LIBRARY OF THE UNIVERSITY OF ADELAIDE
20.7.67

A MATHEMATICAL INVESTIGATION OF THE CONTROL OF TRAFFIC AT AN INTERSECTION

by

MICHAEL C. DUNNE
B.Sc (Hons.)(Adel.)

Thesis submitted for the degree of Doctor of Philosophy in the University of Adelaide, Department of Mathematics, April, 1966.

TABLE OF CONTENTS

SUMMARY

It is the purpose of this thesis to derive the operational characteristics of a signalized intersection controlled by a computer which actuates change of light phase when the competing queue lengths satisfy certain linear relationships. These linear relationships together with minimum and maximum phase time restrictions constitute the "control algorithm" by which the traffic is governed. In Chapter 2 there is given a qualitative discussion of the operation of the intersection under the control of the algorithm.

The case of constant arrival rates is analyzed in Chapter 3. In particular, expressions are derived for the number of vehicles queued in a given arm, the duration of a given phase and so on. Other important topics mentioned are stability and the related question of convergence, delay and the choice of the parameters involved in the control algorithm.

As a central computer may receive data in a binary form the case in which the arriving vehicles are supposed generated by a binomial process is considered in Chapter 4. Analytic expressions are found for the probability generating functions for queue length, phase duration etc. It is shown that a state of statistical equilibrium may be achieved and an investigation is made of the system in this state. The most important operational characteristics of the system are derived, the question of total and average delay is investigated and,

to illustrate the theoretical results obtained in the thesis, a numerical example is given.

Some general aspects of the problem and possible extensions to the theory are discussed in the final chapter.

SIGNED STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge and belief the thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

(Michael C. Dunne)

ACKNOWLEDGEMENT

The author acknowledges the generous support of General Motors-Holden's Pty. Ltd., which has enabled this research project to be carried out. The author is indebted to his supervisor, Professor R. B. Potts, for suggesting such a fruitful line of research and for his encouragement and assistance while the work was being undertaken and to Dr. J. N. Darroch, University of Michigan, for his suggestions which enabled Chapter 4 to be presented in a more refined form. The author has benefited from discussions with Professor G. F. Newell, University of California (Berkeley), and Mr. P. G. Pak-Poy and Mr. R. W. J. Morris of the South Australian Highways Department.

LIBRARY OF THE UNIVE

CHAPTER 1

INTRODUCTION

Most of the papers published in the field of intersection control have been concerned with the derivation of the optimum settings of a fixed cycle traffic light; i.e. with the determination of that cycle time and phase split which optimize the operation of the intersection in some sense. The most widely used criterion is that of minimum average delay per vehicle. The complexity of the mathematics involved in such determinations depends, of course, on the nature of the model being investigated. Unless some simplifying assumptions are made concerning the nature of the traffic stream and of the arrival and departure distributions of the vehicles the problem seems completely intractable. The mathematical analysis is greatly simplified by assuming that the traffic is of a continuous rather than discrete nature and that the traffic arrives at and departs from the intersection at a constant (or uniform) rate. The assumption of random arrivals is also often made following a suggestion by Adams [1] in 1936. The validity of the assumption of random arrivals has been discussed, for example, by Pak-Poy [2].

In 1941 Clayton [3] derived several formulae for cycle time, capacity, and delay at a signalized intersection on the basis of very simple models. Included in these was a formula for the minimum cycle time compatible with the condition that there should be no overflow from one cycle to the next. Clayton [3] also evaluated the average delay to a light flow of vehicles on a minor road at vehicle actuated signals under the assumption of random arrivals. Garwood [4] considered this same problem and generalized it somewhat by assuming the presence of a maximum allowable phase time for vehicles on the major road.

The case of constant arrival rates has also been considered by Wardrop [5] who derived formulae for the average delay to vehicles on each approach and for the average delay to all vehicles using the intersection. From these formulae he determined the cycle times and phase splits which minimize the average delay per vehicle. By computer simulation Wardrop has also obtained delays at a fixed-cycle traffic light for the case of random arrivals and steady departures. The results show good agreement with Clayton's formula for low flows but give much larger average delays as the degree of saturation increases.

This situation, where the arrivals are random and the departures regularly spaced, has been investigated in detail by Webster [6] who obtained an expression for average delay in the form of a theoretical term and an empirical term. This empirical term was reached by fitting curves to data obtained by computer simulation. The work of Webster has provided a sound theoretical and practical basis for the setting of fixed cycle signals. Webster found the optimum cycle time in this case to be approximately double the so-called minimum cycle (as derived by Clayton [3]) which is, in most cases, optimal for steady arrivals. In fact, in the case of random arrivals the minimum cycle is associated with infinite delay as the statistical fluctuations cause a steady build up in the queue resulting from cycle to cycle overflow. More recently Miller [7] has obtained formulae for the optimum cycle times and phase splits in terms of the variance/mean ratio of counts of arrivals per cycle on a given approach. In the particular case of random arrivals the results of Miller are in broad agreement with those of Webster.

Beckman, McGuire and Winsten [8] have analyzed the case of fixed time signals in which the arrivals are generated by a binomial process and obtained an expression for average delay in terms of $E(N_r)$, an unknown average queue at the start of the red time. Newell [9] has obtained an analytic expression for $E(N_r)$ under equilibrium conditions for this model. The case of binomial arrivals applied to a light system in which the control strategy is to switch the lights when the favoured queue empties has been investigated by Dunne and Potts [10] who obtained exact analytic expressions for the probability generating functions for the queue lengths at change of phase, phase times and so on. This subject is developed fully in Chapter 4 of this thesis. Another attempt at an analytic solution to a fixed-cycle traffic light problem was made by Newell [11] who proposed a model in which arrival headways were independent identically distributed random variables with a more or less arbitrary distribution and departures were regularly spaced during the green. He found it possible to evaluate delays approximately provided that the average queue length at the start of the red period was small compared with the average queue length at the start of the green. He showed also that Clayton's formula was a good approximation for any reasonable arrival distribution provided that the flow was not too close to saturation. Darroch [12] has found a formal solution for the stationary distribution of queue length at a fixed cycle traffic light for a fairly general distribution of arrivals and for regularly spaced departures. In the case of statistical equilibrium he also found bounds for $E[X_g]$, the expected value of queue length at the end of the green phase, and for E[D] the expected value of delay.

In another paper, Newell [13] has described some approximate methods for obtaining estimates of queue lengths, delays, etc. For reasonably large queues these estimates of the delays are correct to within a few percent and the method has the advantage that the relevant quantities are much more readily obtained than from the many exact formulae. The paper also contains a brief review of some of the more important recent papers on delays at fixed cycle and vehicle-actuated lights.

Another paper on fixed-cycle lights is that of Uematu [14] who uses a random walk model for apportioning the red and green phases. The criterion used is the minimization of the probability that the queue lengths will exceed a certain maximum allowable length (e.g. the block length).

Tanner [15] has calculated the delays that occur when two opposing streams of vehicles are trying to pass along a length of road only wide enough for one vehicle at a time. This problem is more applicable to the calculation of the delays to pedestrians crossing a road, to minor road vehicles at uncontrolled intersections, to vehicles at uncontrolled intersections where there is no absolute right of way and to intersecting streams of pedestrians. However, with some slight modifications to Tanner's model and some changes in his notation, the problem he has considered is very similar to that considered by Darroch, Newell and Morris [16]. The latter have given a very complete analysis based on the broad assumption of random lost time at each switch of the lights and arbitrary distribution of departure headways, but with the specific assumption of arrival headways being exponentially distributed random variables. As pointed out by them, this assumption of Poisson arrivals, allowing relatively high probabilities for small headways, leads to the unrealistic possibility of a flow during the extension of the green phase which is higher than the saturation flow through the intersection when the queue discharges. In their paper Darroch, et al. obtained estimates for the optimal unit extension on the basis of the usual criterion of minimizing the average delay per vehicle. This work was a follow-up to that of Grace, Morris, and Pak-Poy [17] who used continuum approximations in the construction of models to describe the behavior of vehicle-actuated traffic signals.

With the advent of central computer control of traffic lights the spectrum of possible strategies for the switching of lights has become unlimited, but it is far from obvious which strategies are likely to be most efficient. As indicated above, little more has been done beyond the calculation of some optimal settings for the familiar fixed cycle and vehicle-actuated lights. In many of those cases where exact solutions have been obtained it has proved almost impossible to retrieve any information of a practical nature. In other cases the mathematical difficulties encountered because of the complexity of the model have been insurmountable.

Miller [18] has described some of the main features of computer control systems and has examined one possible form of control, timing control, in detail. This paper is especially noteworthy for the practical points which are raised in it. In particular, Miller indicates the need for an evaluation of the operational characteristics of any intended control system in order to justify its installation, the need for fail-safe devices to guard against the possibility of a failure in the system and the need for detectors which are able to collect accurate information concerning the state of the traffic.

In an attempt to analyze possible algorithms for adaptive control by a computer, Dunne and Potts [19] have considered a simple algorithm and derived its operational characteristics on the assumption of constant arrival and departure rates. These authors have also considered [10] the case in which the departures are regularly spaced and the arrivals are generated by a binomial process. In the latter case the strategy of switching the lights when the favoured queue empties has been analyzed and exact probability generating functions obtained for the phase and cycle times, and for the queue lengths when the lights are switched. These generating functions readily yield a description of the operational characteristics of the control. A detailed investigation of these models is contained in this thesis.

The type of computer control described above has the advantage that it is dynamic rather than static in that it reacts directly to the traffic demand at any instant. As is the case with computerized systems, there is the added advantage that past history can be used in determining possible variations of a particular strategy and that the computer can, if required, perform calculations of a complex nature in a very short period of time.

As there is little theoretical work available to provide a sound basis for the use of computers in the field of intersection control, an exact treatment of these simple models represents a significant step.

CHAPTER 2

CONTROL ALGORITHM

2.1 Intersection Model

Consider a single intersection (Fig. 1) controlled by a two phased traffic light that can be switched from one phase to the next by command from a central computer or master controller. For simplicity attention will be confined to only 2 of the 4 arms, and it will be assumed that each of the 2 arms is of one lane only, and further that right- and left-hand turners are prohibited. This simple model is similar to one used in the Toronto pilot study [20] and is realistic for a properly designed two phase intersection because 2 lanes usually predominate in determining the light settings, especially under tidal flow conditions at peak periods. Right (or left, as the case may be) turning vehicles either filter through the oncoming traffic or make their turn during the amber period, while left (right) turners can turn fairly freely provided that the pedestrian traffic is light.

It will be assumed that each green phase consists of a lost time, during which no vehicles cross the intersection, followed by an effective green time, during which the traffic discharges at a constant rate (the so-called saturation flow).

2.2 Control Functions

At time t (measured from an arbitrary origin), suppose that

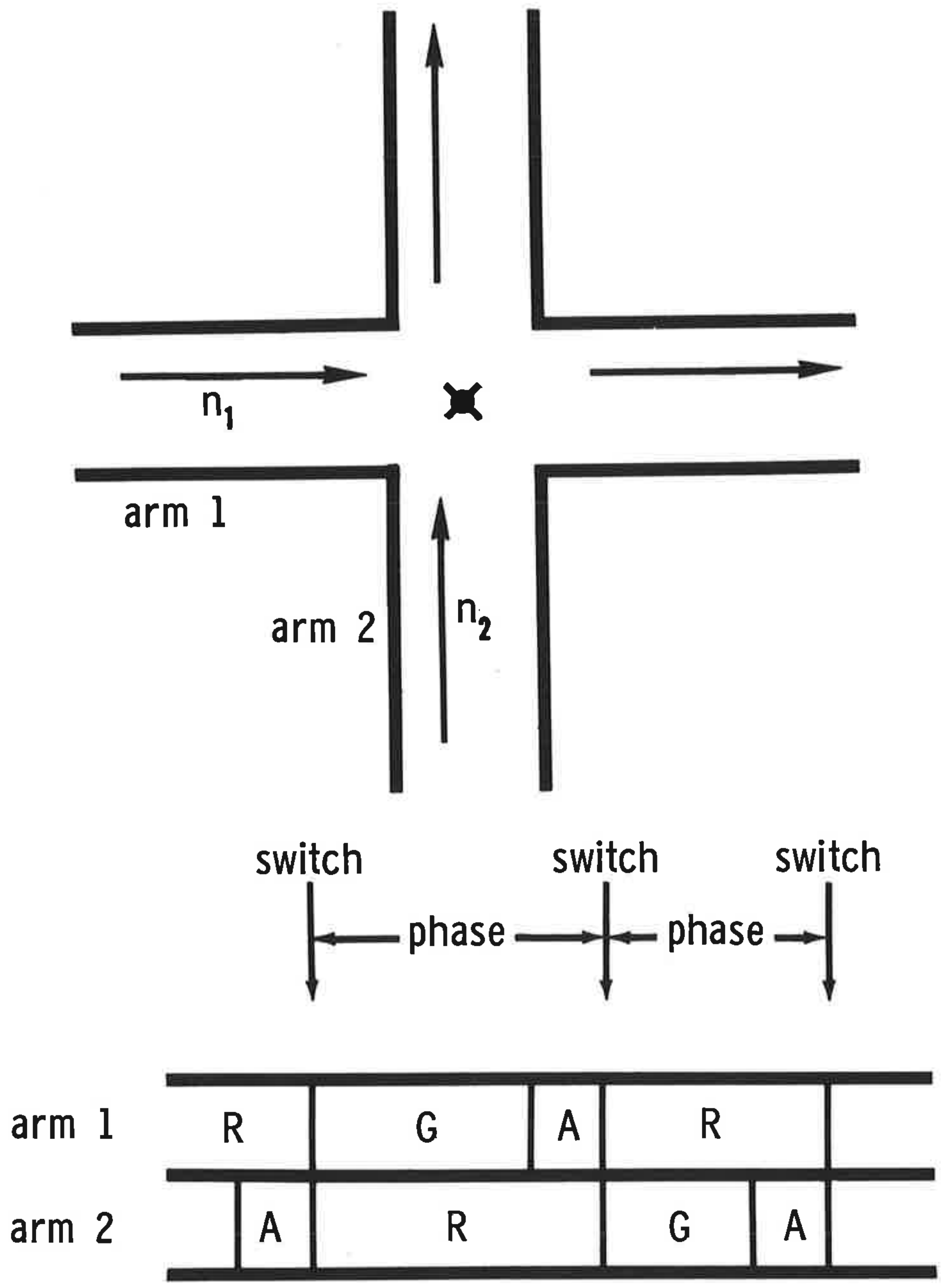


Fig. 1. Single intersection controlled by a two-phased traffic light that can be switched from one phase to the next by remote computer control.

$$n_i(t) = \text{number of vehicles queued in arm } i, \ (i = 1,2) \tag{2.1}$$

$$f_1(t) = \alpha_1 n_1(t) + \beta_1 - n_2(t) \ , \tag{2.2}$$

$$f_2(t) = \alpha_2 n_2(t) + \beta_2 - n_1(t) \ , \tag{2.3}$$

$$T(t) = \text{time elapsed since last change of phase,} \tag{2.4}$$

$$r_i, \ R_i = \text{minimum, maximum allowable durations of phase for arm } i \ . \tag{2.5}$$

The control functions $f_i(t)$ are defined in terms of control constants α_i, β_i which are chosen in the range

$$\alpha_i > 1 \ , \ \beta_i > 0 \ . \tag{2.6}$$

The decision as to whether the phase of the lights should or should not be changed at a given time is governed by the magnitude of the time dependent quantities defined in equations (2.1) - (2.4) subject to the added conditions that no phase can be shorter than a given minimum, r_i, or longer than a maximum, R_i.

2.3 Flow Diagram

The operation of the control algorithm is best illustrated by means of a flow diagram (Fig. 2) that indicates that a change of phase

(i) never occurs if the phase duration is less than the minimum allowable phase time,

(ii) always occurs if the phase is greater than or equal to the maximum,

and further that, if the phase is greater than the minimum and less than

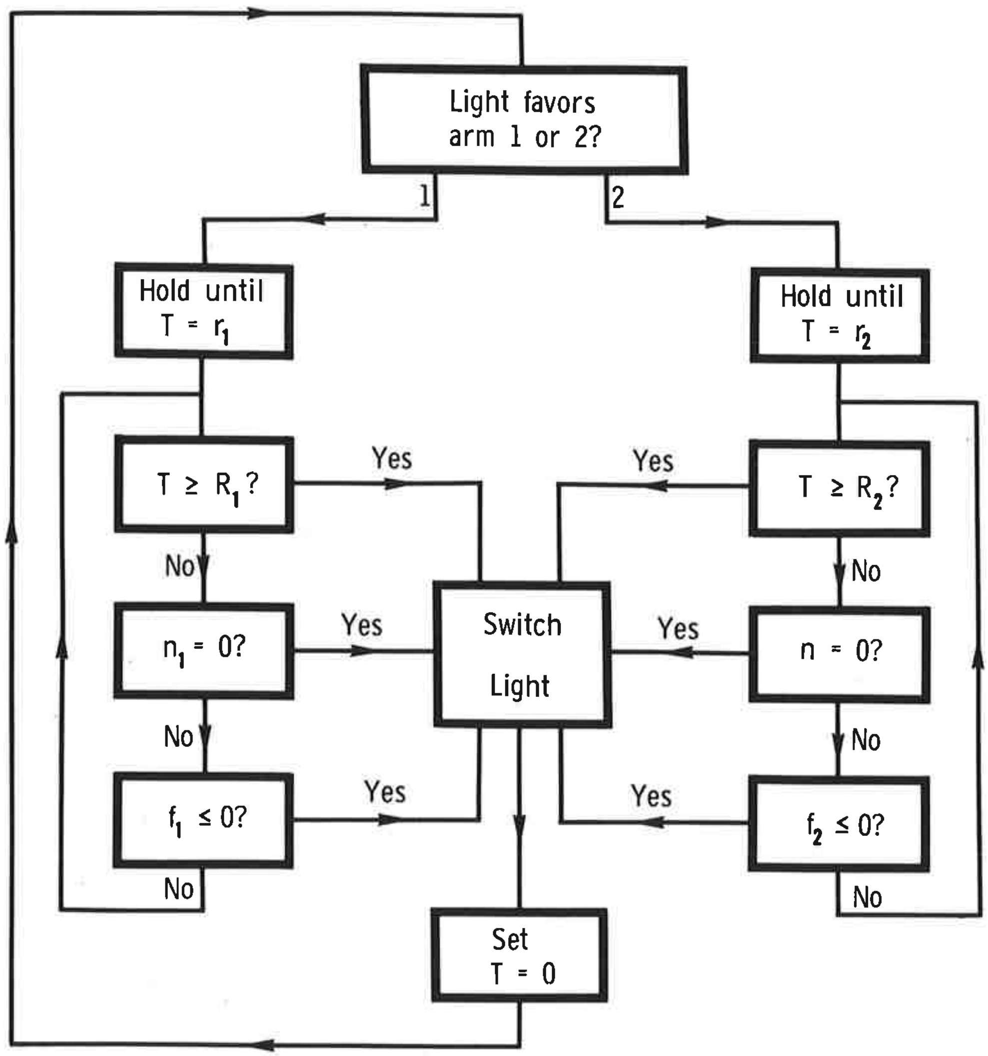


Fig. 2. Flow diagram of control algorithm for switching traffic light. T is the time since the last change of phase, r_i, R_i minimum and maximum allowable phases, n_i, the number of vehicles in queue, f_i, the control functions.

the maximum the lights are switched

(iii) if the queue being favoured is emptied, or

(iv) if there are vehicles in the queue and $f_i(t)$ is negative or zero.

If, as is likely in a practical application, the computer operates only for part of the time it is necessary to incorporate ~~of~~ a "starting" mechanism. This could be accomplished either by a human operator or by the computer itself acting on the basis of the information concerning flow rates which it receives from the detectors. Similarly, to switch back to the local controller, a "stopping" mechanism is necessary.

As the flow diagram is shown, the intersection is scanned continuously. It may however be better (e.g. more economical) to scan the intersection only at certain discrete intervals of time. The length of the scanning interval is somewhat arbitrary but if, for example, the saturation flows are 1,800 v.p.h. (i.e. 1 vehicle every 2 seconds), a scanning interval of the order of 1/2 second seems reasonable.

2.4 Random Walks

The manner in which the queue lengths vary is best described in the theory of random walks on a lattice with reflecting barriers. This is illustrated in Fig. 3 in which $n_2(t)$ is plotted against $n_1(t)$. As vehicles depart from and arrive at the queues the representative point on this graph traces out a "walk" from lattice point to lattice point. If arm 1 is being favoured, $n_1(t)$ can either increase by one

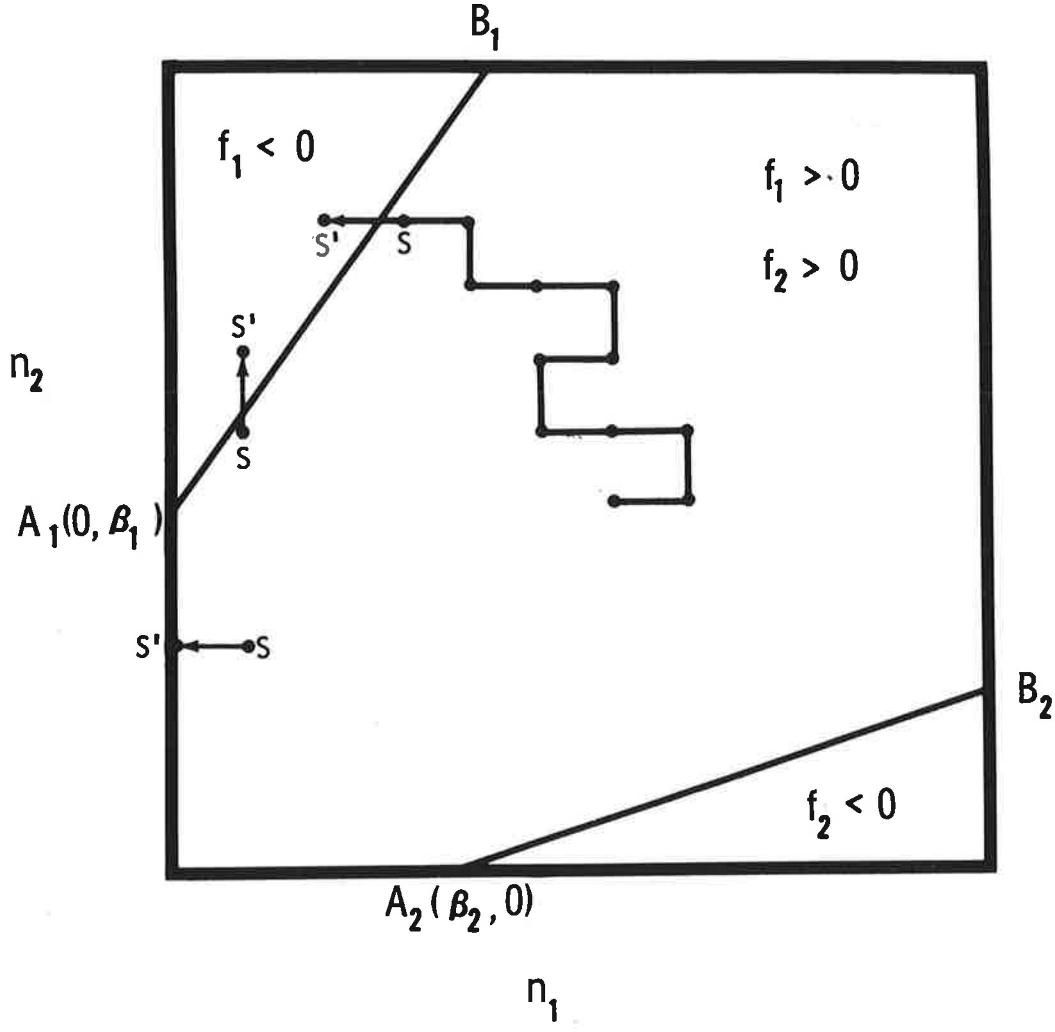


Fig. 3. Graph of n_2 against n_1 for a typical case when arm 1 has the green light. The representative point performs a walk on the lattice and OA_1B_1 (and similarly OA_2B_2) forms a reflecting barrier at which the light is switched in accord with the control algorithm. The three steps SS' shown are the possible last steps before reflection.

(vehicle arrives in arm 1) or decrease by one (vehicle from arm 1 crosses the intersection), while $n_2(t)$ can only increase by one (vehicle arrives in arm 2). The corresponding steps of the walk are to the right, to the left, or up. If the intersection is undersaturated vehicles in arm 1 will depart at a greater rate than they arrive so that the trend will be for $n_1(t)$ to decrease while, since no vehicles depart from arm 2, $n_2(t)$ will steadily increase, a typical walk being shown in Fig. 3.

According to the control algorithm, the lines OA_1 (i.e. $n_1 = 0$) and A_1B_1 ($f_1 \equiv \alpha_1 n_1 + \beta_1 - n_2 = 0$) are "reflecting barriers" for the walk since, when $n_1 = 0$ or $f_1 \leq 0$ the light is switched and the nature of the walk is altered. There are three distinct types of possible last steps before reflection, these being labelled SS' in the diagram. As the time t is not represented explicitly in Fig. 3, that part of the algorithm concerned with minimum and maximum phase duration cannot be illustrated, although it can usually be interpreted as the number of steps during a phase being confined to lie between a lower and upper limit.

2.5 Computer Control

To enable a central computer to control the traffic light it is necessary, as in the Toronto System [20], for the computer to have an accurate clock and for communication links to be provided between computer and light and traffic detectors so that

(i) the computer can monitor the state of the light,

(ii) the computer can determine from its clock the time that has elapsed since the last switch of the light, and

(iii) the computer can carry out the necessary computations to test whether the lights should be switched, and if so, send out the pulse required to change the signal. In practice the intersection would be scanned at regular intervals of time and the computations would be made on the basis of the information received during each scanning. The constants α_i, β_i, r_i, R_i occuring in the control algorithm may be considered as predetermined parameters of the intersection or as quantities that the computer varies according to the existing traffic conditions.

In later chapters the way in which the operational characteristics of the intersection are affected by the choice of these constants will be discussed in detail.

CHAPTER 3

CONSTANT ARRIVAL AND DEPARTURE RATES

3.1 Continuum Model

The mathematical analysis of the behavior of the traffic under the control algorithm is complicated because $n_i(t)$ are integral stochastic variables that depend on the probability distributions describing the arrival and departure of vehicles. An exact solution can be found if it is assumed that the arrival and departure rates are constant and that the $n_i(t)$ are continuous variables (c. f. Newell [13]). This continuum model may be considered as a first approximation to the actual behavior of the traffic.

3.2 Notation

Following the notation of Webster [6], let

q_1, q_2 = vehicular flow (arrival rates) in the two arms, (3.1)

s_1, s_2 = saturation flow (departure rates), (3.2)

$y_1, y_2 = q_1/s_1,\ q_2/s_2$, (3.3)

$Y = y_1 + y_2$, (3.4)

ℓ = lost time for a single phase , (3.5)

g_j = effective green time for the jth phase. (3.6)

If time t is measured from the beginning of the phase j = 1, then $t_j = j\ell + g_1 + g_2 + \ldots + g_j$ is the time at the end of the jth phase since each phase consists of a lost time followed by an effective green time. For simplicity of notation it will be supposed that:

$n_i(t_j) = n_i(j)$ = number of vehicles in arm $i(i = 1,2)$ at the end of the jth phase. (3.7)

3.3 Continuous Walks

Under the assumption that q_i and s_i are constant and that n_i is continuous rather than discrete, the random walks of Fig. 3 are replaced by the straight line continuous walks of Fig. 4, a typical walk being represented by $W_0L_1W_1L_2W_2L_3W_3L_4W_4$ ...... with $[n_1(j), n_2(j)]$ being the coordinates of W_j. The starting point W_0 with coordinates $[n_1(0), n_2(0)]$ has been chosen arbitrarily, and without loss of generality it has been assumed that the first phase favours arm 1. W_0L_1 represents the increase in n_1 and n_2 during the lost time ℓ, and L_1W_1 the decrease in n_1 and increase in n_2 during the effective green time g_1. The light is first switched when the walk reaches W_1, i.e. when the control function f_1 becomes zero. Then follows a lost time, an effective green time for arm 2 and so on. All segments W_jL_{j+1} have length $\ell(q_1^2 + q_2^2)^{1/2}$ and slope q_2/q_1. Segments $L_{2j}W_{2j}$ have lengths $g_{2j}[q_1^2 + (s_2 - q_2)^2]^{1/2}$ and slopes of magnitude $(s_2 - q_2)/q_1$; while segments $L_{2j+1}W_{2j+1}$ have lengths $g_{2j+1}[q_2^2 + (s_1 - q_1)^2]^{1/2}$ and slopes of magnitude $q_2/(s_1 - q_1)$. The minimum and maximum phase duration restrictions imply that a segment such as L_jW_j must be longer than a certain minimum length and shorter than a certain maximum length. As illustrated, the walk may enter the region bounded by OA_1 and OA_2 which become reflecting

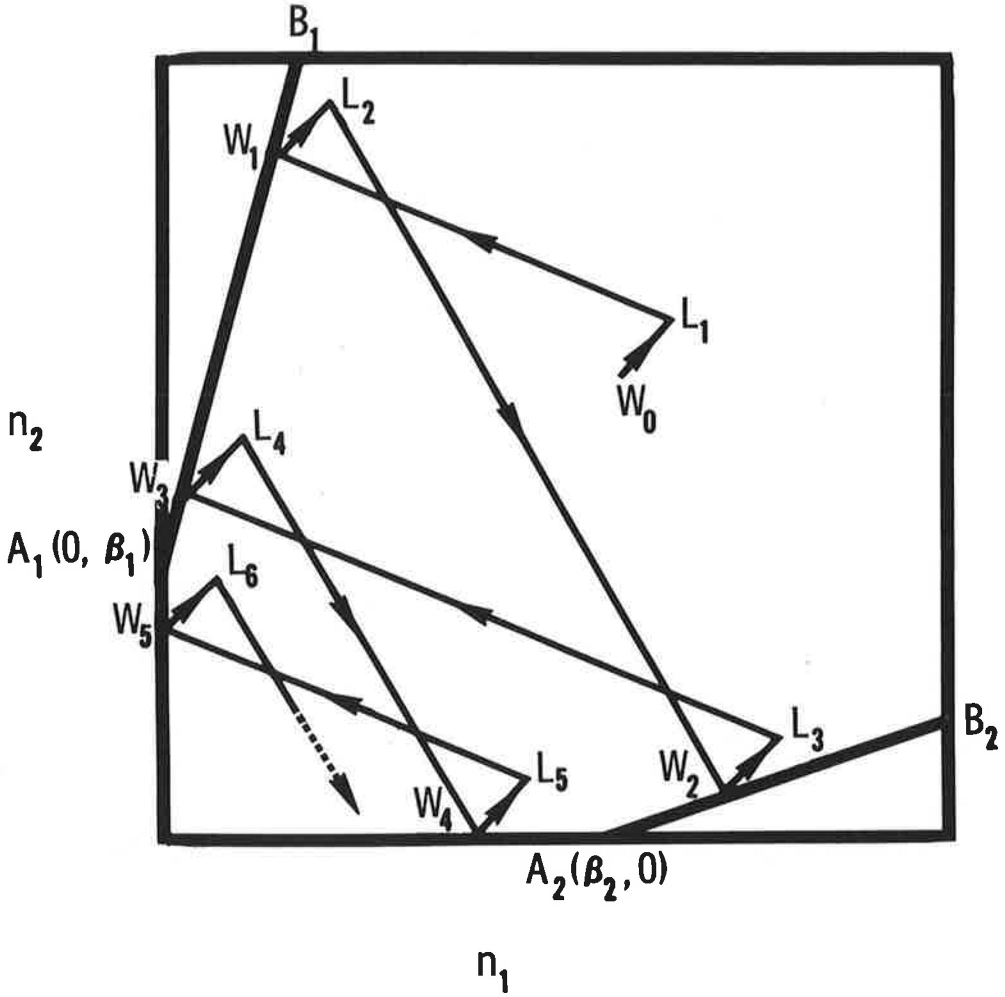


Fig. 4. Graph of n_2 against n_1 for the continuum model with constant arrival and departure rates. $W_0L_1W_1...$ represents a typical walk that has first A_1B_1, A_2B_2 and then OA_1, OA_2 as reflecting barriers.

barriers corresponding to $n_1 = 0$ and $n_2 = 0$ respectively and the light is switched at W_4, W_5, etc.

3.4 Control Equations

Since q_i, s_i are constant the fundamental equations for $n_i(j)$ may be written down immediately assuming initially that there are $n_i(0)$ vehicles in arm i and that arm 1 is favoured:

$$n_1(2j-1) = n_1(2j-2) - g_{2j-1}(s_1-q_1) + \ell q_1 \ , \tag{3.8}$$

$$n_2(2j-1) = n_2(2j-2) + g_{2j-1}q_2 + \ell q_2 \ , \tag{3.9}$$

$$n_1(2j) = n_1(2j-1) + g_{2j}q_1 + \ell q_1 \ , \tag{3.10}$$

$$n_2(2j) = n_2(2j-1) - g_{2j}(s_2-q_2) + \ell q_2 \ , \tag{3.11}$$

where j = 1, 2, 3, ... denotes the phase number. These equations form a set of first order simultaneous difference equations whose solution depends on the particular type of control that determines phase change of the lights.

3.5 Solution of the Control Equations

The control algorithm consists effectively of three types of control, all of which may be relevant at different stages. The solution for the control $T = R_i$ is trivial as the queue lengths either increase or decrease (depending on the relative magnitudes of the arrival and departure rates) by a constant amount each cycle. As α_i tends to infinity the reflecting barriers for $f_i = 0$ control tend to those for the control $n_i = 0$ control, i.e. to the coordinate axes. Thus the solutions under $n_i = 0$ control may be obtained from those under $f_i = 0$ control by allowing α_i to tend to infinity. For these reasons only the solutions for $f_i = 0$ control will be derived in any detail. To indicate the method of solution $n_1(2j-1)$ will be evaluated.

The equations to be solved are (3.8) - (3.11) together with

$$n_2(2j-1) = \alpha_1 n_1(2j-1) + \beta_1 , \tag{3.12}$$

$$n_1(2j) = \alpha_2 n_2(2j) + \beta_2 . \tag{3.13}$$

Elimination of g_{2j-1} from (3.8) and (3.9), of g_{2j} from (3.10) and (3.11) and of n_2 by means of (3.12) and (3.13) leads to the pair of equations

$$n_1(2j-1) = \frac{\alpha_2 q_2 + s_1 - q_1}{\alpha_2[\alpha_1(s_1-q_1)+q_2]} n_1(2j-2) + \frac{\alpha_2 \ell s_1 q_2 - (s_1-q_1)(\alpha_2\beta_1+\beta_2)}{\alpha_2[\alpha_1(s_1-q_1)+q_2]} , \tag{3.14}$$

$$n_1(2j) = \frac{\alpha_2[\alpha_1 q_1 + s_2 - q_2]}{\alpha_2(s_2-q_2)+q_1} n_1(2j-1) + \frac{q_1(\alpha_2 \ell s_2 + \alpha_2\beta_1 + \beta_2)}{\alpha_2(s_2-q_2)+q_1} . \tag{3.15}$$

Substituting for $n_1(2j-2)$ [obtained by replacing j by j-1 in (3.15)] in (3.14) gives

$$n_1(2j-1) = \frac{(\alpha_2 q_2 + s_1 - q_1)(\alpha_1 q_1 + s_2 - q_2)}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]} n_1(2j-3)$$

$$+ \frac{q_1(\alpha_2 q_2 + s_1 - q_1)(\alpha_2 \ell s_2 + \alpha_2\beta_1 + \beta_2)}{\alpha_2[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]} + \frac{\alpha_2 \ell s_1 q_2 - (s_1-q_1)(\alpha_2\beta_1+\beta_2)}{\alpha_2[\alpha_1(s_1-q_1)+q_2]} . \tag{3.16}$$

The terms on the right hand side which are independent of j yield

$$\frac{(\alpha_2\beta_1+\beta_2)[q_1q_2-(s_1-q_1)(s_2-q_2)] + \ell[s_2q_1(\alpha_2q_2+s_1-q_1)+s_1q_2(\alpha_2(s_2-q_2)+q_1)]}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]}$$

$$= \frac{s_1s_2(\alpha_2\beta_1+\beta_2)(Y-1)+s_1s_2\ell[\alpha_2q_2(1+y_1-y_2)+q_1(1-y_1+y_2)]}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]}$$

$$= \frac{s_1s_2(1-Y)\{\alpha_2[\ell q_2 \frac{1+y_1-y_2}{1-Y} - \beta_1]+[\ell q_1 \frac{1-y_1+y_2}{1-Y} - \beta_2]\}}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]}$$

$= \delta$, say .

Hence

$$n_1(2j-1) = an_1(2j-3)+\delta \ , \qquad (3.17)$$

where

$$a = \frac{(\alpha_2q_2+s_1-q_1)(\alpha_1q_1+s_2-q_2)}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]} \ . \qquad (3.18)$$

Substituting the trial solution

$$n_1(2j-1) = (A/\alpha_1)a^{j-1} + \gamma \qquad (3.19)$$

in (3.17) leads to

$$(A/\alpha_1)a^{j-1} + \gamma = a[(A/\alpha_1)a^{j-2} + \gamma] + \delta$$

so that

$$\gamma = \delta(1-a)^{-1} \ .$$

Now

$$1-a = 1 - \frac{(\alpha_2 q_2+s_1-q_1)(\alpha_1 q_1+s_2-q_2)}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]} ,$$

$$= \frac{s_1 s_2(1-Y)(\alpha_1\alpha_2-1)}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]} .$$

Hence,

$$\gamma = (\alpha_1\alpha_2-1)^{-1}\{\alpha_2[\ell q_2 \frac{1+y_1-y_2}{1-Y} - \beta_1]+[\ell q_1 \frac{1-y_1+y_2}{1-Y} - \beta_2]\} = n_1(\infty) \quad .(3.20)$$

To evaluate A, set j=1 in (3.19) giving

$$n_1(1) = (A/\alpha_1) + \gamma = (A/\alpha_1) + n_1(\infty) ,$$

so that

$$(A/\alpha_1) = n_1(1) - n_1(\infty) .$$

Setting j=1 in equations (3.8), (3.9) and (3.12) yields

$$n_1(1) = n_1(0) - g_1(s_1-q_1) + \ell q_1 ,$$

$$n_2(1) = n_2(0) + g_1 q_2 + \ell q_2 ,$$

$$n_2(1) = \alpha_1 n_1(1) + \beta_1 .$$

Proceeding as in the general case leads to

$$n_1(1) = \frac{q_2 n_1(0)+(s_1-q_1)n_2(0)+\ell s_1 q_2-(s_1-q_1)\beta_1}{[\alpha_1(s_1-q_1)+q_2]} .$$

Substituting for $n_1(\infty)$ from (3.20) gives

$$(A/\alpha_1)[\alpha_1(s_1-q_1)+q_2] - q_2n_1(0) - (s_1-q_1)n_2(0)$$

$$= \ell s_1q_2-(s_1-q_1)\beta_1 - \frac{[\alpha_1(s_1-q_1)+q_2]}{(\alpha_1\alpha_2-1)}[\alpha_2(\ell q_2 \frac{1+y_1-y_2}{1-Y} - \beta_1) + (\ell q_1 \frac{1-y_1+y_2}{1-Y} - \beta_2)] ,$$

$$= \ell s_1q_2 - (s_1-q_1)\beta_1 - \frac{(\alpha_2q_2+s_1-q_1)}{(\alpha_1\alpha_2-1)}[(\ell q_2 \frac{1+y_1-y_2}{1-Y} - \beta_1) + \alpha_1(\ell q_1 \frac{1-y_1+y_2}{1-Y} - \beta_2)] - (s_1-q_1)(\ell q_2 \frac{1+y_1-y_2}{1-Y} - \beta_1) + q_2(\ell q_1 \frac{1-y_1+y_2}{1-Y} - \beta_2) ,$$

$$= - (\alpha_2q_2+s_1-q_1)n_2(\infty) - \beta_2q_2 ,$$

$$= - q_2N_1(\infty) - (s_1-q_1)n_2(\infty) ,$$

where

$$n_2(\infty) = (\alpha_1\alpha_2-1)^{-1}[(\ell q_2 \frac{1+y_1-y_2}{1-Y} - \beta_1) + \alpha_1(\ell q_1 \frac{1-y_1+y_2}{1-Y} - \beta_2)]$$

and

$$N_1(\infty) = \alpha_2n_2(\infty) + \beta_2 .$$

Hence

$$(A/\alpha_1) = \frac{q_2[n_1(0)-N_1(\infty)]+(s_1-q_1)[n_2(0)-n_2(\infty)]}{[\alpha_1(s_1-q_1)+q_2]} . \qquad (3.21)$$

Equations (3.20) and (3.21) complete the solution for $n_1(2j-1)$ (see equation (3.19)).

3.6 Summary of Solutions

The complete solution of the continuum model equations for the control $f_i=0$ is

$$a = 1 - \{s_1 s_2(\alpha_1\alpha_2-1)(1-Y)/[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]\} \neq 1 \ , \quad (3.22)$$

$$n_1(2j-1) = (A/\alpha_1)a^{j-1} + n_1(\infty) \ , \quad (3.23)$$

$$n_1(2j) = (Ab)a^{j-1} + N_1(\infty) \ , \quad (3.24)$$

$$n_2(2j-1) = Aa^{j-1} + N_2(\infty) \ , \quad (3.25)$$

$$n_2(2j) = (Ab/\alpha_2)a^{j-1} + n_2(\infty) \ , \quad (3.26)$$

$$g_1 = \{\alpha_1[n_1(0)-N_1(\infty)]-[n_2(0)-n_2(\infty)]\}/[\alpha_1(s_1-q_1)+q_2] + g_{1\infty} \ , \quad (3.27)$$

$$g_{2j} = A(\alpha_1\alpha_2-1)a^{j-1}/\alpha_1[\alpha_2(s_2-q_2)+q_1] + g_{2\infty} \ , \quad (3.28)$$

$$g_{2j+1} = A(\alpha_1\alpha_2-1)a^{j}/\alpha_1[\alpha_2 q_2+s_1-q_1] + g_{1\infty} \ , \quad (3.29)$$

where

$$(\alpha_1\alpha_2-1)n_1(\infty) = \alpha_2[\ell q_2(1+y_1-y_2)/(1-Y)-\beta_1]+[\ell q_1(1-y_1+y_2)/(1-Y)-\beta_2] \ , \quad (3.30)$$

$$(\alpha_1\alpha_2-1)n_2(\infty) = [\ell q_2(1+y_1-y_2)/(1-Y)-\beta_1]+\alpha_1[\ell q_1(1-y_1+y_2)/(1-Y)-\beta_2] \ , \quad (3.31)$$

$$N_1(\infty) = \alpha_2 n_2(\infty)+\beta_2 = n_1(\infty)+\ell q_1(1-y_1+y_2)/(1-Y) \ , \quad (3.32)$$

$$N_2(\infty) = \alpha_1 n_1(\infty)+\beta_1 = n_2(\infty)+\ell q_2(1+y_1-y_2)/(1-Y) \ , \tag{3.33}$$

$$g_{1\infty} = 2\ell y_1/(1-Y) \ , \quad g_{2\infty} = 2\ell y_2/(1-Y) \ , \tag{3.34}$$

$$A = \{\alpha_1/[\alpha_1(s_1-q_1)+q_2]\}\{q_2[n_1(0)-N_1(\infty)]+(s_1-q_1)[n_2(0)-n_2(\infty)] \ , \tag{3.35}$$

$$b = (\alpha_2/\alpha_1)(\alpha_1 q_1+s_2-q_2)/[\alpha_2(s_2-q_2)+q_1] \ . \tag{3.36}$$

The complete solution for the control $n_i=0$ can be written down from those above by letting α_i tend to infinity giving

$$a = 1 - [s_1 s_2(1-Y)/(s_1-q_1)(s_2-q_2)] \neq 1 \ , \tag{3.37}$$

$$n_1(2j-1) = 0 \ , \tag{3.38}$$

$$n_1(2j) = (Ab)a^{j-1} + N_1(\infty) \ , \tag{3.39}$$

$$n_2(2j-1) = Aa^{j-1} + N_2(\infty) \ , \tag{3.40}$$

$$n_2(2j) = 0 \ , \tag{3.41}$$

$$g_1 = [n_1(0) - N_1(\infty)]/(s_1-q_1) + g_{1\infty} \ , \tag{3.42}$$

$$g_{2j} = Aa^{j-1}/(s_2-q_2) + g_{2\infty} \ , \tag{3.43}$$

$$g_{2j+1} = Aa^{j}/q_2 + g_{1\infty} \ , \tag{3.44}$$

where

$$N_1(\infty) = \ell q_1(1-y_1+y_2)/(1-Y) \ , \tag{3.45}$$

$$N_2(\infty) = \ell q_2(1+y_1-y_2)/(1-Y) \ , \tag{3.46}$$

$$g_{1\infty} = 2\ell y_1/(1-Y) \ , \quad g_{2\infty} = 2\ell y_2/(1-Y) \ , \tag{3.47}$$

$$A = \{q_2[n_1(0)-N_1(\infty)] + (s_1-q_1)n_2(0)\}/(s_1-q_1) \ , \tag{3.48}$$

$$b = q_1/(s_2-q_2) \ . \tag{3.49}$$

As is predicted in the usual theory of difference equations, the solutions take a special form in the degenerate case $a = 1$.

The solutions (3.37) - (3.49) can also be obtained directly by solving (3.8) - (3.11) subject to $n_1(2j-1) = 0$, $n_2(2j) = 0$; i.e. by solving

$$0 = n_1(2j-2) - g_{2j-1}(s_1-q_1) + \ell q_1 \ , \tag{3.50}$$

$$n_2(2j-1) = g_{2j-1}q_2 + \ell q_2 \ , \tag{3.51}$$

$$n_1(2j) = g_{2j}q_1 + \ell q_1 \ , \tag{3.52}$$

$$0 = n_2(2j-1) - g_{2j}(s_2-q_2) + \ell q_2 \ . \tag{3.53}$$

Eliminating g_{2j-1} from (3.50) and (3.51), and g_{2j} from (3.52) and (3.53) and then $n_2(2j-1)$ from the resulting pair of equations gives

$$n_1(2j) = \frac{q_1q_2}{(s_1-q_1)(s_2-q_2)}\, n_1(2j-2) + \frac{\ell q_1(s_1s_2+q_2s_1-q_1s_2)}{(s_1-q_1)(s_2-q_2)} \ ,$$

which may be written as

$$n_1(2j) = an_1(2j-2) + (1-a)N_1(\infty) \ , \tag{3.54}$$

where a, $N_1(\infty)$ are constants as defined by (3.37) and (3.45). The equation (3.54) corresponds to (3.17) for $f_i=0$ control.

The maximum lengths of the queues are $M_1(2j) = n_1(2j) + \ell q_1$, $M_2(2j-1) = n_2(2j-1) + \ell q_2$ and the minimum lengths $n_1(2j-1)$, $n_2(2j)$.

3.7 Stability

The transient behavior of the queue lengths, phase times, etc. depends on the value of the constant a. This constant is positive unless one of the quantities $s_i - q_i$ is positive and the other negative as may be seen from equation (3.18). If also $a < 1$, the control is stable and the system converges to a steady state. Further, the convergence is uniform in the sense that it is non-oscillatory, i.e. the sequences of points W_{2j} and W_{2j+1} converge monotonically as illustrated in Fig. 4. Since $s_i > q_i$ and $\alpha_i > 1$ for cases of practical interest the stability criterion $a < 1$, i.e.

$$1 - \frac{s_1 s_2(\alpha_1\alpha_2-1)(1-Y)}{[\alpha_1(s_1-q_1)+q_2][\alpha_2(s_2-q_2)+q_1]} < 1$$

reduces to $Y < 1$, which is simply the criterion for the undersaturation of the intersection [6]. Hence the control algorithm gives stable control provided that the intersection is undersaturated.

The rate of convergence to the steady state is also governed by the value of the constant a. Under $f_i = 0$ control, the value of a depends on the control constants α_1, α_2 and has its maximum value of unity when $\alpha_1 = \alpha_2 = 1$ and decreases with increasing α_1, α_2 to the limiting value

$$a = y_1 y_2/(1-y_1)(1-y_2) \qquad (3.55)$$

which is the value of a for $n_i = 0$ control (see equation (3.37)). In other words, the rate of convergence under $f_i = 0$ control increases with increasing α_1, α_2 and the maximum rate of convergence is obtained when $\alpha_1 = \alpha_2 = \infty$ which corresponds to $n_i = 0$ control.

3.8 Limit Cycle Characteristics

In the limit as time t tends to infinity the walks converge to a closed cycle which may have OA_1, OA_2 [Fig. 5(a)] or A_1B_1, A_2B_2 [Fig. 5(b)] as reflecting barriers corresponding to the controls $n_i = 0$ and $f_i = 0$ respectively. These closed cycles will hereafter be referred to as limit cycles. It should be noted that the possibility of a limit cycle under a mixture of $n_i = 0$ and $f_i = 0$ controls is ignored. Which control actually governs the limit cycle depends on the choice of the control constants α_i, β_i .

In the limit cycle, the phase settings of the light are independent of the type of control and are given by

$$g_{1\infty} = 2\ell y_1/(1-Y) \ , \quad g_{2\infty} = 2\ell y_2(1-Y) \ , \tag{3.56}$$

with corresponding cycle time

$$C_\infty = g_{1\infty} + g_{2\infty} + 2\ell = 2\ell/(1-Y) \ . \tag{3.57}$$

This is the familiar result [6] for that cycle time which is compatible with the number of cars arriving at the intersection being equal to the number leaving, the effective greens being split in the ratio $g_{1\infty}/g_{2\infty} = y_1/y_2$. If the limit cycle is under $n_i = 0$ control

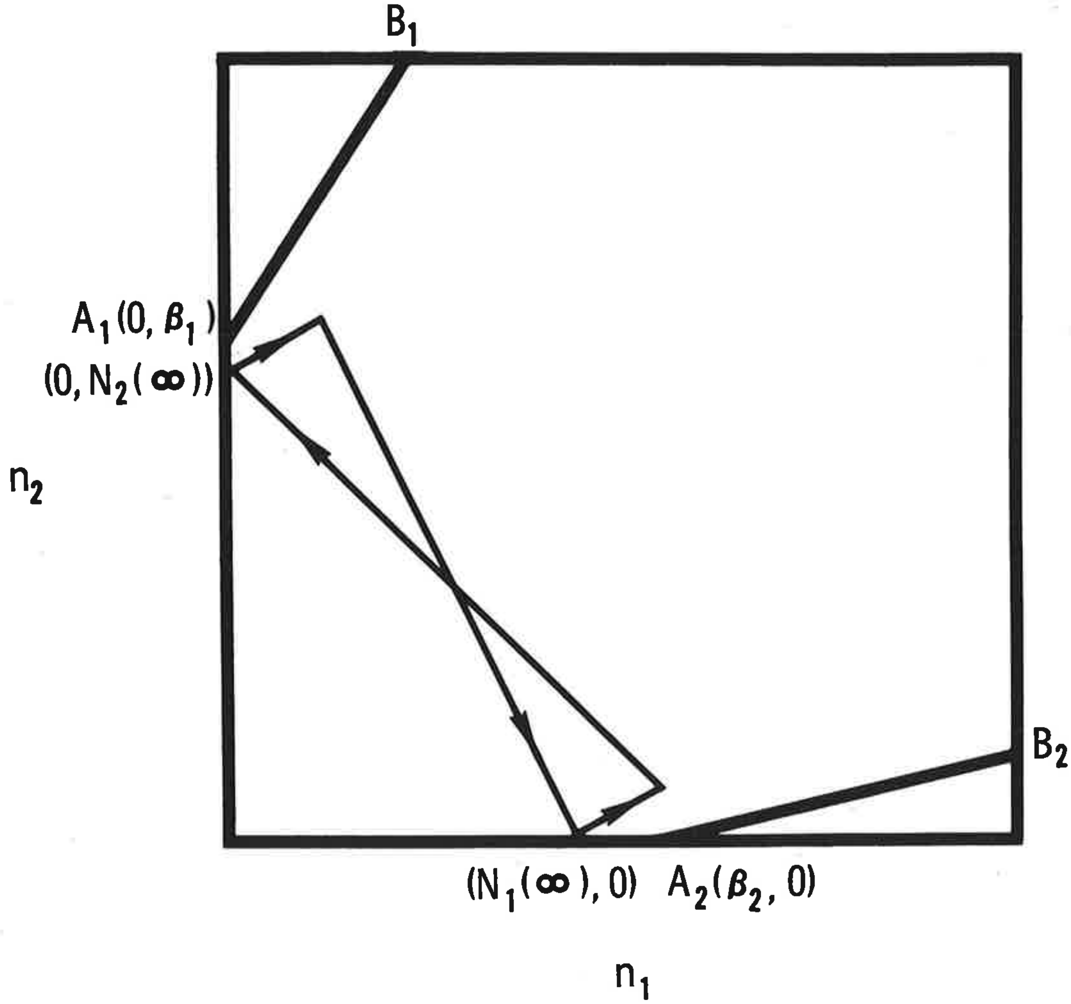


Fig. 5(a). Limit cycle with OA_1, OA_2 as reflecting barriers (n_i=0 control).

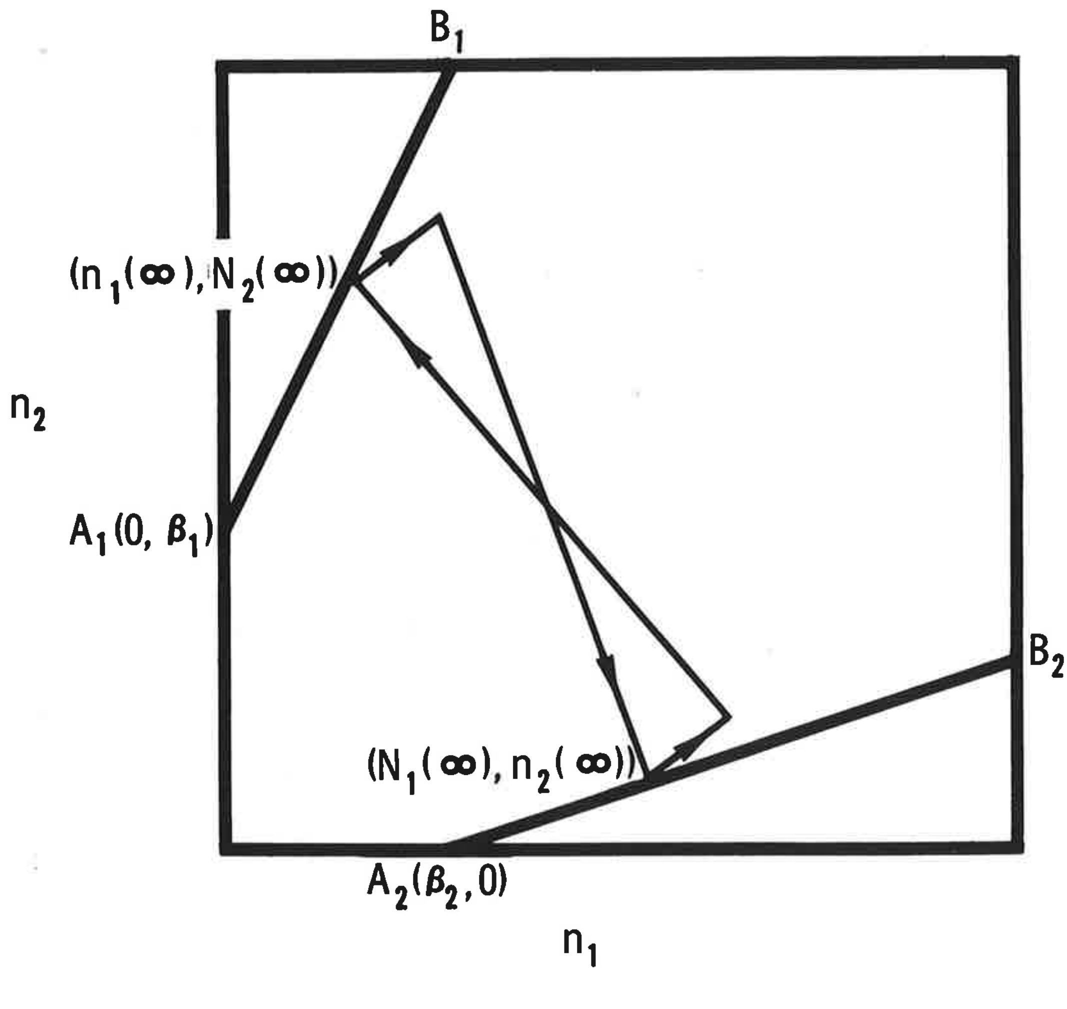


Fig. 5(b). Limit cycle with A_1B_1, A_2B_2 as reflecting barriers (f_i=0 control).

[as in Fig. 5(a)] the lengths of the queues oscillate between $2\ell q_i(1-y_i)/(1-Y)$ and zero. However, if the limit cycle is under $f_i = 0$ control [as in Fig. 5(b)] the queues are never cleared. Although the phase and cycle times, being independent of the control constants, are the same in each case, the uncleared vehicles create an increase in the average delay at the intersection so that a limit cycle under $n_i = 0$ control is preferable.

The limit cycle for $n_i = 0$ control has the property that each car is forced to stop once, and once only; and this is obviously true in the transient state as well. For $f_i = 0$ control, it will be shown that this property is true for the transient and steady states provided that

$$\alpha_1\alpha_2 q_1 > \alpha_2 q_2 + s_1 \ , \quad \alpha_1\alpha_2 q_2 > \alpha_1 q_1 + s_2 \ . \tag{3.58}$$

To prove the first of the above relations consider the $n_1(2j)$ vehicles queued in arm 1 at the end of the $2j^{th}$ phase. The condition that these vehicles will be cleared during the $(2j+1)^{th}$ phase is clearly

$$g_{2j+1}s_1 \geq n_{2j} \ ,$$

i.e.

$$\left[\frac{A(\alpha_1\alpha_2-1)a^j}{\alpha_1(\alpha_2 q_2+s_1-q_1)} + g_{1\infty}\right] s_1 \geq (Ab)a^{j-1} + N_1(\infty) \ .$$

This will certainly be true if

$$\frac{(\alpha_1\alpha_2-1)as_1}{\alpha_1(\alpha_2 q_2+s_1-q_1)} \geq b \tag{3.59}$$

and

$$g_{1\infty}s_1 \geq N_1(\infty) . \tag{3.60}$$

Substituting for a and b the values given in equations (3.18) and (3.36) enables (3.59) to be reduced to

$$\alpha_1\alpha_2q_1 > \alpha_2q_2+s_1 .$$

After some rearrangement (3.60) becomes

$$\frac{2\ell q_1}{1-Y} > \frac{1}{\alpha_1\alpha_2-1}\left[\frac{\alpha_2\ell q_2(1+y_1-y_2)}{1-Y} + \frac{\alpha_1\alpha_2\ell q_1(1-y_1+y_2)}{1-Y} - \alpha_2\beta_1-\beta_2\right] .$$

which will be satisfied if

$$\alpha_1\alpha_2\, q_1(1+y_1-y_2) > \alpha_2\ell q_2(1+y_1-y_2) + 2\ell q_1 ,$$

i.e.

$$\alpha_1\alpha_2q_1 > \alpha_2q_2 + \frac{2q_1}{1+y_1-y_2} .$$

Since $s_1 > \frac{2q_1}{1+y_1-y_2}$ the last inequality will be valid if $\alpha_1\alpha_2q_1 > \alpha_2q_2+s_1$. A similar method may be used to derive the second inequality in (3.58). These inequalities are sufficient, but not necessary, for the control to have the property that any vehicle should be forced to stop no more than once. It is clear that the inequalities will always be satisfied if q_1 and q_2 are not vastly different and if the α_i are "reasonably large".

3.9 Initial Conditions

The sign of the transient terms, e.g., $(Ab)a^{j-1}$ in (3.24) is, as is evident from (3.35), the same as that of

$$q_2[n_1(0)-N_1(\infty)] + (s_1-q_1)[n_2(0)-n_2(\infty)] \ .$$

If the initial values $n_1(0)$, $n_2(0)$ are such that this term is positive, then $n_i(j)$ etc. tend to their limiting values uniformly from above (as illustrated in Fig. 4); if the term is negative, the approach is from below.

3.10 Special Cases

(1) $\alpha_1 = \alpha_2 = 1$

In this case, the phase change under $f_i = 0$ control occurs when the queue lengths differ by some constant number of vehicles. It is evident that a = 1 and it can be shown (by standard methods for the solution of first order difference equations) that the solution is not convergent.

(2) $\beta_1 = \beta_2 = 0$

In this case, the possibility of $n_i = 0$ control is eliminated and the phase is switched, under $f_i = 0$ control, to the next one when the number of vehicles in one arm is a constant multiple of the number of vehicles in the other. As described earlier the corresponding limit cycle will have the undesirable property that the queues never clear.

(3) $\beta_1 \geq \ell q_2(1+y_1-y_2)/(1-Y) \ , \quad \beta_2 \geq \ell q_1(1-y_1+y_2)/(1-Y) \ .$ (3.61)

It is evident from (3.30) and (3.31) that, since n_i is always nonnegative, the corresponding limit cycle must be under $n_i = 0$ control; if the light is ever under $f_i = 0$ control it must therefore converge to $n_i = 0$ control as illustrated in Fig. 4. In this special case, therefore, with the β_i chosen sufficiently large to satisfy (3.61), a limit cycle having the very desirable properties discussed above is obtained.

3.11 Delay

In the steady state the average delay per vehicle for $n_i = 0$ control is less than that for $f_i = 0$ control because of the uncleared vehicles in the latter case. The average delay for $n_i = 0$ control is easily evaluated by estimating the area under the "saw-tooth" graphs of queue length against time (see Fig. 6). If $D^{(1)}$ is the total delay to queue 1 per cycle and $D^{(2)}$ the corresponding quantity for queue 2 then, from the graphs, the total delay to both queues is

$$D = D^{(1)} + D^{(2)} ,$$

$$= \frac{1}{2} C_\infty(N_1(\infty) + \ell q_1) + \frac{1}{2} C_\infty(N_2(\infty) + \ell q_2) .$$

The average delay per vehicle, $\bar{D}$, is found by

$$\bar{D} = D/\text{number of vehicles arriving (or leaving) per cycle} ,$$

$$= D/C_\infty(q_1 + q_2) .$$

or substituting for D this becomes

$$\bar{D} = \frac{\ell}{1-Y} [\frac{q_1(1-y_1)+q_2(1-y_2)}{q_1+q_2}] .$$

It is easily verified (and in fact is a direct consequence of the nature of the $n_i = 0$ control) that the steady state phase times as

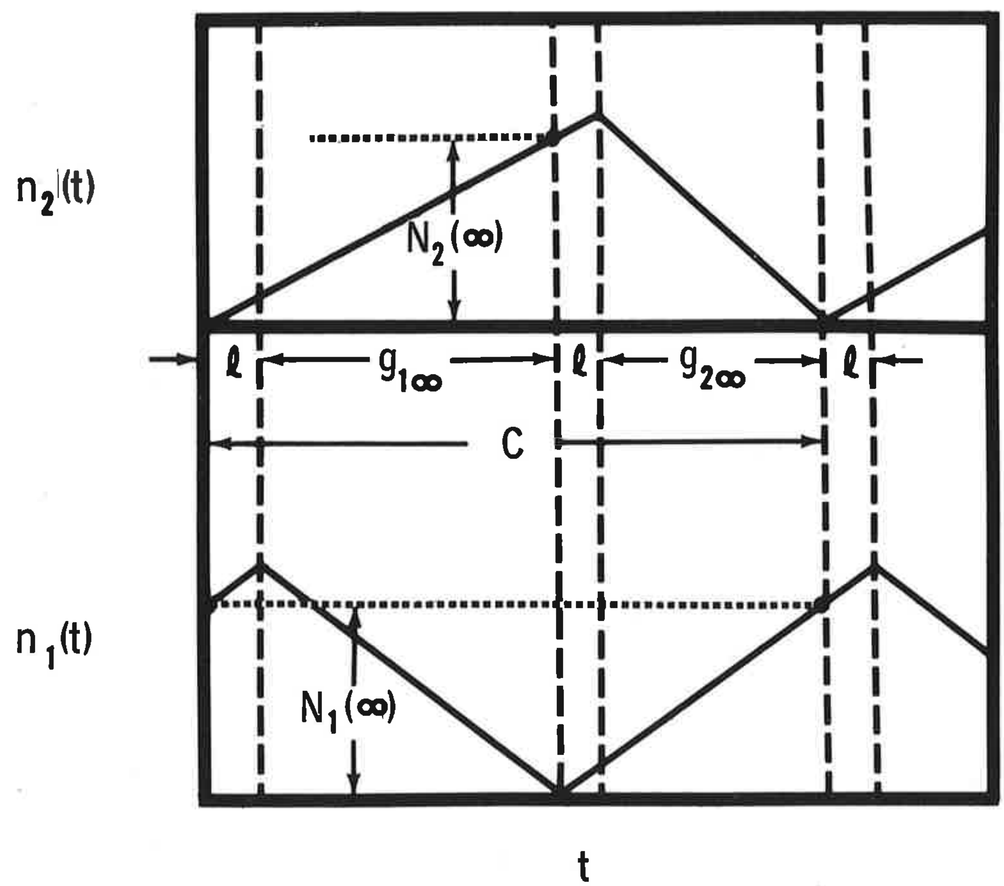


Fig. 6. Graphs of $n_i(t)$ against t for the case of n_i = 0 control in the steady state.

given by (3.56) are the minimum phase times consistent with both queues clearing every cycle. For larger phase times than these the queues will still clear but there will be a change in average delay. It is shown in Appendix I that unless the values of q_1 and q_2 are markedly different the $n_i = 0$ control does give the minimum average delay.

The problem of delay in the transient state for $n_i = 0$ control has been investigated by Grafton and Newell [21] who have shown that if the departure rates on the two arms are equal $n_i = 0$ control gives minimum average delay but if the departure rates are different and in addition the initial queues are very large or very small some modification to the control is necessary to achieve an optimal situation.

3.12 Choice of Control Constants

An optimal control should

(i) ensure rapid convergence to the steady state,

(ii) achieve minimum average delay in both the transient and steady states,

(iii) limit the maximum delay to an individual,

(iv) react to the actual traffic conditions, and

(v) be flexible.

As discussed in section 3.7 the rate of convergence is governed by the value of the constant a and rapid convergence is achieved when α_i are large (certainly somewhat greater than unity).

Minimum average delay in the steady state is achieved whenever the limit cycle is under $n_i = 0$ control. This will be so when the β_i are chosen to satisfy (3.61).

It is clear from the work of Grafton and Newell [21] (see also section 3.11) that $n_i = 0$ control may not be optimal in the transient state and hence that, under certain conditions, it may be necessary to vary the value of the control constants to achieve minimum average delay.

Provided that the α_i satisfy (3.58) each vehicle is forced to stop once only, or equivalently, each vehicle is cleared not later than during the green phase after the (red or green) phase in which it arrives. This ensures that no vehicle is forced to wait for more than one cycle of the lights thus limiting the maximum delay to an individual.

If large queues are present and the intersection is almost saturated the control $T = R_i$ will override the $f_i = 0$ and $n_i = 0$ controls. In these circumstances the phase split may not be optimal; for example, a steady trickle of vehicles on one arm may be sufficient to hold the green phase on that arm for a maximum time although the demand on the other arm may be much greater. For this reason the control $T = R_i$ is to be avoided whenever possible by proper choice of α_i and β_i.

The control is by its nature a flexible one. In practice the computer could use past history and the present traffic state to assist in deciding appropriate values for the constants. For example the computer would know the arrival and departure time headways and could therefore estimate the values of β_i required to satisfy (3.61) and would know the initial queue lengths so that, on the basis of the work of Grafton and Newell [21] could choose the optimum values of α_i for the transient state.

CHAPTER 4

BINOMIAL ARRIVALS

4.1 The Model

Although the assumption of constant arrival rates provides a valuable insight to the nature of the problem it is necessary, particularly with a view to evaluating the operational characteristics of, say, a computer controlled intersection, to use a more realistic type of arrival distribution. In this chapter it will be assumed that the departure rates are constant and are the same for each arm and that the arrivals are generated by a binomial process. In the Toronto pilot study, for example, the arriving traffic is recorded in the central computer as a series of 0's and 1's from a periodical scanning (say every τ seconds) of single lane presence detectors. Each detector registers a pulse if a vehicle crosses it during a τ-second interval and this interval is chosen sufficiently small so that 2 cars, because of their finite length and limited speed, cannot cross a detector during the same interval. The assumption of binomial arrivals, in this case, is equivalent to assuming that the 0's and 1's are uncorrelated which is unrealistic for heavy or pulsed traffic.

Specifically, the vehicle arrivals will be described by denoting by y_i the probability of one arrival in arm i in each of the intervals $(k\tau, k\tau + \tau)$, $k = 0, 1, 2, \ldots$, and by $x_i = 1-y_i$ the probability of no arrival. For convenience, the time interval τ is taken equal to the time interval between vehicle departures, although the analysis could be extended to cover other choices. As a numerical example, suppose that τ = 2 seconds, and y_1 = 0.4, y_2 = 0.3; then the saturation flow (on both arms) is 1,800 v.p.h. and the arriving traffic flows on arms 1 and 2 are 720 v.p.h. and 540 v.p.h. respectively.

As before, each green phase consists of a lost time followed by an effective green time.

The mathematical analysis is confined to the case of $n_i = 0$ control for reasons that will be discussed later.

4.2 Random Walk Equations

During the lost time at the beginning of each green phase, no vehicles cross the intersection. Hence if the number of vehicles queued in arm i at time $t = k\tau$ is denoted by $n_i(k\tau)$, then

(i) $n_i(k\tau + \tau) = n_i(k\tau) + 1$ with probability y_i,

(ii) $n_i(k\tau + \tau) = n_i(k\tau)$ with probability x_i.

If the subsequent effective green favours arm 1 so that vehicles arrive at arms 1 and 2 but discharge only from arm 1, then

(iii) $n_1(k\tau + \tau) = n_1(k\tau)$ with probability y_1,

(iv) $n_1(k\tau + \tau) = n_1(k\tau) - 1$ with probability x_1,

(v) $n_2(k\tau + \tau) = n_2(k\tau) + 1$ with probability y_2,

(vi) $n_2(k\tau + \tau) = n_2(k\tau)$ with probability x_2.

It is clear from equations (iii) - (vi) that, if at time $t = k\tau$ the particle is at (n_1, n_2), then, at time $t = k\tau + \tau$ it will be at one of the points (n_1, n_2), $(n_1 - 1, n_2)$, $(n_1, n_2 + 1)$, $(n_1 - 1, n_2 + 1)$ with associated probabilities y_1x_2, x_1x_2, y_1y_2, x_1y_2 as indicated in Fig. 7.

Equations (i) - (vi), together with four additional ones governing the case when arm 2 is favoured, characterise a two-dimensional binomial random walk [22] which has the coordinate axes $n_1 = 0$ and $n_2 = 0$ as reflecting barriers. The analysis of this walk yields the operational characteristics of the control.

4.3 Notation

The notation used in this chapter closely follows that used in the previous one but the following points should be noted:

(i) for ease of manipulation g_j will be written as $g(j)$,

(ii) all time dependent quantities are measured in units of service time so that the jth phase is of length $g(j)\tau$ seconds and the duration of the lost time is $\ell\tau$ seconds,

(iii) as only $n_i = 0$ control is considered both $n_1(2j + 1)$ and $n_2(2j)$ are zero,

(iv) it will be assumed that the initial queue lengths are $n_1(0)$ and zero respectively.

The most important quantities referred to in this chapter are illustrated in Fig. 8 and are defined as follows:

$M_1(2j)$ = number of vehicles queued in arm 1 at the beginning of the (j + 1) th effective green phase for this arm, (4.1)

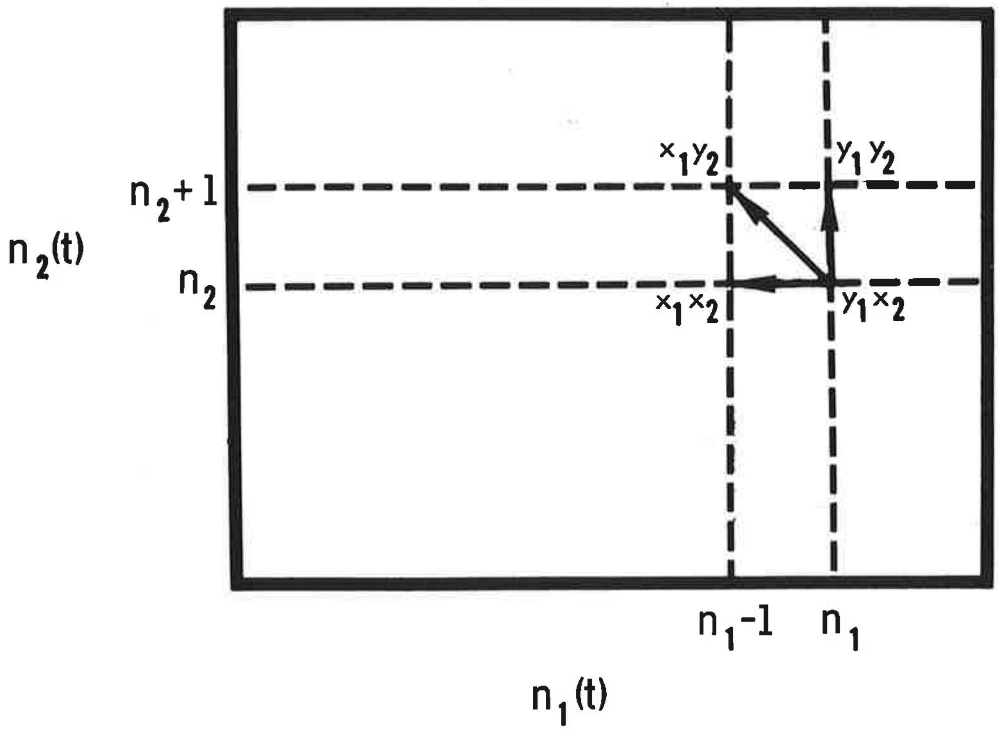


Fig. 7. The steps which the particle may take from the point (n_1, n_2) in the time interval $(k\tau, k\tau+\tau)$ and the probabilities associated with these steps.

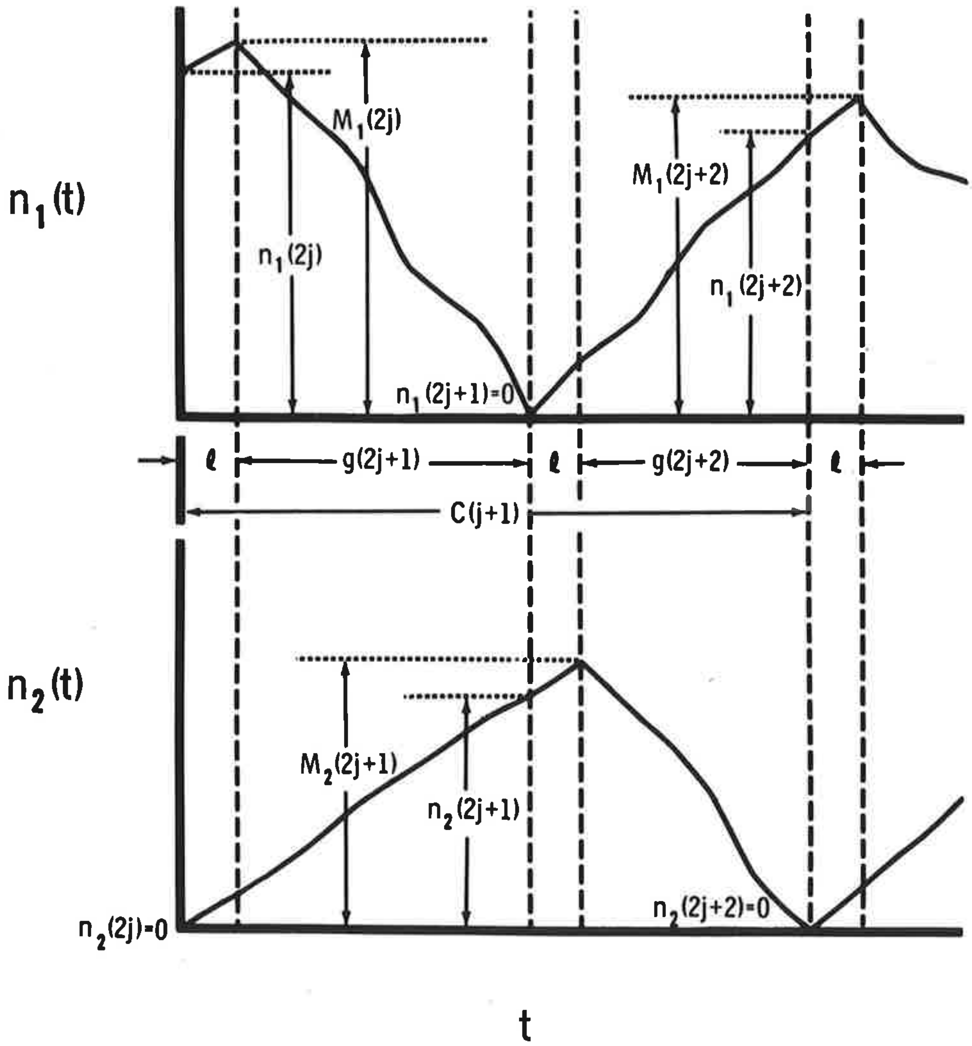


Fig. 8. Graphs of $n_i(t)$ against t illustrating the notation used in Chapter 4.

$M_2(2j + 1)$ = number of vehicles queued in arm 2 at the beginning of the $(j + 1)$ th effective green phase for this arm, (4.2)

$n_1(2j)$ = number of vehicles queued in arm 1 at the beginning of the $(j + 1)$ th green phase for this arm, (4.3)

$n_2(2j + 1)$ = number of vehicles queued in arm 2 at the beginning of the $(j + 1)$ th green phase for this arm, (4.4)

$g(2j + 1)$ = duration of the $(j + 1)$ th effective green phase for arm 1, (4.5)

$g(2j + 2)$ = duration of the $(j + 1)$ th effective green phase for arm 2, (4.6)

$C(j) = g(2j - 1) + g(2j) + 2\ell$ = duration of the jth cycle. (4.7)

4.4 Random Walk Analysis

To derive probability distributions for queue lengths at change of phase, for phase times, and so on, probability generating functions will be found for these random variables and, in the usual way, the distributions will then be found by expressing the generating functions as power series about the origin. In particular:

(i) $\phi_1^{(2j)}(z) = E[z^{M_1(2j)}]$, is the generating function for $M_1(2j)$, (4.8)

(ii) $\phi_2^{(2j+1)}(z) = E[z^{M_2(2j+1)}]$, is the generating function for $M_2(2j + 1)$, (4.9)

(iii) $\theta_1^{(2j)}(z) = E[z^{n_1(2j)}]$, is the generating function for $n_1(2j)$, (4.10)

(iv) $\theta_2^{(2j+1)}(z) = E[z^{n_2(2j+1)}]$, is the generating function for $n_1(2j + 1)$, (4.11)

(v) $\Gamma_1^{(2j+1)}(z) = E[z^{g(2j+1)}]$, is the generating function for $g(2j + 1)$, (4.12)

(vi) $\Gamma_2^{(2j+2)}(z) = E[z^{g(2j+2)}]$, is the generating function for $g(2j + 2)$, (4.13)

(vii) $\Delta^{(j)}(z) = E[z^{C(j)}]$, is the generating function for $C(j)$. (4.14)

The generating function $\phi_1^{(2j)}(z)$ will be evaluated and then relations between this generating function and those defined by (4.9) - (4.14) will be established leading to the evaluation of the latter.

Use will be made of the theory of conditional expectation and the notation used will follow that of Feller [23]. In particular, if X and Y are two random variables then, the conditional distribution of Y for given X will be denoted by $Y|X$, and the conditional expectation of Y for given X will be denoted by $E[Y|X]$.

Evaluation of $\phi_1^{(2j)}(z)$

An expression will be found for the generating function, $E[z^{M_1(2j+2)}|M_1(2j)]$, for $M_1(2j + 2)|M_1(2j)$ and from this a recurrence relation for $\phi_1^{(2j)}(z)$ will be established.

It is clear that $g(2j + 1)|M_1(2j)$ has the negative binomial distribution and therefore has the probability generating function (reference 23 page 252)

$$E[z^{g(2j+1)}|M_1(2j)] = [x_1z/(1-y_1z)]^{M_1(2j)}, \quad (4.15)$$

while $M_2(2j+1)|g(2j+1)$ has the binomial distribution and therefore has the probability generating function (reference 23 page 252)

$$E[z^{M_2(2j+1)}|g(2j+1)] = (x_2+y_2z)^{g(2j+1)+2\ell}. \quad (4.16)$$

Hence

$$E[z^{M_2(2j+1)}|M_1(2j)] = \left[\frac{x_1(x_2+y_2z)}{1-y_1(x_2+y_2z)}\right]^{M_1(2j)}(x_2+y_2z)^{2\ell}$$

$$= [\omega_1(z)]^{M_1(2j)}(x_2+y_2z)^{2\ell}, \quad (4.17)$$

where

$$\omega_1(z) = \frac{x_1(x_2+y_2z)}{1-y_1(x_2+y_2z)}. \quad (4.18)$$

Similarly

$$E[z^{M_1(2j+2)}|M_2(2j+1)] = [\omega_2(z)]^{M_2(2j+1)}(x_1+y_1z)^{2\ell}, \quad (4.19)$$

where

$$\omega_2(z) = \frac{x_2(x_1+y_1z)}{1-y_2(x_1+y_1z)}. \quad (4.20)$$

Hence the probability generating function for $M_1(2j+2)|M_1(2j)$ is

$$E[z^{M_1(2j+2)}|M_1(2j)] = [\omega_1(\omega_2(z))]^{M_1(2j)}[x_2+y_2\omega_2(z)]^{2\ell}(x_1+y_1z)^{2\ell}. \quad (4.21)$$

Taking the expectation of (4.21) with respect to $M_1(2j)$ gives

$$\phi_1^{(2j+2)}(z) = \phi_1^{(2j)}[\omega_1(\omega_2(z))]\,[x_2+y_2\omega_2(z)]^{2\ell}(x_1+y_1z)^{2\ell}. \quad (4.22)$$

On writing

$$\xi(z) = \omega_1(\omega_2(z)),\ \eta(z) = [x_2+y_2\omega_2(z)](x_1+y_1z)\ , \tag{4.23}$$

equation (4.22) becomes

$$\phi_1^{(2j+2)}(z) = [\eta(z)]^{2\ell}\phi_1^{(2j)}[\xi(z)]\ , \tag{4.24}$$

the solution to which may be written down in the form

$$\phi_1^{(2j)}(z) = [(\eta)(\eta\xi)(\eta\xi^{(2)})....(\eta\xi^{(j-1)})]^{2\ell}\phi_1^{(0)}[\xi^{(i)}]\ . \tag{4.25}$$

In this expression,

$$\xi^{(j)} = \xi(\xi(....\ \xi(z)....)) \tag{4.26}$$

is the jth iterate of $\xi(z)$ and $\eta(z)$, $\eta[\xi^{(j)}]$ have been abbreviated to η, $\eta\xi^{(j)}$ respectively.

From the relation $M_1(0) = n_1(0) + \varepsilon_\ell$, where ε_ℓ is the number of arrivals in the lost time (and thus has generating function $(x_1+y_1z)^\ell$), it follows that

$$\phi_1^{(0)}(z) = z^{n_1(0)}(x_1+y_1z)^\ell \tag{4.27}$$

since $n_1(0)$ and ε_ℓ are independent. Thus (4.18) becomes

$$\phi_1^{(2j)}(z) = [(\eta)(\eta\xi)(\eta\xi^{(2)})....(\eta\xi^{(j-1)})]^{2\ell}[\xi^{(j)}]^{n_1(0)}[x_1+y_1\xi^{(j)}]^\ell. \tag{4.28}$$

The form of (4.24) is familiar in the theory of branching processes [24] and for this binomial model its explicit evaluation is possible. From (4.23)

$$\begin{aligned}\xi(z) &= \omega_1[\omega_2(z)]\ ,\\ &= x/(1-yz)\ ,\end{aligned} \tag{4.29}$$

where

$$x = x_1x_2/(x_1x_2 + y_1y_2) \ , \tag{4.30}$$

and

$$y = 1 - x = y_1y_2/(x_1x_2 + y_1y_2) \ . \tag{4.31}$$

Now, as $\xi(z)$ is a bilinear function of z, the iterates $\xi^{(j)}(z)$ will also be bilinear functions of z and as $\xi^{(j)}(1) = 1$ (since $\xi(1) = 1$) it can be assumed that

$$\xi^{(j)} = \frac{x-yz-ya_j(1-z)}{x-yz-yb_j(1-z)} \ , \tag{4.32}$$

where a_j and b_j are constants. The functional equation $\xi^{(j)} = \xi^{(j-1)}(\xi)$ implies that

$$\frac{x-yz-ya_j(1-z)}{x-yz-yb_j(1-z)} = \frac{x-yx(1-yz)^{-1}-y^2a_{j-1}(1-yz)^{-1}(1-z)}{x-yz(1-yz)^{-1}-y^2b_{j-1}(1-yz)^{-1}(1-z)} \ ,$$

$$= \frac{x-yz-y(y/x)a_{j-1}(1-z)}{x-yz-y(y/x)b_{j-1}(1-z)}$$

so that a_j and b_j satisfy the same recurrence relation

$$a_j = (y/x)a_{j-1}, \quad b_j = (y/x)b_{j-1} \ . \tag{4.33}$$

Since $a_1 = 1$ and $b_1 = y/x$,

$$a_j = b_{j-1} = (y/x)^{j-1} \ , \tag{4.34}$$

giving

$$\xi^{(j)} = \frac{x-yz-y(y/x)^{j-1}(1-z)}{x-yz-y(y/x)^{j}(1-z)} . \tag{4.35}$$

Also

$$\eta(z) = [x_2+y_2\omega_2(z)](x_1+y_1z)$$

$$= \frac{x_2(x_1+y_1z)}{1-y_2(x_1+y_1z)} , \tag{4.36}$$

so that

$$\eta\xi^{(j)} = \frac{x_2(x_1+y_1\xi^{(j)})}{1-y_2(x_1+y_1\xi^{(j)})} . \tag{4.37}$$

After some algebraic manipulation this reduces to

$$\eta\xi^{(j)} = \frac{x-yz-(y_1x/x_2)(y/x)^{j}(1-z)}{x-yz-(y_1x/x_2)(y/x)^{j+1}(1-z)} . \tag{4.38}$$

It should be noted that (4.36) may be written in the form

$$\eta(z) = \frac{(x-y)(x_1+y_1z)}{x-yz-(y_1x/x_2)(y/x)(1-z)} . \tag{4.39}$$

Use of (4.38) and (4.39) enables the product $(\eta)(\eta\xi)(\eta\xi^{(2)})\ldots.(\eta\xi^{(j-1)})$ to be evaluated. Since the denominator of $\eta\xi^{(k)}$ cancels with the numerator of $\eta\xi^{(k+1)}$ only the numerator of η and the denominator of $\eta\xi^{(j-1)}$ remain, giving

$$(\eta)(\eta\xi)(\eta\xi^{(2)})\ldots.(\eta\xi^{(j-1)}) = [\frac{(x-y)(x_1+y_1z)}{x-yz-(y_1x/x_2)(y/x)^{j}(1-z)}] .$$

In addition

$$x_1 + y_1\xi^{(j)} = \frac{x-yz-y(y_1/x_2)(y/x)^{j-1}(1-z)}{x-yz-y(y/x)^j(1-z)} ,$$

and these last two results together with (4.35) lead to

$$\phi_1^{(2j)}(z) = [\frac{(x-y)(x_1+y_1z)}{x-yz-x(y_1/x_2)(y/x)^j(1-z)}]^{2\ell}[\frac{x-yz-y(y/x)^{j-1}(1-z)}{x-yz-y(y/x)^j(1-z)}]^{n_1(0)}$$

$$\times [\frac{x-yz-y(y_1/x_2)(y/x)^{j-1}(1-z)}{x-yz-y(y/x)^j(1-z)}]^{\ell} . \tag{4.40}$$

Since $y(y_1/x_2)(y/x)^{j-1} = x(y_1/x_2)(y/x)^j$ this may also be written as

$$\phi_1^{(2j)}(z) = \frac{[(x-y)(x_1+y_1z)]^{2\ell}}{[x-yz-y(y/x)^j(1-z)]^{\ell}} \times \frac{1}{[x-yz-x(y_1/x_2)(y/x)^j(1-z)]^{\ell}}$$

$$\times [\frac{x-yz-y(y/x)^{j-1}(1-z)}{x-yz-y(y/x)^j(1-z)}]^{n_1(0)} . \tag{4.41}$$

This completes the evaluation of the generating function for $M_1(2j)$ and, in the following sections, this result will be used in the determination of further important generating functions.

Evaluation of $\phi_2^{(2j+1)}(z)$

The starting point for the evaluation of $\phi_2^{(2j+1)}(z)$, the generating function for the number of vehicles queued in arm 2 at the onset of the (j + 1) th effective green phase for this arm, is equation (4.17), viz.,

$$E[z^{M_2(2j+1)} | M_1(2j)] = [\omega_1(z)]^{M_1(2j)} (x_2+y_2z)^{2\ell} .$$

Taking the expectation of this equation with respect to $M_1(2j)$ leads to

$$\phi_2^{(2j+1)}(z) = (x_2+y_2z)^{2\ell}\phi_1^{(2j)}[\omega_1(z)] . \tag{4.42}$$

Substituting for $\omega_1(z)$ from (4.18) in the form (4.41) for $\phi_1^{(2j)}(z)$ gives

$$\phi_2^{(2j+1)}(z) = \frac{[(x-y)(x_2+y_2z)]^{2\ell}}{[x-yz-y(y/x)^j(1-z)]^\ell} \times \frac{1}{[x-yz-y(y_2/x_1)(y/x)^j(1-z)]^\ell}$$

$$\times [\frac{x-yz-y(y_2/x_1)(y/x)^{j-1}(1-z)}{x-yz-y(y_2/x_1)(y/x)^j(1-z)}]^{n_1(0)} . \tag{4.43}$$

<u>Evaluation of $\theta_1^{(2j)}(z)$ and $\theta_2^{(2j+1)}(z)$</u>

The relation between $\theta_1^{(2j)}(z)$ and $\phi_1^{(2j)}(z)$ is found by the same method as that used in deriving (4.27). Since $M_1(2j) = n_1(2j) + \varepsilon_\ell$, where, as before, ε_ℓ is the number of arrivals in the lost time, it follows that

$$\phi_1^{(2j)}(z) = (x_1+y_1z)^\ell\theta_1^{(2j)}(z) . \tag{4.44}$$

In a similar manner

$$\phi_2^{(2j+1)}(z) = (x_2+y_2z)^\ell\theta_2^{(2j+1)}(z) . \tag{4.45}$$

Substituting these last two relations in (4.42) leads to a relationship between $\theta_1^{(2j)}(z)$ and $\theta_2^{(2j+1)}(z)$ in the form

$$\theta_2^{(2j+1)}(z) = [\omega_1(z)]^{\ell}\theta_1^{(2j)}[\omega_1(z)] \ . \tag{4.46}$$

Evaluation of $\Gamma_1^{(2j+1)}(z)$ and $\Gamma_2^{(2j+2)}(z)$

The generating function for the effective green times for arm 1 is found by taking the expectation with respect to $M_1(2j)$ of equation (4.1), i.e. of

$$E[z^{g(2j+1)}|M_1(2j)] = [x_1z/(1-y_1z)]^{M_1(2j)} \ ,$$

leading to

$$\Gamma_1^{(2j+1)}(z) = \phi_1^{(2j)}[x_1z/(1-y_1z)] \ . \tag{4.47}$$

Similarly, taking the expectation with respect to $M_2(2j+1)$ of

$$E[z^{g(2j+2)}|M_2(2j+1)] = [x_2z/(1-y_2z)]^{M_2(2j+1)}$$

leads to an expression for the generating function, $\Gamma_2^{(2j+2)}(z)$, for the effective green times for arm 2 viz.,

$$\Gamma_2^{(2j+2)}(z) = \phi_2^{(2j+1)}[x_2z/(1-y_2z)] \ . \tag{4.48}$$

The explicit form for $\Gamma_1^{(2j+1)}(z)$, for example, is found, by use of (4.47) and (4.41), to be

$$\Gamma_1^{(2j+1)}(z) = \frac{(x_2-y_1)^{2\ell}}{[x_2-y_1z-x_2(y/x)^{j+1}(1-z)]^{\ell}} \times \frac{1}{[x_2-y_1z-y_1(y/x)^{j}(1-z)]^{\ell}}$$

$$\times \left[\frac{x_2-y_1z-x_2(y/x)^{j}\ (1-z)}{x_2-y_1z-x_2(y/x)^{j+1}(1-z)}\right]^{n_1(0)} \quad . \qquad (4.49)$$

The relationship between $\Gamma_1^{(2j+1)}(z)$ and $\Gamma_2^{(2j+2)}(z)$ is found as follows:

$$\Gamma_2^{(2j+2)}(z) = \phi_2^{(2j+1)}[x_2z/(1-y_2z)] \ ,$$

$$= \{(x_2+y_2z)^{2\ell}\phi_1^{(2j)}[\omega_1(z)]\}_{z=x_2z/(1-y_2z)} \ ,$$

$$= [x_2/(1-y_2z)]^{2\ell}\{\phi_1^{(2j)}[x_1z/(1-y_1z)]\}_{z=x_2/(1-y_2z)},$$

$$= [x_2/(1-y_2z)]^{2\ell}\Gamma_1^{(2j+1)}[x_2/(1-y_2z)] \ . \qquad (4.50)$$

<u>Evaluation of $\Delta^{(j)}(z)$</u>

To evaluate the generating function, $\Delta^{(j)}(z)$, for the jth cycle time, use is made of the relations

$$E[z^{g(2j+2)}|M_2(2j+1)] = [x_2z/(1-y_2z)]^{M_2(2j+1)} \ ,$$

$$E[z^{M_2(2j+1)}|g(2j+1)] = (x_2+y_2z)^{g(2j+1)+2\ell} \ .$$

These relations imply that

$$E[z^{g(2j+2)}|g(2j+1)] = [x_2+y_2(x_2z/(1-y_2z))]^{g(2j+1)+2\ell} ,$$

$$= [x_2/(1-y_2z)]^{g(2j+1)+2\ell} .$$

Hence,

$$E[z^{g(2j+1)+g(2j+2)+2\ell}|g(2j+1)] = [x_2z/(1-y_2z)]^{g(2j+1)+2\ell} ,$$

i.e.,

$$E[z^{C(j+1)}|g(2j+1)] = [x_2z/(1-y_2z)]^{g(2j+1)+2\ell} . \quad (4.51)$$

Taking the expectation with respect to g(2j+1) of (4.51) gives

$$\Delta^{(j+1)}(z) = [x_2z/(1-y_2z)]^{2\ell}\Gamma_1^{(2j+1)}[x_2z/(1-y_2z)] ,$$

which, after some algebraic manipulation, gives

$$\Delta^{(j)}(z) = \frac{[(1-Y)z]^{2\ell}}{[1-Yz-(y/x)^j(1-z)]^\ell} \times \frac{1}{[1-Yz-(y_1/x_2)(y/x)^{j-1}(1-z)]^\ell}$$

$$\times \left[\frac{1-Yz-(y/x)^{j-1}(1-z)}{1-Yz-(y/x)^j(1-z)}\right]^{n_1(0)} . \quad (4.52)$$

4.5 Statistical Equilibrium

A state of statistical equilibrium is achieved when the intersection is undersaturated. The condition for this is [6]

$$y_1 + y_2 < 1 , \quad (4.53)$$

or, equivalently

$$y < x , \quad (4.54)$$

and if this inequality holds $\xi^{(j)} \to 1$ as $j \to \infty$ and hence $\lim_{j\to\infty} \phi_1^{(2j)}(z)$, the generating function of the so-called steady state probabilities, is independent of $M_1(0)$, the number of vehicles at the onset of the first effective green time. From (4.40) the explicit result for this generating function is

$$\phi_1(z) = \lim_{j\to\infty} \phi_1^{(2j)}(z) = [\frac{(x-y)(x_1+y_1z)}{x-yz}]^{2\ell} , \tag{4.55}$$

It is easy to verify that this expression satisfies the functional equation

$$\phi_1(z) = [\eta(z)]^{2\ell}\phi_1[\xi(z)] , \tag{4.56}$$

which is the limiting form of the recurrence relation (4.~~57~~24). Because there is no dependence on initial conditions, the form for $\phi_2(z)$ may be obtained from that of $\phi_1(z)$ by interchanging the subscripts (and this is true, too, for the phase times and so on), so that

$$\phi_i(z) = [\frac{(x-y)(x_i+y_iz)}{x-yz}]^{2\ell}. \tag{4.58}$$

The limiting forms of the other generating functions defined in the last section are:

$$\theta_1(z) = \lim_{j\to\infty} \theta_1^{(2j)}(z) = [\frac{(x-y)^2(x_1+y_1z)}{(x-yz)^2}]^{\ell} , \tag{4.59}$$

$$\theta_2(z) = \lim_{j\to\infty} \theta_2^{(2j+1)}(z) = [\frac{(x-y)^2(x_2+y_2z)}{(x-yz)^2}]^{\ell}, \tag{4.60}$$

$$\Gamma_1(z) = \lim_{j\to\infty} \Gamma_1^{(2j+1)}(z) = [\frac{x_2-y_1}{x_2-y_1 z}]^{2\ell} , \tag{4.61}$$

$$\Gamma_2(z) = \lim_{j\to\infty} \Gamma_2^{(2j)}(z) = [\frac{x_1-y_2}{x_1-y_2 z}]^{2\ell} , \tag{4.62}$$

$$\Delta(z) = \lim_{j\to\infty} \Delta^{(j)}(z) = [\frac{(1-Y)z}{1-Yz}]^{2\ell} , \tag{4.63}$$

The corresponding steady state probability distributions or stationary distributions are obtained by writing these generating functions as power series about the origin. In the case of $\Gamma_i(z)$ and $\Delta(z)$ this is a very simple matter; for example

$$\Gamma_1(z) = [\frac{x_2-y_1}{x_2-y_1 z}]^{2\ell} = (\frac{x_2-y_1}{x_2})^{2\ell} [1-(y_1/x_2)z]^{-2\ell}$$

$$= (\frac{x_2-y_1}{x_2})^{2\ell} \Sigma_{k=0}^{\infty} \binom{2\ell+k-1}{k}(y_1/x_2)^k z^k , \tag{4.64}$$

(this expansion being valid since $y_1/x_2 = y_1/(1-y_2) < 1$) so that

$$P[g_{1\infty} = k] = \binom{2\ell+k-1}{k}(\frac{x_2-y_1}{x_2})^{2\ell}(\frac{y_1}{x_2})^k , \tag{4.65}$$

where $g_{1\infty} = \lim_{j\to\infty} g(2j+1)$. The probability distributions corresponding to the generating functions $\phi_i(z)$ and $\theta_i(z)$ are obtained as follows:

$$\phi_1(z) = [\frac{(x-y)(x_1+y_1 z)}{x-yz}]^{2\ell} = (x_1x_2-y_1y_2)^{2\ell}[-y_2^{-1}(1+\frac{x_1}{y_1y_2z-x_1x_2})]^{2\ell} ,$$

$$= (\frac{y_1y_2-x_1x_2}{y_2})^{2\ell} \Sigma_{j=0}^{2\ell} \binom{2\ell}{j}(\frac{x_1}{y_1y_2z-x_1x_2})^j ,$$

so that, for $k > 0$,

$$\frac{d^k\phi_1(z)}{dz^k} = \left(\frac{x_1x_2-y_1y_2}{y_2}\right)^{2\ell} \Sigma_{j=1}^{2\ell} \frac{x_1^j(-1)^k(j+k-1)!(y_1y_2)^k}{(j-1)!(y_1y_2z-x_1x_2)^{k+j}} \binom{2\ell}{j} .$$

Hence

$$\frac{1}{k!}\left[\frac{d^k\phi_1(z)}{dz^k}\right]_{z=0} = \left(\frac{x_1x_2-y_1y_2}{y_2}\right)^{2\ell}\left(\frac{y_1y_2}{x_1x_2}\right)^k \Sigma_{j=1}^{2\ell}\binom{2\ell}{j}\binom{j+k-1}{k}\left(\frac{-1}{x_2}\right)^j . \quad (4.66)$$

This form for the coefficient of z^k in the expansion of $\phi_1(z)$ about the origin is particularly suitable for computational purposes as it involves only a summation from $j=1$ to $j=2\ell$ (ℓ being small for cases of practical interest). It is interesting to note that (4.66) may be written in the form

$$\frac{1}{k!}\left[\frac{d^k\phi_1(z)}{dz^k}\right]_{z=0} = \left(\frac{x_1x_2-y_1y_2}{y_2}\right)^{2\ell}\left(\frac{-y_1y_2}{x_1x_2}\right)^k {}_2F_1(-2\ell,k;1;x_2^{-1}) , \quad (4.67)$$

where ${}_2F_1$ is the ordinary hypergeometric function. Polynomials of the form ${}_2F_1(-2\ell,k;1,x_2^{-1})$ have been investigated by Gottlieb [25].

<u>Interpretation as an Eigenvector Problem</u>

If $\underline{t}$ is a stationary distribution then it must satisfy the relation (reference 23 page 356)

$$t_j = \sum_i t_i\lambda_{ij} ,$$

where the λ_{ij} are the corresponding transition probabilities; i.e., $\underline{t}$ is a left eigenvector of the transition matrix $\underline{\Lambda} = (\lambda_{ij})$. As a particular example, suppose that the stationary distribution associated with $\theta_i(z)$ is $\underline{t}_i$. Let $T_i^{(m)}(n,n')$ be the conditional probability that, given there were n vehicles in arm i at the beginning of a particular cycle, there will be n' vehicles in arm i at the end of a further m cycles of the light and let $\underline{T}_i^{(m)}$ be the corresponding transition matrix. Then $\underline{t}_i$ will be a left eigenvector of the matrices $\underline{T}_i^{(m)}$ for $m = 0,1,2,\ldots$. Furthermore, suppose that the random walk performed in the (n_1, n_2) plane is symmetric in the sense that $y_1=y_2$. Let $Q_i(n, n')$ be the conditional probability that, given arm i has the green light and there were n vehicles in arm i at the beginning of this green phase, there will be n' in the other arm at the end of this phase and let $\underline{Q}_i$ be the corresponding transition matrix. Then $\underline{t}_i$ is also a left eigenvector of the matrices $\underline{Q}_i\underline{T}_i^{(m)}(m=0, 1, 2, \ldots)$. In particular, if m=0, $\underline{t}_i$ is a left eigenvector of $\underline{Q}_i$. The matrix $\underline{Q}_1$, for example, can be evaluated from the following relations:

$$E[z^{n_2(2j+1)}|g(2j+1)] = (x_2+y_2z)^{g(2j+1)+\ell} ,$$

$$E[z^{g(2j+1)}|M_1(2j)] = [x_1z/(1-y_1z)]^{M_1(2j)} ,$$

$$E[z^{M_1(2j)}|n_1(2j) = z^{n_1(2j)}(x_1+y_1z)^{\ell} .$$

The first two relations yield

$$E[z^{n_2(2j+1)}|M_1(2j)] = [\omega_1(z)]^{M_1(2j)}(x_2+y_2z)^{\ell} ,$$

which, with the third of the given relations, leads to

$$E[z^{n_2(2j+1)} | n_1(2j)] = [\omega_1(z)]^{n_1(2j)+\ell}$$

where $\omega_1(z) = \dfrac{x_1(x_2+y_2z)}{1-y_1(x_2+y_2z)}$. Thus the generating function for $n_2(2j+1)$ conditional on $n_1(2j)$ is $[\omega_1(z)]^{n_1(2j)+\ell}$ from which the (n, n') element in $\underline{Q}_1$ is found, in the usual manner, by an expansion about z=0.

The matrices $\underline{Q}_i$ $\underline{T}_i^{(m)}$ are, of course, infinite as is the vector $\underline{t}_i$. An approximation to these infinite matrices and corresponding infinite eigenvector is obtained by truncating them at a certain finite dimension. This procedure is such that the ergodic properties of the matrices are preserved [8].

The truncation of $\underline{Q}_1$ and $\underline{t}_1$ and an illustration of the validity of the truncation procedure in a typical case is further discussed in section 4.8.

4.6 Operational Characteristics

From the explicit forms obtained for the various probability generating functions in the last section it is possible to evaluate some of the more important parameters of the corresponding distributions, in particular, the means and variances, and thus to give a complete description of the behavior of the system.

The means

The expected values of $M_1(2j)$, $M_2(2j+1)$ and so on will be denoted

LIBRARY OF THE UNIVERSITY OF ADELAIDE

by $\bar{M}_1(2j)$, $\bar{M}_2(2j+1)$ etc.

A difference equation satisfied by $\bar{M}_1(2j)$ will be found and, by means of the relation connecting $\phi_1^{(2j)}(z)$ and $\theta_1^{(2j)}(z)$, the corresponding difference equation for $\bar{n}_1(2j)$ will be derived. The solution to this equation will be written down by comparing it with that satisfied by the queue lengths in the case of steady arrivals.

The recurrence relation satisfied by $\phi_1^{(2j)}(z)$ is

$$\phi_1^{(2j+2)}(z) = [\eta(z)]^{2\ell}\phi_1^{(2j)}[\xi(z)] ,$$

where $\xi(z) = \omega_1[\omega_2(z)] = x/(1-yz)$ and

$$\eta(z) = [x_2+y_2\omega_2(z)][x_1+y_1z] = \omega_2(z) \text{ with } \omega_2(z) = \frac{x_2(x_1+y_1z)}{1-y_2(x_1+y_1z)} .$$

Logarithmic differentiation of the recurrence relation above gives

$$\frac{\dot{\phi}_1^{(2j+2)}(z)}{\phi_1^{(2j+2)}(z)} = 2\ell\,\frac{\dot{\eta}(z)}{\eta(z)} + \dot{\xi}(z)\,\frac{\dot{\phi}_1^{(2j)}[\xi(z)]}{\phi_1^{(2j)}[\xi(z)]} \tag{4.68}$$

where $\dot{\eta}(z) = \frac{d}{dz}[\eta(z)]$ and so on.

Hence

$$\dot{\phi}_1^{(2j+2)}(1) = 2\ell\dot{\eta}(1) + \dot{\xi}(1)\dot{\phi}_1^{(2j)}(1) . \tag{4.69}$$

Now $\dot{\eta}(1) = (y_1/x_2)$ and $\dot{\xi}(1) = (y/x)$ so that, since $\bar{M}_1(2j) = \dot{\phi}_1^{(2j)}(1)$, equation (4.68) becomes

$$\bar{M}_1(2j+2) = (y/x)\bar{M}_1(2j) + 2\ell y_1/x_2 . \tag{4.70}$$

From (4.44)

$$\phi_1^{(2j)}(z) = (x_1+y_1z)^{\ell}\theta_1^{(2j)}(z)$$

so that, by logarithmic differentiation,

$$\bar{M}_1(2j) = \ell y_1 + \bar{n}_1(2j) \tag{4.71}$$

substitution of which into (4.69) gives

$$\bar{n}_1(2j+2) = (y/x)\bar{n}_1(2j) + \ell y_1(x_1+y_2)/(x_1x_2) \tag{4.72}$$

(use having been made of the identity $2x_1+y_1y_2-x_1x_2 = x_1+y_2$), which may be written as

$$\bar{n}_1(2j+2) = (y/x)\bar{n}_1(2j) + [1-(y/x)]\,\frac{\ell y_1(1-y_1+y_2)}{x_1x_2-y_1y_2} ,$$

or

$$\bar{n}_1(2j+2) = a\bar{n}_1(2j) + (1-a)\bar{N}_1(\infty) . \tag{4.73}$$

This difference equation for $\bar{n}_1(2j)$ is the same as that derived for $n_1(2j)$ in the case of steady arrivals with q_i in the latter being replaced by y_i because of the nature of description of the arrival distribution. It follows that the average phase times etc., satisfy the same equations as the corresponding deterministic quantities in the case of steady arrivals. The solutions may therefore be written down from (3.37) - (3.49):

$$\bar{n}_1(2j) = (Ab)(y/x)^{j-1} + \bar{N}_1(\infty) , \tag{4.74}$$

$$\bar{n}_2(2j-1) = A(y/x)^{j-1} + \bar{N}_2(\infty) \ , \tag{4.75}$$

$$\bar{M}_1(2j) = \bar{n}_1(2j) + \ell y_1 \ , \tag{4.76}$$

$$\bar{M}_2(2j-1) = \bar{n}_2(2j-1) + \ell y_2 \ , \tag{4.77}$$

$$\bar{g}(2j) = A(y/x)^{j-1}/x_2 + \bar{g}_{2\infty} \ , \tag{4.78}$$

$$\bar{g}(2j-1) = A(y/x)^{j-1}/y_2 + \bar{g}_{1\infty} \ , \tag{4.79}$$

$$\bar{C}(j) = A(y/x)^{j-1}/(x_2 y_2) + \bar{C}_\infty \ , \tag{4.80}$$

where

$$\bar{N}_1(\infty) = \ell y_1(1-y_1+y_2)/(1-Y) \ , \tag{4.81}$$

$$\bar{N}_2(\infty) = \ell y_2(1+y_1-y_2)/(1-Y) \ , \tag{4.82}$$

$$\bar{g}_{1\infty} = 2\ell y_1/(1-Y), \quad \bar{g}_{2\infty} = 2\ell y_2/(1-Y) \ , \tag{4.83}$$

$$\bar{C}_\infty = \bar{g}_{1\infty} + \bar{g}_{2\infty} + 2\ell \ , \tag{4.84}$$

$$A = y_2[n_1(0) - \bar{N}_1(\infty)]/x_1 \ , \tag{4.85}$$

$$b = y_1/x_2 \ . \tag{4.86}$$

The variances

In the case of constant arrival and departure rates the moments of order higher than the first are all zero. To make a further comparison of this model with the model of binomial arrivals the second moments of the variables relevant to this latter model will be calculated to give some measure of the statistical fluctuations present. The variances will be evaluated by differentiating the recurrence relation (4.68) which was derived in the last section. Performing the differentiation leads to

$$\frac{\ddot{\phi}_1^{(2j+2)}(z)\phi_1^{(2j+2)}(z)-[\dot{\phi}_1^{(2j+2)}(z)]^2}{[\phi_1^{2j+2}(z)]^2} = 2\ell\,\frac{\ddot{\eta}(z)\eta(z)-[\dot{\eta}(z)]^2}{[\eta(z)]^2}$$

$$+\ \ddot{\xi}(z)\,\frac{\dot{\phi}_1^{2j}[\xi(z)]}{\phi_1^{2j}[\xi(z)]} + [\dot{\xi}(z)]^2\,\frac{\ddot{\phi}_1^{(2j)}[\xi(z)]\phi_1^{(2j)}[\xi(z)]-[\dot{\phi}_1^{(2j)}[\xi(z)]]^2}{[\phi_1^{(2j)}[\xi(z)]]^2},$$

so that

$$\ddot{\phi}_1^{(2j+2)}(1) - [\dot{\phi}_1^{(2j+2)}(1)]^2 = 2\ell[\ddot{\eta}(1)-[\dot{\eta}(1)]^2] + \ddot{\xi}(1)\ \dot{\phi}_1^{(2j)}(1)$$

$$+\ [\dot{\xi}(1)]^2[\ddot{\phi}_1^{(2j)}(1) - [\dot{\phi}_1^{(2j)}(1)]^2]\ .$$

Using the relationship between $\dot{\phi}_1^{(2j+2)}(1)$ and $\dot{\phi}^{(2j)}(1)$ as given in (4.69) leads to

$$\ddot{\phi}_1^{(2j+2)}(1) - [\dot{\phi}_1^{(2j+2)}(1)]^2 + \dot{\phi}_1^{(2j+2)}(1) = 2\ell[\ddot{\eta}(1)-[\dot{\eta}(1)]^2+\dot{\eta}(1)]$$

$$+\ [\ddot{\xi}(1)-[\dot{\xi}(1)]^2+\dot{\xi}(1)]\ \dot{\phi}_1^{(2j)}(1) + [\dot{\xi}(1)]^2[\ddot{\phi}_1^{(2j)}(1)-[\dot{\phi}_1^{(2j)}(1)]^2+\dot{\phi}_1^{(2j)}(1)] \qquad (4.87)$$

Now $\dot{\xi}(1)=y/x$, $\ddot{\xi}(1)=2y^2/x^2$ and $\dot{\eta}(1)=y_1/x_2$, $\ddot{\eta}(1)=2y_1^2y_2/x_2^2$ and as

$\text{var}\,[M_1(2j)] = \ddot{\phi}_1^{(2j)}(1)-[\dot{\phi}_1^{(2j)}(1)]^2 + \dot{\phi}_1^{(2j)}(1)$ equation (4.87) becomes

$$\text{var}[M_1(2j+2)]=(y/x)^2\ \text{var}\,[M_1(2j)] + \frac{2\ell y_1(2y_1y_2-y_1+x_2)}{x_2^2} + \frac{y\bar{M}_1(2j)}{x}\ .$$

Since the initial queue lengths are $n_1(0)$ and zero, respectively, it is convenient to write the last equation in terms of $n_1(2j)$ rather than $M_1(2j)$. Differentiating the relation (4.44), viz

$$\phi_1^{(2j)}(z) = (x_1+y_1z)^{\ell}\theta_1^{(2j)}(z) ,$$

yields successively

$$\bar{M}_1(2j) = \ell y_1 + \bar{n}_1(2j) ,$$

$$\text{var}\ [M_1(2j)] = \ell x_1 y_1 + \text{var}\ [n_1(2j)] .$$

Hence

$$\text{var}\ [n_1(2j+2)] = (y/x)^2\ \text{var}\ [n_1(2j)] + (y/x)\bar{n}_1(2j)$$

$$+ [(y/x)^2-1]\ \ell x_1 y_1 + (y/x)\ \ell y_1 + \frac{2\ell y_1(2y_1y_2-y_1+x_2)}{x_2^2} .$$

$$\text{Now } \bar{n}_1(2j) = Ab(y/x)^{j-1} + \bar{N}_1(\infty)$$

$$= [n_1(0)-\bar{N}_1(\infty)](y/x)^j + \bar{N}_1(\infty)$$

so that the difference equation for $v_{2j+2} = \text{var}\ [n_1(2j+2)]$ is of the form

$$v_{2j+2} = (y/x)^2 v_{2j} + a_1(y/x)^j + a_0, \quad (a_1,\ a_0 \text{ constants})$$

the solution to which may be written down in the form

$$v_{2j} = b_2(y/x)^{2j} + b_1(y/x)^j + b_0 .$$

By substituting this solution into the difference equation and using the initial condition $v_o = 0$, the constants b_o, b_1, b_2 may be evaluated to give

$$v_{2j} = -\left\{ \frac{y}{x-y} [n_1(o)-\bar{N}_1(\infty)] + v_\infty \right\} (y/x)^{2j}$$

$$+ \frac{y}{x-y} [n_1(o)-\bar{N}_1(\infty)] (y/x)^j + v_\infty , \tag{4.88}$$

where

$$v_\infty = \lim_{j\to\infty} v_{2j} = \ell x_1 y_1 \left[1 + \frac{2x_2 y_2}{(x_1-y_2)^2}\right] , \tag{4.89}$$

and

$$\bar{N}_1(\infty) = \frac{\ell y_1 (x_1+y_2)}{(x_1-y_2)} .$$

The variances of the phase times and cycle times may be calculated from the variances of the queue lengths by means of the relationships that exist between the corresponding generating functions: For example, two differentiations of (4.47), i.e. of

$$\Gamma_1^{(2j+1)}(z) = \phi_1^{(2j)}[x_1 z/(1-y_1 z)]$$

will yield a relationship between var $[g_1(2j+1)]$ and var $[M_1(2j)]$. In the transient state these relationships are rather cumbersome despite the fact that their derivations are straightforward. In the steady state, however, the variances of all quantities assume a simple form and are easily derived directly from the relevant generating functions as will be illustrated later in this section.

To further investigate the behavior of v_{2j} it will be shown that $v_\infty > y\bar{N}_1(\infty)/(x-y)$. On substituting for v_∞ and $\bar{N}_1(\infty)$ this becomes

$$\ell x_1 y_1 [1 + \frac{2x_2 y_2}{(x_1-y_2)^2}] > \frac{y_1 y_2}{x_1 x_2 - y_1 y_2} \frac{\ell y_1 (x_1+y_2)}{x_1 - y_2} . \qquad (4.90)$$

Since $x_1 x_2 - y_1 y_2 = x_1 - y_2$ this may be written as

$$x_1 (x_1-y_2)^2 > y_2 [y_1 (x_1+y_2) - 2x_1 x_2] . \qquad (4.91)$$

Now $y_1(x_1+y_2) < 2y_1 x_1$ since $x_1 - y_2 = 1-Y > 0$,

and $2y_1 x_1 < 2x_1 x_2$ since $x_2 - y_1 = 1-Y > 0$.

Hence the left hand side of (4.91) is negative and therefore (4.90) is always satisfied as is the inequality $v_\infty > y\bar{N}_1(\infty)/(x-y)$. It follows that the coefficient of $(y/x)^{2j}$ in (4.88) is always negative whereas the coefficient of $(y/x)^j$ depends on the sign of the quantity $n_1(0) - \bar{N}_1(\infty)$. If $n_1(0) < \bar{N}_1(\infty)$ then v_{2j} is bounded above, for all j, by v_∞; i.e. since $v_o = 0$, v_{2j} increases monotonically to its steady state value. If however $n_1(0) > \bar{N}_1(\infty)$ there will be a maximum value of v_{2j} for some j. It is desirable that (i) the transient variance should not be too large, and (ii) if a maximum occurs the system should near the steady state value of v_∞ in only a few cycles. Because of the form of the expression for v_{2j} it is not possible to derive useful expressions for the value of the maximum, the corresponding value of j, and the time taken for the transient variance to be within some predetermined amount of v_∞. In section 4.8 a numerical example is given indicating the behavior of v_{2j} in a typical case.

The Steady State

The steady state means have already been derived as a result of the theory in the earlier part of this section and are given in equations (4.81) - (4.86).

The steady state variances are most easily calculated from the corresponding steady state generating functions which were derived in the last section (see equations (4.58) - (4.63)). For example, from (4.63),

$$\Delta(z) = [\frac{(1-Y)z}{1-Yz}]^{2\ell} ,$$

$\therefore \log \Delta(z) = 2\ell \log (1-Y) + 2\ell \log z - 2\ell \log (1-Yz)$, so that

$$\dot{\Delta}(z)/\Delta(z) = 2\ell/z + 2\ell Y/(1-Yz) ,$$

and thus

$$\bar{C}_\infty = \dot{\Delta}(1) = 2\ell/(1-Y) .$$

A further differentiation yields

$$[\ddot{\Delta}(z)\Delta(z) - [\dot{\Delta}(z)]^2]/[\Delta(z)]^2 = -2\ell/z^2 + 2\ell Y^2/(1-Yz)^2 ,$$

$$\ddot{\Delta}(1) - [\dot{\Delta}(1)]^2 = -2\ell + 2\ell Y^2/(1-Y)^2 ,$$

so that

$$\text{var } [C_\infty] = \ddot{\Delta}(1) = [\dot{\Delta}(1)]^2 + \dot{\Delta}(1) ,$$

$$= 2\ell Y/(1-Y)^2 .$$

A complete list of the steady state variances is,

$$\mathrm{var}\ [N_1(\infty)] = \ell x_1 y_1\ [1 + \frac{2x_2y_2}{(x_1-y_2)^2}]\ , \tag{4.92}$$

$$\mathrm{var}\ [N_2(\infty)] = \ell x_2 y_2\ [1 + \frac{2x_1y_1}{(x_2-y_1)^2}]\ , \tag{4.93}$$

$$\mathrm{var}\ [M_1(\infty)] = 2\ell x_1 y_1\ [1 + \frac{x_2y_2}{(x_1-y_2)^2}]\ , \tag{4.94}$$

$$\mathrm{var}\ [M_2(\infty)] = 2\ell x_2 y_2\ [1 + \frac{x_1y_1}{(x_2-y_1)^2}]\ , \tag{4.95}$$

$$\mathrm{var}\ [g_{1\infty}] = 2\ell x_2 y_1/(x_2-y_1)^2\ , \tag{4.96}$$

$$\mathrm{var}\ [g_{2\infty}] = 2\ell x_1 y_2/(x_1-y_2)^2\ , \tag{4.97}$$

$$\mathrm{var}\ [C_\infty] = 2\ell Y/(1-Y)^2\ . \tag{4.98}$$

These explicit expressions for the variances, together with those for the mean values, give an accurate description of the behavior of the system. An illustration is given in section 4.8.

4.7 Delay

In this section, using a method very similar to that of Darroch, Morris and Newell [16], a formula will be derived for the expected total wait per cycle for a given arm and the expected delay per vehicle in a given arm.

The "effective" red time for arm 2, i.e. the time during a cycle of length c for which the queue is increasing, will be denoted by R. This time, by its definition, will include two lost times. The total delay for a cycle having an effective red time of length R will be denoted by D_R. The effect on the total delay of beginning this effective red time one time unit earlier will now be considered.

If, in this extra time unit at the beginning of the phase, no vehicle arrives (Fig. 9(a))

$$D_{R+1} = D_R \ , \tag{4.99}$$

the probability of this event being x_2. If, however, there is one arrival, with probability y_2, then

$$D_{R+1} = D_R + \delta_1 + C + \delta_2 \ , \tag{4.100}$$

where

δ_1 = increase in delay at the beginning of the cycle, and

δ_2 = increase in delay at the end of the cycle, as shown in Fig. 9(b).

Equations (4.99) and (4.100) may be combined to give

$$D_{R+1} = D_R + \mu[\delta_1 + C + \delta_2] \ , \tag{4.101}$$

where

$$\mu = \begin{cases} 0 \text{ with probability } x_2 \\ 1 \text{ with probability } y_2 \end{cases} .$$

Taking the expectation of (4.101) gives

$$E[D_{R+1}] = E[D_R] + y_2 \{ E[\delta_1] + E[C|R] + E[\delta_2] \} . \tag{4.102}$$

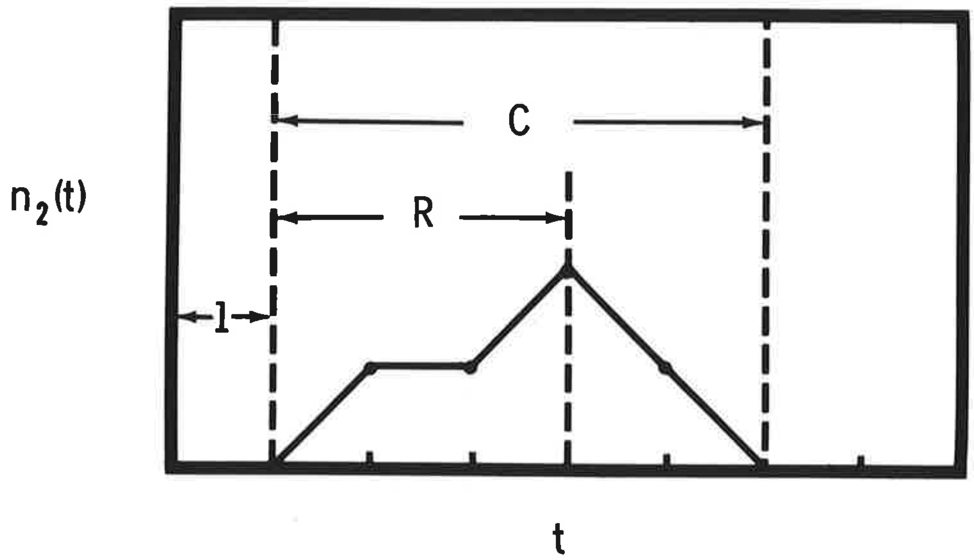


Fig. 9(a). Graph of $n_2(t)$ against t in the case where there is no arrival in the extra time unit preceeding R. $D_{R+1} = D_R$.

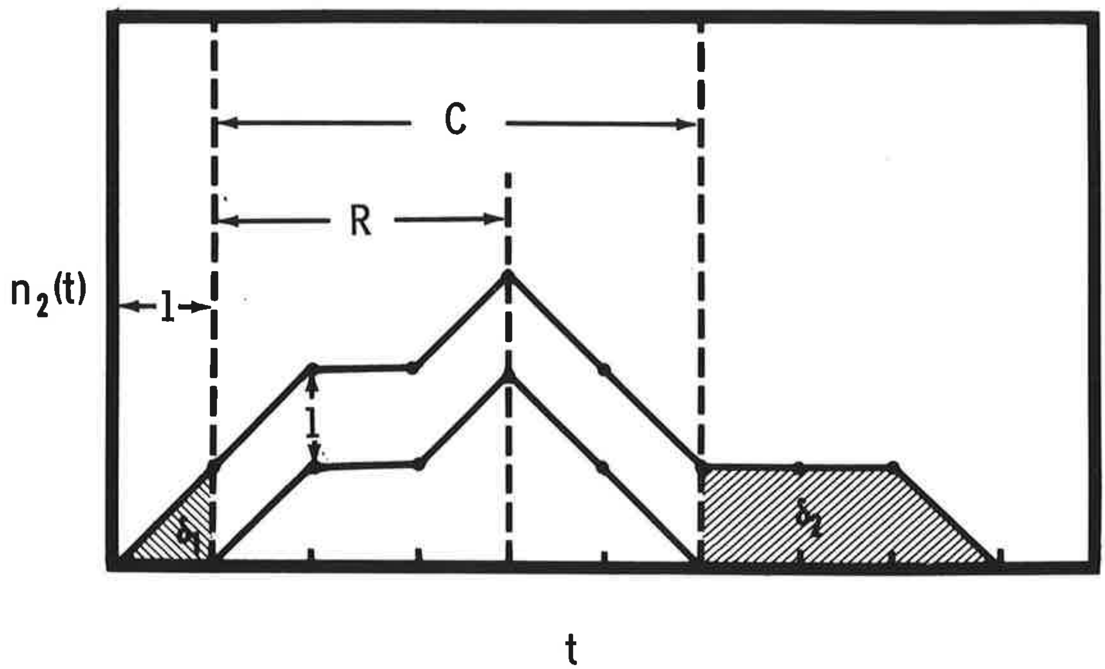


Fig. 9(b). The effect of an arriving vehicle in the extra time unit preceeding R. The quantities δ_1, δ_2 are represented by the shaded areas. In this case $D_{R+1} = D_R + \delta_1 + C + \delta_2$.

Assuming that an arrival may occur at any instant of a service time interval with equal probability, the expected value, $E[\delta_1]$, of δ_1 is 1/2. Now

$$E[C|R] = R + E[\text{time for queue to clear}] ,$$
$$= R + Ry_2/x_2 ,$$
$$= Rx_2^{-1} .$$

If there is no arrival in the first service time interval after the end of the "old" cycle the queue will clear in this service time and in this case

$$\delta_2 = \frac{1}{2} \text{ with probability } x_2 .$$

Similarly if there are arrivals in the k-1 consecutive time intervals after the end of the "old" cycle and no arrival in the kth interval then

$$\delta_2 = (k-1) + \frac{1}{2} \text{ with probability } y_2^{k-1}x_2 .$$

Hence

$$E[\delta_2] = \Sigma_{k=1}^{\infty} (k - \frac{1}{2})y_2^{k-1}x_2 ,$$
$$= x_2^{-1} - \frac{1}{2} .$$

Using these values of $E[\delta_1]$ etc., in (4.102) gives

$$E[D_{R+1}] = E[D_R] + y_2(R+1)/x_2 .$$

This difference equation for $E[D_R]$, with the initial condition that

$D_o=0$, is solved in the standard manner to give

$$E[D_R] = y_2(R^2+R)/(2x_2) \quad . \tag{4.103}$$

Taking the expectation of (4.103) with respect to R gives

$$E[D] = y_2[E(R^2) + E(R)]/(2x_2)$$

$$= y_2[\text{var }(R) + E^2(R) + E(R)]/(2x_2) \quad . \tag{4.104}$$

If R has the probability generating function $\Pi(z)$ then (4.104) may be written as

$$E[D] = y_2[\ddot{\Pi}(1) + 2\dot{\Pi}(1)]/(2x_2) \quad . \tag{4.105}$$

In the case of statistical equilibrium $R = g_{1\infty} + 2\ell$ has generating function $\Pi(z)$ given by

$$\Pi(z) = z^{2\ell}\Gamma_1(z) \quad ,$$

$$= [\frac{(x_2-y_1)z}{x_2-y_1z}]^{2\ell} \quad ,$$

and some simple algebra yields

$$D = D^{(2)} = \frac{\ell(2\ell+1)x_2y_2}{(1-Y)^2} \tag{4.106}$$

as the expected value of the total delay per cycle to vehicles in arm 2. The corresponding quantity, $D^{(1)}$, say, for arm 1 is obtained by interchanging the subscripts. The average delay per vehicle in arm i is obtained by dividing the total delay for that arm by $y_i\bar{C}_\infty$, the expected number of vehicles which arrive in the cycle (or equivalently by dividing by $\bar{g}_{1\infty}$, the expected number of departures) and has the value

$$\bar{D}^{(i)} = \frac{(2\ell+1)x_i}{2(1-Y)} .$$

The average delay, d, per vehicle averaged over both lanes with equal weight to each vehicle is given by

$$\bar{D} = (D^{(1)}+D^{(2)})/[(y_1+y_2)\bar{C}_\infty] ,$$

$$= \frac{(2\ell+1)(x_1y_1+x_2y_2)}{2Y(1-Y)} . \qquad (4.107)$$

As the method used in obtaining these delay formulae is in fact the discrete analogue of that used by Darroch, et al. it is interesting to compare the results although they pertain to different models. By writing $R = R' + \ell$, so that R' corresponds to the X used by Darroch, et al., (4.104) gives

$$E[D] = y_2[\text{var}\ (R'+\ell) + E^2(R'+\ell) + E(R'+\ell)]/(2x_2)$$

and, since ℓ is a constant so that var $(\ell) = 0$,

$$\text{var}\ (R'+\ell) = \text{var}\ (R^+) = \text{var}\ (R') + \text{var}\ (\ell) .$$

Hence

$$E[D] = [y_2/2(1-y_2)][E^2(R'+\ell) + \text{var}\ (R') + \text{var}\ (\ell)]$$

$$+ E[R'+\ell)[y_2/2(1-y_2)] . \qquad (4.108)$$

It is clear that, since y_2 corresponds to α_2 in the notation of the above authors, the first term of (4.108) agrees with equation (28) of reference 16 but that the second terms differ in the coefficients of $E(R'+\ell)$. Further the average delay per vehicle in lane i has a non zero limit as $\ell \to 0$, namely $x_i/[2(1-Y)]$, this feature also having been noticed by Darroch, et al.

4.8 Numerical Example

To illustrate the foregoing theory a numerical example will be given including the evaluation of typical probability generating functions together with the corresponding probability distributions and the first and second moments of these distributions, an example of the truncation of an infinite transition matrix and the corresponding infinite eigenvector, the calculation of the total delay per cycle and the average delay per vehicle and finally some practical considerations.

Except where otherwise stated it will be assumed that the arrival rates on each arm are 720 v.p.h. (so that the random walk in the (n_1, n_2) plane is symmetric), the saturation flows are 1,800 v.p.h. and that the lost time is 6 seconds per phase. Then $\tau=2$ seconds, $y_1=y_2=0.4$, $x_1=x_2=0.6$ and $\ell=3$.

Some generating functions

Substituting the above figures into (4.41) leads to

$$\phi_1^{(2j)}(z) = \frac{(3+2z)^6}{[9-4z-4(4/9)^j(1-z)]^3} \times \frac{1}{[9-4z-6(4/9)^j(1-z)]^3}$$

$$\times \left[\frac{9-4z-4(4/9)^{j-1}(1-z)}{9-4z-4(4/9)^j(1-z)}\right]^{n_1(0)} .$$

Similarly $\phi_2^{(2j+1)}(z)$ is evaluated by substitution into (4.43) and so on.

The limiting form of $\theta_1^{(2j)}(z)$ as $j \to \infty$ can be written down from (4.59) in the form

$$\theta_1(z) = \left[\frac{5(3+2z)}{(9-4z)^2}\right]^3 .$$

The corresponding stationary vector $\underline{t}_1 = \{t_1(m)\}$ is shown in Table 1. In Fig. 10 graphs have been drawn of the steady state probabilities, $t_1(m)$, against m for the cases $\ell = 2$ and $\ell = 3$.

Truncation of $\underline{Q}_1$ and $\underline{t}_1$

The (n, n') element, $Q_1(n, n')$, of the transition matrix $\underline{Q}_1$, is the conditional probability that, given arm 1 has the green light and there were n vehicles in arm 1 at the beginning of this phase there will be n' in arm 2 at the end of this phase. The matrix $\underline{Q}_1$ is given in Table 2. Since $\underline{t}_1$ is an eigenvector of $\underline{Q}_1$ it satisfies

$$\underline{t}_1' = \underline{t}_1'\underline{Q}_1 ,$$

i.e.

m	$t_1(m)$
0	0.00635
1	0.02964
2	0.06868
3	0.10837
4	0.13381
5	0.13963
6	0.12900
7	0.10871
8	0.08514
9	0.06316
10	0.04464
11	0.03034
12	0.01995
13	0.01275
14	0.00795
15	0.00485
16	0.00291
17	0.00171
.	.
.	.
.	.

TABLE 1: The vector $\underline{t}_1 = \{t_1(m)\}$; $t_1(m)$ is the probability that in the steady state there are m vehicles in arm 1 at the end of the red phase.

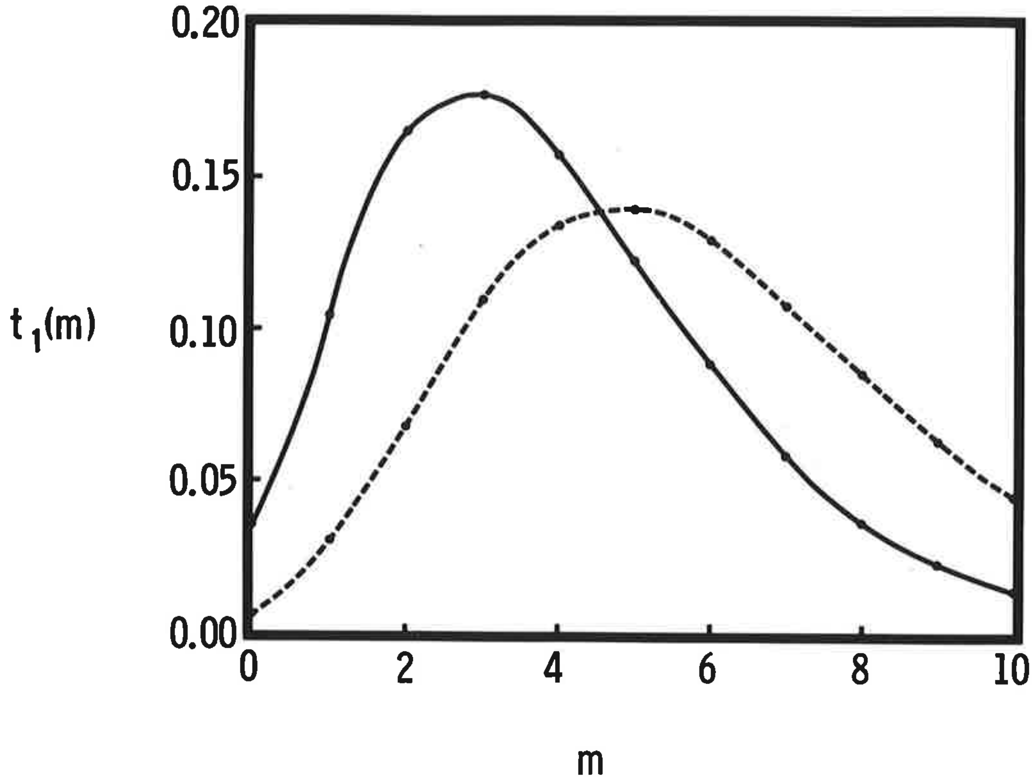


Fig. 10. Plot of the probability $t_1(m)$ that in the steady state there are m vehicles in arm 1 at the end of the red phase. The lost time is 4 seconds ($\ell=2$) for the solid curve and 6 seconds ($\ell=3$) for the dashed curve. The probabilities are defined only for integral values of m but curves have been drawn to facilitate comparison.

n \ n'	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0.10628	0.27969	0.30423	0.18744	0.08054	0.02878	0.00922	0.00275	0.00078	0.00021	0.00006	0.00001	0.00000
1	0.05034	0.17665	0.26962	0.24162	0.14821	0.07027	0.02825	0.01016	0.00337	0.00106	0.00032	0.00009	0.00003
2	0.02385	0.10459	0.20552	0.24286	0.19763	0.12170	0.06119	0.02662	0.01041	0.00376	0.00127	0.00041	0.00013
3	0.01130	0.05945	0.14290	0.21003	0.21453	0.16522	0.10220	0.05344	0.02459	0.01024	0.00395	0.00143	0.00049
4	0.00535	0.03286	0.09338	0.16427	0.20252	0.18864	0.14030	0.08713	0.04687	0.02248	0.00984	0.00399	0.00152
5	0.00253	0.01779	0.05835	0.11959	0.17298	0.18973	0.16597	0.12052	0.07509	0.04126	0.02045	0.00930	0.00394
6	0.00120	0.00948	0.03525	0.08249	0.13707	0.17336	0.17503	0.14635	0.10444	0.06524	0.03645	0.01854	0.00871
7	0.00057	0.00499	0.02074	0.05458	0.10247	0.14697	0.16859	0.16009	0.12939	0.09112	0.05705	0.03230	0.01678
8	0.00027	0.00260	0.01195	0.03492	0.07312	0.11737	0.15098	0.16086	0.14571	0.11470	0.07995	0.05015	0.02871
9	0.00013	0.00134	0.00676	0.02174	0.05024	0.08927	0.12738	0.15069	0.15154	0.13225	0.10173	0.07046	0.04427
10	0.00006	0.00069	0.00377	0.01323	0.03345	0.06519	0.10225	0.13343	0.14740	0.14153	0.11984	0.09079	0.06233
11	0.00003	0.00035	0.00208	0.00789	0.02168	0.04601	0.07871	0.11192	0.13542	0.14210	0.13138	0.10851	0.08104
12	0.00001	0.00018	0.00113	0.00463	0.01374	0.03153	0.05845	0.09017	0.11847	0.13502	0.13553	0.12142	0.09822

TABLE 2: The matrix $\underline{Q}_1$ whose (n, n') element is the conditional probability of there being n' vehicles in arm 2 at the end of a green phase for arm 1 given there were n in arm 1 at the beginning of the phase. The matrix is shown in a truncated form.

$$t_1(n') = \sum_{n=0}^{\infty} t_1(n)\, Q_1(n, n') .$$

In particular let n' = 5 so that

$$t_1(5) = 0.13963 = \sum_{n=0}^{\infty} t_1(n)\, Q_1(n, 5) .$$

If the process of truncation is valid, the sums S_N, defined by

$$S_N = \sum_{n=0}^{N} t_1(n)\, Q_1(n, 5) ,$$

will approach $t_1(5)$ as $N \to \infty$. In Table 3, a few values of S_N are given indicating the convergence to the element $t_1(5)$.

Means

As has previously been remarked, the behavior of the mean values in the case of the binomial arrivals is the same as that of the corresponding deterministic quantities in the case of steady arrivals under the same traffic conditions, so that this example is equally applicable to the case of steady arrivals (under the same control and with the same initial conditions).

The expected values of queue length etc., are found by substituting the given data into equations (4.74) - (4.86). As an example of the behavior of a mean value in the transient state equation (4.74) gives

$$\bar{n}_1(2j) = [n_1(0)-6](4/9)^j + 6$$

as the expected number of vehicles in arm 1 at the beginning of the

N	S_N
10	0.13715
12	0.13917
14	0.13955
17	0.13961
.	.
.	.
.	.
∞	0.13963

TABLE 3: The sums $S_N = \sum_{n=0}^{N} t_1(n)Q_1(n, 5)$ for N = 10, 12, 14, 17, indicating the convergence to $t_1(5) = 0.13963$.

(j+1) th green phase for this arm. It is clear that if $n_1(0) > 6$, i.e. if the initial number of vehicles is greater than the steady state mean number of vehicles, $\bar{n}_1(2j)$ will approach the value 6 monotonically from above. If $n_1(0) < 6$ the approach is from below and if $n_1(0) = 6$ then $\bar{n}_1(2j)$ is constant.

The other variables, such as $\bar{n}_2(2j-1)$, $\bar{M}_1(2j)$, defined in equations (4.74) - (4.80) all exhibit the same type of behavior, as the way in which they approach their limiting values depends on the sign of $n_1(0) - \bar{N}_1(\infty)$ and the rate at which they converge depends on the magnitude of the quantity (y/x).

The limiting forms $\bar{N}_1(\infty)$ etc., defined in (4.81) - (4.84), in this case assume the following values

$$\bar{N}_1(\infty) = \bar{N}_2(\infty) = 6 \ ,$$

$$\bar{g}_{1\infty} = \bar{g}_{2\infty} = 24 \text{ seconds} \ ,$$

$$\bar{C}_\infty = 60 \text{ seconds} \ .$$

Variances

To obtain a measure of the fluctuation about the mean of, for example, $n_1(2j)$ use is made of equation (4.88) for $v_{2j} = \text{var}\,[n_1(2j)]$. With the given data this is

$$\text{var}\,[n_1(2j)] = -\{(4/5)[n_1(0)-6]+9.36\}(4/9)^{2j}+(4/5)[n_1(0)-6](4/9)^{j}+9.36 \ .$$

As discussed in section 4.6 the behavior of var $[n_1(2j)]$ depends on the sign of $n_1(0)-\bar{N}_1(\infty) = n_1(0)-6$ in this case. When $n_1 \leq 6$ the approach to the steady state is monotonic and var $[n_1(2j)]$ is bounded above by var $[N_1(\infty)]$, and when $n_1(0) > 6$ the approach to the steady state is from above but the approach is no longer monotonic. These features are illustrated in Fig. 11 where plots of v_{2j} against j have been made for the two cases $n_1(0) = 5$ and $n_1(0) = 25$. The maximum value of v_{2j} in the latter case is 11.4 and occurs when $j = 2$. The steady state variances are, from (4.92) - (4.98),

$$\text{var } [N_1(\infty)] = \text{var } [N_2(\infty)] = 9.36 \ ,$$

$$\text{var } [M_1(\infty)] = \text{var } [M_2(\infty)] = 10.08 \ ,$$

$$\text{var } [g_{1\infty}] = \text{var } [g_{2\infty}] = 144 \text{ (seconds)}^2 \ ,$$

$$\text{var } [C_\infty] = 480 \text{ (seconds)}^2 \ .$$

Delay

The total expected delay, D, per cycle to all vehicles in one arm is from (4.106)

$$D^{(i)} = 252 \text{ seconds,}$$

and the average delay per vehicle is 21 seconds. In this case the average delay, d, per vehicle averaged over both lanes also has a value of 21 seconds.

Practical Considerations

It is desirable, for the practical application of $n_i = 0$ control, that the phase times should not become too large. For,

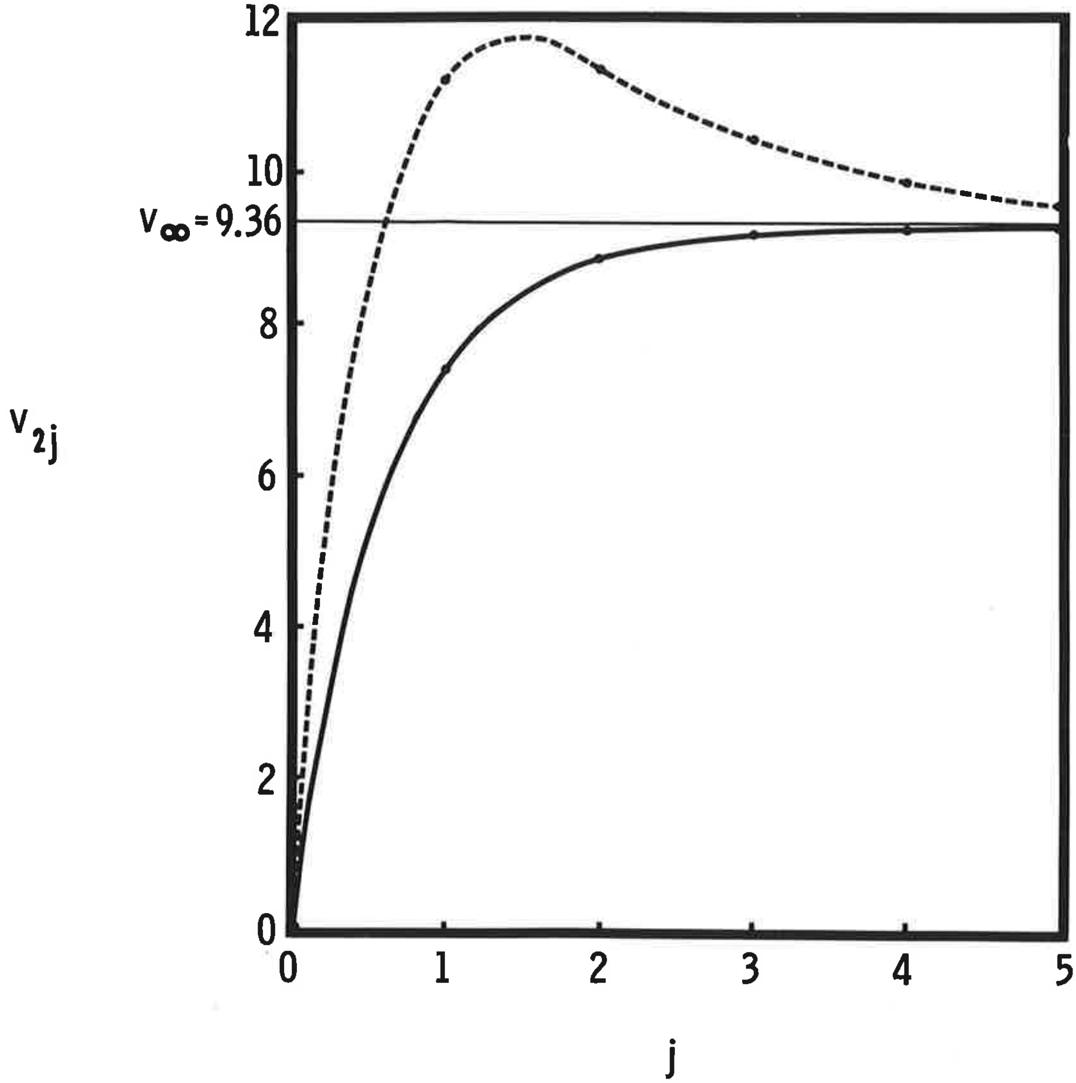
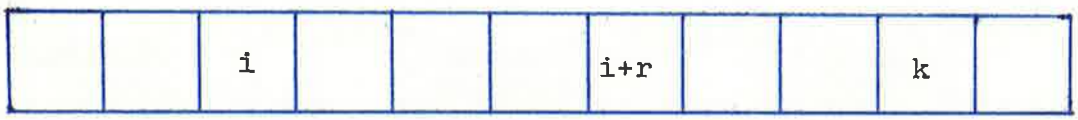


Fig. 11. Plot of the variance, v_{2j}, of the number of vehicles queued in arm 1 at the end of the jth red phase. The initial number of vehicles, $n_1(0)$, is 5 for the solid curve and 25 for the dashed curve. The variances are defined only for integral values of j, but the curves have been drawn to facilitate comparison.

Diagram A



if they should, the maximum phase restriction would be invoked and the control would not necessarily react to the existing traffic conditions. To indicate the frequency with which longer phase times will occur the probabilities of the occurrences of phases longer than those proposed by Webster [6] will be evaluated.

From equation (4.83), $\bar{g}_{i\infty} = 2\ell y_i/(1-Y)$ are the expected values of $g_{i\infty}$ and it is well known that these phase times are those which are consistent with the so called minimum cycle time (see Appendix A). As fluctuations in the arrival rates and, hence, in the queue lengths occur, the phase times will fluctuate about their mean values thus reacting to the actual traffic conditions. Webster [6] has found that, in the case of fixed lights, it is necessary to approximately double the phase times to allow for these fluctuations. From (4.65)

$$P[g_{1\infty} = k] = \binom{2\ell+k-1}{k} \left(\frac{x_2-y_1}{x_2}\right)^{2\ell} \left(\frac{y_1}{x_2}\right)^k$$

so that

$$P[g_{1\infty} \geq M + 1] = \sum_{k=M+1}^{\infty} \binom{2\ell+k-1}{k} \left(\frac{x_2-y_1}{x_2}\right)^{2\ell} \left(\frac{y_1}{x_2}\right)^k .$$

By changing the variable in the usual manner this may be written in the more manageable form

$$P[g_{1\infty} \geq M + 1] = \left(\frac{x_2-y_1}{x_2}\right)^{2\ell} \left(\frac{y_1}{x_2}\right)^M \sum_{k=0}^{2\ell-1} \binom{2\ell+M}{M+k+1} \left(\frac{y_1}{x_2-y_1}\right)^{k+1} . \qquad (4.109)$$

With the other traffic conditions as before (i.e., $\tau = 2$ seconds,

$y_1 = y_2 = 1-x_1 = 1-x_2 = 0.4$) three cases will be considered, viz $\ell = 1, 2, 3$.

Now $\bar{g}_{1\infty} = 2\ell y_1/(1-Y) = 4\ell$ so that Webster would use a phase time of approximately 8ℓ. Setting $M + 1 = 8\ell$ in (4.109) and using the indicated values of y_i, x_i leads to

$$P[g_{1\infty} \geq 8\ell] = (\tfrac{1}{3})^{2\ell} (\tfrac{2}{3})^{8\ell-1} \sum_{k=0}^{2\ell-1} \binom{10\ell-1}{8\ell+k} 2^{k+1} .$$

Setting $\ell = 1, 2, 3$ successively yields

$$P[g_{1\infty} \geq 8] = .143, \quad P[g_{1\infty} \geq 16] = .079, \quad P[g_{1\infty} \geq 24] = .045 .$$

Hence when $\ell = 1$ a green time as long as, or longer than, that proposed by Webster occurs approximately once every seven cycles, for $\ell = 2$ the frequency is about once in thirteen, and for $\ell = 3$ about once in twenty-two.

A further practical requirement is that the queues should never become very long because if this were so a back-up over the previous intersection might occur. From (4.66)

$$P[M_1(\infty)=k] = \left(\frac{x_1x_2-y_1y_2}{y_2}\right)^{2\ell}\left(\frac{y_1y_2}{x_1x_2}\right)^k \sum_{j=1}^{2\ell} \binom{2\ell}{j}\binom{j+k-1}{k}\left(\frac{-1}{x_2}\right)^j (\text{for } k > 0) ,$$

and

$$P[M_1(\infty) \geq M] = \sum_{k=M}^{\infty} P[M_1(\infty) = k] .$$

Inserting the given data leads to

(i) $P[M_1(\infty) \geq 8] = .0082$, for $\ell = 1$, and

(ii) $P[M_1(\infty) \geq 12] = .0094$, for $\ell = 2$.

Hence the maximum queue length, i.e. the number of vehicles queued at the instant when the queue starts to discharge, will exceed seven only once in one hundred cycles in the case $\ell = 1$ and will exceed eleven only once in one hundred cycles in the case $\ell = 2$.

CHAPTER 5

DISCUSSION

The intersection model considered in this thesis is of a simple yet quite practical nature because, as has been mentioned before, in most cases where intersections are controlled by some form of vehicle-actuated traffic lights, 2 lanes usually predominate in determining the phase lengths. At most signalized intersections turning vehicles are allowed to make their manoeuvre whenever the opportunity arises (except where explicit allowance is made for them by the presence of a "turning arrow") and for this reason they have been neglected in the preceding theory. The control has been constructed in such a way that it could readily be incorporated in a computerized control system. Furthermore the control is of a dynamic rather than static nature reacting as it does to the present state of the traffic at all times.

In the first instance a continuum approximation has been used, together with the assumption of constant arrival rates. Although such a system would never be realized in practice it has the advantage that an exact analytic solution is possible and, as has been found by other authors, provides an excellent insight into the behavior of more complex systems. In particular, the system with binomial arrivals behaves, in the mean, exactly as the system with constant arrivals.

As the aim of the thesis is to analyze the behavior of a computer controlled intersection, the case of binomial rather than Poisson arrivals has been investigated. As described in section 4.1 the computer used in the Toronto pilot study receives information concerning the arriving traffic in binary form. More particularly, the computer scans the intersection at regular intervals and if a detector has registered a pulse, denoting that a vehicle has crossed it since the last scanning, the computer stores this information in the form of a "1". Otherwise the computer records a "0" corresponding to no pulse or no arriving vehicle. This binary form of input immediately suggests that the arrivals might be thought of as having been generated by a binomial process. For this assumption to be valid it is necessary that no more than one vehicle should cross the detector in any scanning interval and secondly that the series of 0's and 1's should be uncorrelated.

In the pilot study it was found that for reasonable values of the scanning interval (say, of the order of 1-2 seconds) the event of two or more vehicles crossing a detector in the same interval was quite rare. In fact a separation of 1-2 seconds between vehicles in the same lane would correspond to what is usually recognized as a saturated or almost saturated flow. The assumption that the 0's and 1's are uncorrelated would clearly not be valid for heavy or pulsed traffic.

The analysis of the steady arrival case for both $n_i = 0$ and

$f_i = 0$ controls showed that, in the steady state, the former control has certain desirable features. For instance, vehicles are forced to stop at most once and the control also gives optimization with respect to minimum average delay in most cases. Grafton and Newell [20] showed further that, in the transient state, $n_i = 0$ control is again optimal with respect to minimum average delay except in certain particular cases such as those in which one queue is initially very large. In the transient state also $n_i = 0$ control obviously has the property that no vehicle should stop more than once. For these reasons no attempt has been made to incorporate $f_i = 0$ control into the binomial arrivals model. This has the disadvantage that, in a practical situation, the maximum phase restriction may have to be invoked too often, but this disadvantage is not considered serious since it appears that only under $n_i = 0$ control is an exact solution obtainable. Furthermore, this solution readily yields the operational characteristics of the system. As exact expressions have been found for the important probability generating functions involved in the system, all moments of the corresponding distributions are readily obtainable in contrast to many works in which it has proved impossible to obtain moments of order higher than the first.

For a practical application of the control by a computer it would be important to interpret the number of vehicles in the arms correctly, according to the type of detectors being used, as the control algorithm depends directly upon the queue length. In fact it is rather hard to define what "queue length" might mean in a practical situation and the need for some experimental work to indicate a suitable definition might be desirable. Similarly it may not be easy to determine the exact time at which the queue empties. Here, too, there is a need for some working criterion.

It has been found that a magnetic loop detector is capable of satisfying the requirements of feeding the necessary information concerning queue lengths, etc. to the computer provided that it is located in the correct position. It was suggested during the Toronto pilot study that a magnetic detector placed about 300' from an intersection would be suitable in two respects. Firstly, this placing is such that vehicles which have crossed the detector but are waiting on the red will be cleared during the next green phase and secondly, a queue length of 300' indicates, roughly, a critical condition at the intersection in that observations showed that when queues reach a length of about 300' an intersection is approaching its capacity. In the control strategy suggested by Miller [18], the placement of detectors is again of vital importance. Miller suggests two possible ways of estimating queue lengths viz., the use of one large magnetic loop detector which would give a direct measurement of queue length or of a system of two counters, one on the approach and one on the stop line, giving the queue length by subtraction. Miller also discusses the placement of detectors on the approach arms.

As to the question of when a queue has emptied one could perhaps use the criterion that a queue has cleared when the flow across the intersection drops below saturation flow. An indication as to when this has happened could possibly be obtained by means of another detector placed on the exit arm so that a gap of greater than τ seconds between successive departing vehicles would imply that the queue had cleared (remembering that while a queue persists it discharges at a constant rate of one vehicle every τ seconds).

Other information, such as vehicle speed, which may be determined, for example, as the ratio of average car length (say, 17') to length of registered pulse, may be transmitted from the detectors to the computer for use in determining control constants and so on. Furthermore, it is to be anticipated that the computer will eventually "recognize" many recurring traffic patterns and act on the basis of past experience. However, there will still be some need for the human element as there will always be, outside the computer's experience, events to which it will react slowly.

Several interesting problems are suggested from this work as topics for further research. One could analyze the operation of a system closely resembling the familiar vehicle actuated lights by programming the computer to initiate change of phase when a gap of $k\tau$ seconds between successive departures is detected. Another extension to the theory might be the introduction of some degree of correlation between the vehicles arriving at (or crossing) a point (i.e. between the 0's and 1's). A method to do this could be via a simple 2-state Markov chain with matrix

$$\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} p' & q' \\ q & p \end{bmatrix} \end{array}$$

where the state "0" represents no cars at a point and the state "1" represents one car at a point. Then p is the probability of having one car at a point at an instant after one in which there was one car at the point, and $q = 1 - p$ is the probability of having no car after this instant, etc. Some theory on this subject has been developed by Seth and Gupta [26].

APPENDIX A

OPTIMUM PHASE TIMES IN THE STEADY STATE

The problem of deriving phase times which minimize the average delay per vehicle for the case of constant arrival and departure rates in the steady state has been examined by Wardrop [5]. The method presented here, which differs from that of Wardrop, was suggested by G. F. Newell in a series of lectures given at the University of Adelaide in 1962.

Consider a fixed 2-phase signal in which the total green time (including lost time) for arm 1 is G, and that for arm 2 is R. All other notation is as given in section 3.2.

The conditions that the arms are not saturated (or equivalently that both queues clear each cycle and hence do not accumulate indefinitely) are

$$s_1(G-\ell) \geq q_1(G+R) \ , \tag{A1}$$

$$s_2(R-\ell) \geq q_2(G+R) \ , \tag{A2}$$

which may be rewritten as

$$G \geq (y_1R+\ell)/(1-y_1) \ , \tag{A3}$$

$$R \geq (y_2G+\ell)/(1-y_2) \ . \tag{A4}$$

The minimum values of G and R which satisfy these inequalities are given by

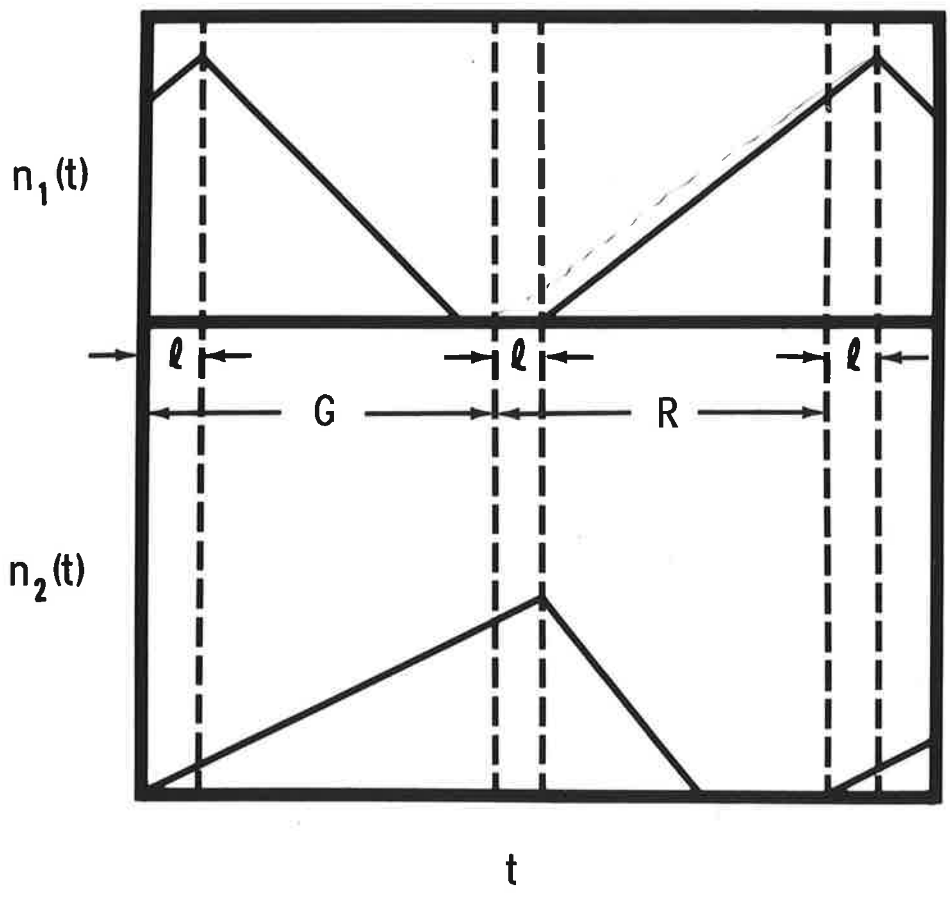


Fig. A1. Number of vehicles queued in arm i (i=1,2) as a function of time for the case of a fixed cycle traffic light with phases G, R respectively.

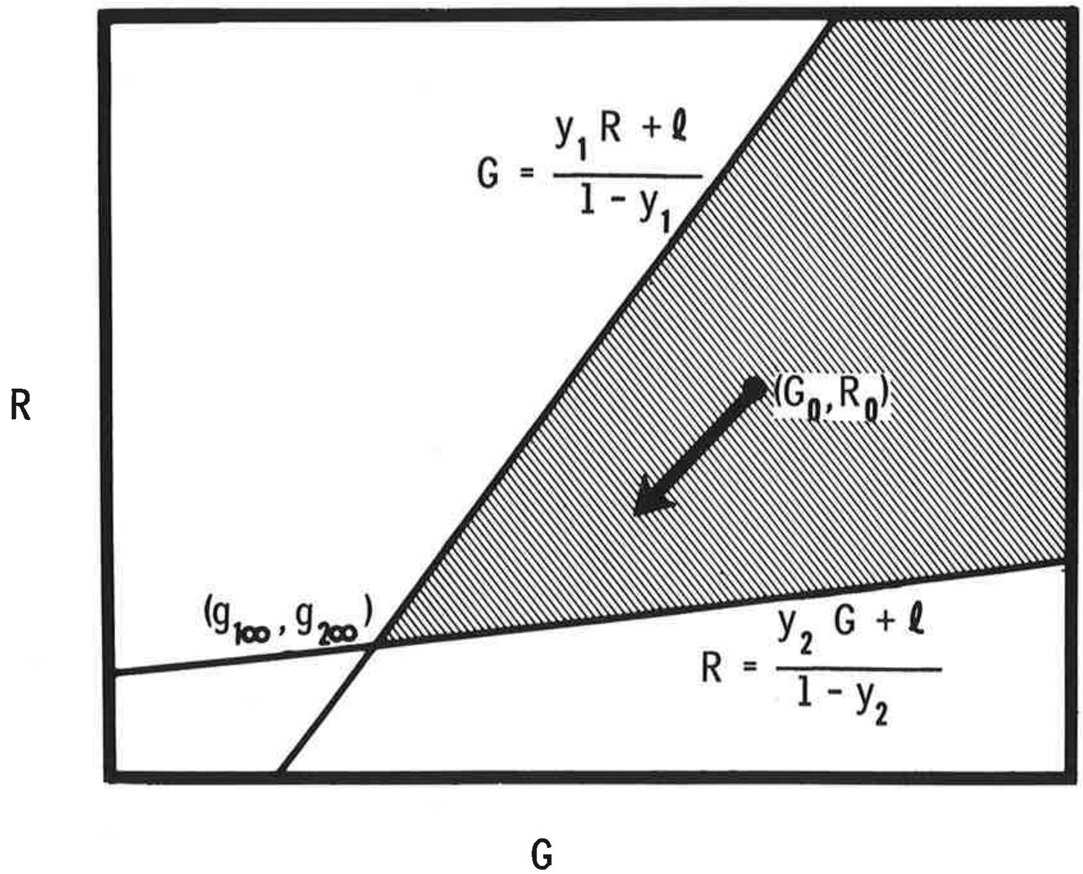


Fig. A2. The wedge, defined by (A3), (A4), in which the average delay, $\bar{D}$, is to be minimized.

$$G - \ell = 2\ell y_1/(1-Y) \ , \tag{A5}$$

$$R - \ell = 2\ell y_2/(1-Y) \ , \tag{A6}$$

which are the same as the steady state phase times given in (3.56) namely $g_{1\infty}$ and $g_{2\infty}$.

The average delay per vehicle, $\bar{D}$, is found by evaluating the area under the graphs of queue length against time in Fig. A1. The total delay to arm 1 over one full cycle is

$$D_1 = q_1(R+\ell)^2/[2(1-y_1)] \ ,$$

and to arm 2 is

$$D_2 = q_2(G+\ell)^2/[2(1-y_2)] \ ,$$

and

$$\bar{D} = (D_1+D_2)/[(q_1+q_2)(G+R)] \ .$$

The problem of finding R and G to minimize the average delay per vehicle thus reduces to:

$$\text{minimize} \quad \bar{D}(G, R) = \frac{(1-y_2)q_1(R+\ell)^2+(1-y_1)q_2(G+\ell)^2}{2(1-y_1)(1-y_2)(q_1+q_2)(G+R)} \ , \tag{A7}$$

subject to (A3) and (A4). The inequalities (A3), (A4) define a wedge in the first quadrant of the G, R plane with vertex at $g_{1\infty}$, $g_{2\infty}$ as shown in Fig. A2.

<u>Theorem</u>: The minimum value of $\bar{D}(G, R)$ occurs at a point on the boundary of the wedge defined by (A3), (A4).

Proof: Suppose the minimum occurs at some interior point (G_o, R_o) of the wedge. Consider the effect on the value of $\bar{D}$ of moving from the point (G_o, R_o) to a point $(G_o-\varepsilon, R_o-\varepsilon)$ where ε is positive. In particular let

$$h(\varepsilon) = (R+\ell-\varepsilon)^2/(R+G-2\varepsilon) .$$

Now

$$h'(\varepsilon) = 2(R+\ell-\varepsilon)(\varepsilon-G+\ell)/(R+G-2\varepsilon)^2$$

so that there are turning points of $h(\varepsilon)$ at the points $\varepsilon = R+\ell$ and $\varepsilon = G-\ell$ both of which are positive since $G > \ell$. If $R + \ell > G - \ell$ the maximum value of $h(\varepsilon)$ (for $\varepsilon > 0$) will occur at either $\varepsilon = 0$ or $\varepsilon = R + \ell$. Since $h(0) > 0$ and $h(R + \ell) = 0$ the maximum occurs at $\varepsilon = 0$. Similarly if $G - \ell > R + \ell$ the maximum will occur at $\varepsilon = 0$ or $\varepsilon = G - \ell$. Now

$$h(0) - h(G-\ell) = (R+\ell)^2/(R+G) - (R-G+2\ell) ,$$

$$= (G-\ell)^2/(R+G) ,$$

so that the maximum again occurs at the point $\varepsilon = 0$. Hence $h(0) > h(\varepsilon)$ for all $\varepsilon > 0$. In a similar manner it can be shown that $k(\varepsilon) = (G+\ell-\varepsilon)^2/(R+G-2\varepsilon)$ has the same property, viz. $k(0) > k(\varepsilon)$ for $\varepsilon > 0$. Thus for any (G_o, R_o) and $(G_o-\varepsilon, R_o-\varepsilon)$ in the wedge

$$\bar{D}(G_o-\varepsilon, R_o-\varepsilon) < \bar{D}(G_o, R_o) \text{ for } \varepsilon > 0 .$$

This contradicts the assumption that the minimum occurs at an interior point.

The property that the minimum occurs on the boundary of the wedge reduces the problem of minimizing $\bar{D}(R, G)$ to a one-dimensional problem. For, suppose the minimum occurs on $G = (y_1 R + \ell)/(1-y_1)$, then

$$\bar{D}(R, G) = \bar{\Delta}(R) = \frac{(1-y_2)q_1(R+\ell)^2+(1-y_1)q_2[\frac{y_1}{1-y_1}(R+\ell)+2\ell]^2}{2(1-y_1)(1-y_2)(q_1+q_2)(\frac{R+\ell}{1-y_1})}$$

$$= [\frac{(1-y_2)q_1+\frac{y_1^2 q_2}{1-y_1}}{2(1-y_2)(q_1+q_2)}](R+\ell) + \frac{2\ell^2(1-y_1)q_2}{(1-y_2)(q_1+q_2)}(R+\ell)^{-1}$$

$$+ \text{ constant.} \qquad \text{(A8)}$$

This will be a minimum when

$$\frac{d\,\bar{\Delta}(R)}{dR} = 0 \; ,$$

or equivalently

$$\frac{d\bar{\Delta}(R+\ell)}{d(R+\ell)} = 0 \; ,$$

i.e. $$\frac{1}{2}[(1-y_2)q_1 + \frac{y_1^2 q_2}{1-y_1}] = 2\ell^2(1-y_1)q_2(R+\ell)^{-2} \; ,$$

i.e. $$(R+\ell)^2 = \frac{4\ell^2(1-y_1)^2}{(1-y_1)(1-y_2)\frac{q_1}{q_2}+y_1^2} \; . \qquad \text{(A9)}$$

This solution will only be valid if

$$\frac{4\ell^2(1-y_1)^2}{(1-y_1)(1-y_2)\dfrac{q_1}{q_2}+y_1^2} \geq (g_{2\infty}+2\ell)^2 \ ,$$

which may be reduced to

$$(1-y_1)(1-q_1/q_2) \geq Y \tag{A10}$$

in which case the optimum cycle is given by

$$C = 2\ell[(1-y_1)(1-y_2)q_1/q_2+y_1^2]^{-\frac{1}{2}} . \tag{A11}$$

Similarly the condition that the minimum occurs on $R = (y_2G+\ell)/(1-y_2)$ leads to the condition

$$(1-y_2)(1-q_2/q_1) \geq Y \tag{A12}$$

with corresponding optimum cycle

$$C = 2\ell[(1-y_1)(1-y_2)q_2/q_1+y_2^2]^{-\frac{1}{2}} \tag{A13}$$

The equivalence of these results to those of Wardrop is easily demonstrated. The inequality (A10) yields

$$q_2(y_1+y_2) \leq (1-y_1)(q_2-q_1) \ ,$$

i.e. $$q_1(1-y_1) - q_2(1-y_2) \geq 2q_1y_2$$

i.e. $$q_1(1 - \frac{q_1}{s_1}) - q_2(1 - \frac{q_2}{s_2}) \geq \frac{2q_1q_2}{s_2} .$$

It follows that the condition that the minimum cycle should be optimal is

$$q_1(1 - \frac{q_1}{s_1}) - q_2(1 - \frac{q_2}{s_2}) < \frac{2q_1q_2}{s_2}$$

and

$$q_2(1 - \frac{q_2}{s_2}) - q_1(1 - \frac{q_1}{s_1}) < \frac{2q_1q_2}{s_1}$$

as obtained by Wardrop.

It should be noticed that if $q_1 = q_2$ neither (A10) nor (A11) can be satisfied hence the minimum cycle is always optimal for equal arrival rates on the two arms. Further, if $q_1 \neq q_2$, one of the quantities $(1-q_1/q_2)$, $(1-q_2/q_1)$ is negative and therefore only one of (A 10), (A11) can be satisfied.

APPENDIX B

SUMMARY OF NOTATION

The following is a list of the more important symbols and definitions used in the thesis listed chapter-wise:

Chapter 2

At time t (measured from an arbitrary origin),

$n_i(t)$ = number of vehicles queued in arm i,

$f_1(t) = \alpha_1 n_1(t) + \beta_1 - n_2(t)$,

$f_2(t) = \alpha_2 n_2(t) + \beta_2 - n_1(t)$,

where α_i, β_i are control constants chosen in the range $\alpha_i > 1$, $\beta_i > 0$,

$T(t)$ = time elapsed since last change of phase,

r_i, R_i = minimum, maximum allowable durations of phase for arm i.

Chapter 3

Following the notation of Webster [6],

q_i = arrival rate in arm i,

s_i = departure rate (saturation flow) in arm i,

$y_i = q_i/s_i$,

$Y = y_1 + y_2$,

ℓ = lost time for a single phase,

g_j = effective green time for the jth phase.

If time t is measured from the beginning of the phase j = 1, then

$$t_j = j\ell + g_1 + g_2 + \dots + g_j ,$$

and, for simplicity of notation

$$n_i(j) = n_i(t_j) .$$

In particular,

$n_1(2j)$ = number of vehicles queued in arm 1 at the beginning of the (j+1)th green phase for this arm,

$n_1(2j+1)$ = number of vehicles queued in arm 1 at the end of this phase,

$n_2(2j+1)$ = number of vehicles queued in arm 2 at the beginning of the (j+1)th green phase for this arm,

$n_2(2j+2)$ = number of vehicles queued in arm 2 at the end of this phase.

In addition,

$M_1(2j)$ = number of vehicles queued in arm 1 at the beginning of the (j+1)th effective green phase for this arm,

$M_2(2j+1)$ = number of vehicles queued in arm 2 at the beginning of the (j+1)th effective green phase for this arm, and

g_{2j+1} = duration of the (j+1)th effective green phase for arm 1,

g_{2j+2} = duration of the (j+1)th effective green phase for arm 2,

$C_j = g_{2j-1}+g_{2j}+2\ell$ = duration of the jth cycle.

The initial lengths of the queues are $n_1(0)$, $n_2(0)$ respectively and a complete summary of solutions to the control equations under these initial conditions is given in section 3.6.

Furthermore,

$D^{(i)}$ = total delay to vehicles in arm i per cycle,

$D = D^{(1)} + D^{(2)}$ = total delay to all vehicles per cycle,

$\bar{D}^{(i)}$ = average delay per vehicle to vehicles in arm i,

$\bar{D}$ = average delay per vehicle averaged over both lanes with equal weight to each vehicle.

Chapter 4

The notation of Chapter 4 follows closely that of Chapter 3 but the following should be noted:

τ = the time between successive departures and is the same for both arms,

y_i = the probability of one arrival in arm i in each of the intervals $(k\tau, k\tau+\tau)$, $k = 0, 1, 2, \ldots$,

$x_i = 1-y_i$ = the probability of no arrival,

$g(j)$ has been written for g_j and $C(j)$ for C_j for ease of manipulation,

because only $n_i=0$ control is considered $n_1(2j+1) = n_2(2j) = 0$,

the initial queue lengths are $n_1(0)$ and zero.

In addition,

$\phi_1^{(2j)}(z) = E[z^{M_1(2j)}]$, is the generating function for $M_1(2j)$,

$\phi_2^{(2j+1)}(z) = E[z^{M_2(2j+1)}]$, is the generating function for $M_2(2j+1)$,

$\theta_1^{(2j)}(z) = E[z^{n_1(2j)}]$, is the generating function for $n_1(2j)$,

$\theta_2^{(2j+1)}(z) = E[z^{n_2(2j+1)}]$, is the generating function for $n_1(2j+1)$,

$\Gamma_1^{(2j+1)}(z) = E[z^{g(2j+1)}]$, is the generating function for $g(2j+1)$,

$\Gamma_2^{(2j+2)}(z) = E[z^{g(2j+2)}]$, is the generating function for $g(2j+2)$, and

$\Delta^{(j)}(z) = E[z^{C(j)}]$, is the generating function for $C(j)$.

The limiting forms of these generating functions as $j\to\infty$ are,

$\phi_1(z)$, $\phi_2(z)$, $\theta_1(z)$, $\theta_2(z)$, $\Gamma_1(z)$, $\Gamma_2(z)$, and $\Delta(z)$ respectively.

Other often used notations in this chapter are,

$$\omega_1(z) = \frac{x_1(x_2+y_2z)}{1-y_1(x_2+y_2z)} ,$$

$$\omega_2(z) = \frac{x_2(x_1+y_1z)}{1-y_2(x_1+y_1z)} ,$$

$\xi(z) = \omega_1[\omega_2(z)]$, $\eta(z) = [x_2+y_2\omega_2(z)][x_1+y_1z]$, and

$\xi^{(j)} = \xi(\xi(\ldots\xi(z)\ldots))$ is the jth iterate of $\xi(z)$ and

$\eta(z)$, $\eta[\xi^{(j)}]$ have been abbreviated to η, $\eta\xi^{(j)}$ respectively.

Much of the notation used in the Chapter follows that of Feller [22].

For example, if X and Y are two random variables then,

the expected value of X will be denoted by $E[X]$ or by $\bar{X}$,

the variance of X by var $[X]$,

the conditional distribution of Y for given X by $Y|X$, and the conditional expectation of Y for given X by $E[Y|X]$.

Appendix A

In this section

G = total green time = effective green time + lost time, and

R = total red time (including one lost time).

REFERENCES

1. W. F. Adams, "Road Traffic Considered as a Random Series", J. Instit. Civil Eng. 4, 121 (1936).

2. P. G. Pak-Poy, "The Limits of the Poisson Distribution in Road Traffic Calculations", Thesis (M. Tech.) Traffic Engineering, Univ. of New South Wales, 1959.

3. A. J. H. Clayton, "Road Traffic Calculations", J. Instit. Civil Eng. 16, 247-284, 558-594 (1941).

4. F. Garwood, "An Application of the Theory of Probability to the Operation of Vehicular-controlled Traffic Signals", Supp. J. Roy. Stat. Soc. 7, 65 (1940-41).

5. J. G. Wardrop, "Some Theoretical Aspects of Road Traffic Research", Proc. Instit. Civil Eng. Part II, 1, 2, 325-362 (1952).

6. F. V. Webster, "Traffic Signal Settings", Road Research Technical Paper No. 39, Road Research Laboratory, H.M.S.O., England, 1958.

7. A. J. Miller, "Settings for Fixed-Cycle Traffic Signals", Opnal. Res. Quart. 14, 373-386 (1963).

8. M. Beckmann, C. B. McGuire, and C. B. Winsten, "Studies in the Economics of Transportation", Yale University Press, 1956.

9. G. F. Newell, "Queues for a Fixed-Cycle Traffic Light", Ann. Math. Stat. 31, 589-597 (1960).

10. M. C. Dunne and R. B. Potts, "Analysis of a Computer Control of an Isolated Intersection", Proceedings of the Third International

Symposium on the Theory of Traffic Flow, New York, June 21-23, 1965 (to be published).

11. G. F. Newell, "Statistical Analysis of the Flow of Highway Traffic Through a Signalized Intersection", Quart. Appl. Math. 13, 353-369 (1956).

12. J. N. Darroch, "On the Traffic-Light Queue", Ann. Math. Stat. 35, 380-388 (1964).

13. G. F. Newell, "Approximation Methods for Queues with Application to the Fixed-Cycle Traffic Light", S.I.A.M. Review 7, 223-240 (1965).

14. Tosio Uematu, "On Traffic Control at an Intersection Controlled by the Repeated Fixed-Cycle Traffic Light", Ann. Instit. Stat. Math. 9, 87-107 (1958).

15. J. C. Tanner, "A Problem of Interference Between Two Queues", Biometrika 40, 58-69 (1953).

16. J. N. Darroch, G. F. Newell, and R. W. J. Morris, "Queues for a Vehicle-Actuated Traffic Light", Opns. Res. 12, 882-895 (1964).

17. M. G. Grace, R. W. J. Morris, and P. G. Pak-Poy, "Some Aspects of Intersection Capacity and Traffic Signal Control by Computer Simulation", Second Conference Australian Road Research Board, 1964.

18. A. J. Miller, "Computers in the Analysis and Control of Traffic", Traff. Quart. 19, 556-572 (1965).

19. M. C. Dunne and R. B. Potts, "Algorithm for Traffic Control",

Opns. Res. 12, 870-881 (1964).

20. J. T. Hewton, "Toronto Installs All-Purpose Electronic Computer for Traffic Signal Control", Traff. Eng. Cont. 5, 589-594 (1964).

21. R. B. Grafton and G. F. Newell, "Optimal Policies for the Control of an Undersaturated Intersection", Proceedings of the Third International Symposium on the Theory of Traffic Flow, New York, June 21-23, 1965 (to be published).

22. F. L. Spitzer, "Principles of Random Walk", Van Nostrand, 1964.

23. W. Feller, "An Introduction to Probability Theory and its Applications", Vol. 1 (2nd. Ed.), Wiley, New York, 1957.

24. T. E. Harris, "The Theory of Branching Processes", Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen, Band 119, Springer-Verlag, 1963.

25. M. J. Gottlieb, "Concerning Some Polynomials Orthogonal on a Finite or Enumerable Set of Points", Amer. J. Math. 60, 453-458 (1938).

26. A. Seth and H. C. Gupta, "Random Walk with and without Correlation in the Presence of Reflecting Barriers", Proc. Nat. Inst. Sci. India 30, A, No. 1 (1964).