TRAFFIC SIMULATION:

A SURVEY

A Thesis

Presented to

the Faculty of Graduate Studies and Research

University of Manitoba

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Computer Science

by

David Abrams

May, 1972

Abstract of Thesis

Within the last decade, traffic flow theory and the rapid improvement of the electronic digital computer have been mutually responsible for the development of simulation as a design tool for examining traffic phenomena. While a considerable quantity of literature has been published concerning the theory and simulation of traffic systems, few general laws have been determined. In addition, little has been accomplished concerning values and relationships suitable for purposes of simulation. This, then, is the subject of this thesis.

Here, selected topics are examined for their application to the theory of traffic flow and for their relationship to modelling traffic behaviour at an intersection. As each mechanism is introduced, a comprehensive state-of-the-art survey is undertaken to evaluate studies and recommendations made in previous research projects. Following this, an attempt is made to establish functional relationships and characteristic distributions for each mechanism.

The last part of this thesis is devoted to traffic signalization. The intersection has always been the cardinal element in street systems that incorporate signalized intersections. To improve the operational effectiveness of these systems, it is necessary to develop traffic signalization to its highest

level of efficiency. To this end, an extensive examination of proposed and existing signal timing schemes has been undertaken. In investigating the different timing methods, behavioral relationships and vehicular operating characteristics have been used which were discussed earlier in this thesis.

Acknowledgements

I wish to thank my supervisor, Jim Wells, for his guidance and assistance in the preparation of this project. I am also indebted to Dr. A. N. Arnason for his helpful criticisms and suggestions concerning the presentation of work in this manuscript. In addition, my thanks to Dr. A. Soliman for his time spent reading this thesis as a Committee Member.

The help and initial support given to me by my former advisor, Don Costin, was instrumental in the planning of this thesis and is gratefully acknowledged.

I would like to extend my thanks to the Metropolitan Corporation of Greater Winnipeg and the Streets and Traffic Department of the City of Winnipeg for the very generous use of their publications and reference manuals.

My appreciation is extended to Margaret Thomson for her patience and cooperation in the typing of the final draft of this thesis.

Finally, to the many people who, by their concern and words of encouragement saw this thesis through many a difficult period of time, I record my sincerest gratitude.

Table of Contents

CHAPTER ONE

INTRODUCTION

The original purpose of this thesis was to examine, through simulation, the effect of bus routes on the flow of traffic at an intersection. Ultimately, it was hoped that the simulation could be expanded to encompass an artery n blocks in length. With this purpose in mind, chapters 3 and 4 were written with the intention of determining actual values for use in the simulation model.

However, as time progressed, it became increasingly obvious that this goal was totally unrealistic. While an excessive amount of time was required to obtain and examine the available literature, the lack of available information made the development of a viable model prohibitive. This was particularly true in the case of information concerning local transit vehicles. Vehicular operating characteristics had either never been investigated or, if investigated, had never been published. Thus, in chapter 8, the deceleration capability of local buses could only be roughly estimated. The acceleration capability was determined from tests conducted on Western Flyer buses and obtained only through private communications.

Vehicular behavior relationships as they relate to transit vehicles were singularly absent from the literature. Although the time spent by buses picking up and discharging passengers at a stop could be expected to have a severe influence on the capacity of an intersection, no thorough investigation has ever been conducted to determine the exact nature of this influence. In fact, in discussing this topic, the Highway Capacity Manual (1965) could offer only nomograms which were "largely rationalizations developed from limited available knowledge."

For this reason, plans to develop an actual simulation model were abandoned. This thesis, then, is devoted to examining a selection of different topics. There is a shift in emphasis from determining values which explicitly pertain to simulation, to an analysis of mechanisms which have a general application to the theory of traffic flow.

Thus, in chapter 5, while the composite distribution is examined for its explicit application to simulation, it is potentially a general distribution for characterizing vehicular headways. In a like manner, the log-normal distribution discussed in chapter 9 is suitable for describing, in general, the gap/lag acceptance phenomena. Finally, in chapter 12, an extensive examination of traffic signalization as it relates to sumulation is presented, and a comparison of different timing methods is performed.

This thesis, then, is an examination of various characteristics of traffic flow and of the models that have been developed for them.

While a great quantity of literature has been published in the last ten years, under the classification of traffic flow theory, few general laws, as yet, exist. Functional relationships and distributions characterizing traffic phenomena are surprisingly absent from the literature. Values, obtained in studies conducted in particular locations for specific purposes, remain just that. These values apply only to the situations studied and contribute little to the theory of traffic flow. Anyone attempting to study traffic behavior through simulation is immediately faced with this problem. Therefore, if there is an underlying objective towards which studies in traffic flow should be directed, then surely it is the development of relationships and mechanisms which characterize traffic phenomena.

In this thesis, an attempt has been made to bring together some of the sources of information which are available, and present some of the relationships and mechanisms which are suitable for simulation. If this has been accomplished, then perhaps a great deal of human effort has not been totally wasted.

CHAPTER TWO

WHY SIMULATE ?

2.1 Introduction

In this chapter, the opinions of different researchers involved in traffic analysis are presented, with particular emphasis given to the relative merits of simulation when compared to other forms of scientific investigation. Their opinions are uniformly pro-simulation, and while the advantages of simulation are, perhaps, a debatable subject, it cannot be denied that these men have contributed much towards the development of traffic flow theory and the solutions of its control.

This is not to suggest that there are no disadvantages associated with simulation. In simulation, it is necessary, at some point, to test both the model and the solution technique used. To do this, actual field data must be gathered and analyzed to ensure that the model both accurately represents the actions and interactions of the designed situation, and generates a realistic solution. Obtaining such data can be very difficult.

2.2 Goode

"There is one aspect of the attack on a problem in which simulation is better than both analysis and trial... In analysis we may use only those criteria which are mathematically tractable (eg. least squares, but not maximum absolute deviation), and in trial, we may choose only one criterion because even one is costly to measure. In simulation, we may select any criterion and as many as we like, measuring them continually if necessary." (Gerlough (1967),p.3)

2.3 Gerlough

"It was decided to develop a microscopic digital simulation model. An analytical model approach was considered unacceptable since either it would be necessary to omit certain requirements, such as the interest in individual vehicles in a heterogeneous stream and the versatility to represent a diverse group of control methods, or else achieve a model with intractable mathematics.

A strictly empirical approach was also considered. Though the requirement of realism would be satisfied, for reasons relating to cost of study, time required, quality of study (precision of control over variables, reproducibility, degree of detail of traffic-performance measurements, and statistical reliability), and safety, the advantages of simulation were overwhelming." (Gerlough (1967),p.4)

2.4 Kell

"A deterrent to field evaluation to determine changes in intersection performance is the necessity for installing or implementing each change - e.g. (traffic signal timing), waiting for traffic to adapt to it, and then collecting additional data to determine the effect. The great amount of effort involved is seldom expended.

Some potential control schemes or timing plans are not attempted in the field because of possible adverse effects on traffic. It is difficult to attempt studies for purely research purposes when such possibilities exist.

The current state of knowledge of mathematical formulae and traffic flow theory do not provide the answers needed to effectively and efficiently operate an intersection. Simulation can help to overcome the above limitations because of the following capabilities:

(1) Traffic conditions (volumes, turning movements, etc.) can be precisely specified for the analysis.

(2) Enough hours of constant traffic conditions can be simulated to provide confidence that the average result is a reasonable estimate of the mean of the entire distribution of values for that set of conditions.

(3) Control conditions can be easily and rapidly altered without requiring a field installation.

(4) Precisely the same traffic can be reproduced to analyze each change. Thus unpredictable field variations are eliminated.

(5) Large amounts of data can be simulated in a relatively short time, permitting detailed evaluation of many control variables under a variety of traffic conditions." (Kell (1964), pp.127-128)

2.5 Lewis and Michael

"One of the foremost problems in the development of traffic warrants* is the difficulty in determining the specific behavior of a general class of intersections. Computer simulation, however, offers tremendous possibilities. Digital simulation possesses unique properties such as the advantage of bringing the traffic facility into the laboratory for study under practically limitless conditions. Precise control of the dynamic traffic process can be maintained and many unwanted variables eliminated. Parameters are varied at the discretion of the programmer, rather than by chance alone." (Lewis and Michael (1963),p.1)

*A warrant is defined as a minimum requirement which should be fulfilled before some policy on traffic control is implemented.

2.6 Bottger

"The laws to which vehicular traffic conform can be investigated

(1) by mathematical computations, and

(2) by practical experiments.

Neither method is adequate. The computational method is unsatisfactory because mathematical descriptions of essential traffic processes are complicated and progress slowly. The practical method is out of the question because of the high costs and the difficulty of experimenting with real-life traffic. The answer is simulation." (Bottger (1966), p.333)

2.7 Buhr, Meserole, and Drew

"Simulation is resorted to when the system under consideration cannot be analyzed using direct or formal analytical methods. Some additional reasons, mostly pertaining to traffic simulation are:

(1) Simulation is a good way to gather data systematically. It makes for a broad education in traffic characteristics and operation.

(2) It is instructive for it gives an intuitive feel for the studied traffic system.

(3) Since the various input parameters can be controlled, the simulation of complex traffic

operations may indicate which variables are important and how they relate.

(4) In some problems, information on the probability distribution of the outcome of a process is desired, rather than only means and variances such as obtained in queueing. Where traffic interaction is involved, the Monte Carlo technique is about the only tool which can give the complete distribution.

(5) Simulation can be used to check an uncertain analytical solution.

(6) Simulation is cheaper than many forms of experiment.

(7) Simulation offers control over time. Real time can be compressed, so that the results of a long time period can be observed in a few minutes of computer time; or expanded, so that all the manifestations of a complex interaction can be comprehended.

(8) Simulation is safe. It provides a means for studying the effect of traffic control measures on existing highways without confusing or alarming drivers.

As a form of model, a simulation model should be compared with analysis which uses analytical and probabilistic models, and trial and error, which involves devising some kind of trial solution and then trying it out in actual traffic.

Table 2.2

Relative Merits

Criterion	Analysis	Simulation	Trial
Cost	least	medium	most
Time	least	medium	most
Reproducibility	most	medium	least
Realism	least	medium	most
Generality of results	most	medium	least

In the past, traffic problems have been tackled using both analysis and trial. Simulation is a combination of both, but unlike analysis, it allows attack on the most complicated of processes. On the other hand, it does not affect traffic until the solution has been reached. It is almost midway between analysis and trial. However, as the studied situation becomes more complex, such as traffic systems, the differences between methods in terms of cost, time, etc., become more pronounced, until finally neither of the extremes can be tolerated and simulation becomes the only feasible method.

The goals achievable by simulation of traffic processes are clear-cut and offer a profound payoff. Simulation is an ideal technique for traffic research. The simulation model is not just another means of accomplishing what we can do today but is a tool for solving problems which cannot be solved today." (Buhr, Meserole, Drew (1968),pp.73-74)

CHAPTER THREE

VEHICULAR LENGTHS

3.1 Introduction

In this chapter, the lengths of the three main components of the traffic stream are investigated. The distribution of vehicular lengths is assumed to correspond to the normal distribution. While it is realized that there is little empirical evidence to support this, it is felt that any simulation results would be relatively insensitive to such an assumption. By selecting an appropriate mean and standard deviation, and using Monte Carlo techniques, traffic with lengths conforming to such a distribution can be generated.

Some researchers attempt to treat arterial traffic as a set of moving points. However, this is felt to be unrealistic. Traffic is composed of vehicles - physical entities which possess length. The distinction becomes apparent as approach volumes near capacity (vehicles do not follow each other at infinitesimal headways).

The length of vehicles in the traffic stream becomes an important consideration in two later chapters - chapter 5 in which the distributions of headways is discussed and chapter 6 in which the generation of vehicles which fit these distributions is examined.

3.2 Definitions (Highway Capacity Manual (1965)):

Traffic--all types of conveyances, together with their load, either singly or as a whole...while using any roadway for the purpose of transportation or travel.

Vehicle--any component of wheeled traffic. The term normally applies to free-wheeled vehicles.

Free-wheeled vehicle--any component of traffic not limited... to rails or tracks.

The above definitions provide the necessary basis for defining the main components of the traffic stream.

Passenger car--a free-wheeled vehicle, self-propelled, generally designed for the transportation of persons. Seating capacity is limited to no more than nine passengers, including taxicabs and station wagons. Also included are two-axle, four-tired pickups, panels and light trucks, whose operating characteristics are similar to those of passenger cars.

Truck--a free-wheeled vehicle having dual tires on one or more axles, or having more than two axles, designed for cargo transportation. Includes tractor-trucks, trailers, and semi-trailers when used in combination.

Bus--a free-wheeled vehicle having a self-contained source of motive power, designed for the transportation of persons, and having a seating capacity of ten or more passengers.

Local transit buses-- buses which pick up and discharge passengers at regular posted stops along a street.

3.3 Passenger Cars

The passenger car constitutes the principal component of the traffic stream. Researchers of traffic control and simulation have used different values as the mean length of cars.

Messrs Buhr, Meserole, and Drew (1968) generated passenger cars with a constant length of 16' for their simulation of freeway traffic, while Gerlough (1967) used 16.76' as the mean car length in his simulation of traffic flow at an intersection.

Lewis and Michael (1963) in their intersection simulation and Glickstein, Findley, and Levy (1961) in their simulation of interchange volume flow, both assumed a mean value of 17'. The same is true of Bellis (1960) in his work on intersection capacity. In a graph, depicting the trend in passenger vehicle lengths presented in the Traffic Engineering Handbook (1965), 17' is the most recent average value illustrated (corresponding to 1963).

Drew (1968), in compiling traffic flow theories, suggested an average length of 17.5', a value almost identical to the 17.6' used by Athol (1965). However, Athol, in examining operational characteristics of a traffic stream, attempted to relate occupancy (a point measurement) to concentration (a measurement recorded within a section). Thus,

concentration = K occupancy

When occupancy is expressed as a percentage, and concentration in veh/mile, K = 5280/L vpm, where L is average vehicular length. Using L = 17.6',

concentration (veh/mile) = 3 occupancy (%)

ie. the value L = 17.6' seems contrived to yield an integer value of K and is, therefore, not seriously considered.

Gerlough (1964), in describing the intersection simulation of Goode, et al. reported that 18' was the average vehicular length. This value was also used by Shumate and Dirksen (1965) in their study of traffic flow behaviour. However, the latter were using SIMCAR, a language "which has a generalized capability for programming traffic simulation problems."

eg. CAR MIX = 86, LENGTH = 18, 1,...

It is suspected that an integer value is required as a parameter for LENGTH and a more accurate real number value cannot be used. Still, an average car length of 18' was used by Forbes (1963) studying driver responses, and Hewton used this value in calculating cycle lengths and phase timings.

From a consideration of the above, a mean passenger car length in the range of 17 feet to 18 feet seems appropriate. Any value selected from this range would reasonably agree with the values used by the above researchers.

Only two examples of the standard deviation of passenger car lengths could be found. Shumate and Dirksen used s = 1' while Gerlough (1967) used a value of 1.83'. Because these values are inconclusive, the following procedure was elected (Southwood (1966)).

Under the assumption of normality, the standard deviation can be determined from the relationship

$$s = \frac{x_2 - x_1}{2} \qquad (3.3.1)$$

where:

x_1 = a value corresponding to the 15.87% mark of the range,

x_2 = a value corresponding to the 84.13% mark of the range.

Thus, it is necessary first to determine the range of vehicular lengths.

If only the values of the maximum and minimum vehicular lengths are known, then

$$\bar{x} = \frac{MAX + MIN}{2} \qquad (3.3.2)$$

In the Traffic Engineering Handbook, the length of the composite longest vehicle is approximately 20'. Using MAX = 20', an average length $\bar{x}$ of 17.5', and substituting into (3.3.2) to solve for MIN yields MIN = 15'.

Thus, using the above values, the range of vehicular lengths is 5'. Determining x_1 and x_2, and applying (3.3.1) yields s = 1.7'.

It should be noted that by using a normal distribution to describe vehicular lengths, there is a finite probability of generating a vehicular length of zero. In a like manner, there is a probability of generating a value of a vehicular length which is exceedingly large. Thus, it is understood that in an actual simulation model, practical values for the upper and lower bounds of vehicular lengths (say 10' to 40') would have to be estabilished and adhered to.

3.4 Trucks

As with passenger cars, the distribution of truck lengths is assumed to be normal. The determination of a representative mean and standard deviation is again necessary in order to assign each truck an appropriate length.

Messrs Buhr, Meserole, and Drew used a constant value of 32' as the length of trucks in their simulation. Oddly enough, this is exactly twice the length of passenger cars in their model. Similarly, Shumate and Dirksen chose a truck length twice that of their passenger cars, namely 36'.

This constant ratio of lengths conforms to the basic premise of Bottger (1966) that:

1 truck = 2 passenger car units.

Gerlough (1967), however, used a mean truck length of 24.34'. This is approximately 1.5 times his mean car length, a ratio often expressed by other researchers (Greenshields (1947), Hewton (?), Webster (1966)).

Since these values for the mean length of trucks seem inconclusive, the following approach is adopted.

In section 3.3, a value of 20' was used as the maximum length of cars. Thus, we now use this same value as describing the minimum length of trucks. This is not to suggest that the minimum length of trucks which could be generated according to a normal distribution is restricted to 20'; but rather, that this is a realistic value to use in order to obtain reasonable values of the mean and standard deviation of truck lengths. Also, the average maximum legal length of trucks in the U.S. is approximately 50' (Traffic Engineering Handbook (1965)).

Using MIN = 20', MAX = 50', and applying (3.3.2), $\bar{x}$ = 35'.

This value generally agrees with the values of the above researchers and, therefore, seems reasonable for simulation purposes. Also, when compared with the previously determined range for mean passenger car lengths, the ratio is approximately 2.0. This, too, is in close agreement with the values discussed earlier.

In determining the standard deviation of truck lengths, the shortage of information is again apparent. Shumate and Dirksen used a standard deviation of 6' while Gerlough (1967) used a value of 3.4'. Neither supply a reason for their choice, and neither value seems any more appropriate than the other. Thus, referring again to (3.3.1) with MAX = 50', MIN = 20' and a range of 30',

$$s = 10.2'$$

This value does not appear to be unreasonable and is, therefore, suitable for our purposes. As in the case of passenger cars, using the normal distribution there is a probability of generating a truck length either exceedingly large or exceedingly small. Again, practical upper and lower bounds would have to be enforced in an actual simulation model.

3.5 Buses

Though local transit buses form an integral component of the traffic stream, one feels it is not necessary to draw such vehicular lengths from a prescribed distribution. Indeed, the limited information concerning average bus lengths makes the task prohibitive. However, some useful information is available.

The Traffic Engineering Handbook recorded a nominal bus length of 40' corresponding to an approximate seating capacity of 51 passengers. In addition, the most common legal maximum length of buses in the U.S. is 40'.

The only other pertinent information is due to Webster (1966). In determining the effect of different types of vehicles on the saturation flow at traffic signals, he equates a bus to 2 1/4 passenger car units. Using the previously determined range of mean passenger car lengths of 17' to 18', the average bus length is approximately 40'.

On the basis of the available information, a constant local transit bus length of 40' is selected as adequately representing the true length of such vehicles.

CHAPTER FOUR

LANE DISTRIBUTION

4.1 Introduction

The distribution of traffic between two lanes for one direction of flow depends on several factors, including medial and marginal friction*, proportion of slow moving vehicles, the number and location of ingress and egress points (important when the origin and destination of drivers is considered), and the volume variations of traffic (Highway Capacity Manual (1965)).

The Traffic Engineering Handbook (1965) includes such factors as curbs, grades, road surface conditions, turn and lane-use regulations, lane markings, and the composition of traffic.

Some, or all, of the above factors combine to influence traffic distribution between lanes. However, the exact way in which these factors combine is undefined.

In this chapter, the distribution of traffic between two lanes is examined. Primary emphasis is placed on the relationship between lane distribution and traffic volume. Available data are analyzed and a linear regression equation, determined by Gerlough (1967), is applied in an

*Definitions of these terms appear in section 7.2.

attempt to add some degree of universality to this equation.

While it may not be sufficient to relate lane distribution to traffic volume alone, there appears to be virtually no quantitative information available relating lane distribution to any of the other factors previously mentioned. In fact, as both Gerlough (1967) and Lewis and Michael (1963) point out, the literature concerning lane selection is surprisingly deficient.

4.2 Discussion

Reilly and Seifert (1970) analyzed three methods of estimating the capacity of signalized intersections. They assumed the lane distribution for two lanes to be 55% and 45%. The maximally used lane was not necessarily the left or right lane. Unfortunately, Reilly and Seifert never explained why they chose these values.

Gerlough (1967) found similar values in his field study but was more specific as to lane usage. His work revealed an average of 58.5% of the total volume use the right lane, while 41.5% use the left.

Two years later, Wagner (1969) performed field studies near the central business district of Los Angeles between 4:30 p.m. and 5:30 p.m. Though traffic flowed east and west in three lanes rather than two, his findings are interesting and warrant analysis.

The null hypothesis tested is that the mean percentage of the total volume in the right lane is equal to the mean percentage in the left lane.* The test criterion is distributed as a t. On the basis of the available data, the null hypothesis is rejected.

Again, using t as a test criterion, both the mean of the right lane and the mean of the left lane are found to be not unequal to the average percentages determined by Gerlough, namely 58.5% and 41.5% resp.

Using the method of least squares, two regression lines are calculated;

Percent using right lane = -.03 x total volume + 79.96 (4.2.1)

Percent using left lane = .02 x total volume + 23.52 (4.2.2)

for a range of approach volumes from 51 vph. to 1780 vph.

Accordingly, the left lane containes a greater percentage of the traffic volume for volumes in excess of 1128 vph. This does not necessarily contradict the previous assumption that fewer vehicles travel in the left lane than the right. The greatest simulated volume in one direction used

* Though the arterial system contained three lanes in each direction, the percentage of traffic using the leftmost lane was very small. Thus, traffic in the middle lane behaved similar to traffic in the left lane of the two-lane artery. It is this lane which is referred to as the left lane.

by Gerlough was 1017 vph., while the greatest observed volume by Wagner (1969) was 1780 vph. Both were drawn from peak hour time periods. Traffic flow during these periods is usually associated with conditions of congestion.

For regression line 4.2.1, the correlation coefficient is -.66, while for regression line 4.2.2, it is .46. Testing each of these values under the hypothesis that $\rho = 0$, the null hypothesis is rejected on the basis of the observed data.

In spite of these results, the correlation coefficients were felt to be too low and it was decided not to use these regression lines. Gerlough determined his own regression equation permitting the estimation of lane distribution percentages given total approach volume.

Percent using left lane = -.015 x total volume +49.5 (4.2.3)*

Equation 4.2.3 is applied to the volume rates observed by Wagner. The mean of the percentages so generated is tested under the null hypothesis $H_0: \mu = 41.5\%$. Using t as the test criterion, the null hypothesis is accepted on the basis of the available evidence.

*The data Gerlough used to determine (4.2.3) are not available. Thus, the range of traffic volume for which this regression line is valid is not known.

It is noted that since only a two-lane approach is involved, the regression line for the percent of the total approach volume using the right lane is the complement of (4.2.3). Accordingly, the percentage using the right lane is always greater than the percentage using the left.

To determine the volume of traffic using the left lane, equation (4.2.3) is multiplied by the total approach volume.

$$\text{Left lane volume} = \text{total volume} \times (-.00015 \times \text{total volume} + .495) \qquad (4.2.4)$$

Thus, the maximum possible volume in the left lane is 408.375 vph. corresponding to a total approach volume of 1650 vph.

This appears to be a realistic maximum. Of the field data recorded by Wagner (1969), only two cases occur in which the volume in the left lane greatly exceeds this maximum.

For simulation purposes, it was decided that regression equation 4.2.3 and its complement could be used to determine the percentage of two-lane approach traffic using the left and right lanes resp.

Gerlough found that trucks do not distribute themselves between lanes according to the same percentages as traffic as a whole. Truck distribution was 68.3% in the right lane and 31.7% in the left. In the absence of any information to the contrary, these values are accepted as being suitable for simulation purposes.

4.3 Conclusion

It is realized that the applicability of equation 4.2.3 is questionable. The data for which Gerlough determined the equation is not available. Furthermore, only a small amount of data was used in an attempt to substantiate (4.2.3). Ideally, of course, it is desirable to have available a quantity of data collected at different locations during different time periods in order to substantiate a particular relationship.

Realistically, then, data which are available are used with the hope that some measure of universality can be attached to this quantitative relationship. While equation 4.2.3 can hardly be construed as a general rule characterizing traffic flow, it is a workable form with a degree of reliability sufficient for purposes of simulation. The alternative is to assume, as did Drew (1968), that vehicles distribute themselves equally between lanes. Such an assumption is felt to be both unsubstantiated and unrealistic.

One further point deserves discussion. As in any linear regression analysis, the parameters characterizing a linear relation have been determined for two variables which may or may not be linearly related. In the absence of any relevant studies, it is difficult to determine if the apparent linearity of relationship 4.2.3 has application beyond the data from which it was derived. The writer has little to add to this except to note that the range of approach volumes to which (4.2.3) is applied (see section 5.4) is quite narrow. Hopefully this range of values does not exceed the range Gerlough used to derive this form.

CHAPTER FIVE

Headway Distribution

5.1 Introduction

Analyses of roadway characteristics have required the investigation of spacing and headway characteristics. Headways have application in predicting arrival rates, testing the randomness of traffic flow, estimating gaps and delays at crossings, developing traffic control warrants, and timing traffic signals.

Spacing and headway give the driver his sense of freedom of movement and of relative safety, and continuously affect his choice of vehicular position and speed. His decisions in weaving and car-following operations are predicated on his judgment of suitable vehicular gaps. The frequency and length of gaps govern his ability to enter or cross the traffic stream (Highway Capacity Manual (1965)).

As a consequence of the importance of headways, considerable effort has been expended developing a sound, general purpose headway model. In this chapter, we examine the distributions for describing vehicular headways and the application of these distributions to simulation models.

Starting with the Poisson distribution in describing vehicular arrivals, section 5.3 examines the problems associated with the negative exponential distribution. In attempting to solve these problems, different modifications have been proposed by various researchers. In all, this has lead to the development of the composite negative exponential distribution.

In section 5.4, a comprehensive analysis of the composite negative exponential distribution is undertaken. The inter-relationships of the parameters of this distribution, as first presented by Kell (1962), are explored and a method presented, whereby these parameters may be determined on the basis of the approach volume.

Finally, in section 5.5, a suitable distribution for describing the headway between local transit vehicles is presented. In order to substantiate the appropriateness of this model, tests using actual field data were conducted. The results of these tests are described at the end of this section.

There are, of course, other theoretical distributions developed by various authors. Lewis and Michael (1963) proposed a modified binomial distribution with different levels of probability.* Thus,

$$Pa = 1-(1-B)^{\frac{1}{r-e-1}}$$

and

$$Pb = \frac{1-B}{\bar{h}-e-\sum_{h=1}^{re}(1-Pa)^{h}}$$

*Lewis proposed $1\frac{1}{2}$ sec for e, $4\frac{1}{2}$ sec for r, and $1-e^{-.00132h}$ for B. A complete discussion of this distribution can be found in Lewis (1963).

where:

Pa = enhanced probability of an arrival at a time increment when $e \leqslant h \leqslant r$;

Pb = diminished probability of an arrival at a time increment when $h > r$;

B = bunching factor; fraction of all headways $\leqslant r$;

r = maximum headway for which the probability of an arrival is enhanced;

e = minimum headway permitted;

$\bar{h}$ = mean of all headways, h.

Cleveland and Capelle (1964) report the development of the Pearson Type III, or Erlang, distribution; while Dawson and Chimini (1960) researched a headway distribution which they call the hyper-Erlang (or hyperlang) function. The latter is very similar to the composite negative exponential model of Kell (1962); however, it involves a modification of the retarded vehicular headways. This modification uses a factor indicating the degree of nonrandomness of the retarded headway distribution.

A summary of applications of mathematical distributions to time headway distributions, and their associated authors, is provided by May (1965).

5.2 Definitions (Highway Capacity Manual (1965)):

Headway--the interval in time between individual vehicles measured from head to head as they pass a given point.

Spacing--the distance measured from head to head of successive vehicles.

5.3 Background

The Poisson distribution is the main theoretical instrument for determining the distribution of vehicular traffic. Cleveland and Capelle (1964) explained that the assumption behind this distribution is that the total number of arrivals, during any given time interval, is independant of the number of arrivals that have occurred prior to the beginning of the interval.

Messrs Wohl and Martin (1967) described the Poisson distribution thus:

For Poisson-distributed arrivals, p(n), the probability of exactly n vehicles arriving at any t-sec. interval is

$$p(n) = \frac{\mu^n e^{-\mu}}{n!} \qquad n = 0,1,2,\ldots$$

where

$q = V/T$ = mean rate of arrivals per unit time;

$\mu = qt$ = average number of vehicle arrivals in an interval of length t, (and μ is a function of time);

V = total volume of vehicles arriving during time interval T.

The probability of no vehicles arriving in the interval t (that is, n=o) is

$$p(0) = e^{-\mu} \text{ or } e^{-qt}, \text{ for } t \geq 0.$$

If no vehicles arrive in the time interval of length t, then there must have been a gap or time headway of at least t sec. Thus, the probability of a headway h being equal to or greater than t is

$$P(h \geq t) = e^{-qt}, \text{ for } t \geq 0. \qquad (5.3.1)$$

The probability density function of (5.3.1) is the negative exponential distribution. Drew (1968) stated the relationship as follows: The distribution of time spacings between Poisson arrivals conforms to a negative exponential distribution.

The first published examples of the Poisson distribution as applied to traffic data, were presented by Adams in the Journal of the Institution of Civil Engineering (as related by Gerlough (1955)). Gerlough, himself, presented examples of fitting the Poisson distribution to field data of vehicular arrivals. However, both he and Adams restricted themselves to approaches of low hourly volume.*

On the assumption that arrivals during light traffic flow are randomly distributed over time, the Poisson distribution

*Adams draws his example from an approach with an hourly volume of 222 vph., while Gerlough cites examples involving hourly volumes of 37 vph. to 418 vph.

can be used to describe vehicle arrivals, and the negative exponential distribution to describe headways. However, Wohl and Martin presented an example of vehicle arrivals during a peak hour. For a volume of 997 vph., the high X^2 value suggested it is unlikely that the hypothesis of exponentially distributed headways is true.

Basically two factors seem to limit the direct application of the negative exponential distribution.

First, the theoretical curve distributes headways continuously over the entire range of time, namely zero to infinity. Dawson and Chimini (1960) point out that there is a minimum headway related; to the length of the lead vehicle, to the minimum intervehicular spacing demanded by the trailing vehicle, and to the speed and acceleration of the trailing vehicle. Inasmuch as each gap must include at least one vehicle, no headways exist at values less than .5 sec. (Highway Capacity Manual (1965)). At 30 mph., this represents a spacing of 22'.

Second, as volumes increase, more and more vehicles adopt short headways as they overtake, but cannot pass, slower moving vehicles. Under this "bunching" condition, a car-following phenomenon develops and vehicular behaviour becomes nonrandom. According to Dawson and Chimini, research into this car-following phenomenon essentially has proven intervehicular dependence.

A corrective measure to the first difficulty is the shifting of the exponential curve to the right by an amount equal to a certain minimum headway τ.

Recalling equation 5.3.1, and setting the mean of the negative exponential distribution $\bar{t} = 1/q$, (5.3.1) may be rewritten

$$P\ (h \geq t) = e^{-t/\bar{t}},\ t \geq 0. \qquad (5.3.2)$$

Introducing the minimum headway τ,

$$P\ (h \geq t) = e^{-\frac{t-\tau}{\bar{t}-\tau}},\ t \geq \tau, \qquad (5.3.3)$$

and the probability of a gap between successive vehicles of less than τ is zero.

However, this modified model does not compensate at all for intervehicular dependence.

To overcome these shortcomings, Schuhl (1955) proposed his own model. He supposed that the entire set of spacings between successive vehicles is a set of random and independent elements composed of two subsets. The first subset of spacings apply to retarded vehicles which are under the influence of other vehicles in the stream. As vehicles have a finite length, two successive vehicles in this set are separated by a time interval with a positive lower bound.

The second subset of spacings apply to free-moving vehicles not under the influence of other vehicles. This condition exists under a variety of circumstances such as exists; when the headway from the free vehicles to preceeding vehicles is of "adequate" duration, when the free vehicle can pass without modifying its time-space trajectory as it approaches preceeding vehicles, or when a passing vehicle has sustained a positive relative speed after passing, so that the free vehicle can still operate as an independent unit.

Thus,

$$P\ (h \geqslant t) = \alpha e^{-\frac{t-\varepsilon}{t_1-\varepsilon}} + (1-\alpha)\ e^{-t/t_2} \qquad (5.3.4)$$

where

$P\ (h \geqslant t)$ = probability of a headway (h) $\geqslant$ to the time (t);

α = proportion of traffic stream in retarded group;

$(1-\alpha)$ = proportion of traffic stream in free-moving group;

t_1 = average headway of retarded group;

t_2 = average headway of free-moving group;

ε = positive lower bound of time spacing.

In this model - which has been variously called a modified exponential or a composite distribution - the traffic stream is not considered a single-lane stream. Thus, the lower limit of the free headway distribution is zero because headways are being measured between the free vehicles occupying the adjacent lane during a passing maneuver.

Kell (1962) modified this model so that both the free and retarded headways are represented by translated functions.

If the traffic stream is a single-lane stream, and headways are measured just between successive vehicles in that stream, there must be real minimums for both the free and retarded headway distributions.

Thus,

$$P\,(h \geqslant t) = \alpha e^{-\frac{t-\tau}{T_2-\tau}} + (1-\alpha)\, e^{-\frac{t-\lambda}{T_1-\lambda}} \qquad (5.3.5)$$

where

$P\,(h \geqslant t)$ = probability of a headway h greater than or equal to the time t;

α = proportion of the traffic stream in retarded group;

$(1-\alpha)$ = proportion of the traffic stream in free-moving group;

T_1 = average headway of free-moving vehicles;

T_2 = average headway of retarded vehicles;

λ = minimum headway of free-moving vehicles;

τ = minimum headway of retarded vehicles;

e = 2.71828.

This has been referred to as the composite negative exponential distribution.

5.4 The Composite Negative Exponential Distribution

In selecting a headway distribution which was suitable for simulation, it was necessary that the distribution be conceptually sound and had been tested over an extensive volume range. Also, the parameters of the distribution had to be solved in terms of volume.

The composite negative exponential model proposed by Kell satisfied these requirements.

Recalling the form of the composite distribution (5.3.5),

$$P\ (h \geq t) = \alpha e^{-\frac{t-\tau}{T_2-\tau}} + (1-\alpha)\ e^{-\frac{t-\lambda}{T_1-\lambda}} .$$

The five parameters $-\alpha$, T_1, T_2, λ, τ- are functions of traffic volume. Kell reduced (5.3.5) to four unknowns using the transformations

$$a = \frac{\lambda}{T_1-\lambda} + \ln\ (1-\alpha),$$

$$K_1 = T_1-\lambda,$$

$$C = \frac{\tau}{T_2-\tau} + \ln\alpha,$$

$$K_2 = T_2-\tau.$$

Thus

$$P\ (h \geq t) = e^{a-t/K_1} + e^{C-t/K_2} \qquad (5.4.1)$$

Using extensive field data, equations for the unknowns in (5.4.1) were empirically determined. Thus

$$K_1 = \frac{4827.9}{V^{1.024}},$$

$$a = -.046 - .000448V,$$

$$K_2 = 2.659 - .0012V,$$

$$C = \left(e^{-10.503 + 2.829\ \ln V - .173\ (\ln V)^2}\right) - 2,$$

*This model was successfully used by Kell (1963) in analyzing intersection delay, and by Gerlough (1967) in analyzing the performance of traffic at an intersection.

where V = approach volume in vph.

Whereas, the parameters of (5.4.1) have been solved in terms of volume, a simulation model requires (5.3.5). Because α, λ, and τ are inter-dependent, the asignment of a value to one determines the other two, and it is impossible to transfer equation 5.4.1 back directly to (5.3.5).

The equations of these parameters are

$$\lambda = K_1 \left(a - \ln (1-\alpha)\right),$$

$$\tau = K_2 \left(C - \ln\alpha\right),$$

$$1-\alpha = e^{\,a-\lambda/K_1},$$

$$\alpha = e^{\,C-\tau/K_2}.$$

Kell found best agreement between the theoretical curve and the observed data where $.9 < \lambda < 1.0$ and $1.2 < \tau < 1.36$. However, attempts to determine "appropriate" values of V and α which would produce values of λ and τ in these ranges proved unsuccessful.* It was decided that if α could be stated in terms of approach volume, much of this difficulty would be eliminated.

*Since the traffic stream is a single-lane stream, and headways are measured between successive vehicles in that stream, the minimum headway of retarded vehicles τ must be less than the minimum headway of free-moving vehicles λ. Any vehicle whose headway was less than λ was no longer a free-moving vehicle, but a retarded one. Thus, the ranges proposed by Kell for λ and τ are inappropriate.

Dawson and Chimini reported that the balance between free and retarded vehicles varies with the flow rate of the traffic stream. That is as flow rate increases, the proportion of free vehicles decreases and the proportion of retarded vehicles increases. To reflect this, they introduced linear coefficients α_1 and $\alpha_2 = (1-\alpha_1)$ before each of the components of the hyperlang function.

It was thus assumed:

1) that there was a linear relation between α and the approach volume,

2) that there was a maximum volume at which all vehicles were retarded, ie. $\alpha=1$,

3) that for the minimum approach volume, namely 0, no vehicles were retarded, ie. $\alpha=0$, and the equation of this linear relation passed through the origin.*

Thus

$$\alpha = V/V_{max} \qquad (5.4.2)$$

where

α = proportion of traffic stream retarded;

V = single-lane approach volume (vph.);

Vmax = maximum single-lane approach volume (vph.).

*These assumptions were first proposed by Adolf May (1965) and Patrick Athol (1965) in their work concerning the headway characteristics of free-moving and retarded vehicles.

To further substantiate the linearity of the relationship between these two quantities, linear regressions were applied to two sets of data used by Dawson and Chimini. For data from the Highway Capacity Manual, the correlation coefficient between the ratio of retarded vehicles and the approach volume was .98, while for data from a study by Purdue University, the correlation coefficient was .96. Thus, it remained to determine an appropriate value for Vmax in order to apply equation 5.4.2.

Greenshields (1947) measured the headway between successive vehicles, queued at a stoplight, as they entered the intersection after the light turned green. As the speeds between successive vehicles became constant and equal to the vehicles' desired speed, the headways between successive vehicles tended towards a constant of 2.1 sec.*

An established relationship, under steady-state conditions, is that

$$\text{Vmax (Veh/hr)} = \frac{\text{3600 (sec/hr)}}{\text{Minimum headway (sec/Veh)}} ,$$

and $\text{Vmax} = 3600/2.1 = 1710$ vph.

However, Hewton (?) points out that a volume of 1710 vph., per lane for an entire hour will only be possible under almost ideal conditions. It is more probable that, due to

*Greenshields's work is widly accepted and was later substantiated, with minor modifications, by Messrs Drew and Pinnell (1962).

marginal friction, uneven speed, etc., volume flow could not reach this figure; or if it did, serious congestion would result. A practical maximum single-lane approach volume was needed.

Since Kell, in solving for the parameters of his composite distribution, analyzed volume rates ranging to almost 1200 vph., it seemed appropriate to assign to this value the maximum volume of a single-lane approach. In fact, there is a strong body of evidence to support this value as the maximum.

Drew and Pinnell (1962) studied the peaking characteristics of traffic at signalized intersections. On 13 intersections of 2-way streets with a speed limit of 30 mph., the highest observed peak hour volume for a 2-lane approach was only 1267 vph., while the average peak hour volume for two lanes was 818 vph. In fact, the highest recorded 5-minute arrival rate was 129 vehicles. This represents an hourly 2-lane approach volume of 1548 vehicles.

Applying equation 4.2.4*

$$\text{Left Lane Volume} = .495 \text{ Vol} - .00015 \text{ Vol}^2.$$

where

Vol = 2-lane approach volume (vph.);

then

left lane volume = 407 vph.,

right lane volume = 1141 vph.

*Recall that using (4.2.4) the maximum possible volume in the left lane is 408 vph., corresponding to a 2-lane approach volume of 1650 vph.

O. K. Normann (1962) discusses the collection of data on 1600 intersections, during periods of peak traffic flow, by the Bureau of Public Roads in preparing to update the original Highway Capacity Manual. Of the volumes recorded at intersections of 2-way streets, with no parking, the highest peak hour volume for a 2-lane approach was only 1427 vph. The average 2-lane approach volume was 1034 vph.

As a final reference, the traffic volumes recorded by Gerlough and Wagner (1967), in the preparation of their simulation model, were examined. Data were collected from six signalized intersections of 4-lane streets in California. The highest hourly volume rate on a 2-lane approach was 1117 vph. Applying equation 4.2.4

left lane volume = 366 vph,

right lane volume = 751 vph.

In light of the above empirical evidence, the assumption of a practical maximum single-lane volume flow of 1200 vph. was felt to be quite realistic. Thus

$$\alpha = \frac{1}{1200} V \qquad (5.4.3)$$

In order to use the composite negative exponential distribution to generate headways, it was necessary to determine a range of single-lane approach volumes which satisfied the following conditions:

1) The volumes, while not exceeding 1200 vph., should have a range sufficient so as to reproduce representative conditions at a typical intersection. To simulate vehicular flow in the left lane, the lower limit of this range must be substantially less than 408 vph.

2) Because negative headways are impossible, the volumes must generate minimum headways greater than zero. In fact, the minimum headway of retarded vehicles τ must be greater than .5 sec.

3) The volumes must generate vehicular headways such that the minimum headway of free-moving vehicles is greater than the minimum headway of retarded vehicles, ie. $\lambda > \tau$.

4) Patrick Athol (1965) investigated the minimum headway of free-moving vehicles, λ. Examining values ranging from 1.5 sec. to 3.9 sec., he noted that values of 2.0 sec. or 2.5 sec. appeared to be the best choice. To accomodate as wide a range of volumes as possible, it was decided to set the maximum value of λ at 3.5 sec.

5) Logically, as approach volume increased, the mean headway of free-moving vehicles T_1 should decrease. Indeed, the data used by Dawson and Chimini bore this out. However, preliminary testing showed that the equation for T_1, was not strictly monotonic decreasing over the entire range of approach volumes. After some critical approach volume, T_1 began to increase with

increasing volume rates. Thus, a final condition was that

$$T_{1,i} < T_{1,i+1} \quad \text{for} \quad V_i < V_{i+1}.$$

A program was written which calculated the parameters of the composite distribution based on volumes ranging from 100 vph. to 1200 vph., in increments of 10 vph. Using the equations developed by Kell, as well as (5.4.3), only those volumes which satisfied the above conditions were accepted as being suitable for purposes of simulation. Table 5.1 shows the single lane and 2-lane approach volumes finally selected, along with the associated parameters of the composite negative exponential distribution. The general trend and ranges of the values α, T_1, T_2, λ, and τ were felt to be in good agreement with the data presented by Dawson and Chimini.

Though all these single-lane approach volumes satisfied the previous restrictions, some of them represented only left-lane volumes while others represented only volumes in the right lane. For example, all volumes greater than 408 vph. represented right-lane volumes since, according to (4.2.4), this was the maximum vehicular flow in the left lane. A program was written which determined the corresponding 2-lane approach flow represented by these single-lane volume rates. The results are presented in table 5.2.

Table 5.1

Single-lane Approach Volumes with Corresponding Composite Distribution Parameters

single lane Vol.	proportion retarded α	avg. headway retarded T_2	min. headway retarded τ	$K_2 = T_2 - \tau$	proportion free-$(1-\alpha)$	avg. headway free T_1	min. headway free λ	$K_1 = T_1 - \lambda$
217	.181	3.497	1.099	2.398	.819	20.65	1.10	19.55
317	.264	3.15	.87	2.28	.736	14.84	1.57	13.27
417	.348	2.94	.79	2.15	.652	11.96	1.94	10.02
517	.431	2.78	.74	2.04	.569	10.337	2.299	8.038
617	.514	2.62	.70	1.92	.486	9.39	2.68	6.71
717	.598	2.47	.67	1.80	.402	8.87	3.12	5.75
788	.657	2.35	.64	1.71	.343	8.719	3.498	5.221

Table 5.2

Approach Traffic Split into Left-lane and Right-lane Volumes.

2-lane approach volume-vph.	left-lane volume-vph.	right-lane volume-vph.
520.5	217	303.5
541	224	317
685.97	268.97	417
822.7	305.7	517
952.4	335.4	617
1075.9	358.9	717
1160.4	372.4	788

5.5 Headway Distribution - Local Transit Vehicles

Having decided on the composite negative exponential distribution to describe the headway of cars and trucks in the traffic stream, it was then necessary to determine an appropriate distribution for the headway between local buses in the same stream. Whereas the negative exponential had originally proved an unsuitable model, it was now reasoned, that because the volume of bus traffic itself was very low, arrivals could be considered randomly distributed and Poisson in nature.

A program was written which tabulated and analyzed bus arrivals at 9 bus stops along a 9-block, 4-lane artery. The street, Graham, is .6 miles long and accommodates ten bus routes, east-bound and west-bound, having stops at every intersection. Data for this program was taken from Metro transit bus schedules.*

Two types of analyses were performed. The first, individually examined each bus stop along Graham. Of the 21 hours of bus service, the peak 60-minute arrival rate was determined. The distribution of arrivals throughout this peak hour was then tested against the theoretical Poisson distribution by means of the chi-square test of goodness of fit.** For each of the 9 bus stops, there was no evidence from which to conclude that the distribution of arrivals differed from the theoretical Poisson distribution. Table 5.3 is a description of the bus stops analyzed, along with the associated bus movement.

*This is not an uncommon method of obtaining information about local transit behaviour. Reilly and Seifert (1970) used bus company schedules in order to examine the Highway Capacity Manual's nomographs which describe the effect of bus stops at signalized intersections.

**Originally, the analysis was performed using the Kolmogorov-Smirnov test for goodness of fit (Ostle (1963)). Massey (1951) suggests that this test may always be more powerful than the chi-square test, since it usually requires less calculations and is able to detect smaller deviations from the theoretical distributions. However, modern consensus among statisticians is that the Kolmogrov-Smirnov test should not be used when parameters of the distribution need to be estimated.

Table 5.3

Local Transit Movement on Graham Avenue.

Location: Cross street of Graham	Direction of Traffic Flow	Type of Stop	Number of Transit Routes Serviced	Local Bus Movement		
				Total No. of Buses/Day	Peak Hour Bus Volume	% Peak Hour of Total
Main St.	west-bound	far-side	2	193	20	10.4
Gary St.	west-bound	near-side	2	193	20	10.4
Donald St.	west-bound	near-side	5	485	45	9.3
Carleton St.	west-bound	near-side	5	342	32	9.4
Kennedy St.	west-bound	near-side	5	463	44	9.5
Edmonton St.	east-bound	near-side	3	375	34	9.1
Hargrave St.	east-bound	near-side	4	461	43	9.3
Smith St.	east-bound	far-side	3	375	35	9.3
Fort St.	east-bound	far-side	2	224	23	10.3

The second analysis was concerned with the distribution of bus arrivals on Graham Avenue. Although Graham Avenue is an artery with multiple ingress and egress points, it was now being treated as a single holding area. Therefore, each bus that entered from any of the cross streets during a unit interval, namely one minute, was tabulated as an arrival.

The null hypothesis first tested was that the distribution of bus arrivals during the peak 60-minute period approximated the theoretical Poisson distribution. Using the chi-square test, the null hypothesis was accepted on the basis of the available data.*

Second, the distribution of bus arrivals throughout the entire day was analyzed. The arrivals were tabulated for time intervals of from one to five minutes. For each of these cases, the null hypothesis tested was that the observed distribution of arrivals approximated the theoretical Poisson distribution. In all cases, the null hypothesis was accepted on the basis of the data.

Based on the above results, the negative exponential distribution is felt to be a suitable model for describing the

*The maximum number of bus arrivals for a 60-minute period was 85. This represented 9% of the 946 buses which entered Graham Avenue throughout the day.

distribution of time spacings. However, because buses have a finite length, it is impossible to have two arrivals in the same lane at the same time. Due to this, the shifted form of the negative exponential disbribution is felt to be a better theoretical model.

Since buses are approximately 40' in length, at 30 mph., the minimum headway between successive arrivals -τ- is .909 sec. Recalling the form of the shifted negative exponential distribution (5.3.3),

$$P\ (h \geqslant t) = e^{-\frac{t-\tau}{\bar{t}-\tau}}\ ,\ t \geqslant \tau.$$

Thus,

$$P\ (h \geqslant t) = e^{-\frac{t-.909}{\bar{t}-.909}}\ ,\ t \geqslant .909,$$

and the probability of a gap between successive buses of less than .909 sec. is zero.

Finally, it was necessary to determine a range of local bus volumes suitable for purposes of simulation. At most intersections studied by O. K. Normann (1962), local bus volume was in the neighbourhood of 2% of the total peak hour traffic. Rainville et al. (1961), in their own study, examined facilities for which local transit accounted for between 1.2% and 8.5% of peak hour approach volume.

Because the volume of peak hour traffic on Graham Ave. was not available, no real comparisons between bus percentages on this artery and those observed by Normann and Rainville could be made. However, using the range of 2-lane approach volumes in table 5.2 as representative traffic flow, the peak hour arrivals at bus stops along Graham Ave. varied from 1.7% to 8.7% of these values.

It was decided, that local transit traffic ranging from 1.5% to 8.5% of a 2-lane approach volume would be appropriate for purposes of simulation. At the same time, all transit activity would be restricted to operating in the right-hand curb lane.

Chapter Six

Vehicle Generation

6.1 Introduction

The basic problem in simulation consists of sampling from statistical distributions for which any associated probability must, by definition, be a fraction of unity. There are at least four well known standard techniques for generating random observations (variates) from continuous probability distributions. These are 1) the inverse probability integral transformation method, 2) the rejection technique, 3) the composition methods, 4) the method of transforming other random variables. As a discussion of these different methods is beyond the scope of this thesis, the reader is referred to chapter 3 of Knuth (1969).

In this chapter, methods for generating the two distributions presented in chapter 5, namely the composite negative exponential distribution and the shifted negative exponential distribution, are discussed. Functional relationships developed in chapter 5 are used, as well as values of vehicular lengths determined in chapter 3.

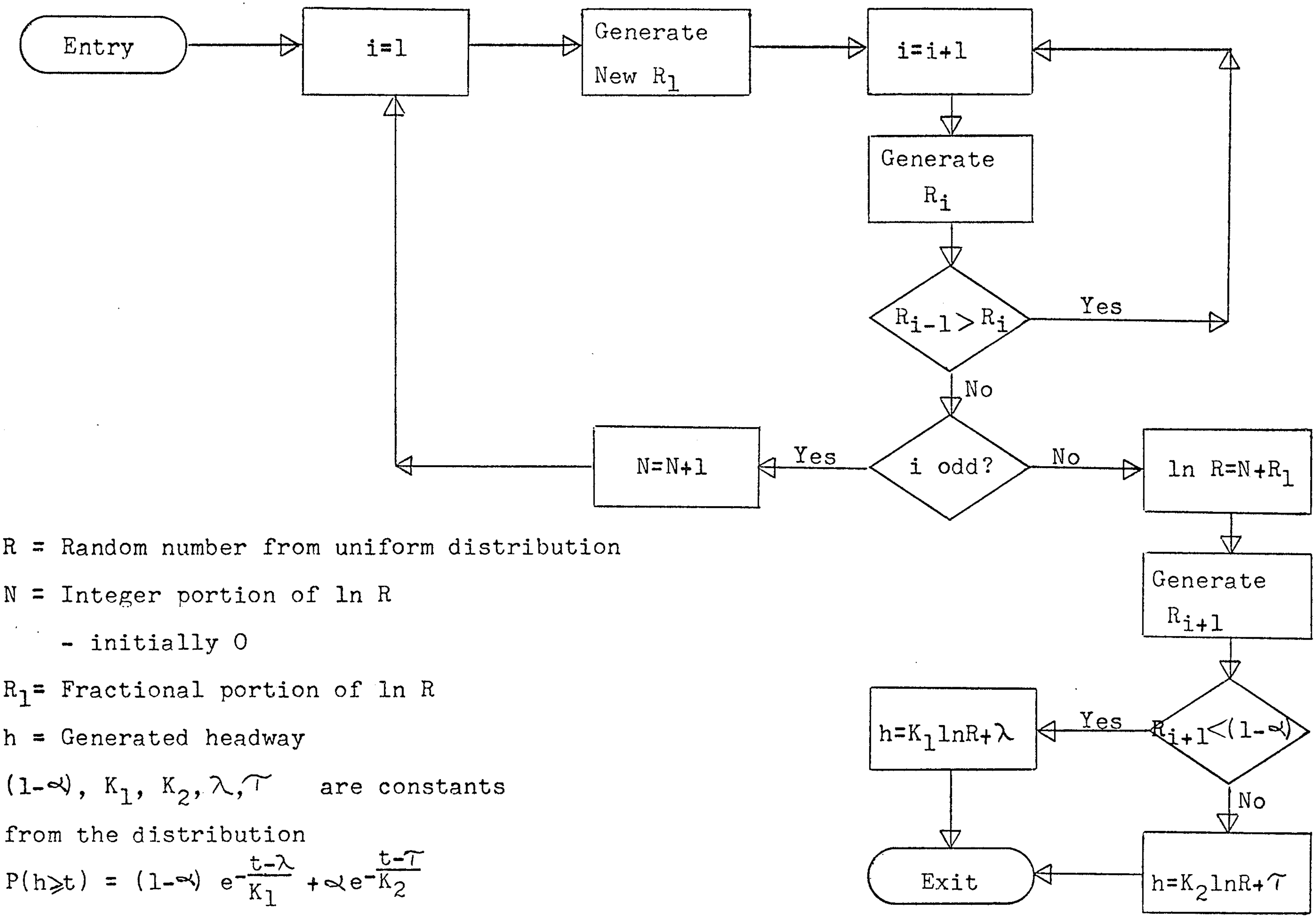
In section 6.4, a method is outlined whereby headways generated according to these two distributions may be successfully merged to form a single traffic stream. A program developed to test this procedure is also discussed. Because this method is suitable only for traffic streams of very low volume levels, and even then only by chance, a procedure is required which guarantees that a traffic stream composed of these two distributions is generated for any appropriate volume level. The development of such an algorithm is presented at the end of this section.

6.2 The Composite Negative Exponential Distribution

Having decided upon the composite negative exponential distribution to describe the headway between cars and trucks in the same lane of a traffic stream, the next step is to generate headways that fit this distribution randomly. Kell (1962) provides a flow diagram of the generator he used in his own simulation model (see fig. 6.1).

Kell is sampling from what is known as a mixture distribution: ie. with probability α a sample is drawn from the shifted exponential distribution with mean $K_2+\tau$ $(x \geqslant \tau > 0)$, and with probability $1-\alpha$, a sample is drawn from the shifted exponential distribution with mean $K_1+\lambda$. The shifted exponential variate is generated by a location and scale transformation of a standard exponential variate (say e) with mean 1. Of the several available methods for generating standard exponential variates (see Knuth (1969)), Kell uses a relatively sophisticated rejection technique. A similar rejection technique is described by Knuth.

Random headway generator (from Kell (1962))



R = Random number from uniform distribution

N = Integer portion of ln R

- initially 0

R_1= Fractional portion of ln R

h = Generated headway

$(1-\alpha)$, K_1, K_2, λ, τ are constants from the distribution

$$P(h\geqslant t) = (1-\alpha)\, e^{-\frac{t-\lambda}{K_1}} + \alpha e^{-\frac{t-\tau}{K_2}}$$

The distribution sampled is selected with the correct probability by first generating a random number R and then branching to the appropriate transformation depending upon the relationship between R and $(1-\alpha)$. Thus, Kell's algorithm may be stated as:

fig. 6.2

Simplified random headway generator

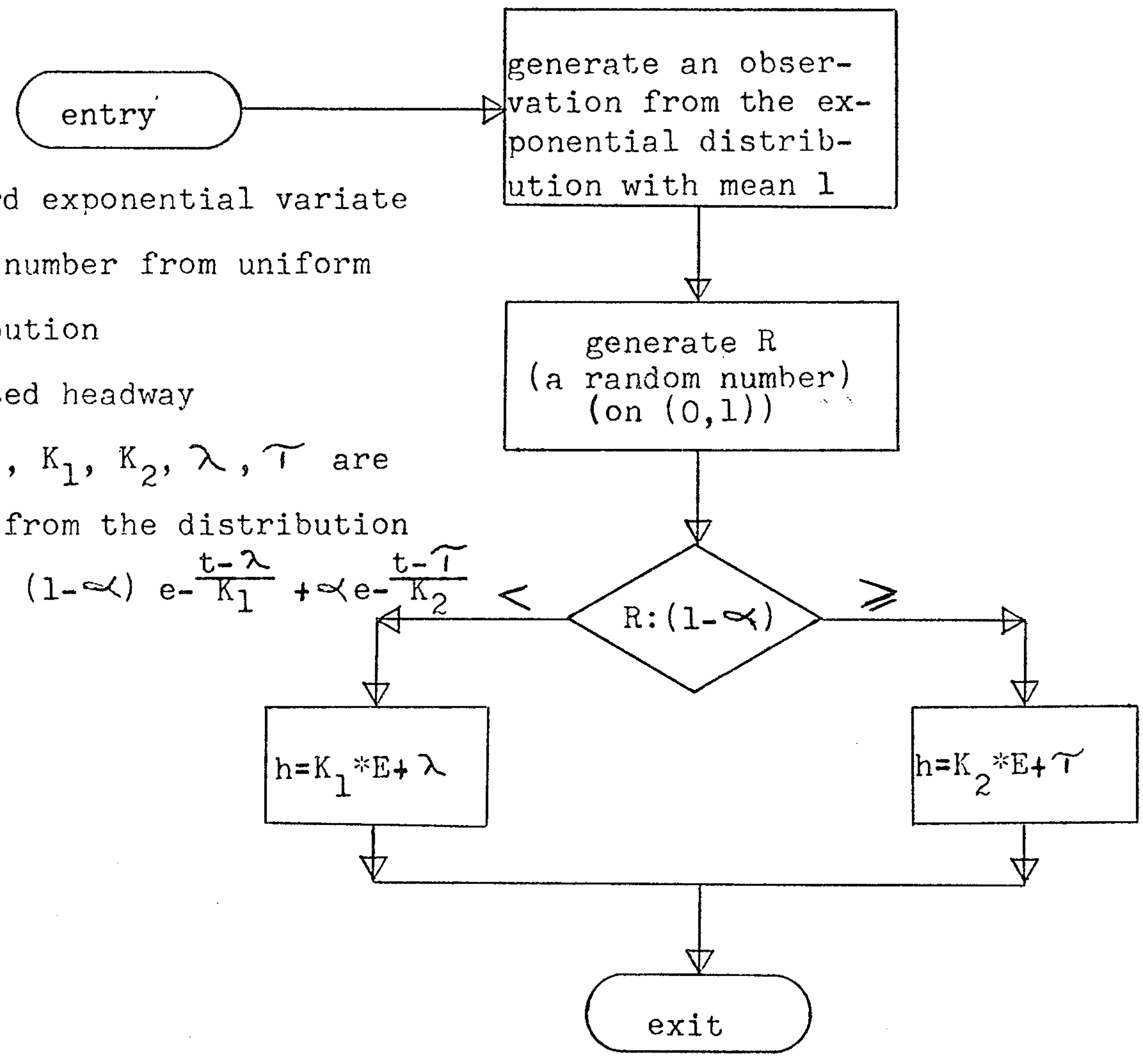
E= Standard exponential variate

R= Random number from uniform distribution

h= Generated headway

α, $(1-\alpha)$, K_1, K_2, λ, τ are constants from the distribution

$$P\,(h \geqslant t) = (1-\alpha)\, e^{-\frac{t-\lambda}{K_1}} + \alpha\, e^{-\frac{t-\tau}{K_2}}$$



In examining fig. 6.1 it was noticed that each time the routine is entered, a new starting value begins the sequence of pseudo-random numbers which ultimately generate only one headway. What is required is that the pseudo-random numbers be iteratively generated from entry to entry, until, say, an hour of headways is produced. Thus, a particular volume of traffic can be reproduced.

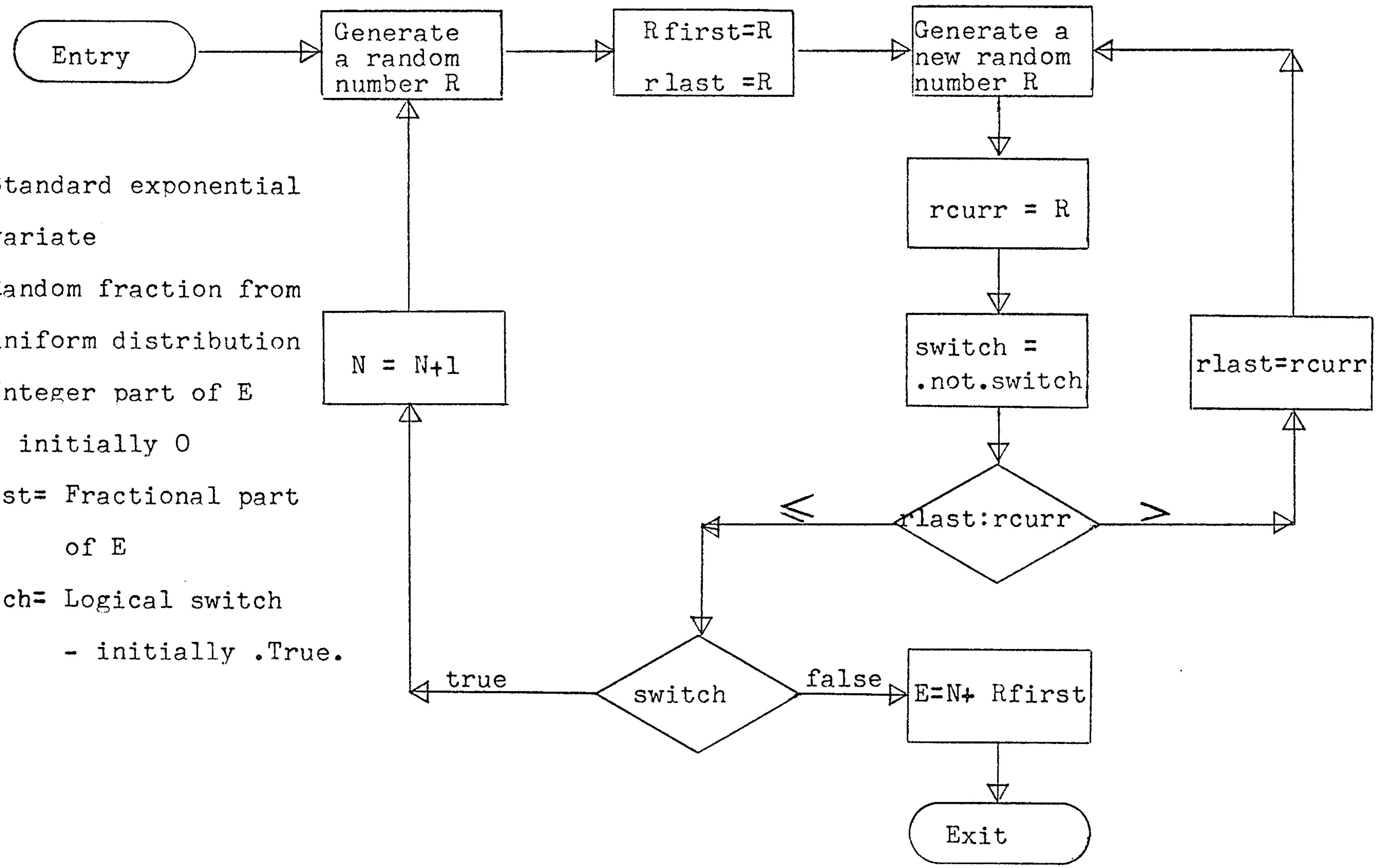
An improved flow diagram for the generation of standard exponential variates which incorporates the above modification is presented in fig. 6.3. It should be noticed that, unlike the flow diagram used by Kell, fig. 6.3 incorporates no subscripting. The basis of the method is that the algorithm generates both the integral part of E (the standard exponential variate) and the fractional part of E independently by two rejection methods.

Kell found that, using his method, different hourly volumes of traffic were generated for a requested volume rate, depending upon the random number used to begin the hour of generation. Because of this variation, he felt that the analysis of the simulation output would be more complex.

To overcome this, a program was written to obtain random numbers that would produce volumes within a certain tolerance (+2% and -1%) of the desired volume. Kell then tested each successfully generated distribution against the theoretical distribution by means of the Kolmogorov-Smirnov test of goodness of fit. When this test was successful, the time into the next hour of the last vehicle, and the time preceding the end

fig. 6.3

Standard exponential variate generator



E= Standard exponential variate

R= Random fraction from uniform distribution

N= Integer part of E - initially 0

Rfirst= Fractional part of E

Switch= Logical switch - initially .True.

of the hour of the last 12 vehicles were output. These times were used to order the random numbers to ensure continuity and volume accuracy. Sufficient random numbers for a particular volume were ordered and used as simulation input to provide the desired number of hours of simulation.

Two things are noticed about the above procedure. First is the use of the Kolmogorov-Smirnov test of goodness of fit. As pointed out in chapter 5, the general consensus amongst statisticians is that this test should not be used in cases where the parameters of a distribution need to be estimated. Instead, the chi-square test is recommended. Second is the selection of only those random numbers which would begin a generation of volumes within a particular tolerance of a requested volume. Kell initially assumes a random distribution of headways and then attempts to do away with this random element. While one can appreciate Kell's attempt to simplify the analysis of his output, complexity alone is not sufficient justification for the shortcut which he attempts to use.

6.3 The Shifted Negative Exponential Distribution

In a like manner, it was necessary to randomly generate headways that fit the shifted negative exponential distribution which describes local transit traffic. This can be accomplished using the inverse probability integral transformation method.

Recalling the cumulative form of (5.3.3),

$$P\,(h \geqslant t) = e^{-\frac{t-\eta}{\bar{t}-\eta}}\ ^{*},$$

where

$\bar{t}$ = average headway, sec;

η = minimum possibile headway between successive transit arrivals, sec.

Using a location and scale transformation of a standard exponential variate, and the fact that if $E=-\ln R$, where R is a random number on (0,1), then E is a standard exponential variate,

$$t = \eta + (\bar{t} - \eta)\, E,$$

and

$$t = \eta - (\bar{t} - \eta) \ln R.$$

t is a random variable with the correct distribution.

6.4 Compatibility

There is one additional problem which needs to be discussed. Following the use of the flow diagram in fig. 6.2, to generate headways which conform to the composite negative exponential distribution, and the inversion method, to generate headways which conform to the shifted negative exponential distribution, one is left with two sets of random numbers. It is now necessary to determine if these two distributions, so generated, are mutually compatible.

*The symbol η is substituted for the previously used symbol τ.

Because the basic philosophy is to interject transit vehicles into the main traffic stream, mutual compatibility means:

$$Tb_i - Tc_j \geqslant Lc_{ji},$$

and

$$Tc_{j+1} - Tb_i \geqslant Lb_{ij} \; *$$

where

Tb_i = arrival time of ith bus;

Tc_j = arrival time of jth car or truck;

Lb_{ij}= minimum headway between the ith bus and the jth car or truck;

Lc_{ji}= minimum headway between the jth car or truck and the ith bus.

As a test of this final condition, a program was written which generated vehicle arrivals according to the composite distribution, and generated transit traffic which conformed to the shifted exponential distribution. For both, the chi-square test of goodness of fit was applied to test the generated distribution against the theoretical distribution.

*It is noted that

$$Tb_i = Hb_i + Tb_{i-1} \quad \text{and } Tb_1 = Hb_1,$$

$$Tc_j = Hc_j + Tc_{j-1} \quad \text{and } Tc_1 = Hc_1$$

where

Hb_i = headway of ith bus;

Hc_j = headway of jth car or truck.

The program used a 2-lane approach volume of 568.819 vph. Of this, the volume of buses was 48.349* vph. while the volume of cars and trucks was 520.47 vph. Since buses were restricted to operating in the right lane, it was only necessary to generate car and truck traffic for this lane to test the merging of these two streams. Car and truck traffic for the curb lane was 303.47 vph.

Parameters for the composite distribution:

$\alpha = .253$ $(1-\alpha) = .747$ $K_1 = 13.87$

$K_2 = 2.295$ $\lambda = 1.52$ $\tau = .894$.

Parameters for the shifted exponential distribution:

$\eta = 40/44 = .909$ $\bar{t} = 3600/48.349 = 74.459$.

To test the compatibility of the two distributions more rigorously, it was assumed that, in all cases, the vehicle preceeding the arrival of a bus was a truck of maximum length. Thus $Lc_{ji} = 1.14$** for all j,i. Also, since all buses are approximately 40' in length, $Lb_{ij} = .909$ for all i,j.

*To make the test more rigorous, local transit traffic accounted for 8.5% of the 2-lane approach volume. This was the maximum percentage suitable for simulation (see section 5.5).

**In section 3.4 it was noted that a workable range of values for the length of trucks is 20' to 50'. Thus, at 30 mph., a truck of maximum length represented a headway of 1.14 sec.

fig. 6.4

Tri-distribution random headway generator

R= Random number from uniform distribution

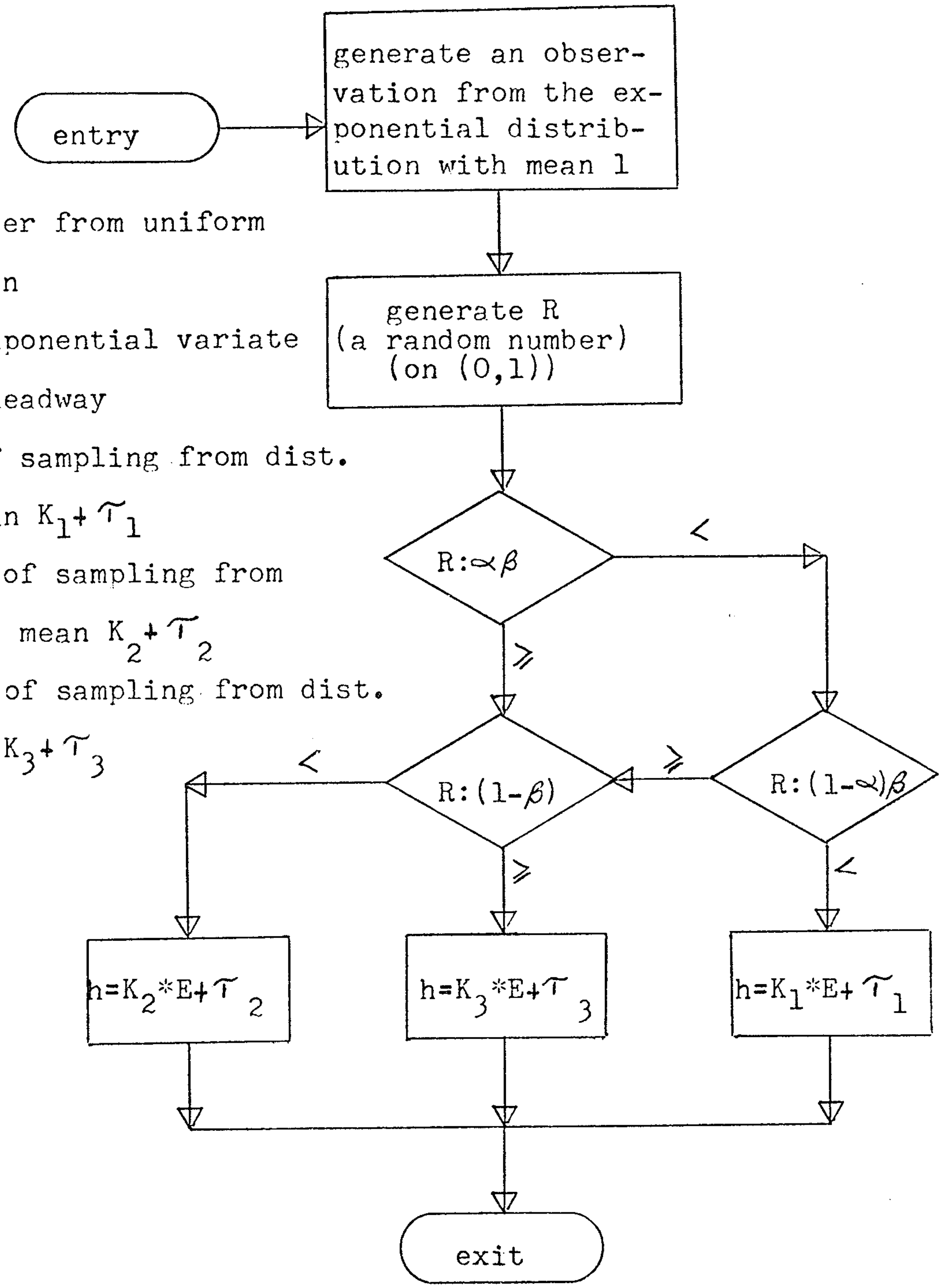
E= Standard exponential variate

h= Generated headway

$\alpha\beta$ = Prob. of sampling from dist. with mean $K_1+\tau_1$

$\beta(1-\alpha)$= Prob. of sampling from dist. with mean $K_2+\tau_2$

$(1-\beta)$= Prob. of sampling from dist. with mean $K_3+\tau_3$



Using the procedure outlined above and comparing different sets of hourly bus traffic to sets of car and truck traffic, it was found that hourly rates of local transit vehicles and corresponding rates of car and truck movement can be generated which are mutually compatible.

However, it should be noted that the two lane approach volume used to test this procedure is quite low. As the volume of car and truck traffic increased, it became increasingly difficult to find hourly rates of bus movement which were compatible with car and truck traffic. Sets of bus movement which were compatible, were so by chance alone.

What is required then is an algorithm which gives the composite distribution for cars and trucks, and the shifted exponential distribution for buses, and which guarantees, a priori, that minimum headways will be achieved. A diagram of such an algorithm is presented in fig. 6.4. Here, the mixture distribution is further exploited to include three headway distributions. Thus, with probability $\alpha\beta$ a shifted exponential distribution with mean $K_1+\tau_1$ is sampled, with probability $(1-\alpha)\beta$ a shifted exponential distribution with mean $K_2+\tau_2$ is sampled, and with probability $1-\beta$ a shifted exponential distribution with mean $K_3+\tau_3$ is sampled.

CHAPTER SEVEN

SPEED

7.1 Introduction

Here the desired or target speed of vehicles in the traffic stream is examined. This value is important in characterizing a roadway since much of a driver's evaluation of a roadway depends on the speed at which he can operate. (By the same token, the speed at which a driver can operate depends on his evaluation of the roadway.)

In addition, the value of a target speed is required in the evaluation of other relationships in later chapters. Thus, in chapters 5 and 6, it is used to determine the values of headways corresponding to vehicular lengths. In chapter 12, most of the signal timing methods examined depend upon a value of the target speed.

7.2 Definitions (Highway Capacity Manual (1965)):

Speed--the rate of movement of vehicular traffic or of specified components of traffic, expressed in miles per hour.

Spot speed--the speed of a vehicle as it passes a specified point on a roadway.

Medial friction--the retarding effect on the free flow of traffic caused by interference among traffic units proceeding in opposite directions.

Marginal friction--the retarding effect on the free flow of traffic caused by interference of any sort at the margin of the highway.

Stream friction--the retarding effect on the free flow of traffic caused by mutual interferences among traffic units proceeding in the same direction. Conflicts are caused primarily by differences in sizes and speeds of traffic units.

7.3 Discussion

In most simulation models, as each vehicle is generated, it is assigned a desired or target speed. A vehicle, placed on a lane approaching the intersection, is assigned a current speed equal to its desired speed. All vehicles attempt to maintain this speed, unless otherwise impeded by traffic conditions.

This approach was used by both Gerlough and Wagner (1967) and Messrs Buhr, Meserole, and Drew (1968) in their simulation work. Speeds were pseudo-random numbers generated so as to be normally distributed.

However, while there is sufficient evidence to suggest that the distribution of spot speeds along a roadway is normal (Traffic Engineering Handbook (1965), Gerlough (1964), Drew (1968)), this follows as a logical consequence of vehicular interaction. The mean and standard deviation of speeds decrease as the flow increases. At low volume rates, speeds depend almost solely on intersection signalization and midblock marginal frictions, while, at higher flows, stream frictions restrict attainable speeds (Highway Capacity Manual (1965)).

However, there is no evidence to suggest that the target or desired velocity of a vehicle is normally distributed. Indeed, Lewis and Michael (1963) postulated that all vehicles operate so as to minimize their delay. Thus, all vehicles in their simulation model attempted to travel at an average velocity of 30 mph. or 44'/sec.

This convention was also followed by James Kell in both of his traffic simulation models (1962, 1963). In addition, since all buses operate according to prescribed time schedules, they are even more inclined to maintain the maximum speed allowable under the speed limit.

Thus, for purposes of simulation, all generated vehicles would, upon introduction onto an approach lane, attempt to operate at the average speed $\bar{V}=30$ mph. Though a vehicle would be affected by the presence of other vehicles in the stream, in no case would $\bar{V}$ be exceeded.

CHAPTER EIGHT

ACCELERATION AND DECELERATION

8.1 Introduction

In this chapter, the maximum potential rates of acceleration and deceleration are investigated for the case of passenger cars, trucks, and local transit vehicles. These rates would not necessarily be explicitly included in a simulation model. Much of the spectrum of vehicular behaviour at an intersection involves a tracking or following process. The acceleration and deceleration rates inherent in such behaviour can be described by the stimulus-response equations of car following theory. The maximum rates of acceleration and deceleration play no real part in the application of these equations.*

However, they are important for making qualitative judgments about other relationships and for evaluating other quantities. Thus, in chapter 10, the maximum rate of

*It should be noted that, in most models, if the acceleration or deceleration rate calculated by one of the stimulus-response equations exceeds the maximum rate of acceleration or deceleration, the maximum value is used.

acceleration becomes a useful criterion for judging the reasonableness of results presented by Miller (1968) concerning driver reaction to the amber signal. Later, in chapter 12, the maximum rate of deceleration becomes a prime consideration in determining an appropriate value for use in evaluating the nondilemma yellow period.

8.2 Passenger Cars

According to the Traffic Engineering Handbook (1965), practical values of deceleration used in everyday traffic conditions rarely exceed 8 to 9 ft/sec^2. A deceleration rate of 11'/sec^2 is considered undesirable, but not alarming to passengers. At decelerations of 14'/sec^2, packages slide off the seat, and the occupants of the vehicle find this rate uncomfortable. Decelerations of 20'/sec^2 are used only in emergencies while, at decelerations above this value, the occupants brace themselves to avoid being thrown off the seat.

Aside: Vehicles being stopped by brake action are decelerated by the force of friction between tires and road. The maximum amount of braking force is limited by the maximum coefficient of friction, f between road surface and tires. Deceleration caused by braking is equal to 32.2 f. At a speed of 30 mph. and on dry pavement, $f \cong .62$. Thus, maximum deceleration is approximately 20'/sec^2.

In fact, Saal and Petring (1957) found that, although about 98% of passenger vehicles are capable of deceleration rates greater than 20'/sec^2, only 50% are capable of deceleration rates greater than 30'/sec^2. Rates greater than this are mechanically possible, however only rarely would they occur under everyday traffic conditions.

Unfortunately, not as much data are available concerning maximum acceleration rates. Data presented by the Ford Motor Company (Traffic Engineering Handbook (1965)) showed that the maximum acceleration rates of 1963 passenger cars vary from 8.8'/sec^2 for a compact 6-cylinder car with automatic transmission to 14.93'/sec^2 for a low-priced 8-cylinder car with standard transmission. These values generally support the research of Beakey (1938) who found that typical acceleration rates are approximately half the value of deceleration rates.

8.3 Trucks

By their very nature, trucks have different operating characteristics than passenger cars. Greenshields (1947) found that trucks, on the whole, travel more slowly than the rest of the vehicles in the traffic stream. Also, trucks require greater stopping distance and tend toward lower uniform rates of deceleration than passenger cars.

Data gathered in New York City indicates that the maximum rate of deceleration for trucks is approximately $9'/sec^2$ (Greenshields (1947)).

The Traffic Engineering Handbook (1965) depicts the percentage of trucks capable of a given or a greater deceleration rate. Different graphs are presented for different types of trucks. Because the types of trucks range from very light 2-axle trucks to trucks and full trailers with 4, 5, and 6 axles, it is difficult to determine one value as a maximum deceleration rate. However, using the graph of heavy 2-axle trucks as an average, 90% of all such trucks are capable of deceleration rates greater than $15'/sec^2$, while 50% are capable of deceleration rates greater than $23'/sec^2$.

Surprisingly, virtually no research results have been published concerning the maximum acceleration capability of trucks. Data which is available - mostly the work of Deen (1957) and Greenshields (1947) - concerns the normal acceleration performance of trucks from a standing start. Because such acceleration rates average around $2'/sec^2$ to $3'/sec^2$ (Greenshields (1947), Traffic Engineering Handbook (1965)), it was felt that a maximum acceleration rate of approximately $10'/sec^2$ would be appropriate.

8.4 Local Transit Buses

The lack of information concerning the operating characteristics of local transit vehicles makes it difficult to determine accurately their maximum acceleration or deceleration capabilities. Greenshields (1947) observed, that between five and ten feet from a stop, the maximum average deceleration for buses is about 6.3'/sec^2. Furthermore, the cumulative distribution of maximum deceleration rates for a sample of buses shows that the limit of bus deceleration is 10.5'/sec^2.

Unfortunately, little importance can be attached to these findings. Greenshields examined only a limited number of buses, and the deceleration ability of transit vehicles has undoubtedly improved since he conducted his study.

What is interesting, however, is that the operating characteristics of both buses and passenger cars approaching a stop appear to be extremely similar. In fact, there is reason to believe that the deceleration capabilities of the two are roughly comparable. A maximum deceleration rate of approximately 20'/sec^2 seems representative of local transit vehicles.

While a high rate of deceleration is often necessary to avoid an accident, seldom is there a similar need for a high rate of acceleration. In addition, it must be remembered that a bus accelerating too quickly from a stop can result in injury to passengers standing inside the bus. For this reason,

fig. 8.1
ESTIMATED ACCELERATION CURVES

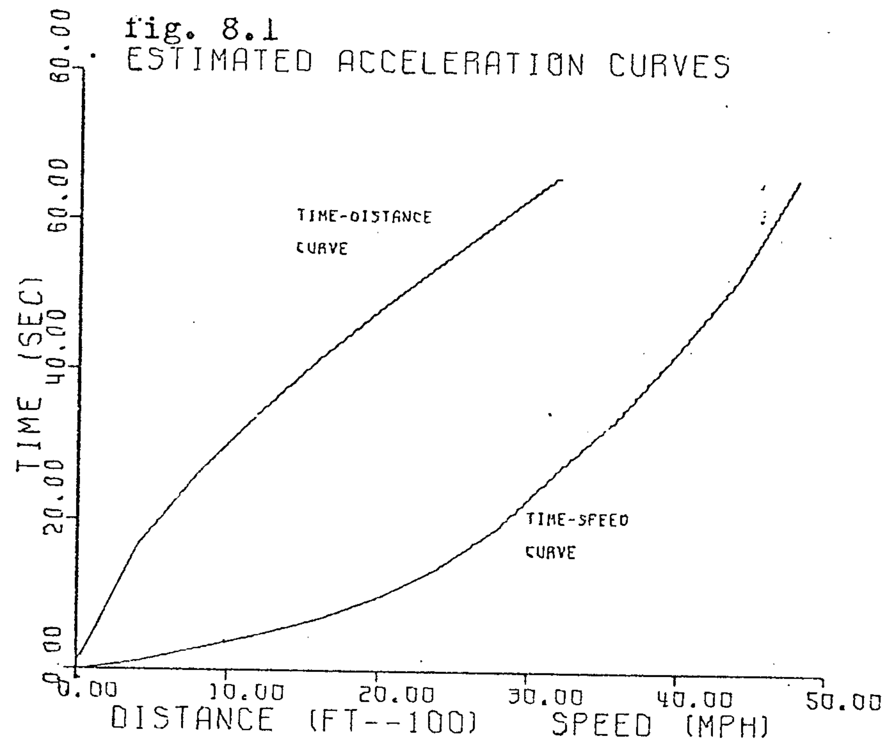
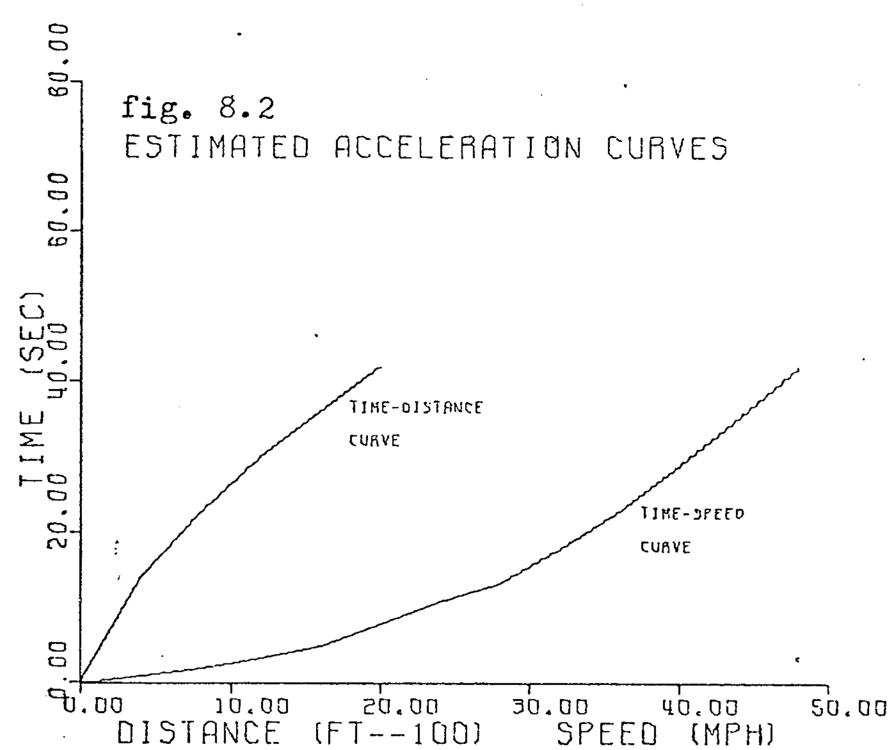


fig. 8.2
ESTIMATED ACCELERATION CURVES



it is suspected that the maximum rate of acceleration for local transit vehicles is quite low.

Tests conducted by Western Flyer Coach Ltd. on 196 of their buses produced two sets of estimated acceleration curves (see figures 8.1 and 8.2).* From these data, the maximum rate of acceleration was 6'/sec^2. While it is realized that different bus manufacturers produce buses with different operating characteristics, these data provide a representation of the acceleration capability of local transit vehicles.

*The Metropolitan Corporation of Greater Winnipeg maintains a transit fleet supplied by Western Flyer Coach Ltd., General Motors Ltd., and Mitsubishi. No comparable acceleration curves are available for either G.M. or Mitsubishi buses.

CHAPTER NINE

LEFT TURN GAP/LAG ACCEPTANCE

9.1 Introduction

When a left turn is made at a signalized intersection, one or more lanes of opposing traffic are crossed. The opposing stream is evaluated and a time interval selected which is considered long enough to cross safely into the cross street. The interval for this maneuver is known as the acceptable gap/lag.

Early field observations designed to examine this phenomenon were concerned primarily with stop sign controlled intersections. It was not until Dart (1968) conducted field studies to determine left turn characteristics that this phenomenon was thoroughly investigated for signalized intersections. The results of his work, along with the results of earlier studies, are presented at the beginning of section 9.3, wherein gap acceptance is discussed, and section 9.4, which is devoted to lag acceptance.

Some researchers found it sufficient to examine only the average values of accepted time gaps/lags, while others sought to determine a distribution which characterized this phenomenon. To this end, the log-normal distribution was widely investigated. It emerges as a generalized form for describing gap/lag acceptance. The latter parts of sections 9.3 and 9.4

are devoted to these studies and their results.

The shortage of information concerning gap/lag acceptance was very apparent in the case of trucks. The results of relevant studies which were available are presented at the beginning of section 9.5. The latter part of this section is devoted to a procedure whereby the gap/lag acceptance mechanism may be successfully modelled for the case of truck traffic.

Together gap and lag acceptance form an important characteristic of left turning traffic and have a severe influence on intersection capacity. If the average gap/lag accepted by left turn vehicles is very short, through traffic is forced to decelerate to avoid a collision. This increases delay to through vehicles and reduces the capacity of the intersection. Conversely, if the average gap/lag accepted is too long, left turn vehicles themselves endure undue delay in the intersection, and consequently delay other turning and nonturning vehicles in the same stream. Again, the intersection capacity is reduced.

In a simulation model, when a left turn vehicle reaches a point* near the intersection, a decision is made probabilistic-

*Gerlough (1967) called this point the start-free-left-turn position. Though its exact location depends upon the geometry of the intersection, it corresponds approximately to the near-side boundary of the cross-walk.

ally whether to accept or reject the lag in the opposing stream. A lag is considered rather than a gap because the first gap encountered is unimportant. The lead vehicle , which made up the gap, may have passed long before the turning vehicle was positioned to execute the turn.

If the lag is accepted, the vehicle continues through the turn. If rejected, the vehicle begins stopping at the constant rate required to stop at the left-turn waiting point. Thereafter, the driver views each time gap in the opposing stream and, when an acceptable gap is found, the vehicle negotiates a delayed left turn.

Kell (1962) outlined a table-look-up procedure to determine gap/lag acceptance. When a gap/lag of size x is presented, the probability of its acceptance is obtained from a table of probability values. A random number (between 0 and 1) is generated and compared with the probability value. If the random number is larger, the gap/lag is rejected; otherwise it is accepted.

In sections 9.3 and 9.4, tables of probability values, drawn from log-normal distributions, are presented. These tables are thought to be suitable for determining gap/lag acceptance in an actual simulation model.

Using Kell's procedure, there is a chance that a gap of a certain size will be rejected, while one even shorter will later be accepted. Bissell (1960) found this occurred for about 5% of vehicles observed, while Dart (1968) recorded this as occurring about 10% of the time.*

9.2 Definitions (Greenshields (1947), Lewis and Michael (1963), Dart (1968), Drew (1968)):

Gap--the time interval between the leaving of a point by one vehicle and the arrival at the same point by the next succeeding vehicle travelling in the same direction. Measurements made are from the rear of one vehicle to the head of the succeeding vehicle.

Lag--the time interval between the arrival of one vehicle on an approach and the arrival of another vehicle on an opposing approach at a point of conflict. The point of conflict is the intersection of the centre lines of the two vehicle paths.

Critical gap--the size of the gap whose number of accepted gaps shorter than it is equal to the number of rejected gaps longer than it.

Critical lag--the size of the lag whose number of accepted lags shorter than it is equal to the number of rejected lags longer than it.

*It must be remembered that each driver can reject only one lag, while he can reject many gaps. Thus, the phenomenon of a driver accepting a certain time interval shorter than the one previously rejected does not apply to lag acceptance.

Acceptable average - minimum time gap (AAMG)--that time gap which will be accepted by more than 50% of the drivers.
Acceptable average - minimum time lag (AAML)--that time lag which will be accepted by more than 50% of the drivers.

9.3 Passenger Car Gap Acceptance

Most previously conducted gap acceptance studies have not specifically considered left turns at 4-lane signalized intersections in the central business district. Kaiser (1951) analyzed 158 left turns at a stop sign controlled intersection. He found the smallest gap accepted was 3.75 sec. while the largest gap rejected was 4.75 sec. The critical gap was 4.25 sec. At a similarly controlled intersection, Clark (1960) found a critical gap of 3.2 sec.

The critical gap observed at a stop sign controlled intersection would not, however, be the same as that observed at a signalized intersection. A vehicle, executing a left turn from a stop sign, intersects the paths of vehicles approaching from the left and converges with those coming from the right. In short, two different types of gaps are involved. At a signalized intersection, his movement intersects only the paths of on-coming vehicles.

Apparently, only Greenshields (1947) chooses to acknowledge the distinction between the two types of gaps encountered at a stop sign. He found that the acceptable average-minimum gap for a crossing maneuver was 6.1 seconds while, for a converging maneuver, it was 4.1 seconds.

By contrast, Betz (1967) analyzed 260 left turns at a 4-lane signalized intersection, and obtained a critical gap of 5.5 seconds. A more extensive field study was conducted by Dart (1968) at a 4-lane signalized intersection. He considered four types of gaps in the opposing stream. There were two types of lane gaps, ie., gaps formed by two successive inside-lane vehicles or two successive outside-lane vehicles. In addition, there were two types of offset gaps, ie., gaps formed by a vehicle in one lane trailed by a vehicle in the adjacent lane.

Although the sample size of each type of turn was small, there did appear to be evidence that a gap formed by an outside lead and an inside lag (known as a Type 4 gap) is more likely to be accepted than any of the other types of gaps of the same size. Dart did not determine critical values for the different types of gaps. However, from the cumulative distributions, it is noted that the AAMG for a Type 4 gap is approximately 3.75 sec., as compared to 4.9 sec. for the other types of gaps.

Using the critical gap concept, some researchers assumed waiting drivers evaluate each intercar gap, turning if the gap is greater than some predetermined time gap T. A more realistic model was suggested by Herman and Weiss (1961) who associated with each time gap t, a gap acceptance probability P (t). Thus, there exists a probability P (t) that a driver, faced with a gap t, accepts it and completes his turn.

Drew (1968) presented several representative forms of gap acceptance distributions, including the log-normal distributions. This distribution, in particular, has been found to provide the best description of the gap acceptance phenomenon.

Bissell (1960) collected and analyzed data for those vehicles leaving a cross street under stop sign control. He found that the gap acceptance distributions were of a log-normal form. The distributions are shown in fig. 9.1. An independent study conducted in Australia verified these findings.

These studies provided Kell (1962) with the impetus to conduct his own investigation. Analyzing over 500 field observations on 2-lane 2-way streets, he found the left turn gap acceptance distribution given in

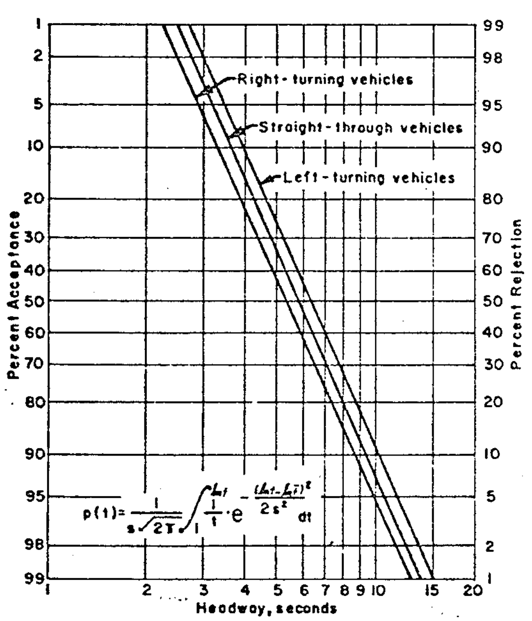


fig.9.1 Gap acceptance distributions at a stop sign.

table 9.1, and pointed out the log-normal nature of these data. Later, Gerlough (1967) developed a log-normal probability distribution for left turn gap acceptance from data supplied by Kell (see table 9.2). These data were modified to apply to a 4-lane signalized intersection. Unfortunately, Gerlough does not explain what these modifications are, or by what criteria they were applied.

Table 9.1

Gap Acceptance on Two-Lane, Two-Way Streets

Gap Size (sec)	Cumulative % Accepting	Gap Size (sec)	Cumulative % Accepting
1.0	0	4.5-5.0	94.7
1.0-1.5	1.4	5.0-5.5	96.4
1.5-2.0	10.2	5.5-6.0	97.9
2.0-2.5	18.3	6.0-6.5	98.2
2.5-3.0	31.3	6.5-7.0	98.5
3.0-3.5	50.0	7.0-7.5	99.3
3.5-4.0	64.6	7.5-8.0	99.4
4.0-4.5	85.3	8.0+	100.0

mean = 3.27 st. dev. = 1.13

Unlike Dart (1968), Gerlough did not distinguish among different types of gap formations. Betz (1967) agreed. Betz acknowledged that gap acceptability depends to a certain extent on the lane position of on-coming cars.

However, he also felt that "determining different acceptable gaps for different positions of cars in the through lanes introduces unnecessary complexity."

This is debatable. If the consequences are detrimental to a model then complexity alone is not a sufficient reason to sacrifice realism. However, because so little information is available concerning different gap acceptance distributions for different gap formations, using one distribution to describe the entire range of gap acceptance circumstances is, if not justifiable, then at least excusable. In all probability, the net effect of such a decision on a simulation model is slight.

Table 9.2

Left Turn Gap Decision Probability

Gap Size (sec)	Probability of Accepting Gap	Gap Size (sec)	Probability of Accepting Gap
3.0	0		
3.0-3.5	.150	6.5-7.0	.970
3.5-4.0	.320	7.0-7.5	.986
4.0-4.5	.520	7.5-8.0	.993
4.5-5.0	.690	8.0-8.5	.997
5.0-5.5	.820	8.5-9.0	.998
5.5-6.0	.900	9.0-9.5	.999
6.0-6.5	.950	9.5+	1.000

mean = 4.35 st. dev. = 1.06

It was finally decided that table 9.2, developed by Gerlough, would be suitable for purposes of simulation. There are several reasons for this decision.

First, the acceptable average-minimum time gap for this distribution is approximately 4.2 seconds. This compares favourably with those values determined by Dart (1968) for different types of gaps, but is greater than the AAMG of 3.25 sec., used by Kell (1962), for a 2-lane 2-way street (table 9.1).

Secondly, the variance is quite low. Drew (1968) felt this is desirable in maintaining consistency of driver behavior. In the field, consistency is important in achieving a high degree of safety. In simulation models, it is important in minimizing the effect of the gap acceptance phenonemon on the other interactions being studied.

Thirdly, although Gerlough fails to explain how he modified Kell's data to meet the conditions of his own model, the fact remains that Gerlough's model is highly compatible with the one discussed here, namely an isolated, 4-lane, bi-directional, signalized intersection. In fact, it is the only table of values in the literature which is totally applicable.

9.4 Passenger Car Lag Acceptance

Much of the foregoing discussion on gap acceptance applies equally to the acceptance of lags by moving, left turn vehicles. The main difference is the amount of information available concerning a lag acceptance distribution.

Dart (1968) investigated two types of lags; one with the nearest opposing vehicle in the outside lane (Type 2), and the other with the nearest opposing vehicle in the inside lane (Type 1). He found, for a given lag size, a lower percentage of drivers accept the lag if it is formed by an opposing vehicle in the outside lane, as compared with one in the inside lane. For a Type 1 lag, the acceptable average-minimum lag was about 4.0 sec. while, for a Type 2 lag, it was 4.5 sec.*

Table 9.3

Left Turn Lag Decision Probability

Lag Size (sec)	Probability of Accepting Lag	Lag Size (sec)	Probability of Accepting Lag
3.0	0		
3.0-3.5	.030	6.0-6.5	.940
3.5-4.0	.124	6.5-7.0	.970
4.0-4.5	.300	7.0-7.5	.990
4.5-5.0	.530	7.5-8.0	.996
5.0-5.5	.730	8.0-8.5	.999
5.5-6.0	.860	8.5+	1.000

mean = 4.77 st. dev. = .92

*These values were read from distribution curves which Dart visually fitted to his data, after attempts to fit different mathematical distributions proved unsuccessful. He did, however, acknowledge that such data usually takes a log-normal form.

Using data supplied by Kell, Gerlough (1967) constructed a table of lag acceptance probabilities from a log-normal distribution (see table 9.3). The acceptable average-minimum time lag for this distribution was 4.69 sec. As before, Gerlough does not distinguish between the different lag formations.

For much the same reasons that the gap acceptance distribution determined by Gerlough was accepted, it was decided to adopt his lag acceptance distribution as well. Wohl and Martin (1967) point out that determination of a gap/lag in any particular situation should be based on an empirical study, since the acceptable gap/lag length is probably a function of driver and vehicular characteristics, physical layout of the intersection, and general traffic conditions. In the absence of empirical data, these gap and lag acceptance probability tables offer the most appropriate alternatives.

9.5 Truck Gap/Lag Acceptance

Lack of available information makes it impossible to construct tables of gap and lag acceptance probability distributions for trucks, similar to those adopted for passenger cars. The only related research is that reported by Noblitt (1959). He found gap acceptance for left turn truck combinations to be 1.4 to 1.8 times as large as the required gap for cars, and 1.2 to 1.5 times as large as the required gap for single-unit trucks. Unfortunately, it is not known if this refers to signalized or unsignalized intersections. In any case, this is hardly a sufficient base from which to construct a gap acceptance distribution. A different approach was elected. Gerlough (1967) defined a left turn waiting point as being one-quarter way through the turn. From here a stopped vehicle analyzes gaps in the on-coming traffic stream. At this waiting point, a gap was either accepted or rejected by reference to the gap acceptance probability table. However, Gerlough made no provision for a separate left turn gap acceptance distribution for trucks.

Dart (1968) further researched the position at which a delayed left turn vehicle waits. He found that 59.6% waited near the centre of the intersection, while 32.5% waited at a location about half-way into the intersection from the stop line, ie. one-quarter way through the turn.

The percentage of vehicles which chose each position is not really significant, since only about 100 turns were sampled. What is important, is that there appears to be roughly two different left-turn wait positions.

With this in mind, the following is felt to be an appropriate procedure. In a simulation model, all passenger cars could be stopped at a waiting point one-quarter way through the turn, while all delayed left turn trucks could be stopped one-half way through the turn. Compensation could thus be provided for the greater length and lower acceleration capability of trucks, without developing a different gap acceptance probability distribution.

The problem of a lag acceptance probability distribution for trucks can be similarly overcome. Whereas passenger cars decided either to accept or reject a lag at a start-free-left-turn position, trucks could make their decision one-quarter way through the turn.

CHAPTER TEN

Response to the Amber Signal

10.1 Introduction

There is a problem associated with the amber signal which arises when a driver is too close to the intersection to stop safely, and too far from the intersection to proceed and pass completely through before the red phase commences. Thus, an improperly timed amber phase leads to the formation of a "dilemma zone."

In this chapter, the problem of the amber signal and drivers' responses to the amber are examined. An equation developed by Gazis et al. (1960) for eliminating dilemma zones in the field is presented in the beginning of section 10.3. Following this, the results of different research projects designed to understand driver response to the amber, and the applicability of these investigations to models of simulation, are presented.

While a reasonable degree of research has been conducted concerning this problem, as it relates to passenger cars, almost no studies have been reported concerning the response of truck traffic to the amber. This is discussed in section 10.4 and a method presented, whereby tables of distribution functions developed for describing car behaviour may be modified to apply to truck traffic as well.

In section 10.5, the logic involved in modelling this behavior, using mechanisms presented in the previous sections, is discussed. Separate attention is reserved for the case of local transit vehicles, which are examined in subsection 10.5.2.

10.2 Definitions (Manual of Uniform Traffic Control Devices for Canada (1966)):

Clearance interval (amber)--allow (s) time for approaching vehicles to stop in safety. The length...should be sufficient to allow a comfortable stop from the normal approach speed, and at the same time, it should allow a vehicle, which is too close to the intersection to stop safely, the time necessary to pass through the intersection before the appearance of the red.

10.3 Passenger Cars

Gasiz et al (1960) presented criteria for constructing an amber phase which would eliminate "dilemma zones." They suggest that the yellow period should be such, that a vehicle which could just stop (on seeing the yellow light) before entering the intersection, could also continue at uniform speed and cross the intersection just before the light turns red. This non-dilemma yellow period is

$$\tau_{min} = \delta + \frac{Vo}{2d} + \frac{W+L}{Vo} \qquad (10.3.1)$$

where

τ_{min} = minimum amber phase duration, sec;

δ = reaction-decision-making time of driver, sec;

Vo = approach speed of vehicle, ft/sec;

W = width of intersection to be traversed, ft;

L = vehicle length, ft;

d = constant rate of deceleration (in practice this represents the maximum average deceleration to which it is desirable or practical to subject drivers.), ft/sec^2.

Equation 10.3.1 is widely accepted in publications of signal design. A summary of these publications, along with the values which were assumed for parameters δ, d, and L, is presented in table 10.1. In each publication, charts are presented for the minimum clearance interval according to different approach speeds and cross-street widths.

For purposes of simulation, there is still a problem associated with the amber signal. If the light changes to amber as a vehicle approaches the intersection, a decision has to be made as to whether the vehicle will stop or not.

The first to investigate this problem was Greenshields (1947) who studied those drivers forced to make close decisions.

Table 10.1

Publications using Equation 10.3.1 Along with Assumed Parametric Values.

Publication	Parameter δ (sec)	d (ft/sec^2)	L (ft)
Traffic Engineering Handbook (1965)	1	15	20
Manual of Uniform Traffic Devices (1966)	.795	12.41	18.13
Hewton (?)	1.5	8.8	*

He reasoned that, of the vehicles passing through an amber, the last is most likely to have stopped, while of the vehicles stopping for the red, the first is most likely to have passed through on the preceeding amber. The distance to the stop line, divided by the speed of the vehicle when the signal changes to amber, is the "potential time" to the intersection.

*Hewton considered intersection width as the greatest distance a vehicle covered to be completely free of traffic entering the intersection on the next phase. Thus, vehicular length L is included as part of the width of the intersection.

Greenshields found that no vehicle whose "potential time" was greater than 3.99 sec. passed through the intersection, while the majority of vehicles passing through were less than 3.0 sec. from the intersection. Basically, the two groups were distinct, having an overlap around the three to four second groups. Unfortunately, only 47 close decision vehicles were included in this study. Thus, no accurate examination can be made of the probability that a vehicle, caught in this zone, proceeds through the intersection.

Another study was conducted by Webster (1959). He constructed a mock-up light signal wherein vehicles approaching the light at specific speeds, triggered the light to the amber themselves. From this somewhat artificial situation, Webster constructed a table showing the probability of stopping for different approach speeds (see table 10.2).

Table 10.2

Probability of Cars Stopping as a Function of Their Speed and Distance From Signal At Onset of Amber[1]

Speed (mph)	Distance from Signal (ft)		
	0.50[2]	0.80[2]	0.95[2]
30	100	120	135
40	160	190	210
50	225	275	300

[1]Data from Webster (1959)
[2]Probability of Stopping.

Perhaps the most important investigation was performed by Olson and Rothery (1962). They obtained data from intersections representing low-, medium-, and high-speed approaches, and found three distinct probability distributions, each a function of the distance from the intersection. The curves representing these distributions are shown in fig. 10.1. For comparison, the curves based on Webster's data are also shown.

The distribution curves obtained by Olson and Rothery were readily adopted by Kell (1962) in his simulation of vehicular delay at intersections. However, since his work centered around California, he reported that additional data would be gathered in this area to refine the distributions.

What is curious, is that Kell felt justified in adopting these distributions at all. The functions presented by Olson and Rothery are applicable only for vehicles travelling at speeds of 30 mph, 40 mph, or 50 mph. If a vehicle travels at some other speed, there is no way to determine whether or not this vehicle stops when the light changes.

The problem is not that their study is too narrow, but rather that excessive demands are being placed upon their results. What Kell had overlooked was the reason Olson and Rothery had conducted this research in the first place.

Continuing the theoretical analysis and observations reported by Gazis et al (1960), they sought to determine whether driver behaviour changes as a function of different amber phase durations. Finding it does not, they intended to use their distribution functions to adjust amber phase settings to suit human performance. The scope of their investigation was adequate for these intentions, however for purposes of simulation it was not.

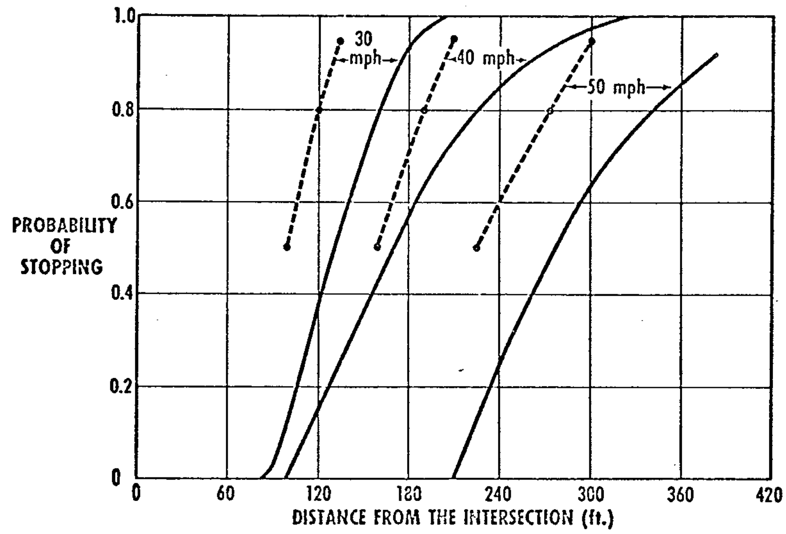


fig. 10.1 Curves for probabiliby of stopping for three speed zones. Dotted curves based on Webster data (table 10.2)

Olson and Rothery based their findings on measurements made in the field. Miller (1968) reports the work of Webster and Ellson who based their work on experiments. They found that the proportion of drivers who stop may be described by

$$Y = 12.5 \log_{10} d - .2125V - 14 \quad (10.3.2)$$

where

Y = the probit corresponding to the proportion of drivers who stop;

d = distance of vehicle from stop line when signal changed, ft;

V = speed of vehicle, mph.*

By converting from distances to times t, table 10.3 was obtained.

Table 10.3

Proportion of Drivers Who Stop When They are Within t Seconds of the Stop Line at the Signal Change.[1]

V (m.p.h.)	Time t (sec)				
	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$
20	.003 (19.6)	.126 (14.7)	.526 (11.7)	.854 (9.8)	.971 (8.4)
25	.005 (24.4)	.160 (18.3)	.585 (13.7)	.886 (12.2)	.979 (10.5)
30	.004 (29.3)	.143 (22)	.556 (17.6)	.871 (14.7)	.976 (12.6)
35	.002 (34.2)	.098 (25.7)	.467 (20.5)	.818 (17.1)	.959 (14.7)

1 Values in paranthesis are associated deceleration rates from $d = V/t$ (ft/sec^2).

Two things are noticed about this study. The first is the use of the probit method in determining the proportion of drivers who stop. The second is the extremely high rates of deceleration associated with some of the entries in table 10.3. In several cases, these rates approach the limit of vehicular deceleration associated with everyday traffic conditions.**

*A probit is a normal distribution transformation. The probit Y corresponding to a proportion P is defined as the solution of

$$P = \int_{-\infty}^{Y} \exp\left(-(x-5)^2/2\right)/\sqrt{2\pi}\, dx \quad \text{(Miller (1968))}$$

**Previously, in section 8.2, it was observed that the maximum vehicular deceleration rate associated with everyday traffic conditions is approximately 30'/sec^2.

Rather than consider the probability of stopping as a function of distance or travel time from the intersection, Gerlough (1967) computed the deceleration (negative acceleration) rates required to stop at the stop line. In making these computations, a 1-second reaction time was assumed. Using data from Olson and Rothery (1962), and supplementing it with data from several other sources, he noticed that the composite data formed a single distribution (see table 10.4). Plotting the data on normal probability paper and using the probit method, an excellent linear regression (R= .967) was obtained.

This distribution function was selected as being particularly suited for describing driver response to the amber in a simulation model. There are two reasons for this decision.

First, by presenting the probability of stopping as a function of deceleration, Gerlough removes the restriction which limited the application of the distribution functions developed by Olson and Rothery. Second, by presenting a range of deceleration rates which reflect the behaviour of the road user, rather than the limitations of the vehicle, Gerlough offers a degree of pragmatism singularly absent in the experiments of Webster and Ellson.

10.4 Truck Traffic

Gerlough never specified whether his amber decision probability table (ie. table 10.4) was designed for truck traffic

Table 10.4

Amber Decision Probability

Acceleration Required to Stop at Stop Line (ft/sec^2)	Probability of Stopping
-0 to -0.99	1.000
-1.00 to -1.99	0.994
-2.00 to -2.99	0.989
-3.00 to -3.99	0.982
-4.00 to -4.99	0.972
-5.00 to -5.99	0.956
-6.00 to -6.99	0.935
-7.00 to -7.99	0.905
-8.00 to -8.99	0.867
-9.00 to -9.99	0.820
-10.00 to -10.99	0.762
-11.00 to -11.99	0.700
-12.00 to -12.99	0.624
-13.00 to -13.99	0.548
-14.00 to -14.99	0.468
-15.00 to -15.99	0.390
-16.00 to -16.99	0.318
-17.00 to -17.99	0.250
-18.00 to -18.99	0.190
-19.00 to -19.99	0.140

as well as passenger cars. If it was, then it suffers from the same shortcoming that characterized the work of Webster and Ellson.

Table 10.4 includes deceleration rates as high as 20'/sec^2. While this is well within the capability of nearly all passenger cars (98% are capable of rates greater than this), only 50% of all trucks can attain deceleration rates this high. Some modification of this decision probability table is required.

In developing table 10.4, Gerlough allowed a 1-second reaction time. If this reaction time is not applied to truck traffic, then stopping times are increased by one second and lower deceleration rates result.

Using the equation for uniform deceleration motion (final velocity is zero)

$$Vo = d_c t_c,$$

or

$$t_c = Vo/d_c.$$

$$Vo = d_T t_T,$$

or

$$t_T = Vo/d_T;$$

where

d_c = deceleration rate for passenger cars, ft/sec^2;

d_T = deceleration rate for trucks, ft/sec^2;

t_c = time for passenger cars to decelerate to a stop, sec;

t_T = time for trucks to decelerate to a stop, sec;

Vo = initial velocity, ft/sec.

(Vo is the average speed $\bar{V}$ = 44'/sec.)

Table 10.5

Amber Decision Probability for Trucks

Acceleration Required to Stop at Stop Line (ft/sec^2)	Probability of Stopping
- 0 to - 0.97	1.000
- 0.98 to - 1.90	0.994
- 1.91 to - 2.80	0.989
- 2.81 to - 3.66	0.982
- 3.67 to - 4.48	0.972
- 4.49 to - 5.27	0.956
- 5.28 to - 6.03	0.935
- 6.04 to - 6.76	0.905
- 6.77 to - 7.46	0.867
- 7.47 to - 8.14	0.820
- 8.15 to - 8.79	0.762
- 8.80 to - 9.42	0.700
- 9.43 to -10.03	0.624
-10.04 to -10.61	0.548
-10.62 to -11.18	0.468
-11.19 to -11.72	0.390
-11.73 to -12.25	0.318
-12.26 to -12.76	0.250
-12.77 to -13.26	0.190
-13.27 to -13.75	0.140

Setting $t_T = t_c + 1$,

$$\frac{Vo}{d_T} = \frac{Vo}{d_c} + 1 .$$

Solving for d_T:

$$d_T = \frac{Vo\ d_c}{Vo + d_c} \qquad (10.4.1)$$

Applying (10.4.1) to the upper and lower class bounds of the deceleration rates of table 10.4, shifts these bounds by the amount of $d_c^2 / (44 + d_c)$ (see table 10.5).

The maximum deceleration rate in table 10.5 is 13.75'/sec^2, a rate well within the capability of more than 90% of heavy 2-axle trucks. Though most trucks are capable of deceleration rates greater than this, these are required for only a very small percentage of driving situations.

10.5 Logic

10.5.1 Passenger cars and truck traffic

In modelling this behavioural pattern, the deceleration rate required for a vehicle to stop at the stop line is calculated when the signal changes to amber. Using this rate, an amber decision probability table is entered and the probability of stopping is obtained. A random number (between 0 and 1) is generated and compared with the probability value. If the random number is greater, the vehicle proceeds; if less, it stops.

Initially, only a vehicle first in line on an approach is considered. If this vehicle proceeds, then the second-in-line vehicle will be examined. This process continues until a vehicle is found which stops. Of course, all vehicles behind it stop as well.

The above procedure applies to both passenger cars and trucks. The former could refer to an amber decision probability table (say table 10.4) in deciding whether to stop or proceed, while the latter would refer to a shifted table of probability values (say table 10.5).

10.5.2 Local transit vehicles

The response to the amber signal by local transit vehicles depends upon the type of bus stop. Two basic types of stops were felt to be appropriate for simulation: the near-side curb stop and the far-side curb stop.

Aside: There are two other types of stops, namely the mid-block bus stop (a relatively common type) and the mid-street platform. The latter was felt to be too specialized to be considered and the former was felt to be irrelevant, the intention being to examine the behavior of traffic at an intersection, not in the middle of a block.

1) Near-side curb stop

The near-side curb stop is located at the curb on the approach in advance of the intersection proper. Where bus volumes are appreciable, this type of stop has a greater adverse effect on intersection capacity than a far-side stop. Both right turns and through movements are affected (Highway Capacity Manual (1965)).

In a simulation model, all buses approaching the intersection would stop to load and unload passengers, regardless of the status of the traffic signal. In fact, buses which stop during an amber or red interval, combine traffic signal delay with delay due to loading and unloading. This improves the overall transit operation.

It is necessary to clarify one particular situation. Suppose, in a simulation model, a bus completed its service operation and was about to depart when the light changed to amber?

In this case, the bus would remain stationed at the stop until the light again changed to green. Such behaviour is felt to be realistic. Buses possess a poor acceleration capability, and thus would be unable to clear the intersection in time.

2) Far-side curb stop

The far-side curb stop is located at the curb immediately beyond the intersection proper on the straight-through exit from the approach under consideration. The adverse effect of this type of stop is due primarily to buses pulling back into the moving traffic lanes, not due to the loss of that space which they occupy during loading. The percentage of turning traffic is also a factor. Where streets are wide and turning movement high, the loss of volume due to the turning movement sometimes nullifies any adverse effect (Highway Capacity Manual (1965)).

If there is a difference in response to the amber signal between local transit buses and passenger cars, then surely it is a function of the difference in their lengths. Deceleration capability is also a factor. However, as pointed out earlier, the deceleration ability of the two is approximately the same.

As noted in section 3.3, an approximate range for passenger car lengths is 15' to 20'. By contrast, all buses are approximately 40' long. The actual effect of this difference is not known. However, if these vehicular lengths are substituted into (10.3.1), and all other factors are assumed equal, the difference in the minimum amber time is approximately .5 seconds.

There is another factor to consider. Buses approaching a far-side curb stop need to clear the intersection before loading and unloading passengers. Thus, there is an urgency associated with crossing the intersection which is not present in the case of passenger cars in a similar situation. In addition, if the intersection is cleared on the amber, the service operation can be conducted while traffic on the approach is stopped for the red signal. If the light as yet has not changed when the bus finishes servicing the stop, the bus can continue without the difficulty of merging into a moving traffic stream.

This motivational factor compensates for the effect of the difference between the length of buses and the length of passenger cars. Therefore, it was decided that table 10.4 could be applied to the case of local transit vehicles approaching a far-side curb stop.

CHAPTER ELEVEN

TURNING PERFORMANCE

11.1 Introduction

Here, the maximum speeds associated with turning maneuvers for both cars and trucks are examined. Values used by several researchers are presented and an example of the type of conflict which can occur between intersection design and vehicular operating characteristics is illustrated. In section 11.3, part of the logic involved in simulating turning maneuvers is briefly outlined.

Unlike many of the variables and parameters discussed in preceding chapters, the values of the turning speeds decided herein are not explicitly required for determining relationships in other chapters. However, chapter 11 is important because, along with the discussion concerning left turn gap/lag acceptance (chapter 9), it forms a necessary component in the treatment of the problems and characteristics associated with turning maneuvers.

11.2 Discussion

In his simulation model, Gerlough (1967) supplied the turning radii as input data. Using the centrifugal force equilibrium formula, he computed the maximum speeds associated with right turns, free left turns, and delayed left turns.

Thus,

$$v = \sqrt{rg\ (e+f)} \qquad (11.2.1)$$

where

v= maximum turning speed, ft/sec;

r= radius of horizontal curve, ft;

g= acceleration of gravity, ft / sec^2;

e= superelevation in feet per foot of horizontal width;

f= coefficient of side friction.

Setting g= 32.2'/sec^2, e=o, and f=.3*

$$v = \sqrt{9.66r} \qquad (11.2.2)$$

However, by relying on this functional relationship, Gerlough created an unnecessary complication.

In one particular simulation run, the free left turn radius was 50', while the right turn and delayed left turn radii were 30'. The maximum turning speeds associated with these radii were 22'/sec. and 17'/sec. respectively. The problem is that trucks, because of their length, require a wider turning radius than passenger cars. A medium-sized truck 30' long and making a 90° turn, requires a turning radius of 50' (Traffic Engineering Handbook (1965)).

*The 95 percentile turning speed is associated with f=.3 for medium low-speed turns (American Association of State Highway Officials (1954)).

Theoretically, then, the trucks in this simulation run can execute only free left turns.

To allow truck traffic unrestricted turning movement, the minimum turning radius has to be 50'. Right turns require a shorter turning radius than left turns. Thus, the left turn radius had to be greater than 50' - say 70'. Substituting r = 70' into (11.2.2), results in a maximum turning speed of 26'/sec or 18 mph. Such a speed is both excessive and unrealistic.

The solution is simply to assign maximum vehicular speeds to the different turning maneuvers. This does not imply that maximum turning speed is not associated with the radius of the turn; but that the relationship is tacit in nature, not a controlling factor. The intention is not to design an intersection, but rather to simulate traffic behavior at one.

In fact, Gerlough's approach is unique. Lewis and Michael (1963) assumed equal turning speeds for both left and right turn maneuvers. The maximum velocity at the turn point - a point approximately midway through the turn - was 15'/sec or 10.2 mph. In both the simulations conducted by Kell (1962, 1963), vehicles turning right slowed to 8 mph. while those turning left slowed to 9 mph.

The Traffic Engineering Handbook (1965) supports these figures. Here, the speed of turning maneuvers, under nearly all conditions of turning radii and vehicular types, is listed as 10 mph.

The only other available research is that conducted by George (1952). He examined the general characteristics of left turning passenger vehicles at rural and urban stop sign controlled intersections. Trucks were not included in this study.

For the turning maneuvers discussed herein, however George's work is not applicable. In his study, the speed of the turning movements varied from zero, at the start of the turn, to a desired speed, at the end of the turn. This desired or tangent speed was measured at a point where curvilinear motion became linear motion. Because these turning vehicles were accelerating to merge with cross-street traffic, the speeds recorded were higher than speeds observed at signalized intersections. The mean target speed was about 16 mph.

For purposes of simulation, it was decided that the maximum speed appropriate for right turning and delayed left turning passenger cars is 9 mph; while the maximum speed for free left turning cars is 10 mph.

Because of their greater length and poorer operating characteristics, turning trucks are equivalent to 1½ passenger cars (Greenshields (1947), Webster (1966), Wohl and Martin (1967)). Thus, the maximum speed for right turning and delayed left turning trucks is 6 mph., while the maximum speed for free left turns is 7 mph. These values were felt to be suitable for simulation.

11.3 Logic

In a simulation model, turning vehicles undertake a fixed turning schedule at the intersection. Each turning vehicle approaching the intersection is examined. If its current velocity is greater than the maximum turning velocity, the vehicle begins slowing to the maximum turning velocity, until reaching the start turn position. Maximum turning speed is maintained throughout the turn.

The foregoing does not imply that all vehicles make their turns at maximum speed. Some are affected by other queued vehicles such that they turn at speeds considerably lower. However, in no case, are these maximum turning speeds exceeded.

CHAPTER TWELVE

TRAFFIC SIGNALIZATION

12.1 Introduction

The primary function of traffic signals is to alternate the right-of-way to various traffic movements at an intersection. In so doing, they exert a profound influence on traffic flow and operate either to the advantage or disadvantage of the vehicles they control. A properly designed installation provides for orderly movement of traffic and increases the traffic handling capacity of the intersection (Traffic Engineering Handbook (1965)).

In this chapter, traffic signalization is examined for its relationship to simulation. The components of a traffic signal system are deterministic in nature. Functional relationships between the input variables (approach volume, speed, etc.) and the signal times are constant. In other words, one and only one signal time will occur for a particular set of values of input variables. Because of this, signal timing is the easiest element to control in a simulation model. In addition, the profound influence of signal timing on the behavior of vehicles within the intersection makes it important that signal times be optimally selected.

The objective of signal timing is to alternate the right-of-way between traffic streams so as to minimize both average delay to all vehicles, and total delay to any single group of vehicles. While long cycles accommodate more vehicles per hour, cycles longer than necessary produce higher average delays (Kell (1963), Webster (1966)).

Stopped vehicles at a traffic light do not instantaneously enter the intersection at minimum headways when the light changes, but rather incur a starting delay. In addition, the clearance interval of each phase is not fully utilized. Thus, over any time period, a short cycle with a greater number of starting delays and clearance intervals produces more lost time and accommodates fewer vehicles.

Conversely, a long cycle is efficient only when there is a constant demand during the entire green period on each approach. As cycle and green periods lengthen, the probability of a demand during the latter portions of the green decreases. Vehicles waiting on the corresponding red phase now incur greater delay. In all, a higher total delay results. Thus, an optimally timed signal implies the shortest cycle and phase time which provides orderly traffic movement and still occasions the lowest average delay.

There is a distinction to be made between analytic studies and simulation studies. In analysis, suitable algebraic equations are used to explain particular relationships. Variables are related in precise mathematical terms from which values may be determined exactly. However, most situations in the real world are very complicated and cannot be represented exactly by tractable mathematical models. The procedure then is to apply mathematics to ideal situations to obtain approximations having only the important features of the real phenomena. Providing the approximations are verified by experiment and are close to the real phenomena, they have important applications to simulation models. Some of these applications will now be discussed.

Analytic formulations are composed of combinations of mechanisms. The relative influence of a particular mechanism on an analytic form indicates the importance of the mechanism to the phenomenon which the form characterizes. Therefore, analytic equations suggest directions towards which simulations should progress.

For example, the nondilemma yellow period determined by (10.3.1) is

$$\tau \text{ min} = \delta + \frac{Vo}{2d} + \frac{W+L}{Vo}$$

If the perception-reaction time δ is incremented by an amount K, the difference to τ min is K sec. However, if vehicular length L is incremented by this same amount, then the difference to τ min is K/Vo, which is less than K for Vo$>$1.

Therefore, there is an indication that, in studying driver reaction to the amber signal in a simulation model, an understanding of the effect of driver perception-reaction time is more important than examining the effect of vehicular length.

Analytic studies are important to simulation because they suggest optimum control policies. In order to optimize a traffic system in a simulation model, a criterion must be established to determine the quality of the operation. The measure of effectiveness inherent in the criterion of optimization must be expressed as a function of the mechanisms entering into the problem. Because the different timing procedures were developed with the intent of optimizing particular service conditions (the degree of congestion or delay, the length of the queue, the probability of a vehicle being stopped, the average number of vehicles stopped, etc.), and because these service conditions are expressed as functions of particular mechanisms, the timing procedures are important for suggesting the types of statistics necessary for determining an optimal policy of signalization.

Besides suggesting the mechanisms relevant to a method of traffic signalization, analytic studies suggest the type of data to collect and analyze in order to optimize a simulation model of a traffic system. The methods of collecting this data and some of the problems involved in this collection process are examined in section 12.6.

In previous chapters, most of the mechanisms relevant to a policy of optimal traffic signalization have already been presented. In chapter 3, vehicular lengths are examined. Vehicles are not points but physical entities which possess length. As approach volumes near capacity and intersections become congested, the distributions characterizing vehicular behavior change. The distribution which describes vehicular headways now reflects a greater proportion of retarded vehicles. The proportion of the traffic stream in free-moving groups approaches zero. Vehicles are no longer able to change lanes freely nor travel at their desired speed. The traffic stream becomes simply a line of vehicles which must be directed through the intersection. Any traffic timing procedure should account for this mechanism.

The acceleration and deceleration of cars, trucks, and buses is examined in chapter 8. Both these mechanisms are necessary components of an appropriate timing procedure. When a traffic light changes to green, waiting vehicles incur a starting delay which is, in part, a function of the acceleration of the lead vehicle. In addition, a portion of the amber is unused due to the first vehicle in line stopping before the red signal. This lost time is a function of the deceleration ability of the stopping vehicle. Any traffic timing procedure must account for these lost times.

The reaction-decision making time is another important mechanism. For a vehicle first in line in a queue, the reaction-decision time is the time between the beginning of the green signal and the beginning of the vehicle's forward motion. For a vehicle approaching an intersection, it is the time taken by the driver to react to the amber signal in deciding to stop. Therefore, as a relevant component of the lost time per cycle, it is a necessary mechanism in signal timing.

The speed of vehicles approaching the intersection is discussed in chapter 7. Vehicles approaching an intersection at high speeds must be given sufficient time to stop on seeing the amber signal. If the amber is too short for vehicles to stop using normal rates of deceleration, then these vehicles will attempt to cross the intersection before the red. This is a dangerous situation and one to be avoided if possible.

The width of the intersection is also important for determining the length of the amber. The amber light must be long enough to allow vehicles which cannot stop, the time necessary to continue at uniform speed and cross the intersection before the light turns red. This time is a function of both approach speed and intersection width.

When a signal turns green, the first vehicle incurs a starting delay and is then followed by the rest of the queued vehicles moving at nearly equal time headways. This pattern is complicated by commercial vehicles. Because of their greater length and generally lower acceleration ability, commercial vehicles require more time to clear an intersection than do passenger cars. Some adjustment must be incorporated into any timing procedure to account for this.

The characteristics of turning vehicles are discussed in chapters 9 and 11. Turning vehicles require more time to clear an intersection than straight through movements, particularly when blocked by other vehicles in their path. They, in turn, block other vehicles in the same lane. When traffic is light, there may be little blocking but when it is heavy, the blocking may continue to the end of the green phase, forcing turning vehicles to proceed on the yellow or even into the beginning of the red phase. Therefore, any timing procedure must provide some adjustment for the effect of turning movements.

Finally, it is necessary for any traffic signalization policy to account for the pattern of vehicular arrivals at the intersection. This mechanism is thoroughly examined in chapter 5. Using average arrival rates is not satisfactory since, in some cycles, more than the average number of vehicles arrive. These vehicles would then be unable to pass through the intersection. Congestion would result.

While analytic studies have contributed much towards a better understanding of traffic phenomena, there is a need to supplement these studies using simulation. As explained earlier, analysis describes only the gross features of a system, based on ideal situations and a particular set of input mechanisms. Mechanisms which are easily measured are usually considered the "independent variables" while those which are difficult to measure are considered the "dependent" ones. This is an oversimplification of the problem as, in fact, there is no single "dependent variable". Mechanisms react in a complex fashion and a change in one results in a combination of changes in the others. These changes, in turn, further affect the behavior of the first.

In traffic control, seldom can the mechanisms of a system be represented by a single deterministic value. Most mechanisms must be represented by a stochastic variable which has a specified distribution of values if the results are to be realistic. In simulation, random sampling is applied in finding the occurance of events. Realism is provided by allowing probabilistic events to occur randomly (Drew (1968)). These random characteristics have a considerable effect on the micro-behavior of traffic and through simulation, we are able to examine this behavior. Traffic can be represented in the quantities desired and of the particular characteristics desired. Information on the probability distribution of an outcome can be determined, rather than just the measures of central tendency and dispersion as obtained in analysis.

In section 12.3, a brief description of major analytic traffic signalization investigations is presented. For each, particular attention is given to the relevant mechanisms used and the performance criteria optimized. Only fixed time signalization is discussed.*

*Fixed time signals direct traffic in accordance with a single predetermined time schedule, or series of such schedules (Traffic Engineering Handbook (1965)).

In developing a simulation model of a signalized intersection, it is important that a signalization procedure be used which provides for orderly traffic movement and minimizes delay. In the field, signal times are not set arbitrarily but are set in accordance with an established signalization procedure. In the same way, a suitable signalization procedure should be incorporated into a simulation model. Apart from providing a degree of realism by using such a procedure, the fact is that it is not sufficient to set cycle times and phase times arbitrarily. Signal times must change in accordance with approach volumes and the operating characteristics of traffic. This can only be accomplished by an appropriate timing procedure. Otherwise, congestion will occur, and the inter-relationships examined will be severly distorted.* To this end, a comparative evaluation of the major traffic signalization methods is conducted in section 12.4 to determine a procedure suitable for simulation.

While section 12.3 is devoted to examining only fixed time signalization methods, there is another important signalization policy which is highly suitable for examination by simulation. This signalization policy is known as traffic actuated. In section 12.5, we discuss traffic actuated signals and their application to simulation.

*This is assuming that one is not examining the effect of traffic signals on a particular phenomenon.

12.2 Definitions (Traffic Engineering Handbook (1965), Manual of Uniform Traffic Control Devices for Canada (1966)):

Traffic control signal (traffic signal)--any power operated traffic control device, except a sign, by which traffic is warned or directed to take some specific action.

Right-of-way--the privilege of the immediate use of the roadway.

Phase--a part of the total time cycle allocated to any traffic movement, or combination of traffic movements, receiving the right-of-way during one or more intervals.

Interval--any one of several divisions of the time cycle during which signal indications do not change.

Cycle length (time cycle)--the total time necessary for one complete sequence of signal indications.

Clearance interval--the time of display of the signal indication following the right-of-way interval.

Green interval--the portion of a signal phase during which the green indication is displayed.

12.3 Major Analytic Investigations

12.3.1 Preliminary remarks

While a number of timing methods have been developed, too often, they are presented in "cook book" form and there is little opportunity to analyze the assumptions from which they were developed. For example, Davidson (1961) presents a set of cycle capacity probability design curves for use in traffic signal timing. No related theory accompanies these curves.

There are, however, several major analytic investigations which are well documented and thoroughly explained. These will now be examined.

The purpose of this section is to show what mechanisms are accounted for and to find what criteria for optimization were used (and hence, what data would need to be collected).

12.3.2 Webster's method

Dealing with the random nature of traffic flow, Webster (1966) determined equations for the average delay per vehicle (and total intersection delay), the approximate cycle length Co which minimized total intersection delay, average queue length, and the probability of queues reaching some maximum value.

Subsequently, Bone, Martin, and Harvey (1962)* developed

*Webster first published his results in 1958. A more comprehensive publication followed in 1966. It is this subsequent publication which is referenced herein.

a general approach to signal design. This included: 1) determination of the yellow period, 2) consideration of vehicle performance, 3) apportionment of green time, 4) determination of delay, 5) determination of queue lengths, 6) determination of the probability that vehicles cleared the intersection upon first encountering the green phase.

These will now be discussed briefly. The length of the yellow period was determined from (10.3.1). As such, most of the mechanisms important in the design of a signalized system are herein incorporated, namely speed, vehicular length, intersection width, perception-reaction time, and vehicular deceleration.

In order to account for the presence of commercial vehicles and turning movements, actual vehicular volumes are increased. Volumes are now in terms of "equivalent passenger car units." Admittedly, this is an approximate procedure, but does seem to supply satisfactory results. However, Wohl and Martin (1967) point out that if turning vehicles have unusual characteristics, this should be examined separately.

The green time available for vehicular movement is the cycle time minus the lost time (time lost getting queued vehicles moving and while vehicles clear the intersection). This effective time is distributed among the phases in direct proportion to the volume on each phase. To determine the actual length of the green phase, the lost time is added back.

To examine the performance of vehicles within the intersection, and to evaluate the signal timing, Webster determined expressions for the average delay per vehicle, the queue length at the start of the green, and the length of the queue equaled or exceeded 5% of the time. The queue length is an indication of the delay incurred by vehicles in the queue.

Wohl and Martin (1967) felt that drivers who are delayed more than one cycle become irritated by the delay and inevitably retard other vehicles as well. Thus, the probability of a vehicle entering the intersection on first encountering the green is a measure of the comfort and convenience associated with a signal setting. Webster outlined a method to determine this probability on the assumption of Poisson distributed arrivals.

Aside: It is noted that the assumption of Poisson arrivals does not account for the variability of arrivals with time. As volumes increase, vehicles adopt shorter headways as they overtake slower vehicles. Under this bunching condition, vehicular behavior becomes nonrandom and the assumption of Poisson arrivals will be in error. (This is fully discussed in Chapter 5).

To select a cycle which optimizes vehicular performance, a number of alternative cycles are analyzed. The cycle which optimizes a majority of the performance criteria (average delay per vehicle, maximum queue in 19 out of 20 cycles, average queue, and probability of entering intersection on first green) is the one selected.

Examining Webster's method, one notices that all the mechanisms relevant to an optimum timing procedure are incorporated herein. As well, Webster is particularly conscious of the interreaction between cycle times and performance criteria, seeking to optimize as many of these criteria as possible. While his assumption of random flow can be criticized, Wohl and Martin note that random flow can be assumed for low volumes, although some studies indicate that very high flows, but not medium flows, are also random.

12.3.3 Bellis's method

Bellis (1960) developed an equation for the time for a car to move a given distance from the stop line after a light turns green. Thus,

$$T = PN + \frac{E}{U}\sqrt{(W+L(N-1))(W+L(N-1) + \frac{U^2}{4})} \qquad (12.3.1)$$

where

T = time after the beginning of the green signal to arrive at the distance W, sec;

P = reaction-decision making time, sec;

N = Nth vehicle stopped in line;

E = acceleration constant;

U = approach speed, mph;

W = distance from stop line of first car, ft;

L = front-to-front spacing of standing vehicles, ft.

Additionally, Bellis assumed vehicular arrivals to be Poisson distributed. Initially, a cycle length is assumed and the expected number of arrivals for a probability of .95 are computed from

$$.95 = \sum_{X=0}^{N} \frac{e^{-\Theta}\Theta^{x}}{x!} \qquad (12.3.2)$$

where

Θ = average arrivals per cycle

= VC;

V = volume, vehicles/sec;

C = cycle length, sec.

For different average flows and cycle lengths, the maximum number of vehicles expected per cycle can be computed for any desired level of probability.

Bellis accounts for the level of service in two ways. First, he computes the number of vehicles to be cleared during the green for a probability level of .95. (Obviously, he is making an assumption about which level of service is most economical). Secondly, a level of service is implied by setting the green time long enough to clear all expected arrivals. As the level of probability increases, a higher level of service is implied. Though delays and queues are not directly computed, there is some measure of control over the service offered.

It is noticed that Bellis employs nearly all the mechanisms important in the development of a signalized procedure, namely speed, vehicular length (Bellis uses effective vehicular length), intersection width (Bellis calls this the distance from the stop line of first car), reaction-decision time, and the Poisson distribution to describe vehicular arrivals. Supposedly, the acceleration constant E functions much the same as the rate of vehicular acceleration. Unfortunately, Bellis makes no provision for either turning vehicles or commercial vehicles in the traffic stream.

12.3.4 Hewton's method

Hewton reasoned that an interrupted traffic stream suffers a certain lost time composed of: the time lost decelerating to a stop, reaction-decision making time, and the time lost accelerating from a stop to the average speed. Thus,

$$Z = \delta + \frac{Vo}{2aa} + \frac{Vo}{2d} \qquad (12.3.3)$$

where

Z = lost time, sec;

δ = reaction-decision time, sec;

Vo = approach speed, ft/sec;

aa = average acceleration, ft/sec^2;

d = average deceleration, ft/sec^2.

Interrupting the traffic stream n times in a cycle of length C results in a lost time of $\frac{3600\ nZ}{C}$ sec/hr. The time available for traffic movement is 3600 $(1-\frac{nZ}{C})$ and the maximum number of vehicles (equivalent vehicles) passing through the intersection in all directions in one hour is

$$Ve = 3600\ (1-\frac{nZ}{C})/h \qquad (12.3.4)$$

where

Z = lost time, sec;

C = cycle time, sec;

n = number of phases;

h = minimum headway, sec;

Ve = maximum number of equivalent vehicles per cycle.

Rearranging (12.3.4),

$$C = nZ + hM \qquad (12.3.5)$$

where

M = average arrivals per cycle

$= \frac{VeC}{3600}$.

Like Webster (1966) and Bellis (1960), Hewton assumed vehicular arrivals to be Poisson distributed. He thus calculated the expected number of vehicular arrivals, for a probability of .95, during a cycle of length C. Thus,

$$C = nZ + hX \qquad (12.3.6)$$

where X was determined from

$$.95 = \sum_{0}^{X} \frac{e^{-M} M^{X}}{X!} \qquad \text{and} \qquad M = \frac{VeC}{3600} \; *$$

Hewton added that the assumption of Poisson distributed arrivals is applicable only to randomly moving traffic. When free movement is disrupted due to congestion or traffic signals, the above method will be in error.

To determine individual phase times, Hewton used the same conventions as above.

Thus,

$$c = Z + hx$$

where

c = phase time, sec;

x is determined from

$$.95 = \sum_{o}^{x} \frac{e^{-m} m^{x}}{x!}$$

$m = \frac{veC}{3600}$ (average arrivals per cycle on that phase)

and where C = cycle time, sec;

ve = equivalent vehicles arriving during that phase.

Like Webster, Hewton does not explicitly account for the influence of commercial vehicles or turning vehicles but uses, instead, the somewhat artificial form;

$$Ve = (V + \frac{H}{2} + \frac{L}{6})/n \qquad (12.3.7)$$

* Hewton used the following form to determine the minimum headway h:

$h = .7 + 10^{-4} S^{2.3} + \frac{.68L}{S}$ for $10 \leqslant S \leqslant 40$ and S = speed, mph;

L = vehicular length, ft.

where

Ve = equivalent volume, vph;

V = total volume, vph;

H = number of commercial vehicles;

L = number of left turning vehicles;

n = number of lanes.

Equation 12.3.7 assumes traffic is equally distributed between the available lanes. This is probably not the case except at very high volumes.

Hewton also provides an equation for calculating the clearance interval. Basically, his equation is the same as 10.3.1, except that vehicular length is not explicitly included. This is because intersection width is assumed to be the greatest distance a vehicle must cover to be completely free of traffic entering the intersection on the next phase. Implicitly, this includes vehicular length.

Hewton's procedure, thus, incorporates all the mechanisms which are relevant to a signalization method. Unfortunately, he makes no provision for examining any performance criteria. At best, a level of service is implied by establishing the probability of the occurance of a back-up, under the assumption of Poisson distributed arrivals, as .05 (ie. the intersection clears 95 times out of 100).

In all, it would seem that Hewton's primary objective is to maximize the number of vehicles passing through the intersection. In doing so, he is attempting to minimize the number of vehicles delayed in the intersection and to prevent congestion.

12.3.5 Dunne and Potts's method

Dunne and Potts (1965) analyzed the operational characteristics of a signalized intersection for which the control strategy is to switch the lights when the favoured queue empties. A 2-phase traffic light was considered.

At the start of each green, there is a lost time composed of the minimum vehicular headway h and the number of vehicles which first enter the intersection before the entire queue begins to move L. This is followed by an effective green time during which queued vehicles discharge at saturation flow rate. Vehicle arrivals are assumed generated as a binomial process.* Thus, q_i is the probability of one arrival in arm i = 1,2 in each of the intervals (mh, m h+h), m = 0,1,2,...; and the probability of no arrivals is $p_i = 1-q_i$.**

Dunne and Potts developed a bivariate Poisson distribution (in discrete time) to describe vehicular arrivals. In addition, they determined a set of equations characterizing a two-dimensional random walk, the analysis of which yields the operational characteristics of the control of the intersection.

*The idea of vehicular arrivals generated as a binomial process was first presented by Beckman, McGuire, and Winston (1956). In Chapter 5, we present an example of a binomial distribution.

**Only two arms are considered since two lanes usually predominate in determining the light setting.

Providing the intersection is undersaturated, the operation approaches statistical equilibrium and the steady-state probabilities give the phase durations and the mean queue lengths when the light is switched.

Note that $q_i = v_i/V$ where v_i = approach volume on arm i and V = total approach volume on all arms. The phase durations are functions of the lost time and q_i, while the mean queue lengths are functions of q_i and L.

It was noticed that very few of the mechanisms discussed as being relevant to a signal timing procedure are incorporated herein. Dunne and Potts make no provision for vehicular length, intersection width, acceleration, deceleration, approach speed, or perception-reaction time. Additionally, there is no provision for either turning vehicles or commercial vehicles. Dunne and Potts felt that left turning vehicles either filter through on-comming traffic or make their turns during the amber. Right turning cars can turn freely, providing pedestrian traffic is light.

In developing their procedure, Dunne and Potts sought to minimize delay. Phase times were calculated so as to provide maximum efficiency (ie. minimum delay) for constant arrival rates, and to shorten or to lengthen about their average as changes in traffic demand occur.

The only other performance standard which is examined is the mean and the variance of the queues when the light changes. Long queues may block cross traffic at adjacent intersections or interfere with activities at near-by access points. Also, excessive queues may interfere with the proper alignment of vehicles making left or right turns. Therefore, knowledge of queues is an important criterion for evaluating signal timing. Unfortunately, Dunne and Potts do not examine any of the other, equally important, performance standards.

12.4 Application to Simulation

12.4.1 Comparative evaluation

In order to compare the timing procedures presented in section 12.3 to determine a method suitable for simulation, a computer analysis was performed. For each timing method, the same combinations of approach volumes were generated, and cycle times, phase times, and different operating characteristics were calculated.

To maintain control in comparing these different methods, any mechanism common to two or more timing schemes retained a constant value. Thus, whenever possible, the value of a parameter determined in one method was used throughout the other methods as well. Related parameters were altered accordingly, unless a constraint, unique to a particular method was violated.

To determine the length of the amber signal, the following values were used: perception-reaction time δ = 2 sec; approach speed Vo = 44'/sec; vehicular length L = 20'; deceleration d = 20'/sec^2; intersection width W = 60'. Substituting these values into equation 10.3.1, $\tau_{min} \simeq 5$ sec. This value was used throughout the analysis.

Generally speaking, minimum green times are calculated on the basis of two objectives:

a) proportioning the green time among the approaches so that the ratio of capacity to demand on each of the approaches is equal;

b) choosing the green time so that, during peak demand periods, all traffic arriving during the phase is accommodated (ie. severe congestion does not occur).

It was noted earlier the importance of maintaining as short a cycle length as possible. In addition, both Gerlough (1967) and Webster (1966) indicate the adverse effect of traffic flow imbalance. These two criteria were thus included with the criterion of incorporating the relevant mechanisms, in evaluating the results of the analysis.

In analyzing Webster's procedure, it was necessary to use a number of cycle lengths to determine one which optimized a majority of the performance criteria. Cycle lengths from 30 sec. to 140 sec. were analyzed at intervals of 10 sec. Table 12.1 is an example of the type of results obtained.

It was noticed that for cycle lengths less than the "optimum", the average delay per vehicle increases very rapidly and, theoretically becomes infinite as the approach volume approaches saturation flow. For cycle lengths greater than the "optimum", the average delay increases, but at a much slower rate.

The maximum queue length is similar in form to the average delay; increasing rapidly for cycle lengths less than Co, but increasing slowly for cycle lengths longer.

The values for the average queue increase more slowly than those for the maximum queue for cycle times less than Co, but generally have the same form.

Values for the probability of entering the intersection during the first green are slightly different in form. Here, the probabilities increase with increasing cycle times up to a maximum, after which the probability remains constant.

Table 12.1 is typical of the type of results obtained and is in strong agreement with the observations of both Webster and Wohl and Martin.

In analyzing Bellis's method, it was found that, in some cases, the sum of the volumes for the intersecting approaches was so large that no consistent combination of required green times and cycle times could be determined, for a probability level of .95. This occurred when the sum of the volumes on the two approaches exceeded 900 vph. and, as such, no further calculations were conducted. When this occurs in the field, it is necessary to set the signal timing for at least one of the approaches in accordance with a lower level of probability (and assume that a higher level of delay is satisfactory).

Examining the cycle times calculated by the four different timing methods, it was found that cycle times calculated by Bellis's method were greater than those calculated by Webster's procedure. Times calculated by Hewton's technique were shorter than Webster's, while those of Dunne and Potts are the shortest of all.

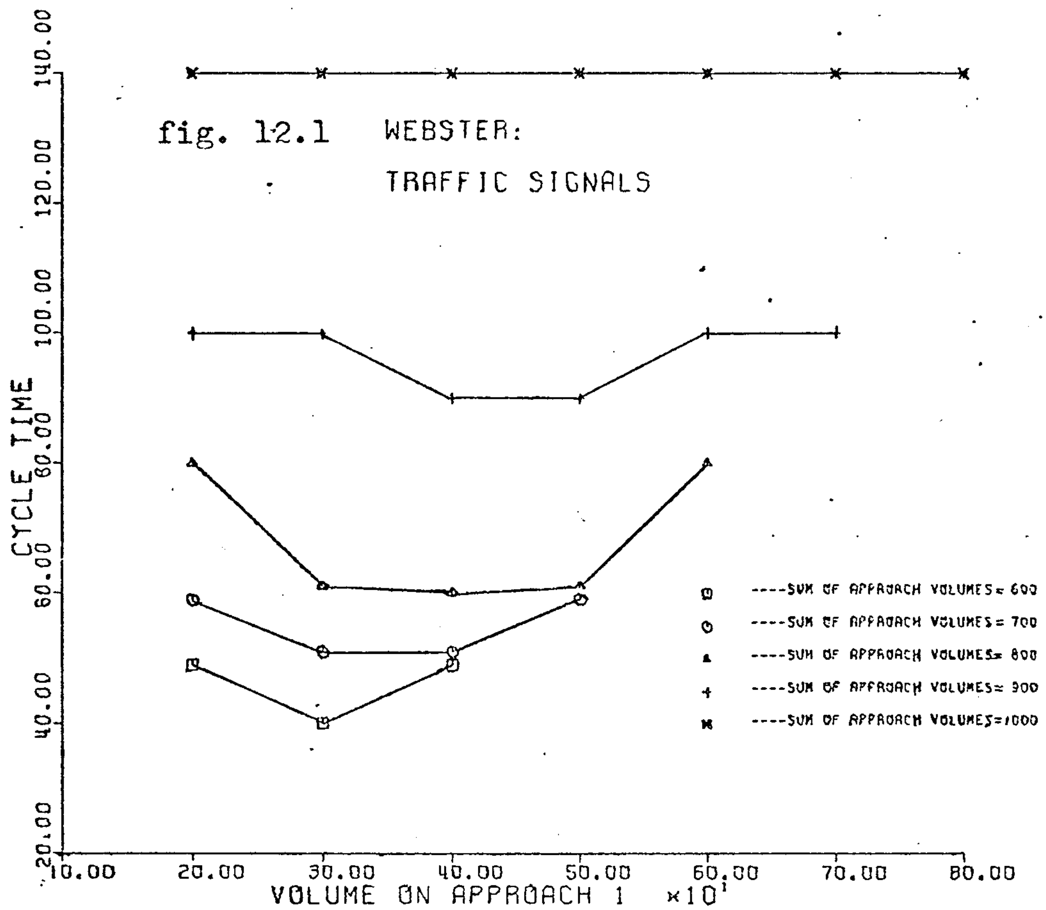
Performance of Design Volume for Different Cycle Times

Cycle Length, sec.	Average Delay Per Vehicle, sec.		Average Queue At Start Of Green		Maximum Queue (19 of 20 cycles).		Prob. of Entering on 1st. Green.	
	App 1	App 2	App 1	App 2	App 1	App 2	App 1	App 2
30	92	39	6	6	25	21	.77	.76
40	36	17	3	3	9	10	.81	.80
50	31	14	3	3	8	9	.85	.84
60	31	13	3	3	8	9	.88	.92
70	32	13	3	3	8	9	.90	.93
80	34	14	3	4	8	9	.92	.94
90	37	14	4	4	8	10	.93	.97
100	39	15	4	4	9	11	.89	.97
110	42	15	4	5	10	11	.91	.99
120	45	16	5	5	10	12	.92	.99
130	47	16	5	5	11	12	.94	.99
140	50	17	6	6	11	13	.95	.99

Design Volume: Approach 1= 200 vph.

Approach 2= 500 vph.

Optimum Cycle Length, C_o = 51 sec.



It should be noted that, using Webster's method, different approach volumes resulted in different cycle times, even though the sum of the approach volumes remained constant. This is in contrast to the other methods wherein cycle time remained constant. Thus, in comparing Webster's procedure to the others, the mean of the cycle times which corresponded to a particular sum of approach volumes was used.

Fig. 12.1 is a graph of cycle times, corresponding to different combinations of approach volumes, calculated using Webster's method. This graph is presented to illustrate the effect of traffic flow imbalance on Webster's timing scheme.

Examining the green times as calculated by the four methods, it was found that the same form that characterized the relationship among the cycle times persists here as well. Thus, green times calculated by Bellis's method are longer than those calculated by Webster. Times calculated by Hewton are shorter than those of Webster, while those of Dunne and Potts are the shortest of all.

12.4.2 Discussion

In all, the procedure developed by Webster is felt to be superior to the other three. Basically, Bellis's method suffers too many critical shortcomings.

First, Bellis's scheme is applicable only for a limited range of approach volumes. When the total intersection volume becomes too large, no cycle time satisfies the required green times.

Second, the calculated cycle times were the highest of all the timing methods. It is important to obtain the shortest cycle possible in order to minimize delay and increase the effectiveness of the traffic signals. This was stressed earlier and is reaffirmed here.

Third, cycle times did not fluctuate with traffic flow imbalance but remained constant for a particular approach volume.*

*In truth, this is an unexpected result since there is little in the mechanics of Bellis's method to indicate that this would occur.

Both the results of our analysis and the work of Gerlough indicate the adverse effect of directional imbalance, and it is important for any timing scheme to account for this.

Finally, though bellis's method does include almost all of the mechanisms considered important in a timing procedure, his method does not provide much control of the level of service.

Some of the criticisms of Bellis's method apply equally to Hewton's procedure. Like Bellis, Hewton does not account fully for imbalanced flow rates. While phase times do fluctuate, cycle times remain unchanged for a constant total approach volume. Also, while Hewton's method does incorporate all the relevant mechanisms, like Bellis, he provides almost no control of the level of service.

The procedure developed by Dunne and Potts if perhaps the worst. Though phase times and cycle times are the shortest, Dunne and Potts employ almost none of the mechanisms relevant to a timing procedure. As well, few performance criteria are examined and there is very little control of the level of service.

Thus, for one reason or another, each of these three methods proves untenable as a suitable timing scheme. By contrast, Webster's method appears to be highly suited. There are several reasons for this:

1) Traffic flow imbalance has been shown to adversely effect intersection performance. Webster's method is the only one which allows for adjustments to both cycle and phase times in order to compensate for this imbalance.

2) While the other timing schemes permit some control of the level of service, the method developed by Webster provides for a much broader and more thorough examination of different performance standards. In addition, Webster incorporates all the mechanisms which were felt to be important in a timing procedure.

In all, much of Webster's work has found a substantial degree of acceptance by others. His techniques were used by Gerlough (1967) to investigate optimum cycle lengths for different signal-phasing schemes, and were thoroughly examined by Wohl and Martin (1967) in their discussion on pretimed signalized intersection performance.

12.5 Other Control Policies

In section 12.3 we presented several major analytic traffic signalization investigations and discussed their mechanisms and their optimization policies. These investigations were fixed time signalization methods. By contrast, some signalization procedures operate in accordance with the varying demands of traffic as registered by the actuation of vehicle detectors. These are called traffic actuated controllers. They provide maximum efficiency (ie. minimize delay) at intersections where traffic arrivals cannot be adequately anticipated in advance with pretimed controllers (Traffic Engineering Handbook (1965)).

For a traffic actuated controller, an initial green time is allocated for each approach. This initial time is then incremented until either all the vehicles on the approach are accommodated (ie. the queue dissipates or no more vehicles arrive at the intersection), or the maximum green time is reached. The signal indication is then changed and the right-of-way is allocated to the other approaches. Thus, cycle times and phase times are continually adjusted to meet the varied requirements of traffic.

Because the mechanisms relevant to an optimization policy for fixed time signals (see section 12.3) are also relevant to the optimization of traffic actuated signals, it is not necessary to reiterate their importance in a model of simulation. However, some additional remarks are, perhaps, in order.

In developing procedures for fixed time signalization, most researchers realized the importance of vehicular arrivals, and tried to account for this by assuming vehicular arrivals to be Poisson distributed. It was also noted that, as volumes increase, bunching occurs. Vehicular behavior becomes nonrandom and the assumption of Poisson arrivals is in error. Because vehicle actuated signals are designed to account for the varying demands of traffic, it is important that the headway distribution in the simulation model reproduce these variations.

In simulating a traffic actuated control system, vehicular operating characteristics acquire considerable importance because of the interaction between the operating characteristics of the traffic stream and the signal timing. Whereas the analytic timing procedures used deterministic values to account for a particular mechanism, traffic actuated controllers should, whenever possible, use probabilistic values to provide realism.

For example, phenomena such as queue-discharge involve mechanisms relevant to fixed time signalization. These include acceleration and deceleration rates, vehicle lengths, and perception-reaction times. However, because vehicle actuated policies are sometimes designed to alternate the right-of-way after the queue dissipates, these mechanisms become very important. If they are chosen deterministically, queue dissipation will be unrealistic. Phase times and delay times will be in error, and this traffic control method will be inaccurately evaluated.

In a model of simulation, mechanisms such as acceleration and deceleration rates, perception-reaction times, etc. can be drawn randomly from a distribution of values. Each vehicle can then be assigned a unique set of values of operating characteristics. Times for queues to dissipate will then be unique, even for two queues equal in length. Such times can then be used to determine the delay to queued vehicles, and thus provide for a more realistic evaluation of the traffic control policy.

Because of the complexity of alternating the right-of-way according to the varying demands of traffic, generated according to a relatively sophisticated distribution, few deterministic models of traffic actuated signalization have been attempted.

On the contrary, traffic actuated systems have, for the most part, been evaluated using simulation.

Perhaps the most extensive evaluation of traffic actuated signalization was conducted by Gerlough (1967). He developed a computer simulation model of an intersection controlled by traffic actuated signals. Three types of control policies were investigated: the basic queue control, the queue length arrival rate control, and the space presence control. (It is noted that the level of service provided by fixed time signalization served as the control factor.) Control policies such as these are important because they suggest the type of data which needs to be gathered for determining an optimum control policy (see section 12.6) and they suggest the criteria which need to be optimized.

According to the basic queue control policy, the duration of the green was dependent upon directly sensed queue lengths at the beginning of each phase. The green time was set equal to the time required on the average to discharge the longest queue at the beginning of the green. This control policy suggests the importance of minimizing queue lengths in order to help minimize delay and to improve the level of service offered.

The queue length arrival rate control policy was similar to the policy of basic queue control, except that now the green time was dependent upon the queue lengths and the upstream arrival rates. This control policy suggests the necessity to account for the effect of upstream arrival rates on the queue of waiting vehicles. Compensating for this effect by appropriately timed phases can aid in discharging the queue of waiting vehicles. This minimizes the time spent by vehicles in the system and improves the level of service offered.

The space presence control policy operated so as to clear vehicles within the intersection proper. The rate of vehicular arrivals was not a direct part of this control policy. In the field, passage detectors would be replaced by presence detectors. This control policy is important because it emphasizes that vehicles need not be queued within the intersection to incur delay. The length of time vehicles remain within the intersection proper is a measure of the delay incurred by these vehicles, and is thus a measure of the level of service provided.

12.6 Data Collection

The majority of the traffic signalization methods examined in section 12.4 do not provide for a sufficient analysis of the level of service they offer. By comparison, Webster's procedure best provides for this type of analysis. In doing so, Webster suggests the type of data necessary to evaluate his timing scheme, using a micro-simulation program.

The problem now is to collect the relevant data. This can be a severe problem. As simulations become larger and more complex, it becomes increasingly difficult to obtain the necessary statistics with which to evaluate the control mechanism. 'Flags' must be set, checked, and then reset to keep track of the behavior of the vehicles within the system. Gerlough (1967) termed this the behavior profile and sought to determine the proportion of time vehicles operate according to the different classifications of behavior. These classifications include: free-behavior, car following, stopping, slowing for a turn, stopping for a left turn, stopped for a left turn, travelling in a delayed left turn, travelling in a free left turn, executing a right turn, and stopped. Using this behavior profile, and tabulating the time spent by each vehicle in each of these behavior patterns, the statistics Webster required could be obtained.

To determine the average queue lengths during each phase, it is necessary first to store the lengths of all the queues occuring during the phase. This can be accomplished using the behavior profile. Vehicles stopped for a turn or simply stopped are considered as members of a queue (although the queue might be just one vehicle in length). The number of successive vehicles with this behavior classification, at any particular time, represent the length of the queue at that time. These queue lengths would be stored and later averaged.

To determine the maximum queue in 19 out of 20 cycles, a similar record of queue lengths would need to be compiled. A counter would record the passage of at least 20 cycles. Then, the lengths of the queues which occurred throughout this period would be scanned to determine the maximum one.

To determine the probability of a vehicle entering the intersection on the first green, it would be necessary to tag each vehicle which approached the intersection. The number of red phases which each vehicle encountered would have to be tabulated. If a vehicle encountered more than one red signal, then this vehicle had not cleared the intersection on the first green. If P is the number of vehicles generated during any time period T, and Q is the number of vehicles encountering more than one red phase, then (P-Q)/P is the probability of a vehicle entering on the first green. Note that a vehicle completing a left turn at the start of the red would not be considered as having encountered another red signal.

To determine the delay incurred by vehicles within the intersection, there is the problem of the definition of the word "delay". Total or mean delay is often used as an appropriate objective measure of the effectiveness of a system; and is considered equal to the actual travel time minus some theoretical undelayed travel time. If an undelayed trip is considered one made at average free-flow approach speed, then the delay of individual vehicles can be in error. Desired free-flow speeds are dispersed around their mean and a vehicle with a desired speed less than the mean would be considered delayed, even if it progressed unhindered through the system.

In a like manner, vehicles with desired speeds greater than the mean might have a negative delay. To overcome this problem, Gerlough generated a target velocity for individual vehicles in order to compute a unique undelayed travel time for each. Furthermore, a vehicle which decelerates to maximum turning velocity before a turn, and then accelerates to target velocity after the turn, is not considered delayed.

While such refinements in the definition of delay may be a matter of hair-splitting, it does illustrate the type of problems which can occur in attempting to measure this statistic.

Of greater concern, perhaps, is the difficulty measuring this form of delay in the field. Such measurements are necessary for evaluating installations in the field, and for calibrating and refining simulation models which attempt to duplicate this behavior.

To overcome this difficulty, Gerlough found performance characteristics which are highly correlated with delay, but more easily measured in the field. To do this, he compiled a list of primary measures believed to be adequate representations of effectiveness, but are difficult to collect. These included a) mean travel time, b) mean system delay, c) mean travel speed.

To these were added a list of secondary measures, which were less representative of over-all effectiveness, but strongly correlated with the primary measures and more easily collected. These included a) maximum individual system delay, b)mean stopped delay per stopped vehicle, c) mean delay in queue, d) mean queue length, e) proportion of vehicles stopped.

Finally, a list of measures of "worst" conditions were added. While not strongly correlated with the primary measures, they were to be used in conjunction with them. These included a) maximum individual stopped delay, b) maximum queue length, c) minimum individual travel speed.

Using a simulation program and statistical analyses, Gerlough sought to determine pairs of performance standards which were highly correlated. In all, he was concerned with finding measurements which were easily gathered in the field and yet would be strongly associated with primary measures of effectiveness not so easily gathered.

Subsequently, additional simulations were performed in order to validate the previous regression models.

The best method of collecting data will of course, be highly dependent upon the way in which vehicles are represented in a simulation model. It is necessary, then, to have a mechanism for manipulating these 'representations' to reflect the desired characteristics of traffic flow and its interactions. Wohl and Martin (1967) presented two ways to represent and manipulate vehicles, namely physical representation and memorandum representation.

With the physical representation, an address is set aside to represent the actual traffic on a roadway. A 1 bit denotes the presence of a vehicle, a 0 bit denotes the space following a vehicle. In addition, the number of 1 bits describes the length of vehicles. The arrangement of the bits indicates the position and length of vehicles at some particular time. The movement of vehicles is from the higher numbered address to the lower numbered one. To accomplish this movement, the bits are multiplies by 2^b where b is the number of spaces the vehicle is to be moved ahead per second; b being related to the speed of the vehicles.

If the vehicles are moving at different speeds, they are grouped into channels (blocks of addresses) so that each channel containes only those vehicles travelling at the same speed. Each channel is multiplied by 2^{b_i} where b_i corresponds to the speed of that group. After, the various channels are added together to compose the entire traffic stream.

With the memorandum notation, the location of the vehicle, its speed, its length, its entry time, etc. are described in code. Each vehicle's characteristics are identifiable as it moves through the system. This makes it possible to compute delays to individual vehicles. The distance along the roadway is quantified by using a unit block. This block is one lane wide and has a length equivalent to some fractional part of the length of an average vehicle. A vehicle occupies only a limited number of discrete positions. Vehicles are advanced by multiplying their speed by the time increment and adding the product to their position. The record of the vehicle's position is then updated.

There are, of course, many programming techniques available for data manipulation. Techniques such as push down-pop up stacks, pointers, list structures, circular arrays, etc.

allow storage to be conserved and execution time to be reduced. Their scope is limited only by the skill and ingenuity of the programmer. In addition, languages have been developed specifically for purposes of simulation. Most notable of these languages is GPSS-General Purpose Simulation System. GPSS is a computer program for conducting evaluations and experiments of systems, methods, processes, and designs. A description of such languages lies outside the scope of this thesis.

CHAPTER THIRTEEN

CONCLUSION

In this thesis, an attempt is made to examine some of the more important mechanisms of a simulation model which describes the flow of traffic at a signalized intersection in the central business district. Some of these mechanisms are strictly quantitative descriptions of vehicular operating characteristics. Treatment is limited to a survey of the literature published to date, and to an analysis of the values cited in this literature for their applicability to a simulation model. Examples of this are; chapter 3 wherein vehicular lengths are examined, and chapter 8 wherein the acceleration and deceleration capabilities of vehicles are discussed. In each of these chapters, attention is given to the three primary components of the traffic stream, namely cars, trucks, and local transit vehicles.

While the determination of appropriate values of vehicular operating characteristics should have been a relatively simple task, it was found that very little had been published concerning these mechanisms. This was especially true in the case of local transit vehicles whose operating characteristics are, for the most part, poorly defined.

A complete definition of the range of values of these mechanisms as observed in everyday traffic conditions would greatly benefit the state-of-the-art of many behavior phenomena. An example of this is the left turn gap/lag acceptance phenomenon. Values of the critical gap/lag surely must be a function of the maximum rate of acceleration, driver perception-reaction time, and approach speed. However, such a formulation has never been presented. Perhaps this is due to the fact that the values of these operating characteristics are not well defined.

It seems, therefore, that there is an urgent need for a clearer definition of the range of different vehicular operating characteristics, under a variety of traffic conditions, in order to better explain behavior phenomena.

Because of the importance of the distribution of headways in predicting arrival rates and in timing traffic signals, a considerable quantity of literature has been published on this subject. As a result, an extensive examination of the distribution of headways is conducted in chapter 5. The composite negative exponential distribution, suitable for describing the distribution of headways of cars and trucks, and the shifted negative exponential distribution, suitable for describing the distribution of headways between local transit vehicles, are selected herein. For each, a method is outlined in chapter 6 whereby headways that fit these two distributions may be randomly generated. In order to properly merge these two sets of generated headways into a single traffic stream, an algorithm exploiting the mixture distribution is presented.

While a considerable portion of this thesis is devoted to an examination of vehicular operating characteristics, an understanding of vehicular behavior within the intersection proper is also important. Two of the more important behavioral phenomena are examined in chapters 9 and 10, namely left turn gap/lag acceptance and driver response to the amber signal. Driver behavior is implicitly dependent upon vehicular operating characteristics, but other factors are involved also. These include drivers' emotions and motivation. Understandably, it is difficult to determine functional relationships which characterize driver behavior. However, if simulation models are to be successfully constructed which are capable of realistically modelling traffic movement, this is a necessary development.

In chapter 12, traffic signalization is examined for its relationship to simulation. Traffic timing methods are dependent upon a variety of vehicular operating characteristics and are responsible for a number of vehavior phenomena. Because of the complexity of this inter-relationship, it is difficult to evaluate the different timing methods. Simulation can be used to determine the importance of the different mechanisms comprising a control method, and to analyze the combined effect of these mechanisms to the various behavior phenomena; the intention being to develop a viable timing procedure capable of optimizing traffic performance. A development such as this is urgently needed in the field of traffic control and is recommended herein as a fruitful area for future research.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. A Policy on Geometric Design of Rural Highways, American Association of State Highway Officials, Washington, D.C., 1954.

2. Athol, Patrick, "Interdependence of Certain Operational Characteristics Within a Moving Traffic Stream.", H R Record No. 72, pp.58-87 (1965).

3. Athol, Patrick, "Headway Groupings.", H R Record No. 72, pp.137-155 (1965).

4. Beakey, J., "Acceleration and Deceleration Characteristics of Private Passenger Vehicles.", H R B Proc., 18: 81-89, pt. I (1938).

5. Beckman, M., McGuire, C.B., and Winsten, C.B., "Studies in the Economics of Transportation.", Yale University Press, New Haven, Connecticut, 1956.

6. Bellis, W.R., "Capacity of Traffic Signals and Traffic Signal Timing.", H R B Bull.271, pp. 45-67 (1960).

7. Betz, M.J., "Driver Characteristics at Intersections.", H R Record No. 195, pp. 34-51 (1967).

8. Bissell, H.H., "Traffic Gap Acceptance from a Stop Sign.", Unpublished Graduate Research Report, Institute of Transportation and Traffic Engineering, University of California, Berkeley (1960).

9. Bone, A.J., Martin, B.V., and Harvey, L.N., "The Selection of a Cycle Length for Fixed Time Traffic Signals.", M.I.T. Dept of Civil Engineering, Research Report R 62-37, Cambridge, Mass., 1962.

10. Bottger, R., "Possibilities of Simulating Complicated Road Traffic Processes.", Central Laboratories, Munich, pp. 333-341 (1966).

11. Buhr, J.H., Meserole, T.C., and Drew, D.R., "A Digital Simulation of a Section of Freeway with Entrance and Exit Ramps.", H R Record No. 230, pp. 15-31 (1968).

12. Clark, T.R., "The Left Turn, A Criteria (sic) for Intersection Control.", Thesis Manuscript, Bureau of Highway Traffic, Yale Univ., May 1960.

13. Cleveland, D.E., and Capelle, D.G., "Queueing Theory Approaches.", H R B Special Report 79, pp. 49-96 (1964).

14. Dart, O.K., "Left-turn Characteristics at Signalized Intersections on Four-lane Arterial Streets.", H R Record No. 230, pp. 45-59 (1968).

15. Davidson, B.M., "Traffic Signal Timing Utilizing Probability Curves.", Traffic Engineering, Institute of Traffic Engineers, Vol. 32, No. 2, November, 1961.

16. Dawson, R.F., and Chimini, L.A., "The Hyperlang Probability Distribution - A Generalized Traffic Headway Model.", H R Record No. 230, pp. 1-14 (1960).

17. Deen, L.B., "Acceleration Lane Lengths for Heavy Commercial Vehicles.", Traffic Engineering, Institute of Traffic Engineers, Vol. 27, No. 5, February, 1957.

18. Drew, D.R., Traffic Flow: Theory and Control, McGraw-Hill, 1968.

19. Drew, D.R., and Pinnell, C., "A Study of Peaking Characteristics of Signalized Urban Intersections as Related to Capacity and Design.", H R B Bull. 352, pp. 1-54 (1962).

20. Dunne, M.C., and Potts, R.B., "Analysis of a Computer Control of an Isolated Intersection.", Paper presented at Third Internat. Symposium on Theory of Traffic Flow, New York (June 1965).

21. Forbes, T.W., "Human Factor Considerations in Traffic Flow Theory.", H R Record No. 15, pp. 60-66 (1963).

22. Gazis, D.C., Herman, R., and Maradudin, A.A., The Problem of the Amber Signal Light in Traffic Flow.", Operations Research, 8: 112-132 (1960).

23. George, L.E., "Characteristics of Left-Turning Passenger Vehicles.", H R B Proc. 31, pp. 374-385 (1952).

24. Gerlough, D.L., "Use of Poisson Distribution in Highway Traffic.", Poisson and Traffic, The Eno Foundation for Highway Traffic Control, pp. 1-58 (1955).

25. Gerlough, D.L., "Simulation of Traffic Flow.", H R B Special Report 79, pp. 97-118 (1964).

26. Gerlough, D.L., and Wagner, F.A., "Improved Criteria for Traffic Signals at Individual Intersections.", Nat. Cooperative Highway Research Program Report., Vol. 32 (1967).

27. Glickstein, A., Findley, L.D., and Levy, S.L., "Application of Computer Simulation Techniques to Interchange Design Problems.", H R B Bull. 291, pp. 139-162 (1961).

28. Greenshields, B.D., Shapiro, D., and Eriksen, E.L., Traffic Performance at Urban Street Intersections. Tech. Rpt. No. 1, Yale Bureau of Highway Traffic (1947).

29. Herman, R., and Weiss, G., "Comments on the Highway-Crossing Problem.", Operations Res., Vol. 9, pp. 828-840 (1961).

30. Hewton, J.L., "Traffic Signals.", Traffic Training Course of the Ontario Traffic Conference.

31. Highway Capacity Manual, H R B Special Report 87, 1965.

32. Institute of Traffic Engineering. Traffic Engineering Handbook 1965.

33. Kaiser, F.J., "Left Turn Gap Acceptance at an Unsignalized Intersection.", Thesis manuscript, Bureau of Highway Traffic, Yale Univ., May 1951.

34. Kell, J.H., "Analyzing Vehicular Delay at Intersections Through Simulation.", H R B Bull. 356, pp. 28-39 (1962).

35. Kell, J.H., "Intersection Delay Obtained by Simulating Traffic on a Computer.", H R Record No. 15, pp. 73-97 (1963).

36. Kell, J.H., "Results of Computer Simulation Studies as Related to Traffic Signal Operation.", Proceedings, Institute of Traffic Engineering, pp. 70-107 (1963).

37. Kell, J.H., "Simulation of the Intersection.", Conference on Traffic Surveillance, Simulation and Control, Wash., D.C., pp. 127-143 (1964).

38. Knuth, D.E., The Art of Computer Programming, Vol. 2, Addison-Wesley Publishing Company, 1969.

39. Lewis, R.M., "A Proposed Headway Distribution for Traffic Simulation Studies.", Traffic Engineering, Institute of Traffic Engineers, Vol.33, No. 5, February, 1963.

40. Lewis, R.M., and Michael, H.L., "Simulation of Traffic Flow to Obtain Volume Warrants for Intersection Control.", H R Record No. 15, pp. 1-43 (1963).

41. Manual of Uniform Traffic Control Devices for Canada. Canadian Good Roads Association (1966).

42. Massey, F.J. Jr., "The Kolmogorov-Smirnov Test for Goodness of Fit.", Jour. Amer. Statistical Assoc., 46:68 (1951).

43. May, A.D., "Gap Availability Studies.", H R Record No. 72, pp. 101-136 (1965).

44. Miller, A.J., "The Capacity of Signalized Intersections in Australia.", Australian Road Research Board, Bull. 3 (1968).

45. Noblitt, J.T., "A Study of Gaps Associated with a Left-turning Truck.", Thesis manuscript, Yale Univ., May 1959.

46. Normann, O.K., "Variations in Flow at Intersections as Related to Size of City, Type of Facility, and Capacity Utilization.", H R B Bull. 352, pp. 55-99 (1962).

47. Olson, P.L., and Rothery, R., "Driver Response to Amber Phase of Traffic Signals.", H R B Bull. 330, pp. 40-51 (1962).

48. Ostle, B., Statistics in Research, Iowa State Univ. Press, 1963.

49. Rainville, W.S., Wolfgang, S.H., Hyde, D.C., and Strickland, R.I., "Preliminary Progress Report of Transit Subcommittee, Committee on Highway Capacity.", H R B Proc. 40, pp. 523-540 (1961).

50. Reilly, E.F., and Seifert, J., "Capacity of Signalized Intersections.", H R Record No. 321, pp. 1-15 (1970).

51. Saal, C.C., and Petring. W.F., "Braking Performance of Motor Vehicles.", Public Roads, June, 1957.

52. Schuhl, A., "The Probability Theory Applied to Distribution of Vehicles on Two-Lane Highways.", Poisson and Traffic, The Eno Foundation for Highway Traffic Control, pp. 59-75 (1955).

53. Shumate, R.P., and Dirksen, J.R., "A Simulation System for Study of Traffic Flow Behavior.", H R Record No. 72, pp. 19-39 (1965).

54. Southwood, L.R.E., Ecological Methods, Methuen & Co. Ltd., 1966.

55. Wagner, F.A., and May, A.D., "Volume and Speed Characteristics at Seven Study Locations.", H R B Bull. 281, pp. 48-67 (1961).

56. Wagner, F.A., Gerlough, D.L., and Barnes, F.C., "Improved Criteria for Traffic Signals on Urban Arterials.", Nat. Cooperative Highway Research Program Report., Vol.73 (1969).

57. Webster, F., "Progress of Work on a New Type of Controller for Traffic Signals on High Speed Roads.", Road Research Laboratory, Research Note, R H/3634 (Dec. 1959).

58. Webster, F.A., Traffic Signals, Road Res. Tech. Paper 56, Road Research Laboratory, London, 1966.

59. Western Flyer Coach Ltd., Metropolitan Corporation of Greater Winnipeg - Private Communications.

60. Wohl, M., and Martin, B.V., Traffic System Analysis for Engineers and Planners, McGraw-Hill, 1967.