

From (2.8) and (2.9) the first two moments of the completion time distributions are

$$\bar{\psi}_2(b) = -\psi_2'(0,b) = \lambda_1^{-1}(1+\lambda_1\bar{\varphi}_1)(e^{\lambda_1 b}-1)$$

$$\bar{\zeta}_2(b) = -\zeta_2'(0,b) = \bar{\psi}_2(b)+\bar{v}_2$$

$$\psi_2''(0,b) = \varphi_1''(0)(e^{\lambda_1 b}-1)+2\lambda_1^{-2}(1+\lambda_1\bar{\varphi}_1)[(e^{\lambda_1 b}-1)$$
$$\{(1+\lambda_1\bar{\varphi}_1)e^{\lambda_1 b}-\lambda_1\bar{\varphi}_1-\lambda_1 b\}-\lambda_1 b]$$

$$\zeta_2''(0,b) = \psi_2''(0,b)+2\bar{v}_2\,\bar{\psi}_2(b)+\vartheta_2''(0). \tag{2.10}$$

We now suppose that non-priority customers require an uninterrupted service with d.f. $F_2(x)$ having LST $\alpha_2(\theta) = \int_0^\infty e^{-\theta x}\, dF_2(x)$, which exists and is finite for all real $\theta \geqq -R_2$, where $R_2 > 0$; this is equivalent to the d.f. $F_2(x)$ having an analytic characteristic function. The LST's $\psi_2(\theta) = \int_0^\infty e^{-\theta x}\, dB_2(x)$ and $\zeta_2(\theta) = \int_0^\infty e^{-\theta x}\, dD_2(x)$ (Rl $\theta \geqq 0$) of the d.f. $B_2(x)$ and $D_2(x)$ $(0 \leqq x < \infty)$ respectively of the completion times for non-priority customers are given by

$$\int_{x=0}^{\infty} \psi_2(\theta,x)dF_2(x) = \int_{x=0}^{\infty}\int_{y=0}^{\infty} e^{-\theta y}\, dB_2(y,x)dF_2(x)$$

$$= \int_{y=0}^{\infty} e^{-\theta y} \int_{x=0}^{\infty} dB_2(y,x)dF_2(x)$$

$$= \int_{y=0}^{\infty} e^{-\theta y} dB_2(y)$$

$$= \psi_2(\theta)$$

$$\zeta_2(\theta) = \int_{x=0}^{\infty} \zeta_2(\theta,x)dF_2(x)$$

$$= \psi_2(\theta)\vartheta_2(\theta), \tag{2.11}$$

where the inversion of the order of integration may be justified by Fubini's theorem (see Lukács (1960)). Thus we have

$$\psi_2(\theta) = \int_{x=o}^{\infty} \frac{(\lambda_1+\theta)e^{-(\lambda_1+\theta)x}}{\lambda_1+\theta-\lambda_1\varphi_1(\theta)\left\{1-e^{-(\lambda_1+\theta)x}\right\}} dF_2(x) \tag{2.12}$$

$$\zeta_2(\theta) = \psi_2(\theta)\vartheta_2(\theta); \tag{2.13}$$

(2.12) has also been obtained as (4.18) of Gaver (1962).

Except in some special cases the integral in (2.12) cannot be evaluated explicitly; however, its moments may be found. If $R_2 \geqq \lambda_1$ we have

$$\overline{\psi}_2 = -\psi_2'(0) = \lambda_1^{-1}(1+\lambda_1\overline{\varphi}_1)\{\alpha_2\ (-\lambda_1)-1\} \tag{2.14}$$

$$\overline{\zeta}_2 = -\zeta_2'(0) = \overline{\psi}_2 + \overline{\psi}_2,$$

and for $R_2 \geqq 2\lambda_1$

$$\psi_2''(0) = \varphi_1''(0)\{\alpha_2(-\lambda_1)-1\}+2\lambda_1^{-2}(1+\lambda_1\overline{\varphi}_1)[(1+\lambda_1\overline{\varphi}_1)\{\alpha_2(-2\lambda_1)$$

$$-\alpha_2(-\lambda_1)\}-\lambda_1\overline{\varphi}_1\{\alpha_2(-\lambda_1)-1\}+\lambda_1\alpha_2'(-\lambda_1)]$$

$$\zeta_2''(0) = \psi_2''(0)+2\overline{v}_2\psi_2+\vartheta_2''(0)\ , \tag{2.15}$$

where $\alpha_2'(-\lambda_1) = \lim_{\theta\to-\lambda_1} \alpha_2'(\theta)$.

(ii) The preemptive priority repeat different policy

Suppose that the first time a non-priority customer occupies the server, it requires uninterrupted service of time b_1; if preempted before this time elapses then it requires uninterrupted service of time b_2 when it returns to actual service; after this it requires service times

b_3, b_4,... etc until there is a completed service, when it leaves the system. The LST $\psi_2(\theta,b_1,b_2,\ldots) = \int_{x=0}^{\infty} e^{-\theta x} dB_2(x,b_1,b_2,\ldots)$ (Rl $\theta \geqq 0$) of the d.f. $B_2(x,b_1,b_2,\ldots)$ of the time a non-priority customer spends from reaching the counter to completing service is

$$\psi_2(\theta,b_1,b_2,\ldots) = \sum_{n=0}^{\infty} \frac{\lambda_1{}^n \varphi_1(\theta)^n}{(\lambda_1+\theta)^n} e^{-(\lambda_1+\theta)b_{n+1}} \prod_{j=2}^{n} (1-e^{-(\lambda\ +\theta)b_j}) \tag{2.16}$$

where we define $\prod_{j=2}^{0} = \prod_{j=2}^{1} = 1$. When non-priority customers have a general service time requirement, i.e. all the $b_i (i = 1,2,\ldots)$ are random variables with identical d.f.'s $F_2(x)$ $(0 \leqq x < \infty)$ and LST $\alpha_2(\theta)$ (Rl $\theta \geqq 0$), then (2.16) reduces to

$$\psi_2(\theta) = \int_{x_1=0}^{\infty} \int_{x_2=0}^{\infty} \ldots \psi_2(\theta,x_1,x_2,\ldots) dF_2(x_1) dF_2(x_2) \ldots$$

$$= \frac{(\lambda_1+\theta)\alpha_2(\lambda_1+\theta)}{\lambda_1+\theta-\lambda_1\varphi_1(\theta)\{1-\alpha_2(\lambda_1+\theta)\}} . \tag{2.17}$$

For definition (b) of the waiting time the LST $\zeta_2(\theta)$ of the completion time distribution for a non-priority customer arriving at a queue free of other non-priority customers is given by

$$\zeta_2(\theta) = \psi_2(\theta)\vartheta_2(\theta). \tag{2.18}$$

The first two moments of $\psi_2(\theta)$ and $\zeta_2(\theta)$ are

$$\bar{\psi}_2 = \frac{(1+\lambda_1\varphi_1)(1-\alpha_2(\lambda_1))}{\lambda_1\alpha_2(\lambda_1)}$$

$$\bar{\zeta}_2 = \bar{\psi}_2 + \bar{v}_2$$

$$\psi_2''(0) = [\lambda_1{}^2\alpha_2(\lambda_1)\varphi_1''(0)\{1-\alpha_2(\lambda_1)\}+2(1+\lambda_1\overline{\varphi}_1)\{1-\alpha_2(\lambda_1)\}$$

$$\{1+\lambda_1\overline{\varphi}_1-\lambda_1\overline{\varphi}_1\alpha_2(\lambda_1)\}+2\lambda_1(1+\lambda_1\overline{\varphi}_1)\alpha_2'(\lambda_1)][\lambda_1{}^2\alpha_2(\lambda_1)^2]^{-1}$$

$$\zeta_2''(0) = \psi_2''(0)+2\bar{\bar{v}}_2\overline{\psi}_2+\vartheta_2''(0) \qquad (2.19)$$

The two preemptive repeat policies may exhsbit some rather different characteristics. For a repeat identical policy, long service requirements tend to be interrupted often; however, if a repeat different policy is operative then a long service requirement may be replaced after preemption by a shorter one, thus reducing the completion time. In fact $\overline{\psi}_2$ for a repeat identical policy is never less than for a repeat different policy with the same inter-arrival and service time distributions. From inspection of $\overline{\psi}_2$ from (2.14) and (2.18) this is so if $\alpha_2(-\lambda_1)\alpha_2(\lambda_1) \geqq 1$, which can be proved from the properties of a symmetric distribution on $(-\infty,\infty)$. Further both of these policies yield a completion time distribution whose mean is never less than for a fixed service time requirement with the same mean; this is intuitively obvious as a long service time requirement may delay several customers, while a short service time requirement is not likely to save much time for other customers.

(iii) The preemptive priority resume policy

This problem is dealt with by the same method as for the previous policies (i) and (ii); however the problem is now rather simpler, as the inherent difficulties are not as great as before. If non-priority customers require actual service for a total time b then the LST $\psi_2(\theta,b)$ of the distribution of the time a non-priority customer spends in the system after reaching the counter for the first time is

$$\psi_2(\theta,b) = e^{-(\lambda_1+\theta)b} + \int_{y_1=0}^{b} \lambda_1 e^{-\lambda_1 y_1} dy_1 \int_{z_1=y_1}^{\infty} dG_1(z_1-y_1)\{e^{-\lambda_1(b-y_1)-\theta(z_1-y_1+b)}$$

$$+ \int_{y_2=0}^{b-y_1} \lambda_1 e^{-\lambda_1 y_2} dy_2 \int_{z_2=y_2}^{\infty} dG_1(z_1-y_2)\{e^{-(\lambda_1+\theta)(b-y_1-y_2)-\theta(z_1+z_2)}$$

$$+ \ldots\ldots\}\}$$

$$= e^{-\{\lambda_1+\theta-\lambda_1\varphi_1(\theta)\}b} . \tag{2.20}$$

From this we readily obtain for a general d.f. $F_2(x)$ with LST $\alpha_2(\theta)$ of the time a non-priority occupies the server that

$$\psi_2(\theta) = \alpha_2(\lambda_1+\theta-\lambda_1\varphi_1(\theta)) \tag{2.21}$$

$$\zeta_2(\theta) = \psi_2(\theta)\vartheta_2(\theta). \tag{2.22}$$

The LST $\psi_2(\theta)$ has also been found by Miller (1960) and Gaver (1962). The first two moments of the completion time distributions are

$$\bar{\psi}_2 = (1+\lambda_1\bar{\varphi}_1)\mu_2$$

$$\psi_2''(0) = \lambda_1\mu_2\varphi_1''(0)+(1+\lambda_1\bar{\varphi}_1)\alpha_2''(0)$$

$$\bar{\zeta}_2 = \bar{\psi}_2+\bar{v}_2$$

$$\zeta_2''(0) = \psi_2''(0)+2\bar{v}_2\,\bar{\psi}_2+\vartheta_2''(0). \tag{2.23}$$

(iv) The head-of-the-line priority policy

Once a customer has reached the server there is no preemption so that the actual time in service of priority and non-priority customers has the IST's $\alpha_1(\theta)$ and $\alpha_2(\theta)$ respectively. However we are also interested in the time from the commencement of service of one non-priority customer to the point where the server is next available to serve another non-priority customer. This completion time is identical with that for a preemptive resume policy, so that its IST $\psi_2(\theta)$ is given by (2.21); similarly for $\zeta_2(\theta)$. These completion times are the same for the two policies, as the order of service is irrelevant to the total time that the server is occupied; for this reason the busy period distributions are identical.

The waiting time distribution for priority customers is no longer independent of the non-priority customers as the latter may now delay the service of the priority customers; this is not possible for a preemptive priority policy.

Considering only the Markov chain imbedded at points of departure of customers Miller (1960) has shown that the stationary waiting time distribution at these points has the IST $\mathcal{H}_1(\theta)$ given by

$$\mathcal{H}_1(\theta) = \frac{(1-\rho_1\rho_2)\theta+\lambda_2-\lambda_2\alpha_2(\theta)}{\theta-\lambda_1+\lambda_1\alpha_1(\theta)} \tag{2.24}$$

and the p.g.f. $R_1(z)$ of the stationary priority queue size distribution at these points of departure may be found from (2.18) of Miller as

$$R_1 \ (z) = \frac{(1-\rho_1-\rho_2)(1-z)\lambda_1\alpha_1(\lambda_1-\lambda_1 z)+\lambda_2\alpha_1(\lambda_1-\lambda_1 z)-\lambda_2 z\alpha_2(\lambda_1-\lambda_1 z).}{\nu_2\{\alpha_1(\lambda_1-\lambda_1 z)-z\}} \qquad (2.25)$$

The joint p.g.f. of the stationary number of priority and non-priority customers at arbitrary times has been obtained as the limit of the time-dependent solution by Jaiswal (1962). After a slight correction to his equation (41) we find that the p.g.f. $R_1(z)$ of the stationary priority queue size distribution is

$$R_1(z) = \frac{\alpha_1(\lambda_1-\lambda_1 z)}{\alpha_1(\lambda_1-\lambda_1 z)}\left\{(1-z)(1-\rho_1-\rho_2)+\frac{\lambda_2}{\lambda_1}\ (1-\alpha_2(\lambda_1-\lambda_1 z))\right\} \qquad (2.26)$$

which is different from (2.25); this indicates that in this case the result for an imbedded process is not sufficient to give the result for arbitrary points in time.

We now find the stationary waiting time distribution for priority customers by using the formula (3.9.1) due to Finch (1959). The probability P_{oo} that the server is idle is $P_{oo} = 1-\rho_1-\rho_2$ and from (2.26) the probability that there are no priority customers in the queue is

$$R_1(0) = 1-\rho_1-\rho_2+\frac{\lambda_2}{\lambda_1}\ (1-\alpha_2(\lambda_1)).$$

If a priority customer arrives to find the queue free of other priority customers it may not be able to obtain service immediately owing to the presence of non-priority customers. If it arrives to find a non-priority customer being served, its wait has the LST $\chi\ (\theta)$ given by

$$\chi\ (\theta) = \int_{x=o}^{\infty}\int_{y=o}^{x} \frac{\lambda_1 e^{-\lambda_1 y}\, e^{-\theta(x-y)}}{1-\alpha_2(\lambda_1)} \qquad dydF_2(x)$$

$$= \frac{\lambda_1\{\alpha_2(\theta)-\alpha_2(\lambda_1)\}}{(\lambda_1-\theta)\{1-\alpha_2(\lambda_1)\}} .$$

Thus the IST $\vartheta_1(\theta)$ of the wait of a priority customer arriving at a queue free of other priority customers is

$$\vartheta_1(\theta) = \frac{1-\rho_1-\rho_2+\lambda_2\{\alpha_2(\theta)-\alpha_2(\lambda_1)\}\{(\lambda_1-\theta)(1-\alpha_2(\lambda_1))\}^{-1}}{1-\rho_1-\rho_2+\lambda_2\lambda_1^{-1}(1-\alpha_2(\lambda_1))} . \tag{2.27}$$

Substitution in (3.9.1) shows the IST $\Phi_1(\theta)$ of the stationary waiting time distribution to be identical with $\mathcal{H}_1(\theta)$ given by (2.24). The stationarity condition shows that a stationary distribution exists whenever $0 \leqq \rho_1 + \rho_2 - \lambda_2\lambda_1^{-1}(1 - \alpha_2(\lambda_1)) < 1$, i.e. $\rho_1 < 1$ and $\bar{v}_1 < \infty$. It is seen below that a proper stationary distribution exists for non-priority customers when $\rho_1 + \rho_2 < 1$. When $\rho_1 + \rho_2 - \lambda_2\lambda_1^{-1}(1 - \alpha_2(\lambda_1)) < 1 \leqq \rho_1 + \rho_2$ a proper stationary distribution exists for priority customers but not for non-priority customers.

A slightly more general model is obtained by supposing that instead of there being priority customers there are interruptions of a general type to the service of non-priority customers. We suppose that the interruptions have a general distribution, which is not restricted to being that of a busy period, and the time between the end of one interruption and the commencement of the next has a negative exponential distribution. This problem may be dealt with by exactly the same method used for the priority non-priority queueing model, and we may consider each of the four types of priority policy discussed above;

the LST $\psi_2(\theta)$ of the completion time distribution is given by (2.12), (2.17), (2.21) depending on the priority policy required, where $\varphi_1(\theta)$ is the LST of the interruption time distribution. The LST $\vartheta_2(\theta)$ of the distribution of delay up to the commencement of service when a customer arrives at a queue free of other customers is given by (2.4).

If interruptions, or breakdowns to the server, can occur only when a customer is present, the problem is similar to the above except that now a busy period can begin only with the arrival of a customer and not with the start of an interruption; thus $\vartheta_2(\theta) = 1$.

We now make use of the formulae of Chapter 3, and the completion time distributions found above to obtain the stationary waiting time and queue size distributions for non-priority customers. The four priority policies may here be discussed together if we take $\psi_2(\theta)$ and $\zeta_2(\theta)$ appropriately, i.e. $\psi_2(\theta)$ is given by (2.12) for a preemptive priority repeat identical policy, by (2.17) for a preemptive priority repeat different policy and by (2.21) for the remaining two policies. Then for definitions (a) and (b) of the waiting time the LST's $\Phi_2(\theta)$ and $\Omega_2(\theta)$ respectively of the stationary waiting time distributions, which exist as proper distributions when $\lambda_2\bar{\psi}_2 < 1$, are given by

$$\Phi_2(\theta) = \frac{(1-\lambda_2\bar{\psi}_2)\{\lambda_2-(\lambda_2-\theta)\vartheta_2(\theta)\}}{(1+\lambda_2\bar{v}_2)\{\theta-\lambda_2+\lambda_2\psi_2(\theta)\}} \tag{2.28}$$

$$\Omega_2(\theta) = \frac{(1-\lambda_2\bar{\psi}_2)\{\theta+\lambda_2\psi_2(\theta)(1-\vartheta_2(\theta))\}}{(1+\lambda_2\bar{v}_2)\{\theta-\lambda_2+\lambda_2\psi_2(\theta)\}} \tag{2.29}$$

For the three preemptive policies the p.g.f. $r_2(z)$ of the stationary queue size distribution is given by (3.6.4) as

$$r_2(z) = \frac{(1-\lambda_2\overline{\psi}_2)\psi_2(\lambda_2-\lambda_2 z)\{1-z\vartheta_2(\lambda_2-\lambda_2 z)\}}{(1+\lambda_2\overline{v}_2)\{\psi_2(\lambda_2-\lambda_2 z)-z\}} \tag{2.30}$$

This last result is invalid for a head-of-the-line policy as a non-priority customer may depart from the system before its completion time is concluded. The p.g.f. $r_2(z)$ has been found by Gaver (1962) as

$$r_2(z) = \frac{(1-\lambda_2\overline{\psi}_2)\alpha_2(\lambda_2-\lambda_2 z)\{1-z\vartheta_2(\lambda_2-\lambda_2 z)\}}{(1+\lambda_2\overline{v}_2)\{\psi_2(\lambda_2-\lambda_2 z)-z\}} \quad . \tag{2.31}$$

Some of these results have been obtained by other authors. Miller (1960, (3.23) and (3.11)) has found $\Phi_2(\theta)$ for a preemptive resume and a head-of-the-line priority policy. Heathcote (1959) has obtained $r_2(z)$ for the exponential queue where $\alpha_1(\theta) = \mu_1(\mu_1+\theta)^{-1}$ and $\alpha_2(\theta) = \mu_2(\mu_2+\theta)^{-1}$, while recently Gaver (1962) has found $r_2(z)$ for all four priority policies.

For priority policies (iii) and (iv) we may also consider the total service time in the system at a particular instant, i.e. we neglect the effect of all later arrivals, which may upset the order of service of customers. We have already noted that for these policies the order of service is irrelevant to the total service time in the system; thus this is identical for these two policies. We use the results of Takács (1955) in the form (2.5.4) with K = 2; the IST $\mathfrak{H}_2(\theta)$ of the distribution of the total service time in the system, including that of the customer being served, is given by

$$\mathcal{H}_2(\theta) = \frac{(1-\rho_1-\rho_2)\theta}{\theta-\nu_2+\lambda_1\alpha_1(\theta)+\lambda_2\alpha_2(\theta)}\ , \tag{2.32}$$

which has also been obtained by Miller (1960, (3.18) and (3.24)).

4.3 Busy period distributions for K = 2

We begin by considering the busy period distribution of the priority non-priority system for a preemptive priority resume and a head-of-the-line priority policy. The length of a busy period has an identical distribution in both cases, as we have noted previously. The following method is not applicable to a preemptive repeat policy as the time a non-priority customer occupies the server now depends on later arrivals of priority customers.

The LST $\varphi_2(\theta) = \int_0^\infty e^{-\theta x}\, dG_2(x)$ (Rl $\theta \geq 0$) of the d.f $G_2(x)$ $(0 \leq x < \infty)$ of the length of a busy period may be determined from (2.5.7) with K = 2 as

$$\varphi_2(\theta) = \lambda_1\nu_2^{-1}\alpha_1(\nu_2+\theta-\nu_2\varphi_2(\theta))+\lambda_2\nu_2^{-1}\alpha_2(\nu_2+\theta-\nu_2\varphi_2(\theta)). \tag{3.1}$$

This has been obtained by Miller (1960) and by Heathcote (1959) for negative exponential service time distributions and a resume policy. The mean length of a busy period is

$$-\bar{\varphi}_2'(0) = \frac{\rho_1+\rho_2}{\nu_2(1-\rho_1-\rho_2)}\ . \tag{3.2}$$

We now provide a method where the busy period distributions for all four priority policies may be considered together, if we take the completion time distributions appropriately from §4.2. A busy period may commence with either a priority or a non-priority customer, and we separate the two cases. In the latter case the LST $\varphi_{22}(\theta) = \int_0^\infty e^{-\theta x} dG_{22}(x)$ (Rl $\theta \geqq 0$) of the d.f. $G_{22}(x)$ $(0 \leqq x < \infty)$ of the length of a busy period is the solution with $\lim_{\theta\to\infty} \varphi_{22}(\theta) = 0$ of

$$\varphi_{22}(\theta) = \int_{y=0}^{\infty} \sum_{n=0}^{\infty} \int_{x=0}^{\infty} e^{-\lambda_2 y} \frac{(\lambda_2 y)^n}{n!} e^{-\theta(y+x)} dB_2(y) dG_{22}^{n*}(x)$$

$$= \psi_2(\lambda_2+\theta-\lambda_2\varphi_{22}(\theta)), \tag{3.3}$$

which has been obtained independently by Gaver (1962, (9.1)). The LST $\varphi_{12}(\theta) = \int_0^\infty e^{-\theta x} dG_{12}(x)$ (Rl $\theta \geqq 0$) of a busy period distribution starting with a priority customer is similarly found as

$$\varphi_{12}(\theta) = \int_{y=0}^{\infty} \sum_{n=0}^{\infty} \int_{x=0}^{\infty} e^{-\lambda_2 y} \frac{(\lambda_2 y)^n}{n!} e^{-\theta(y+x)} dG_1(y) dG_{22}^{n*}(x)$$

$$= \varphi_1(\lambda_2+\theta-\lambda_2\varphi_{22}(\theta)). \tag{3.4}$$

The LST $\varphi_2(\theta)$ of the busy period distribution for the priority non-priority system is thus given by

$$\varphi_2(\theta) = \lambda_1 \nu_2^{-1} \varphi_{12}(\theta) + \lambda_2 \nu_2^{-1} \varphi_{22}(\theta), \tag{3.5}$$

which is equivalent to (3.1) for a preemptive priority resume and a head-of-the-line priority policy. The mean length of a busy period is from (3.5)

$$\bar{\varphi}_2 = -\varphi_2'(0) = \frac{\rho_1+\lambda_2(1-\rho_1)\bar{\psi}_2}{\nu_2(1-\rho_1)(1-\lambda_2\bar{\psi}_2)} \ . \tag{3.6}$$

The above properties of the busy period distributions are sufficient to extend our results to $K(\geqq 2)$ classes; however, for non-priority customers let us consider the more general problem of the joint d.f. $G_2(n,t)$ of the length t of a busy period and the number n of non-priority customers served in this period. Here a busy period starts with the arrival of a non-priority customer at a counter free of other non-priority customers. The transform $\varphi_2(\theta,z) = \sum_{t=0}^{\infty} z^n e^{-\theta t} dG_2(n,t)$ $(0 \leqq z \leqq 1,\ \mathrm{Rl}\ \theta \geqq 0)$ is given by (3.7.14) as

$$\varphi_2(\theta,z) = z\vartheta_2(\lambda_2+\theta-\lambda_2 x)\psi_2(\lambda_2+\theta-\lambda_2 x), \tag{3.7}$$

where x is the unique solution in $0 < z \leqq 1$, $\mathrm{Rl}\ \theta > 0$, of

$$x = z\ \psi_2(\lambda_2+\theta-\lambda_2 x).$$

4.4 The general case $K \geqq 2$

In this section we generalise the results of §4.2 and §4.3 to the case where there are $K \geqq 2$ priority classes of customers. From the point of view of class j customers ($2 \leqq j \leqq K$, with $j = K$ only for a head-of-the-line priority policy), the server appears to be alternately present for a time which has a negative exponential distribution with mean ν_{j-1}^{-1}, and absent for a time which is the busy period distribution for the first j-1 classes of customers. Thus the problem is similar to that for just two classes of customers; we employ the same methods as before, but do not give as many details as previously.

The LST $\vartheta_K(\theta)$ of the distribution of the time from the arrival of a class K customer in the queue when it is free of other class K customers (and after the end of the last class K completion time) to its reaching the server for the first time is

$$\vartheta_K(\theta) = \frac{\lambda_K[\lambda_K-\theta+\sum_{i=1}^{K-1} \lambda_i\{\varphi_{iK-1}(\theta)-\varphi_{iK-1}(\lambda_K)\}]}{\nu_K(\lambda_K-\theta)(1-g_{K-1})} , \tag{4.1}$$

where $\varphi_{ij}(\theta)$ $(1 \leq i \leq j \leq K)$ is the LST of a busy period distribution for the first j priority classes in isolation given that it commences with a class i customer, and are obtained below, and

$$g_{K-1} = \nu_K^{-1}\sum_{i=1}^{K-1} \lambda_i\varphi_{iK-1}(\lambda_K).$$

This has mean

$$\bar{v}_K = -\vartheta_K'(0) = -\lambda_K^{-1} + \frac{(1+\nu_{K-1}\bar{\varphi}_{K-1})}{\nu_K(1-g_{K-1})} . \tag{4.2}$$

We now proceed to the LST's $\psi_K(\theta) = \int_0^\infty e^{-\theta x}dB_K(x)$ and $\zeta_K(\theta) = \int_0^\infty e^{-\theta x}dD_K(x)$ (Rl $\theta \geq 0$) of the completion time distributions for class K customers for each of the four priority policies. In all cases we have

$$\zeta_K(\theta) = \vartheta_K(\theta)\psi_K(\theta) .$$

(i) The preemptive priority repeat identical policy

Suppose initially that class K require uninterrupted service of time b (> 0), while the first K-1 classes have general service time d.f.'s $F_i(x)$ $(1 \leq i \leq K-1,\ 0 \leq x < \infty)$. Then the LST $\psi_K(\theta,b)$ of the distribution of the time from a class K customer reaching the server for the first time to his departure from the system is

$$\psi_K(\theta,b) = e^{-(\nu_{K-1}+\theta)b} + \sum_{i=1}^{K-1} \int_{y=o}^{\infty} \lambda_i e^{-\nu_{K-1}y}\, dy \int_{x=o}^{\infty} dG_{iK-1}(x)$$

$$\left\{e^{-(\nu_{K-1}+\theta)b-\theta x} + \sum_{i=1}^{K-1} \lambda_i \int_{y_1=o}^{\infty} e^{-\nu_{K-1}y_1}\, dy_1 \int_{x_1=y_1}^{\infty} dG_{iK-1}(x_1-y_1)\right.$$

$$\left.\left\{e^{-(\nu_{K-1}+\theta)b-\theta(x+x_1)} + \ldots\ldots\right\}\right\}$$

$$= e^{-(\nu_{K-1}+\theta)b}\left[1-\sum_{i=1}^{K-1} \frac{\lambda_i}{\nu_{K-1}+\theta}\, \varphi_{iK-1}(\theta)\left\{1-e^{-(\nu_{K-1}+\theta)b}\right\}\right]^{-1} \qquad (4.3)$$

If class K customers require an uninterrupted service which has a d.f. $F_K(x)$ $(0 \leqq x < \infty)$ with LST $\alpha_K(\theta) = \int_o^{\infty} e^{-\theta x}\, dF_K(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$, then the completion time distribution has LST $\psi_K(\theta)$ given by

$$\psi_K(\theta) = \int_{x=o}^{\infty} e^{-(\nu_{K-1}+\theta)x}\left[1-\sum_{i=1}^{K-1} \frac{\lambda_i}{\nu_{K-1}+\theta}\, \varphi_{iK-1}(\theta)\left\{1-e^{-(\nu_{K-1}+\theta)x}\right\}\right]^{-1} dF_K(x)$$

(4.4)

The first two moments are for $R_K \geqq \nu_{K-1}$ and $R_K \geqq 2\nu_{K-1}$ respectively

$$\bar{\psi}_K = -\psi_K{}'(0) = \nu_{K-1}{}^{-1}(1+\nu_{K-1}\bar{\varphi}_{K-1})\{\alpha_K(-\lambda_{K-1})-1)\} \qquad (4.5)$$

$$\psi_K{}''(0) = \varphi_{K-1}{}''(0)\{\alpha_K(-\nu_{K-1})-1\}+2\nu_{K-1}{}^{-2}(1+\nu_{K-1}\ \bar{\varphi}_{K-1})$$

$$[(1+\nu_{K-1}\bar{\varphi}_{K-1})\{\alpha_K(-2\nu_{K-1})-\alpha_K(-\nu_{K-1})\}+\nu_{K-1}\alpha_K'(-\nu_{K-1})$$

$$-\nu_{K-1}\bar{\varphi}_{K-1}\{\alpha_K(-\nu_{K-1})-1\}] \ . \tag{4.5}$$

The same method may be used for the other priority policies; we simply quote the following results.

(ii) The preemptive priority repeat different policy

$$\psi_K(\theta) = \frac{\alpha_K(\nu_{K-1}+\theta)(\nu_{K-1}+\theta)}{\nu_{K-1}+\theta-\sum_{i=1}^{K-1}\lambda_i\varphi_{iK-1}(\theta)\{1-\alpha_K(\nu_{K-1}+\theta)\}} \ , \tag{4.7}$$

$$\bar{\psi}_K = \frac{(1+\nu_{K-1}\bar{\varphi}_{K-1})\{1-\alpha_K(\nu_{K-1})\}}{\nu_{K-1}\ \alpha_K(\nu_{K-1})} \tag{4.8}$$

$$\psi_K''(0) = [\nu_{K-1}^2\alpha_K(\nu_{K-1})\varphi_{K-1}''(0)(1-\alpha_K(\nu_{K-1}))+2(1+\nu_{K-1}\bar{\varphi}_{K-1})$$

$$(1-\alpha_K(\nu_{K-1}))\{1+\nu_{K-1}\bar{\varphi}_{K-1}-\nu_{K-1}\bar{\varphi}_{K-1}\alpha_K(\nu_{K-1})\}$$

$$+2\nu_{K-1}(1+\nu_{K-1}\bar{\varphi}_{K-1})\alpha_K'(\nu_{K-1})][\nu_{K-1}\alpha_K(\nu_{K-1})]^{-2} \ . \tag{4.9}$$

(iii) - (iv) The preemptive priority resume policy and the head-of-the-line priority policy

$$\psi_K(\theta) = \alpha_K(\nu_{K-1}+\theta-\sum_{i=1}^{K-1}\lambda_i\varphi_{iK-1}(\theta)) \tag{4.10}$$

$$\bar{\psi}_K = \mu_K(1+\nu_{K-1}\bar{\varphi}_{K-1}) \tag{4.11}$$

$$\psi_K''(0) = \nu_{K-1}\mu_K\varphi_{K-1}''(0)+(1+\nu_{K-1}\bar{\varphi}_{K-1})\alpha_K''(0) \tag{4.12}$$

Substitution of $\psi_K(\theta)$ and $\zeta_K(\theta)$ in (3.9.1) and (3.5.1) now gives the stationary waiting time distributions (for definitions (a) and (b) respectively) for class K customers as

$$\Phi_K(\theta) = \frac{(1-\lambda_K\bar{\psi}_K)\{\lambda_K-(\lambda_K-\theta)\vartheta_K(\theta)\}}{(1+\lambda_K\bar{\psi}_K)\{\theta-\lambda_K+\lambda_K\psi_K(\theta)\}} \tag{4.13}$$

$$\Omega_K(\theta) = \frac{(1-\lambda_K\bar{\psi}_K)\{\theta+\lambda_K\psi_K(\theta)(1-\vartheta_K(\theta))\}}{(1+\lambda_K\bar{v}_K)\{\theta-\lambda_K+\lambda_K\psi_K(\theta)\}} \tag{4.14}$$

For priority policies (iii) and (iv) $\Phi_K(\theta)$ has been found by Miller (1960). These are proper distributions for $\lambda_K\bar{\psi}_K < 1$. The p.g.f. $r_2(z)$ of the stationary queue size distribution for class K customers for preemptive and head-of-the-line priority policies are respectively

$$r_2(z) = \frac{(1-\lambda_K\bar{\psi}_K)\psi_K(\lambda_K-\lambda_K z)\{1-z\vartheta_K(\lambda_K-\lambda_K z)\}}{(1+\lambda_K\bar{v}_K)\psi\ _K(\lambda_K-\lambda_K z)-z\}} \tag{4.15}$$

$$r_2(z) = \frac{(1-\lambda_K\bar{\psi}_K)\alpha_K(\lambda_K-\lambda_K z)\{1-z\vartheta_K(\lambda_K-\lambda_K z)\}}{(1+\lambda_K\bar{v}_K)\{\psi_K(\lambda_K-\lambda_K z)-z\ \}} \tag{4.16}$$

We have used the busy period distributions for the first K-1 classes in isolation to find the completion time distributions for class K customers; we now obtain the busy period distributions for the system of K classes.

As in §4.3 we may use (2.5.7) to find the LST $\varphi_K(\theta)$ for the busy period distribution for priority policies (iii) and (iv) of the system as the solution with $\lim_{\theta\to\infty} \varphi_K(\theta) = 0$ of

$$\varphi_K(\theta) = \nu_K^{-1} \sum_{i=1}^{K} \lambda_i \alpha_i (\nu_K + \theta - \nu_K \varphi_K(\theta)). \tag{4.17}$$

In general a busy period may commence with a customer from any class i ($1 \leqq i \leqq K$). The LST $\varphi_{KK}(\theta) = \int_0^\infty e^{-\theta x}\, dG_{KK}(x)$ (Rl $\theta \geqq 0$) of a busy period distribution starting with a class K customer is

$$\varphi_{KK}(\theta) = \int_{y=o}^{\infty} \sum_{n=o}^{\infty} \int_{x=o}^{\infty} e^{-\lambda_K y} \frac{(\lambda_K y)^n}{n!} e^{-\theta(y+x)} dG_{KK}^{n*}(x) dB_K(y)$$

$$= \psi_K(\lambda_K + \theta - \lambda_K \varphi_{KK}(\theta)) \tag{4.18}$$

For a busy period beginning with a class K-1 customer the LST $\varphi_{K-1K}(\theta)$ of its duration is

$$\varphi_{K-1K}(\theta) = \int_{y=o}^{\infty} \sum_{n=o}^{\infty} \int_{x=o}^{\infty} e^{-\lambda_K y} \frac{(\lambda_K y)^n}{n!} e^{-\theta(y+x)} dG_{K-1\ K-1}(y) dG_{KK}^{n*}(x)$$

$$= \varphi_{K-1\ K-1}(\lambda_K + \theta - \lambda_K \varphi_{KK}(\theta))$$

and in general we obtain

$$\varphi_{iK}(\theta) = \varphi_{ii}(\nu_K - \nu_i + \theta - \sum_{j=i+1}^{K} \lambda_j \varphi_{jK}(\theta)) \qquad 1 \leqq i \leqq K-1. \tag{4.19}$$

Thus the LST $\varphi_K(\theta)$ of the busy period distribution for the system of K classes is given by

$$\varphi_K(\theta) = \nu_K^{-1} \sum_{i=1}^{K} \varphi_{iK}(\theta) , \tag{4.20}$$

which is equivalent to (4.17) for a preemptive priority resume and a head-of-the-line priority policy. The mean duration of a busy period is

$$-\varphi_K'(0) = \left\{\nu_K \prod_{i=1}^{K} (1-\lambda_i\bar{\Psi}_i)\right\}^{-1} -1. \tag{4.21}$$

We see that from the solutions obtained for the completion time and busy period distributions, we can find the completion time distributions for the j-th priority class in terms of the busy period distributions for the (j-1)th class in isolation. We may then obtain the busy period distributions for the j-th class from the completion time distributions for the j-th class. Thus the completion time distributions, and hence the waiting time, and busy period distributions can be constructed by iteration from one class to the next.

PART 3

ROAD TRAFFIC THEORY

CHAPTER 5

DELAY AT AN INTERSECTION: DISCRETE TIME

5.1 General remarks

Mathematical problems related to road traffic theory have recently been receiving a large amount of attention; this is hardly surprising as road transport now plays an important part in the structure of any mechanised economy. For example, Tanner (1961a) points out that in Britain it absorbs 10 to 15 percent of the total national expenditure; further, it is estimated that in 1956 over four hundred million vehicle hours were lost in delays arising from road and traffic conditions, and this is increasing rapidly every year.

Several types of traffic problems, such as the formation of traffic on roads and delays at junctions and in networks, have been considered by a number of different methods: kinematics, hydrodynamics, network theory, percolation processes and queueing theory, each having been applied to some extent. The Operations Research bibliography on road traffic theory (1961) gives a detailed bibliography of all phases of mathematical applications to road traffic up to 1960.

In this Part 3 of the thesis we are concerned only with the application of queueing theory to problems of delay at junctions and on long, uninterrupted roads. Early work on the delay at an intersection, with and without traffic lights, was carried out by Adams (1936) and Garwood (1940), and has been continued by Tanner (1951, 1953, 1961a,b, 1962), Raff (1951), Newell (1955), Winsten (1956) and others. In the earlier works it was usually assumed that vehicles on both roads acted independently of each other and that the time between any two successive vehicles had a negative exponential distribution, so that the passage of vehicles past a point formed of homogeneous Poisson process. This has been found to be a reasonable approximation for low densities of vehicles, but is unsatisfactory for medium or high densities when the interaction between vehicles tends to be greater. For the latter case some more complicated formations of traffic, where bunching may take place, have been used; for example Tanner (1953, 1961a, 1962) has made use of the Borel-Tanner distribution, and the random queues model of A. Miller (1961, 1962a) is of a similar form. Winsten (1956) has described a discrete time model for intersection delay.

In the next two chapters we discuss two problems, one in discrete and the other in continuous time, which represent two different ways of looking at intersection problems. The first discrete time model considers a comparatively simple construction of traffic flow, which may be applied to quite a wide range of conditions, such as a two-way major road; the second, which gives a more highly sophisticated formation of traffic, may be applied to only a comparatively narrow range of practical situations. In most practical cases there are several lanes of traffic, each of which may have considerable internal interaction, and it is necessary to make a decision as to which aspect is the more important. We do not discuss the comparative importance of these two approaches, but will restrict ourselves to describing an example of each.

5.2 Introduction

We suppose a one-lane one-way minor road intersects with either a one-way one-lane or a two-way one-lane major road; the traffic in the major road has absolute right of way over that in the minor road. Except in §5.6 there are no traffic lights or policemen controlling the junction, but there may be a stop sign in the minor road; the gap acceptance time distributions described below depend on the presence or absence of such a sign, but we do not explicitly differentiate between them. Arrivals of vehicles at the junction in both major and minor roads occur at discrete points of time $t = 0,1,2,\ldots$ (in an arbitrary unit). In the minor road, vehicles arrive at the junction with a geometric inter-arrival time distribution, such times having the probabilities bd^{j-1} $(0 < b < 1,\ d = 1-b;\ j = 1,2,\ldots)$.

In the major road it is desirable to allow some degree of dependence between vehicles close to each other; an account of this dependence is given in Chapter 6. For a one-way major road we can find the stationary waiting time distribution for a quite general formation of road traffic; however, it has not been found possible to extend this to a two-way major road in any general form, and we consider only some special cases. In a simple case where dependence between successive vehicles is allowed, we may suppose there is a Markov chain relation between them. We call a time point in a major road lane green (G) if there is not a vehicle there, and red (R) if there is. Let us consider the first major road stream encountered by minor road vehicles, this being the only stream for a one-way major road. We suppose that the presence or absence of vehicles at successive time points in this lane is governed by a Markov chain relation with the transition matrix

$$\underline{A} = \begin{array}{cc} & \begin{array}{cc} G & R \end{array} \\ \begin{array}{c} G \\ R \end{array} & \begin{bmatrix} a & c \\ \alpha & \gamma \end{bmatrix} \end{array}, \tag{2.1}$$

where $0 < a < 1$, $c = 1-a$, $0 < \alpha < 1$, $\gamma = 1-\alpha$; thus the probability that a green point is followed by another green point is a, and so on. When there is a two-way major road we define the traffic in the second major road stream in a similar manner; the transition matrix $\underline{B}$ governing this second stream is

$$\underline{B} = \begin{array}{cc} & \begin{array}{cc} G & R \end{array} \\ \begin{array}{c} G \\ R \end{array} & \begin{bmatrix} x & y \\ \beta & \delta \end{bmatrix} \end{array}, \tag{2.2}$$

where $0 \leqq x \leqq 1$, $y = 1-x$, $0 \leqq \beta \leqq 1$, $\delta = 1-\beta$. In the special case where vehicles in each stream are independent of each other we have $a = \alpha$ and $x = \beta$.

A minor road vehicle may arrive at the intersection to find that there is a queue of minor road vehicles waiting to enter the intersection, or that there are no other minor road vehicles waiting so that it proceeds immediately to the head of the 'queue'. We formulate slightly different rules for each of these two cases. If a minor road vehicle arrives at an empty queue it waits until a critical gap of at least w green time points appears in the major road (or roads if there is two-way traffic), where w is a random variable with probabilities $\Pr\{w = i\} = \xi_i$ $(i = 1,2,\ldots)$, and p.g.f. $\xi(s) = \sum_{i=1}^{\infty} \xi_i s^i (|s| \leq 1)$, the mean $\bar{\xi} = \xi'(1)$ being finite; it enters the junction at the w-th time point of this gap. A vehicle joining a non-empty queue waits until all the vehicles ahead of it have departed. If there are at least v more green points, where v is a random variable with probabilities $\Pr\{v = i\} = \chi_i$ $(i = 1,2,\ldots)$, p.g.f. $\chi(s) = \sum_{i=1}^{\infty} \chi_i s^i$ $(|s| \leq 1)$ and finite mean $\bar{\chi}$, then it enters the junction at the v-th of these time points. However, if these v time points are not all green, then it waits until the first gap of at least w appears and enters the junction at the w-th of these.

We wish to obtain the stationary waiting time and delay distributions. Winsten (1956) has considered a general model for major road traffic when there is one-way traffic and rules for entering the intersection which are rather more restricted than ours; however, it has not been possible to extend Winsten's results to a two-way major road, except in the special case of independent traffic with a geometric distribution between successive vehicles.

5.3 The service time

Throughout this Part 3 we make use of the queueing terminology employed in Part 2. Waiting time is defined as the time a vehicle spends from joining the minor road queue to reaching the head of the queue while the service time lasts from this instant until the vehicle enters the junction. The delay is the total time a vehicle spends at the junction, which is just the sum of the waiting and service times. The service time distributions depend on whether a vehicle arrives at an empty or a non-empty queue, so that formulae of the type discussed in Chapter 3 are appropriate.

We begin by considering a one-way major road. As an indication of the increase in difficulty which follows when there is dependence between vehicles, we consider the simple case where major road vehicles are independent, i.e. $a = \alpha$, $c = \gamma$, and then extend our argument to the more general case. We suppose initially that a minor road vehicle requires critical gap acceptance times of w and v, and later generalise this case.

If a minor road vehicle arrives at the junction when there are no other minor road vehicles waiting, its service time distribution is identical with the distribution of the recurrence time for a success run of length w in a sequence of Bernoulli trials, each having the probability of success a. The p.g.f. $g(s,w) = \sum_{i=1}^{\infty} g_i(w)s^i \quad (|s| \leq 1)$ of the probabilities $g_i(w)$ $(i = 1,2,\ldots)$ of this distribution has been obtained by Feller (1957, p.299) as

$$g(s,w) = \frac{a^w s^w (1-as)}{1-s+ca^w s^{w+1}} . \tag{3.1}$$

A vehicle joining a non-empty queue has a service time v if there are v further green points when it reaches the head of the queue; otherwise it waits for a gap of size w. The p.g.f. $f(s,v,w)$ of the service time distribution in this case is thus

$$\begin{aligned} f(s,v,w) &= a^v s^v + \sum_{j=1}^{v} a^{j-1} c \sum_{i=1}^{\infty} g_i(w) s^{i+j} \\ &= a^v s^v + \frac{(1-a^v s^v)\; g(s,w)}{1 - as} \\ &= \frac{(1-s)a^v s^v + ca^w s^{w+1}}{1 - s + ca^w s^{w+1}} . \end{aligned} \tag{3.2}$$

When the gap acceptance times are random variables we obtain from (3.1) and (3.2) that

$$g(s) = \sum_{i=1}^{\infty} g(s,i)\, \xi_i = \sum_{i=1}^{\infty} \frac{a^i s^i (1-as)}{1-s+ca^i s^{i+1}} \; \xi_i \tag{3.3}$$

$$f(s) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f(s,j,i)\xi_i \chi_j = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left\{ \frac{(1-s)a^j s^j + ca^i s^{i+1}}{1-s+ca^i s^{i+1}} \right\} \xi_i \chi_j . \tag{3.4}$$

Except for some special cases it is not possible to evaluate these summations explicitly; however, moments of the service time distributions may be found by differentiation. This is similar to a discrete analogue of the completion time distributions for a preemptive priority repeat identical system discussed in the previous chapter. We suppose $\xi(s)$ has a radius of convergence $S \geqq 1$. In our case the means of the service time distributions are finite if $a^{-1} \leqq S$ and the second moments are finite if $a^{-2} \leqq S$. We obtain from (3.3) and (3.4) that

$$\begin{aligned} g'(1) &= c^{-1}\{\xi(a^{-1})-1\} \\ f'(1) &= c^{-1}\ \xi(a^{-1})\{1-X(a)\} \end{aligned} \tag{3.5}$$

$$\begin{aligned} g''(1) &= 2c^{-2}[\xi(a^{-2})-(2-a)\xi(a^{-1})+ac\xi'(a^{-1})+c] \\ f''(1) &= 2c^{-2}[\{1-X(a)\}\{\xi(a^{-2})-c\xi(a^{-1})+ac\xi'(a^{-1})\}-acX'(a)\xi(a^{-1})]. \end{aligned} \tag{3.6}$$

Let us now extend the above example to the case where the major road vehicles are no longer independent of each other, but are related by the Markov chain relation (2.1). We must modify our previous argument, as the service time distribution for a minor road vehicle arriving at an empty queue depends on the length of time since the departure of the last minor road vehicle. Let us consider the probabilities

$$G_{22}(i) = \Pr\{\text{time point } i \text{ is } G \mid \text{time point } 0 \text{ is } G\}$$
$$G_{21}(i) = \Pr\{\text{time point } i \text{ is } R \mid \text{time point } 0 \text{ is } G\}.$$

Obviously $G_{22}(i)+G_{21}(i) = 1$, $G_{22}(0) = 1$, $G_{21}(0) = 0$, and the following relations hold:

$$G_{22}(i+1) = aG_{22}(i)+\alpha G_{21}(i)$$
$$G_{21}(i+1) = cG_{22}(i)+\gamma G_{21}(i), \tag{3.7}$$

from which we obtain the generating functions

$$G_{22}{*}(s) = \sum_{i=o}^{\infty} G_{22}(i)s^i = \frac{1 - as}{(1-s)(1-(a-\alpha)s)}$$

$$G_{21}{*}(s) = \sum_{i=o}^{\infty} G_{21}(i)s^i = \frac{\alpha s}{(1-s)(1-(a-\alpha)s)} \ . \tag{3.8}$$

We wish to determine the p.g.f. $H_R(s,w) \equiv H(s,w) = \sum_{i=1}^{\infty} H_i(w)s^i$ ($|s| \leqq 1$) of the distribution $H_i(w)$ ($i = w, w+1, \ldots$) of the time from a red point at zero to the first gap of at least w green points in the major road traffic; we call this a w block. The method of Winsten (1956, p.38) for the special case $a = \alpha$, $c = \gamma$, may be generalised to find this distribution. If a given time point is red the probability that it is the last red point of the w block is $\alpha a^{w-1} = L$, and the probability ℓ_j that the block continues with the arrival of a vehicle within the next w time points is

$$(1-L)\ell_j = \begin{cases} \gamma & j = 1 \\ \alpha\ ca^{j-2} & j = 2,3,\ldots,w, \end{cases} \tag{3.9}$$

which has the p.g.f.

$$L(s,w) = \sum_{i=1}^{w} \ell_j s_j = \left\{ \gamma s + \frac{\alpha cs^2(1-a^{w-1}s^{w-1})}{1-as} \right\} \Big/ (1-L) \ . \tag{3.10}$$

The probability that a block has i vehicles in it is $L(1-L)^{i-1}$ ($i = 1,2,\ldots$), so that with (3.10) we obtain

$$H(s,w) = \sum_{i=1}^{\infty} L(1-L)^{i-1} L(s,w)^{i-1} s^{w}$$

$$= \frac{Ls^{w}}{1-(1-L)L(s,w)}$$

$$= \frac{(1-as)\, a^{w-1} s^{w} \alpha}{1-(a+\gamma)s+(a\gamma-\alpha c)s^{2}+\alpha c a^{w-1} s^{w+1}} \tag{3.11}$$

As in (3.2) the p.g.f. $f(s,v,w)$ for a vehicle joining a non-empty queue is

$$f(s,v,w) = a^{v}s^{v} + \frac{cs(1-a^{v}s^{v})}{1-as} H(s,w). \tag{3.12}$$

If a vehicle joins an empty queue we average over the time since the departure of the last minor road vehicle to obtain

$$g(s,w) = \sum_{i=1}^{\infty} bd^{i-1}\{G_{22}(i)f(s,w,w) + G_{21}(i)H(s,w)\}$$

$$= \frac{(1-\gamma d-bc)f(s,w,w)+cH(s,w)}{1-\gamma d+cd}. \tag{3.13}$$

Thus we obtain for general gap acceptance time distributions that

$$g(s) = \sum_{i=1}^{\infty} g(s,i)\xi_{i} \tag{3.14}$$

$$f(s) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f(s,j,i)\xi_{i}\chi_{j}, \tag{3.15}$$

from which we derive by differentiation

$$g'(1) = \frac{(c+\alpha)}{\alpha c}\,\xi(a^{-1}) - \frac{a(c+\alpha)}{\alpha c} + \frac{c(\alpha-a)}{\alpha(1-\gamma d+cd)}$$

$$f'(1) = \frac{a(c+\alpha)}{\alpha c}\,\xi(a^{-1})\{1-\chi(a)\} \tag{3.16}$$

$$g''(1) = \frac{2(c+\alpha)^2a^2}{\alpha^2c^2}\,\xi(a^{-2}) + \frac{2a^2(c+\alpha)}{\alpha c}\xi'(a^{-1}) + \frac{2\{c\alpha+a(1-\gamma d-bc)\}}{\alpha c(1-\gamma d+cd)}$$

$$- \frac{2a\{((c+\alpha)^2a+\alpha c)(1-\gamma d-bc)+\alpha c(2c+\alpha)\}}{\alpha^2c^2(1-\gamma d+cd)}\xi(a^{-1})$$

$$f''(1) = \frac{2a}{(\alpha c)^2}\,[\{1-\chi(a)\}\{a(\alpha+c)^2\xi(a^{-2})-\alpha c\xi(a^{-1})$$

$$+a(\alpha+c)\alpha c\xi'(a^{-1})\}-(\alpha+c)a\alpha c\chi'(a)\xi(a^{-1})] \tag{3.17}$$

We now turn to a two-way major road with the minor road vehicles requiring to cross both major road streams. We suppose that a minor road vehicle requires gaps of at least w(and v) in both major road streams before being able to enter the intersection. A time point in the major road may be red in both streams, red in the first and green in the second, green in the first and red in the second, or green in both; we denote these respectively by RR, RG, GR, GG or 1,2,3,4. These are the four states of a Markov chain with the transition matrix $\underline{C} = (c_{ij})$:

$$\underline{C} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\!\!\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \left[\begin{array}{cccc} \gamma\delta & \gamma\beta & \alpha\delta & \alpha\beta \\ \gamma y & \gamma x & \alpha y & \alpha x \\ c\delta & c\beta & a\delta & a\beta \\ cy & cx & ay & ax \end{array}\right] \end{array}. \tag{3.18}$$

We define

$$G_{ij}(k) = \Pr\{\text{time point } k \text{ is } j \,|\, \text{time point } 0 \text{ is } i\} \qquad i,j = 1,2,3,4$$

and

$$G_{ij}^*(s) = \sum_{k=0}^{\infty} G_{ij}(k)s^k \qquad |s| \leqq 1,$$

with $\underline{G}(s)$ as the (4×4) matrix whose elements are $G_{ij}^*(s)$. Forming difference equations for the $G_{ij}(k)$ and summing we obtain

$$\begin{aligned}\underline{G}(s) &= \underline{I} + \underline{C}\, s\, \underline{G}(s)\\ &= (\underline{I} - \underline{C}\, s)^{-1},\end{aligned}$$

from which we obtain

$$G_{ij}^*(s) = \frac{|\underline{K}_{ij}(s)|}{|\underline{I} - \underline{C}\, s|}, \tag{3.19}$$

where $\underline{K}_{ij}(s)$ is similar to the matrix $(\underline{I} - \underline{C}s)$ with the j-th column replaced by the i-th column of $\underline{I}$; evaluation of the determinants in (3.19) yields the solutions for $G_{ij}^*(s)$.

We wish to determine the p.g.f. $H_i(s,w) = \sum_{j=1}^{\infty} h_i(j,w)s^j$ ($|s| \leqq 1$; $i = 1,2,3$) of the distribution of the time from a type i time point to the first instant at which there are w consecutive time points which are green in both streams. Any type of time point other than green in both streams is regarded by a minor road vehicle as a red point, so a w block ends when there are w successive GG points. The p.g.f. $H(s,w)$ of a block is readily seen to be

$$H(s,w) = \sum_{i=1}^{3} \frac{c_{4i}}{1-c_{44}} H_i(s,w). \tag{3.20}$$

If a time point is of type i $(i = 1,2,3)$, a block ends w points later with probability $P_i = c_{i4}\, c_{44}^{w-1}$ and continues with another vehicle arriving at point $k(1 \leqq k \leqq w)$ with probability

$$\ell_{ij}(k) = \begin{cases} c_{ij} & k = 1 \\ c_{i4}c_{44}^{k-2}c_{4j} & k = 2,3,\ldots,w, \end{cases}$$

with

$$L_{ij}(s,w) = \sum_{k=1}^{w} \ell_{ij}(k)s^{w}$$

$$= c_{ij}s + \frac{c_{i4}c_{4j}s^{2}(1-c_{44}^{w-1}s^{w-1})}{1-c_{44}s}$$

We write $\underline{H}' = (H_1(s,w),\ H_2(s,w),\ H_3(s,w))$, $\underline{P}' = (P_1,\ P_2,\ P_3)$ and $\underline{L}(s) = (L_{ij}(s,w))$; then we can extend the argument used for (3.11) to obtain

$$\underline{H} = \underline{P}\ s^{w} + \underline{L}(s)\underline{H}$$

$$= (\underline{I} - \underline{L}(s))^{-1}\ \underline{P}\ s^{w}, \tag{3.21}$$

so that

$$H_i(s,w) = \frac{c_{44}^{w-1}s^{w}}{|\underline{I} - \underline{L}(s)|} \sum_{j=1}^{3} c_{j4}(-1)^{i+j}L_{ij}^{*} \tag{3.22}$$

where L_{ij}^{*} is the determinant of the cofactor of $(\delta_{ij} - L_{ij}(s))$ in $(\underline{I} - \underline{L}(s))$. $|\underline{I} - \underline{L}(s)|$ is a (3×3) determinant and the L_{ij}^{*} are (2×2) determinants, so that (3.22) is easily evaluated. $H(s,w)$ is found from (3.20) and (3.22).

If a minor road vehicle joins a non-empty queue there are green points in both major road streams when it reaches the head of the minor road queue. If there are a further v green (GG) points the vehicle has a service time of v; otherwise there is a delay until a gap of size w appears for the first time. The p.g.f. $f(s,v,w)$ of the service time distribution for a vehicle joining a non-empty queue is

$$f(s,v,w) = c_{44}{}^{v} s^{v} + \frac{s(1-c_{44}{}^{v}s^{v})}{1-c_{44}s}(1-c_{44})\, H(s,w). \tag{3.23}$$

If a vehicle arrives at an empty queue the service time distribution has the p.g.f.

$$g(s,w) = \sum_{i=1}^{\infty} bd^{i-1} \left\{ \sum_{j=1}^{3} G_{4j}(i) H_j(s,w) + G_{44}(i) f(s,w,w) \right\} . \tag{3.24}$$

We finally obtain the service time p.g.f.'s for general gap acceptance time distributions as

$$g(s) = \sum_{i=1}^{\infty} g(s,i)\, \xi_i \tag{3.25}$$

and

$$f(s) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f(s,j,i)\, \xi_i\, \chi_j \, . \tag{3.26}$$

5.4 <u>The stationary waiting time distribution</u>

This problem is a special case of a discrete time queueing (or dam) model where customers arrive independently in a binomial process with parameter b, and the service time distribution depends on whether or not there are any other customers in the queue when a customer arrives. This is a discrete analogue of the problem discussed in Chapter 3. In the discrete case the waiting time Z_t $(= 0,1,2,\ldots;\ t = 0,1,2,\ldots)$ is subject to a discrete jump when a customer (vehicle) arrives, this being the service time for the customer, and decreases uniformly with time unless the queue is empty. When a customer arrives at an empty queue its service time distribution is g_i $(i = 1,2,\ldots)$ with p.g.f. $g(s) = \sum_{i=1}^{\infty} g_i s^i$ $(|s| \leqq 1)$ and mean $g'(1) < \infty$ while if the customer joins a non-empty queue its service time distribution is f_i $(i = 1,2,\ldots)$ with p.g.f. $f(s) = \sum_{i=1}^{\infty} f_i s^i$ $(|s| \leqq 1)$ and mean $f'(1) < \infty$. Z_t is a Markov chain with a countable infinity of states.

The stationary probabilities $P_i = \lim_{t\to\infty} \Pr\{Z_t = i\}$ of the waiting time, which constitute a proper distribution only if $bf'(1) < 1$ and $g'(1) < \infty$, satisfy the difference equations

$$P_o = (d+bg_1)P_o+dP_1$$

$$P_i = \sum_{j=1}^{i} bf_{i+1-j}P_j+bg_{i+1}P_o\ dP_{i+1}. \tag{4.1}$$

By multiplying the i-th relation of (4.1) by s^{i+1} and summing both sides we obtain the p.g.f. $P(s) = \sum_{i=o}^{\infty} P_i s^i$ $(|s| \leqq 1)$ as

$$P(s) = \frac{P_o\{(s-1)d+bg(s)-bf(s)\}}{s-d-bf(s)}\ . \tag{4.2}$$

The probability P_o that the queue is empty is found from (4.2) by using the condition $P(1) = 1$ so that

$$P_o = \frac{1-bf'(1)}{d+bg'(1)-bf'(1)}. \tag{4.3}$$

The moments of the stationary waiting time distribution may be obtained from (4.3) by differentiation; the mean is

$$P'(1) = \frac{b\,P_o g''(1) + b(1-P_o)f''(1)}{2\{1-bf'(1)\}} \tag{4.4}$$

The distribution of delay may be found from that of the waiting time by convoluting the waiting and the service time distributions. The mean delay $\bar{d}$ is

$$\begin{aligned}\bar{d} &= P'(1) + P_o g'(1) + (1-P_o)f'(1)\\ &= \frac{(1-bf'(1))\{bg''(1) + 2g'(1) - 2bf'(1)\} + b\{g'(1) - 1\}f''(1)}{2\{1-bg'(1)\}\{d+bg'(1)-bf'(1)\}}.\end{aligned} \tag{4.5}$$

The stationary distributions of waiting time and delay for our traffic models are obtained by substituting the appropriate formulae of §5.3 in those of this section.

5.5 The continuous analogue

As in Chapter 1 we may pass from the discrete time model to one in continuous time by taking our unit of measurement as Δ (> 0) and letting $\Delta \to 0$ in an appropriate manner. As we take this limit the dependence between vehicles in the discrete case is lost, for as $\Delta \to 0$, the dependence is allowed only over an infinitesimal time. For a one-way major road we put

$$\begin{aligned} a &= \lambda\Delta + o(\Delta),\ b = \mu\Delta + o(\Delta)\\ w &= W/\Delta,\ v = V/\Delta\ , \end{aligned} \tag{5.1}$$

in such a way that as $w \to \infty$, $v \to \infty$, $\Delta \to 0$, $w\Delta \to W$, $v\Delta \to V$; we have $\lim \sum_{j=1}^{w} \xi_{j\Delta} = F(W)$, $\lim \sum_{j=1}^{v} \chi_{j\Delta} = G(V)$ $(0 \leq V, W < \infty)$, where $F(W)$ and $G(V)$ are d.f.'s on $(0,\infty)$. As $\Delta \to 0$ the duration between arrivals at the intersection in both major and minor roads has a negative exponential distribution with mean μ^{-1} and λ^{-1} respectively; thus vehicles in both streams arrive at the intersection in Poisson processes.

The IST's $\zeta(\theta) = \int_0^\infty e^{-\theta x}\, dD(x)$ and $\psi(\theta) = \int_0^\infty e^{-\theta x} dB(x)$ $(\mathrm{Rl}\ \theta \geq 0)$ of the service time distributions of vehicles joining empty and non-empty queues respectively are determined from those in discrete time using (5.1) and the limit $\Delta \to 0$ as

$$\zeta(\theta) = \lim_{\Delta\to 0} g(e^{-\theta\Delta}) = \int_0^\infty \frac{(\lambda+\theta)}{\lambda+\theta e^{(\lambda+\theta)x}} dF(x) \tag{5.2}$$

$$\psi(\theta) = \lim_{\Delta\to 0} f(e^{-\theta s}) = \int_{x=0}^\infty \int_{y=0}^\infty \left\{ e^{-(\lambda+\theta)y} + \frac{\lambda-\lambda e^{-(\lambda+\theta)y}}{\lambda+\theta e^{(\lambda+\theta)x}} \right\} dG(y)dF(x) \tag{5.3}$$

In the special case where $G(x-w)$ is the unit step function, $\zeta(\theta)e^{\theta w}$ has been obtained by Tanner (1951) for the delay to pedestrians wishing to cross a road.

The stationary waiting time distribution may be obtained by substituting (5.2) and (5.3) in (3.5.1) or by the limiting process above from the discrete result (4.2).

5.6 Traffic lights

Winsten (1956) has considered the delay to a stream of traffic at a set of fixed interval traffic lights. Let us suppose that vehicles arrive at a traffic light intersection independently in a binomial process with parameter b $(0 < b < 1)$. There are alternate red periods of integral time $R(\geqq 1)$, when vehicles cannot cross the intersection, and green periods of integral time $G(\geqq 1)$ when vehicles can enter the intersection. The periods R and G are arranged so that at most G vehicles can enter the intersection in a red and green period; G may not represent the actual green period, since there may be vehicles which are waiting at the end of a red period taking time to accelerate. For this model Winsten (1956) has shown that the expected delay to vehicles is given by

$$\bar{d} = \frac{R}{2bd(R+G)} \{(R+1)b + 2\bar{g}\} , \tag{6.1}$$

where $\bar{g}$ is the expected delay at the end of a green period; the times t_i are regeneration points of the process. Considering only these times t_i the process is identical with that for an infinite dam in discrete time with maximum release G and with input distribution

$$p_i = \binom{R+G}{i} b^i \, d^{R+G-i} \qquad i = 0,1,\ldots,R+G \tag{6.2}$$

which has been discussed in §1.2. We must have $\rho = (R+G)b < G$ for a stationary distribution to exist. The probabilities P_i represent the number of vehicles in the queue at the end of a green period; the mean number of vehicles $P'(1)$ is given by (1.2.14) with $p(z) = (bz+d)^{R+G}$ and $G = m$. The mean delay $\bar{g}$ at these points is given by $\rho^{-1}\ P'(1)$ as vehicles must depart at the same rate as they arrive. The G roots of $z^G - p(z) = 0$ in $|z| \leqq 1$ may be obtained, and thus also $P_o, P_1, \ldots, P_{G-1}$ and $P'(1)$, so that we can obtain $\bar{d}$.

A similar expression may be found for the delay in the other stream, where vehicles arrive independently in a binomial process with parameter b. The red and green intervals now last R_1 and G_1 respectively. R_1 and G_1 may be different from G and R due to the period taken up by the amber light and the possibly artificial definition of G and G_1 to allow for the acceleration times of vehicles having to stop at the intersection. We have $R_1 + G_1 = R + G$. The traffic lights should normally be set to minimise the total delay to traffic at the intersection.

For an uncontrolled intersection, delay is caused only in the minor road. It is useful to compare the mean delay in this case with the total mean delay for an optimum setting of traffic lights. For low traffic densities it is to be expected that there will be less delay at an uncontrolled intersection than at one using traffic lights; for high traffic densities the delay will be greater. A comparison of the delays in the two cases would allow an estimate to be made as to when traffic lights should be installed.

5.7 Further problems

For a two-way major road some minor road vehicles may wish to turn into the first major road traffic stream, and so require gaps in only one stream. In this case the gap acceptance times are likely to be less than for the crossing of both streams; we can readily define these gap acceptance times for each case. The main difficulty in this problem is when we consider a vehicle wishing to cross both streams after a vehicle has turned into the first stream; here we do not have a starting point for determining whether a point in the second stream is red or green, but we may make some approximations. If the dependence between vehicles is not very great in the second major road stream we may reasonably take the point where a minor road vehicle reaches the head of the queue from behind a left turning vehicle as red or green according to the stationary probabilities independent of the initial condition. We can then find the service time distributions, and hence the stationary delay. In some cases it may be reasonable to suppose that the vehicles in the second major road stream are independent of each other; here there is a light density of traffic in the second stream, which is possible on a main road leading to or from a city centre.

When we consider more complex junctions, such as a two-way minor road and several lanes in the major road streams, we can obtain results only by making more restrictive assumptions about the interaction of the various streams of traffic. If we assume that there is a negligible degree of interference between the streams in a two-way minor road (e.g. virtually no right turning) we can extend our previous results to cover this case. When there is more than one lane in a major road stream we may approximate by lumping these lanes into a single one.

CHAPTER 6

DELAY AT AN INTERSECTION: CONTINUOUS TIME

6.1 Introduction

In the previous chapter we described some discrete time traffic models in which there was limited or no interaction between vehicles in the major road. In the continuous analogue, the vehicles in all streams travelled past a point in a homogeneous Poisson process. This assumption is reasonable only for low densities of traffic, and breaks down when the traffic densities become at all heavy; we would then expect an increasing degree of interaction between vehicles. The traffic density in the major road is likely to be greater than that in the minor, so that the description of traffic in the major road should be of a more general form than that in the minor road.

We consider two one-way one lane roads, where traffic in the minor road yields absolute right of way to that in the major road; some other cases are discussed in §6.4-§6.6. There are no traffic lights but there may be a stop sign on the minor road; it is possible to adapt our model to a case where a stop sign is either present or absent. Vehicles in the minor road wish to cross the major road; they must wait for sufficiently large gaps in the major road traffic before they are able to do so.

Our rules for minor road vehicles entering the intersection are of a form similar to, but more general than for the discrete model in Chapter 5. A minor road vehicle may arrive at the intersection to find that there are or there are not other minor road vehicles in a queue waiting to enter the intersection. In the former case the vehicle waits until all those ahead of it have departed. If there is then a further free period of at least β (the first gap acceptance time) in the major road stream, then the vehicle enters the intersection a time β after the previous minor road vehicle; β is a random variable with d.f. $b(x)$ $(0 \leqq b_1 \leqq x \leqq b_2)$, IST $\beta(\theta) = \int_{b_1}^{b_2} e^{-\theta x} db(x)$ $(\text{Rl } \theta \geqq 0)$ and finite mean $\bar{\beta} = -\beta'(0) < \infty$. If the period β is not free then a major road vehicle must arrive at the junction in this time. The minor road vehicle then waits until a first period of at least α (the second gap acceptance time) free of vehicles appears in the major road traffic, where α is a random variable with d.f. $a(x)$ $(0 \leqq a_1 \leqq x \leqq a_2)$, IST $\alpha(\theta) = \int_{a_1}^{a_2} e^{-\theta x} da(x)$ $(\text{Rl } \theta \geqq 0)$ and finite mean $\bar{\alpha} = -\alpha'(0) < \infty$; the minor road vehicle enters the intersection at time α after the commencement of this last free period.

When a minor road vehicle arrives at the intersection to find no other minor road vehicles waiting, we distinguish three different sets of rules for the entrance of the minor road vehicle into the intersection.

(i) If at the moment the vehicle arrives there is a free period of at least γ in the major road traffic, then the vehicle enters the intersection after a delay of time γ; γ is a random variable with d.f. $c(x)$

$(0 \leq c_1 \leq x \leq c_2)$, IST $\gamma(\theta) = \int_{c_1}^{c_2} e^{-\theta x} dc(x)$ $(\mathrm{Rl}\ \theta \geq 0)$ and finite mean $\bar{\gamma} = -\gamma'(0) < \infty$. If this period γ is not free, the vehicle waits until the next time there is a free time at at least α, defined as above, and enters the intersection α after the beginning of this last free period. (ii) We may simplify the problem by taking $\gamma = \alpha$. This seems reasonable when there is a stop sign in the minor road. (iii) We suppose a minor road vehicle has first and second gap acceptance times β and α defined as above. This vehicle cannot enter the intersection within a time β of the previous minor road vehicle or within α of a major road vehicle passing the intersection. Subject to these two restrictions a minor road vehicle is delayed by a time γ, as above, if there is a free time of at least γ in the major road traffic just when the minor road vehicle arrives at the junction. Otherwise it waits for the first free time of at least α. As a special case of this we may take $\gamma = 0$, and use this last model to compare our results with Tanner's (1961b, 1962), where the major road traffic has a Borel-Tanner distribution and there are fixed gap acceptance times.

The process of minor road vehicles' entering the intersection is reminiscent of the service of non-priority customers under a preemptive priority repeat identical policy; the priority busy period distribution is replaced by the distribution of a bunch of vehicles (described below) less than time c apart, where $c = \min(a_1, c_1)$ is the smaller of a_1, c_1, for rule (i) and $c = a_1$ for rules (ii) and (iii). We have assumed that there may be variation in gap acceptance times between vehicles, but that the second gap acceptance time α for a particular vehicle does not alter every time a decision entering the intersection is made. An alternative formulation, analogous to a preemptive priority repeat different policy, assumes that each time there is a decision on entering the intersection, it is made with the same distribution, independently of the particular vehicle making the decision. We suppose that a minor road vehicle has in the first instance a second gap acceptance time α_1; if this is larger than the first free time greater than c in the major road, the gap acceptance time is replaced by α_2, then α_3, α_4,... until there is a gap sufficiently large for the minor road vehicle to enter the intersection. All the α_i have the same d.f. $\Pr\{\alpha_i \leqq x\} = a(x)$ $(a_1 \leqq x \leqq a_2)$. The further problem of the impatience of drivers, who may accept smaller gaps after a long wait, is discussed in §6.4.

Let us now consider the formation of traffic in the major road; this description is similar to that in the next chapter. We would expect some interaction between vehicles or groups of vehicles which are relatively close together, and we wish to allow for this in a general manner. We suppose that the major road traffic passes the intersection in alternate bunches and gaps. The gaps have a negative exponential distribution with mean μ^{-1}, so that the time from the rear of one bunch to the front of the next follows this distribution. The time length of a bunch is a random variable with the d.f. $F_a(x)$ $(a \leqq x < \infty)$, LST $\xi_a(\theta) = \int_0^\infty e^{-\theta x}\, dF_a(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$ and finite mean $f = -\xi_a'(0) < \infty$.

A bunch is measured from the front of the first vehicle to time a behind the rear of the last vehicle of the bunch, so that the terminal time a of a bunch is actually free of vehicles. Thus the time from the rear of the last vehicle of a bunch to the commencement of the next bunch is a + X where X follows a negative exponential distribution. We may give a practical interpretation of the major road traffic model in the following way. A bunch consists of a group of one or more vehicles less than time a apart, the last vehicle in the bunch being that followed by a time-gap greater than a to the next vehicle in the stream. Vehicles less than time a apart may be dependent on each other; this dependence decreases as successive vehicles become further apart and ceases altogether when they are separated by a time-gap larger than a. It would be desirable to determine a entirely from the conditions in the major road; however, it must also satisfy certain conditions for the minor road traffic. For example, we must not allow a minor road vehicle to enter the intersection during a major road bunch.

The shortest time between two major road vehicles allowing a minor road vehicle to enter the intersection is $c = \min(a_1, c_1)$ for rule (i) and $c = a_1$ for rules (ii) and (iii). We might therefore define a bunch in the major road as a group of one or more vehicles less than c apart, and we have a = c. On two lane roads interaction appears to cease for a free period of about 6-8 seconds (Miller (1961)), although it is quite small in the last few seconds of this interval. From data collected by the School of Traffic Engineering at the University of

New South Wales it has been found at several junctions that (second) gap acceptance times at the end of bunches are of the order of 4-5 seconds on average but vary over a much wider range. Thus c may be as low as 2-3 seconds; major road vehicles are likely to travel at about 15-20 yards per second so that the amount of interference allowed in this formulation is rather less than may actually occur in practice. We have not made any estimate of the errors caused by reducing the interaction between major road vehicles in the above manner, but our model still seems more general than any previously formulated for similar problems.

We may describe the occurrence of bunches and gaps in terms of the minor road vehicles if we are considering a repeat identical type policy; such a formulation is inapplicable for a repeat different policy. It is intrinsically undesirable to introduce factors external to the major road stream in defining the bunches and gaps, but it does in some cases allow of a greater degree of interaction between the major road vehicles. From the point of view of a minor road vehicle with gap acceptance times α, β and γ waiting to enter the intersection, the major road traffic appears to consist of alternate gaps and blocks, where the blocks have a connotation differing slightly from the bunches we have already described. A y block ($y \geq c$) is defined as the time from the front of a bunch until the first appearance of a free time of at least y; the block ends a period y after the rear of the last vehicle of the block. If the vehicle arrives at an empty queue there needs to be a free time of at least γ before it can enter the intersection; if it joins a non-empty queue or has to wait while some major road vehicles pass the junction then there must be a free time of at least α between two major road vehicles. For rules (ii) and (iii) the free time must be at least α, and our present formulation is most readily applied to these two examples. Here we may consider a block to be made up of a group of vehicles less than α apart, and ending α after the last vehicle of the block. This is a formulation more general than that for the bunches discussed above, and allows a greater amount of interference between major road vehicles, e.g. α varies from 4-5 seconds and possibly over a wider range of perhaps 3-10 seconds. In case (i) we must consider vehicles less than γ apart (if $\gamma < \alpha$) if a minor road vehicle joins an empty queue.

Although the previous two formulations of major road traffic are physically different, they may be considered together mathematically. Given the first formulation, the time from the beginning of a bunch to the first free time of at least $y(\geqq c)$ is given by the same argument as for (4.2.1); the IST $\xi y(\theta) = \int_0^\infty e^{-\theta x}\, dF_y(x)$ (Rl $\theta \geqq 0$) of this distribution is

$$\xi_y(\theta) = \frac{(\mu+\theta)\xi_c(\theta)\; e^{-(\mu+\theta)(y-c)}}{\mu+\theta-\mu\xi_c(\theta)\{1-e^{-(\mu+\theta)(y-c)}\}}. \qquad (1.1)$$

A block concluding with a free time of $y(\geqq c)$ may be physically considered as arising from either of the two major road traffic formulations we have described.

We must finally describe the process of arrival of minor road vehicles at the intersection. We should like to have a general process similar to that for the major road traffic, but we are here restricted by the mathematical techniques at our disposal. Let us suppose that vehicles travel towards the intersection in bunches with gaps, these having a negative exponential distribution with mean ν^{-1}. We define a bunch as consisting of a group of vehicles less than $h(0 \leqq h \leqq \min(a_1, b_1, c_1))$ apart when measured front to front; a bunch length (in time) is a random variable with d.f. $h(x)$ $(h \leqq x < \infty)$, LST $\mathcal{H}(\theta) = \int_h^\infty e^{-\theta x} dh(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$ and mean $\bar{h} < \infty$. We choose $h \leqq \min(a_1, b_1, c_1)$ as we do not wish to allow the possibility of a vehicle's following behind another vehicle at a gap larger than the minimum gap acceptance time (although this may be generalised as in the second formulation of major road traffic).

Unfortunately the generality of this model makes its solution difficult, and it becomes necessary to provide some further simplifying assumptions. We know that the time between the front of two successive bunches is the length of the first bunch plus a time with a negative exponential distribution; we would like this total time to have a negative exponential distribution, so we must modify the formation of the bunches. We suppose that the number of vehicles in a bunch is a discrete random variable with probabilities $g_i (i = 1,2,\ldots)$, p.g.f. $g(s) = \sum_{i=1}^\infty g_i s^i$ $(|s| \leqq 1)$ and finite mean $\bar{g} = \sum_{i=1}^\infty i g_i < \infty$, and the actual mean (time) length of a bunch is $\bar{h}$. We assume that bunches of vehicles, of size i with probability g_i, arrive at the intersection in a time homogeneous Poisson process with parameter $\lambda = (\nu^{-1}+\bar{h})^{-1}$; thus bunches are considered mathematically to be of zero length, and the time between any two successive bunches now has a negative exponential distribution. We have chosen λ independently of the number of vehicles in a particular bunch, but if this were not so, we should not have an independent inter-arrival time distribution. A special case of this formulation is that for which vehicles arrive singly, i.e. $g_1 = 1$, in a Poisson process; this has been used by Tanner (1962) and others.

It has been shown by Miller (1961) that the Borel-Tanner distribution gives a good fit to some traffic models. If we assume the minor road traffic to be of this type we have

$$g_i = e^{-ir_1}(ir_1)^{i-1}(i!)^{-1} \qquad i = 1,2,\ldots,$$

where $r_1 = \ell_1\nu$, $\ell_1 (\leqq h)$ being the effective time length of a vehicle. A bunch has mean size $\bar{g} = (1-r_1)^{-1}$ $(r_1 < 1)$ and mean length $\bar{h} = \ell(1-r_1)^{-1}$, so that $\lambda^{-1} = \nu^{-1}+\ell(1-r_1)^{-1}$.

With rule (iii) for minor road vehicles entering the intersection our model is a generalisation of Tanner's (1962), who considers Poisson minor road traffic, a Borel-Tanner distribution in the major road and constant gap acceptance times. Tanner obtains the mean delay to minor road vehicles using a regeneration point method; our method is more general in that we obtain the transforms of the stationary waiting time distribution, and from it that of the delay, so that their higher moments may be found.

Herman and Weiss (1961) and Weiss and Maradudin (1962) have considered an alternative formulation of the major road traffic in terms of headway distributions. Here the time between the front of two successive vehicles is a random variable with a general distribution, so that the position of a vehicle is affected only by the nearest vehicle in front of it. In our model we do not make this restriction but we assume that the distribution of time from one vehicle to the next has a negative exponential tail. The delay to a single minor road vehicle arriving at random at the intersection is discussed in the last two papers; however, this formulation of major road traffic makes it very difficult to consider the delay for a flow of minor road vehicles. When a vehicle reaches the head of the queue, the distribution of time to the end of the major road gap is dependent on the previous minor road vehicles which have entered the intersection during the same gap, so that the service time distribution for vehicles joining a non-empty queue is not identical for all vehicles. Consequently little progress has been made using this model.

Our problem is a special case of that described in §3.10; the service time distribution depends on whether a vehicle joins an empty or a non-empty queue. Thus, we have to find the service time distributions for each of these two possibilities; if a bunch arrives at an empty queue all vehicles of the bunch except the first are regarded as joining a non-empty queue. Once we have found the service time distributions, we can substitute these in the formulae of §3.10 to obtain stationary results for our process.

6.2 The service time

We wish to find the service time distributions for minor road vehicles, namely the time taken from their reaching the head of the queue to entering the intersection. This is very similar to finding the completion time distributions for a preemptive priority repeat identical queueing system, although there are now additional difficulties for a vehicle joining an empty queue. Particular values α, β and γ of the gap acceptance times are considered, and appropriate integration yields the results for the general case.

When a minor road vehicle joins a non-empty queue there must still be a gap in progress on its reaching the head of the queue, otherwise the previous vehicle could not have entered the intersection. If there is a further gap of at least β its service time is β; otherwise service continues until the first free time of at least α appears in the major road traffic: Thus the d.f. $B_{\alpha\beta}(x)$ $(0 \leqq x < \infty)$ of this service time distribution is

$$B_{\alpha\beta}(x) = e^{-\mu\beta}H(x-\beta) + \int_{y=0}^{\min(x,\beta)} \mu e^{-\mu y}F_{\alpha}(x-y)dy,$$

where $H(x)$ is the unit step function. This has IST

$$\psi_{\alpha\beta}(\theta) = \int_{x=0}^{\infty} e^{-\theta x}\, dB_{\alpha\beta}(x) \qquad \mathrm{Rl}\ \theta \geqq 0$$

$$= e^{-(\mu+\theta)\beta} + \frac{\mu\xi_{\alpha}(\theta)}{\mu+\theta}\{1-e^{-(\mu+\theta)\beta}\}. \tag{2.1}$$

Integration over the distribution of α and β yields

$$\psi(\theta) = \int_{\beta=b_1}^{b_2}\int_{\alpha=a_1}^{a_2} \psi_{\alpha\beta}(\theta)da(\alpha)db(\beta) = \int_{x=0}^{\infty} e^{-\theta x}\, dB(x)$$

$$= \beta(\mu+\theta) + \frac{\mu\xi(\theta)}{\mu+\theta}\{1-\beta(\mu+\theta)\}, \tag{2.2}$$

where

$$\xi(\theta) = \int_{y=a_1}^{a_2} \xi_y(\theta)da(y). \tag{2.3}$$

Differentiation of (2.2) yields the first two moments of the service time distribution:

$$\overline{\psi} = -\psi'(0) = \mu^{-1}(1+\mu\rho)(1-\beta(\mu)) \tag{2.4}$$

$$\psi''(0) = \sigma^2(1-\beta(\mu)) + \frac{2(1+\mu\rho)}{\mu^2}\{1-\beta(\mu)+\mu\beta'(\mu)\} \tag{2.5}$$

where $\rho = -\xi'(0)$, $\sigma^2 = \xi''(0)$ and $\beta'(\mu) = \lim_{\theta\to\mu} \frac{d}{d\theta}\beta(\theta)$. Comparison with (4.2.14) and (4.2.15) shows that $\rho < \infty$ if $\alpha(-\mu) < \infty$ and $\sigma^2 < \infty$ if $\alpha(-2\mu) < \infty$.

We now turn to the service time distribution for the first vehicle of a bunch which arrives at the intersection when there are no other vehicle waiting. The problem is similar to those described in §3.8 and §4.2, where we integrated out over the time since the departure of the previous customer. We consider in turn each of the three rules for vehicles entering the intersection after arriving at an empty queue, in the first instance for particular values α β and γ of the gap acceptance times.

(i) A minor road vehicle m joining an empty queue is delayed by a time γ if there is a free time of at least γ in major road traffic at the moment of m's arrival. Otherwise m waits for the next free time of at least α in the major road traffic. We split up the major road traffic into parts, one part being the time from the end of one bunch to the end of the next. We can enumerate the possible points of arrival of any m during a single part, and the service time distribution is obtained by combining all the parts. During the first part we may have (a) m arrives before the gap is completed and the gap continues at least a further γ, (b) the gap lasts less than γ and m arrives at least c before the end of the bunch, (c) the gap lasts more than γ and m arrives between the gap time less γ, and gap plus block time less c, (d) m arrives in the last c of the bunch, and (e) m does not arrive during the part. Enumerating the possibilities we have that the LST $\zeta_{1\alpha\gamma}(\theta)$ of the (improper) service time distribution for the first part is

$$\zeta_{1\alpha\gamma}(\theta) = \frac{\lambda}{\lambda+\mu}\, e^{-(\mu+\theta)\gamma} + \left\{ \int_{y=o}^{\gamma} \int_{z=y}^{\infty} \int_{x=o}^{z-c} + \int_{y=\gamma}^{\infty} \int_{z=y}^{\infty} \int_{x=y-\gamma}^{z-c} \right\}$$

$$\Big[\mu\lambda e^{-\mu y} dy \; dF_c(z-y) e^{-\lambda x-\theta(z-x)} \; \{ e^{-(\mu+\theta)(\alpha-c)}$$

$$+\int_{u=o}^{\alpha-c}\int_{v=u}^{\infty} \mu e^{-\mu u}\, dF_{\alpha}(v-u)e^{-\theta v}\, du\}\Big]$$

$$+\int_{y=o}^{\infty}\int_{z=y}^{\infty}\int_{x=z-c}^{z} \lambda\mu e^{-\mu y}\, dydF_{c}(z-y)e^{-\lambda x-\theta(z-x)}$$

$$\left\{e^{-(\mu+\theta)(\gamma-z+x)} + \int_{u=o}^{\gamma-z+x}\int_{v=u}^{\infty} \mu e^{-\mu u-\theta v}\, dudF_{\alpha}(v-u)\right\}$$

$$= \frac{\lambda\mu\xi_{\alpha}(\theta)}{(\lambda-\theta)(\mu+\theta)} + \frac{\lambda e^{-(\mu+\theta)\gamma}}{\lambda+\mu}\left\{1-\frac{\mu\xi_{\alpha}(\theta)}{\mu+\theta}\right\}$$

$$- \frac{\lambda\mu\xi_{c}(\lambda)}{\mu+\lambda}\left[\frac{\mu\xi_{\alpha}(\theta)}{(\lambda-\theta)(\mu+\theta)} + \left(1-\frac{\mu\xi_{\alpha}(\theta)}{\mu+\theta}\right)\left\{\frac{e^{-(\mu+\theta)\alpha+(\lambda+\mu)c}}{\lambda-\theta}\right.\right.$$

$$\left.\left.- \frac{e^{-(\mu+\theta)\gamma}(e^{(\lambda+\mu)c}-1)}{\lambda+\mu}\right\}\right], \qquad (2.6)$$

where $\zeta_{1\alpha\gamma}(0) = 1 - \mu\xi_c(\lambda)/(\mu+\lambda)$, $g_o(c) = \mu\xi_c(\lambda)/(\mu+\lambda)$ being the probability that no minor road vehicles arrive during the first part. The service time distribution is identical for m arriving during later parts, so that the LST $\zeta_{\alpha\gamma}(\theta)$ of the service time distribution for any m with gap acceptance times γ and α, arriving at an empty queue is

$$\zeta_{\alpha\gamma}(\theta) = \zeta_{1\alpha\gamma}(\theta)\ \{1+g_o(c)+g_o(c)^2+\ldots\}$$

$$= \zeta_{1\alpha\gamma}(\theta)/(1-g_o(c))$$

$$= \frac{\lambda\mu\xi_{\alpha}(\theta)}{(\lambda-\theta)(\mu+\theta)} + \frac{\lambda e^{-(\mu+\theta)\gamma}}{\lambda+\mu}\left\{1-\frac{\mu\xi_{\alpha}(\theta)}{\mu+\theta}\right\}$$

$$+ \frac{\lambda g_o(c)e^{(\lambda+\mu)c}}{(1-g_o(c))} \left\{ 1 - \frac{\mu\xi_\alpha(\theta)}{\mu+\theta} \right\} \left\{ \frac{e^{-(\mu+\theta)\gamma}}{\mu+\lambda} - \frac{e^{-(\mu+\theta)\alpha}}{\lambda-\theta} \right\}. \tag{2.7}$$

When the gap acceptance times γ and α have d.f.'s $c(x)$ and $a(x)$ respectively with IST's $\gamma(\theta)$ and $\alpha(\theta)$ the IST $\zeta(\theta)$ of the service time distribution for any m joining an empty queue is

$$\zeta(\theta) = \int_{x=a_1}^{a_2} \int_{y=c_1}^{c_2} \zeta_{xy}(\theta)dc(y)da(x)$$

$$= \frac{\lambda\mu\xi(\theta)}{(\lambda-\theta)(\mu+\theta)} + \frac{\lambda\gamma(\mu+\theta)}{\lambda+\mu} \left\{ 1 - \frac{\mu\xi(\theta)}{\mu+\theta} \right\}$$

$$+ \frac{\lambda g_o(c)e^{(\lambda+\mu)c}}{(1-g_o(c))} \left\{ \frac{\gamma(\mu+\theta)}{\mu+\lambda} \left(1 - \frac{\mu\xi(\theta)}{\mu+\theta}\right) \right.$$

$$\left. - \frac{\alpha(\mu+\theta)}{\lambda-\theta} \left(1 - \frac{\mu A(\theta)}{\mu+\theta}\right) \right\}, \tag{2.8}$$

where

$$A(\theta) = \frac{1}{\alpha(\mu+\theta)} \int_{x=a_1}^{a_2} \xi_x(\theta)e^{-(\mu+\theta)x}da(x). \tag{2.9}$$

(ii) When $\gamma \equiv \alpha$ the IST $\zeta_\alpha(\theta)$ of the service time distribution for any m with second gap acceptance time α, arriving at an empty queue is given by (2.2) with $\gamma = \alpha$. In the general case we then obtain

$$\zeta(\theta) = \frac{\lambda\mu\xi(\theta)}{(\lambda-\theta)(\mu+\theta)} + \frac{\lambda\alpha(\mu+\theta)}{\mu+\lambda} \left\{ 1 - \frac{\mu A(\theta)}{\mu+\theta} \right\}$$

$$+ \frac{\lambda g_o(c)e^{(\mu+\lambda)c}\alpha(\mu+\theta)}{(1-g_o(c))} \left\{ \frac{1}{\mu+\lambda} - \frac{1}{\lambda-\theta} \right\} \left\{ 1 - \frac{\mu A(\theta)}{\mu+\theta} \right\}. \tag{2.10}$$

(iii) Unless a minor road vehicle m with gap acceptance times α, β and γ arriving at an empty queue enters the intersection during the same gap as the previous minor road vehicle the minimum free time required for m to enter the intersection is α. As in (i) we break up the major road traffic into parts. Instead of considering one part as the time of a gap and a bunch we could take a part as the time from the commencement of a gap (or end of an α block) to the end of the next α block; this gives us an alternative way of looking at the major road traffic in terms of the minor road vehicles.

The service time distribution depends on whether γ is greater than or less than β; we first consider the case $\gamma < \beta$, and also suppose $\gamma < \alpha$ which is quite reasonable. From the time of departure of the last minor road vehicle until the end of the next α block a minor road vehicle until the end of the next α block a minor road vehicle m will be confronted with one of the following possibilities: (a) m arrives before β-γ has elapsed and the gap continues at least β, (b) m arrives after β-γ has elapsed and the gap continues for at least this time and also a further γ, (c) the gap continues less than β and m arrives before the gap plus block time less γ has elapsed, (d) the gap continues at least β and m arrives in the interval between gap time less γ and gap plus block time less γ, (e) m arrives in the last γ of the gap plus block time, (f) no vehicles arrive during this first part.

From the first five of these cases we can find the LST ${}_1\zeta_{\alpha\beta\gamma}(\theta)$ of the (improper) service time distribution of a minor road vehicle m with gap acceptance times α, β and γ, which arrive during the first gap and block period since the departure of the previous minor road vehicle as

$$
\begin{aligned}
{}_1\zeta_{\alpha\beta\gamma}(\theta) &= \int_{x=0}^{\beta-\gamma} \lambda e^{-\lambda x-\mu\beta-\theta(\beta-x)}dx + \int_{y=\beta-\gamma}^{\infty} \lambda e^{-\lambda y-\mu(y+\gamma)-\theta\gamma}dy \\
&+ \left\{\int_{y=0}^{\beta}\int_{z=y}^{\infty}\int_{x=0}^{z-\gamma} + \int_{y=\beta}^{\infty}\int_{z=y}^{\infty}\int_{x=y-\gamma}^{z-\gamma}\right\}\left\{\lambda\mu e^{-\mu y}dydF_\alpha(z-y)\right. \\
&\quad \left. e^{-\lambda x-\theta(z-x)}dx\right\} \\
&+ \int_{y=0}^{\infty}\int_{z=y}^{\infty}\int_{x=z-\gamma}^{z} \lambda\mu e^{-\mu y}\; dydF_\alpha(z-y)e^{-\lambda x-\theta(z-x)} \\
&\quad \left\{e^{-(\mu+\theta)(\gamma-z+x)} + \int_{u=0}^{\gamma-z+x}\int_{v=u}^{\infty} \mu e^{-\mu u}due^{-\theta v}\; dF_\alpha(v-u)\right\} \\
&= \frac{\lambda e^{-(\mu+\theta)\beta}}{\lambda-\theta} + \left(\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda-\theta}\right)\; e^{-(\mu+\lambda)\beta+(\lambda-\theta)\gamma} + \frac{\lambda\mu\xi_\alpha(\theta)}{\lambda-\theta} \\
&\quad \left\{\frac{1-e^{-(\mu+\theta)\beta}}{\mu+\theta} + \frac{e^{-(\mu+\lambda)\beta+(\lambda-\theta)\gamma}}{\mu+\lambda}\right\} - \frac{\lambda\mu\xi_\alpha(\lambda)}{(\mu+\lambda)(\lambda-\theta)}. \qquad \gamma < \beta,
\end{aligned} \tag{2.11}
$$

where ${}_1\zeta_{\alpha\beta\gamma}(0) = 1 - \mu\xi_\alpha(\lambda)/(\mu+\lambda)$, $g_0(\alpha) = \mu\xi_\alpha(\lambda)/(\mu+\lambda)$ being the probability that no minor road vehicles arrive during a gap and an α block. When $\gamma \geqq \beta$ and $\alpha \geqq \gamma$ the IST ${}_2\zeta_{\alpha\beta\gamma}(\theta)$ of the service time distribution may be obtained in a similar manner as above:

$$
{}_2\zeta_{\alpha\beta\gamma}(\theta) = \frac{\lambda e^{-(\mu+\theta)\gamma}}{\lambda+\mu} + \frac{\lambda\mu\xi_\alpha(\theta)}{\lambda-\theta}\left\{\frac{1-e^{-(\mu+\theta)\gamma}}{\mu+\theta} + \frac{e^{-(\mu+\theta)\gamma}}{\mu+\lambda}\right\}
$$

$$
- \frac{\lambda\mu\xi_\alpha(\lambda)}{(\mu+\lambda)(\lambda-\theta)} \qquad \gamma \geqq \beta, \tag{2.12}
$$

with ${}_2\zeta_{\alpha\beta\gamma}(0) = 1 - g_o(\alpha)$. Except during the first part, the service time distribution for the minor road vehicle is of the form (2.12) as it is not possible to follow directly behind another minor road vehicle so that $0 = \beta \leqq \gamma$.

By simply enumerating the possible points of arrival we find that the IST $\zeta_{\alpha\beta\gamma}(\theta)$ of the service time distribution for a minor road vehicle, with gap acceptance times α, β and γ, joining an empty queue is

$$
\zeta_{\alpha\beta\gamma}(\theta) = {}_1\zeta_{\alpha\beta\gamma}(\theta) + g_o(\alpha)\,{}_2\zeta_{\alpha\beta\gamma}(\theta) + g_o(\alpha)^2\,{}_2\zeta_{\alpha\beta\gamma}(\theta) + \cdots
$$

$$
= \begin{cases} {}_1\zeta_{\alpha\beta\gamma}(\theta) + g_o(\alpha)\,{}_2\zeta_{\alpha\beta\gamma}(\theta)/(1-g_o(\alpha)) & \gamma < \beta \\ {}_2\zeta_{\alpha\beta\gamma}(\theta)/(1-g_o(\alpha)) & \gamma \geqq \beta, \end{cases} \tag{2.13}
$$

so that

$$
\zeta_{\alpha\beta\gamma}(\theta) = \frac{\lambda}{\lambda-\theta}\, e^{-(\mu+\theta)\beta} + \left(\frac{\lambda}{\mu+\lambda} - \frac{\lambda}{\lambda-\theta}\right) e^{-(\mu+\lambda)\beta+(\lambda-\theta)\gamma} + \frac{\mu\lambda\xi_\alpha(\theta)}{\lambda-\theta}
$$

$$
\left\{\frac{1-e^{-(\mu+\theta)\beta}}{\mu+\theta} + \frac{e^{-(\mu+\lambda)\beta+(\lambda-\theta)\gamma}}{\mu+\lambda}\right\} + \frac{g_o(\alpha)}{1-g_o(\alpha)}
$$

$$
\left\{\left(\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda-\theta}\right) e^{(\lambda-\theta)\gamma} + \frac{\lambda\mu\xi_\alpha(\theta)}{(\lambda-\theta)(\mu+\lambda)}\right\} \qquad \gamma < \beta \tag{2.14}
$$

and

$$\zeta_{\alpha\beta\gamma}(\theta) = \frac{\lambda}{\lambda+\mu}\, e^{-(\mu+\theta)\gamma} + \frac{\lambda\mu\xi_\alpha(\theta)}{\lambda-\theta}\left\{\frac{1-e^{-(\mu+\theta)\gamma}}{\mu+\theta} + \frac{e^{-(\mu+\theta)\gamma}}{\mu+\lambda}\right\}$$

$$+ \frac{\lambda g_o(\alpha)(\mu+\theta)}{1-g_o(\alpha)}\left\{1 - \frac{\mu\xi_\alpha(\theta)}{\mu+\theta}\right\}\left\{\frac{e^{-(\mu+\theta)\gamma}}{\lambda+\mu} - \frac{1}{\lambda-\theta}\right\} \quad \gamma \geqq \beta. \tag{2.15}$$

From (1.1) we have that

$$\frac{g_o(\alpha)}{1-g_o(\alpha)} = \frac{\mu\xi_c(\lambda)e^{-(\mu+\lambda)(\alpha-c)}}{\mu+\lambda-\mu\xi_c(\lambda)}$$

$$= \frac{g_o(c)e^{(\mu+\lambda)c}}{1-g_o(c)}\; e^{-(\mu+\lambda)\alpha} \tag{2.16}$$

The IST $\zeta(\theta)$ of the service time distribution for a vehicle joining an empty queue is given from (2.14) and (2.15) by integrating over the range of values of α, β and γ, so that

$$\zeta(\theta) = \int_{x=a_1}^{a_2}\int_{y=b_1}^{b_2}\int_{z=c_1}^{c_2} \zeta_{xyz}(\theta)\; dc(z)db(y)da(x). \tag{2.17}$$

In general this is difficult to integrate explicitly, but may be done in some special cases. It seems reasonable to suppose that γ is small compared with α so that we can neglect the parts of γ where $\gamma > a_1$. Also we might expect that γ is a function of β for if one gap acceptance time is large then the other should also be comparatively large, and vice-versa. We thus put $\gamma = f(\beta)$; if $\gamma = d_1\beta + d_2 < \beta$ then we obtain from (2.14) and (2.17) that

$$\zeta(\theta) = \frac{\lambda\beta(\mu+\theta)}{\lambda-\theta} + \left(\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda-\theta}\right) e^{(\lambda-\theta)d_2} \beta(\mu+\lambda-\lambda d_1-\theta d_1)$$

$$+ \frac{\mu\lambda\xi(\theta)}{\lambda-\theta} \left\{\frac{1-\beta(\mu+\theta)}{\mu+\theta} + \frac{e^{(\lambda-\theta)d_2}\beta(\mu+\lambda-\lambda d_1-\theta d_1)}{\mu+\lambda}\right\}$$

$$- \frac{\lambda(\mu+\theta)g_o(c)e^{(\mu+\lambda)c}}{(1-g_o(c))(\mu+\lambda)(\lambda+\theta)} \alpha(\mu+\lambda) \left\{e^{(\lambda-\theta)d_2}\beta(\lambda d_1-\theta d_1) - \frac{\mu E(\theta)}{\mu+\theta}\right\}, \tag{2.18}$$

where

$$E(\theta) = \frac{1}{\alpha(\mu+\lambda)} \int_{x=a_1}^{a_2} e^{-(\mu+\lambda)x}\xi_x(\theta)da(x). \tag{2.19}$$

When $\gamma \equiv 0$ then (2.18) reduces to

$$\zeta(\theta) = \frac{\lambda\beta(\mu+\theta)}{\lambda-\theta} + \left(\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda-\theta}\right) \beta(\mu+\lambda) + \frac{\lambda\mu\xi(\theta)}{\lambda-\theta} \left\{\frac{1-\beta(\mu+\theta)}{\lambda-\theta} + \frac{\beta(\mu+\lambda)}{\mu+\lambda}\right\}$$

$$- \frac{\lambda g_o(c)e^{(\mu+\lambda)c}(\mu+\theta)}{(1-g_o(c))(\mu+\lambda)(\lambda-\theta)} \alpha(\mu+\lambda) \left\{1 - \frac{\mu E(\theta)}{\mu+\theta}\right\}. \tag{2.20}$$

The first two moments of the service time distribution in this last case ($\gamma \equiv 0$) are found by differentiating (2.20):

$$\overline{\zeta} = -\zeta'(0) = -{}^1/_\lambda + (1+\mu\rho) \left\{\frac{1-\beta(\mu)}{\mu} + \frac{\beta(\mu+\lambda)}{\mu+\lambda}\right\} + \frac{D(1+\mu\overline{E})}{(\mu+\lambda)} \tag{2.21}$$

$$\zeta''(0) = \frac{\mu}{\mu+\lambda} \{\beta(\mu+\lambda)\sigma^2 + DE''(0)\}$$

$$+ \frac{2(1+\mu\rho)}{\mu} \left\{(\frac{1}{\mu} - \frac{1}{\lambda})(1-\beta(\mu)) + \beta'(\mu) - \frac{\mu\beta(\mu+\lambda)}{\lambda(\mu+\lambda)}\right\}$$

$$+ \frac{2D}{\lambda(\mu+\lambda)}(1 + \mu\overline{E}), \tag{2.22}$$

where

$$D = \frac{g_o(c)e^{(\mu+\lambda)c}}{1-g_o(c)}\alpha(\mu+\lambda) = \frac{\mu\xi_c(\lambda)e^{(\mu+\lambda)c}\alpha(\mu+\lambda)}{\mu+\lambda-\mu\xi_c(\lambda)}$$

$$\bar{E} = -E'(0) = -\int_{x=a_1}^{a_2} e^{-(\mu+\lambda)x}\ \xi_x'(0)da(x)/\alpha(\mu+\lambda)$$

$$E''(0) = \int_{x=a_1}^{a_2} e^{-(\mu+\lambda)x}\xi_x''(0)\ da(x)/\alpha(\mu+\lambda).$$

Before completing this section let us consider the service time distributions when the equivalent of a preemptive priority repeat different policy operates. By the same type of arguments as used in §4.2 we can show that the LST $\psi^*(\theta)$ of the service time distribution for a customer joining a non-empty queue is

$$\psi^*(\theta) = \beta(\mu+\theta) + \frac{\mu\xi^*(\theta)}{\mu+\theta}\{1 - \beta(\mu+\theta)\}, \tag{2.23}$$

where

$$\xi^*(\theta) = \frac{(\mu+\theta)\xi_c(\theta)\alpha(\mu+\theta)e^{(\mu+\theta)c}}{\mu+\theta-\mu\xi_c(\theta)\{1-\alpha(\mu+\theta)e^{(\mu+\theta)c}\}}\ . \tag{2.24}$$

When a vehicle arrives at an empty queue we take as an example rule (iii) with $\gamma = 0$; other cases may be dealt with in a similar manner. In place of $E(\theta)$ we have

$$E(\theta) = \frac{1}{\alpha(\mu+\lambda)}\int_{x=a_1}^{a_2}\int_{y=a_1}^{a_2} e^{-(\mu+\lambda)x}da(x)\ \xi^*_y(\theta)da(y)$$

$$= \xi^*(\theta).$$

We obtain the LST $\zeta^*(\theta)$ of the service time distribution of a minor road vehicle joining at empty queue as

$$\zeta^*(\theta) = \frac{\lambda\beta(\mu+\theta)}{\lambda-\theta} + \left(\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda-\theta}\right)\beta(\mu+\lambda) + \frac{\lambda\mu\xi^*(\theta)}{\lambda-\theta}\left\{1-\frac{\beta(\mu+\theta)}{\mu+\theta} + \frac{\beta(\mu+\lambda)}{\mu+\lambda}\right\}$$

$$- \frac{\lambda g_o(c)e^{(\mu+\lambda)c}(\mu+\theta)}{(1-g_o(c))(\mu+\lambda)(\lambda-\theta)}\left\{\alpha(\mu+\lambda) - \frac{\mu\xi^*(\theta)}{\mu+\theta}\right\}. \tag{2.25}$$

Finally, differentiation of (2.23) and (2.25) yields

$$\bar{\psi}^* = -\psi^{*\prime}(0) = \mu^{-1}(1+\mu\rho^*)(1-\beta(\mu))$$

$$\psi^{*\prime\prime}(0) = \sigma^{*2}(1-\beta(\mu)) + \frac{2(1+\mu\rho^*)}{\mu^2}\{1-\beta(\mu)+\mu\beta'(\mu)\} \tag{2.26}$$

$$\bar{\zeta}^* = -\zeta^{*\prime}(0) = -\frac{1}{\lambda} + (1+\mu\rho^*)\left\{\frac{1-\beta(\mu)}{\mu} + \frac{\beta(\mu+\lambda)}{\mu+\lambda}\right\} + \frac{D(1+\mu\rho^*)}{\mu+\lambda}$$

$$\zeta''(0) = \frac{\mu\sigma^2}{\mu+\lambda}\{\beta(\mu+\lambda)+D) + \frac{2D}{\lambda(\mu+\lambda)}(1+\mu\rho^*) + \frac{2(1+\mu\rho^*)}{\mu}$$

$$\left\{\left(\frac{1}{\mu} - \frac{1}{\lambda}\right)(1-\beta(\mu)) + \beta'(\mu) - \frac{\mu\beta(\mu+\lambda)}{\lambda(\mu+\lambda)}\right\}, \tag{2.27}$$

where

$$\rho^* = -\xi^{*\prime}(0) = \frac{(1+\mu\xi_c)e^{-\mu c}}{\mu\alpha(\mu)} - 1/\mu$$

$$\xi^{*\prime\prime}(0) = \frac{\xi_c''(0)e^{-c\mu}}{\alpha(\mu)} + \frac{2(1+\mu\bar{\xi}_c)^2e^{-2c\mu}}{\mu^2\alpha(\mu)^2} + \frac{2(1+\mu\bar{\xi}_c)e^{-c\mu}}{\mu^2\alpha(\mu)^2}$$

$$\left\{\mu\alpha'(\mu) - \alpha(\mu)(1+\mu\bar{\xi}_c-\mu c)\right\}. \tag{2.28}$$

6.3 The stationary distributions

We are now in a position to find the stationary distributions for the waiting time, delay, queue size and busy period. We substitute $\psi(\theta)$ from (2.3) and the required version of $\zeta(\theta)$ ((2.17), (2.18) or (2.20) in the formulae (3.10.2), (3.10.3) and (3.10.4) to obtain the LST $\Omega(\theta)$ of the stationary waiting time distribution and its mean, while the mean delay is given by (3.10.5). As an example let us take rule (iii) with $\gamma = 0$, so that $\zeta(\theta)$ is given by (2.20), and $\bar{g} = 1$ so that minor road vehicles arrive in a Poisson process; then the mean delay is

$$\begin{aligned}\bar{d} = [&\mu^3\sigma^2\{\beta(\mu+\lambda)+D\} + \lambda\mu^2 D(1+\mu\rho)(1-\beta(\mu))(E''(0)-\sigma^2)\\ &+ 2\lambda(1+\mu\rho)\{1-\beta(\mu)+\mu\beta'(\mu)\}\{(1+\mu\rho)\beta(\mu+\lambda) + (1+\mu\bar{E})D\}]\\ &[2\mu\{\mu-\lambda(1+\mu\rho)(1-\beta(\mu))\}\{(1+\mu\rho)\beta(\mu+\lambda)+(1+\mu\bar{E})D\}]^{-1}. \qquad (3.1)\end{aligned}$$

The stationary queue size distribution is given by (3.10.6) and the mean queue size by $\lambda\, \bar{d}\, \bar{g}$. In the special case where the gap acceptance times α and β are constant, i.e. $\alpha(\theta) = e^{-\alpha\theta}$, $\beta(\theta) = e^{-\beta\theta}$, (3.1) reduces to

$$\bar{d} = \frac{\mu^3\sigma^2 + 2\lambda(1+\mu\rho)^2\{1-\beta(\mu) + \mu\beta'(\mu)\}}{2\mu(1+\mu\rho)\{\mu-\lambda(1+\mu\rho)(1-\beta(\mu))\}} \qquad (3.2)$$

which is of the same form as Tanner's (1962) equation (17) for a Borel-Tanner distribution of major road traffic; we see that his result holds for general major road traffic conditions and constant gap acceptance times.

A sufficient condition that (3.1) should reduce to the form (3.2) is that $\overline{E} = \rho$ and $E''(0) = \sigma^2$. In general this is not true, although it does of course hold for constant gap acceptance times.

Under the alternative formulation of gap acceptance times, analogous to a preemptive priority repeat different policy, and with rule (iii) with $\gamma = 0$ we find that the mean delay $\overline{d}*$ is

$$\overline{d}* = \frac{\mu^3\sigma^{*2} + 2\lambda\overline{g}(1+\mu\rho^*)\{1-\beta(\mu)+\mu\beta'(\mu)\}}{2\mu(1+\mu\rho^*)\{\mu-\lambda\overline{g}(1+\mu\rho^*)(1-\beta(\mu))\}}$$

$$+ [\lambda\mu(1+\mu\rho^*)(1-\beta(\mu))\beta(\mu+\lambda)\{2(\overline{g}-1)(1-\lambda\overline{\psi}^*\overline{g})-\lambda\overline{\psi}^{*2}g''(1)\}$$

$$-2(\mu+\lambda)(\overline{g}-1)(1-\lambda\overline{\psi}^*\overline{g})\{1-\beta(\mu)+\mu\beta'(\mu)\}][2\mu^2(1-\lambda\overline{\psi}^*\overline{g})$$

$$\{\beta(\mu+\lambda) + D\}]^{-1}, \qquad (3.3)$$

where ρ^* and σ^{*2} are given by (2.28). When $g = 1$ this reduces to

$$\overline{d}* = \frac{\mu^3\sigma^{*2}+2\lambda(1+\mu\rho^*)\{1-\beta(\mu)+\mu\beta'(\mu)\}}{2\mu(1+\mu\rho^*)\{\mu-\lambda(1+\mu\rho^*)(1-\beta(\mu))\}}\ , \qquad (3.4)$$

which is also of the form (3.2) similar to Tanner's (1962) result. We see that this result may be directly generalised for a repeat different policy but not for a repeat identical policy.

The LST $\gamma(\theta)$ of the busy period distribution for minor road vehicles (a busy period is in progress whenever there is at least one vehicle in the queue at the intersection) is given by substitution of the appropriate values of $\psi(\theta)$ and $\zeta(\theta)$ (or $\psi^*(\theta)$ and $\zeta^*(\theta)$) in (3.10.7); this distribution is proper if and only if $\lambda\,\overline{\psi}\overline{g} < 1$ (or $\lambda\overline{\psi}^*\overline{g} < 1$). The mean length of a busy period is

$$\bar{\gamma} = -\gamma'(0) = \frac{\bar{g}\,\bar{\zeta}}{1-\lambda\bar{g}\,\bar{\psi}}\ , \tag{3.5}$$

and the mean number of vehicles in a busy period is

$$\bar{r} = \frac{(1-\lambda\bar{\psi}+\lambda\ \bar{\zeta})}{1-\lambda\bar{\psi}\ \bar{g}}\ . \tag{3.6}$$

6.4 Driver impatience

In the next three sections we describe some extensions of our traffic model. If a vehicle has been waiting to enter the intersection for a long time it is likely that the driver will become impatient, and tend to accept smaller gaps in the major road traffic.

There are several ways of taking account of driver impatience. One is to suppose that the rate of service, i.e. the rate of decrease of waiting time, increases with the waiting time in the queue; this has been partially dealt with by Miller and Gaver (1961). However, this is difficult to apply, and also needs to be extended to the case where a customer joining an empty queue may have a service time distribution different from one joining a non-empty queue. We might suppose that a vehicle joining a long queue has shorter gap acceptance times than if it joined a short queue, so that the impatience is transferred back along the queue; however, this concept is also very difficult to apply.

One useful method of measuring impatience is to suppose that for a minor road vehicle the second gap acceptance time reduces every time a gap appears in the major road traffic until finally a gap is accepted. We suppose that in the first instance a minor road vehicle has a second gap acceptance time α_1; if the first gap is less than α_1-c then the gap acceptance time is replaced by $\alpha_2 (\geqq c)$, then α_3, α_4,... until a gap is finally accepted. The LST $\xi(\theta|\alpha_1, \alpha_2,\ldots)$ of the distribution of the time from the commencement of a bunch to the first time an acceptable gap appears in the major road traffic (c.f. (4.2.16)) is given by

$$\xi(\theta|\alpha_1,\alpha_2,\ldots) = \int_{z_1=0}^{\infty} e^{-\theta z_1} dF_c(z_1) \Big\{ e^{-(\mu+\theta)(\alpha_1-c)} + \int_{y_2=0}^{\alpha_1-c} \int_{z_2=y_2}^{\infty} \mu e^{-\mu y_2-\theta z_2} \, dF_c(z_2-y_2)$$

$$\Big\{ e^{-(\mu+\theta)(\alpha_2-c)} + \int_{y_3=0}^{\infty} \int_{z_3=y_3}^{\infty} \mu e^{-\mu y_3-\theta z_3} \, dF_c(z_3-y_3)$$

$$\Big\{ e^{-(\mu+\theta)(\alpha_3-c)} + \ldots \Big\}\Big\}\Big\}$$

$$= \sum_{n=1}^{\infty} \frac{\mu^{n-1}\xi_c(\theta)^n e^{-(\mu+\theta)(\alpha_n-c)}}{(\mu+\theta)^{n-1}} \prod_{j=1}^{n-1} (1-e^{-(\mu+\theta)(\alpha_j-c)}), \tag{4.1}$$

where we define $\prod_{j=1}^{0} = 1$. For these fixed values of $\{\alpha_i\}$ the LST's $\psi(\theta|\alpha_1, \alpha_2, \ldots)$ and $\zeta(\theta|\alpha_1, \alpha_2,\ldots)$ of the service time distributions may readily be found as before. We now wish to integrate $\xi(\theta|\alpha_1, \alpha_2,\ldots)$ over the distributions of the $\{\alpha_i\}$ $(i = 1, 2, \ldots)$. When $\alpha_i = \alpha$ and

$\Pr\{\alpha_i \leq x\} = a(x)$ then $\xi(\theta) = \int_{\alpha=a_1}^{a_2} \xi(\theta|\alpha_1, \alpha_2, \ldots)da(\alpha)$ is given by (1.1) and (2.3), and when the α_i are independently and identically distributed $\xi(\theta)$ is given by (2.24). If we take account of driver impatience we would expect that α_{n+1} is dependent on α_n in some manner and thus on α_1; we put $\alpha_j = f_j\alpha \equiv f_j\alpha_1$, where $0 < f \leq f_j \leq 1$, $f\alpha \geq c$ (in practice the f_j would be non-increasing functions of j) and $\Pr\{\alpha \leq x\} = a(x)(a_1 \leq x \leq a_2,\ fa_1 \geq c)$. From (4.1) we obtain

$$\xi(\theta) = \int_{\alpha=a_1}^{a_2} \xi(\theta|\alpha_1, \alpha_2, \ldots)da(\alpha)$$

$$= \sum_{n=1}^{\infty} \left(\frac{\mu}{\mu+\theta}\right)^{n-1} \xi_c(\theta)^n \int_{\alpha=0}^{\infty} e^{-(\mu+\theta)(\alpha f_n - c)} \prod_{j=1}^{n-1} (1-e^{-(\mu+\theta)(\alpha f_j - c)})\, da(\alpha) \tag{4.2}$$

which is not in a very useful form. We consider one simple case where we put $f_j = 1$ for $j \leq N$ are $f_j = f(0 < f \leq 1)$ for $j > N$ so that

$$\alpha_j = \begin{cases} \alpha & j \leq N \\ f\alpha & j > N. \end{cases}$$

In this case

$$\xi(\theta) = \xi_c(\theta) \int_{x=a_1}^{a_2} \frac{e^{-(\mu+\theta)(x-c)}\left\{1 - \left(\frac{\mu\xi_c(\theta)}{\mu+\theta}(1-e^{-(\mu+\theta)(x-c)})\right)^N\right\} da(x)}{1 - \frac{\mu\xi_c(\theta)}{\mu+\theta}(1-e^{-(\mu+\theta)(x-c)})}$$

$$+ \left(\frac{\mu}{\mu+\theta}\right)^{N-1} \xi_c(\theta)^N \int_{x=a_1}^{a_2} \frac{e^{-(\mu+\theta)(fx-c)}(1-e^{-(\mu+\theta)(x-c)})^{N-1} da(x)}{1 - \frac{\mu\xi_c(\theta)}{\mu+\theta}(1-e^{-(\mu+\theta)(fx-c)})} \tag{4.3}$$

Moments of this distribution, and hence of the service time distributions, may be found explicitly by differentiating (4.3), although this becomes very complicated if N is large (e.g. $N \geqq 5$).

6.5 A two way major road

Consider a two way major road with one lane in each direction. The minor road vehicles have to pass through both streams of traffic in order to cross the intersection; we suppose that the gap acceptance times for minor road vehicles are as above, except that now free times must exist in both streams before a vehicle can accept a gap. We assume that the traffic in th first major road stream encountered by minor road vehicles consists of alternate bunches and gaps exactly as for a one way major road. In the second major road stream we also assume that the traffic consists of alternate bunches and gaps; each gap has a negative exponential distribution with mean $\nu^{-1} (< \infty)$, and each bunch is made up of a group of vehicles not greater than $c (\leqq \min (a_1, c_1)$ for rule (i) and $\leqq a_1$ for rules (ii) and (iii)) apart. A bunch is measured from the front of the first vehicle to c behind the rear of the last vehicle, and its length is a random variable with d.f. $G_c(x)$ $(0 \leqq x < \infty)$, LST $X_c(\theta)$ $= \int_{x=0}^{\infty} e^{-\theta x} dG_c(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$ and mean $\bar{X}_c = -X'_c(0) < \infty$. The two major road streams flow independently of each other.

From the point of view of minor road vehicles the major road traffic may be condensed into a single lane of traffic consisting of alternate gaps and bunches; there is a gap when there are gaps in both major road streams and a bunch everywhere else. As the two gap distributions are independent negative exponentials the gap distribution of the combined process is negative exponential with mean $\eta^{-1} = (\mu+\nu)^{-1}$. If we could find the d.f $H_c(x)(0 \leq x < \infty)$, with LST $\Phi_c(\theta) = \int_0^\infty e^{-\theta x} dH_c(x)$ (Rl $\theta \geq 0$), of the combined bunch distribution then we could derive the service time and hence the stationary delay distribution for minor road vehicles. However, this is a difficult problem and has previously been solved only in the special case of homogeneous Poisson traffic in both streams; in this case all vehicles are of zero length and independent of each other, $\Phi_o(\theta) = 1$ and an α block has the LST

$$\Phi_\alpha(\theta) = \frac{(\eta+\theta)e^{-(\eta+\theta)\alpha}}{\theta+\eta e^{-(\eta+\theta)\alpha}} . \tag{5.1}$$

In order to obtain the mean delay to minor road vehicles it is sufficient to determine the first two moments, $\overline{\Phi}_c = -\Phi'_c(0)$ and $\Phi''_c(0)$, of the combined bunch distribution. We can find the mean $\overline{\Phi}_c$ by a simple averaging argument, which unfortunately cannot be applied for higher moments. The probabilities that at a random instant of time there are gaps in progress in the first and second major road streams are

$$\Pr\{\text{gap in first stream}\} = \frac{\mu^{-1}}{\mu^{-1}+\xi_c} = \frac{1}{1+\mu\overline{\xi}_c}$$

$$\Pr\{\text{gap in second stream}\} = \frac{\nu^{-1}}{\nu^{-1}+\overline{X}_c} = \frac{1}{1+\nu\overline{X}_c} .$$

As the two streams are independent the probability of a joint gap is

$$P_{gg} = \Pr\{\text{gap in both streams}\} = \frac{1}{(1+\mu\bar{\xi}_c)(1+\nu\bar{X}_c)} ,$$

and the probability that at a random instant there is a bunch in the combined streams is $1 - P_{gg}$. The mean length of a gap is $(\mu+\nu)^{-1}$ so that

$$\bar{\Phi}_c = \frac{1-P_{gg}}{P_{gg}} \cdot \frac{1}{\mu+\nu} = \frac{(1+\mu\bar{\xi}_c)(1+\nu\bar{X}_c)-1}{\mu+\nu} . \tag{5.2}$$

An alternative method is necessary to obtain $\Phi_c''(0)$. Each major road stream and the combined stream consists of an alternative sequence of gaps and bunches, for which the end or beginning of bunches are regeneration points of the particular process. In each stream the times between successive regeneration points are independently and identically distributed random variables X_n, Y_n and Z_n $(n = 1,2,\ldots)$, where X_n, Y_n and Z_n are the sum of a gap and a bunch time in the first, second and combined streams respectively. The IST's of these distributions are given by

$$\alpha^*(\theta) = \int_{x=0}^{\infty} e^{-\theta x}\, d\,\Pr\{X_n \leqq x\} = \frac{\mu\xi_c(\theta)}{\mu+\theta}$$

$$\beta^*(\theta) = \int_{x=0}^{\infty} e^{-\theta x}\, d\,\Pr\{Y_n \leqq x\} = \frac{\nu X_c(\theta)}{\nu+\theta}$$

$$\gamma^*(\theta) = \int_{x=0}^{\infty} e^{-\theta x}\, d\,\Pr\{Z_n \leqq x\} = \frac{(\mu+\nu)}{\mu+\nu+\theta}\,\Phi_c(\theta). \tag{5.3}$$

Over a long period L we have approximately that

$$L = X_1+\ldots+X_n \cong Y_1+\ldots+Y_\ell \cong Z_1+\ldots+Z_m \qquad (n \geqq m, \ell \geqq m).$$

We write p_n^m as the probability that given n regeneration points in the first stream in time L there are $m(\leqq n)$ regeneration points in the combined stream; this probability is also a function of the second minor road stream. Over the period L we can approximately equate the LST's of the distribution in the various streams, so that

$$(\alpha^*(\theta))^n \cong \sum_{m=1}^{n} p_n^m(\gamma^*(\theta))^m = p_n(\gamma^*(\theta)), \tag{5.4}$$

where $p_n(\theta) = \sum_{m=1}^{n} p_n^m\theta^m$ $(|\theta|\leqq 1)$. If we could find p_n^m (or $\lim_{n\to\infty} p_n^m$) we could obtain $\gamma^*(\theta)$, and hence $\Phi_c(\theta)$, from (5.4). However, it has not been found possible to determine p_n^m, although some approximations may be made. If m n tends to a limit such that $\lim_{n\to\infty} p_n^{qn} = 1$, $\lim_{n\to\infty} p_n^m = 0$, $m \neq qn$ $(0 < q < 1)$ then

$$\alpha^*(\theta) = \gamma^*(\theta)^q, \tag{5.5}$$

and using (5.2) we have

$$q = \frac{\mu+\nu}{\mu(1+\nu\bar{X}_c)} .$$

The approximations for $\Phi_c''(0)$ in this case are not always very useful, even in a few simple cases which have been computed (c.f.(5.1)). If we suppose that as $n \to \infty$ p_n^m tends to a more general distribution, such as a normal or a Gamma(for which we have to estimate a second parameter), then some useful results may be obtained, although these have not been checked against practical data.

We now consider our problem in terms of renewal and queueing theory. The gaps in a stream are equivalent to the idle periods and the bunches are similar (but not identical) to the busy period in a single server queueing system with Poisson arrivals. We suppose that Z_t^i $(i = 1,2,3)$ is the time to the end of the current bunch and $Z_t^i = 0$ during a gap; we define

$$W_i(x,y,t) = \Pr\{Z_t^i \leqq y \,|\, Z_0^i = x\} \qquad (i = 1,2)$$

$$W(x,y,z,t) = \Pr\{Z_t^3 \leqq z \,|\, Z_0^1 = x,\ Z_0^2 = y\}$$

as the d.f. of the time to the end of a bunch at time t for first and second major road streams, and the combined stream respectively. The probability that there is a gap in the i-th $(i = 1,2,3)$ stream at time t is $W_i(x,0,t)$ $(i = 1,2)$ and $W(x,y,0,t)$; we have by definition that $W_i(x,0,t) = W_i(0,0,t-x)$ $(t \geqq x,\ i = 1,2)$ $(W_i(x,0,t) = 0$ for $t < x)$. We also define

$$V(x,y,t) = \Pr\{Z_\tau^3 = 0 \text{ for some } \tau < t;\ Z_v^3 > 0,\ 0 < v < \tau \,|\, Z_0^1 = x, Z_0^2 = y\}$$

as the probability that a bunch in the combined stream ends by time t given that at time zero, x and y remain of bunches in the first and second streams respectively. We have by analogy with emptiness in a queueing system that

$$W(x, y,0,t) = \int_{\tau=0}^{t} W(0,0,0,t-\tau)d_{\tau}V(x,y,\tau). \tag{5.6}$$

We define the LST's

$$\Xi_i(x,\theta) = \int_{t=0}^{\infty} e^{-\theta t} \, d_t W_i(x,0,t) \qquad \text{Rl } \theta \geqq 0 \qquad i = 1,2$$

$$\Xi(x,y,\theta) = \int_{t=0}^{\infty} e^{-\theta t} \, d_t W(x,y,0,t) \qquad \text{Rl } \theta \geqq 0$$

$$\Theta(x,y,\theta) = \int_{t=0}^{\infty} e^{-\theta t} \, d_t V(x,y,t) \qquad \text{Rl } \theta \geqq 0.$$

It may be shown from equations similar to (5.6) for the two major road streams in isolation, or by renewal theory, that

$$\Xi_1(x,\theta) = \frac{\theta e^{-\theta x}}{\mu+\theta-\mu\xi_c(\theta)}$$

$$\Xi_2(y,\theta) = \frac{\theta e^{-\theta y}}{\nu+\theta-\nu\chi_c(\theta)} , \tag{5.7}$$

and by inverting these that

$$W_1(x,0,t) = \sum_{n=0}^{\infty} \int_{u=0}^{t-x} e^{-\mu(t-x-u)} \frac{(\mu(t-x-u))^n}{n!} \, dF_c^{n*}(u).$$

$$W_2(y,0,t) = \sum_{n=0}^{\infty} \int_{u=0}^{t-y} e^{-\nu(t-y-u)} \frac{(\nu(t-y-u))^n}{n!} \, dG_c^{n*}(u). \tag{5.8}$$

Taking transforms in (5.6) yields

$$\Theta(x,y,\theta) = \frac{\Xi(x,y,\theta)}{\Xi(0,0,\theta)} . \tag{5.9}$$

As the two major road streams are independent and there is a gap in the combined stream only if there is a gap in both streams it follows that

$$W(x,y,0,t) = W_1(x,0,t)W_2(y,0,t). \tag{5.10}$$

As a bunch in the combined stream may commence with a bunch in either single major road stream we have from (5.9)and (5.10) that

$$\begin{aligned}(\mu+\nu)\Phi_c(\theta) &= \mu\int_{x=o}^{\infty}\Theta(x,0,\theta)dF_c(x)+\nu\int_{y=o}^{\infty}\Theta(0,y,\theta)dG_c(y)\\ &=\Big\{\mu\int_{x=o}^{\infty}\int_{t=x}^{\infty}e^{-\theta t}\,d_t[W_1(0,0,t-x)W_2(0,0,t)]dF_c(x)\\ &\quad+\nu\int_{y=o}^{\infty}\int_{t=y}^{\infty}e^{-\theta t}\,d_t[W_1(0,0,t)W_2(0,0,t-y)]dG_c(y)\Big\}\\ &\quad\Big\{\int_{t=o}^{\infty}e^{-\theta t}\,d_t[W_1(0,0,t)W_2(0,0,t)]\Big\}^{-1}.\end{aligned} \tag{5.11}$$

This gives a formal solution to our problem; however, I have been unable to evaluate the integrals in (5.11), and have not found any useful expression for $\Phi_c''(0)$. If the expressions on the right hand side of (5.8) converge fairly rapidly a reasonable approximation for $\Phi_c''(0)$ may possibly be obtained by numerical methods.

If some minor road vehicles wish to turn left into the first major road stream, then for a vehicle wishing to cross both streams after a vehicle has turned left we do not know if there is a gap in progress in the second major road stream. Some further simplifications or approximations must thus be made, as was necessary in §5.6

6.6 Some other problems

The methods used in the previous sections may be extended to some more complex road situations. If the major road has more than one lane, its traffic may still be viewed from the minor road as an alternate sequence of blocks and gaps, although the block distribution may be more complicated than for a single lane. However, when there is more than one lane of traffic in the major road, it becomes possible to pass slow vehicles and the traffic may not tend to bunch up so heavily.

If some minor road vehicles wish to turn left and others wish to cross a one lane major road then these two groups of vehicles may require gap acceptance times with different distributions. Knowing the probability of a vehicle turning left and the gap acceptance times for this manoeuvre, we can readily extend our formulae to cover this situation. If the major road has more than one lane and vehicles turning left require a gap in only the left lane of the major road, then the problem is complicated by our not being able to consider the major road as condensed into one lane (as with a two way major road); also there may be a block in the second major road lane immediately after a vehicle has turned left into the major road. Some further assumptions have to be made in this case, but it is possible to use approximations.

It would be useful if we could find the total mean delay to vehicles in both major and minor roads if the intersection were controlled by fixed interval traffic lights, given the previous formulation of the major and minor road traffic streams. We could then compare the mean delay for the two systems and, subject to some other economic considerations, an estimate could be made as to when traffic lights should be installed at the intersection. However, the problem is much more complex when there are traffic lights, and it has not been found possible to obtain an expression for the delay for the general arrival time distributions considered in the previous sections. An estimate of delay may be made by Monte Carlo methods; several values of the length of the 'red' and 'green' intervals would then need to be taken to obtain the mean delay for an optimum setting of the lights.

For one stream the intersection is alternately clear and blocked for fixed intervals. This is similar to a preemptive repeat identical queueing system where priority customers have a constant inter-arrival time, thus being the length of a red plus a green interval, and a constant service time which is the length of a red interval. For an uncontrolled intersection the time between the end of one bunch and the commencement of the next has a negative exponential distribution. The latter case is easier to handle as all vehicles which have to queue up before entering the intersection have the same service time distribution; this is not true when there are traffic lights, as the service time of a vehicle joining a non-empty queue depends on the number of vehicles which have been served in the green period during which the vehicle reaches the head of the queue.

If we consider the end of green periods as regeneration points, it is possible to write down a set of difference equations for the distribution of the number of vehicles waiting in a stream. In general, however, it has not been found possible to solve these equations, except for some simple cases; under quite restrictive conditions the problem may be reduced to that of an infinite dam in discrete time with maximum release $m(>1)$ (c.f. §1.2 and §5.6).

6.7 Numerical results

A series of calculations have been carried out on the Australian National University's I.B.M. 1620 computer. The main purpose of these has been to compare results for variable gap acceptance times with those of Tanner (1961b) for constant gap acceptance times. We assume that (a) the one way major road traffic forms a Borel-Tanner distribution with $c = 0$ and the distribution of bunches has the LST

$$\xi_c(\theta) \equiv \xi_o(\theta) = e^{(\xi_o(\theta)-1)r-\theta\ell}, \tag{7.1}$$

where ℓ is the effective length of a major road vehicle and $r = \mu\ell < 1$; (b) rule (iii) holds for minor road vehicles entering the intersection with $\gamma = 0$ and $\bar{g} = 1$; and (c) the gap acceptance times have truncated Erlangian distributions with

$$db(x) = \frac{e^{-(x-b_1)/b} \cdot (x-b_1)^{m-1}\, dx}{b^m(m-1)!\left\{1 - \sum_{r=o}^{m-1} e^{-(b_2-b_1)/b} \dfrac{(b_2-b_1)^r}{r!b^r}\right\}} \qquad b_1 \leqq x \leqq b_2 \tag{7.2}$$

$$da(x) = \frac{e^{-(x-a_1)/a} \cdot (x-a_1)^{m-1}\, dx}{a^m(m-1)!\left\{1 - \sum_{r=o}^{m-1} e^{-(a_2-a_1)/a} \dfrac{(a_2-a_1)^r}{r!}\right\}} \qquad a_1 \leqq x \leqq a_2\,, \tag{7.3}$$

where b and a (different from that in §6.1) are so chosen as to make the mean gap acceptance times $G = \int_{a_1}^{a_2} x\, da(x)$ and $H = \int_{b_1}^{b_2} x\, db(x)$ close to the α and β of Tanner (1961b). Blunden, Clissold and Fisher (1962) have shown that in several practical examples the Erlangian distribution gives a good fit for the gap acceptance times, so that (7.2) and (7.3) should fit many practical cases. Initially we take $a_2 = b_2 = \infty$; this considerably simplifies the calculations and readily enables us to set $G = \alpha$ and $H = \beta$; in this case

$$G = ma + a_1 = \alpha$$

$$H = mb + b_1 = \beta$$

so that we chose a and b as

$$a = (\alpha - a_1)/m, \; b = (\beta - b_1)/m. \tag{7.4}$$

The first two moments of the bunches of the Borel-Tanner distribution (7.1) are

$$\bar{\xi}_o = \ell(1 - r)^{-1}$$

$$\xi_o''(0) = \ell^2(1 - r)^{-3} \tag{7.5}$$

From (7.5), (2.3) and (1.1) we obtain

$$\rho = \frac{\alpha(-\mu)e^{-\mu\ell}}{\mu(1-r)} - 1/\mu$$

$$\sigma^2 = \frac{2e^{-\mu\ell}}{\mu^2(1-r)^2} \left[e^{-\mu\ell}\alpha(-2\mu) + \mu(1-r)\alpha'(-\mu) + \left\{(1-r) - \frac{r^2(1-2r)}{2(1-r)}\right\}\alpha(-\mu)\right] . \qquad (7.6)$$

A necessary and sufficient condition for the existence of a stationary distribution is that $\lambda\,\bar{\psi} = \mu^{-1}\alpha(-\mu)e^{-\mu\ell}(1-\beta(\mu))/(1-r) < 1$. When $G = \alpha$ and $H = \beta$ saturation point is reached sooner for variable gap acceptance times than for constant gap acceptance times as $\alpha(-\mu) \geqq e^{\mu\alpha}$.

When $a_2 = b_2 = \infty$ the mean delay $\bar{d}$ given by (3.1) has been computed for $(m,a_1,b_1) = (1,0,0), (3,0,0), (1,1,1), (3,1,1), (7,0,0), (7,1,1), (7,3,2)$, $\ell = 0,1,2$, $G = 4,6,8,10$, $H = 1,3,5,7$, $\lambda = 0.02\ (0.04)\ 0.50$ and $\mu = 0.10\ (0.10)\ 1.00$ or until $\lambda\,\bar{\psi} < 1$ is no longer satisfied. Only a selection of the results are presented here, but these are sufficient to indicate the trends involved; Tanner's results are included for comparison.

As one would expect intuitively, the mean delay is always greater for variable gap acceptance times; a vehicle with a large gap acceptance time may delay not only itself but several other vehicles as well, whereas one small gap acceptance time is unlikely to save other vehicles much time. We wish to find how much larger the delay is for variable gap acceptance times. When $m = 1$, so that gap acceptance times have a truncated negative exponential distribution, the mean delay is considerably larger for variable gap acceptance times, even for quite low densities of traffic. For the more realistic case of $m = 7$ with $(a_1, b_1) = (1,1)$ and $(3,2)$ the mean delay is only slightly greater for low densities, but is considerably larger for high densities, particularly so near saturation level. For medium values of the mean delay, e.g 10-25 seconds, a decision on whether the mean delay is significantly larger for variable gap acceptance times depends on the importance attached to marginal extra delays (e.g. 5-20%).

If we choose $a_2 < \infty$, $b_2 < \infty$ then

$$G = ma + a_1 - \frac{(a_2-a_1)^m e^{-(a_2-a_1)/a}}{a^m(m-1)!\left\{1 - \sum_{r=o}^{m-1} e^{-(a_2-a_1)/a} \frac{(a_2-a_1)^r}{r!a^r}\right\}}$$

$$H = mb + b_1 - \frac{(b_2-b_1)^m e^{-(b_2-b_1)/b}}{b^m(m-1)!\left\{1 - \sum_{r=o}^{m-1} e^{-(b_2-b_1)/b} \frac{(b_2-b_1)^r}{r!b^r}\right\}}$$

and it becomes more difficult to choose a and b so that $G = \alpha$, $H = \beta$; however, the choice of a and b as in (7.4) gives reasonable approximations for G and H to α and β. It has been found for $m = 3$ and $m = 7$ that making a_2 and b_2 finite produces very little difference in the mean delay as compared with $a_2 = b_2 = \infty$, while it involves a much lengthier computational procedure. The mean delay $\bar{d}$ has been computed for

$(m, a_1, a_2, b_1, b_2) = (3,1,12,1,10)$, $(7,1,12,1,10)$, $(7,3,20,2,16)$ and $(7,3,12,2,10)$; for the last two cases the mean delay agrees with that for $a_2 = b_2 = \infty$ to four significant figures and it has consequently not been thought worthwhile to include these calculations.

In conclusion, it would seem that for low traffic densities it is reasonable to take fixed gap acceptance times; this becomes less true as the traffic density increases, and is quite unreasonable for high densities.

We have formulated the problem in more general terms than that for which we have computed results, and it may be found that major road traffic distributions other than the Borel-Tanner, and other gap acceptance time distributions, are more applicable in different circumstances.

					m	3	7	1	3	7	7
					a_1	0	0	1	1	1	3
					a_2	0	0	1	1	1	2
λ	ℓ	μ	H	G	$\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
.02	0	.10	3	4	1.11	2.19	1.48	4.55	1.65	1.31	1.14
.06					1.63	2.98	2.09	5.84	2.30	1.87	1.65
.10					2.40	4.15	3.01	7.79	3.27	2.72	2.43
.14					3.68	6.09	4.54	11.06	4.89	4.14	3.74
.18					6.28	9.93	7.58	17.71	8.17	7.00	6.36
.22					14.26	21.13	16.68	38.54	18.09	15.69	14.37
.26					31.11	540.76	308.37				
.02	1	.10	3	4	1.15	2.24	1.52	4.61	1.69	1.34	1.17
.06					1.67	3.04	2.14	5.92	2.35	1.92	1.70
.10					2.46	4.23	3.08	7.91	3.34	2.79	2 50
.14					3.78	6.22	4.65	11.26	5.02	4.25	3.84
.18					6.48	10.20	7.80	18.14	8.40	7.20	6.56
.22					14.97	22.05	17.46	40.24	18.94	16.45	15.07
.02	2	.10	3	4	1.30	2.42	1.68	4.85	1.85	1.50	1.31
.06					1.86	3.27	2.35	6.24	2.56	2.12	1.89
.10					2.72	4.56	3.37	8.37	3.64	3.06	2.76
.14					4.19	6.75	5.10	12.03	5.49	4.68	4.25
.18					7.28	11.24	8.69	19.77	9.34	8.05	7.36
.22					17.93	25.77	20.66	47.08	22.44	19.60	18.02

				m	3	7	3	7	7
				a_1	0	0	1	1	3
				a_2	0	0	1	1	2
λ	ℓ	μ	H	G	$\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
.02	0	.20	3	4	2.52	10.20	4.21	5.45	3.37	2.60
.06					3.63	13.90	5.89	7.54	4.76	3.73
.10					5.64	21.00	8.96	11.46	7.32	5.79
.14					10.36	40.12	16.48	21.47	13.49	10.61
.18					34.78	271.15	63.04	100.51	49.67	35.45
.02	1	.20	3	4	2.71	10.64	4.47	5.74	3.59	2.79
.06					3.92	14.62	6.28	8.00	5.11	4.03
.10					6.15	22.42	9.68	12.33	7.93	6.31
.14					11.64	44.77	18.41	24.00	15.11	11.91
.18					47.18	676.33	91.06	165.06	70.30	47.98
.02	2	.20	3	4	3.79	12.70	5.78	7.21	4.79	3.88
.06					5.52	18.01	8.30	10.32	6.93	5.65
.10					9.04	29.59	13.50	16.89	11.31	9.23
.14					19.89	74.79	30.64	40.19	25.49	20.27
.02	0	.30	3	4	4.49	96.45	10.39	18.46	7.14	4.71
.06					6.99	162.88	16.20	29.24	11.10	7.32
.10					13.32	509.91	33.38	65.82	22.03	13.95
.14					60.94		1720.27		195.81	65.09
.02	1	.30	3	4	5.15	103.86	11.52	20.19	8.02	5.39
.06					8.22	184.24	18.59	33.22	12.83	8.58
.10					16.87	783.96	42.64	86.21	27.97	17.64
.14					204.23					231.36

					m	3	7	3	7	7
					a_1	0	0	1	1	3
					a_2	0	0	1	1	2
λ	ℓ	μ	H	G	$\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
.02	2	.30	3	4	10.77	148.91	19.90	32.09	14.91	11.11
.06					20.23	374.23	39.82	67.98	29.02	20.89
.10					98.59		663.70		232.73	103.62
.02	0	.40	3	4	7.37		27.49	120.33	14.84	7.88
.06					13.54		58.44	304.38	29.16	14.52
.10					51.43				28.06	57.20
.02	1	.40	3	4	9.24		32.70	140.55	17.97	9.85
.06					18.87		83.10	467.33	40.65	20.20
.10					252.05					336.06
.02	2	.40	3	4	55.28		131.42	479.88	83.75	57.20
.02	0	.50	3	4	11.73		93.50		33.14	12.88
.06					30.73		1082.05		134.02	34.53
.02	1	.50	3	4	16.97		128.25		45.84	18.51
.06					78.34				2869.62	91.17
.02	0	.60	3	4	18.80		571.34		88.36	21.36
.06					171.38					259.60
.02	1	.60	3	4	35.16		1267.66		167.17	39.73

				m	3	7	3	7	7
				a_1	0	0	1	1	3
				a_2	0	0	1	1	2
λ	ℓ	μ	H	G $\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$

λ	ℓ	μ	H	G	$\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
.02	0	.10	3	6	2.60	7.06	3.88	5.25	3.44	2.88
.06					3.66	9.48	5.33	7.12	4.76	4.02
.10					5.48	13.73	7.85	10.41	7.04	6.00
.14					9.34	23.13	13.25	17.69	11.96	10.20
.18					22.89	61.59	33.13	47.57	30.19	25.33
.02	1	.10	3	6	2.66	7.15	3.94	5.33	3.50	2.94
.06					3.74	9.61	5.43	7.23	4.85	4.11
.10					5.61	13.96	8.00	10.60	7.19	6.13
.14					9.60	23.66	13.58	18.13	12.27	10.48
.18					23.99	64.70	34.72	50.03	31.67	26.56
.02	0	.20	3	6	6.83	149.46	15.93	43.45	12.23	8.36
.06					11.27	291.46	27.04	78.76	20.61	13.88
.10					26.34	4904.83	77.67	373.65	55.74	33.96
.02	1	.20	3	6	7.21	154.15	16.61	44.98	12.79	8.79
.06					12.09	309.11	28.70	83.39	21.91	14.82
.10					29.69	64119.90	89.88	490.15	63.94	38.46
.02	0	.30	3	6	14.64		78.05		43.97	20.93
.06					41.67		648.21		217.75	68.25
.02	1	.30	3	6	16.35		85.57		48.35	23.21
.06					53.54		1862.25		352.00	91.08

					m	3	7	3	7	7
					a_1	0	0	1	1	3
					a_2	0	0	1	1	2
λ	ℓ	μ	H	G	$\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
.02	0	.40	3	6	31.75		1159.66		267.76	58.75
.02	1	.40	3	6	39.84		1627.55		349.49	73.87
.02	0	.50	3	6	87.49					300.85
.02	1	.50	3	6	171.26					1028.67
.02	0	.10	5	6	2.98	7.71	4.33	5.79	3.87	3.27
.06					5.63	13.04	7.77	10.06	7.04	6.09
.10					13.65	29.49	18.17	23.20	16.68	14.57
.02	1	.10	5	6	3.04	7.81	4.41	5.88	3.94	3.34
.06					5.75	13.24	7.91	10.23	7.18	6.22
.10					14.05	30.24	18.66	23.82	17.14	14.98
.02	0	.20	5	6	7.64	164.68	17.56	47.62	13.52	9.29
.06					18.67	618.21	45.45	141.38	34.16	22.68
.02	1	.20	5	6	8.08	170.35	18.35	49.41	14.16	9.78
.06					20.36	700.86	49.58	155.85	37.23	24.70
.02	0	.30	5	6	16.64		90.51		50.31	23.70
.06					210.11					1197.74
.02	1	.30	5	6	18.72		1957.68		55.93	26.50
.06					2632.60					

					m	3	7	3	7	7
					a_1	0	0	1	1	3
					a_2	0	0	1	1	2
λ	ℓ	μ	H	G	$\bar{w}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
.02	0	.40	5	6	37.97				365.92	71.56
.02	1	.40	5	6	49.38				534.57	94.30
.02	0	.50	5	6	126.91					621.87
.02	1	.50	5	6	399.61					
.02	0	.10	5	8	5.65	22.66	9.40	16.18	8.35	6.89
.06					11.50	44.48	18.60	31.87	16.65	13.85
.10					50.38	334.38	91.07	209.58	80.73	63.09
.02	1	.10	5	8	5.74	22.89	9.52	16.36	8.47	6.99
.06					11.74	45.17	18.93	32.40	16.96	14.11
.10					53.04	367.55	96.53	227.29	85.51	66.57
.02	0	.20	5	8	18.08		70.16	5987.20	49.55	29.31
.06					226.07					
.02	1	.20	5	8	19.01		73.25	6263.85	51.78	30.70
.06					368.16					
.02	0	.30	5	8	61.74		8905.23		1221.51	201.31
.02	1	.30	5	8	72.20		64996.20		1789.85	245.10

CHAPTER 7

DELAYS ON A TWO-LANE ROAD

7.1 Introduction

In a Royal Statistical Society symposium (and its following discussion) on traffic theory Tanner (1961a) and A. Miller (1961) have described some statistical problems in the formation of and delay to traffic on a long two-way road. Miller's formulation of traffic as consisting of alternate gaps having a negative exponential distribution, and of bunches, is similar to that described below and in the previous chapter. Tanner has used the special case of traffic forming a Borel-Tanner distribution to obtain the mean delay to a single extra fast vehicle. Another model of the flow and delay to vehicles on a two way road has been described by A. Miller (1962). Here vehicle velocities may ha a probability distribution, and the process of catching up and overtaking in one lane of traffic is considered for very simple overtaking rules. We consider a generalisation of Tanner's model to a more general description of the traffic flow. This description is similar to that for the major road traffic in the previous chapter; however, it now becomes more convenient to describe the flow in units of distance rather than time. These are in fact equivalent as we assume that all vehicles, except one, have a constant speed.

We suppose there is a long, straight two-way road having one lane in each direction, with vehicles travelling at a constant speed v in one direction and a constant speed V in the other direction. From a fixed point on the road the streams in both lanes may be considered as passing the point in alternate 'bunches' of vehicles and of gaps free of vehicles. The lengths of the gaps are independently and identically distributed random variables having negative exponential distributions with means g^{-1} and G^{-1} in the v- and V- streams respectively. We suppose the length of a bunch in the v- stream is a random variable with the d.f. $F_\epsilon(x)$ $(0 \leqq x, \epsilon < \infty)$, where the first interval ϵ of the total distance judged to be occupied by a bunch is actually free of vehicles, so that a 'v-bunch' commences ϵ before the front of the first vehicle in it; $F_\epsilon(x)$ has LST $\xi_\epsilon(\theta) \equiv \xi(\theta) = \int_0^\infty e^{-\theta x} dF_\epsilon(x)$ with finite mean $\overline{\xi}_\epsilon = \overline{\xi} = -\xi'_\epsilon(0) < \infty$. The LST $\xi_\epsilon(\theta)$ is assumed to exist and be finite for all real $\theta \geqq -S$ where S is a positive constant; this is equivalent to the condition that the d.f. $F_\epsilon(x)$ have an analytic characteristic function. In the V-stream we suppose the length of a bunch is a random variable with the d.f. $G_\eta(x)$ $(0 \leqq x, \eta < \infty)$, where the last interval η of a bunch is free of vehicles, LST $\chi_\eta(\theta) = \int_0^\infty e^{-\theta x} dG_\eta(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$ and finite mean $\overline{\chi} \equiv \overline{\chi}_\eta < \infty$. The distances ϵ and η, which may be chosen arbitrarily subject to the practical restrictions discussed below, are included at the beginning and the end, respectively, of the bunches as for these, unlike for the gaps, the distribution of the distance to the next vehicle may not be negative exponential. The formation of gaps and blocks is illustrated in Figure 1.

We are concerned with the speed of a single additional vehicle travelling in the v direction. When uninterrupted it travels at speed $u(> v)$, and when it catches up to a bunch and wishes to overtake, it behaves in the following way similar to that described by Tanner (1961a). Bunches of v-vehicles are overtaken in a single manoeuvre. Consider the moment when the u-vehicle reaches a v-bunch of length γ. If there is a free distance in the opposing stream of at least $\tau = \varphi\gamma + d$, where $\varphi = (u + V)/(u - v)$ and d is a constant which is a safety margin for the u-vehicle, then it overtakes without slowing down. Otherwise it reduces speed immediately to v and waits for a free distance of at least $D = \tau + \zeta = \tau + (v + V)t$ in the opposing stream, waits a further time t, where this is a positive constant, accelerates instantaneously to speed u and overtakes. The last distance η of a V-bunch is free of vehicles and is included in τ and D.

The practical interpretation of the formation of bunches is similar to that in the previous chapter. As before we must restrict the maximum distance between two successive vehicles if they are to be included in the same bunch. In the v-stream vehicles forming a bunch must be sufficiently close together so that another vehicle travelling in the same direction and overtaking cannot fit between two vehicles of the bunch, but must overtake in one manoeuvre. Thus we suppose that

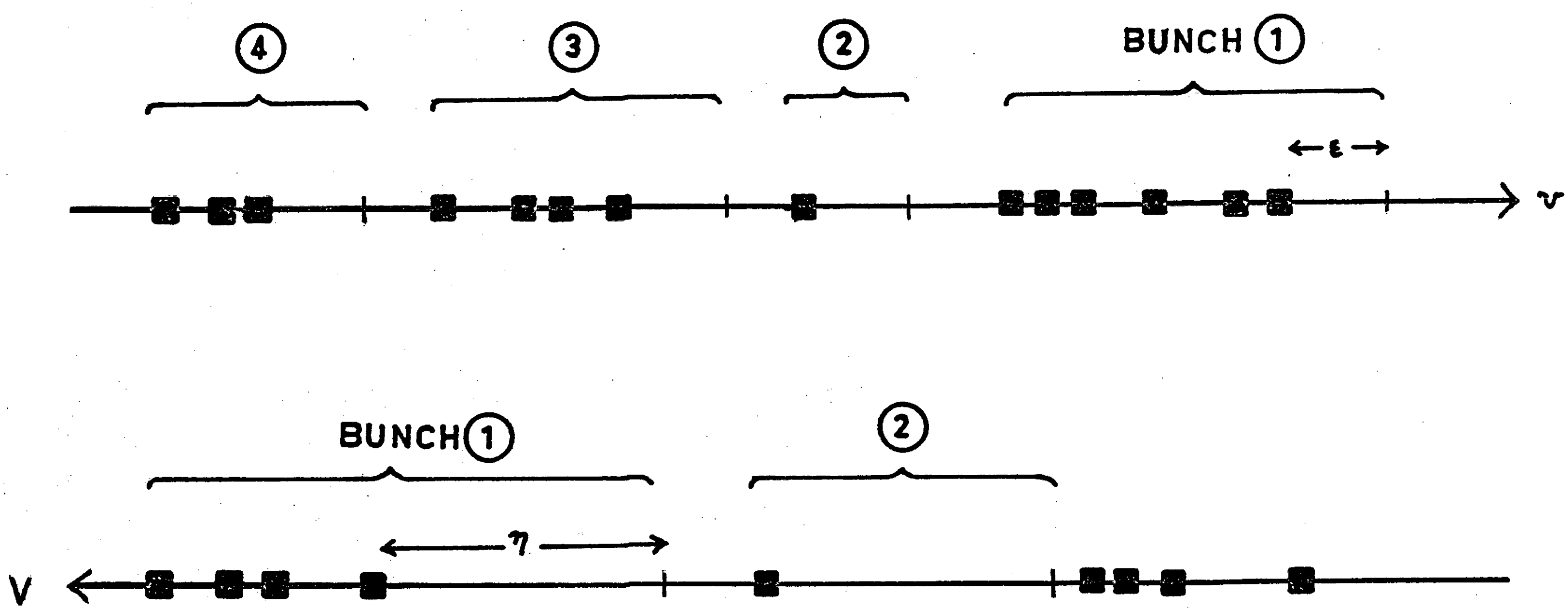

Figure 1. Formation of bunches and gaps.

the vehicles of a v-bunch must be less than $\epsilon(\leq \epsilon_1)$ apart, where ϵ_1 is the effective minimum space an overtaking vehicle requires to fit between two v-vehicles.

The minimum distance that the u-vehicle ever requires in order to pass a v-bunch is $\eta_1 = \varphi\gamma_1 + d$, where γ_1 is the shortest possible length of a v-bunch which includes just one vehicle, i.e. the effective length of a v-vehicle plus ϵ_1. If two successive vehicles of a V-bunch were allowed to be more than η_1 apart then it would be possible for the u-vehicle to pass a v-bunch during a V-bunch in the opposing stream, which we do wish to permit. Thus we suppose that a V-bunch consists of a group of vehicles in which successive vehicles are less than $\eta(0 \leq \eta \leq \eta_1)$ apart.

As we have seen, a 'bunch' may actually end or begin with a space free of vehicles, but until there is a distance η(or ϵ as the case may be) free of vehicles the distribution of the distance to the next vehicle is not necessarily negative exponential. We thus allow interaction between successive vehicles which are less than η(or ϵ) apart, but this interaction decreases as the distance between vehicles becomes greater and cease when it reaches η(or ϵ). In practice η should be of the order of 100 yards or more, so that we may allow a reasonable degree of interference between V-vehicles. Now ϵ_1 is considerably less than η_1, so that less interference is allowed in the v- than the V-stream. However, we can improve the formulation in the v-stream and still find the mean speed of the u-vehicle. Instead of the gap distribution in the v-stream having a negative exponential distribution, we suppose it has a d.f. $A(x)$ $(0 \leqq x < \infty)$, with LST $\alpha(\theta) = \int_0^\infty e^{-\theta x} dA(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$ and finite mean $a = -\alpha'(0) < \infty$. This combines the advantages of both our model and of the headway distribution model of Weiss and Maridudin (1962) and Buckley (1962). Once there is a free distance of at least ϵ in the v-stream the distribution of the distance to the next vehicle is independent of the characteristics of a bunch, but need not be of the simple negative exponential form.

The u-vehicle travels alternately for free runs at speed u and waits, possibly for time zero, at speed v. One of the runs at speed u lasts for the length of a v-bunch plus the length of a gap, so that if $\gamma(\theta) = \int_0^\infty e^{-\theta x} dC(x)$ $(\mathrm{Rl}\ \theta \geqq 0)$ is the LST of the d.f. $C(x)$ of this time, we have

$$\gamma(\theta) = \alpha(\theta/(u-v))\ \xi(\theta/(u-v)), \tag{1.1}$$

the mean of the distribution being

$$\bar{x} = -\gamma'(0) = \frac{\bar{\alpha}+\bar{\xi}}{u-v}\ . \tag{1.2}$$

Let $\bar{w}$ be the mean of the distribution of delay to the u-vehicle at a v-bunch and $\bar{u}$ the mean speed travelled by the u-vehicle. If the u-vehicle travels for a time x at speed u and a time w at speed v then its speed over x + w is $u_1 = (ux + vw)/(x + w)$. Over the time taken to travel N times at speed u and N times at speed v (for times x_i and w_i $(1 \leqq i \leqq N)$), the speed of the u-vehicle is $u_N = \sum_{i=1}^N (ux_i + vw_i) / \sum_{i=1}^N (x_i + w_i)$; this converges in probability to the average speed

$$\bar{u} = \frac{u\bar{x} + v\bar{w}}{\bar{x} + \bar{w}}, \tag{1.3}$$

over a long journey. Thus to obtain this it is sufficient to find the mean $\bar{w}$ of the distribution of waiting times at a v-bunch, as we shall do in the next section.

7.2 The mean speed

The distance the u-vehicle travels from the front of one v-bunch to the rear of the next, relative to the V-stream, is X/φ where X is a random variable with the d f. $A(x)$. When the u-vehicle completes the passing of one v-bunch, there is a further gap of at least d in the opposing stream The relative distance ν to the next v-bunch may be less than or greater than d; we consider the two cases separately.

We require the distribution of the distance from the beginning of a V-bunch up to the first free distance (including the last η of the bunches) of at least τ(or D) in the V-stream; this is similar to (6.1.1). The LST $\chi_y(\theta) = \int_0^\infty e^{-\theta x} dG_y(x)$ (Rl $\theta \geqq 0$) of the d.f. $G_y(x)$ ($y \geqq \eta$, $0 \leqq x < \infty$)of the distance from the commencement of a V-bunch to the first point where there is a distance of at least y (including y) free of v-vehicles is

$$\chi_y(\theta) = \frac{(G+\theta)\, \chi_\eta(\theta) e^{-(G+\theta)(y-\eta)}}{G+\theta - G\chi_\eta(\theta)\{1-e^{-(G+\theta)(y-\eta)}\}} . \tag{2.1}$$

When $\nu < d$ there is no delay if there is a further free distance of at least $\nu + \tau - d$ in the V-stream; otherwise the vehicle waits at speed v until the first break of at least D appears in the V-stream, waits a further distance ζ, accelerates to speed u and passes. The LST $\psi_1(\theta, \nu, \gamma)$ of the distribution of the distance for this wait is

$$\psi_1(\theta,\nu,\gamma) = e^{-G(\nu+\tau-d)} + \int_{y=o}^{\nu+\tau-d} \int_{z=y}^{\infty} Ge^{-Gy-\theta(z-\tau+d-\nu)} dy dG_D(z-y)$$

$$= e^{-G(\nu+\tau-d)} + \frac{GX_D(\theta)}{G+\theta}\left\{e^{\theta(\nu+\tau-d)} - e^{-G(\nu+\tau-d)}\right\}. \qquad (2.2)$$

Integrating (2.2) over the distribution of ν between 0 and d, we obtain that

$$\psi_1(\theta,\gamma) = \int_{x=o}^{d} \psi_1(\theta,x,\gamma)dA(x/\varphi)$$

$$= \alpha(G\varphi,d/\varphi)e^{-G(\tau-d)} + \frac{GX_D(\theta)}{G+\theta}\left\{e^{\theta(\tau-d)}\alpha(\theta\ \varphi,d/\varphi)\right.$$

$$\left. - e^{-G(\tau-d)}\alpha(G\varphi,d/\varphi)\right\}, \qquad (2.3)$$

where

$$\alpha(\theta,d) = \int_{y=o}^{d} e^{-\theta y}dA(y).$$

When $\alpha(\theta) = g/(g+\theta)$, then (2.3) reduces to

$$\psi_1(\theta,\gamma) = \frac{GCX_D(\theta)}{G+\theta}\left\{\frac{e^{\theta(\tau-d)} - e^{\theta\tau-Cd}}{C-\theta} + \frac{e^{-G\tau-Cd} - e^{-G(\tau-d)}}{G+C}\right\}$$

$$+ \frac{Ce^{-G(\tau-d)}}{G+C}\left\{1 - e^{-(G+C)d}\right\}, \qquad (2.4)$$

where $C = g/\varphi$.

When $\nu > d$ the problem is more complicated as we have to consider what has happened in the opposing traffic further than the distance d from the front of the last v-bunch. We use the same method as in §3.8, §4.2 and §6.2. We consider the V-stream in terms of gaps and τ blocks, where a block is defined as in §6.1. Enumerating all the possible points of arrival of the u-vehicle at a v-bunch we derive the IST $\psi_2(\theta,\gamma)$ of the distribution of the distance of this wait $(\nu > d)$ as

$$\psi_2(\theta,\gamma) = \sum_{n=0}^{\infty} \Bigg[\int_{y=0}^{\infty} e^{-G(y+\tau)}$$

$$+ \int_{w=0}^{\infty} \int_{z=w}^{\infty} \int_{y=z-\tau}^{z} Ge^{-G(w+\tau-z+y)} dw dG_\tau(z-w)$$

$$+ \Bigg\{ \int_{w=0}^{\tau} \int_{z=w}^{\infty} \int_{y=0}^{z-\tau} \int_{v=0}^{\zeta} \int_{\alpha=v}^{\infty} + \int_{w=\tau}^{\infty} \int_{z=w}^{\infty} \int_{y=0}^{z-\tau} \int_{v=0}^{\zeta} \int_{\alpha=v}^{\infty}$$

$$+ \int_{w=0}^{\infty} \int_{z=w}^{\infty} \int_{y=z-\tau}^{z} \int_{v=0}^{y+\tau-z} \int_{\alpha=v}^{\infty} \Bigg\} \Bigg\{ G^2 e^{-G(w+v)-\theta(z+\alpha-y-\tau)}$$

$$dw dG_\tau(z-w) dv dB_D(\alpha-v) \Bigg\}$$

$$+ \int_{w=0}^{\infty} \int_{z=w}^{\infty} \int_{y=\max(0,w-\tau)}^{z-\tau} Ge^{-G(w+\zeta)-\theta(z-y-\tau+\zeta)} dw dG_\tau(z-w) \Bigg]$$

$$d_y h_n(\gamma,\tau,y+d), \qquad (2.5)$$

where

$$d_y h_n(\gamma,\tau,y+d) = \prod_{i=1}^{n} \int_{w_i=0}^{\infty} \int_{z_i=w_i}^{\infty} Ge^{-Gw_i} dG_\tau(z_i-w_i) d_y A\left(\frac{z_1+\ldots+z_n+y+d}{\varphi}\right)$$

$h_n(\gamma,\tau,d)$ being the probability that the u-vehicle does not reach the next v-bunch during the first n gaps and τ blocks in the opposing stream. When the v-stream gaps have a negative exponential distribution with $\alpha(\theta) = g/(g+\theta)$, then (2.5) simplifies to

$$e^{Cd}\psi_2(\theta,\gamma) = (1-g(\gamma,\tau))^{-1}\left[\left\{\frac{G^2C\chi_D(\theta)\chi_\tau(\theta)(1-e^{-(G+\theta)\zeta})}{(G+\theta)^2}\right.\right.$$

$$\left.+\frac{GC\chi_\tau(\theta)e^{-(G+\theta)\zeta}}{G+\theta}\right\}\left\{\frac{e^{\theta\tau}}{C-\theta}-\frac{e^{-G\tau}}{G+C}\right\}+\frac{Ce^{-G\tau}}{G+C}\Bigg]$$

$$+\frac{g(\gamma,\tau)}{1-g(\gamma,\tau)}\left[\frac{GC\chi_D(\theta)}{G+\theta}\left\{\frac{e^{-G\tau}}{G+C}+\frac{e^{C\tau-(G+\theta)\zeta}}{C-\theta}-\frac{e^{\theta\tau}}{C-\theta}-\frac{e^{C\tau}}{G+C}\right\}\right.$$

$$\left.-\frac{Ce^{C\tau-(G+\theta)\zeta}}{C-\theta}+\frac{Ce^{C\tau}-Ce^{-G\tau}}{G+C}\right], \tag{2.6}$$

where

$$g(\gamma,\tau) = G(G+C)^{-1}\chi_\tau(C).$$

The LST $\vartheta(\theta,\gamma)$ of the distribution of the distance which the u-vehicle spends waiting at a v-bunch of length γ is

$$\vartheta(\theta,\gamma) = \psi_1(\theta,\gamma) + \psi_2(\theta,\gamma), \tag{2.7}$$

and the LST $\Phi(\theta,\gamma) = \int_0^\infty e^{-\theta x}\, dW(x,\gamma)$ (Rl $\theta \geq 0$) of the d.f. $W(x,\gamma)$ $(0 \leq x < \infty)$ of the time the u-vehicle waits at a v-bunch of length γ is

$$\Phi(\theta,\gamma) = \vartheta(\frac{\theta}{v+V},\gamma), \tag{2.8}$$

as one stream is travelling at a relative speed of v + V compared with the other stream.

We wish to obtain the LST $\Phi(\theta) = \int_0^\infty e^{-\theta x} dW(x)$ (Rl $\theta \geq 0$) of the d.f. $W(x)$ $(0 \leq x < \infty)$ of the wait for the u-vehicle at a v-bunch. From Lukács (1960) (c.f.(4.2.11)) we have

$$\begin{aligned}\int_{y=0}^{\infty} \Phi(\theta,y)dF(y) &= \int_{y=0}^{\infty} \int_{x=0}^{\infty} e^{-\theta x} d_x W(x,y)dF(y) \\ &= \int_{x=0}^{\infty} \int_{y=0}^{\infty} e^{-\theta x} d_x W(x,y)dF(y) \\ &= \int_{x=0}^{\infty} e^{-\theta x} dW(x) = \Phi(\theta).\end{aligned} \tag{2.9}$$

For the remainder of this chapter we consider the special case of gaps in the v-stream having a negative exponential distribution with $\alpha(\theta) = g/(g+\theta)$, although the mean speed $\bar{u}$ may be found explicitly for some other cases. Substituting (2.2) and (2.6) into (2.8) and differentiating, we find that

$$\begin{aligned}-\Phi'(0,\gamma) = &\frac{(1+G\bar{X}_\eta)e^{G(d+\zeta-\eta)}}{G(v+V)}\left\{e^{G\varphi\gamma} - \frac{C}{G+C}\right\} - \frac{(\varphi\gamma+C^{-1})}{v+V} \\ &+ \frac{X_\eta(C)(1+G\bar{X}_\eta)e^{C(\eta-d)}}{(G+C-GX_\eta(C))(v+V)}\left\{1 - \frac{Ce^{G\zeta}}{G+C}\right\}.\end{aligned} \tag{2.10}$$

We now obtain the mean $\overline{w}$ of the distribution of the wait of the u-vehicle at a v-bunch as $\overline{w} = -\Phi'(0) = -\int_{x=0}^{\infty} \Phi'(0,x)dF(x)$, where a sufficient condition for this to be finite is that $S \geqq G\varphi$, i.e. $\xi_\epsilon(\theta)$ exists and is finite for all real $\theta \geqq -G\varphi$. From (2.10) we obtain

$$\overline{w} = \frac{(1+G\overline{X}_\eta)e^{G(d+\zeta-\eta)}}{G(v+V)}\left\{\xi(-G\varphi)-\frac{C}{G+C}\right\} - \frac{(\varphi\overline{\xi}_\epsilon+C^{-1})}{v+V}$$

$$+\frac{X_\eta(C)(1+G\overline{X}_\eta)e^{C(\eta-d)}}{(G+C-GX_\eta(C))(v+V)}\left\{1-\frac{Ce^{G\zeta}}{G+C}\right\} . \tag{2.11}$$

Substitution of (2.11) and (1.2) into (1.3) yields the mean speed of the u-vehicle over a long stretch of road, which we can write in the form

$$\frac{v+V}{\overline{u}-v} = -1 + \frac{g(1+G\overline{X}_\eta)}{G(1+g\overline{\xi})}\left[e^{G(d+\zeta-\eta)}\left\{\xi(-G\varphi)-\frac{C}{G+C}\right\}\right.$$

$$\left.+\frac{GX_\eta(C)e^{C(\eta-d)}}{G+C-GX_\eta(C)}\left\{1-\frac{Ce^{G\zeta}}{G+C}\right\}\right] . \tag{2.12}$$

The mean wait $\overline{w}$ at a v-bunch becomes infinite when $\xi(-G\varphi)$ becomes infinite, i.e. at the point where $G\varphi$ reaches S. Thus over an infinitely long journey the mean speed $\overline{u}$ approaches v asymptotically, although over a short period $\overline{u}$ is greater than v as there is a non-zero probability of part of the journey being made at speed u.

If the v- and V-streams form Borel-Tanner distributions with effective vehicle lengths b and B respectively then

$$\xi_0(\theta) = e^{r(\xi_0(\theta)-1)-\theta b}$$

$$\chi_0(\theta) = e^{R(\chi_0(\theta)-1)-\theta B} ,$$

where $r = bg$, $R = BG$. The results (20) and (21) of Tanner (1961a) follow from (2.11) and (2.12) with $\epsilon = \eta = 0$, $\xi_0(-G\varphi) = N$, $\chi_\eta(C) \equiv \chi_0(C) = K$. We have thus generalised Tanner's result to a wider class of bunching distributions. The practical interpretations given for his model by Tanner may in general be applied to the above more general case.

7.3 A flow of fast vehicles

Let us now consider a flow of fast vehicles, whose desired speed is $u > v$, travelling in the same direction as the v-vehicles and fitting in the gaps between the bunches of v-vehicles. If the u-vehicles were originally placed at random in the gaps, they would eventually form bunches and gaps (not necessarily negative exponential) due to the interference of the v-bunches. We wish to find the mean speed of the u-vehicles; however, this is very difficult in general as the u-vehicles do not behave independently of each other. The process of arrivals of u-vehicles at a v-bunch is dependent on the output from the previous v-bunch and this is not in general of a Poisson type. The model is most readily applicable when there is only a light flow of v-vehicles, and particularly when there are just a few v-vehicles spaced at random with a large mean distance between them; this may be applicable to the delay caused by semi-trailers and heavy transports on an inter-city run. In the further special case of just a single v-vehicle the problem is similar to that discussed in Chapter 6 for delay at an intersection.

We assume that in a stationary state bunches of u-vehicles arrive at a v-bunch in a Poisson process with parameter ν; the size of a bunch is a random vvariable with probabilities b_i $(i = 1,2,\dots)$, p.g.f. $\beta(s) = \sum_{i=1}^{\infty} b_i s^i$ $(|s| \leq 1)$ and finite mean $\bar{b} = \beta'(1) < \infty$. A reasonable assumption might be that these are bunches of a Borel-Tanner distribution, so that

$$\beta(s) = se^{(\beta(s)-1)r} ,$$

where $r = L\nu$, L being the effective length of a u-vehicle.

We make a further simplifying assumption, namely that there is always sufficient room between two v-bunches for any number of u-vehicles to fit into this gap; this seems reasonable if there is only a light flow of v-vehicles.

It is also assumed that the v-bunches are sufficiently far apart for the wait at two successive v-bunches to be considered independently, i.e. the V-stream moves so far along while a u-vehicle is travelling between bunches that the effect of a particular gap or block in the V-stream cannot be connected with the wait at successive v-bunches.

If a u-vehicle arrives at a v-bunch where there are no other u-vehicles waiting, then we suppose that for overtaking the v-bunch it behaves exactly as for the single u-vehicle above. If it arrives to find other u-vehicles waiting then it queues up behind these vehicles; when it reaches the head of the u-queue there is a gap in the V-stream. If this gap lasts a further distance h, where h is a random variable with d.f. $H(x)$ $(0 \leq x < \infty)$ and LST $\mathcal{H}(\theta) = \int_0^\infty e^{-\theta x} dH(x)$ $(\mathrm{Rl}\ \theta \geq 0)$ and mean $\bar{h} = -\mathcal{H}'(0) < \infty$, then it waits for this distance h, accelerates instantaneously to speed u and overtakes; otherwise it remains at speed v until the first free distance of at least D appears in the opposing stream and behaves as for a single u-vehicle.

At a single v-bunch the distribution of the wait for u-vehicles is similar to that described in the previous chapter and is an application of the results of §3.10. We require only to find the service time distributions for u-vehicles reaching an empty and a non-empty queue of u-vehicles. In the latter case the LST $\psi^*(\theta,\gamma)$ of the service time distribution for a v-bunch of length γ is given by

$$\psi^*(\theta,\gamma) = \mathcal{H}(G+\theta) + \frac{G}{G+\theta}\, \chi_D(\theta)\, e^{\theta\tau}\{1-\mathcal{H}(G+\theta)\} \,. \tag{3.1}$$

Relative to the V-stream, u-vehicles arrive (in bunches) at a v-bunch in a Poisson process with parameter $\lambda = \nu/\varphi$. Let us consider a u-vehicle arriving at a v-bunch with no other u-vehicles waiting. We break up the V-stream into parts, one part consisting of a gap and a block. The first gap must last at least a distance τ as this distance is free when the previous u-vehicle overtakes the bunch. Enumerating the possible points of arrival of the u-vehicle at a v-bunch of length γ with no other u-vehicles waiting we find that the LST $\zeta^*(\theta,\gamma)$ of the distance a u-vehicle spends at the v-bunch is

$$\zeta^*(\theta,\gamma) = \zeta_1(\theta,\gamma) + \frac{g_o(\gamma)}{1-g_1(\gamma)}\,\zeta_2(\theta,\gamma) \tag{3.2}$$

where

$$\begin{aligned}
\zeta_1(\theta,\gamma) = {} & \frac{\lambda e^{-(\lambda+D)\tau}}{\lambda+G} + \frac{\lambda G X_D(\theta)}{(\lambda+\theta)(G+\theta)}\,(1-e^{-(\lambda+\theta)\tau}) \\
& + \frac{\lambda}{\lambda+G}\,(1 - \frac{GX_D(\theta)}{\lambda+\theta})\,(1 - e^{-(\lambda+G)\tau}) \\
& + \left\{ \frac{GX_D(\theta)}{G+\theta} + (1 - \frac{GX_D(\theta)}{G+\theta})e^{-(G+\theta)\zeta} \right\} \Bigg[\frac{\lambda G e^{-(\lambda+G)\tau}}{(\lambda-\theta)(\lambda+G)} \\
& \left\{X_\tau(\theta) - X_\tau(\lambda)\right\} + \frac{\lambda G}{\lambda-\theta} \left\{ \frac{X_\tau(\theta)}{G+\theta}(e^{-(\lambda-\theta)\tau} - e^{-(\lambda+G)\tau}) \right. \\
& \left. - \frac{X_\tau(\lambda)}{\lambda+G}(1 - e^{-(\lambda+G)\tau}) \right\} \Bigg] \\
& + \frac{\lambda G e^{-\lambda\tau} X_\tau(\lambda)}{\lambda+G} \left\{ \frac{GX_D(\theta)(e^{\lambda\tau}-e^{\theta\tau})}{(\lambda-\theta)(G+\theta)} + \left(1 - \frac{GX_D(\theta)}{G+\theta}\right)\left(\frac{e^{\lambda\tau}-e^{-G\tau}}{\lambda+G}\right)\right\}
\end{aligned} \tag{3.3}$$

$$\begin{aligned}
\zeta_2(\theta,\gamma) = {} & \left\{ \frac{GX_D(\theta)}{G+\theta} + (1 - \frac{GX_D(\theta)}{G+\theta})\,e^{-(G+\theta)\zeta} \right\} \Bigg[\frac{\lambda G e^{-G\tau}}{(\lambda-\theta)(\lambda+G)} \\
& \left\{X_\tau(\theta)-X_\tau(\lambda)\right\} + \frac{\lambda G}{\lambda-\theta} \left\{ \frac{X_\tau(\theta)}{G+\theta}\,(e^{-(\lambda-\theta)\tau} - e^{-(\lambda+G)\tau}) \right. \\
& \left. - \frac{X_\tau(\lambda)}{\lambda+G}(1 - e^{-(\lambda+G)\tau}) \right\} \Bigg] + \frac{\lambda}{\lambda+G}\,e^{-G\tau}
\end{aligned}$$

$$+ \frac{\lambda G X_\tau(\lambda)}{\lambda+G} \left\{ \frac{G_D(\theta)}{(\lambda-\theta)(G+\theta)} (e^{\lambda\tau}-e^{\theta\tau}) + (1 - \frac{G X_D(\theta)}{G+\theta}) \frac{e^{\lambda\tau}-e^{-G\tau}}{\lambda+G} \right\}, \tag{3.4}$$

$$g_o(\gamma) = 1 - \zeta_1(0,\gamma) = \frac{G X_\tau(\lambda)e^{-\lambda\tau}}{\lambda+G}$$

$$g_1(\gamma) = 1 - \zeta_2(0,\gamma) = \frac{G X_\tau(\lambda)}{\lambda+G}.$$

Integration of (3.1) and (3.2) over the distribution of the length of a v-bunch yields the LST's $\psi^*(\theta)$ and $\zeta^*(\theta)$ of the distributions of the distance a u-vehicle spends waiting at a v-bunch. The LST's $\psi(\theta)$ and $\zeta(\theta)$ of the distributions of the lengths of time of these waits are given by

$$\psi(\theta) = \psi^*(\theta/(v+V))$$
$$\zeta(\theta) = \zeta^*(\theta/(v+V)) . \tag{3.5}$$

Substitution of (3.5) into (3.10.2) yields the stationary waiting time distribution for u-vehicles at a v-bunch; the mean delay is given by (3.10.5). Finally substitution of this last result and of (1.3) with $\alpha(\theta) = \mu/(\mu+\theta)$ into (1.2) yields the mean speed of u-vehicles.

REFERENCES

ADAMS, W.F. (1936) Road traffic considered as a random series, J. Inst. Civil Eng., 4, 121-130

BAILEY, N.T.J. (1954) On queueing with bulk service, J. Roy. Statist. Soc., B16, 80-87

BARRY, J.Y. (1956) A priority queueing problem, Opns. Res., 4, 385-386.

BARTLETT, M.S. (1955) An Introduction to Stochastic Processes, Cambridge U.P.

BLUNDEN, W.R., CLISSOLD, C.M., and FISHER, R.B. (1962) Distribution of acceptance gaps for crossing and turning manoeuvres, First Biennial Conference of the Australian Road Research Board.

BUCKLEY, D.J. (1962) Road traffic headway distributions, First Biennial Conference of the Australian Road Research Board.

CHUNG, K.L. and FUCHS, W.H.J. (1951) On the distribution of values of sums of random variables, Mem. Amer. Math. Soc., 6, 1-12.

COBHAM, A. (1954) Priority assignment in waiting line problems, Opns. Res., 2, 70.

CRAMÉR, H. (1955) Collective Risk Theory, Försakringsahtielologet Skandia.

EWENS, W.J. and FINCH, P.D. (1962) A generalised single server queue with Erlang inputs, Biometrika, 49, 242-245.

FELLER, W. (1957) An Introduction to Probability Theory and its Applications, Vol. 1, Second edition, John Wiley.

FINCH, P.D. (1959) A probability limit theorem with application to a generalisation of queueing theory, Acta Math. Acad. Sci. Hung., 10, 317-325.

FOSTER, F.G. (1961) Queues with batch arrivals, Acta Math. Acad. Sci. Hung., 12, 1-10.

GANI, J. (1957) Problems in the probability theory of storage systems,
J. Roy. Statist. Soc., B19, 181-206.

GANI, J. (1958) Elementary methods in an occupancy problem of storage,
Math. Annalen, 136, 454-465.

GANI, J. (1960) Time-dependent results for a dam with ordered Poisson inputs,
Nature, 188, 341-342.

GANI, J. (1961) First emptiness of two dams in parallel,
Ann. Math. Statist., 32, 219-229.

GANI, J. (1962a) A stochastic dam process with non-homogeneous Poisson inputs,
Studia Math. 21, 307-315.

GANI, J. (1962b) The time-dependent solution for a dam with ordered inputs,
Studies in Applied Probability and Management Science, ed. by Arrow, Karlin and Scarf. Stanford U.P., 101-109.

GANI, J. and PRABHU, N.U. (1959a) Remarks on the dam with Poisson type inputs,
Aust. J. Appl. Sci., 10, 113-122.

GANI, J. and PRABHU, N.U. (1959b) The time-dependent solution for a storage model with Poisson input,
J. Math. and Mech., 8, 653-664.

GANI, J. and PYKE, R. (1962) Inequalities for first emptiness of a dam with ordered inputs,
J. Roy. Statist. Soc., B24, 102-106.

GARWOOD, F. (1940) An application of the theory of probability to the operation of vehicular-controlled traffic signals,
J. Roy. Statist. Supplement, 7, 65-77.

GAVER, D.P. (1959) Imbedded Markov chain analysis of a waiting line process in continuous time,
Ann. Math. Statist., 30, 698-720.

GAVER, D.P. (1962) A waiting line with interrupted service, including priorities,
J. Roy. Statist. Soc., B24, 73-90.

GAVER, D.P. (1963) Accommodation of second-class traffic,
Opns. Res., 11, 73-87.

GAVER, D.P. and MILLER, R.G. (1962) Limiting distributions for some storage problems,
Studies in Applied Probability and Management Science, ed. by Arrow, Karlin and Scarf. Stanford U.P., 110-126.

GHOSAL, A. (1960) Emptiness in the finite dam,
Ann. Math. Statist., 31, 863-868.

HEATHCOTE, C.R. (1959) The time-dependent problem for a queue with preemptive priorities,
Opns. Res., 7, 670-680.

HEATHCOTE, C.R. (1960) A simple queue with several preemptive priority classes,
Opns. Res., 8, 630-638.

HOLLEY, J. (1954) Waiting lines subject to priorities,
Opns. Res. 2, 341.

JAISWAL, N.K. (1962) Time-dependent solution of the head-of-the-line priority queue,
J. Roy. Statist. Soc., B24, 91-101.

KENDALL, D.G. (1957) Some problems in the theory of dams,
J. Roy. Statist. Soc., B19, 207-212.

LINDLEY, D.V. (1952) The theory of queues with a single server,
Proc. Cambridge Phil. Soc., 48, 277-289.

LUKÁCS, E. (1960) Characteristic Functions,
Charles Griffen.

MEISLING, T. (1958) Discrete-time queueing theory,
Opns. Res., 6, 96-105.

MILLER, A.J. (1961) A queueing model for road traffic flow,
J. Roy. Statist. Soc., B23, 64-75.

MILLER, A.J. (1962) Road traffic flow considered as a stochastic process,
Proc. Cambridge Phil. Soc., 58, 312-325.

MILLER, R.G. (1960) Priority queues,
Ann. Math. Statist., 31, 86-103.

MORAN, P.A.P. (1954) A probability theory of dams and storage systems,
Aust. J. Appl. Sci., 5, 116-124.

MORAN, P.A.P. (1956) A probability theory of a dam with a continuous release,
Quart. J. Math. (Oxford, 2), 7, 130-137.

MORAN, P.A.P. (1959) The Theory of Storage,
Methuen.

NEWELL, G.F. (1955) Statistical analysis of the flow of highway traffic through a signalised intersection,
Quart. of Appl. Math., 13, 353-369.

PRABHU, N.U. (1958) Some exact results for the finite dam,
Ann. Math. Statist., 29, 1234-1243.

PRABHU, N.U. (1960) Some results for the queue with Poisson arrivals,
J. Roy. Statist. Soc., B22, 104-107.

RAFF, M.S. (1951) The distribution of blocks in an uncongested stream of automobile traffic,
J. Amer. Statist. Assoc., 46, 114-123.

REICH, E. (1958) On the integro-differential equation of Takács, I,
Ann. Math. Statist, 29, 563-570.

REICH, E. (1959) On the integro-differential equation of Takács, II,
Ann. Math. Statist., 30, 143-148.

STEPHAN, F.F. (1958) Two queues under preemptive priority with Poisson arrival and service rates,
Opns. Res., 6, 399-418.

TAKÁCS, L. (1955) Investigation of waiting time problems by reduction to Markov processes,
Acta Math. Acad. Sci. Hung., 6, 101-129.

TAKÁCS, L. (1961) Transient behavior of single-server queueing processes with Erlang input,
Transactions Amer. Math. Soc., 100, 1-28.

TANNER, J.C. (1951) The delay to pedestrians crossing a road,
Biometrika, 38, 383-392.

TANNER, J.C. (1953) A problem of interference between two queues,
Biometrika, 40, 58-69.

TANNER, J.C. (1961a) Delays on a two-lane road,
J. Roy. Statist. Soc., B23, 38-63.

TANNER, J.C. (1961b) A theoretical analysis of delays at an uncontrolled intersection,
D.S.I.R. Road Research Laboratory Research Note RN/3937.

TANNER, J.C. (1962) A theoretical analysis of delays at an uncontrolled intersection,
Biometrika, 49, 163-170.

WEESAKUL, B. (1961a) First emptiness in a finite dam,
J. Roy. Statist. Soc., B23, 343-351.

WEESAKUL, B. (1961b) The time-dependent solution for a finite dam with geometric inputs,
Research Report of the first Summer Research Institute of the Australian Mathematical Society.

WEESAKUL, B. and YEO, G.F. (1963) Some problems in finite dams with an application to insurance risk,
Z. Wahrkeinlichkeitstheorie, 2, (to appear).

WHITE, H. and CHRISTIE, L.S. (1958) Queueing with preemptive priorities or with breakdown,
Opns. Res., 6, 79-96.

WIDDER, D.V. (1946) The Laplace Transform,
Princeton U.P.

WINSTEN, C.B. (1956) (BECKMANN, M., McGUIRE, C.B. and WINSTEN, C.B.)
Studies in the Economics of Transportation,
Yale U.P.

YEO, G.F. (1961a) The time-dependent solution for an infinite dam with discrete additive inputs,
J. Roy. Statist. Soc. B23, 173-179.

YEO, G.F. (1961b) Dams in discrete time,
M.Sc. Thesis (W.A.)

YEO, G.F. (1961c) A discrete dam with ordered inputs,
Research Report of the first Summer Research Institute of the Australian Mathematical Society.

YEO, G.F. (1962) Single server queues with modified service mechanisms, J. Aust. Math. Soc., 2, 499-507.

YEO, G.F. (1963a) Preemptive priority queues, J. Aust. Math. Soc., 3 (to appear).

YEO, G.F. (1963b) Traffic delays on a two-lane road, Biometrika (to appear).

YEO, G.F. and WEESAKUL, B. (1964) Delays to road traffic at an intersection, (submitted to Applied Probability).

Bibliography on theory of traffic flow and related subjects (1961), Opns. Res., 9, 568-573.

Reprinted from *The Australian Journal of Statistics*, Vol. 4, No. 2,
pages 55-60, August, 1962

ON THE AGE DISTRIBUTION OF n RANKED ELEMENTS AFTER SEVERAL REPLACEMENTS

J. GANI and G. F. YEO

Department of Statistics, Australian National University

ON THE AGE DISTRIBUTION OF n RANKED ELEMENTS AFTER SEVERAL REPLACEMENTS[1]

J. GANI and G. F. YEO

Department of Statistics, Australian National University

1. Consider a system of n elements whose ranked initial ages at time $t=T_0=0$ are $0=x_1<x_2<\ldots<x_n$. At certain instants of time $T_1=\tau_1$, $T_2=\tau_1+\tau_2,\ldots$, $T_m=\tau_1+\ldots+\tau_m,\ldots$, where the intervals $\tau_i>0$ $(i=1, 2,\ldots)$, one element in the system is replaced at random by another element of age zero. We are concerned with the age distributions of the n ranked elements after $m\geqslant 1$ replacements. This problem arises in a biological model of phage reproduction, where from time to time one of the phages in the reproductive pool of fixed size n leaves the pool on receiving its protein envelope, and is replaced by a new phage. It could equally well be considered as a model for the replacement of parts in a system of machines.

Let us suppose that at each instant of change (or regeneration point) there is a probability $p_i(1\leqslant i\leqslant n, p_1+p_2+\ldots+p_n=1)$ that the i^{th} ranked element is replaced. The distribution of the $\{\tau_i\}$ may in general take quite a complicated form ; we shall, however, consider only the case where the τ_i are independently and identically distributed, and further the special case where this distribution is negative exponential.

2. Let us define

$$F_i(x,T_j)=Pr\{i^{th} \text{ ranked element has age} \leqslant x \text{ at } t=T_j+0\}$$

as the distribution function (d.f.) of the i^{th} ranked element after the j^{th} regeneration point, and

$$\psi_i(\theta,T_j)=\int_{x=0-}^{\infty} \mathrm{e}^{-\theta x} d_x F_i(x,T_j) \quad Rl\,(\theta)>0$$

as its Laplace transform. It will readily be seen that $F_i(x,T_j)$, and thus $\psi_i(\theta,T_j)$, is a function not only of T_j, but also $T_1,T_2,\ldots,T_{j-1}$. Let us further denote by $\mathbf{F}(x,T_j)$ and $\boldsymbol{\psi}(\theta,T_j)$ the column vectors with elements $\{F_i(x,T_j)\}$ and $\{\psi_i(\theta,T_j)\}$, and write $q_i=p_1+p_2+\ldots+p_i$. From a consideration of the possible changes at each regeneration point we readily obtain for the d.f.'s that

$$(1)\quad \left.\begin{aligned} F_1(x,T_{j+1})&=F_1(x,T_j)=1\\ F_2(x,T_{j+1})&=q_1F_2(x-\tau_{j+1},T_j)+(1-q_1)F_1(x-\tau_{j+1},T_j)\\ &\;\vdots\\ F_i(x,T_{j+1})&=q_{i-1}F_i(x-\tau_{j+1},T_j)+(1-q_{i-1})F_{i-1}(x-\tau_{j+1},T_j)\\ &\;\vdots\\ F_n(x,T_{j+1})&=q_{n-1}F_n(x-\tau_{j+1},T_j)+(1-q_{n-1})F_{n-1}(x-\tau_{j+1},T_j)\end{aligned}\right\} x\geqslant 0,$$

where it is understood that $F_i(x,T_j)=0$ for $x<0$ for all i,j.

[1] Manuscript received February 17, 1962.

The set of equations (1) may be written in the matrix form

$$F(x,T_{j+1})=\mathbf{A}F(x-\tau_{j+1},T_j)+\mathbf{f}_{j+1} \tag{2}$$

where $\mathbf{f}_{j+1}$ is the column vector with first element

$$F_1(x,T_{j+1})-F_1(x-\tau_{j+1},T_j)$$

and all other elements zero, and $\mathbf{A}=(a_{ij})$ is the matrix

$$\mathbf{A}=\begin{bmatrix} 1 & & & & & \\ 1-q_1 & q_1 & & & & \\ & 1-q_2 & q_2 & & & \\ & \cdot & \cdot & \cdot & \cdot & \\ & & & 1-q_{n-2} & q_{n-2} & \\ & & & & 1-q_{n-1} & q_{n-1} \end{bmatrix} \tag{3}$$

with zero elements in all positions not along the diagonal or lower off-diagonal. Taking Laplace transforms in (2) we obtain

$$\psi(\theta,T_{j+1})=\mathbf{A}\psi(\theta,T_j)e^{-\theta\tau_{j+1}}+\mathbf{L}_{j+1}(\theta), \tag{4}$$

where the transpose of $\mathbf{L}_{j+1}(\theta)$ is given by the row vector

$$\mathbf{L}'_{j+1}(\theta)=(1-e^{-\theta\tau_{j+1}},\ 0,\ .\ .\ .,\ 0)\quad j\geqslant 0.$$

On repeating the above process, we obtain for the time $t=T_m+0$

$$\psi(\theta,T_m)=\sum_{i=0}^{m} e^{-\theta(T_m-T_i)}\mathbf{A}^{m-i}\mathbf{L}_i(\theta) \tag{5}$$

where

$$\mathbf{L}'_0(\theta)\equiv\psi'(\theta,0)=(1,\ e^{-\theta x_2},\ .\ .\ .,\ e^{-\theta x_n}).$$

The matrix $\mathbf{A}$ may be written in canonical form as

$$\mathbf{A}=\mathbf{B}\mathbf{\Lambda}\mathbf{C}$$

where $\mathbf{\Lambda}$ is the diagonal matrix of roots $1,q_1,\ldots,q_{n-1}$ of $\mathbf{A}$, and $\mathbf{B}=(b_{ij})$ and $\mathbf{C}=(c_{ij})\equiv\mathbf{B}^{-1}$ are lower triangular matrices. Their elements are found to be

$$b_{i+j,i}=\prod_{k=1}^{j}\frac{(1-q_{i+j-k})}{(q_{i-1}-q_{i+j-k})}\qquad j\geqslant 0$$

$$c_{i+j,i}=\prod_{k=1}^{j}\frac{(1-q_{i+j-k})}{(q_{i+j-1}-q_{i+j-1-k})}\qquad j\geqslant 0,$$

where by defining $\prod_{k=1}^{0}=1$, we have that the diagonal elements are unity. The elements of $\mathbf{A}^m=(a_{ij}^{(m)})$ are clearly

$$a_{ij}^{(m)}=\begin{cases}\sum_{k=j}^{i} b_{ik}c_{kj}q_{k-1}^{m} & i\geqslant j\\ 0 & i<j \text{ and } i>j,\ i-j>m.\end{cases}$$

From equation (5) the elements of $\psi(\theta,T_m)$ may be written as

$$\psi_r(\theta,T_m)=\sum_{i=1}^{m} a_{r1}^{(m-i)}e^{-\theta(\tau_{i+1}+\cdots+\tau_m)}(1-e^{-\theta\tau_i})$$
$$+\sum_{i=1}^{n} e^{-\theta(\tau_1+\cdots+\tau_m)-\theta x_i}\,a_{ri}^{(m)}\qquad 1\leqslant r\leqslant n. \tag{6}$$

3. We are now in a position to deduce some of the properties of the process. Let us first consider the age distribution function $F_{rm}(x)$, of the r^{th} ranked element at the m^{th} regeneration point, regardless of the time taken for these m changes. We suppose that the $\{\tau_i\}$ are independently and identically distributed with d.f. $G(x)$. The Laplace transform $\psi_{rm}(\theta)$ of $F_{rm}(x)$ is then obtained by integrating $\psi_r(\theta,T_m)$ over all possible lengths of time required for m changes to occur, so that

$$\begin{aligned}
\psi_{rm}(\theta)
&=\int_{\tau_1=0}^{\infty}\cdots\int_{\tau_m=0}^{\infty}\psi_r(\theta,T_m)dG(\tau_1)\ldots dG(\tau_m)\\
&=\sum_{i=1}^{m}a_{r1}^{(m-i)}\int_{\tau_1=0}^{\infty}\cdots\int_{\tau_m=0}^{\infty}e^{-\theta(\tau_{i+1}+\cdots+\tau_m)}(1-e^{-\theta\tau_i})dG(\tau_1)\ldots dG(\tau_m)\\
&\qquad+\sum_{i=1}^{r}a_{ri}^{(m)}e^{-\theta x_i}\int_{\tau_1=0}^{\infty}\cdots\int_{\tau_m=0}^{\infty}e^{-\theta(\tau_1+\cdots+\tau_m)}dG(\tau_1)\ldots dG(\tau_m)\\
&=\sum_{i=0}^{m-1}a_{r1}^{(i)}\varphi(\theta)^i(1-\varphi(\theta))+\sum_{i=1}^{r}a_{ri}^{(m)}e^{-\theta x_i}\varphi(\theta)^m.
\end{aligned}\tag{7}$$

where $\varphi(\theta)$ is the Laplace transform of $G(x)$.

The result (7) is readily inverted to give

$$F_{rm}(x)=\sum_{i=0}^{m-1}a_{r1}^{(i)}\{G^{i*}(x)-G^{(i+1)*}(x)\}+\sum_{i=1}^{r}a_{ri}^{(m)}G^{m*}(x-x_i),\tag{8}$$

where $G^{m*}(x)$ is the m-fold convolution of $G(x)$, and $G^{m*}(x)=0$ for $x<0$. As an example, when the $\{\tau_i\}$ have a negative exponential distribution, then $F_{rm}(x)$ is a linear combination of gamma and truncated gamma-type distributions.

4. Let us now consider the problem of the age distribution of the r^{th} ranked element if the m^{th} change occurs in the time interval $(t,t+dt)$. If $\psi_{rm}(\theta,t)$ is the Laplace transform of the d.f. $F_{rm}(x,t)$, then

$$\begin{aligned}
&d_t\psi_{rm}(\theta,t)=\\
&\int_{\tau1=0}^{t}\int_{\tau2=0}^{t-T1}\cdots\int_{\tau_{m-1}=0}^{t-T_{m-2}}\psi_r(\theta,T_m)dG(\tau_1)\ldots dG(\tau_{m-1})dG(t-T_{m-1}).
\end{aligned}\tag{9}$$

Substituting for $\psi_r(\theta,T_m)$ from (6), and simplifying, this yields

$$\begin{aligned}
d_t\psi_{rm}(\theta,t)=\sum_{i=1}^{m}a_{r1}^{(m-i)}\int_{s=0-}^{t}&e^{-\theta(t-s)}\{d_tG^{(m-i)*}(t-s)dG^{i*}(s)\\
&-d_tG^{(m-i+1)*}(t-s)dG^{(i-1)*}(s)\}\\
&+\sum_{i=1}^{r}a_{ri}^{(m)}e^{-\theta t-\theta x_i}dG^{m*}(t).
\end{aligned}\tag{10}$$

As an example suppose $G(\tau_i)=1-e^{-\mu\tau_i}(\tau_i\geqslant 0)$; then (10) reduces to

$$d_t\psi_{rm}(\theta,t)=\frac{\mu^m t^{m-1}e^{-(\mu+\theta)t}dt}{(m-1)!}\left[\sum_{i=1}^{m} a_{r1}^{(m-i)}\{\beta_{i,m-i}(-\theta t) -\beta_{i-1,m-i+1}(-\theta t)\}+\sum_{i=1}^{r} a_{ri}^{(m)}e^{-\theta x_i}\right], \tag{11}$$

where

$$\beta_{a,b}(\theta)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_{x=0}^{1} e^{-\theta x}x^{a-1}(1-x)^{b-1}dx$$

is the Laplace transform of the Beta distribution

$$\beta_{a,b}(\theta)=\frac{\Gamma(a+b)}{\Gamma(a)}\sum_{k=0}^{\infty}\frac{\Gamma(a+k)(-1)^k\theta^k}{\Gamma(a+b+k)k!}.$$

5. Let us finally consider the stationary age distribution $F_r(x)(1\leqslant r\leqslant n)$ of the r^{th} ranked element directly after a regeneration point, where we assume for the moment that it exists. For any particular length τ of time between regeneration points, it will satisfy the matrix equation

$$\mathbf{F}(x)=\mathbf{A}\mathbf{F}(x-\tau)+\mathbf{W}_{j+1}, \tag{12}$$

where $\mathbf{W}_{j+1}$ is a column vector with one non-zero element $F_1(x)-F_1(x-\tau)$ in the first position, and $\mathbf{F}(x)$ is the column vector with elements $F_r(x)$. Let τ have a distribution function $G(\tau)$ with Laplace transform $\varphi(\theta)$, then the Laplace transform $\boldsymbol{\psi}(\theta)$ of elements $\{\psi_r(\theta)\}$ may be obtained from (12) as

$$\boldsymbol{\psi}(\theta)=[\mathbf{I}-\mathbf{A}\varphi(\theta)]^{-1}\mathbf{W}(\theta) \tag{13}$$

where $\mathbf{W}(\theta)'=(1-\varphi(\theta),\ 0,\ .\ .\ .,\ 0)$.

To obtain $\boldsymbol{\psi}(\theta)$ we require only the first column of $[\mathbf{I}-\mathbf{A}\varphi(\theta)]^{-1}=(d_{ij})$; this is given by

$$d_{11}=(1-\varphi(\theta))^{-1}$$

$$d_{j+1,1}=\prod_{i=1}^{j}\frac{(1-q_i)\varphi(\theta)^j}{(1-q_i\varphi(\theta))(1-\varphi(\theta))}\qquad 1\leqslant j\leqslant n-1,$$

so that

$$\begin{aligned}&\psi_1(\theta)=1\\ &\psi_r(\theta)=d_{r1}(1-\varphi(\theta))=\prod_{i=1}^{r-1}\frac{(1-q_i)\varphi(\theta)^{r-1}}{1-q_i\varphi(\theta)}\qquad 2\leqslant r\leqslant n.\end{aligned} \tag{14}$$

Equation (14) may be reduced to the sum of $r-1$ partial fractions; when all $p_i>0(1\leqslant i\leqslant n)$ we obtain

$$\begin{aligned}&\psi_1(\theta)=1\\ &\psi_r(\theta)=\left\{\prod_{i=1}^{r-1}(1-q_i)\right\}\sum_{i=1}^{r-1}\frac{\varphi(\theta)^{r-1}}{(1-q_i\varphi(\theta))\prod_{\substack{j=1\\ j\neq i}}^{r-1}(1-q_j/q_i)}\qquad 2\leqslant r\leqslant n.\end{aligned} \tag{15}$$

This may be inverted, for $\max_{1 \leqslant i \leqslant r-1} |q_i\varphi(\theta)| < 1$, in the form of an infinite series with rapidly decreasing terms (see §6):

$$F_1(x) = 1$$

$$F_r(x) = \left\{\prod_{i=1}^{r-1}(1-q_i)\right\}\sum_{i=1}^{r-1}\left\{\frac{1}{\prod_{\substack{j=1\\j\neq i}}^{r-1}(1-q_j/q_i)}\right\}\sum_{k=0}^{\infty} q_i^k G^{(k+j)*}(x) \quad 2 \leqslant r \leqslant n.$$

When $\varphi(\theta) = \mu(\mu+\theta)^{-1}$, then $F_r(x)$ is in the form of a finite linear sum of gamma-type distributions:

$$F_r(x) = \left\{\prod_{k=1}^{r-1}(1-q_k)\right\}\sum_{i=1}^{r-1}\left\{\frac{1}{\prod_{\substack{k=1\\k\neq i}}^{r-1}(1-q_k/q_i)}\right\}\left[\frac{b_{i0}(1-e^{-\mu(1-q_i)x})}{\mu(1-q_i)}\right.$$
$$\left.+\sum_{j=1}^{r-2}\frac{b_{ij}}{\mu^j}\left\{1-\sum_{k=0}^{j-1}e^{-\mu x}\frac{(\mu x)^k}{k!}\right\}\right] \qquad 2 \leqslant r \leqslant n,$$

with $F_1(x) = 1$, where the b_{ij} are constants obtained from partial fractioning of (15).

We now obtain this stationary distribution as the limit of $F_{rm}(x)$ as $m \to \infty$. By applying the theorem of Zygmund (1951) concerning the limit of characteristic functions, we have that a necessary and sufficient condition for the existence of a complete stationary age distribution is that $-\varphi'(0) < \infty$ and $p_n > 0$. Taking the limit of (8) as $m \to \infty$ we find that

$$\psi_1(\theta) = 1$$

$$\psi_r(\theta) = \varphi(\theta)^{r-1}\left\{1+\sum_{j=1}^{r-1}\frac{\prod_{i=1}^{r-1}(1-q_i)(1-\varphi(\theta))}{\{1-q_j\varphi(\theta)\}(1-q_j^{-1})\prod_{\substack{i=1\\i\neq j}}^{r-1}(1-q_i/q_j)}\right\} \qquad 2 \leqslant r \leqslant n,$$

which, when all $p_i > 0 (1 \leqslant i \leqslant n)$, is in fact equal to (15).

The moments of the stationary distribution may be found from (14) by differentiation; for example the mean is

$$(17) \qquad -\psi_r'(0) = \rho\left\{r-1+\sum_{j=1}^{r-1}q_j(1-q_j)^{-1}\right\}$$

where $\rho = -\varphi'(0)$ is the mean time between regeneration points.

6. The following numerical examples will illustrate some of the points previously made. If there are four elements with probabilities $p_1 = 0{\cdot}1$, $p_2 = 0{\cdot}2$, $p_3 = 0{\cdot}3$, $p_4 = 0{\cdot}4$ of replacement at any regeneration point, we obtain for $\mathbf{A}^4$ and $\mathbf{A}^8$:

$$\mathbf{A}^4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ {\cdot}9999 & {\cdot}0001 & 0 & 0 \\ {\cdot}9891 & {\cdot}0280 & {\cdot}0081 & 0 \\ {\cdot}6880 & {\cdot}2044 & {\cdot}1620 & {\cdot}1296 \end{bmatrix} \quad \mathbf{A}^8 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ {\cdot}9997 & {\cdot}0002 & {\cdot}0001 & 0 \\ {\cdot}9299 & {\cdot}0310 & {\cdot}0223 & {\cdot}0168 \end{bmatrix}$$

of which only the first columns are relevant to the approach to the stationary distribution. For a process with eight elements having probabilities $p_1=0\cdot02$, $p_2=0\cdot03$, $p_3=0\cdot05$, $p_4=0\cdot08$, $p_5=0\cdot12$, $p_6=0\cdot17$, $p_7=0\cdot23$, $p_8=0\cdot30$ of replacement the first column of $\mathbf{A}^8$ and $\mathbf{A}^{12}$ are given respectively by

$$\begin{array}{l}(1,\ 1,\ 1,\ 1,\ 0\cdot9995,\ 0\cdot9763,\ 0\cdot7457,\ 0\cdot4318)\\(1,\ 1,\ 1,\ 1,\ \quad 1 \quad,\ 0\cdot9998,\ 0\cdot9845,\ 0\cdot9458).\end{array}$$

For these examples, and for others of a similar type which might be found in practice, it is seen that the coefficients $a_{r1}^{(i)}=0$ for $r<i-1$ we have that $a_{r1}^{(i)}$ reaches $0\cdot99$ within approximately n steps, so that the stationary age distribution may be determined accurately from the limit of (8) by taking approximately n terms of the first summation of (8).

We are indebted to the referee for his suggestion which led to a great improvement in Section 4.

Reference

Zygmund, A. (1951). "A remark on characteristic functions." Proceedings of the Second Berkeley Symposium, pp. 369–372.

AUSTRALASIAN MEDICAL PUBLISHING CO. LTD.
SEAMER AND ARUNDEL STS., GLEBE, SYDNEY