HIGHWAY RESEARCH BOARD
Bulletin 356

Theory of Traffic Flow

National Academy of Sciences—
National Research Council

publication 1051

HIGHWAY RESEARCH BOARD

Officers and Members of the Executive Committee

1962

NRC HIGHWAY RESEARCH BOARD

Bulletin 356

Theory of Traffic Flow

Presented at the

41st ANNUAL MEETING

January 8-12, 1962

National Academy of Sciences—
National Research Council
Washington, D.C.
1962

Department of Traffic and Operations

Fred W. Hurd, Chairman
Director, Bureau of Highway Traffic
Yale University, New Haven, Connecticut

COMMITTEE ON THEORY OF TRAFFIC FLOW

Daniel L. Gerlough, Chairman
Head, Automobile Traffic Control Section
Thompson Ramo Wooldridge, Inc.
Canoga Park, California

John L. Barker, General Manager, Automatic Signal Division, Eastern Industries, Inc., East Norwalk, Connecticut
Martin J., Beckmann, Department of Economics, Brown University, Providence, Rhode Island
J. Douglas Carroll, Jr., Director, Chicago Area Transportation Study, Chicago, Illinois
A. Charnes, Technological Institute, Northwestern University, Evanston, Illinois
Donald E. Cleveland, Assistant Research Engineer, Texas Transportation Institute, Texas A & M College, College Station
Theodore W. Forbes, Department of Psychology and Engineering Research, Michigan State University, East Lansing
Herbert P. Galliher, Assistant Director, Operations Research, Massachusetts Institute of Technology, Cambridge
Bruce D. Greenshields, Assistant Director, Transportation Institute, University of Michigan, Ann Arbor
Frank A. Haight, Associate Research Mathematician, Institute of Transportation and Traffic Engineering, University of California, Los Angeles
Robert Herman, Research Laboratories, General Motors Corporation, Warren, Michigan
A. W. Jones, Bell Telephone Laboratories, Holmdel, New Jersey
James H. Kell, Assistant Research Engineer, Institute of Transportation and Traffic Engineering, University of California, Berkeley
Sheldon L. Levy, Director, Mathematics and Physics Division, Midwest Research Institute, Kansas City, Missouri
E. W. Montroll, IBM Research Center, Yorktown Heights, New York
Karl Moskowitz, Assistant Traffic Engineer, California Division of Highways, Sacramento
Joseph C. Oppenlander, Assistant Professor, School of Civil Engineering, Purdue University, West Lafayette, Indiana
Carlton C. Robinson, Director, Traffic Engineering Division, Automotive Safety Foundation, Washington, D. C.
Martin C. Stark, Operations Research Analyst, Data Processing Systems Division, National Bureau of Standards, Washington, D. C.
Asriel, Taragin, Chief, Traffic Performance Branch, U. S. Bureau of Public Roads, Washington, D. C.
William P. Walker, Chief, Geometric Standards Branch, Highway Standards and Design Division. U. S. Bureau of Public Roads, Washington, D. C.
Albert G. Wilson, Senior Staff—Planetary Sciences, The RAND Corporation, Santa Monica, California
Martin Wohl, Transportation Consultant to Assistant Secretary for Science and Technology, U. S. Department of Commerce, Washington, D. C.

Contents

Some Mathematical Aspects of The Problem of Merging

FRANK A. HAIGHT, E. FARNSWORTH BISBEE, and CHARLES WOJCIK, Respectively, Associate Research Mathematician, Graduate Research Engineer, and Associate Research Engineer, Institute of Transportation and Traffic Engineering, University of California, Los Angeles

• AS ROADS and highways become capable of carrying higher and higher traffic volumes, the perturbations introduced by vehicles traveling at speeds or in paths that differ substantially from the norm become increasingly harmful to safe and efficient operation of the road network. Some of this individual variation is undoubtedly fortuitous and can be removed, or at least diminished, by sensible efforts to educate and control drivers.

Even in the best of circumstances, however, there remains the necessity for accelerating, decelerating, weaving, and merging; namely, the need of each car to enter the system and leave the system where it wishes. Not only is this each driver's prerogative, but it is also one that in many cases he exercises without specific traffic control.

Perhaps the most important example of such a situation is the freeway on-ramp and acceleration lane. At these points, which must be provided fairly frequently in urban areas, the smooth flow of traffic is perpetually harassed by new arrivals.

Although it does not seem practical at the moment to imagine an automatic merging control device having the ability to synchronize effectively the multitude of individual merges that occur in a day, this does not mean that the traffic engineer need go to the other extreme and abandon any idea of controlling the merging process.

Indeed, the literature contains ample evidence that location and design of on-ramps and acceleration lanes are closely connected with the influence they exert on traffic stability. If the exact nature of this influence is imperfectly understood today, it is only because the relative complexity of the merging situation has made a completely scientific treatment of the subject too difficult.

A complete mathematical model for merging cannot be claimed, but it is hoped instead to point out the problems in formulation of such a model, and solve a few of them. From the purely mathematical point of view, the merging problem has some interest beyond the simple question of waiting for a suitable gap in traffic. It might be supposed, for example, that a car traveling along an acceleration lane while waiting for the opportunity to merge is mathematically equivalent to a car waiting at a stop sign, or that the difference resides only in the moving coordinate system. However, the driver on the acceleration lane is able to control the traffic stream with which he wishes to merge by changing his own speed, thereby increasing or decreasing his headway and spacing relative to the main stream. The stop sign problem (which has been very fully analyzed by mathematicians) does not contain this important ingredient, and therefore questions of driving policy do not arise. There is only one possible policy at a stop sign: wait for a suitable gap. Therefore, a mathematical model for a stop sign is purely descriptive, and its principal result consists of a probability distribution for delay.

It is hoped to show that there are a much more varied and interesting collection of problems available when the driver is allowed to alter (within limits) his attitude with respect to the main stream.

NOTATION AND TERMINOLOGY

Three fundamental maneuvers performed by vehicles in traffic may be identified as follows:

1. Weaving. The process of changing lanes within a flow where more than one flow lane exists.

2. Branching. The process of leaving a flow. It does not include any preliminary weaving necessary to get into position for the maneuver.

3. Merging. The process of entering and establishing constituency within a flow. In this part, a vehicle is said to merge during the time when it moves from one acceleration lane to the lane of principal flow, excluding its travel time in the acceleration lane. The particular vehicle under study is called the merging vehicle.

In actual practice, the merging lane can be regarded as an extreme case of the uncontrolled intersection, as shown in Figures 1 and 2. In an idealized model, this is simplified as in Figure 3, where the car shown is said to be in entry position. The length of the merging lane is called L, and of the merging vehicle D. In some circumstances, it will be assumed that L is infinite, or that D is zero.

At any moment, the merging car will define its leading car and following car, meaning simply those cars in the flow lane nearest to the merging car, and respectively ahead of and behind it. The possibility of merging depends largely on these two cars, but may also be influenced by other cars in the main stream. Therefore, the n^{th} car ahead of the merging car is defined as his n^{th} leader, and similarly his n^{th} follower. The merging car cannot have any effect on his leaders, but can compel deceleration among his followers if he wishes to do so. Also, these definitions refer to different cars whenever the merging car passes or is passed.

There are several categories of merging problems:

1. What information is the merging driver assumed to possess? Is he instantly aware of the dynamic characteristics of all cars in the flow lane, or only of his leader and follower, or perhaps only of certain cars' positions, or positions and velocities, or positions, velocities and accelerations? By varying slightly the degrees of information available to the merging driver, new variations can be created on the merging problem.

2. What is assumed the merging driver is attempting to do? Is he trying to merge as quickly as possible, or as far downstream as possible, or as safely as possible? How much deceleration among the following cars is he willing to tolerate?

3. What constraints exist on the merging driver's behavior? Clearly it must be assumed that he cannot accelerate or decelerate his own car beyond the known range of vehicle performance; also, that he will not collide with other cars. This last point is slightly ambiguous, however. If he is not permitted to collide, can it be said that he can merge in such a way as to produce a collision between other cars? It is a well-known consequence of several theories of traffic flow that fairly modest interference with high density traffic can produce shock waves which in certain ranges of parameter values lead quickly to a rear-end collision. Is one to assume the merging driver's familiarity with such theories?

4. What stochastic process shall be assumed governs the flow of traffic in the main stream? An answer to this question might vary from the specification of a separate function x(t) for each car in that stream to some relatively simple idea such as random arrangement and equal velocities.

5. Description of the best policy for a

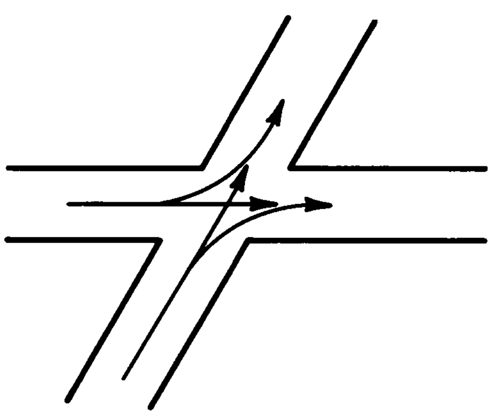

Figure 1. Single lanes intersecting.

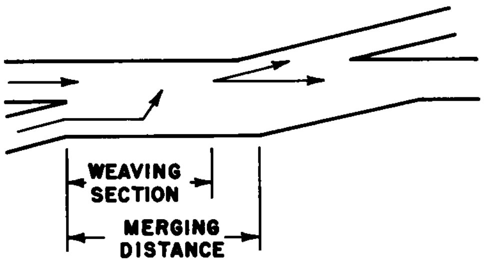

Figure 2. High flow merge.

driver to follow in order to satisfy some particular merging criterion.

6. Description of the probable effect (distribution of delay, for example) on a driver who pursues such a policy.

7. Description of the operation of the system if each driver pursues such a policy.

It is easy to see that by varying 1 to 4, different answers should obtain for 5 to 7. This paper deals with certain special cases.

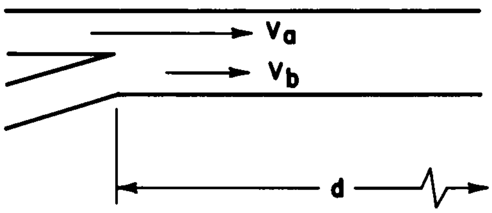

Figure 3. Infinite merging lane.

SAFE MERGING

Before entering the flow of traffic on a freeway, a driver must select the right moment for merging. This selection is based on his judging whether a gap which he intends to enter is large enough for a safe merge.

It is assumed that the driver's primary concern is the distance between his vehicle and the one in front. This distance should be large enough so that, in the event that the vehicle in front makes an emergency stop, there is enough room for the second car to make a safe stop. The distance between two cars could be small if the driver of the following vehicle had all information (i.e., position, velocity, and acceleration) about the vehicle in front and had the means of controlling the acceleration of his vehicle instantaneously. However, this is not the case in practice.

Knowledge about the vehicle in front is limited. The gap and the rate at which it opens or closes can only be roughly estimated. Then, the responses are delayed. The most helpful information in case of emergency stop is the instantaneous appearance of the tail lights on the car in front. This light indicates that the brakes are applied, yet, it is not known how hard. In defining the "safe distance," all of these facts should be taken into consideration. Before defining this "safe distance," the basic mechanics of a vehicle on a straight path should be reviewed.

According to Newton's first law, every body continues in its state of rest or in uniform motion in a straight line until it is compelled by force to change that state. In the case of an automobile traveling on a straight road at a constant speed, the sum of all forces acting on it is equal to zero. Two types of forces are distinguished here: (a) driving forces and (b) motion-resisting forces. The driving force usually is derived from the torque generated by the power plant; sometimes it may be a grade of highway (actually the gravity force) or a wind. The motion resisting forces are caused by friction, rolling friction, wind, highway grade, etc. For a car to travel at a speed of, say, 50 mph, a certain amount of power has to be delivered to overcome the resisting forces. A decrease in supply of power will make a vehicle slow down, until it reaches a new velocity for which the driving and motion-resisting forces are in equilibrium (steady state). In some cases, for example in highway driving, such a control (supply of power) of speed is sufficient for extended periods of time. However, for changes in speed as encountered in city driving, brakes are used to slow the vehicle at a much greater rate than the motion resisting forces would do. In either case (i.e., whether using or not using gas or brake pedal), the driver actually does not control the speed of the car directly; he controls some forces (driving and braking) in such a way that the resultant of all forces makes the vehicle accelerate or decelerate. Both these controlling forces are limited by car and road characteristics. Knowing these characteristics, all forces acting on the vehicle can be evaluated and thus its motion defined. However, in this limited scope of defining the safe distance, it will be sufficient to consider only the maximum values of the controlling forces; in other words, vehicle performance limits. Further, it will be much more convenient to express these limits in terms of acceleration or deceleration rather than in terms of forces. The values of acceleration can be easily measured and are convenient to use in equations of motion. In further discussions, it is assumed that all cars considered have the same acceleration and deceleration capabilities.

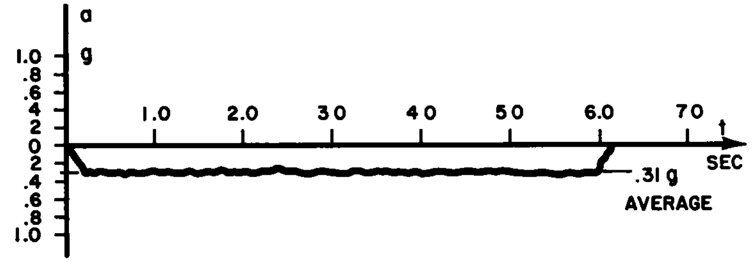

Figure 4. Maximum acceleration-time curve, up to 40 mph, dry surface.

Figures 4 and 5 show typical recordings of maximum acceleration and maximum deceleration taken by the Institute. To simplify the analysis, the average values are used. Thus, in Figure 4, the average value of the maximum acceleration is 0.31 g = 10 ft per sq sec, and in Figure 5, the average maximum deceleration is 0.63 g = 20 ft per sq sec.

An absolute safe distance is a gap between two cars in a lane which will allow the following car to stop safely, even if deceleration of the car in front is maximum. Also, it is assumed that the driver of the merging vehicle will use his brakes to full capacity in order to avoid collision.

For example, two cars on a straight and level path can be represented in the mathematical model as the x-axis (see Fig. 6). The position of car 1 is denoted as x_1 and that of car 2 as x_2. The quantities $\dot{x}_1$, $\ddot{x}_1$ and $\dot{x}_2$, $\ddot{x}_2$ are the respective velocities and accelerations (or decelerations). It is assumed that at time t = 0, the driver of car 1 applies brakes and at the same instant, the tail lights light up. Further, it is assumed that at $t = 0$, $x_2 = 0$ and therefore, $x_1(0) = y(0) = y_o$ (see Fig. 6), also $\dot{x}_1(0) = v_{1o}$ and $\dot{x}_{2o} = v_{2o}$

The driver of vehicle 2 will respond to the signal (tail lights) and will apply his brakes. However, there is always some time required for a driver to move his foot from the gas to brake pedal. This amount of time, called a "time delay," or "reaction time," varies greatly for different people. Figure 7 shows a distribution of reaction times for a group of drivers. The average reaction time according to this figure is 0.73 sec. This time can be defined by T = 0.73 sec. Therefore, at time t = T, car 2 will start to decelerate (neglecting a small variation in speed due to the removal of foot from the gas pedal, an action that precedes the application of brakes by a fraction of a second).

The positions of the cars, for time t > T, are defined by:

$$x_1 = y_o + v_{1o}t - \frac{\ddot{x}_1 t^2}{2} \tag{1}$$

and

$$x_2 = v_{2o}t - \frac{\ddot{x}_2 (t - T)^2}{2} \tag{2}$$

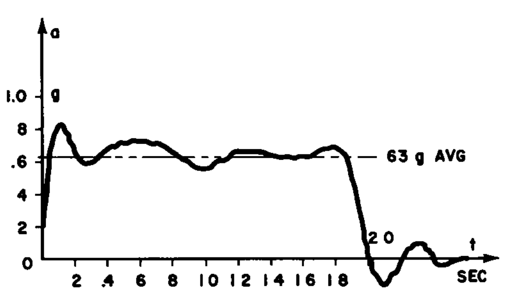

Figure 5. Deceleration-time curve for emergency stop from 30 mph on dry surface.

Because only the emergency stop is considered $\ddot{x}_1 = \ddot{x}_2 = a = 20$ ft per sq sec to find $x_{1max.}$ and $x_{2max.}$, or in other words, the positions at which vehicles 1 and 2 come to a full stop, Eqs. 1 and 2 are differentiated and equated to zero. Thus, $dx_1/dt = v_{1o} - at = 0$ and $dx_2/dt = v_{2o} - a(t-T) = 0$, which gives $t = v_{1o}/a = t_1$ and $t = v_{2o}/a + T = t_2$ in which t_1 and t_2 are stopping times of vehicles 1 and 2. Substituting t_1 and t_2 for t Eqs. 1 and 2,

$$x_{1max.} = y_o + \frac{v_{1o}^2}{2a} \tag{3}$$

and

$$x_{2max.} = v_{2o}\,T + \frac{v_{2o}^{2}}{2a} \tag{4}$$

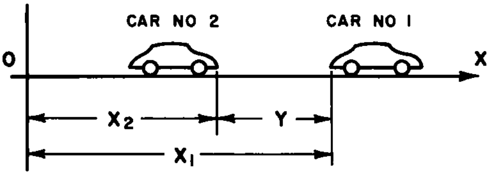

Figure 6. Model of two cars in lane.

Subtracting Eq. 4 from 3,

$$x_{1max.} - x_{2max.} =$$

$$y_o - v_{2o}T - \frac{v_{2o}^2 - v_{1o}^2}{2a} \tag{5}$$

For a safe stop (no collision),

$$x_{1max.} - x_{2max.} \geqq 0 \tag{6a}$$

Therefore,

$$y_o - v_{2o}T - \frac{v_{2o}^{2} - v_{1o}^{2}}{2a} \geqq 0 \tag{6b}$$

For y_o to be a safe distance,

$$y_o \geqq v_{2o}T + \frac{v_{2o}^{2} - v_{1o}^{2}}{2a} \tag{6c}$$

<u>Case 1</u>

If

$$v_{2o} = v_{1o} = v_o, \text{ then,}$$

$$y_o \geqq v_oT \tag{7a}$$

The "safe distance" here, y_o, is a function of the initial velocity v_o and the response time T. It is independent of the decelerations as long as both $\ddot{x}_1$ and $\ddot{x}_2$ are equal. As an example, if v_o = 50 mph = 73 ft per sec and T = 0.73 sec, then $y_o \geqq$ 53.3 ft.

<u>Case 2</u>

In the case when the following car is "catching up" with the car in front, or $v_{2o} > v_{1o}$, Eq. 6c obtains. In comparison with Case 1, the "safe distance," y_o, is increased by $(v_{2o}^2 - v_{1o}^2)/2a$. As an example, if v_{2o} = 50 mph = 73 ft per sec, v_{1o} = 40 mph = 58.4 ft per sec, a = 20 ft per sec, and T = 0.73 sec, then substituting these values in Eq. 6, gives $y_o \geqq$ 101.3 ft. As seen, y_o, is nearly doubled.

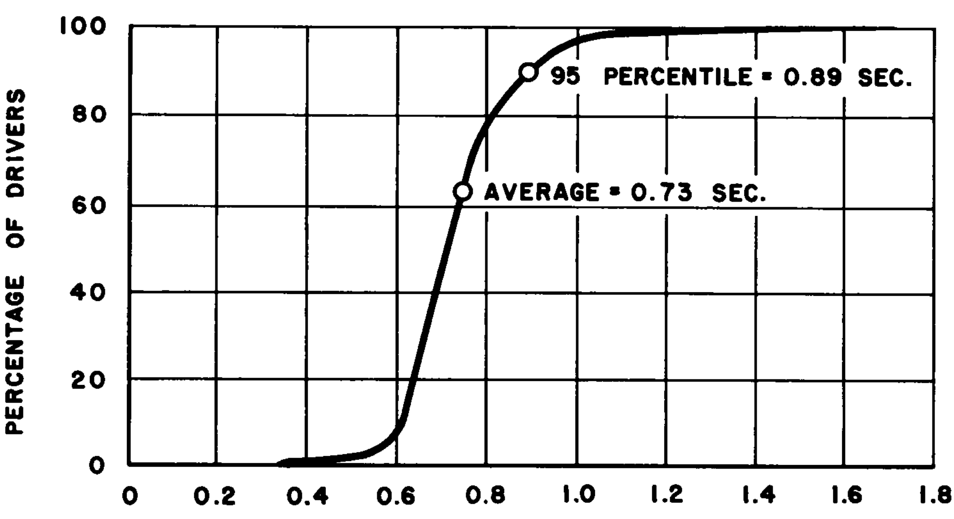

Figure 7. Reaction time in seconds (first stop).

For more general cases, Eqs. 1 and 2 would be used and in similar manner the safe distance, y_o, derived for various $\ddot{X}_1(t)$ and $\ddot{X}_2(t)$.

So far, the safe distance discussed here refers to the gap between two vehicles when one is following the other. In case of merging into oncoming traffic on a freeway, the gap must be large enough to include safe distances between the merging car and the cars in front and behind, and the length of the merging car. Assuming again marginal conditions—i.e., use of brakes to their full capacity on all cars (emergency stop)—the safe gap for merging is

$$s_o = y_{1o} + y_{2o} + L \tag{7b}$$

in which L is the length of the merging car, and y_{1o} and y_{2o} are safe distances (front and rear) computed in the same way as y_o.

In the selection of gap for merging and "placing" the vehicle within this gap, the following three conditions have to be satisfied: $s \geqq s_o$, $y \geqq y_{1o}$, and $y \geqq y_{2o}$.

If $s = s_o + b$, then b is a distance within which the merging vehicle should be placed (see Fig. 8).

From previous discussion (Eqs. 6 and 7) it follows that, for the same value of velocities of vehicles 1 and 3, $y_{1o} \geqq y_{2o}$ for $v_{2o} \geqq v_{1o}$, and $y_{1o} \leqq y_{2o}$ for $v_{2o} \leqq v_{1o}$; when $v_{2o} = v_{1o}$, then $y_{2o} = y_{1o}$. These facts should be remembered by the driver so he can place his car at the right distances, depending on whether his velocity is greater or smaller than that of traffic.

MERGING AS A SIMPLE DELAY PROBLEM

This section makes the simplifying assumptions that the merging vehicle maintains the constant speed v with which it arrives at the entry position, and that the vehicles in the flow lane travel with constant speed V and random placement. This means that the spacing or headway between consecutive vehicles in the flow lane will be governed by the negative exponential distribution. There is then a flow which is Poisson for a moving or stationary observer in either the number of vehicles passing in a given time, or the number of vehicles contained in a given length of road.

When a gap appears that is large enough to allow the entering vehicle to merge safely, taking into account the difference in velocity between the merging vehicle and its leading and following vehicles, then the merge is executed. The distance traveled while waiting for this gap is simply

$$d = vt \tag{8}$$

in which t is the time elapsed after passing the entry point P_1 until a gap appears. The distance d is measured from P_1. A gap sufficient for a safe merge will be at least T time units in length.

It is the intent to construct a theory of merging based on known results in delay theory. These latter treat the wait that a vehicle must endure to enter or cross a stream of traffic when the entering vehicle is at a stop. Under these conditions, the probability distribution of waiting time W(t) has been discussed and is well known (6, 7, 8). However, when the merging vehicle is moving, two difficulties arise. In the first place, a safe gap must be defined more carefully because at some relative velocity for the entering vehicle with respect to the major stream velocity, a time criterion for merge must give way to a space criterion. At a very low relative velocity between major stream and entering vehicle, the time between the transits of two successive vehicles past the entering vehicle may be very long without the existence of sufficient physical space between them for a merge.

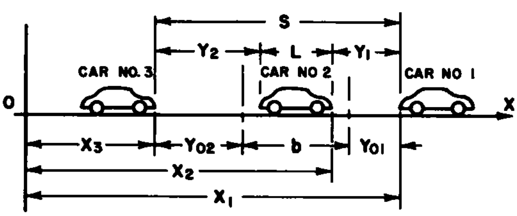

Figure 8. Model of three cars in lane.

The second difficulty arises in the changed rate of flow of gaps past a moving vehicle from the rate of gap flow past a stationary vehicle. It is necessary to be able to characterize the velocity of the vehicles in the flow lane which cannot be done from a mere statement of the flow rate.

As the merging vehicle arrives at the point P_1 imagine all traffic to be stopped instantly as in a photograph. Two points, P_2 and P_3 in the flow lane, in the upstream and downstream direction, respectively, are defined: P_2 is the first point upstream from P_1 for which the distance to the next upstream vehicle is greater than, or equal to, a value S (see Fig. 9). P_3 is defined similarly. If s is the distance between P_1 and P_2, then s has a distribution which is called g(s). The distance P_1 to P_3 also has the same distribution g. If the arrival of the merging vehicle at P_1 occurs at an arbitrary time, then at that instant the location of the other vehicles with respect to P_1 is also arbitrary.

The distance S is more explicitly defined in terms of the safe gap T. If a value is assumed for the quantity T, then to a stationary observer at P_1, the safe gap T can be transformed to a minimum distance VT, and to an observer in the merging vehicle, the distance is (V - v) T. The foregoing expression is only valid in the case where v is less than V. If a safe gap T is required at a relative velocity V - v, then spacing $|V - v|T$ is required.

When relative velocities are small and approach zero, then the physical requirement of a certain minimum space must be accounted for. If S_0 is the length of a vehicle plus minimum maneuvering clearance, the expression for S may now be written,

$$S = \max \left[|V - v|\, T,\ S_0\right] \tag{9}$$

In Figure 10, S is plotted against relative velocity.

It is important to relate the well-known distribution of wait for a gap, w(t), with the distribution g(s). If a minimum gap time is assumed,

$$W(t;T) = \text{Prob}\left[\text{Wait for (gap} \geq T) \text{ is } \geq t\right] \tag{10}$$

and assuming a minimum gap distance,

$$G(s;S) = \text{Prob}\left[\text{Distance to (gap} \geq S) \text{ is } \geq s\right] \tag{11}$$

When each unit of the traffic stream has a velocity V, the minimum time gap T = S/V and, t = s/V. Then,

$$\begin{aligned} G(s;S) &= \text{Prob}\ [\text{Distance to (gap} \geq S) \text{ is } \geq s] \\ &= \text{Prob}\ [v \cdot \text{time to (gap} \geq S) \text{ is } \geq s] \\ &= \text{Prob}\ [\text{Time to (gap} \geq S/V) \text{ is } \geq s/V] \\ &= W(s/V;S/V) \end{aligned} \tag{12}$$

The same result may be obtained by making a change of variable in the density and integrating:

$$g(s;S) = w(t;T)\ |dt/ds| = (1/V)\ w(s/V;S/V) \tag{13a}$$

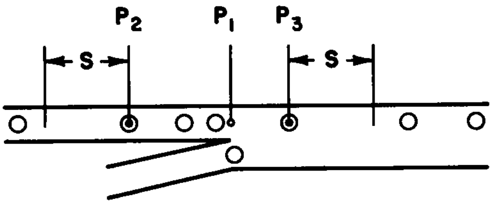

Figure 9. Safe merging space.

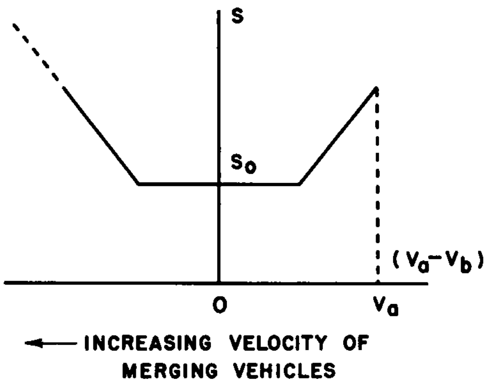

Figure 10. Safe merging space vs relative velocity.

$$G(s;S) = \int_S^\infty g(u;S)\,du = (1/V)\int_S^\infty w(u/V;S/V)\,du = W(s/V;S/V) \tag{13b}$$

If the probability that the merging vehicle travels a distance greater than or equal to d before being able to merge is F(d), the time which the merging vehicle waits for a safe gap is just

$$t = s\ |V - v|^{-1} \tag{14}$$

and the distance traveled by the merging vehicle while waiting this time, assuming V - v is not small, is

$$d = vt = \left(\frac{v}{|V - v|}\right)s \tag{15}$$

The merging vehicle may travel faster or slower than V, but if V = v, then the distance traveled before merge is zero with probability e^{-aS_o}, where a is the average flow rate, and infinite with probability $1 - e^{-aS_o}$.

Substituting in F(d),

$$\begin{aligned} F(d) &= \text{Prob}\ [\text{Distance to merge} > d] \\ &= \text{Prob}\ [(v/\ |V - v|\ )s > d] \\ &= \text{Prob}\ [s > (\ |V - v|\ /v)d] \\ &= G(\ |V - v|\ d/v) \\ &= W\left(\frac{|V - v|\ d}{V_v}\right) \end{aligned} \tag{16}$$

For the particular case of exponential spacings, the distribution of wait is given by the following expression which has been tabulated by Raff (6). The distribution of F(d) then proceeds from the substitution indicated in Eq. 16.

$$\begin{aligned} W(t;T) &= \text{Prob}\ [\text{Wait} > t] \\ &= \sum_{i=0}^{j-1} (-1)^i e^{-(i+1)aT} \left\{ \frac{[a(t - iT)]^i}{i!} = \frac{[a(t - iT)]^{i+1}}{(i+1)!} \right\} \end{aligned} \tag{17}$$

$$\text{for}\ (j - 1)T \leqq t \leqq jT$$

The gap criterion enters as a parameter in this distribution. F(d) has been plotted in Figure 11. The probability of zero wait has also been plotted (see Fig. 12).

In Figure 13 the most interesting results of this section are plotted. The length of the merging lane is considered fixed at 500 ft and merging vehicles travel at constant velocity on the merging lane until either a merge is completed or the lane ends. The stopping behavior considered in the next section is simplified here to instantaneous braking in zero distance. The figure shows the probability of success in merging with the elementary policy of constant velocity and two features are of interest; the anomaly due to the minimum distance requirement in the vicinity of the flow lane velocity V, and the minimum probability of success for moderate merging velocity.

For this simple merging policy, Figure 11 shows the point of view of design length of merging lane so that a given fraction of vehicles will merge before stopping, and Figure 13 shows the alternate point of view, which is the best constant velocity to choose for an existing merging lane and given flow lane velocity.

VARIABLE SPEEDS, COORDINATE SYSTEMS

If the length of the merging lane is L, and the position of the leading point of the merging car is denoted by x, $0 \leqq x \leqq L$, where the origin is taken at the beginning of the merging lane, the time origin at the moment when the merging car appears in the lane may also be conveniently taken so that t = 0 when x = 0. If the velocity of the merging car is v (no longer constant), if v_0 be the value on entering the lane, and v_m the largest value obtainable in the distance L, then with exponential acceleration, the

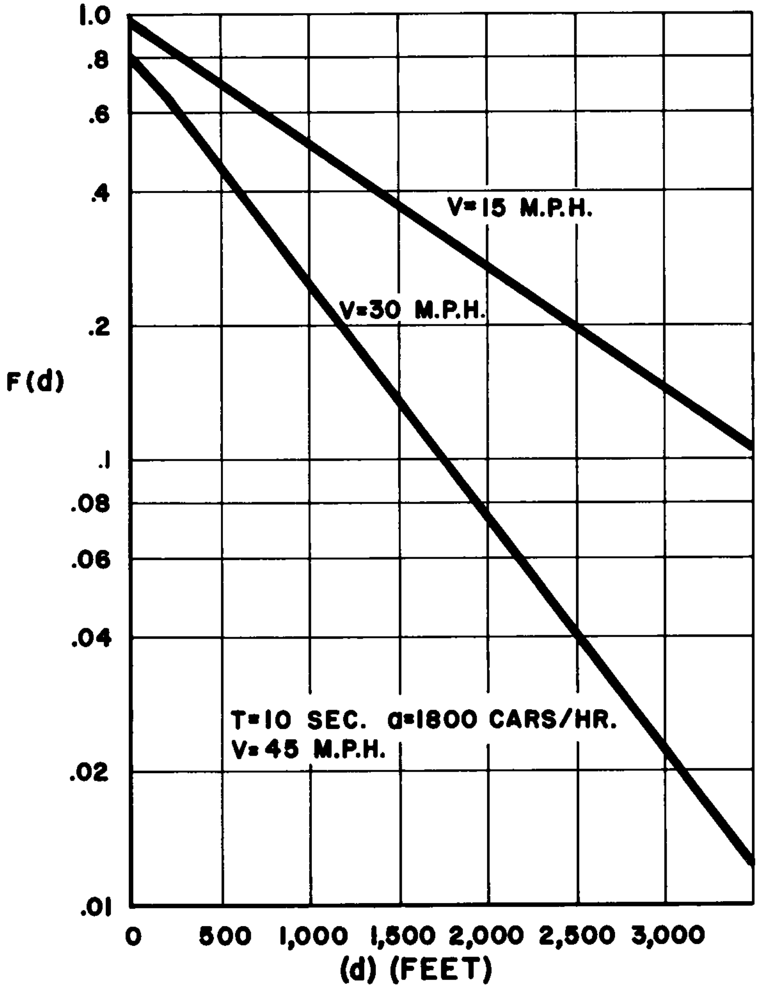

Figure 11. Probability that merge requires at least d feet.

best velocity achievable in time t in the merging lane would be

$$v = v_m - (v_m - v_o)\, e^{-\beta t} \qquad (18)$$

in which β is either a constant or the mean value of a random variable. Some information on values of β and v_m could be obtained from the drag races, which are now widely held. In these contests, both v_m and elapsed time are announced in every case, apparently in recognition of the independence of these quantities. In fact, if Eq. 18 is integrated,

$$x = v_m t + (1/\beta)(v_m - v_o)(e^{-\beta t} - 1) \qquad (19)$$

Setting x = L in this equation yields a relationship involving the elapsed time, denoted by t_L. In the drag races, $v_o = 0$, and L = ¼ mi; therefore, β could be conveniently computed for various given values of v_m and t_L.

The minimum velocity permitted the merging car is v_o. The reason for this apparently arbitrary restriction is quite simple; if that car were permitted to have very small velocities, this would be equivalent to allowing an infinitely long acceleration

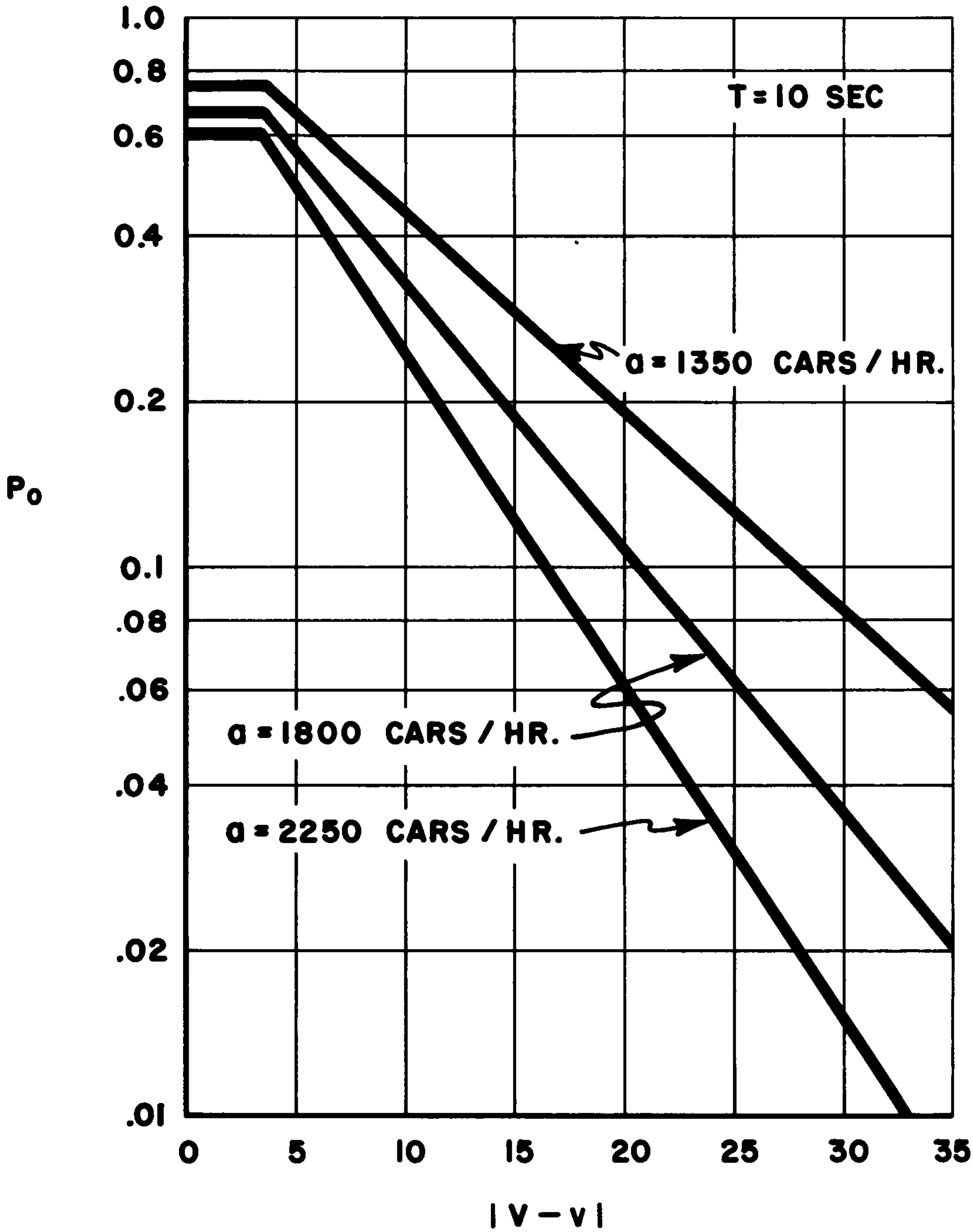

Figure 12. Probability of zero wait (P_0) vs relative velocity $|V-v|$.

lane. Because it is intrinsic to problem 5 that L is finite (for with an infinitely long merging lane, a best policy might be never to merge), the velocities allowed must be bounded away from zero, and v_0 is a convenient and not wholly unrealistic bound.

There is however, an exception to this statement. When the merging car approaches the end of the available lane, he must stop if he has been unsuccessful in merging. Assuming linear braking, the allowed velocity variation for the merging driver is shown in Figure 14, where the equation of the curved portion of the boundary is Eq. 18. The driver will be permitted to cross from the vertically shaded region into the horizontally shaded region only if he is able to merge before coming to the end of the merging lane. Otherwise, he must apply maximum braking at the line of maximum braking, and merge from a standstill at the end of the lane.

If it is assumed that the traffic in the adjoining lane is all going at the same speed, then an auxiliary coordinate system can be defined in the adjoining lane relative to the merging car, and this system can be used to measure the degree of success in merging. If y denotes the position of the merging car relative to the rigidly moving adjoining lane; if the value of y is taken as zero at the point of entry into the acceleration lane when the merging car first arrives there; and if the position of the merging car relative to the adjoining stream is positive downstream from $y = 0$ and negative upstream from $y = 0$; then one can refer to a positive or negative merge, depending on whether at the instant of merging, the merging car has improved his position relative to the adjoining stream or not.

If the constant speed of the adjoining stream is V, then the merging car moves at speed $v - V$ relative to that stream, and in time t changes his y coordinate by an amount $t(v - V)$. Therefore, if he obtains maximum acceleration in the merging lane, and is able to merge at the last moment, he will have gained on the traffic stream an amount,

$$t_L (v - V),$$

which represents the best value y can have. If the criterion of success that he shall merge as far downstream as possible, is adopted, this can be expressed numerically

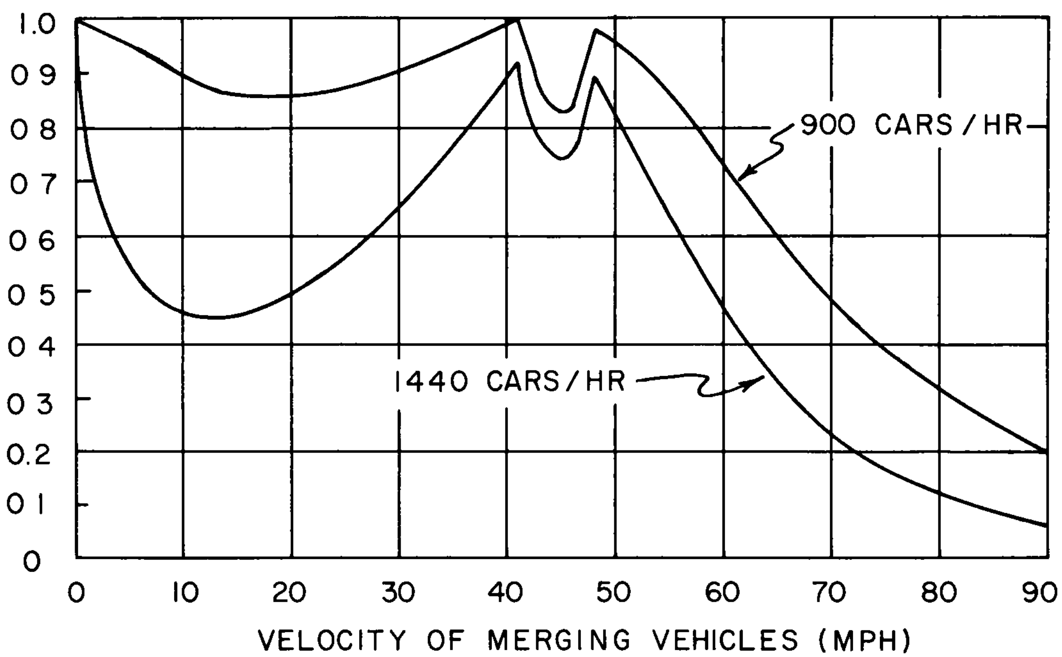

Figure 13. Probability of merging in 500 ft or less for flow lane velocity of 45 mph.

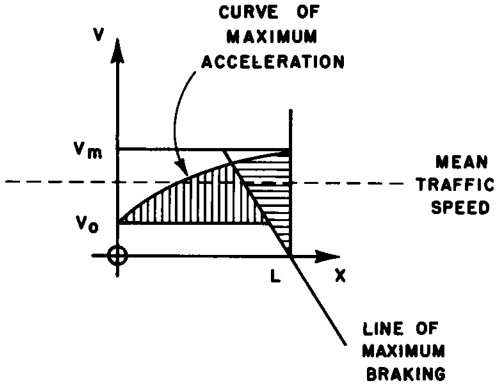

Figure 14. Allowed velocity variance for merging driver.

according to the distance of the y value obtained from its maximum given by $t_L(v-V)$.

PROBABILITY OF MERGING IN N^{TH} GAP

If at the moment when the merging car arrives at the entry position, the length of the gap from L back to the first vehicle flow lane is d_1, from that one to the next is d_2, etc., what conditions need to be fulfilled for the merging vehicle to merge into d_r, and what is the probability of these conditions being fulfilled? First, considering d_1, which is supposed to extend back from $x = L$ to $x = L - d_1$, if the merging vehicle is barely going to fit into this gap, then its nose must arrive at $x = L$ exactly when the gap has shrunk to a length D (a car length). This has taken an elapsed time t_L, during which the traffic flow has traveled Vt_L, and this is the amount by which d_1 has shrunk. Therefore,

$$L - d_1 + t_L V = L - D \tag{20}$$

or

$$d_1 = D + t_L V \tag{21}$$

for the barely possible merge into the first gap. Consequently, if Eq. 20 is improved by enlarging d_1, merge in that gap will be possible. Therefore, if it is said that

$$C = D + t_L V \tag{22}$$

then it follows that merging into d_1 is possible when $d \geqq C$ and impossible otherwise. Because the distribution of gaps is negative exponential, the probability of a merge into d_1 is

$$P_1 = \int_C^\infty \lambda e^{-\lambda x} dx = e^{-\lambda C} \tag{23}$$

in which λ is the traffic density in the flow lane.

The probability that the first gap d_1 will be unsatisfactory but the second one d_2 will be satisfactory is the probability that all the following inequalities will be satisfied: $d_1 < C$, $d_1 + d_2 \geqq C$, and $d_2 \geqq D$, which, in the d_1-d_2 plane represents the area shaded in Figure 15. Setting up the integral for this area,

$$P_2 = \int_C^\infty \int_0^C \lambda^2 e^{-\lambda(x+y)} dx\, dy + \int_D^C \int_{C-y}^C \lambda^2 e^{-\lambda(x+y)} dx\, dy$$

$$= e^{-\lambda C} \left[1 - e^{-\lambda D} + \lambda(C - d)\right] \tag{24}$$

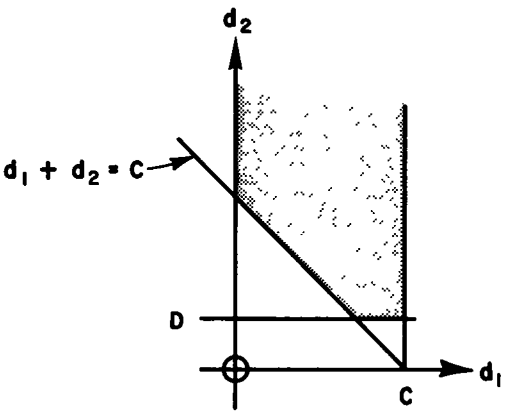

Figure 15.

It is easy to calculate successive forms of this equation, using in each case the previous form, together with the fact that the distribution of sums of exponential variables obey the gamma-type distribution. For example, the third stage uses the inequalities $d_1 + d_2 < C$, $d_1 + d_2 + d_3 \geq C$, and $d_3 \geq D$, which can be reduced to $x < C$, $x + y \geq C$, and $y \geq D$, in which $x = d_1 + d_2$ and $y = d_3$. In this way, the probability of fitting into the n^{th} slot is given by

$$p_n = \int_C^\infty \int_0^C \lambda e^{-\lambda y} \frac{\lambda^{n-1} x^{n-2}}{(n-2)!} e^{-\lambda x} \, dx \, dy$$

$$+ \int_D^C \int_{C-y}^C \lambda e^{-\lambda y} \frac{\lambda^{n-1} x^{n-2}}{(n-2)!} \, dx \, dy \tag{25}$$

So far the constraints mentioned in the other parts of this paper have not been applied to Eq. 25. It appears that the best way to proceed from this point would be to use programed digital computers to analyze and compare the various approaches proposed in this paper, and hope to present further results in this direction. Meanwhile, the analysis does not seem to have been completely worn out, and the authors wish to encourage other workers to carry it further.

REVIEW OF MERGING LITERATURE

There have not been many studies of merging as a distinct model from delay at traffic lights to stop signs. The game-theory aspect of the problem, in which the merging driver is able to control to some extent the process in which he wishes to merge, has been recognized by Huemer (2) but not carried far except by computer simulation.

As long ago as 1954, Ho (1) proposed a primitive merging model in which n_2 cars in the merging lane are waiting to merge with n_1 cars in the flow lane separated randomly with mean headway $1/\lambda$. If T is the time required for a single car to merge, then Ho gives the density function of the total time to complete the merging by

$$f(t) = C e^{-\lambda t} \sum a_i (t + b_i)^n \tag{26}$$

in which C, a_i, b_i, and n are functions of the parameters defining the system. When n_1 and n_2 are equal, and some approximations are used, Eq. 26 simplifies drastically to

$$f(t) = \frac{\lambda^{n-1} (t - nT)^{n-2}}{(n-2)!} \exp \left[-(t - nT) \lambda \right] \tag{27}$$

Little (3) compares the advantages of merging just before and just after the main stream has passed through a signalized intersection, and obtains formulas for the average delay in each case. He also treats a number of other maneuvers in Poisson traffic near intersections.

By far the best mathematical treatment of merging is due to Oliver (5). He first considers equally important lanes merging with each other and allows the possibility of queues in either branch. He then develops the classical queue probability equations for both branches and solves these to obtain the steady state queue levels. The system considered here is called by Oliver "priority merging," in the sense that vehicles in the merging lane are always at the mercy of traffic in the flow lane. In this case, Oliver finds both the stationary queue length probabilities and the distribution of delay.

REFERENCES

1. Ho, E., "A Statistical Analysis of Congested Merging Traffic." In Trautman, D. L., et al., "Analysis and Simulation of Vehicular Traffic Flow." Inst. of Trans. and Traffic Engineering, UCLA, Research Report 20 (Dec. 1954).
2. Huemer, D. A., "The Monte Carlo Method." M. A. thesis, Claremont College (1961).
3. Little, J. D. C., "Approximate Expected Delays for Several Maneuvers by a Driver in Poisson Traffic." Operations Res., 9:39-52 (Jan.-Feb. 1961).
4. Normann, O. K., "Braking Distances of Vehicles from High Speeds." HRB Proc., 32:421-436 (1953).
5. Oliver, R. M., "On High Speed Two-Lane Traffic Merges." (In preparation)

6. Raff, M. S., "The Distribution of Blocks in an Uncongested Stream of Traffic." Jour. of Amer. Statistical Assoc., 46:114-23 (1951).
7. Tanner, J. C., "The Delay to Pedestrians Crossing a Road." Biometrika, 38: 383-92 (1953).
8. Weiss, G. H., and Maradudin, A. A., "Some Problems in Traffic Delay." Jour. Soc. Indus. and Appl. Math. (in preparation). Presently Tech. Note BN-224, AFOSR-125, Inst. for Fluid Dynamics and Appl. Math., Univ. of Maryland (Dec. 1960).

A High-Flow Traffic-Counting Distribution

ROBERT M. OLIVER and BERNARD THIBAULT, Respectively, Associate Research Engineer and Graduate Research Engineer, Institute of Transportation and Traffic Engineering, University of California, Berkeley

Although many observations have been made on intervehicle headways and traffic volumes, it is important to improve the theoretical bases for predicting a number of flow and density characteristics from a limited number of observations. Whereas considerable attention has been given to the theoretical and experimental evaluation of the statistical distributions of intervehicle spacings, there has been much less information available about the discrete counting distributions. The principal effort has been devoted to Poisson-like counting distributions.

The purpose of this paper is to review and present counting distributions which take into account two fundamental characteristics of medium- and high-density traffic flows: (a) platooning or bunching, and (b) minimum spacing, jam-density of the so-called maximum-pack situations. These counting distributions are derived from intervehicle spacing distributions, which have been studied both theoretically and experimentally; in the low-density or low-flow case it is shown that these distributions have the limits of the well-known Poisson case.

• IN AN ATTEMPT to understand conditions that affect traffic flow, engineers have applied probability theory to the analysis of many traffic-counting problems. Although it has sometimes been difficult to predict the exact behavior of any one vehicle or driver, experiments have demonstrated that departures from an average behavior may follow predictable and relatively stable patterns.

In the theory of traffic flow, several authors have studied the probability distributions of spacings between vehicles and the related problem of the distribution of vehicle counts in an interval of time or space. With a reasonably accurate description of intervehicle headways and the distribution of vehicle counts, it should be possible to answer a large number of flow and congestion problems that arise in and around traffic streams. The relations between flow, density, road capacities, delays, and the effect of queueing on the velocity distributions of free-moving vehicles will undoubtedly depend on the basic assumptions about intervehicle spacings.

In studying the arrangement of cars on a road, early writers discussed the combinatorial aspects of random arrangements of points on a line. The well-known Poisson counting law was then derived as a limiting (low-density) case. More recently, the counting problems have been studied as time-dependent processes. By formulating the probability that an intervehicle spacing lies between certain limits, it is theoretically possible to find the probability distribution of spacings between nonadjacent vehicles and from them the discrete distributions of vehicle counts in an interval of time or space.

A cursory review of the literature suggests that the theoretical as well as the experimental work in this area has focused on at least two major problems. The first of these is the effect of bunching or queueing within the traffic stream. It is not uncommon to find several vehicles following a slow or unusually large vehicle; the probability that spacings between cars lie between limits that are of the order of several car lengths has been observed to be higher than that predicted by the exponential distribution. Equivalently, the probability of counting several vehicles close to one another is higher than terms of the Poisson distribution would predict.

The second of these problems has to do with the size of the vehicles; this size forbids them from occupying the same road space. To replace cars on a road by points on a line is not always realistic because high flow or jam-density situations inevitably lead to the conclusion that there is an upper bound to the number of vehicles that can be counted in an interval. If vehicles also have an upper limit to their velocities this statement applies as well to time counts as it does to counts over a length of road. In these cases, the probability of finding vehicles within a fraction of their respective lengths is zero and is, of course, smaller than the Poisson law would predict.

Although there has not been complete agreement, either theoretically or experimentally, as to the structure of the distributions of intervehicle spacings at their origins, there does seem to be general acceptance of the exponential shape of the distribution for large arguments; that is to say, the probability of finding intervehicle spacings greater than a large value decreases exponentially with the size of the spacing.

This paper is divided into seven sections. The second section briefly reviews the historical background of the statistical analysis of intervehicle headways. The third section describes a limiting form of Schuhl's double-exponential distribution. The fourth section reviews some of the mathematical properties of the geometrically compounded Poisson process (Stuttering Poisson); these results are then used in the fifth section to obtain a discrete counting distribution. The sixth section discusses the probability of "maximum pack" and the final section presents some numerical results and a discussion of qualitative features of these distributions.

HISTORICAL BACKGROUND

As early as 1936 Adams (1) pointed out that the distribution of cars on a road could be formulated mathematically. By assuming that the vehicles were randomly distributed points on a line and by making certain limiting assumptions he and at least two other authors (6, 7) showed that the Poisson distribution was applicable to some traffic counting experiments.

By 1955 several distributions of intervehicle spacings had been proposed. One of these is the double-exponential distribution (see Eqs. 1, 2, and 3) derived from geometrical arguments by Schuhl (22, 23, 24). He also obtained certain relations for the discrete counting distributions associated with an arbitrary distribution of intervehicle spacings. An important aspect of Schuhl's distribution is that one limiting case represents the exponential distribution, whereas a second limiting case represents the class of distributions found in certain high-density situations.

To study the flow of traffic through a signalized intersection Newell (18) in 1956 discussed a translated exponential distribution for intervehicle headways. The main feature of this distribution was that it could account for the size and finite velocity of a vehicle as well as some experimental evidence which supported maximum or capacity flow rates. Under certain medium flow conditions, Kinzbruner (13) obtained further experimental evidence to support this distribution. Oliver (19) published some theoretical results for the various counting distributions associated with this translated exponential distribution. Feller (3) in 1948 had already formulated the basic problems associated with the count of nuclear particles. The so-called type I counter resulted in a counting distribution which, except for the distribution of spacings to the first count, was in many respects identical to that one posed in the context of traffic flows.

In 1958 Haight and several collaborators (10) analyzed traffic flow data and came to the conclusion that realistic distributions could be classified as intermediate between (a) random and (b) equally spaced models. In the former case, the exponential intervehicle spacings led to the Poisson counting distributions; the second, to a deterministic count that is just equal to the integral part of the interval of interest divided by the fixed spacing between vehicles. Haight showed that a family of distributions satisfying certain theoretical and experimental requirements were the Erlang or Pearson type III distributions. Counting distributions which correspond to this assumption for intervehicle spacings are the generalized Poisson functions described by Haight (9) or various state probabilities calculated by Morse (17) and Jewell (11). Whittlesey and Haight (30) have also obtained certain approximations and numerical results for these counting distributions.

In 1959 Kell (12) produced experimental evidence to show that the double exponential distribution suggested by Schuhl accurately described intervehicle spacings in certain medium flow situations. An extensive number of experiments was made and four unknown parameters in the Schuhl distribution were expressed in terms of the flow rate or volume of traffic. Extrapolation of these parameters for high volumes indicate that a limiting form of the double exponential distribution may be appropriate for high-flow situations (see Eqs. 4 and 5).

By 1960 Miller (16) had reached the important conclusion that the random variables describing successive intervehicle spacings might not be independently sampled. He proposed a model of traveling queues which took specific account of bunching or queueing effects. Mathematically, this was a generalization of a special bunching configuration suggested by Bartlett (2) and derived independently from overtaking rules by Oliver (20). The important new consideration brought into all of these studies was the dependence of gaps between adjacent vehicles. Not only is it necessary to resolve the distributions of gaps between queued vehicles, but also one must discuss the spacings between queues, the distribution of queue lengths, and the formation of queues as the result of flow around slow-moving vehicles.

In 1960 May and Wagner (14) published an extensive list of data gathered in the vicinity of Detroit and Lansing, Mich. In the case of extremely high flow rates, the probability density distributions of intervehicle headways showed a marked tendency to rise sharply from zero and then decrease exponentially from this peak or modal value. Minimum headways were seldom evident for flow rates exceeding 30 per min but were almost always present for flow rates less than this value.

In the same year Weiss and Maradudin (24) published some new results in the theory of vehicle delays at the stop-sign type of intersection. In deriving numerical results they made use of a probability distribution of intervehicle spacings which was a translated version of the geometric-exponential distribution discussed by Jewell (11) and which, as shown later, is the same limiting case of Schuhl's distribution observable in some high-volume samples of Kell's data.

Although vehicles in a dense traffic stream are obviously restrained by each other's movements and although the independence assumption of spacings between successive vehicles may be unrealistic in some respects, a large body of theoretical and experimental research has supported Schuhl's distribution. Either in its own right or as a limiting version of more general cases, the mixture of two exponentials has been used to describe vehicle behavior in medium density traffic streams. A limiting version of Schuhl's distribution is discussed in the following section. The counting distributions that correspond to it are the major subject of the remainder of this paper.

DISTRIBUTION OF INTERVEHICLE SPACINGS

Schuhl obtained a description of the distribution of spacings between vehicles on purely theoretical grounds. By considering two types of vehicles—slow and fast— and by requiring that the sum of their respective flow rates equal the total vehicular flow rate, he obtained the probability distribution of the spacings between adjacent vehicles as the mixture of two exponential functions, each with its own decay constant. An observer picks a fast vehicle with probability α and a slow vehicle with probability $1 - \alpha$; if the choice results in a fast vehicle the probability that the spacing to the next vehicle (either slow or fast) is greater than t is just $e^{-\lambda_1 t}$ if the choice results in a slow vehicle the probability that the spacing to the next vehicle is greater than t being equal to $e^{-\lambda_2 t}$. The mixture of these probabilities results in

$$A(t) = \alpha e^{-\lambda_1 t} + (1 - \alpha) e^{-\lambda_2 t} \tag{1}$$

for the probability that the spacing between any two vehicles is greater than or equal to t. Though the words "slow" and "fast" may not be appropriate in the sense that vehicle velocities may themselves be distributed over a wide range of values, it may be helpful to think in terms of retarded and unretarded vehicles. That is to say, the slow vehicles travel at their free or desired speed, whereas the fast are restricted in their ability to maneuver; because of heavy flows in an adjacent lane, the latter may not have opportunities to perform the passing maneuvers that lead to unrestrained flow conditions.

Morse (17) has called the distribution of Eq. 1 the hyper-exponential distribution, and one of its counting distributions the hyper-Poisson.

Although Eq. 1 might apply to a set of points with restricted motion along a line, it is clear that vehicles occupy a finite amount of space in a traffic stream; if there is an upper bound to the free velocity then there is at least this same upper limit to the velocity of the constrained or retarded group of vehicles and the minimum time or headway between successive vehicles is simply the ratio of the minimum spacing between vehicles to the maximum velocity at which they can travel. Even if the domain of definition of velocity values were $(0, \infty)$ and the lower bound on intervehicle headways were zero, many experimental results indicate that near-zero headways are highly improbable; hence, the assumption of a lower bound on intervehicle headways serves as an approximation of a real probability distribution where the density function is small for small headways, increases sharply to a maximum, and then decreases exponentially for large values of the argument.

To account for this feature of minimum headways, Schuhl modified Eq. 1 to include a term for minimum separations between vehicles. In this translated version,

$$\begin{aligned} A^{\Delta}(t) &= 1 && 0 \leq t < \Delta \\ &= \alpha e^{-\lambda_1(t - \Delta)} + (1 - \alpha) e^{-\lambda_2(t - \Delta)} && \Delta \leq t \end{aligned} \tag{2}$$

Δ refers to the minimum headway or spacing between vehicles. By a simple change of the constant terms in Eq. 2, it is possible to write the translated probability distribution as

$$\begin{aligned} A^{\Delta}(t) &= 1 && 0 \leq t < \Delta \\ &= c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t} && \Delta \leq t \end{aligned} \tag{3}$$

The constants c_1 and c_2 must add to a number greater than one.

Experiments by Kell (12) and May and Wagner (14) indicate that as traffic volumes increase there will be an increase in the fraction of restrained vehicles relative to the free-moving group. This seems reasonable because passing maneuvers generally become difficult as traffic volumes and densities increase. With very high-flow conditions the constrained vehicles travel closer and closer to the free-flowing leader of a platoon or bunch. Kell's data indicate that the exponential decay constant, λ_1 in Eq. 2, increases sharply with increasing flow rates, whereas that of the free-moving group (λ_2 in Eq. 2 is less than λ_1). The composite curve $A^{\Delta}(t)$ tends to look like an exponential with a large decay constant for small headways and like an exponential with a much smaller decay constant for large headways.

In the limit as $\lambda_1 \longrightarrow \infty$ and $\lambda_2 \longrightarrow 0$ one obtains a probability distribution of intervehicle spacings

$$\begin{aligned} A^{\Delta}(t) &= 1 && 0 \leq t < \Delta \\ &= (1 - \alpha) e^{-\lambda(t - \Delta)} && \Delta \leq t \end{aligned} \tag{4}$$

where λ replaces λ_2 in Eq. 2. The probability density distribution

$$a^{\Delta}(t) = -\frac{dA(t)}{dt} = \alpha\,\delta(t - \Delta) + \lambda(1 - \alpha) e^{-\lambda(t - \Delta)} \tag{5}$$

also points out the fact that the probability of finding vehicles queued at the minimum separation is α.

The probability distribution of Eqs. 4 and 5 forms the basis for the counting distributions obtained in this paper. When $\Delta = 0$, one obtains the special case—called the geometric exponential distribution (11)—of the hyper-exponential distribution which has been discussed by Jewell (11). Although many results have been published on the counting distributions associated with this special case, it may help to review some of their properties in the following section. It is important to point out that many of the analytical expressions obtained for the $\Delta > 0$ case can be obtained in terms of those obtained for the $\Delta = 0$ case; hence, numerical computations made for the $\Delta = 0$ case can be used as building blocks for the $\Delta > 0$ cases.

The mean spacing, ν_Δ, of the translated distribution of Eq. 4 is Δ plus the mean spacing of the untranslated case, and the variance, σ_Δ^2, of the former is identical to the variance of the latter. If constants in Eq. 4 are renormalized so that $\mu = \lambda/1 - \alpha$,

$$\nu_\Delta = \int_0^\infty A^\Delta(t)\,dt = \Delta + \mu^{-1} \tag{6a}$$

$$\sigma_\Delta^2 = \int_0^\infty 2(t - \bar{t})\,A(t)\,dt = \frac{1+\alpha}{\mu^2(1-\alpha)} \tag{6b}$$

It has been shown by several authors (11, 24, 25) that if one randomly selects a point in time, the probability density distribution of spacings to the first car, $u^\Delta(t)$, is the product of the stationary flow rate and the probability that the spacing between two cars is greater than t. Because the stationary flow rate, μ_Δ, is the reciprocal of the average intervehicle headway in Eq. 6a the "starting-at-random" density distribution is obtained:

$$u^\Delta(t) = \mu_\Delta \qquad 0 \leq t < \Delta \tag{7a}$$

$$= (1-\alpha)\,\mu_\Delta e^{-\frac{\mu_\Delta(1-\alpha)(t-\Delta)}{1-\Delta\mu_\Delta}} \qquad \Delta \leq t \tag{7b}$$

μ is the stationary vehicle flow rate for the special case where minimum headways are zero. The expected wait to the first vehicle from a random origin is

$$\int_0^\infty t u^\Delta(t)\,dt = \frac{\sigma_\Delta^2 + \nu_\Delta^2}{2\nu_\Delta}$$

GEOMETRICALLY COMPOUNDED OR STUTTERING POISSON PROCESS

Several authors (4, 5, 11) have studied the geometrically-compounded or Stuttering Poisson process which corresponds to the special case $\Delta = 0$ in Eq. 4. As already mentioned, the mathematical structure of the discrete counting distributions which corresponds to the more general case $\Delta > 0$ is similar to that of the Stuttering Poisson.

The Stuttering Poisson distribution may arise in the following way: consider vehicles replaced by points on a road. One group of these vehicles is a free-flowing or unrestrained group, the probability density distribution of spacings (headways) between vehicles in this group is exponential with mean value λ^{-1}. This is equivalent to the statement that the count of vehicles in an interval is Poisson-distributed with average flow rate equal to λ. The second group of vehicles are queued behind unrestrained vehicles. The probability that the queue of restrained vehicles is of length n-1 is a geometric distribution;

$$a_n = (1-\alpha)\,\alpha^{n-1} \qquad n = 1, 2, \ldots \tag{8}$$

in which $(1 - \alpha)$ is the probability of no restrained vehicle following an unrestrained one. The probability p(n/m) of finding n - m restrained vehicles behind m unrestrained vehicles or a total of n vehicles with m unrestrained is the m-fold convolution of Eq. 7, the negative binomial,

$$p(n/m) = \binom{n-1}{m-1}(1-\alpha)^m\, a^{n-m} \qquad n \geq m \geq 1 \tag{9a}$$

$$= 1 \qquad n = m = 0 \tag{9b}$$

Consequently, the probability $g_n(t)$ of finding n vehicles in an interval t chosen at random is the sum over the possible number of unrestrained vehicles:

$$g_n(t) = \sum_{m=1}^{n} \binom{n-1}{m-1} (1-\alpha)^m \alpha^{n-m} \frac{e^{-\lambda t}(\lambda t)^m}{m!} \qquad n \geq 1 \tag{10a}$$

The probability that no vehicles are observed in t is just $e^{-\lambda t}$, the probability that the spacing to the first unrestrained vehicle is greater than t. For $n \geq 1$ the distribution $g_n(t)$ can be written in terms of the associated Laguerre polynomials of order 1,

$$g_n(t) = \frac{\lambda t(1-\alpha)\alpha^{n-1}}{n} e^{-\lambda t} L_{n-1}^{(1)}\left(\frac{(\alpha-1)\lambda t}{\alpha}\right) \tag{10b}$$

in which the Laguerre polynomial of order a is

$$L_n^{(a)}(x) = \sum_{i=0}^{n} \binom{n+a}{i+a} \frac{(-x)^i}{i!} \tag{11}$$

The distribution of Eqs. 10a and 10b corresponds to the case where the counting of traffic begins at random. If one starts to count just after one vehicle has passed, the discrete counting distribution differs only slightly from $g_n(t)$. This new distribution is labeled $h_n(t)$; it can also be derived by arguments similar to those used for Eq. 8 through 10.

The distinction between these two cases can be well illustrated by means of their generating functions. In the "starting-at-random" case, the generating function

$$G(z,t) = \sum_{n=0}^{\infty} g_n(t)\, z$$

of the Stuttering Poisson is just that of the geometrically compounded Poisson process discussed by Feller (3). By substituting the generating function of Eq. 8 for the variable z in the generating function $e^{-\lambda t(1-z)}$ of the Poisson process,

$$G(z,t) = e^{-\frac{\lambda t(1-z)}{1-\alpha z}} \tag{12}$$

Expansion of Eq. 12 in powers of z leads to formulas for the counting distribution. By substituting the stationary vehicle flow rate $\mu = \lambda/1-\alpha$,

$$g_0(t) = e^{-\mu t(1-\alpha)} \tag{13a}$$

$$g_1(t) = \mu t(1-\alpha)^2 e^{-\mu t(1-\alpha)} \tag{13b}$$

$$g_2(t) = \left[\alpha(1-\alpha)^2 \mu t + \frac{(1-\alpha)^4(\mu t)^2}{2}\right] e^{-\mu t(1-\alpha)} \tag{13c}$$

These are identical to the expressions already obtained in Eq. 10a and 10b.

When the counting experiment starts with the passing of a vehicle, the generating function

$$H(z;t) = \sum_{n=0}^{\infty} h_n(t) z^n$$

is obtained by a slight modification of the geometrically-compounded Poisson process previously discussed. The total count can now be expressed as the sum of two random variables X and Y. X is the number of vehicles in the first bunch exclusive of the vehicle which begins the counting experiment. Hence, from Eq. 8

$$\Pr\left\{x = n\right\} = a_{n+1} \qquad n \geq 0 \tag{14}$$

Because the first bunch of vehicles is located at the origin, Y represents the remaining count with probability distribution $g_n(t)$. Hence, the distribution of X + Y is the convolution of Eq. 14 with $g_n(t)$ in Eq. 10:

$$h_n(t) = \sum_{m=0}^{n} (1-\alpha)\alpha^m g_{n-m}(t) \qquad n \geqq 0 \tag{15}$$

Its generating function is therefore equal to the product of the generating function of Eq. 14 and G(z;t):

$$H(z;t) = \frac{1-\alpha}{1-\alpha z} e^{-\frac{\mu t(1-\alpha)(1-z)}{1-\alpha z}} \tag{16}$$

By summing Eq. 15 or by expanding H(z;t) in powers of z, the coefficient of z^n becomes

$$h_n(t) = \sum_{m=1}^{n} \sum_{i=0}^{m-1} \binom{m-1}{i} \frac{(1-\alpha)^{i+2}}{(i+1)!} \alpha^{n-i-1}(\lambda t)^{i+1} e^{-\lambda t} + (1-\alpha)\alpha^n e^{-\lambda t} \tag{17a}$$

By making use of Eq. 11 $h_n(t)$ can also be expressed in terms of Laguerre polynomials as

$$h_n(t) = \alpha^n(1-\alpha)\, e^{-\mu t(1-\alpha)} L_n^{(0)}\left(-\frac{(1-\alpha)^2 \mu t}{\alpha}\right) \tag{17b}$$

These discrete counting distributions can also be obtained by considering the probability distribution of spacings between nonadjacent vehicles. If a(t) is the probability density distribution of intervehicle spacings, $a_n(t)$ is used to denote the probability density distribution of spacings between every nth vehicle. (Throughout this paper the absence of the subscript n is identical to the n=1 case.) The probability that the spacing between every nth vehicle is greater than or equal to n will be denoted by $A_n(t)$. It is fortunate that the discrete counting distribution, $h_n(t)$, can always be expressed in terms of $A_n(t)$ and $A_{n+1}(t)$. Arguments formally developed by Feller (3) show that

$$h_n(t) = A_{n+1}(t) - A_n(t) \tag{18a}$$

Because this is a linear first-order difference equation in n,

$$A_{n+1}(t) = \sum_{i=0}^{n} h_i(t) \tag{18b}$$

provided one starts with the boundary condition $h_0(t) = A_1(t) = A(t)$. This last equation again points up a result that can be obtained by a direct line of argument: the probability that n or fewer vehicles are observed in t equals the probability that the spacing between n+1 vehicles is greater than or equal to t.

Further, the random variable that measures the spacing from the time origin to the nth vehicle is the sum of the spacing to the first vehicle plus the spacing between the first and the (n-1)th. The probability density distribution $a_n(t)$ is therefore the convolution of the probability that the first lies between r and r+dr with the probability that the spacing to the (n-1)th vehicle lies between (t-r) and (t-r+dt). Hence,

$$A_n(t) = \int_t^\infty \int_0^t a(r)\, a_{n-1}(t-r)\, dr\, dt \tag{19}$$

Equating the Laplace transform of $A_n(t)$,

$$\tilde{A}_n(s) = \int_0^\infty e^{-st} A_n(t)\, dt \tag{20a}$$

to the Laplace transform of the right-hand side of Eq. 19 gives

$$\tilde{A}_n(s) = \frac{1 - [\tilde{a}(s)]^n}{s} \tag{20b}$$

in which $\tilde{a}(s)$ is the Laplace transform of the intervehicle density distribution. When $\Delta = 0$ in Eq. 4,

$$\tilde{a}(s) = \int_0^\infty a(t)\, e^{-st}\, dt = \alpha + \frac{\mu(1-\alpha)^2}{s + \mu(1-\alpha)} \tag{21}$$

Expressing the nth power of $\tilde{a}(s)$ by the binomial expansion and making use of the transform pair,

$$f(t) = c^{-n} \sum_{j=n}^{\infty} \frac{e^{-ct}(ct)^j}{j!}$$

$$\tilde{f}(s) = \frac{1}{s}\, \frac{1}{(s+c)^n}$$

one can invert Eq. (20b):

$$A_n(t) = \sum_{i=1}^{n} \sum_{i=0}^{j-1} \binom{n}{j} (1-\alpha)^j \alpha^{n-j}\, \frac{[\mu t(1-\alpha)]^{i}}{i!}\, e^{-\mu t(1-\alpha)} \tag{22a}$$

$A_n(t)$ can be expressed in several alternate forms: (a) in terms of Laguerre polynomials of order (0)

$$A_n(t) = (1-\alpha)\,\alpha^{n-1} \sum_{j=0}^{n} \binom{n+1}{j} \left(\frac{\alpha-1}{\alpha}\right)^{n-j} \left(\frac{1}{\alpha}\right)^{j} e^{-\mu t(1-\alpha)}\, L_{n-j}^{(0)}\left[\mu t(1-\alpha)\right] \tag{22b}$$

or (b) in terms of an incomplete integral, of the Laguerre polynomial of order (1)

$$A_n(t) = (1-\alpha)\,\alpha^{n-1} \int_{\mu t(1-\alpha)}^{\infty} e^{-x}\, L_{n-1}^{(1)}\left[\frac{(\alpha-1)x}{\alpha}\right] dx \tag{22c}$$

Substitution of Eq. 22a into 18a and use of Eq. 11 and the identity

$$\binom{n+1}{j+1} - \binom{n}{j+1} = \binom{n}{j} \qquad j+1 \leq n$$

again leads to Eq. 17b.

COUNTING DISTRIBUTION

In this section the discrete counting distributions that correspond to the intervehicle distribution of Eq. 4 are derived. If $A_n^{\Delta}(t)$ is labeled as the probability distribution of intervehicle spacings greater than or equal to t when the minimum headway is Δ, and $\Delta_n(t)$ for the $\Delta = 0$ case, Eq. 4 can be rewritten

$$A^{\Delta}(t) = 1 \qquad 0 \leq t < \Delta \tag{23a}$$

$$= A(t-\Delta) \qquad \Delta \leq t \tag{23b}$$

It follows from the defining Eq. 19 for the distribution of spacings between nonadjacent vehicles that

$$A_n^{\Delta}(t) = 1 \qquad 0 \leq t < n\Delta \tag{24a}$$

$$= A_n(t - n\Delta) \qquad n\Delta \leq t \tag{24b}$$

Hence, the discrete counting distribution

$$p_n(t) = A_{n+1}^{\Delta}(t) - A_n^{\Delta}(t) \tag{24c}$$

which is obtained by substituting $p_n(t)$ for $h_n(t)$ and $A_n^{\Delta}(t)$ for $A_n(t)$ in Eq. 18a can be written in terms of $A_n(t)$:

$$p_n(t) = 0 \qquad 0 \leq t < n\Delta \qquad (25a)$$

$$= 1 - A_n(t - n\Delta) \qquad n\Delta \leq t < (n+1)\Delta \qquad (25b)$$

$$= A_{n+1}(t - (n+1)\Delta) - A_n(t - n\Delta) \qquad (n+1)\Delta \leq t \qquad (25c)$$

By substituting the translated versions of Eq. 22a into 24

$$A_n^{\Delta}(t) = \sum_{j=1}^{n} \sum_{i=0}^{j-1} \binom{n}{j} (1-\alpha)^j \alpha^{n-j} \frac{[\mu(1-\alpha)(t-n\Delta)]^i}{i!} e^{-\mu(1-\alpha)(t-n\Delta)} \qquad (26)$$

and

$$p_n(t) = 0 \qquad 0 \leq t < n\Delta \qquad (27a)$$

$$= 1 - \sum_{j=1}^{n} \sum_{i=0}^{j-1} \binom{n}{j} (1-\alpha)^j \alpha^{n-j} \frac{[\mu(1-\alpha)(t-n\Delta)]^i}{i!} e^{-\mu(1-\alpha)(t-n\Delta)}$$

$$n\Delta \leq t < (n+1)\Delta \qquad (27b)$$

$$= \sum_{j=1}^{n+1} \sum_{i=0}^{j-1} \binom{n+1}{j} (1-\alpha)^j \alpha^{n+1-j} \frac{[\mu(1-\alpha)(t-(n+1)\Delta)]^i}{i!} e^{-\mu(1-\alpha)[t-(n+1)\Delta]} -$$

$$\sum_{j=1}^{n} \sum_{i=0}^{j-1} \binom{n}{j} (1-\alpha)^j \alpha^{n-j} \frac{[\mu(1-\alpha)(t-n\Delta)]^i}{i!} e^{-\mu(1-\alpha)(t-n\Delta)}$$

$$(n+1)\Delta \leq t \qquad (27c)$$

There are, of course, many equivalent ways of writing these counting distributions. One of the simpler analytic forms, and possibly one that will be computationally useful, expresses the counting distribution in cumulative form:

$$P_n(t) = \sum_{j=0}^{n} P_j(t)$$

$P_n(t)$ is the probability that n or fewer vehicles are counted in the interval t. From Eqs. 18 and 24,

$$P_n(t) = A_{n+1}^{\Delta}(t) = A_{n+1}[t - (n+1)\Delta]$$

can be written in the form of Eq. 26. By making use of the incomplete gamma function notation,

$$\gamma(u;x) = \int_0^x t^{n-1} e^{-t} dt = \sum_{j=n}^{\infty} \epsilon^{-x} x^j$$

one can write $P_n(t)$ in the form

$$P_n(t) = \sum_{j=1}^{n+1} \sum_{i=0}^{j-1} \binom{n+1}{j} (1-\alpha)^j \alpha^{n+1-j} \frac{x^i}{i!} \epsilon^{-x}$$

$$= \sum_{k=0}^{n} \binom{n+1}{k} (1-\alpha)^{n+1-k} \alpha^k \frac{\gamma(n+1-k;x)}{(n-k)!} + \alpha^{n+1} \qquad (28)$$

in which $x = \mu(1-\alpha)(t-(n+1)\Delta)$. This alternate form for the counting distribution may be useful in view of the well-known properties of the incomplete gamma function and programs which are currently available for high-speed computation (30).

PROBABILITY OF MAXIMUM-PACK

The probability of maximum pack equals the probability that in an interval chosen at random the maximum number of vehicles are counted in that interval. The probability that exactly N vehicles are counted in an interval $t = N\Delta$, which begins just after the passing of a vehicle, equals α^N. This is the probability that a queue of N restrained vehicles follows the car that began the counting experiment. This probability falls off rapidly with small α. However, if one randomly picks the origin of an interval $N\Delta$ units long and asks for the probability that exactly N vehicles will be counted, a number which is larger than α^N is obtained because the first vehicle can occupy any position from the moment after counting begins to an instant just before the end of the first interval Δ. From the definition of the starting-at-random density distribution in Eq. 7 the probability of counting n vehicles in a randomly chosen interval of length t is

$$q_n(t) = \int_0^t u^{\Delta}(r)\, p_{n-1}(t - r)\, dr \tag{29}$$

Because $p_n(t)$ and $u^{\Delta}(t)$ can be expressed in terms of $A_n(t)$ for the $\Delta = 0$ case, one can rewrite $q_N(N\Delta)$, the probability that N vehicles will be observed in a randomly located interval exactly $N\Delta$ units long as

$$q_N(N\Delta) = \mu_\Delta \int_0^{\Delta} \left[1 - A_{N-1}(\Delta - t)\right] dt$$

The probability of maximum-pack is obtained by substituting Eq. 22a for $1 - A_{N-1}(\Delta - t)$:

$$q_N(N\Delta) = \frac{\mu\Delta}{1 + \mu\Delta} - \frac{1}{1 - \alpha + \beta} \sum_{j=1}^{N-1} \sum_{i=0}^{j-1} \sum_{k=i+1}^{\infty} \binom{N-1}{j} (1 - \alpha)^j \alpha^{N-1-j} \frac{e^{-\beta}\beta^k}{k!} \tag{30}$$

in which $\beta = \mu\Delta(1 - \alpha)$. When $N = 1$ one obtains the probability that a vehicle is observed in an interval equal to the minimum spacing:

$$q_1(\Delta) = \frac{\mu}{1 + \mu\Delta}$$

This is equal to the average flow rate of Eq. 6a divided by the maximum flow rate Δ^{-1} which would be observed if all vehicles were spaced regularly Δ units apart.

SUMMARY

Figure 1 shows the fraction of intervehicle headways greater than or equal to the values indicated on the horizontal scale. In Figure 1a the experimental data are denoted by the circled points and the solid line is a plot of the theoretical curve of Eq. 2 with parameters calculated from observed flow rates. These data were observed and fitted to Schuhl's distribution by Kell (12) in experiments that observed single-lane flow rates ranging from 150 to 1,200 vehicles per hour. In Figure 1b the theoretical curve is decomposed in two separate exponential terms; the large decay constant refers to the restrained vehicles and the smaller decay constant refers to the free-moving group.

Figure 2 shows the probability distribution $p_n(t)$ of Eq. 28 for several values of α when $t = 10\Delta$; i.e., if the minimum headway were $\Delta = 2$ sec, the period of observation would be 20 sec. In Case a, the probability of finding no restrained vehicle following a free-moving vehicle is 0.1, in Case b, 0.3 and in Case c, 1.0. Case c is also the counting distribution which corresponds to the translated exponential distribution discussed earlier by Newell (18) and Oliver (19).

In comparison to the Poisson law, these counting distributions point up at least two distinct effects. The first of these is the finite number of terms in the distribution. The probability of a count which exceeds the integral part of $t\Delta^{-1}$ is identically zero. For high flow rates and more particularly for average counts which are close to the integral part of $t\Delta^{-1}$ the difference between these and the Poisson distribution is pronounced.

A second feature is due to the effects of queueing. As the fraction of constrained

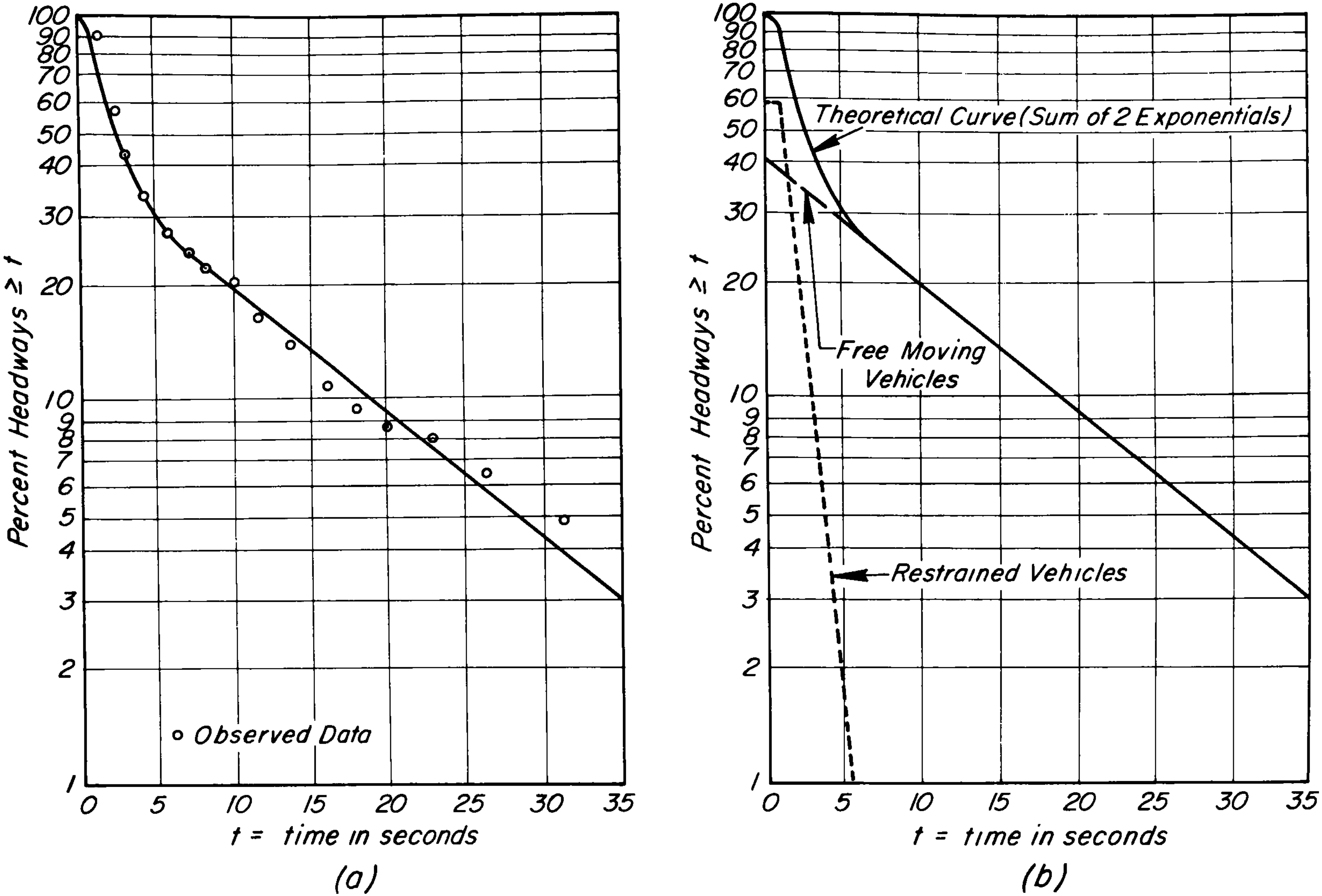

Figure 1. Distribution of intervehicle headways.

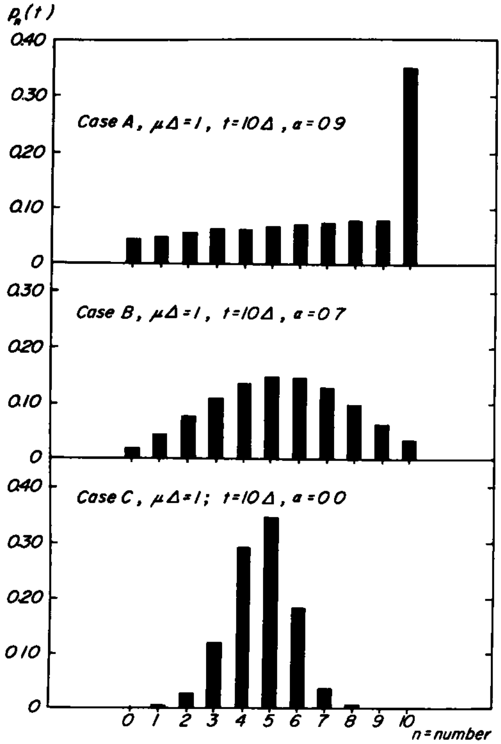

Figure 2. Counting distribution.

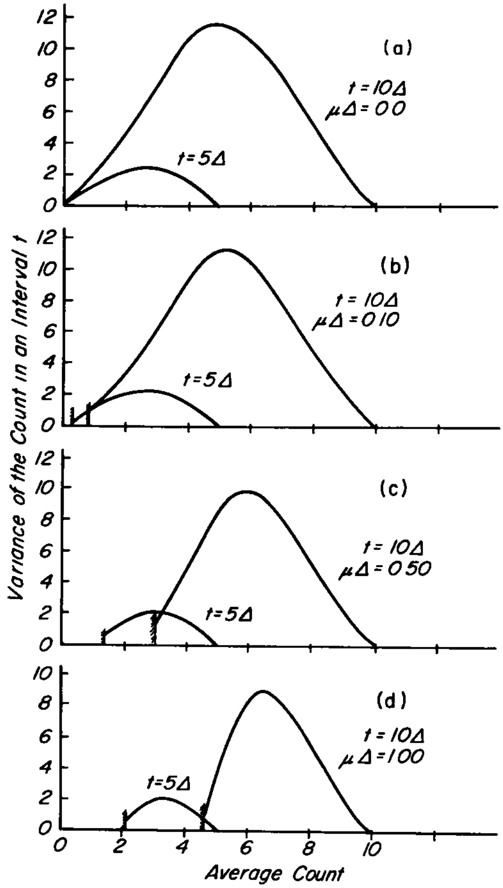

Figure 3. Variance vs average count.

vehicles increases, the average count in a small interval increases even though the stationary flow rate remains constant. This effect is due to the increase in the average size of the first bunch of restrained vehicles located at the counting origin.

Figure 3 shows the variance, Var n(t), as a function of the average count, $\bar{n}(t)$, for several values of the fraction of constrained vehicles and $t = 5\Delta$, 10Δ. The parameter $\mu\Delta$ rather than α is varied. With each curve, specification of a minimum headway, Δ, automatically specifies the counting interval, a value of μ, and the steady-state flow rate, $\mu/1 + \mu\Delta$. At least one author (16) has argued that in certain regions traffic counting distributions should have a larger variance than the Poisson due to the vehicles which concentrate in random queues. This feature can be observed in Figure 3 because a straight line Var n(t) $= \bar{n}(t)$ would intersect the peaked curves corresponding to $t = 10\Delta$. In those cases where small average counts are observed, the variance approaches that of the Poisson distribution. For average counts close to the integral part of $t\Delta^{-1}$ the variance decreases because the probability of maximum-pack increases. In the limit, this probability equals one, vehicles are spaced regularly Δ units apart, and the variance is zero. The slanted lines in Figure 3 indicate that it is not possible to obtain certain average counts for the given values of μ and Δ. This should not be interpreted, however, as a statement that small average counts cannot be observed in t, but rather that the probability of queueing must lie below certain values if small average counts are observed.

REFERENCES

1. Adams, W. F., "Traffic Considered as a Random Series." Jour., Inst. Civil Eng., vol. 4 (1936).

2. Bartlett, M. S., "Some Problems Associated with Random Velocity." Publ., L'Institut de Statistique de Paris. vol. 6, fasc. 4 (1957).
3. Feller, W., "On Probability Problems in the Theory of Counters Studies and Essays Presented to R. Courant." Interscience (1948).
4. Feller, W., "An Introduction to Probability Theory and Its Applications." 2nd ed., Wiley (1958).
5. Galliher, H. P., "Ordnance Logistics Studies—II." Interim Technical Report 9, Project in Fundamental Investigations in Methods of Operations Research, Mass. Inst. Technol. (Aug. 1958).
6. Gerlough, D. L., "Use of Poisson Distribution in Highway Traffic." In "Poisson and Traffic." Eno Foundation (1955).
7. Greenshields, B. D., "Application of Statistical Sampling Methods to Traffic Performance at Urban Intersections." HRB Proc., 26:377-389 (1946).
8. Haight, F., "Towards a Unified Theory of Road Traffic." Operations Res. 6:No. 6 (1958).
9. Haight, F., "The Generalized Poisson Distribution." Ann. Inst. Statist. Math., Tokyo, 11:No. 2 (1959).
10. Haight, F., Whisler, B. F., and Mosher, W. W., "New Statistical Method for Describing Highway Distribution of Cars." HRB Proc., 40:557-564 (1961).
11. Jewell, W. S., "The Properties of Recurrent-Event Processes." Operations Res. 8:No. 4 (1960).
12. Kell, J. H., "A Theory of Traffic Flow on Urban Streets." Proc. Western Section, Inst. Traffic Eng., vol. 13, (1960).
13. Kinzbruner, P. K., "A Statistical Analysis of Vehicular Traffic." M. Sc. thesis, Department of Electrical Engineering, M.I.T. (1960).
14. May, A. D., Jr., and Wagner, F. A., Jr., "Headway Characteristics and Interrelationships of Fundamental Characteristics of Traffic Flow." HRB Proc., 39:524-527 (1960).
15. Miller, A. J., "Traffic Flow Treated as a Stochastic Process." In "Theory of Traffic Flow." Elsevier (1961).
16. Miller, A. J., "A Queueing Model for Road Traffic Flow," J. Royal Stat. Soc., Series B, 23:No. 1 (1961).
17. Morse, P. M., "Queues, Inventories and Maintenance." Wiley (1958).
18. Newell, G. F., "Statistical Analysis of the Flow of Highway Traffic Through a Signalized Intersection." Quart. App. Math., 13:4 (1956).
19. Oliver, R. M., "A Traffic Counting Distribution." Operations Res. 9:No. 6(1961).
20. Oliver, R. M., "A Note on a Traffic Counting Distribution." To appear in Operations Res. Quart. (1962).
21. Petigny, B., "Le Calcul des Probabilities et la Circulation des Vehicules sur les Chaussees a Deux ou Trois Voies." Ann. Ponts et Chaussees, 131:2 (1961).
22. Schuhl, A., "Le Calcul des Probabilites et la Repartition des Vehicules sur les Routes a Deux Voies de Circulation." Travaux (Paris), 39:No. 243 (1955).
23. Schuhl, A., "Repartition des Espacements de Temps Entre Vehicules Successifs sur Une Route a Deux Voies." Travaux(Paris) 39:No. 249 (1955).
24. Schuhl, A., "The Probability Theory Applied to Distribution of Vehicles on Two-Lane Highways." In "Poisson and Traffic." Eno Foundation (1955).
25. Takacs, L., "Stochastic Processes." Wiley (1960).
26. Tanner, J. C., "A Problem of Interference Between Two Queues." Biometrika, vol. 40 (June 1953).
27. Tanner, J. C., "A Simplified Model for Delays in Overtaking on a Two-Lane Road." J. Royal Stat. Soc., Series B, vol. 20 (1958).
28. Tanner, J. C., "Delays on a Two-Lane Road." J. Royal Stat. Soc., Series B, vol. 23 (1961).
29. Weiss, G. H., and Maradudin, A. A., "Some Problems in Traffic Delay." Operations Research, 10:No. 1 (1962).
30. Whittlesey, J. R. B., and Haight, F. A., "Counting Distributions for an Erland Process." To appear Ann. Inst. Statist. Math., Tokyo (1961).

Analyzing Vehicular Delay at Intersections Through Simulation

JAMES H. KELL, Assistant Research Engineer, Institute of Transportation and Traffic Engineering, University of California, Berkeley

The first section of this paper describes the development of a simulation model for the intersection of two 2-lane two-directional streets, with one street being controlled by stop signs. The lack of adequate mathematical distributions describing traffic behavior and the field studies performed to obtain these distributions are discussed. The elaborate techniques used to test the logic of the model before beginning the analysis are also reviewed.

The second portion of the paper presents the simulation results. The variability of vehicular delay under constant traffic conditions is described and the relationships between vehicular delay and individual approach volumes and turning movements are formulated.

A brief discussion of the value of this research and the general applicability of simulation techniques to solving traffic problems is also presented.

• A GREAT MANY traffic engineering decisions are made on the basis of experience and engineering judgment at the present time. In this growing profession, there is an increasing need for factual information concerning the effect of these decisions. The simulation of traffic situations on high-speed digital computers can be used to provide a large amount of data under controlled laboratory conditions which would be difficult, if not impossible, to obtain through field studies. The techniques and methods of vehicular simulation on computers are relatively new, and the lack of basic mathematical distributions describing traffic behavior has slowed progress. As these techniques are established, simulation can become one of the more valuable research tools available to the profession.

Almost any traffic situation is capable of simulation and, as techniques improve, can be rapidly programed. Variables or controls can be changed and their effects analyzed. Before-and-after studies can be performed in hours or days without disturbing traffic in the field. Situations can be simulated and observed which could not be risked in field installations. Peak traffic flows can be simulated for hundreds of hours under the precise conditions desired instead of obtaining one or two hours of data per day in the field under uncontrolled conditions.

One of the most important problems facing traffic engineers is congestion on city streets. There has been much discussion in the last few years concerning network analysis of street systems to provide maximum efficiency in moving traffic. This certainly is an ultimate objective and will use computers to solve the problem. But, congestion begins primarily at intersections. A network analysis combines a great number of intersections and studies the over-all operation of the system. To do this, even on large computers, the individual intersection must be analyzed in a macroscopic manner. In other words, only the significant features of intersection operation can be included in the analysis if it is to be accomplished in a reasonable time. To determine which operational features are significant, it is necessary to study intersection operation in fine detail or microscopically. The microscopic simulation of vehicular traffic at intersections is the subject of this paper.

The current research in simulation at the Institute of Transportation and Traffic Engineering is divided into two phases—intersections controlled by stop signs and intersections controlled by traffic signals. The first phase is nearly complete and the second phase has begun. Numerous objectives are to be obtained from this research project, including the following:

1. Development of "models" by which vehicular traffic at intersections can be simulated on high-speed computers.
2. Determination of the total vehicular delay experienced at intersections with respect to approach volumes and turning movements. In the case of signalized intersections, signal timing is the third major variable.
3. Evaluation of the effect of installing a signal at an intersection on vehicular delay which will provide a basis for examining and refining existing traffic signal warrants.
4. Determination of the effect of turning movement restrictions (and signal timing in the second phase of the research) on intersection operation.

STOP INTERSECTION

The programing for this phase of the research is essentially complete. The model described was coded for an IBM 701 computer available at the University of California. Unfortunately, this computer has been out of operation since July 1, 1961. Programs coded for the 701 computer are not readily translated for use on later generation computers and, due to the complexity of this simulation program, no attempt has been made to recode the model for another computer. The 701 is expected to be in operation by January 1962. This phase of the research has, therefore, been tabled temporarily pending the availability of the computer.

Model Description

The programed model for the first phase of this project consists of the time simulation of an orthogonal intersection of two two-lane, two-way streets with the minor street being controlled by stop signs. The model (shown in Fig. 1) includes the approaches to the intersection for sufficient length that vehicles enter the system before they are influenced by any condition existing at the intersection.

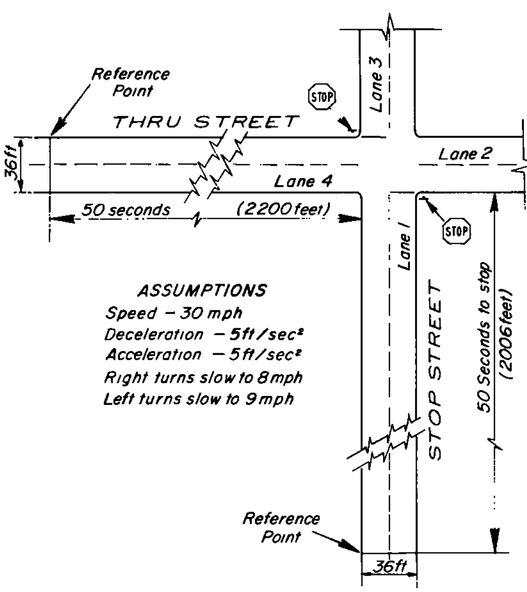

Figure 1. Simulation model for stop sign intersection.

Vehicles entering an approach are generated randomly by Monte Carlo techniques from a given distribution at a preselected hourly volume. Each of the four approaches has a preselected volume and is generated independently. As a vehicle is generated, it is randomly assigned a turning movement (right, left, or thru) based on a requested distribution of these movements. Vehicles travel from the point of generation to the intersection (or to a point where they must decelerate) at a predetermined velocity.

Minor street vehicles decelerate to a stop either at the intersection or in queue (in which case they move forward as the queue is released). A vehicle at the intersection accepts or rejects available gaps in conflicting traffic streams from given acceptance distributions based on their respective turning movements. On accepting a gap, the minor street vehicle accelerates to recovery driving velocity.

Major street vehicles are given the right-of-way over minor street vehicles when conflicts in turning movements and/or time exist. Through major street vehicles are also given the right-of-way over opposing left-turning major street vehicles when conflicts in time occur. Major street queues, due to stopped left-turning vehicles, and delays caused by vehicles slowing to make turns, are also included in the model.

Briefly, the simulation is accomplished as follows. Each time a major street vehicle enters the intersection, the model is analyzed. If the major street vehicle is not delayed, the minor street traffic is brought up to this time (vehicles delayed or released as appropriate) and the major street vehicle is released. If the major street vehicle is delayed, the minor street traffic is also delayed and the system is checked to see when the major street vehicle might be released. This process is repetitive, generating new traffic as necessary.

Acceleration, deceleration, slowing, stopping, and queueing delays are computed for both major and minor street vehicles. Stopped time delays for minor street vehicles while waiting for an acceptable gap are also computed. All delays are accumulated and stored for each simulated hour of real time.

At the termination of each simulation run, the results of each simulated hour are printed and include the following items for each approach, for the minor street, for the major street, and for the entire intersection: elapsed real time, vehicular volume entering the system, vehicular volume released from the system, turning movement counts, turning movement percentages, number of vehicles delayed, percent of vehicles delayed, total vehicular delay, average delay per vehicle, average delay per delayed vehicle, and maximum queues experienced.

Vehicle Generation

As mentioned earlier, vehicles are generated randomly from a given distribution for each approach independently. One of the first problems encountered in the development of the simulation model was the absence of a satisfactory mathematical distribution describing gaps or headways in a traffic stream. The Poisson distribution which has been used in the analysis of other traffic problems does not adequately describe the headway distribution. A number of theoretical distributions have been proposed by various authors but no information could be found to indicate that these distributions had been tested over an extensive volume range or that the parameters had been solved in terms of volume. It was necessary, therefore, to select a distribution, test it, and solve for its parameters.

A composite distribution proposed by Andre Schuhl (1) was selected. He theorized that a traffic stream is divided into two groups. A certain proportion of the vehicles in the stream travel as they wish and are not influenced by the vehicle in front of them. For convenience, this group shall be referred to as the free-flowing vehicles. The remaining vehicles have been influenced by the vehicle in front of them and shall be called the restrained vehicles. Each of these groups has a distinct mean and obeys some Poisson-type law. The theoretical distribution for the total stream is a composite or summation of these two subdistributions. Figure 2 shows this composite distribution along with the two individual curves which have been summed. Figure 3 shows the same curves replotted so that the ordinate p is the probability of a headway or gap that is less than or equal to the time t indicated.

The equation for this composite distribution is

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

in which

$p(h \geq t)$ = probability of a headway (h) greater than or equal to the time (t);
α = proportion of the traffic stream in restrained group;
$(1 - \alpha)$ = proportion of traffic stream in free-moving group;
T_1 = average headway of free-moving vehicles;
T_2 = average headway of restrained vehicles;
λ = minimum headway of free-moving vehicles;

τ = minimum headway of restrained vehicles; and
e = natural or Naperian base of logarithms.

In this equation, there are five parameters—α, T_1, T_2, λ, and τ—which are functions of the traffic volume. These can be reduced to four unknowns by transforming Eq. 1 to give

$$p(h \geqq t) = e^{a - \frac{t}{K_1}} + e^{c - \frac{t}{K_2}} \tag{2}$$

in which

$$a = \frac{\lambda}{T_1 - \lambda} + \ln(1 - \alpha) \tag{3}$$

$$K_1 = T_1 - \lambda \tag{4}$$

$$c = \frac{\tau}{T_2 - \tau} + \ln \alpha \tag{5}$$

$$K_2 = T_2 - \tau \tag{6}$$

The problem, therefore, is to find these four unknowns in terms of volume. Fortunately, Eq. 6 still describes the two subdistributions separately. Returning to the plot of the composite curve (Fig. 2), the restrained subdistribution does not affect the composite curve for larger headways. Therefore, the unknowns corresponding to the free-moving subdistribution can be determined by fitting an exponential curve to the equivalent portion of field data. Once these two unknowns are determined, the contribution of the free-moving vehicles in the lower portion of the curve can be calculated and subtracted from the original data. The residuals form the distribution of the

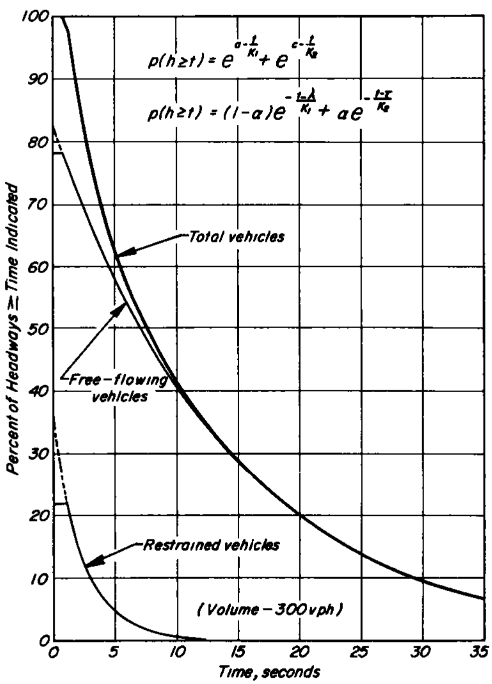

Figure 2. Composite exponential distribution.

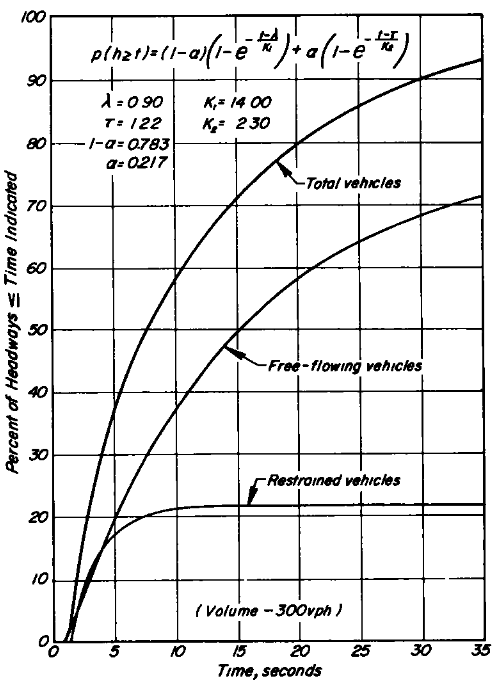

Figure 3. Composite exponential distribution.

restrained vehicles and, therefore, the remaining unknowns can be determined by fitting a second exponential curve to these points.

To evaluate these unknowns, extensive field data were collected. These data, obtained on two-lane urban streets, resulted in 585 samples with volume rates ranging from a little over 100 vph to almost 1,200 vph. Eighteen different fits were computed and tested by means of χ^2 tests for each sample.

To indicate the magnitude of the computations involved, a man with a desk calculator and a book of log tables requires from 12 to 14 hr to compute the values for one data sample. Needless to say, these computations were done on a computer which performed the same operations in approximately 18 min.

Equations for the unknowns in Eq. 2 have been made:

$$K_1 = \frac{4827.9}{V^{1.024}} = e^{8.48 - 0.024 \ln V} \tag{7}$$

$$a = -0.046 - 0.0448 \frac{V}{100} \tag{8}$$

$$K_2 = 2.659 - 0.120 \frac{V}{100} \tag{9}$$

$$c = \left[e^{-10.503 + 2.829 \ln V - 0.173 (\ln V)^2} \right] -2 \tag{10}$$

Whereas the parameters of Eq. 2 have been solved in terms of volume, the simulation model requires the use of Eq. 1. It is impossible to transform Eq. 2 back directly to the form of Eq. 1 because α, λ, and τ are dependent on one another. The assignment of a value to one of these three parameters determines the other two. The equations of these parameters are

$$\lambda = K_1 \left[a - \ln(1 - \alpha) \right] \tag{11}$$

$$\tau = K_2 (c - \ln \alpha) \tag{12}$$

and

$$1 - \alpha = e^{a - \frac{\lambda}{K_1}} \tag{13}$$

or

$$\alpha = e^{c - \frac{\tau}{K_2}} \tag{14}$$

λ and τ are the minimum headways of the subdistributions and, because negative headways are impossible, as these minimum headways approach zero, they define the two boundary conditions. In fact, these parameters cannot even approach zero because vehicles have a finite length. A more realistic minimum headway is 0.5 sec which, at 30 mph, is a distance of 22 ft between the front bumper of the lead car and the bumper of the following car.

Figure 4 is a plot of the two conditions where λ and τ equal 0.5 sec. When t is greater than the largest λ or τ, the summation curve is identical. The only problem is the shaded area on Figure 4 between λ and τ or τ and λ as the case may be. To solve this problem, the original field data were analyzed by first grouping the data into volume ranges and then examining the leading portion of the cumulative curve for each volume range. It became apparent that λ and τ were relatively constant throughout the volume range. Best agreement between the theoretical curve and the observed data occurred where $0.9 < \lambda < 1.0$ and $1.20 < \tau < 1.36$.

Once the parameters of this distribution are determined for a given volume, the next step is to generate headways that fit this distribution randomly. This is accomplished by a technique similar to that described in an earlier paper by Gerlough (2). A flow diagram of the random generator used in this simulation model is shown in Figure 5. A separate random generator is used for each approach.

Using these random generators to generate many consecutive hours of traffic at a particular volume yields a distribution of volumes that has some spread on either side of the requested volume. This spread becomes larger as the volume increases. Because

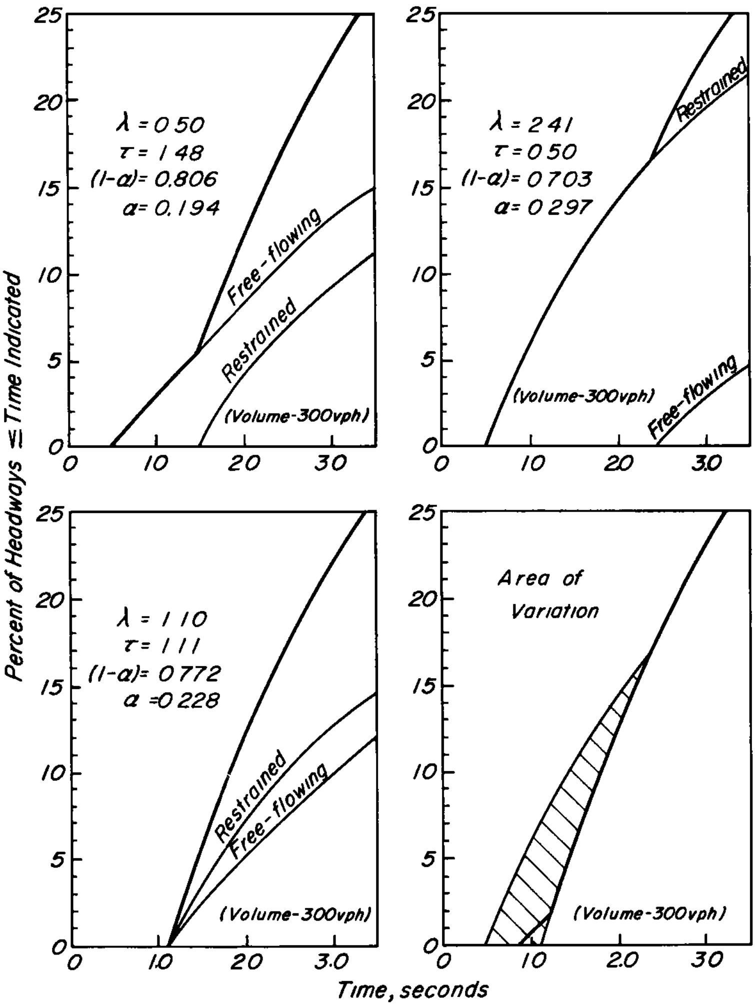

Figure 4. Effect of λ, τ, and α on composite distribution.

of this variation in generated volumes, the analysis of the simulation output becomes more complex. This problem can be overcome if the random number used to begin each hour of simulation is known to produce a volume within a small tolerance of the desired volume. To obtain random numbers that would produce the desired volumes, a separate computer program was developed, incorporating the same random generator as used in the simulation, to pregenerate vehicular volumes. As each hour is generated, the generated volume is compared to the desired volume. If the generated volume is within the tolerance (presently +2 percent and - 1 percent), the generated distribution is tested against the theoretical distribution by means of the Komolgorov-Smirnov test of goodness of fit (3). When this test is successful, the computer prints the random number used to start the hour of generation, the time into the next hour of the last vehicle, and the time preceding the end of the hour of the last twelve vehicles. These vehicle times are used in ordering the random numbers for the simulation to ensure continuity and volume accuracy. Essentially, the time over the hour of the last vehicle in the preceding hour is an offset for the new hour of generation. This may result in the loss of one or more vehicles in the new hour of generation (i.e., the simulation hour may end before all vehicles in the expected volume are generated). This is the reason

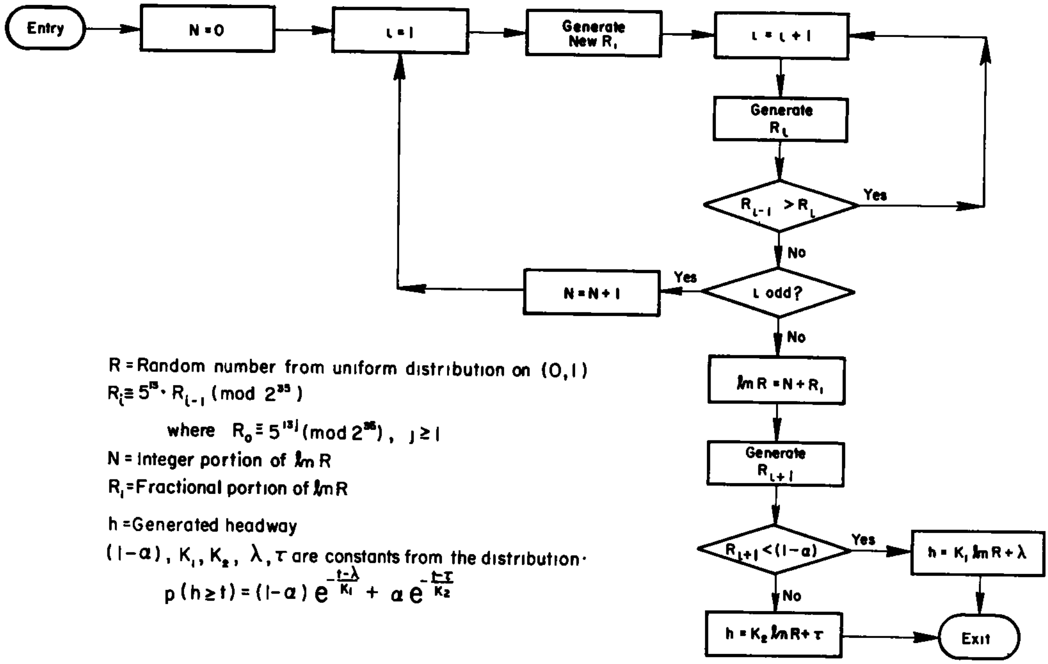

Figure 5. Random headway generator using Monte Carlo technique.

for using a larger plus than minus tolerance in the generation. Sufficient random numbers for a particular volume are ordered and used as simulation input to provide the desired number of hours of simulation.

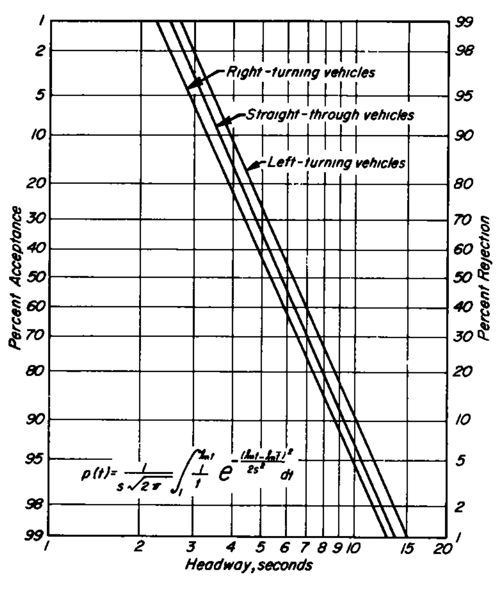

Figure 6. Gap acceptance distributions for vehicles passing a stop sign.

Gap Acceptance

Vehicles waiting at the stop sign in the simulation model accept or reject available gaps in conflicting traffic streams based on a gap acceptance distribution for their respective turning movement. These gap acceptance distributions had to be determined from field data. Howard Bissell, a graduate student at the Institute, undertook this study as a graduate research project (4). He collected and analyzed data for over 10,000 gaps and found that the gap acceptance distribution was of a lognormal form. The distributions found are shown in Figure 6.

The validity of using a lognormal distribution to describe gap acceptance was confirmed when this distribution was fitted to some data obtained in an independent study in Australia. The fit was exceptionally good. The only difference between the United States' data and the Australian data was the mean gap accepted. Apparently, Australian drivers accept shorter gaps on the average than drivers in this country.

The time required by the computer to

generate individual values of a lognormal distribution was considered excessive. Therefore, the distributions are inserted in the simulation program in tabular form and a table-look-up procedure is used to determine gap acceptance. This is accomplished in the following manner: A gap of size x is available. The probability of accepting a gap of size x is obtained from the table. A random number (between 0 and 1) is generated and compared with the probability value. If the random number is larger, the gap is rejected. Otherwise, it is accepted. With this method, there is a possibility of a vehicle rejecting a gap of a certain size and then accepting a shorter gap later. Bissell (4) found that this occurred for approximately 5 percent of the vehicles observed.

Simulation Procedure

To utilize the simulation model to provide data (in this case, vehicular delay), a number of simulation "runs" were performed. A simulation run consists of simulating a predetermined number of hours of traffic under constant conditions. After the model has been programed and tested on the computer, the first step in any simulation project is to determine the number of simulated hours under constant conditions that are required to obtain valid estimates of the output factors. In this project, vehicular delay is being measured. Sufficient hours of traffic are simulated with constant approach volumes and turning movements to analyze the delay pattern and determine the variance that occurs. On the basis of this analysis, the length of a simulation run is determined.

A series of runs are then established to begin the analysis of vehicular delay. The approach volumes used in the first 15 runs of the first series are given in Table 1. Turning movements from all four approaches are held constant with left and right turns being 10 percent each. The first five runs of this series have approach volumes which correspond to the minimum vehicular volume warrant for traffic signal installation (5).

An evaluation of the results of this first series of runs provides an insight into the relationship between delay and approach volumes and indicates the additional runs necessary to obtain sufficient data to define this relationship.

TABLE 1

SIMULATION RUNS, SERIES 1

Run Number	Approach Volume[a]						Total Volume
	Major Street			Minor Street			
	Lane 2	Lane 4	Total	Lane 1	Lane 3	Total	
1	250	250	500	150	100	250	750
2	300	200	500	150	100	250	750
3	200	300	500	150	100	250	750
4	400	100	500	150	100	250	750
5	100	400	500	150	100	250	750
6	200	200	400	150	100	250	650
7	300	100	400	150	100	250	650
8	300	300	600	150	100	250	850
9	400	200	600	150	100	250	850
10	500	100	600	150	100	250	850
11	400	200	600	200	100	300	900
12	400	200	600	100	200	300	900
13	400	200	600	200	200	400	1,000
14	400	200	600	300	100	400	1,000
15	400	200	600	400	100	500	1,000

[a]Turning movements in this series are constant with 10 percent right turns and 10 percent left turns.

After the volume relationship is determined, a further series of runs is made to evaluate the effect of turning movements. Approach volumes are held constant at various volume levels for both the major and minor streets while the percentage of turns is varied. The results of the initial runs, when analyzed, determine areas where further data are needed. This defines the runs to be included in the next series.

This procedure is continued until the various relationships under investigation are determined for the range of variables desired.

SIGNAL INTERSECTION

Programing of the second phase of this research is underway. The completed program will be coded in computer language for use on any of the later generation computers to avoid any repetition of delays experienced in the first phase of this project. A grant of $3,250 has been received from the Technical Development and Research Fund of the Institute of Traffic Engineers to support this project. These monies are to defray the cost of computer time for simulating the signalized intersection.

Model Description

The signalized model is quite similar to the first model. Physical conditions are the same—an orthogonal intersection of two two-lane, two-way streets. Vehicles are generated in the same manner as previously described. In programing the second phase, care is being taken to insure that the exact same traffic generated in the stop sign model can be reproduced in the signalized model.

Intersection control in this model is programed as a separate subroutine. There will be at least three different types of signal controllers programed as subroutines—fixed-time, semi-actuated, and full-actuated. During any simulation run, the model is preset to select the proper subroutine for the type of signal being simulated.

The signal sequence is main street (lanes 2 and 4) green, main street amber, all red, cross-street green, cross-street amber, all red. The all-red intervals will normally be set equal to zero. They have been included in the model to permit the analysis of all-red clearance periods and also to allow the inclusion and evaluation of separate pedestrian phases (scramble system). It may also be possible to include additional phases for left-turning vehicles during these intervals.

An optional feature to be incorporated into the program is the ability of vehicles waiting on a red phase to make right turns. This maneuver is legal in several western States and its inclusion in the model will permit an evaluation of its effect on vehicular delay. Vehicles permitted to make this turn will accept gaps in the cross-traffic in a similar manner as right-turning vehicles at a stop sign.

Operation of the model is quite similar to the through street in the stop sign model. Vehicles are generated at the reference point and approach the intersection. If no barrier (red signal, queue, or turning vehicle) impedes its progress, the vehicle continues through the intersection. Otherwise, it decelerates and stops, if necessary, until such time as it can proceed. Left-turning vehicles accept gaps in the opposing stream based on a gap-acceptance distribution. If left-turning vehicles are waiting as the signal changes, a maximum of two will be permitted to turn after the signal change.

Each street will be analyzed in detail during its green and amber phase. Vehicles are generated and brought to the intersection until the arrival time of the next vehicle is later than the end of the amber phase. In the case of an actuated signal, this last vehicle is used to "call the green" back to that phase. The cross-street is then brought "up-to-date" through the preceding red phase and the current green and amber phases. At the conclusion of each simulated hour, results are tabulated and stored until the completion of the run, when they are printed. The same items are obtained in this output as were obtained in the stop sign simulation to facilitate analysis through direct comparisons.

Vehicular Approach

As previously stated, vehicles are generated at the reference point and proceed

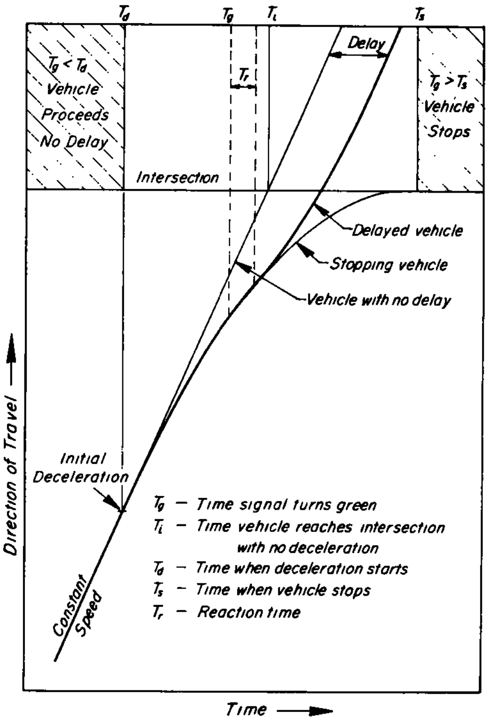

Figure 7. Vehicle approacning signal.

through the intersection unless a barrier impedes their progress. When a prior vehicle is already stopped in the approach, there is no question—the approaching vehicle stops in queue behind the stopped vehicle. Neither is there any question when the signal is red as the vehicle approaches and remains red until the vehicle stops, nor when the signal is green and remains green until the vehicle enters the intersection. But, there are two situations where an approaching vehicle's behavior is varied: (a) where the signal changes from red to green as an approaching vehicle is decelerating, and (b) where the signal changes from green to amber before an approaching vehicle reaches the intersection.

The first case is fairly simple to resolve. As a vehicle approaches an intersection where a red signal is displayed, there is some point (in distance and time) where the vehicle begins to decelerate to a stop. If the signal turns green before this point (in time), then no deceleration occurs. Similarly, if the signal does not turn green before the vehicle arrives at the intersection, the vehicle stops. Therefore, two boundary conditions have been defined. If a vehicle is between these two boundaries when the light turns green, he

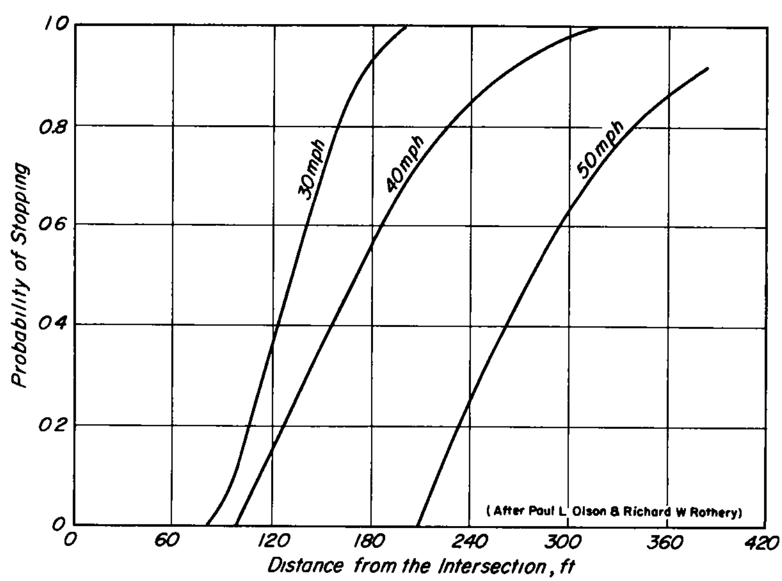

Figure 8. Stopping distributions for vehicles approaching amber signal.

is decelerating but has not stopped. After reacting to the light change, he accelerates and returns to normal driving speed. This critical interval of a vehicle's approach is shown in Figure 7. If the signal turns green during this interval and there is no other delaying factor, the vehicle proceeds without stopping and has a slowing delay proportional to the time spent in deceleration.

The signal change from green to amber (or yellow) is more complex. In this case, only one boundary condition is fixed—when the amber light appears after the vehicle arrives at the intersection (in which case, there is no delay). But, if the light changes to amber as a vehicle approaches, then the driver must decide whether he is going to stop or not. To simulate this in the model, a probability distribution must be used. Initially, a distribution curve presented by Olson and Rothery (6) will be used (see Figure 8). These curves are based on data from five intersections in Michigan. Additional data will be gathered in California to refine the distribution. Also, an attempt will be made to evaluate the difference in the legal definition of the yellow light. In California, the yellow is a "warning" period, whereas in most other States it is a "clearance" period.* It may be necessary to use two different distributions.

The procedure is similar to that using gap acceptance distributions except that instead of using a gap to enter the probability table, the distance the vehicle is from the intersection is used. A random number (between 0 and 1) is generated and compared with the probability value from the table. If the random number is greater, the vehicle proceeds; if less, it stops.

CONCLUDING REMARKS

Originally, this paper was to include results from the first phase of this research. The unfortunate shutdown and continued lack of operation of the particular computer for which the program was coded has prevented this. However, the techniques presented should be of value to other researchers concerned with intersection simulation.

A generation routine has been provided which will result in a realistic and random distribution of traffic that can be reproduced as often as required. A technique has been described by which it is possible to control the volume per hour to values within a small tolerance of the requested volume in order to simplify analysis of the simulation output. Gap acceptance distributions based on field data have been presented along with a technique for utilizing the distributions.

Two intersection models and their general operation have been described to illustrate some of the possibilities that are available through the use of simulation. As simulation techniques improve and mathematical distributions become available, almost any situation will be capable of simulation. The potential of this tool in evaluating the efficiency of any system of control is practically unlimited.

REFERENCES

1. Schuhl, A., "The Probability Theory Applied to Distribution of Vehicles on Two-Lane Highways." In "Poisson and Traffic." Eno Foundation for Highway Traffic Control, 59-75 (1955).
2. Gerlough, D. L., "Traffic Inputs for Simulation on a Digital Computer." HRB Proc., 38:480-492 (1959).
3. Massey, F. J., Jr., "The Kolmogorov-Smirnov Test for Goodness of Fit." Jour. Amer. Statistical Assoc., 46:68 (1951).

*Sec 11-202(b) of the Uniform Vehicle Code defines the yellow light ("steady yellow alone") and reads in part:

> "1. Vehicular traffic facing the signal is thereby warned that the red or 'Stop' signal will be exhibited immediately thereafter and such vehicular traffic shall not enter or be crossing the intersection when the red or 'Stop' signal is exhibited." (emphasis added)

Sec. 21452(a) of the California Vehicle Code has the same provision except that the emphasized portion is replaced by "vehicular traffic will be required to stop."

4. 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).
5. "Manual on Uniform Traffic Control Devices for Streets and Highways." U. S. Department of Commerce, Bureau of Public Roads (1961).
6. Olson, P. L., and Rothery, R. W., "Driver Response to the Amber Phase of Traffic Signals." Operations Res. 9:650-663 (Sept. - Oct. 1961).

Computer Simulation of Traffic on Nine Blocks of a City Street

MARTIN C. STARK, National Bureau of Standards, Washington, D. C.

A computer model has been constructed which simulates the volume and movement of traffic on a nine-block section of a city street. The simulated cars are reviewed every quarter-second and are moved according to rules for movement which have been built into the computer program. The simulation run on the computer produces two outputs. The quarter-second car positions are plotted on an oscilloscope and photographed. The result is a moving picture which can be shown in real time. The effect is comparable to viewing the traffic flow from a helicopter. The other output is a series of tables that catalogs all vehicles as they enter the model, clock and count them as they pass a key intermediate point, and, finally, check them out at the end of the course, counting them again and noting their individual running times. Other information is also furnished, such as type of vehicle, speed, and lane use. The tables thus furnish an abundance of quantitative data for measuring and evaluating the performance of the model.

• THIS is a report on a digital computer simulation of automobile traffic at an existing location on city streets. The study was made for the Bureau of Public Roads by the Data Processing Systems Division, National Bureau of Standards, over a period of three years from July 1958 to June 1961.

PURPOSE OF STUDY

The purpose of the work was to simulate the volume and movement of cars with a digital computer, using as the test site a real location where abundant field data were available for control and checking purposes. The test course selected was a nine-block section of 13th Street, N. W., Washington, D. C., from Euclid Street to Monroe Street, in the afternoon rush-hour period when all four lanes are operated one-way northbound (see Fig. 1).

A standard computer-simulation technique involving the use of random numbers "generates" cars at each entering lane in such a manner that the total number entering at each point over a period of time has an assigned expected value. Cars are moved each quarter-second according to detailed "rules of the road" built into the computer program.

Successive car positions have been plotted on an oscilloscope and photographs taken so that the simulated operation can be viewed from moving pictures. Printout tables furnish detailed quantitative data about the volumes, running times, and characteristics of the cars involved.

To the extent that the simulated model can be made to reproduce the known real conditions truly, the volumes and characteristics of traffic and the operating rules then can be changed and the results of a run will represent a prediction of what would happen on the street if the indicated changes were really made. The immediate area of application relates to the use and timing of traffic signals. Simulation runs can be made to study the sensitivity of the traffic flow to altered signal settings, to measure the effect of changed offsets, cycle length, and splits with a view to arriving at optimal timing and to explore the capacity of the signal system to handle increased volumes of traffic. The use of a generalized model can be extended to many other traffic engineering situations.

ANALYSIS OF RESULTS

Several simulation runs have been made. One 4-min real-time run (three complete 80-sec signal cycles) has been selected and is the basis for most of the detail that is to follow.

A moving picture was made of the oscilloscope display. The computer issued printout sheets furnishing detailed numerical data to permit analysis of the behavior of the simulated cars. From this information three summary tables were made.

Tables 1 and 2 summarize the count of cars generated during each cycle for each of the entry points. The figures are shown by cycle, summed for the 4 min, and expanded to an hourly volume. These tables represent the traffic "inputs", whereas Tables 3 and 4 represent "outputs" and are the means for measuring the performance of the simulated cars.

Table 3, the Station B count, gives the results of a count about two-thirds of the way along the 13th Street course, at the maximum load point. In the model, the expansion of the 4-min count at Station B produced an hourly volume of 3,330 simulated cars. This is comparable to the average field volume of 3,168 cars per hour.

Table 4, the vehicle retirement table, relates to cars leaving the end of the 13th Street test course north of Monroe Street. Again, the simulated cars are counted by cycles and by lanes, and are expanded to an hourly rate. Table 3 also records the running times of those cars which have traversed the full length of the test course from Euclid Street.

The distribution of running times of the simulated cars showed that some of the

TABLE 1

SUMMARY OF CARS GENERATED BY STREET AND BY HOURLY RATE[a]

Street	Hourly Rate (no.)			
	Cycle 1	Cycle 2	Cycle 3	Total Period
13th	3,060	3,105	2,565	2,910
Fairmont	90	90	45	75
Girard	45	90	90	75
Harvard	270	630	90	330
Columbia	540	630	630	600
Irving	360	360	180	300
Kenyon	450	180	90	240
Lamont	0	90	0	30
Park	1,035	630	675	780
Monroe	135	315	45	165

[a]Signal cycle length is 80 sec. Cycle 1 represents 320 quarter-seconds from simulation-run time 1,036 to 1,355; cycle 2, from time 1,356 to 1,675; and cycle 3, from time 1,676 to 1,995.

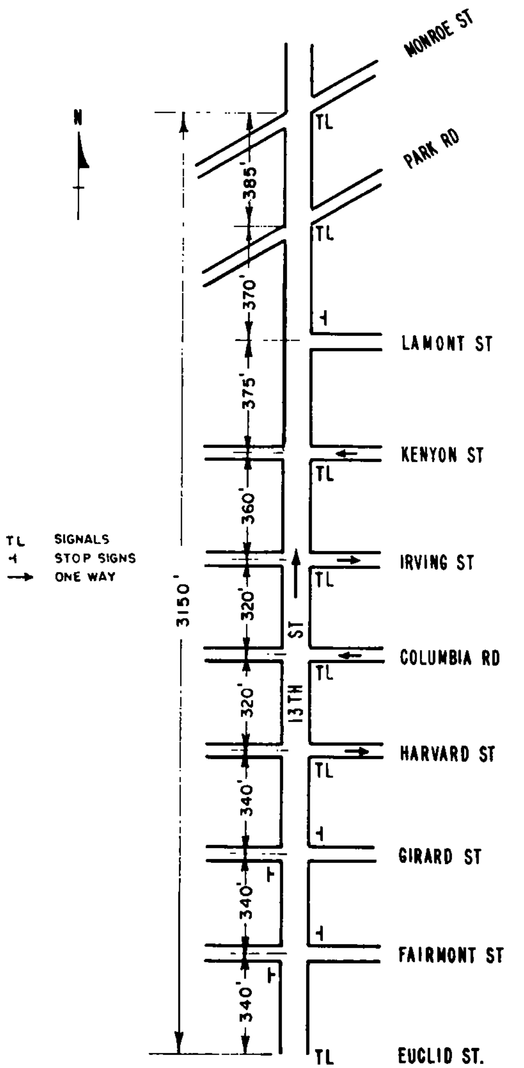

Figure 1. 13th Street layout.

TABLE 2

SUMMARY OF CARS GENERATED BY GENERATION POINT AND BY CYCLE

Street	Lane	Signal	Generation Point	No. of Cars Generated Cycle 1	Cycle 2	Cycle 3	Total
13th	1	Green	1	17	15	9	41
		Red	2	2			2
	2	Green	3	19	23	15	57
		Red	4	1		2	3
	3	Green	5	19	18	16	53
		Red	6				
	4	Green	7	10	13	15	38
		Red	8				
Fairmont	1		9	1	2	1	4
	2		10	1			1
Girard	1		11			1	1
	2		12	1	2	1	4
Harvard	1		13	3	8	1	12
	2		14	3	6	1	10
Columbia	1		15	4	7	8	19
	2		16	8	7	6	21
Irving	1		17	3	5	1	9
	2		18	5	3	3	11
Kenyon	1		19	6	1	2	9
	2		20	4	3		7
Lamont	2		21		2		2
Park	1		22	13	8	11	32
	2		23	10	6	4	20
Monroe	1		24	1	3		4
	2		25	2	4	1	7

cars were able to stay in pace with the signal progression but many fell behind. Tables 1, 2, 3, and 4 are summaries of voluminous, full computer printouts which identify individual vehicles, thus making it possible to trace through the movement of any particular vehicles.

CONCLUSIONS

A working simulation model of an existing, fairly complex traffic location has been constructed. A computer program causes the cars to behave in what seems to be a realistic manner. The cars stop at red lights, they yield the right-of-way at stop signs, they maneuver into correct positions for turns, they move at different speeds, they accelerate and decelerate, faster cars shift lanes to overtake slower cars, they form queues, and they do most of the definable things that cars can be expected to do in city traffic.

The results in no sense indicate a rigorous validation of the model. Up to the present point, reasonableness is the only criterion for judging the performance. Approximately the correct number of cars are accounted for at key points; their characteristics

TABLE 3

SUMMARY OF STATION B COUNTS[a]

Lane	Cars Passing Station B			
	Cycle 1	Cycle 2	Cycle 3	Total
1	20	22	23	65
2	13	18	24	55
3	11	15	21	47
4	18	17	20	55
Total	62	72	88	222
Hourly rate	2,790	3,240	3,960	3,330

[a]Station B is located on 13th Street just north of Lamont Street. Cycle 1 is from simulation-run time 1,036 to 1,355; cycle 2, from time 1,356 to 1,675; and cycle 3, from time 1,676 to 1,955.

as to speed category, type of vehicle, and intended turns correspond with known input data; their average running times are expectedly somewhat slower than that required to keep up with the progressively timed traffic lights. (see Fig. 2)

To get more information bearing on the validity of the model, two steps may still be done. One is to study the movie display carefully to see whether a "helicopter" view of the cars verifies that they are performing correctly. The other is to compare the simulation running times with actual running times from the field.

A point worth bearing in mind is that even though the simulated running times may not be entirely valid in total, a difference in running time to reflect a changed parameter may be highly significant. The reverse is also true. A particular detail of the simulation may not check completely with reality and yet the total result can still furnish a useful measure. Ideally, the simulation would correspond with reality both in detail and in total, but it has value even if one of these objectives is not immediately accomplished.

AREAS FOR FURTHER RESEARCH

The question remains: What constitutes validation of the model? So far, the test of reasonableness is the only criterion that has been applied. When the

TABLE 4

SUMMARY OF VEHICLE RETIREMENT DATA

Cycle	Volume[a]						Distribution of Running Times in ¼ Sec[b]										
	Lane 1	Lane 2	Lane 3	Lane 4	Total	Hourly Rate	200-249	250-299	300-349	350-399	400-449	450-499	500-549	550-599	600-649	Total Cars	Avg. Running Time
1	4	6	11	11	32	1,440		4	3	12	2	3				24	365
2	20	16	17	17	70	3,150			3	12	23	9	7	4	2	60	444
3	11	13	17	17	58	2,610			2	15	9	7	5	11	2	51	464
Total	35	35	45	45	160	2,400		4	8	39	34	19	12	15	4	135	437

[a]Of cars retiring at end of 13th Street.
[b]Of cars traversing entire 13th Street course.

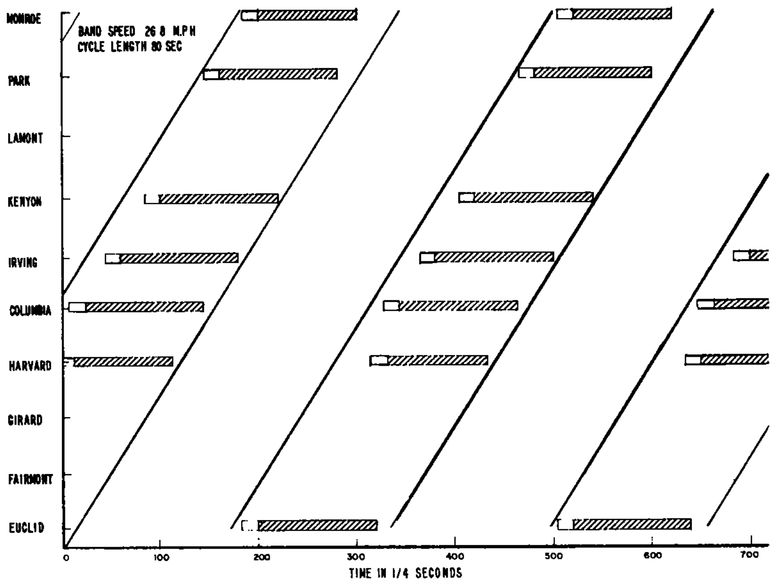

Figure 2. 13th Street signal timing diagram.

performance of the model is accepted as corresponding reasonably closely with actual field conditions, it will be possible to change the parameters and study the new results. From a practical point of view, it is the ability to test untried conditions and to make predictions of likely results that will be the real payoff of simulation as a tool to traffic engineers.

Apart from the immediate objective of getting practical answers for 13th Street are several broader objectives. Study should be made of how to generalize the model in various ways. A model should be made where the main street is two-way rather than one-way. Additional features should be added such as random delay factors, standing vehicles, bus stops, wider range of speeds and acceleration rates, pedestrians, and additional count stations.

Study should be made of what is required to make the model applicable to other locations by "plugging in" different basic data at key points in the program.

Another area of study is the question of how fine the model needs to be to furnish good answers. The present model is very fine. The basic time unit is one-quarter second and the basic distance unit is one-hundredth part of 12 ft (1.44 in). These small units lead to an enormous number of computations even for a high-speed electronic computer. To what extent, if at all, would the usefulness of the results be jeopardized if the model used larger time and distance units? To answer this important research question, it is necessary to use a model that is capable of a fine breakdown. To seek an answer to this question using a coarse model would be impossible.

HISTORY OF PROJECT

The late Professor H. H. Goode of the University of Michigan was one of the first persons to stress the possibilities of digital computer simulation as an aid to engineers in solving traffic engineering problems. He gave a paper on this theme before the Highway Research Board in 1956 (1).

A simple model was constructed under his direction. Each street carried a single lane of traffic in each direction. When cars moved, they all moved at the same speed. In the computer model each car was represented by a single bit. Thus, all cars were regarded as alike and were processed uniformly.

Concept of More Sophisticated Model

It was Goode's idea that instead of describing each car by only one bit it would be possible to represent each car by a whole computer word. In this way the car could be assigned individual characteristics as to speed, type, and destination.

The Bureau of Public Roads entered into an initial agreement with the National Bureau of Standards in July 1958 under which NBS would develop a considerably more advanced simulation model, utilizing several of Goode's concepts.

Selection of 13th Street Site

The 13th Street test site selected is 3,240 ft in length and comprises nine blocks. Seven of the intersections are controlled by traffic lights and three by stop signs. The operation relates to the peak hour of the afternoon rush, when the four lanes of 13th Street are operated one-way northbound.

DESCRIPTION OF METHOD

Each lane of each street is divided into 12-ft sections called unit blocks. Computer storage reserves a place for information about each unit block (UB). If there is a car in a UB, full information about its exact location and its physical characteristics is stored. Another portion of the storage word furnishes any necessary information about the road at that point.

The time cycles for searching all UB's for cars, moving the cars, generating new cars, and preparing any outputs is one quarter-second of simulated real time.

COMPUTER PROGRAM

The basic working program for the IBM-704, including working constants and the input parameters, contains about 6,000 instruction words. The program searches methodically for cars to be processed. Starting at UBO, the first UB in lane 1, the search continues through lanes 1, 2, 3, and 4 of 13th Street, then the lanes of all the cross-streets, and finally the diagonal UB's (for turns).

A Layout and B Layout

The cars are found on what has been called the A layout. To keep matters straight, because it is impossible to process all the cars simultaneously, each car as it is processed is moved to its new position on the B layout. For the remainder of the review cycle the car continues to appear on the A layout in its old position.

When all the cars found in the A layout have been moved to new positions in the B layout, the scanning is completed. Then the A layout is erased and the B layout becomes the starting point for the next scan.

Generation

At the end of each cycle, the car generation routine is performed. If a car is generated, its characteristics are also determined including its destination or "exit." A newly generated car will be launched if this can be done safely. Otherwise, it will be retained on a backlog list for the particular generation point in question until it can be safely launched. Finally the clocks are advanced one quarter-second and the program is ready to repeat the cycle.

Permissible Speed

When a car is found for processing, virtually the first task of the program is to consider the car's desired speed in relation to its present speed and its allowable acceleration rate. Each car carries with it an information package describing various physical characteristics and details.

Sight Distance

When the permissible speed has been determined (in terms of a jump per quarter-second), the equivalent of required sight distance is determined by a table look-up. The program then probes ahead, UB by UB, attempting to achieve the "goal points" necessary to satisfy the sight distance requirement.

Two prime considerations are whether there is a car ahead and whether there is any irregularity about the roadway (such as a traffic signal or a turn). In every case, the key to the information appears in the UB word format (Fig. 3) and can be found by systematic checking of every UB involved (ahead, behind, right, or left as required).

If the goal points can finally be verified, the stated jump can be made (onto the B layout). If the goal points are not adequate, then a table is consulted to determine what reduced jump can be made safely.

Irregular Unit Blocks

During the processing, the program is constantly on the alert to comply with the requirements of any of the roadway "irregularities." If a UB is responsive to a traffic signal, the program must check the signal indication. If there is a turn ahead, the program must test whether the car is intending to turn. If the car passes a count station, it must be properly tallied and clocked. If the car reaches the end of a lane, it must be checked out. A number of other special situations may occur, singly or together. In general, each situation has one or more subroutines which can be called on to determine the proper move. There are 37 main routines, subroutines, and table look-up routines.

ASSUMPTIONS AND PARAMETERS

In many instances, in the absence of specific answers from field data, it was necessary to make certain assumptions or to assign certain arbitrary values to parameters. In most cases, these can be readily changed if desired. These assumptions relate to either one of two areas: the characteristics of the car or the rules governing the movements. Examples of these parameters are the initial lane distribution, the assignment

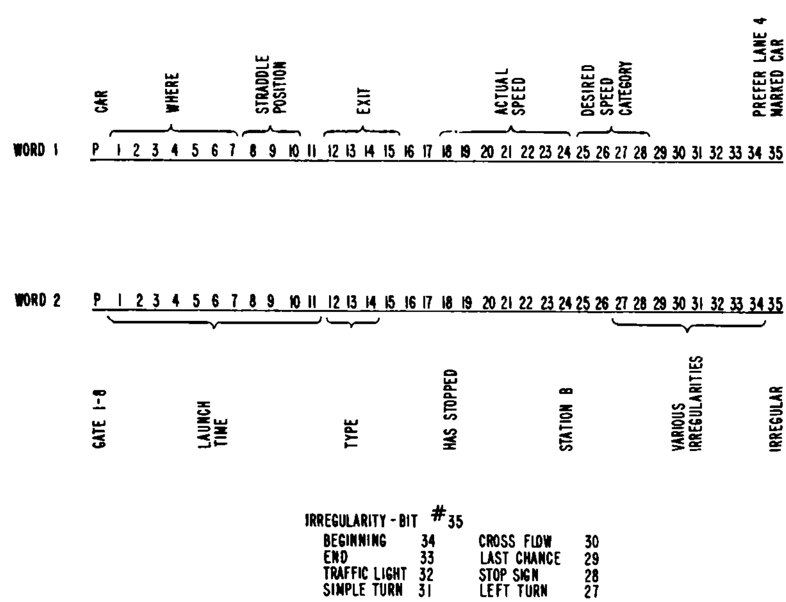

Figure 3. Two-word format.

of desired speeds, acceleration and deceleration rates, clearances, criteria for overtaking, reduced speed for going around turns, acceptable gaps, and reaction time.

It is not presumed that this is a validated model but it is a device that works mechanically and, in general, it would be very easy to change any of the parameters when there are authoritative values to substitute.

ACKNOWLEDGMENT

This paper is a greatly condensed version of NBS Technical Note 119, PB 161620, by the same author, entitled "Computer Simulation of Street Traffic."

REFERENCE

1. Goode, H. H., Pollmar, C. H., and Wright, J. B., "The Use of a Digital Computer to Model a Signalized Intersection." HRB Proc., 35:548-557 (1956).

A LaGrangian Approach to Traffic Simulation on Digital Computers

J. R. WALTON, Instructor, Department of Civil Engineering; and
R. A. DOUGLAS, Associate Professor, Department of Engineering Mechanics
North Carolina State College

A technique for simulating traffic movement on digital computers is described. In this technique computations are performed on an as required basis rather than on an incremental time basis. This treatment of the time parameter may reduce the computational effort required to simulate traffic movement.

• SINCE PUBLICATION of the traffic simulation studies of Gerlough (1) and of Goode, Pollmar, and Wright (2) in 1956, there has been increasing interest in the simulation method of analyzing traffic flow. Subsequent investigations (3, 4) in this area have been concerned with the simulation of increasingly complex traffic situations, based on the simulation techniques presented by Gerlough and by Goode, et al.

Briefly, techniques employed heretofore may be classified according to the way in which vehicles are represented in a computer, and from the manner in which motion and the time parameter are handled.

In the technique of physical representation a roadway or system of roadways is represented by a group of storage locations within a computer. Individual vehicles are represented by binary ones and the spacing between vehicles by binary zeros. The binary ones, representing vehicles, are moved stepwise along the simulated roadway by simple mathematical operations. Before each movement of a simulated vehicle, the existing "traffic" and "roadway" conditions are examined to determine the permissible movement. The result of repeated application of the process is a flow of binary ones along a system of computer locations in a manner analogous to a flow of vehicles along a roadway.

In the memorandum method each vehicle is represented by information stored in a computer word or words. This information includes the vehicle's velocity, its location, its desired velocity, and any other characteristics attributed to an individual vehicle and driver. The parameter of time is treated as an incremental function so that movement is simulated by adjusting, at specified increments of time, the position of each vehicle and such other characteristics attributed to the vehicle that may be influenced by time and position.

Both the memorandum method and the method of physical representation, as used, have been based on an Eulerian viewpoint of motion in that traffic is represented in each as it would appear to outside observers in positions fixed with respect to the roadway. Such observers would see each vehicle moving along the roadway and could describe the phenomena by recording, at successive instants of time, the position, velocity, and attitude of individual vehicles.

LAGRANGIAN APPROACH

In a LaGrangian approach to the simulation of traffic flow, the traffic is described as it would appear to an observer in each vehicle. By moving the observer from outside to within the moving traffic system, the necessity of considering the behavior of each vehicle at each of regular intervals of time is eliminated. Computer programs written from the La Grangian viewpoint need be concerned only with those values of the time parameter when the behavior of a vehicle would change and only with those vehicles actually affected by the change in behavior.

As employed by the authors, (5), the LaGrangian treatment of the time parameter has been incorporated in the memorandum method of simulation. The following information is stored for an individual vehicle:

1. The characteristics assigned to an individual vehicle and its driver.
2. An index number corresponding to the vehicle's relative position in its lane of travel.
3. The equation of motion of the vehicle.
4. The value of the time parameter at which the equation of motion may change.

Boundary conditions at the beginning of a simulation period may be established with the roadway system either empty or bearing vehicles. If vehicles are in the system, their equations of motion and their individual characteristics will be stored in appropriate storage locations. Also stored will be a value of time, T_e, when the first additional vehicle is to enter the roadway. These additional vehicles are introduced into the system by means of a function generator sub-routine of either random or controlled nature. In the remainder of the discussion, it is assumed that vehicles are on a roadway when simulation begins.

The beginning of a period of simulation is taken as time zero, T_0. At T_0, each vehicle's equation of motion is solved for the time, T_i, when the first condition will exist that will require a change in the vehicle's behavior. The computations required involve the solution of each vehicle's equation of motion to determine the time a certain position will be reached; and/or the time at which the vehicle's velocity or acceleration reaches some critical value; and/or the simultaneous solution of several equations of motion for the time when the distance between the vehicle and other vehicles attains some preset critical value. When more than one of these computations is performed for a single vehicle, the resulting set of times is examined and the least value taken as T_i.

The value selected as T_i is stored with the information pertaining to a particular vehicle, and represents the time at which, under the existing traffic and roadway conditions, something may cause the behavior of the vehicle to change.

The least value, T, of the stored times, considering T_e and all the T_i, then determines the first event that will cause a change to be made in the stored information relating to the individual vehicles. To determine the changes required, the situation corresponding to time T must be identified, either by assigning, at the time of the original computation of the T_i, code numbers identifying the events predicted, or by examining the equation of motion and desired behavior pattern of the vehicle whose T_i became T.

If, at T, the event is a vehicle entering the system, the individual characteristics of the vehicle are generated and its equation of motion formed. This new equation of motion then is solved for the T_i of the vehicle. Also determined and stored is a new value T_e which, with the new T_i, is compared with all the previously determined T_i to find the time T of the next event.

When the event at T is a vehicle leaving the system, any desired information relating to that vehicle is stored for compilation or print-out and remaining information is removed from further consideration.

If, at time T, any event other than a vehicle entering or leaving the system is to take place, the change called for is introduced into the vehicle's equation of motion and a new time, T_i determined and stored for that vehicle.

Any vehicle entering the roadway, leaving the roadway, or changing its behavior, may cause changes in the times when other vehicles are to alter their behavior. These changes will effect any vehicle for which the computation of T_i involved the previously affected vehicle. Frequently, only the immediately following vehicle and the immediately oncoming vehicle are involved. In any case, the values of T_i of the vehicles affected are recomputed and stored.

The behavior of all vehicles, as defined by the new set of equations of motion, remains constant in the interval from the original value of T until the least value of the new set of stored times. The event associated with the new T is identified and, again, any required changes are made in the equations of motion of the affected vehicles. The process of determining the time of occurence, T of an event, of adjusting equations of motion dependent upon that event, and determining a new value, T, is continued until the value of T exceeds some predetermined time of simulation and the process is stopped.

SUMMARY

The LaGrangian approach to the simulation of traffic may be recognized as an inspection of events rather than a continuous survey of traffic. Here, computations are performed only for those times when interactions occur, and only for those vehicles affected by the interactions. If the traffic situation being simulated is extremely complex (i.e., a large number of interactions is involved), the computational effort required by this technique approaches that of the Eulerian techniques. If, however, the number of interactions is reduced, as is the case with either moderate or very high traffic density, this technique is indicated.

REFERENCES

1. Gerlough, D. L., "Simulation of Freeway Traffic by an Electronic Computer." HRB Proc., 35:543-547 (1956).
2. Goode, H. H., Pollmar, C. H., and Wright, J. B., "The Use of a Digital Computer to Model a Signalized Intersection." HRB Proc., 35:548-557 (1956).
3. Wong, S. Y., "Traffic Simulation With a Digital Computer," Proc., Western Joint Computer Conf., 92-94 (1956).
4. Perchonok, P., and Levy, S. L., "Application of Digital Simulation Techniques to Freeway On-Road Traffic Flow." HRB Proc., 39:506-523 (1960).
5. Walton, J., and Douglas, R., "Simulation of Traffic Movement on Two Lane Roadways." Final Report, Engineering Research Department and North Carolina State Highway Commission (unpublished).
6. Helly, W., "Dynamics of Single-Lane Vehicular Traffic Flow." Research Report 2, Center for Operations Research, Mass. Inst. Technol. (Oct. 1949).
7. Gerlough, D. L., "A Comparison of Techniques for Simulating the Flow of Discrete Objects." Paper, Nat. Simulation Conf., Dallas, Texas (Oct. 23-25, 1958).

THE NATIONAL ACADEMY OF SCIENCES—NATIONAL RESEARCH COUNCIL is a private, nonprofit organization of scientists, dedicated to the furtherance of science and to its use for the general welfare. The ACADEMY itself was established in 1863 under a congressional charter signed by President Lincoln. Empowered to provide for all activities appropriate to academies of science, it was also required by its charter to act as an adviser to the federal government in scientific matters. This provision accounts for the close ties that have always existed between the ACADEMY and the government, although the ACADEMY is not a governmental agency.

The NATIONAL RESEARCH COUNCIL was established by the ACADEMY in 1916, at the request of President Wilson, to enable scientists generally to associate their efforts with those of the limited membership of the ACADEMY in service to the nation, to society, and to science at home and abroad. Members of the NATIONAL RESEARCH COUNCIL receive their appointments from the president of the ACADEMY. They include representatives nominated by the major scientific and technical societies, representatives of the federal government, and a number of members at large. In addition, several thousand scientists and engineers take part in the activities of the research council through membership on its various boards and committees.

Receiving funds from both public and private sources, by contribution, grant, or contract, the ACADEMY and its RESEARCH COUNCIL thus work to stimulate research and its applications, to survey the broad possibilities of science, to promote effective utilization of the scientific and technical resources of the country, to serve the government, and to further the general interests of science.

The HIGHWAY RESEARCH BOARD was organized November 11, 1920, as an agency of the Division of Engineering and Industrial Research, one of the eight functional divisions of the NATIONAL RESEARCH COUNCIL. The BOARD is a cooperative organization of the highway technologists of America operating under the auspices of the ACADEMY–COUNCIL and with the support of the several highway departments, the Bureau of Public Roads, and many other organizations interested in the development of highway transportation. The purposes of the BOARD are to encourage research and to provide a national clearinghouse and correlation service for research activities and information on highway administration and technology.