QUARTERLY OF APPLIED MATHEMATICS

Vol. XIII JANUARY, 1956 No. 4

STATISTICAL ANALYSIS OF THE FLOW OF HIGHWAY TRAFFIC THROUGH A SIGNALIZED INTERSECTION*

BY

G. F. NEWELL
Brown University

Abstract. The following is a report on some calculations of the statistical distribution of delay times due to a fixed-time traffic signal on a single lane highway.

In Sec. 1, a model of a traffic light is proposed leading to a set of dynamical equations describing a relation between the times at which cars leave the light in terms of the times at which they arrive. In Sec. 2, some equations are derived for the conditional probabilities that a car will leave at any specified time if it enters at some given time. For this, it is assumed that the time intervals between incoming cars form a set of independent random variables and that one seeks only the equilibrium solutions for which the arrival time of any individual car has a constant probability density.

In Sec. 3, a procedure for obtaining approximate solutions of these equations is derived which actually gives exact solutions for the special case in which the cars arrive at equally spaced time intervals, discussed in Sec. 4. In Secs. 5 and 6 this procedure is also applied to obtain first and second approximations in the special case in which cars arrive with the maximum disorder in spacing possible for this model.

It is found that to a first approximation, it makes very little difference what statistical assumptions are made if one wishes to calculate the average delay.

1. Introduction. As an illustration of how one may apply statistical methods to the study of traffic problems, we consider a relatively simple model intended to simulate the flow of automobiles through a single traffic light on a single lane highway.

The motions of individual cars are described graphically in Fig. 1. The trajectory of each car is represented by plotting its position x as a function of time t. The position $x = 0$ is chosen to be the position of the signal and we represent the state of the traffic light by a dark line for those times during which the light is red and by a thin line when it is green. (Actually the red and green intervals will be effective red and green intervals defined more precisely for the particular model in terms of their influence on the trajectories of the individual automobiles).

We make the following as our initial postulate:

I. The cars approach the traffic light in an ordered sequence, without passing. At sufficiently large distance from the light all cars move with the same velocity v_0 .

In view of this assumption, the trajectories of Fig. 1 must approach parallel straight lines asymptotically for $x \to +\infty$ or $x \to -\infty$. No trajectories will cross anywhere.

*Received October 4, 1954.

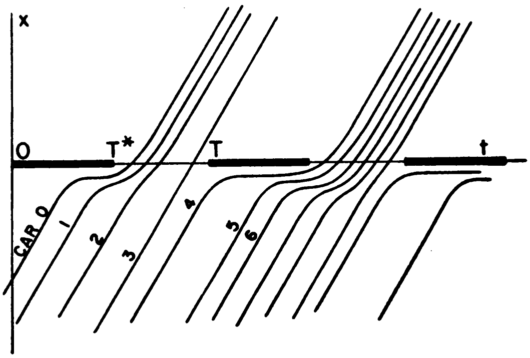

FIG. 1. An illustration of a set of trajectories in which the position x is plotted against t for each car, numbered 0, 1, 2, $\cdots$. Heavy lines represent red intervals of the light at $x = 0$. Light lines at $x = 0$ denote green intervals of the light.

As a result of I, a somewhat more convenient graphical representation of the trajectories is obtained if we choose a "time", $\tau = t - x/v_0$. τ represents graphically a coordinate perpendicular to the asymptotes of Fig. 1. If one plots x vs. τ as in Fig. 2,

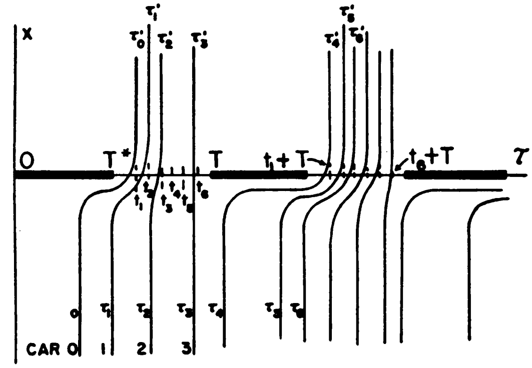

FIG. 2. The trajectories of Fig. 1 are represented by a plot of x vs. $\tau = t - x/v_0$. The "arrival time", the intercept at $x = 0$ of the asymptote of the lower part of each trajectory, is represented by τ_j . The "leaving time", the intercept at $x = 0$ of the asymptote of the upper part of the trajectory, is represented by τ_j' . t_1 , t_2 , $\cdots$; $t_1 + T$, $t_2 + T$ $\cdots$ denote the possible values of τ_j' for delayed cars. Car 3 in Figs. 1 and 2 is not delayed.

the asymptotes becomes vertical lines. At $x = 0$, $\tau = t$. Thus the traffic light itself is represented graphically in Fig. 2 exactly as in Fig. 1. Also the values of τ at which the trajectories cross the light in Fig. 2 correspond exactly to the values of t at $x = 0$ in Fig. 1. For $x \leq 0$, τ represents the time at which the car would arrive at the light if it were to proceed toward the light at a constant velocity v_0 . For $x \geq 0$, τ represents the time at which the car would have left the light had it been moving at the velocity v_0 from the time it was at $x = 0$ until arriving at the position x. Differences in τ along the trajectory thus measure the delay due to the traffic light; the total delay is simply the difference between τ for $x = +\infty$ and for $x = -\infty$.

The detailed shape of the trajectories near the traffic light is of much less practical significance than the difference between the initial and final values of τ, the total delay.

Rather than attempt to propose detailed dynamical laws of motion for each car, we shall propose instead a model which describes only the final values of τ in terms of the initial values and postulates nothing about the intermediate states.

The dynamics of this problem is thus completely described in terms of the variables

$$\tau_j = \text{initial value of } \tau \text{ for the } j\text{th car},$$

$$\tau_j' = \text{final value of } \tau \text{ for the } j\text{th car}.$$

Since there is to be no passing, we can number the cars according to the values of τ so that

$$\tau_j < \tau_{j+1} \qquad \text{and} \qquad \tau_j' < \tau_{j+1}' .$$

We shall for convenience describe the values of τ_j and τ_j' as the times of arrival at and departure from the light respectively and we shall further postulate that:

II. $$\tau_{j+1} - \tau_j \geq \delta \qquad \tau_{j+1}' - \tau_j' \geq \delta.$$

Far in front of or behind the light, cars remain at least a certain minimum distance apart. Since all cars move with the same initial and final velocities, we can say equivalently that they arrive or leave at times separated by a certain minimum interval of time δ.

There are certain qualitative features of the motion through the traffic light which should be considered as a basis for future postulates that will yield a detailed description of the relation between τ_j' and τ_j .

Imagine that some car, which we shall number as car 0, approaches the light soon after it has turned red and that all preceding cars have already cleared the intersection. The car comes to a full stop and proceeds after the light turns green. The value of τ_0' is determined by the time at which the light turned green and is essentially independent of τ_0 . We extend this notion somewhat and postulate the following:

IIIa. If the jth car is delayed in its motion either by the car preceding it or the traffic light or both and if it is the first car to leave the intersection after some red period, then τ_j' is independent of τ_j .

Although the postulate conforms to reality in most respects, it is undoubtedly in some error when applied to a car that approaches the light just before it turns green. Such a car decelerates, but before it comes to a complete stop, the light might turn green and it would proceed through the intersection leaving earlier than if it had arrived somewhat earlier and been forced to stop completely. Such events undoubtedly will in most cases be relatively rare and have little effect upon the overall picture.

In a similar manner we postulate:

IIIb. If the jth car is delayed and is the second car to leave the intersection after a red period, then $\tau_j' = \tau_{j-1}' + \delta$, independent of τ_j . Similarly for the third car to leave, fourth cars etc. We also assume that if all cars leaving during some green period have been delayed, then exactly n cars leave at discreet times τ_j' independent of the manner in which they arrived.

This assumption also contains some error in that the spacing of the cars leaving the light does in practice depend somewhat on whether or not the cars are stopped completely by the light or just slowed down. Also n may depend somewhat on the values of τ_j . To make any more detailed assumptions than these would however make the model extremely complicated and probably add very little to the general picture.

To complete the model we have yet to consider the car that is delayed neither by the light nor the preceding car.

IV. If τ_j occurs during a green period and if $\tau_j > \tau'_{j-1} + \delta$, i.e. if the jth car arrives at the light at least δ later than the j-1th car has left the light, we postulate that $\tau'_j = \tau_j$. A car does not decelerate unless it is forced to do so either by a red light or its proximity to the car in front of it.

Postulates I to IV give essentially a complete description of the dynamics of our model.

In order to express these postulates symbolically, we introduce the following additional notation (see Fig. 2):

T = duration of cycle,

T^* = length of the red interval.

We choose the zero of time at the beginning of some arbitrary red period so that if τ mod $T \leq T^*$, the light is red at time τ.

In accordance with postulate IIIa and IIIb we let t_1 = value of τ' for the first car to leave the light after time T^*, assuming that it was delayed, t_2 = corresponding value of τ' for the second car etc. In general, a delayed car will leave only at the time $t_j + kT$, $j = 1, \cdots, n$, k = some integer, $t_j - t_{j-1} = \delta$.

The dynamical equations of the system expressed by III and IV specify in effect that if car j enters at time τ_j, it will leave at the earliest possible time thereafter. The easiest way of expressing these conditions by equations is to give τ'_j in terms of τ_j and τ'_{j-1}. If we let k'_{j-1} denote the cycle in which the j-1th car leaves the light and k_j the cycle in which the jth car enters so that

$$k'_{j-1}T < \tau'_{j-1} < (k'_{j-1} + 1)T, \qquad k_jT < \tau_j < (k_j + 1)T,$$

then the equations of motion are:

(i) If $\tau_j < \tau'_{j-1} + \delta$, the jth car is delayed by car $j - 1$. This implies that car $j - 1$ has also been delayed for some reason since by II, $\tau_j > \tau_{j-1} + \delta$ and therefore $\tau'_{j-1} > \tau_{j-1}$. We can therefore write τ'_{j-1} in the form

$$\tau'_{j-1} = k'_{j-1}T + t_l \qquad \text{for some } l.$$

τ'_j is then given by

$$\tau'_j = k'_{j-1}T + t_{l+1} \qquad \text{if} \qquad l + 1 \leq n,$$

or

$$\tau'_j = (k'_{j-1} + 1)T + t_1 \qquad \text{if} \qquad l = n.$$

(ii) For $\tau_j > \tau'_{j-1} + \delta$ (jth car not delayed by car $j - 1$) then

$$\tau'_j = \tau_j \qquad \text{if} \qquad \tau_j \bmod T > T^*,$$

$$\tau'_j = k_jT + t_1 \qquad \text{if} \qquad \tau_j \bmod T < T^*.$$

(If the light is green, it is not delayed; if the light is red, it leaves in the first position after the red.)

It follows from the above description, that if one knows the trajectory of one car, it is a rather simple matter, for any particular values of the entering times τ_j, to construct the trajectories for subsequent sequences of cars. By iteration, one simply sends each car in order through free or puts it in the first available "bin" according to the conditions left by the preceding car.

To try to catalog all possible motions for the entire collection of cars is, however, a task of another magnitude. It is a rather uninteresting task also because the questions of primary interest are concerned not with the detailed behavior of the system but rather with the average behavior resulting from some typical distribution of incoming cars. To try to analyze the problem exactly would only yield a catalog of irrelevant information which would probably completely obscure the relevant features of the system. For this reason it is necessary that we approach the problem now from a statistical point of view.

As a final note on the dynamics of this model, it is interesting to note that the equations of motion are not reversible. Although the τ_i' are determined by the τ_i, the converse is not true. Many possible sets of τ_i will lead to the same τ_i'.

2. Some statistical postulates. We shall be concerned here primarily with the problem of determining the distribution of delay times $\tau' - \tau$, assuming that the incoming cars are distributed in such a way that the times between incoming cars $\tau_i - \tau_{i-1}$ form a set of independent random variables. Although much of the study of queues of various sorts have dealt with incoming flows of the above type [1, 2], in fact usually with a Poisson distribution, and have usually considered only questions of delay times or lengths of the queues, these are by no means the only problems of practical interest. The assumption of statistical independence of differences in arrival times is a reasonable one for a long highway free of other obstacles but it is apparent that the flow *leaving* the traffic light is not of this type. In order to study the very interesting problems that would arise when the flow from one obstacle encounters a new obstacle, one must not only relax the restriction on the type of incoming flow that one considers but one must calculate other things than just delay times or queue lengths; in particular, one must calculate the statistical distribution of leaving times. This is obviously a problem of much greater difficulty.

To return to the simpler problem, we concentrate our attention on car 0 as some arbitrary car arriving at time τ_0 in the cycle $k_0 T < \tau_0 < (k_0 + 1)T$ and we seek to calculate the probability that this car will leave at a time τ_0' in the cycle $k_0' T < \tau_0' < (k_0' + 1)T$. We let $P_0(\tau_0' \mid \tau_0)$ = probability that car 0 leaves at τ_0' if it enters at time τ_0. τ_0 will be a continuous variable, but except for the special case $\tau_0' = \tau_0$, τ_0' will have only discreet values $k_k' T + t_l$. We also define $P_k(\tau_k' \mid \tau_k)$ = probability that car k leaves at time τ_k' if it enters at time τ_k.

By assuming that $(\tau_k - \tau_{k-1})$ are independent random variables, we guarantee that the problem will be a Markov process of order 1. In so doing, we let $f(y)dy$ = probability that two consecutive cars arrive at times differing by an amount between y and $y + dy$. The function f completely defines the distribution of incoming cars. Actually we have also assumed here that f is independent of k. To do otherwise would be of questionable physical interest and a great mathematical inconvenience. We will not at this time specify any particular form of the function f but it is to be considered as known and to be consistent with assumption II

$$f(y) = 0 \qquad \text{for} \qquad y < \delta.$$

The general procedure of attack on this problem is first to find a series of integral equations expressing $P_k(\tau_k' \mid \tau_k)$ in terms of $P_{k-1}(\tau_{k-1}' \mid \tau_{k-1})$ and $f(\tau_k - \tau_{k-1})$ using the dynamical equations. We then seek to find an equilibrium solution, assuming one exists, by imposing the additional condition that $P_k(\tau_k' \mid \tau_k)$ is not a function of k. In this way we obtain a set of integral equations involving only $P_0(\tau' \mid \tau)$ as the unknown.

We begin the derivation of the equations by considering $P_1(\tau_1' \mid \tau_1)$ for $\tau_1' = k_1'T + t_l$ and $l > 1$, i.e. we seek to find the probability that car 1 will be delayed but will leave at various τ_1' other than the first position after some red light if it enters at τ_1. According to the dynamical equations, this implies that car 0 left at $\tau_0' = k_1'T + t_{l-1}$, $(k_0' = k_1')$. Thus

$$P_1(k_1'T + t_l \mid \tau_1) = \int_{-\infty}^{\tau_1} d\tau_0\, P_0(k_1'T + t_{l-1} \mid \tau_0) f(\tau_1 - \tau_0) \quad \text{for} \quad \tau_1 < k_1'T + t_l, \quad l > 1. \tag{1a}$$

$$P_1(k_1'T + t_l \mid \tau_1) = 0 \quad \text{for} \quad \tau_1 > k_1'T + t_l, \quad l > 1. \tag{1b}$$

[We assume here that τ_0 can have any value with equal probability. In general one obtains an equation of this type for the joint distribution of τ_1' and τ_1 in terms of the joint distribution of τ_0' and τ_0. If τ_0 is uniformly distributed, so also is τ_1 since f depends only upon $(\tau_1 - \tau_0)$. Thus one obtains an equation involving only condition probabilities.]

The upper limit of the integral in (1a) is somewhat arbitrary since $f(\tau_1 - \tau_0) = 0$ for $\tau_0 > \tau_1 - \delta$.

The case $l = 1$ and the case $\tau_1' = \tau_1$ must be considered separately. For $l = 1$, car 1 can leave in a first position $k_1'T + t_1$, if either of the following applies.

(i) $\tau_1 < k_1'T$ and $\tau_0' = (k_1' - 1)T + t_n$.

(ii) $k_1'T < \tau_1 < k_1'T + T^*$ and $\tau_0' < k_1'T$.

From this we obtain

$$P_1(k_1'T + t_1 \mid \tau_1) = \int_{-\infty}^{\tau_1} d\tau_0\, P_0[(k_1' - 1)T + t_n \mid \tau_0] f(\tau_1 - \tau_0) \quad \text{for} \quad \tau_1 < k_1'T, \tag{2a}$$

$$P_1(k_1'T + t_1 \mid \tau_1) = \int_{-\infty}^{\tau_1} d\tau_0 \{ \sum_{\tau'_0 < k'_1 T} P_0(\tau_0' \mid \tau_0) \} f(\tau_1 - \tau_0) \quad \text{for} \quad k_1'T < \tau_1 < k_1'T + T^*, \tag{2b}$$

$$P_1(k_1'T + t_1 \mid \tau_1) = 0 \quad \text{for} \quad k_1'T + T^* < \tau_1. \tag{2c}$$

The upper limits of these integrals are again somewhat arbitrary.

Finally, for the case $\tau_1' = \tau_1$, we obtain

$$P_1(\tau_1 \mid \tau_1) = \int_{-\infty}^{\tau_1} d\tau_0 \{ \sum_{\tau'_0 < \tau_1 - \delta} P(\tau_0' \mid \tau_0) \} f(\tau_1 - \tau_0) \tag{3a}$$

$$\text{if} \quad \tau_i \bmod T > T^*. \tag{3b}$$

$$P_1(\tau_1 \mid \tau_1) = 0 \quad \text{if} \quad \tau_i \bmod T < T^*. \tag{3c}$$

Equations 1, 2 and 3 give a complete description of the function $P_1(\tau_1' \mid \tau_1)$ in terms of $P_0(\tau_0' \mid \tau_0)$, assuming that τ_0 is uniformly distributed (consequently also τ_1).

The next step is to seek an equilibrium distribution. This is done simply by dropping the subscripts on the function P so that P_1 is the same function of its argument as P_0 is of its argument. Equations 1, 2 and 3 then become a set of simultaneous linear integral equations. Rather than rewrite these equations without the subscripts, we shall hereafter use them with this minor alteration implied.

As the equations now stand, we have an infinite set of equations, for determining an infinite set of functions (one for each of the discrete values of τ_k') of a continuous variable τ_1, $-\infty < \tau_1 < \tau_1'$.

This can be reduced to more manageable form by taking advantage of the invariance of the equations to displacements of the time coordinate by multiples of T. Although this invariance does not in itself necessarily guarantee that the only solutions of the equations are themselves invariant to such a transformation, i.e.

$$P(\tau_1' + T \mid \tau_1 + T) = P(\tau_1' \mid \tau_1),$$

such solutions are certainly the only ones that would seem to make sense physically.

In view of this, one need calculate $P(\tau_1' \mid \tau_1)$ only for $0 < \tau_1' < T$, i.e. for $k_1' = 0$, but for arbitrary τ_1. The values of $P(\tau_1' \mid \tau_1)$ for other values of τ_1' are then obtained by using the periodicity.

There is still one other fact that we will use to advantage. Equations 1, 2, 3 are linear homogeneous integral equations. The normalization should eventually be chosen so that

$$\sum_{\tau'_1} P(\tau_1' \mid \tau_1) = 1$$

since these are conditional probabilities. By using this, we can eliminate one of the unknowns [for example, $P(\tau_1 \mid \tau_1)$] thus reducing the number of integral equations by one but by so doing making them inhomogeneous.

3. Analysis of equations. We begin by considering Eq. 1. To simplify notation we let

$$y_1 = \tau_1 - k_1'T,$$

$$y = \tau_1 - \tau_0,$$

and use the periodicity to obtain

$$P(t_l \mid y_1) = \int_0^\infty dy\, P(t_{l-1} \mid y_1 - y) f(y) \qquad \text{for} \qquad y_1 < t_l \tag{4a}$$

$$= 0 \qquad \text{for} \qquad y_1 > t_l,\ l > 1. \tag{4b}$$

We can iterate this equation by successive substitutions of the left side into the right side for appropriate t_l until we obtain an expression for $P(t_l \mid y_1)$ in terms of $P(t_1 \mid y)$. By so doing, we obtain the following equation which corresponds to the statement: if car 0 leaves in the lth position, car $(-l + 1)$ must leave in the corresponding first position.

$$P(t_l \mid y_1) = \int_0^\infty dy\, P(t_1 \mid y_1 - y) f^{(l-1)}(y) \qquad y_1 < t_l, \tag{5a}$$

$$= 0 \qquad y_1 > t_l,\ l > 1 \tag{5b}$$

in which

$$f^{(1)}(y) \equiv f(y),$$

$$f^{(i)}(y) = \int_0^y dy_1 f(y - y_1) f^{(i-1)}(y_1), \tag{6}$$

$f^{(j)}(y)$ represents simply the probability density for a zeroth car and a jth car to be separated by a time interval y. The above integrals are multiple convolutions of the function $f(y)$.

Equation (2a) is quite similar in form to (1a) and by the same reasoning gives

$$P(t_1 \mid y_1) = \int_0^\infty dy\, P(t_1 \mid y_1 + T - y) f^{(n)}(y) \qquad \text{for} \qquad y_1 < 0. \tag{7}$$

Equation (2b) is less pleasant. By using the normalization condition we can avoid considering Eq. (3) and write in (2b)

$$\{ \sum_{\tau'_0 < k'_1 T} P(\tau'_0 \mid \tau_0) \} = 1 - \sum_{\tau'_0 > k'_1 T} P(\tau'_0 \mid \tau_0).$$

The case $\tau'_0 = \tau_0$ is now excluded from the sum because in (2b) we integrate only over those τ_0 for which

$$\tau_0 < \tau_1 \leq k'_1 T + T^*.$$

But $\tau'_0 > k'_1 T$ means $\tau'_0 > k'_1 T + T^*$, since τ'_0 cannot lie in a red interval. Thus $\tau'_0 > \tau_0$ and so the above sum includes only discrete values of τ'_0.

$$\{ \sum_{\tau'_0 < k'_1 T} P(\tau'_0 \mid \tau_0) \} = 1 - \sum_{k'_0 = k'_1}^{\infty} \sum_{l=1}^{n} P(k'_0 T + t_l \mid \tau_0). \tag{8a}$$

Furthermore since $P(\tau'_0 \mid \tau_0) = 0$ for $\tau'_0 < \tau_0$, all terms on the left side vanish if $k'_1 T < \tau_0$; i.e. if a car enters after $k'_1 T$ it cannot leave before $k'_1 T$.

$$\{ \sum_{\tau'_0 < k'_1 T} P(\tau'_0 \mid \tau_0) \} = 0 \qquad \tau_0 > k'_1 T. \tag{8b}$$

In terms of the new notation (2b) and (8) give

$$P(t_1 \mid y_1) = \int_0^\infty dy \left\{ 1 - \sum_{k=0}^{\infty} \sum_{l=1}^{n} P(t_l \mid y_1 - y - kT) \right\} f(y) \quad \text{for} \quad 0 < y_1 < T^*. \tag{9}$$

According to (8b), the quantity in the brackets of (9) must vanish for $y_1 > y$. In the final solution, Eq. (8b) and (8a) must be self-consistent and we apparently have the freedom of using (8b) or not here, as we wish. If we do use (8b), we take y_1 as the lower limit of the integral in (9). This leads to a less elegant form of the integral equation but one which is more suitable, at least for approximate calculations. In either case, substitution of (5) and (7) will reduce (9) to a simple form.

Equation (9), without (8b) gives

$$P(t_1 \mid y_1) = 1 - \int_0^\infty dy\, P(t_1 \mid y_1 - y) \left\{ \sum_{l=1}^{\infty} f^{(l)}(y) \right\}, \qquad 0 < y_1 < T^*. \tag{10}$$

Equation (9) with (8b) gives

$$P(t_1 \mid y_1) = \int_{y_1}^\infty dy\, f(y) - \int_{y_1}^\infty dy\, P(t_1 \mid y_1 - y) \sum_{l=1}^{\infty} f^{(l)}(y, y_1), \qquad 0 < y_1 < T^*, \tag{11}$$

in which

$$f^{(1)}(y, y_1) \equiv f(y)$$

$$f^{(l)}(y, y_1) \equiv \int_0^{y - y_1} dy'\, f^{(l-1)}(y') f(y - y'). \tag{12}$$

Equation (10) or (11) expresses the fact that $P(t_1 \mid y_1)$ equals the probability that no previous car left in position t_1. In (10) this is expressed as one minus the probability that a car, l cars ahead left in position t_1, summed over l and integrated over all arrival times of the lth car ahead of the reference car. In (10) this is expressed as the probability that the reference car is the first to arrive after the light turned red, less the probability that if all previous cars arrive before the red light, the lth one left in position t_1; summed over all l and integrated over all arrival times before the red light.

Finally we rewrite Eq. (2c) in the form

$$P(t_1 \mid y_1) = 0 \qquad \text{for} \qquad y_1 > T^*. \tag{13}$$

To summarize, we see that Eq. (5) gives an expression for $P(t_l \mid y_1)$ in terms of $P(t_1 \mid y)$ for all $l > 1$, whereas equations (7), (10) or (11) and (13) together form an equation for $P(t_1 \mid y)$ in terms of itself. The problem is essentially reduced to solving this latter set of equations for $P(t_1 \mid y)$.

The set of equations for $P(t_1 \mid y_1)$ could be incorporated into a single inhomogeneous integral equation with a discontinuous kernel, although this would be of doubtful value. It might be of some help, however, to substitute (11) into (7). Equation (11) gives $P(t_1 \mid y)$ for $y > 0$ in terms of $P(t_1 \mid y)$ for $y < 0$. Substitution of (11) and (13) into (7) would then give an equation for $P(t_1 \mid y)$ for $y < 0$ in terms of itself, namely

$$\begin{aligned} P(t_1 \mid y_1) = {} & \int_0^{T^*} dy\, f^{(n)}(y_1 + T - y) \int_y^\infty dy'\, f(y') \\ & + \int_0^\infty dy\, P(t_1 \mid -y)\Big\{ f^{(n)}(y + y_1 + T) \\ & - \int_0^{T^*} dy'\, f^{(n)}(y_1 + T - y') \sum_{l=1}^\infty f^{(l)}(y + y', y') \Big\} \qquad \text{for} \qquad y_1 < 0. \end{aligned} \tag{14}$$

The above set of integral equations is rather disagreeable in any form. At first one might be strongly inclined to express everything in terms of Laplace or Fourier transforms. This would certainly lead to very simple expressions for the convolutions in Eq. (6) and (12) and even for the sums over l in (10), (11) and (14). One also observes that Eq. (7) would be invariant to displacements in y_1 except for the fact that (7) is valid only for $y_1 < 0$, a condition that is not invariant to such displacements. Similarly Eq. (10) and (13) "almost" have this translational symmetry. These supplementary conditions are just enough to make this approach rather unpleasant and as yet no way has been found to overcome certain difficulties. The other forms of the equations, namely (11) and (14) show this lack of symmetry more clearly.

It seems quite unlikely that one will be able to obtain even a formal closed form solution of these equations except possibly for special functions $f(y)$. One can, however, obtain a useful approximation procedure based upon the assumption that the average number of incoming cars per cycle is small compared with the critical number n.

In the limit zero flux, no interaction between cars, $P(t_1 \mid y_1)$ must obviously be either one or zero accordingly as y is in the interval $0 < y < T^*$ or not. In the above formulas, this approximation obtains when the integrals in (7) and (10) vanish. A much better approximation for low flows is obtained, however, from (7) and (11).

We take as a first approximation*

$$P(t_1 \mid t_1) \sim P_1(t_1 \mid y_1) = 0 \qquad \text{for} \qquad y_1 < 0, \tag{15a}$$

i.e. there is zero probability that a car entering in one cycle will leave in a later cycle. Substitution of this into (11) then gives as a first approximation for $y_1 > 0$

$$\begin{aligned} P_1(t_1 \mid y_1) &= \int_{y_1}^{\infty} dy\, f(y) = 1 - \int_0^{y_1} dy\, f(y) \qquad \text{for} \qquad 0 < y_1 < T^* \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad y_1 > T^*. \end{aligned} \tag{15b}$$

This first approximation corresponds to the statement that a car leaves in a position t_1 if it arrives during the corresponding red period and if it was also the first car to arrive in this period.

One can make a sequence of approximations based upon this first approximation by successive substitution of each new approximation back into the equations to obtain the next approximation. Each new approximation will give a correction due to the influence of cars arriving in cycles further and further removed from the reference car.

If one expresses all the P's in terms of the solution of (14), it suffices to apply this procedure simply to this equation. Thus we obtain for $y_1 < 0$

$$\begin{aligned} P_1(t_1 \mid y_1) &= 0, \\ P_2(t_1 \mid y_1) &= \int_0^{T^*} dy\, f^{(n)}(y_1 + T - y) \int_y^{\infty} dy'\, f(y'), \\ P_j(t_1 \mid y_1) &= \int_0^{\infty} dy\, P_{j-1}(t_1 \mid -y)\Big\{ f^{(n)}(y + y_1 + T) \\ &\quad - \int_0^{T^*} dy'\, f^{(n)}(y_1 + T - y') \times \sum_{l=1}^{\infty} f^{(l)}(y + y', y')\Big\} \qquad \text{for} \qquad j > 2, \end{aligned} \tag{16}$$

and

$$P(t_1 \mid y_1) = \sum_{j=1}^{\infty} P_j(t_1 \mid y_1). \tag{17}$$

Successive corrections generally become increasingly more difficult to evaluate and so the value of this procedure will be limited in practice to situations in which the first few approximations already give results of sufficient accuracy.

4. Ordered flow. To illustrate the above scheme, we consider two extreme cases. The simplest case is that of a completely ordered arrangement of incoming cars, uniformly spaced in time; to be considered in this section. The second case, to be considered in the next section represents the opposite extreme, namely that with the maximum disorder possible for this model.

In the former case, d is chosen to be the time interval between consecutive equally spaced cars and $f(y)$ is chosen to be

$$\begin{aligned} f(y) &= \delta(y - d), \qquad \delta = \text{Dirac } \delta\text{-function} \\ f^{(l)}(y) &= \delta(y - ld). \end{aligned} \tag{18}$$

*The subscripts on P will be used here to denote successive corrections in this approximation scheme and are not to be confused with those used in Sec. 1 (and later discarded).

It follows immediately from (16) that if $nd > T$, i.e. if the incoming flow is less than the critical flow, then $P_2(t_1 \mid y_1) = 0$ for $y_1 < 0$ and consequently all higher corrections also vanish.

$P(t_1 \mid y_1)$ is given exactly from (11) by

$$P(t_1 \mid y_1) = \begin{cases} 1 & \text{if} \quad 0 < y_1 < d \text{ and } T^* \\ 0 & \text{otherwise.} \end{cases} \tag{19a}$$

From (5) one then finds

$$\begin{aligned} P(t_l \mid y_1) &= \begin{cases} 1 & \text{if} \quad (l-1)d < y_1 < ld \quad \text{and} \quad y_1 < t_l \\ 0 & \text{otherwise} \quad \text{for} \quad d < T^*, \end{cases} \\ &= \begin{cases} 1 & \text{if} \quad (l-1)d < y_1 < (l-1)d + T^* \quad \text{and} \quad y_1 < t_l \\ 0 & \text{otherwise} \quad d > T^*. \end{cases} \end{aligned} \tag{19b}$$

Thus the complete solution is in this case quite elementary. We can pursue the analysis further and investigate the distribution of delay times. We define $p(t)$ such that $p(t)dt$ is the probability of a delay of t to t plus dt. With the uniform distribution of arrival time for the reference car, in accordance with earlier postulates, one finds that

$$p(t) = \frac{1}{T} \sum_{l=1}^{n} P(t_l \mid t_l - t). \tag{20}$$

There will always be a non-zero probability for no delay. $p(t)$ is therefore defined as above only for $t > 0$. The probability of zero delay is related to $P(\tau \mid \tau)$ which was eliminated from the integral equations by the normalization condition. The probability of zero delay is obtained also from a normalization condition and is given by

$$1 - \int_0^\infty p(t)\, dt.$$

Substitution of (19) into (20) gives first of all

$$p(t) = 0 \qquad \text{for} \qquad t < 0 \qquad \text{or} \qquad t > t_1 .$$

For $0 < t < t_1$, there are numerous special cases depending upon the relative size of d, T^* and t, but in any case $p(t)$ usually has, for various t, values of 0, 1 or 2 depending upon whether 0, 1 or 2 of the step functions in (19) overlap for that t when substituted into (20). In any particular case it is a simple task to find the appropriate explicit form. In a typical case $p(t)$ will oscillate between the values 1 and 2 in this interval as illustrated in Fig. 3 for a particular choice of parameters.

From $p(t)$ one can calculate averages of any function of t. In particular, one can evaluate the average delay

$$\langle t \rangle = \frac{1}{T} \sum_{l=1}^{\infty} \int_0^\infty t\, dt\, P(t_l \mid t_l - t). \tag{21}$$

For any given set of constants for the traffic light itself, one can easily calculate $\langle t \rangle$ as a function of d or as a function of $\beta \equiv T/nd$, the ratio of the flow rate to the critical rate. Figure 4 shows the results of an exact calculation for a particular choice of T^*, δ and t_1. Although the graph appears on this scale to be quite smooth, it really has a

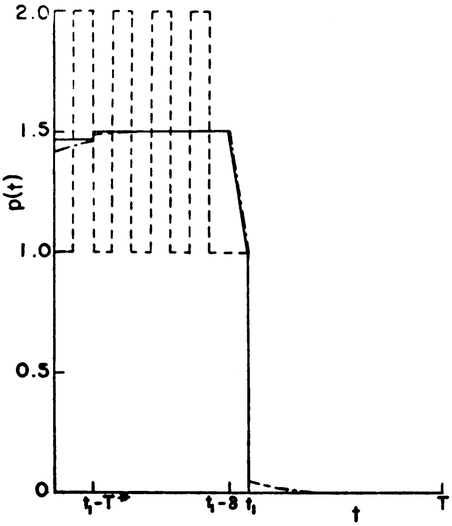

FIG. 3. Distribution of delay times $p(t)$ plotted vs. t for a traffic light with $T = 1$, $T^* = 0.4$, $\delta = 1/20$, $t_1 = 0.5$, $n = 10$ and a flow parameter $\beta = 2/3$ ($\alpha = 10$, $d = 1/20$).

- - - - - ordered flow

———— disordered flow, first approximation

— — — - disordered flow, second approximation.

large number of discontinuities in the derivatives resulting from the numerous discontinuities in the function $p(t)$.

As a practical procedure for obtaining reasonable estimates of the average delay for equally spaced cars, the above method gives much more detail and requires much more work than is justified by the crudeness of the model. In most cases, particularly when n is not a small integer, the above formulation will give results differing only slightly from the formulas for $\langle t \rangle$ given by Clayton [3] on the basis of somewhat simpler assumptions. Clayton calculates $\langle t \rangle$ on the basis of a completely specified sequence of arrival times whereas it has been assumed here that any given car arrives at a randomly distributed time but that the differences in arrival times of consecutive cars are specified.

In terms of the notation of this paper, Clayton's formula reads

$$\langle t \rangle = \frac{(t_1 - \delta/2)^2}{2T(1 - n\,\delta\beta/T)}. \tag{22}$$

The above formula is also plotted in Fig. 4 for comparison, using the same parameters. It can best be described perhaps as a smooth out version of the previous results. The differences between the two curves would however be more accentuated if we had chosen to compare them for a light described by a smaller value of n or one for which δ was not so small compared with t_1 .

5. Disordered flow. The postulates made in Sec. 1 exclude the possibility of applying the above procedure directly to a flow which is completely disordered, i.e. one for which the spacing of incoming cars obeys a Poisson distribution. Although this type of flow has been investigated by several people, it is not obvious that it is representative of the true situation especially for high density of cars near critical flow, the only circumstance in which the statistical distribution of cars may have appreciable consequences.

The assumption has been made in Sec. 1 that cars cannot arrive at time intervals less than δ. This assumption has been used at several places in the derivation and, at least for certain single lane highways, is perhaps a reasonable one even for high density flows.

Instead of considering a Poisson distribution which corresponds to a distribution of arrival times with maximum disorder consistent with a known value of the average flow or average time interval between cars, we shall consider here a distribution with maximum disorder consistent with the minimum separation assumption and a known value of the average flow or time interval between cars.

The distribution satisfying these specifications must be of the form

$$\begin{aligned} f(y) &= \alpha e^{-\alpha(t-\delta)} && \text{for} \quad y > \delta \\ &= 0 && \text{for} \quad y < \delta \end{aligned} \tag{23}$$

in which α is to be found from the known value of the average of y, denoted again by d as in the previous case

$$d = \delta + 1/\alpha, \qquad \alpha = (d - \delta)^{-1}; \tag{24}$$

$f(y)$ is essentially a Poisson distribution with a displaced origin.

One easily calculates from (6) that

$$\begin{aligned} f^{(i)}(y) &= \alpha^{i} \frac{e^{-\alpha(y-j\delta)}}{(j-1)!} (y - j\delta)^{i-1} && \text{for} \quad y > j\delta \\ &= 0 && \text{for} \quad y < j\delta \end{aligned} \tag{25}$$

and from (12)

$$f^{(i)}(y, y_1) = \begin{cases} f^{(i)}(y) & \text{for} \quad y_1 < \delta, j > 1 \\ e^{-\alpha(y_1-\delta)} f^{(i)}(y - y_1 + \delta) & \text{for} \quad y_1 > \delta, j > 1. \end{cases} \tag{26}$$

The sequence of approximations represented by Eqs. (16) and (17) does not yield a simple exact solution as it did for the ordered flow, but it does give a very rapidly convergent series of approximations except for flows very close to the critical value.

From the first approximation in (16), one obtains

$$P_1(t_1 \mid y_1) = \begin{cases} e^{-\alpha(y_1-\delta)} & \text{for} \quad \delta < y_1 < T^* \\ 1 & \text{for} \quad 0 < y_1 < \delta \\ 0 & \text{for} \quad y_1 < 0 \quad \text{or} \quad y_1 > T^* \end{cases} \tag{27}$$

and for $l \geq 2$, $P_1(t_l \mid y_1)$ can be found in terms of the function*

$$\begin{aligned} I[u, p] &\equiv \frac{1}{\Gamma(p+1)} \int_0^{u(p+1)^{1/2}} v^p e^{-v}\, dv && \text{for} \quad u > 0 \\ &\equiv 0 && \text{for} \quad u < 0. \end{aligned} \tag{28}$$

To calculate $p(t)$, we again use Eq. (20) but it is convenient to modify the approximation scheme slightly. Substitution of (5) into (20) gives

$$p(t) = \frac{1}{T} \left\{ P(t_1 \mid t_1 - t) + \sum_{l=2}^{n} \int_0^{\infty} dy\, P(t_1 \mid t_l - t - y) f^{(l-1)}(y) \right\}.$$

*This function has been tabulated for various values of u and p in *Tables of the incomplete* Γ*-function* edited by K. Pearson, London, 1922.

If one substitutes (25) into this, one obtains

$$p(t) = T^{-1}\left\{P(t_1 \mid t_1 - t) + \int_0^\infty \alpha\, dz\, P(t_1 \mid t_1 - t - z)e^{-\alpha z} \sum_{l=2}^{n} \frac{(az)^{l-2}}{(l-2)!}\right\}.$$

The summation now represents the first $(n - 2)$ terms in the expansion of $e^{\alpha z}$. It is convenient to write

$$e^{-\alpha z} \sum_{l=2}^{n} (\alpha z)^{l-2}/(l-2)! = 1 - I[\alpha z(n-1)^{-1/2}, n-2].$$

If now one calculates a sequence of approximate values of $p(t)$ from the series of approximate values of $P(t_1 \mid y_1)$, the contribution to $p(t)$ arising from $I[\alpha z(n - 1)^{-1/2}, n - 2]$ is usually small in each such approximation and can logically be incorporated at each step into the next higher approximation. Thus we calculate approximate values of $p(t)$ according to the following scheme.

$$\begin{aligned} p_1(t) &= \frac{1}{T}\left\{P_1(t_1 \mid t_1 - t) + \int_0^\infty \alpha\, dz\, P_1(t_1 \mid t_1 - t - z)\right\}, \\ p_j(t) &= \frac{1}{T}\left\{-\int_0^\infty \alpha\, dz\, P_{j-1}(t_1 \mid t_1 - t - z) I[\alpha z(n-1)^{-1/2}, n-2]\right. \\ &\qquad \left. + P_j(t_1 \mid t_1 - t) + \int_0^\infty \alpha\, dz\, P_j(t_1 \mid t_1 - t - z)\right\} \quad \text{for} \quad j > 1 \\ p(t) &= \sum_{j=1}^{\infty} p_j(t). \end{aligned} \tag{29}$$

Substitution of (27) into (29) for the first approximation gives only elementary integrations resulting in the expression

$$p_1(t) = \begin{cases} 0 & \text{for} \quad t_1 < t \quad \text{or} \quad t < 0 \\ T^{-1}\{1 + \alpha(t_1 - t)\} & \text{for} \quad t_1 - \delta < t < t_1 \\ T^{-1}\{1 + \alpha\delta\} & \text{for} \quad t_1 - T^* < t < t_1 - \delta \\ T^{-1}\{1 + \alpha\delta - e^{-\alpha(T^* - \delta)}\} & \text{for} \quad 0 < t < t_1 - T^* \end{cases} \tag{30}$$

This function is plotted in Fig. 3 where it is compared with the corresponding curve for ordered flow.

The evaluation of the first approximation to $\langle t \rangle$ from $p_1(t)$ is also elementary and gives

$$\langle t \rangle_1 = (1 + \alpha\delta)\frac{t_1^2}{2T} - \frac{(t_1 - T^*)^2}{2T} e^{-\alpha(T^* - \delta)} - \frac{\alpha\delta^2}{2T}\left(t_1 - \frac{\delta}{3}\right). \tag{31}$$

In most practical cases, this is also very close to Clayton's formula. If we assume in (31) that δ and $t_1 - T^*$ are both small compared with t_1, (31) gives

$$\langle t \rangle_1 \sim (1 + \alpha\delta)t_1^2/2T$$

and if we again let

$$\beta = \frac{T}{nd} = \frac{T}{n(\alpha^{-1} + \delta)},$$

we obtain

$$\langle t \rangle_1 \sim \frac{t_1^2}{2T(1 - n\delta\beta/T)}.$$

This is the same as Clayton's formula (22), in which the same approximation of neglecting δ as compared with t_1 is made.

Equation (31) is plotted in Fig. 4 as a function of β along with the corresponding results from the last section.

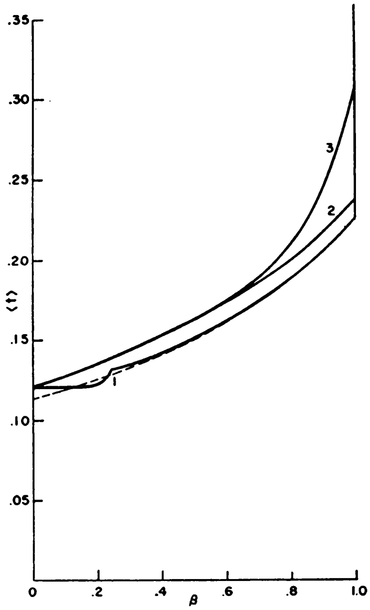

FIG. 4. The average delay $\langle t \rangle$ is plotted as a function of the dimensionless traffic volume β for the same traffic light constants as in Fig. 3. The broken line is Clayton's formula. Solid curve 1, is the curve for ordered flow (Sec. 4). Curve 2 is the first approximation for the disordered flow (Sec. 5) and Curve 3 is the second approximation for disordered flow (Sec. 6).

6. Second approximation—disordered flow. From Eq. (16) and those of the last section, one obtains the following expression for the second approximation $P_2(t_1 \mid y_1)$ for $y_1 < 0$.

$$P_2(t_1 \mid y_1) \begin{cases} = I[\alpha(y_1 + T - n\delta)n^{-1/2}, n - 1] \\ \qquad - I[\alpha(y_1 + T - (n + 1)\delta)(n + 1)^{-1/2}, n] \\ \qquad\qquad \text{for} \quad y_1 < -T + T^* + n\delta \\ = I[\alpha(y_1 + T - n\delta)n^{-1/2}, n - 1] \\ \qquad - I[\alpha(y_1 + T - (n + 1)\delta)(n + 1)^{-1/2}, n] \\ \qquad - \exp[-\alpha(y_1 + T - (n + 1)\delta)]\alpha^n(y_1 + T - T^* - n\delta)^n/n! \\ \qquad\qquad \text{for} \quad -T + T^* + n\delta < y_1 < 0. \end{cases} \tag{32}$$

The substitution of (32) into the various other formulas will lead to some rather cumbersome expressions. Fortunately, the term causing most of the formal complications, namely the term of (32) with the exponential, can usually be neglected.

If one discards this small term, (32) becomes

$$P_2(t_1 \mid y_1) = I[\alpha(y_1 + T - n\delta)n^{-1/2}, n - 1] - I[\alpha(y_1 + T - (n+1)\delta)(n+1)^{-1/2}, n] \quad \text{for} \quad y_1 < 0. \tag{32a}$$

To obtain $P_2(t_1 \mid y_1)$ for $y_1 > 0$, one must substitute this into (11) and take only the terms not already included in the first approximation (27). Such a calculation gives

$$\begin{aligned} P_2(t_1 \mid y_1) &= -I[\alpha(y_1 + T - (n+1)\delta)(n+1)^{-1/2}, n] && \text{for} \quad 0 < y_1 < \delta, \\ &= -e^{-\alpha(y_1-\delta)} I[\alpha(T - n\delta)(n+1)^{-1/2}, n] && \text{for} \quad \delta < y_1 < T^*, \\ &= 0 && T^* < y_1 . \end{aligned} \tag{32b}$$

From (32a, b), one can now compute the next approximation to $P(t_l \mid y_1)$ for $l > 1$, and $p_2(t)$, however we shall pass over this somewhat unpleasant and uninteresting phase of the calculation.

The evaluation of $\langle t \rangle_2$, the second order term for the average delay can be evaluated exactly from (33) with no trouble, but we give below only a much simplified approximate expression which is correct to within a fractional error of $(\alpha\delta)^2/2n$ or $\alpha(t_1 - T)^2 \times \exp[-\alpha(T^* - \delta)]/2n$. In the example considered for illustration, the error is actually less than 3% of the second order term alone. This expression is

$$\langle t \rangle_2 \sim \alpha^{-1}(T - n\delta)(1 + \alpha\delta) \int_0^{\alpha(T-n\delta)} dy\, I[yn^{-1/2}, n - 1] \tag{33}$$

which can be evaluated numerically with the aid of the relation

$$\int_0^z dy\, I[yn^{-1/2}, n - 1] = zI[zn^{-1/2}, n - 1] - nI[z(n+1)^{-1/2}, n].$$

$\langle t \rangle_1 + \langle t \rangle_2$ is plotted as Curve 3 in Fig. 4 for comparison with previous results.

7. Conclusions. Even though the analysis so far does not include an estimate of the errors arising from the mathematical approximations (this can be done, however), certain qualitative conclusions are already quite apparent.

Figure 3 shows that delays for the ordered flow, the first approximation to the disordered flow, and Clayton's formula differ by an amount which would in most cases be considered as negligible in view of the crudeness of the model. Clayton's formula, being the simplest to evaluate, still remains the most useful expression at least for low density flows.

It is apparent that the approximation scheme must break down for $\beta \to 1$ and that $\langle t \rangle$ must become infinite. Even the second approximation gives a correction that is quite sensitive to changes in β near $\beta = 1$ and is very small over a considerable range of β, being of order β^n for small β. The third approximation will be even more sensitive to β near $\beta = 1$, will be of order β^{2n} for small β and will be negligible over a larger range of β than the second approximation.

The assumption that the cars approach the light with a certain minimum separation, has a considerable influence on the behavior of $\langle t \rangle$ near $\beta = 1$. As cars become more densely packed, keeping a certain minimum separation, the uncertainty in the position of the cars decreases at a much faster rate than if cars were allowed to have any separation with the usual Poisson distribution. The delays calculated here will lie somewhere between those one would calculate for the Poisson distribution and those for a completely ordered distribution.

Which type of flow is most representative of the true situation is a difficult question to answer. A completely ordered flow is certainly unrealistic even though it will in many problems, such as that considered here, give simple, qualitatively correct estimates of certain features of the traffic. For low volumes of traffic, the difference between the two possible types of disordered flow is small, but, in any case, as far as delays are concerned, it makes little difference what kind of flow pattern exists.

For high density flows, the degree of disorder does make a difference. Quite likely, the disordered flow considered here is more realistic for single-lane highways and the completely disordered flow more realistic for a multiple-lane highway although the dynamics of the latter situation is not too well defined. A real experimental test of this, however, would not be possible except for nearly critical flows and since the delay is very sensitive to the density, it would be necessary that one have an accurately defined volume of traffic in an equilibrium situation with no transients or disturbances of any kind not specifically taken into consideration here. This is particularly important for high volumes because any slight disturbance of the equilibrium will cause queues to build up very rapidly and dissipate very slowly.

One result of these calculations does seem rather surprising, namely that even to a second approximation, in the example considered, the average delay is less than a third of a cycle. It is also apparent that one will not obtain average delays of more than about one cycle on the basis of this model until the volume of flow has reached a value extremely close to the critical value.

Acknowledgment. The author is indebted to Professor William Prager for instigating this study of traffic problems and for performing himself most of the unpleasant tasks necessary to get such an investigation started.

References

1. D. G. Kendall, *Some problems in the theory of queues*, J. Royal Soc. **B13**, 151 (1951)
2. J. S. Wardrop, *Some theoretical aspects of road traffic Research*, Proc. Inst. Civil Engrs., **II 1**, 325, 1952, and fairly extensive references included in these articles
3. A. J. H. Clayton, *Road traffic calculations*, J. Inst. Civil Engrs. **16**, 247 (1941); see also Wardrop, *loc. cit.*