GAP ACCEPTANCE CHARACTERISTICS
FOR RAMP-FREEWAY SURVEILLANCE AND CONTROL

by

Dr. Donald R. Drew
Assistant Research Engineer

Research Report Number 24-12

Freeway Surveillance and Control
Research Project Number 2-8-61-24

Sponsored by

The Texas Highway Department
In Cooperation with the
U. S. Department of Commerce, Bureau of Public Roads

July, 1965

TEXAS TRANSPORTATION INSTITUTE
Texas A&M University
College Station, Texas
ACKNOWLEDGEMENTS

This study was conducted within the scope of the objectives of the Gulf Freeway Surveillance and Control project. Research in this project is being conducted by the Texas Transportation Institute in cooperation with the Texas Highway Department and Bureau of Public Roads. Sincere appreciation is expressed to these agencies and to Dr. Charles Pinnell, Project Director.

Gratitude is also expressed to Mr. Conrad Dudek, Research Assistant, for his contribution in organizing and conducting the field studies.
SYNOPSIS

The purpose of this investigation was to determine gap acceptance characteristics and merging delay characteristics for six inbound entrance ramps on the Gulf Freeway Surveillance and Control Project. A 20-pen graphic recorder operated by four men stationed adjacent to the merging area between the freeway and the frontage road was used in collecting the data.

Merging vehicles were divided into two groups--those in which the driver rejected gaps before finally accepting a gap and those in which the driver of a ramp vehicle accepted the first gap. The former was referred to as "stopped" vehicles and the latter "moving vehicles." The critical gap for stopped vehicles was found to be about 20 percent higher than for moving vehicles. In addition, it was concluded that the critical gap (median) for the merging maneuver from an entrance ramp is independent of the freeway volume but is apparently affected by ramp geometrics and ramp controls.

A distribution of critical gaps was formed and fitted to a gamma distribution. Merging delay values calculated using this distribution were shown to be higher than those calculated assuming that all drivers have the same fixed critical gap. The calculated values were also compared to observed merging delays.
If the distribution of time spent by merging vehicle at the head of the queue is approximated by a gamma distribution, the entrance ramp merging operation may be considered within the context of classical queueing theory. Based on this queueing model, a ramp metering technique was developed which takes into account the individuality of entrance ramps. Finally, it is shown how the need exists for an automatic ramp control technique combining the microscopic approach developed in this paper combined with the systems or macroscopic approach which has been used as the basis for the past manual ramp metering experiments on the Gulf Freeway.
INTRODUCTION

Freeway Surveillance and Control

The Gulf Freeway Surveillance and Control project sponsored by the Texas Highway Department and U. S. Bureau of Public Roads has the basic objective of developing criteria for the design and operation of automatic surveillance and control systems which would permit the attainment of acceptable levels of service on urban freeways during periods of peak traffic demand.

A traffic surveillance system should involve the continuous sampling of basic traffic characteristics for interpretation by established control parameters, in order to provide a quantitative knowledge of operating conditions necessary for immediate rational control and future design. The chart in Figure 1 illustrates some relationships between many of these control parameters, the basic traffic characteristic, and their ultimate applications. Most of the freeway characteristics and freeway control parameters shown in Figure 1 were evaluated in previous studies and have been discussed in various project research reports.

The Freeway Merging Maneuver

One of the most important elements of freeway operation is the merging maneuver from an entrance ramp onto the shoulder lane of a freeway. Merging, in this case, is simply the absorption of the stream of ramp traffic. A vehicle is said to merge when it moves from the acceleration lane to the shoulder lane. The driver in the merging vehicle must, of course, select the proper moment to execute the merging maneuver based on his judgement as to whether or not a gap which he can enter is large enough for a safe merge.

A merging gap is the interval of time between the arrival of two successive shoulder lane vehicles at a point in the merging area. The concept of gap acceptance is important in describing the interaction of the freeway and ramp streams. It is assumed that the ramp driver measures each gap \( t \) in the shoulder lane and either merges (accepts the gap if \( t > T \)) or waits (rejects the gap if \( t < T \)) where \( T \) may be assumed to be the driver's critical gap for that decision (Figure 2).
ANALYSIS OF FREEWAY TRAFFIC CONGESTION

FIGURE 1
TIME-SPACE RELATIONSHIPS OF FREEWAY MERGING MANEUVER

FIGURE 2
Most of the work on gap acceptance applies to the crossing maneuver at intersections. Greenshields\textsuperscript{4} defined the "acceptable average-minimum time gap" as a gap accepted by half of the drivers. Raff\textsuperscript{5} used a slightly different parameter, the "critical lag". A lag is the interval from the arrival of a minor stream vehicle at the point of conflict to the arrival of the next major stream arrival (See Figure 2). The critical lag is the size lag which has the property that the number of accepted lags shorter than the critical lag is equal to the number of rejected lags longer than the critical lag. For the data presented in Raff's report the Greenshields gap quantity averages 0.2 seconds longer than Raff's critical lag. The principal use of both parameters, or any such parameter, is to simplify the computation of the number of cars delayed by permitting the assumption that all intervals shorter than the critical value (lag or gap) are rejected while all intervals longer are accepted.

It is known that the speed of the merging vehicle is important in considering the distribution of gaps that is acceptable to the merging driver at a freeway entrance ramp. By changing the speed of his own car, a ramp driver also changes his time criterion (critical gap) for a safe merge. In simulating entrance ramp operation the Midwest Research Institute\textsuperscript{6} used a distribution of gaps for nonstop merges with a range of values roughly equal to half those values used for vehicles merging from a stop. Actually, because of the scarcity of published characteristics on entrance ramp operation, the exact relationship between relative speed and gap acceptance is still one of conjecture.

Although numerous data are available regarding merging at an intersection, this information would probably not be applicable as a model for describing ramp merging. The availability of an acceleration lane not only increases the ability of the merging vehicle to minimize its relative speed but provides the driver with a limited opportunity to inspect different parts of the freeway stream for places to merge. Such considerations as a driver's desire to merge as safely, as quickly, or as directly as possible make the problem a question of policy and complicate the exact formulation of a mathematical model.\textsuperscript{7}

**Purpose and Objectives**

The purpose of this investigation is to determine the gap acceptance characteristics for entrance ramps on the Gulf Freeway Surveillance and Control Project. This information will be used to calibrate some of the ramp metering models slated for future use on this project.\textsuperscript{12} It is also believed that this research will help to fulfill the very definite need for gap acceptance data for ramp simulation inputs,\textsuperscript{6} as well as the formulation and testing
of ramp-freeway merging models.\textsuperscript{8}

Specific objectives of this investigation are:

(1) To measure gap acceptance characteristics,

(2) To determine delay characteristics at the merging area, and

(3) To relate these characteristics to freeway operations and design.
STUDY PROCEDURE

All study sites were located on the inbound entrance ramps of the Gulf Freeway in Houston, Texas (see Figure 3). This facility has three 12-foot lanes in each direction. The through lanes of the Gulf Freeway overpass the intersecting roadways producing a "roller-coaster" effect. This design seems to compromise the merging maneuver to the extent that where ramps enter before the overpass some of the acceleration lanes terminate at the overpass structures, and where ramps enter after the overpass the sight distance may be reduced.

A 20-pen graphic recorder operated by four men stationed adjacent to the merging area between the freeway and the frontage road was used in collecting the data. The code for recording traffic behavior is summarized in Figure 4. One pen was assigned to the decision point on the entrance ramp and one pen to a similar point opposite the ramp nose on the outside lane of the freeway. The other two pens were assigned to a point in the merging area: one pen recording both outside lane and merging vehicles and one pen merging vehicles only.

The studies were planned so as to be conducted during periods of congestion-free operation, yet during periods when the ramp demand was high enough to reduce study duration.

The clock times for the four channels shown in Figure 4 were placed on punch cards. A program was written for the IBM 7094 which would output the ramp vehicle arrival time, outside lane arrival time, and time all vehicles passed the designated point in the merging area. The output also included all the gaps rejected by a ramp vehicle plus the gap finally accepted. The study included over 12,000 gaps for six entrance ramps.
LOCATION OF STUDY RAMPS ON GULF FREeway SURVEILLANCE PROJECT

FIGURE 3
SECTION OF RECORDING CHART SHOWING DATA COLLECTION PROCEDURE

FIGURE 4
GAP ACCEPTANCE CHARACTERISTICS

Determination of the Critical Gap

Most studies of gap acceptance have been concerned with the determination of the "critical gap"—a time gap that is just as likely to be accepted as it is to be rejected. The probability that it will be accepted is thus equal to one-half.

In evaluating the critical gap it is apparent that a given gap must be either accepted or rejected by a given driver and as such is binomial in nature. Yet each driver can accept only one gap while he can reject several of them. This means that if all rejected gaps are given the same weight as accepted gaps, then the percentage of intervals accepted for a particular size will not be a true measure of the percentage of drivers who find such an interval acceptable. If the percentage of intervals accepted is to be used to determine the percentage of drivers who are willing to accept them, then the same number of intervals must be counted for each driver. Raff accomplished this by counting only lags and ignoring the gaps.

For this investigation, merging vehicles were divided into two groups—those in which the driver rejected gaps before finally accepting a gap and those in which the driver of a ramp vehicle accepted the first gap. The former is referred to as "stopped" vehicles and the latter "moving vehicles"; the implication being that if a ramp vehicle rejects the first gap he must stop. Theoretically, this would not have to be true where there is an acceleration lane present. Yet, on the Gulf Freeway, it was observed that when traffic is heavy, ramp vehicles tend to shy away from using the acceleration lane, perhaps, for fear of being "trapped" at the end while other ramp vehicles behind them accept the available gaps.

This arbitrary classification of stopped and moving merges afforded the means of eliminating any bias due to the inclusion of all rejected gaps. For example, in the case of stopped vehicles only two gaps were considered for each vehicle—the largest rejected gap and the gap finally accepted. To evaluate the gap acceptance characteristics for moving vehicles only the first gap available to those ramp vehicles not delayed by previous ramp vehicles was considered, much in the same manner that Raff treated lags.

At the Dumble entrance ramp, the number of gaps accepted and rejected have been tabulated in cumulative form in Table 1 for stopped and moving vehicles, and then for all vehicles. The critical gap may be determined algebraically as indicated in the table or graphically as illustrated...
TABLE 1
ACCEPTED AND REJECTED GAPS AT DUMBLE RAMP

<table>
<thead>
<tr>
<th>Length of Gap (Seconds)</th>
<th>Stopped Vehicles</th>
<th>Moving Vehicles</th>
<th>All Vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. Accepted</td>
<td>No. Rejected</td>
<td>No. Accepted</td>
</tr>
<tr>
<td></td>
<td>Gaps&lt; t</td>
<td>Gaps&gt; t</td>
<td>Gaps&lt; t</td>
</tr>
<tr>
<td>0.0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Δt = 0.5</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>1.0</td>
<td>0</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>1.5</td>
<td>0</td>
<td>71</td>
<td>1</td>
</tr>
<tr>
<td>2.0</td>
<td>2</td>
<td>49</td>
<td>7</td>
</tr>
<tr>
<td>2.5</td>
<td>11</td>
<td>34</td>
<td>a = 13</td>
</tr>
<tr>
<td>3.0</td>
<td>a = 15</td>
<td>c = 20</td>
<td>b = 26</td>
</tr>
<tr>
<td>3.5</td>
<td>b = 23</td>
<td>d = 10</td>
<td>38</td>
</tr>
<tr>
<td>4.0</td>
<td>32</td>
<td>5</td>
<td>46</td>
</tr>
<tr>
<td>4.5</td>
<td>41</td>
<td>4</td>
<td>55</td>
</tr>
<tr>
<td>5.0</td>
<td>48</td>
<td>2</td>
<td>63</td>
</tr>
<tr>
<td>5.5</td>
<td>57</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>10.0</td>
<td>100</td>
<td>0</td>
<td>106</td>
</tr>
</tbody>
</table>

Critical Gap, $T = t + \frac{(c-a) \Delta t}{(b+c)-(a+d)}$; $T$ (Stopped) = 3.1; $T$ (Moving) = 2.5; $T$ (All) = 2.8
in Figure 5. Two cumulative distribution curves are shown which depict the number of accepted gaps shorter than $t$ and the number of rejected gaps longer than $t$. The value of $t$ for which these two curves intersect is the critical gap.

Bissell\textsuperscript{9} calculated the percentage of acceptance observed at an intersection for each time group, plotted the data on log probability paper, and then drew a straight line through the point. Solberg and Oppenlander\textsuperscript{10} also calculated the percentage acceptance, but converted the probability scale to "probits" and then fitted a regression line with confidence limits through the points. The critical gap was taken as the median gap--the 50 percentile in the case of Bissell, the probit of acceptance equal to 5.0 in the case of Solberg and Oppenlander.

Figure 6 illustrates a procedure combining some of the features of the Raff method on the one hand and the Bissell method on the other. Using Table 1 the number of accepted gaps less than $t$ was divided by the sum of the number of accepted gaps less than $t$ plus the number of rejected gaps greater than $t$. A straight line is drawn through these points on the log-probability graph (Figure 6) with the 50 percentile or median gap used to establish the critical gap.

The reason for using this new procedure for establishing the critical gap is that it serves to define the parameter in a way that relates directly to the manner in which it is used. The probability of accepting a smaller gap than the critical gap is exactly equal to the probability of rejecting a longer gap, thus equating the critical gap to the median gap for a distribution of critical gaps. This concept will be discussed in more detail later in the paper.

Factors Affecting the Critical Gap

Raff\textsuperscript{5} investigated factors influencing the critical lag at urban intersections and concluded that the critical lag (1) varies directly with main street speeds and main street width; (2) is independent of the main street volume; and (3) is influenced by sight obstructions and the side-street traffic pattern.

The effect of relative speed on the critical gap for merging ramp vehicles is evident in Figure 6. At the Dumble entrance ramp, the critical gap for moving vehicles (based on the first gap evaluated) is seen to be 2.5 seconds as against a critical gap of 3.1 seconds for stopped vehicles (vehicles rejecting at least one gap before finally accepting a gap).

The graph of the percent acceptance for merging vehicles at different ramps
NUMBER OF REJECTED GAPS $> t$

NUMBER OF ACCEPTED GAPS $< t$

ACCEPTED AND REJECTED GAPS FOR ALL VEHICLES AT DUMBLE RAMP

FIGURE 5
PERCENT ACCEPTANCE FOR MERGING VEHICLES

DUMBLE ENTRANCE RAMP

PERCENT ACCEPTANCE

GAP (SECONDS)
(Figure 7) indicates that the critical gaps were not the same, but varied from ramp to ramp just as they do among intersections. In order to determine to what extent this variability was due to ramp geometrics, the relationship between critical gap and entrance ramp elements has been summarized in Table 2. First of all, it is apparent from the table that the assumption that the critical gap for a ramp is independent of the freeway outside lane volume is justifiable. Secondly, analyzing the table objectively, there is no significant correlation between any single element and the critical gap. Subjectively however, to the person familiar with the study ramps, it was not surprising that the Cullen, Dumble and Wayside ramps exhibited lower critical gaps than the Griggs, Mossrose and Woodridge ramps. The Griggs and Mossrose ramps have taper-type acceleration lanes which terminate abruptly at overpass structures, whereas operation at the Woodridge ramp is complicated by a high grade differential between the ramp (upgrade) and freeway (downgrade).

In order to be certain that the variability in the critical gaps between ramps was due to differences in the ramps rather than random fluctuations in the critical gap parameter, the gap acceptance studies were repeated at a single ramp at different times. Figure 8 illustrates that the critical gap is a stable measure of the operational performance of a given ramp. The parameter, like capacity, is a measure of the ability of a traffic stream to absorb merging vehicles from a given ramp.

In the case of intersections, the side street traffic pattern (that is whether the side streets carried one-way or two-way traffic) seemed to affect the critical gap appreciably. Of course, ramps are limited to one-way operation, but ramp metering is a form of operational control which could change the minor street traffic pattern. One objective of this study was to observe some of the characteristics associated with semi-automatic metering in order to determine equipment requirements for future automatic ramp control systems. This phase of the study was conducted at the Dumble on-ramp. A post-mounted traffic signal with red, amber, and green lenses was installed on the frontage road at the P. C. of the on-ramp. Overhead signals mounted over each lane of the frontage road were employed to separate the two movements (ramp usage and frontage road usage). The signal phasing was designed for bulk-service metering utilizing a three-dial pre-timed controller with a 30-second cycle length. The three dials were set to give 10 1/2, 8 and 13 1/2 seconds of green with a constant amber of 2 1/2 seconds capable of dispatching platoons of 4, 3, and 5 vehicles per cycle.

The effect of this technique of ramp metering on the critical gap for moving vehicles is illustrated in Figure 9. The critical gap for bulk service metering is seen to be 3.1 seconds as against 2.5 seconds for normal operation. The reason for this is not clear, but it is suspected that (1) metered vehicles have a greater relative speed and (2) metered drivers are more conscious of the merging maneuver and are therefore more cautious.
VARIATIONS IN PERCENT ACCEPTANCE FOR MERGING VEHICLES AT DIFFERENT RAMPS

FIGURE 7
### TABLE 2
RELATIONSHIP BETWEEN CRITICAL GAP AND OTHER ENTRANCE RAMP ELEMENTS

<table>
<thead>
<tr>
<th>Ramp</th>
<th>Cullen</th>
<th>Dumble</th>
<th>Wayside</th>
<th>Griggs</th>
<th>Mossrose</th>
<th>Woodridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical Gap (Sec.)</td>
<td>2.7</td>
<td>2.8</td>
<td>3.3</td>
<td>3.7</td>
<td>4.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Avg. Freeway Volume Outside Lane During Period of Observation (veh. per hour)</td>
<td>1350</td>
<td>1240</td>
<td>1100</td>
<td>700</td>
<td>1060</td>
<td>970</td>
</tr>
<tr>
<td>Percent Grade on Freeway</td>
<td>0.0</td>
<td>+0.5</td>
<td>-1.5</td>
<td>+1.5</td>
<td>+1.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>Ramp Length (feet)</td>
<td>320</td>
<td>320</td>
<td>330</td>
<td>230</td>
<td>190</td>
<td>300</td>
</tr>
<tr>
<td>Ramp Width (feet)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Angle of Entry (deg.)</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Acceleration Lane Length (feet)</td>
<td>--</td>
<td>330</td>
<td>--</td>
<td>310*</td>
<td>250*</td>
<td>370</td>
</tr>
<tr>
<td>Auxilliary Lane Length (feet)</td>
<td>350</td>
<td>--</td>
<td>510</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

*Taper type acceleration lane terminating at an overpass structure.*
VARIATIONS IN PERCENT ACCEPTANCE FOR MERGING VEHICLES AT THE SAME RAMP

FIGURE 8
COMPARISON OF PERCENT ACCEPTANCE FOR MOVING Merging VEHICLES DURING NORMAL AND CONTROLLED OPERATION

FIGURE 9
Distribution of Critical Gaps

The critical gap for merging indicates how large a time interval is required for the typical ramp driver to enter a freeway. The use of a single typical figure rather than the whole range of observed human behavior has the advantage of simplicity in such advanced stages of the analysis of ramp behavior as delay calculations. In spite of the importance of the critical gap concept, very little consideration has been given to the properties of the parameter. True, Bissell estimated the standard deviation of the gap acceptance distribution by taking the difference between the median acceptance value and the time corresponding to an acceptance of 15.9 per cent. Yet, it is a well known fact that few traffic phenomena can be described by a single measure of central tendency and dispersion, but can be described by a frequency distribution.

One technique explored in this investigation to obtain a frequency distribution for critical gaps is illustrated in Figure 10. Only drivers who rejected at least one gap before merging were considered. Assuming that a driver who accepts a given gap size at a certain ramp can be expected to accept any gaps of great length and that a driver who rejects a given gap size can in a similar manner be expected to reject all gaps of shorter length, it is then evident that a driver's critical gap must lie somewhere in the range between the largest gap he rejected and the gap finally accepted. The number of times the critical gap range included a certain gap interval divided by the total number of gap intervals establishes the observed frequency.

The histogram of observed frequencies obtained from Figure 10 is shown in Figure 11. Since the task of gap selection is primarily a function of the human component of the composite driver-vehicle population, a logical choice in fitting a probability distribution to the observed data is the normal distribution. Since the normal distribution is completely determined by its mean and standard deviation and these quantities are rather accurately estimated from the histogram, a normal curve was easily fitted to the observed data (See Figure 11).

Although there is reasonably good agreement between the normal curve and the data, there are certain theoretical not to mention practical considerations which suggest a less restrictive distribution. The principal characteristics of the normal distribution are (1) it is symmetrical, (2) it assigns a finite probability to every finite deviation, and (3) the mode or most probable result is equal to the mean. Of course, possession of these three requisites is by no means general in considering vehicular traffic characteristics. It is unrealistic to think of a driver with a negative critical gap. Yet, if critical gaps were distributed normally, the second property of the distribution would assign a finite probability a negative critical gap.
### Limits of Critical Gaps for Stopped Vehicles at Dumble Inbound Ramp

**Figure 10**

<table>
<thead>
<tr>
<th>Length of Gap (Seconds)</th>
<th>1</th>
<th>6</th>
<th>11</th>
<th>16</th>
<th>21</th>
<th>26</th>
<th>31</th>
<th>35</th>
<th>36</th>
<th>41</th>
<th>46</th>
<th>51</th>
<th>56</th>
<th>61</th>
<th>66</th>
<th>Sum of Gaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0-0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0.5-0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>1.0-1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>1.5-1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>2.0-2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>2.5-2.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>3.0-3.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>47</td>
</tr>
<tr>
<td>3.5-3.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>4.0-4.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>4.5-4.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>5.0-5.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>5.5-5.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

A gap of between 4.0-4.4 seconds was accepted after rejecting a maximum gap of 0.5-0.9 sec.
COMPARISON OF OBSERVED AND THEORETICAL CRITICAL GAP DISTRIBUTIONS

FIGURE II
Among the families of distributions which have been formulated for the purpose of enabling the statistician to deal with a wide variety of data are those developed by Karl Pearson. Pearson noted that the binomial and Poisson distributions satisfy the differential equation

\[
\frac{dP}{dx} = \frac{(a + x)P}{b + cx + dx^2}
\]

for some set of values of the constants \(a\), \(b\), \(c\) and \(d\). Certain solutions of this equation have been sorted out, largely because of their algebraic simplicity to form the Pearson frequency distributions.

Because of the non-negative restriction, we are looking for a curve limited at one end. The Pearson Type III distribution is of this type (see Appendix). Because, this distribution utilizes the gamma function it is also known as the gamma distribution. The element of randomness inherent in this distribution, as exemplified by its formulation from the Poisson distribution, gives it some conceptional appeal for describing the critical gap phenomenon. In Figure 11, the Pearson Type III (Gamma) curve has been superimposed on the histogram along with the normal curve.

The cumulative type curves of Figures 6-9 suggest a second procedure for obtaining the desired frequency distribution. A distribution of minimum or critical gaps can be estimated by finding the proportion of drivers who accept a certain sized gap and the proportion who accept a slightly longer gap. The difference in these proportions then gives an estimate of the proportion of drivers who will accept the slightly longer gap but not the shorter one, i.e., the proportion of drivers whose critical gap is the slightly longer gap. A summary of the critical gap distribution moments obtained using this second method appears in Table 3. The Pearson Type III distribution was fitted to these frequency distributions; the estimates of the parameters of the theoretical distribution are also summarized in the Table 3.

There is a subtle yet distinct difference between the frequency distributions obtained by the two methods. The frequency distribution obtained using the method of Figures 10 and 11 enables one to predict the probability of a driver having a certain critical gap. The area under the frequency distribution obtained from the cumulative distribution, on the other hand, gives the probability of a driver accepting a certain size gap.

What is desired is a description of the critical gaps among a population of drivers. The mode, mean and variance of the critical gap distribution add significantly to this description and therefore must be considered in the applications along with the median critical gap, as shall be shown.
TABLE 3
SUMMARY OF CRITICAL GAP MOMENTS AND PARAMETERS OF THE PEARSON TYPE III DISTRIBUTION APPROXIMATING THE OBSERVED CRITICAL GAP DISTRIBUTION

<table>
<thead>
<tr>
<th>Ramp</th>
<th>Critical Gap Parameters</th>
<th>Pearson Type III Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>(Seconds)</td>
<td></td>
</tr>
<tr>
<td>Cullen</td>
<td>2.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Dumble</td>
<td>2.8</td>
<td>3.2</td>
</tr>
<tr>
<td>Wayside</td>
<td>3.3</td>
<td>3.6</td>
</tr>
<tr>
<td>Griggs</td>
<td>3.7</td>
<td>4.1</td>
</tr>
<tr>
<td>Mossrose</td>
<td>4.7</td>
<td>5.2</td>
</tr>
<tr>
<td>Woodridge</td>
<td>3.8</td>
<td>4.2</td>
</tr>
</tbody>
</table>
DELAY CHARACTERISTICS

Gap Availability

In using mathematics to estimate delay due to merging, it is necessary to have a description of the distribution of time headways or gaps in the outside freeway lane as well as a description of the critical gap.

The Erlang frequency distribution, a special case of the Pearson Type III distribution in which the parameter "a" must be a positive integer, combines many of the advantages of the parent distribution plus simplicity in application. The cumulative form of the Pearson Type III distribution must be evaluated by either numerical methods or tables of the incomplete gamma function unless "a" is a positive integer, making it difficult to determine the probability of a ramp driver finding a gap larger than the critical gap, for example.

The cumulative frequency curves in Figure 12 are offered to show the theoretical relationship between gap availability and the critical gap, assuming that the distribution of headways on the outside lane of the freeway may be described by an Erlang distribution. Although use of the Erlang distribution affords the opportunity of considering the distribution of freeway vehicles for all cases from randomness (a=1) to complete uniformity (a=∞), experience has shown that the first four curves (a = 1, 2, 3, 4) are adequate for the description of vehicular headways in a traffic stream.

Determination of Merging Delays

In the theoretical approach to describing a system composed of two interacting streams, it is necessary to determine the average time for vehicles in the minor stream to merge. It may be assumed that the waiting driver measures each time gap, $t$, in the traffic on the major highway until he finds an acceptable gap, $T$, which he believes to be of sufficient length to permit his safe entry. If he accepts the first gap ($t>T$), his waiting time is zero. If he rejects the first gap ($t<T$), but accepts the second gap, his expected waiting time would be one interval. If we assume that the driver's gap acceptance policy does not change with time, then by induction, the individual waiting periods form a geometric distribution and the probability, $P_n$, if any driver having to wait for $n$ intervals each less than $T$ seconds before merging is

$$P_n = p^n (1-p), \quad n = 1, 2, \ldots,$$ (1)
CUMULATIVE PROBABILITY CURVES FOR THE ERLANG DISTRIBUTION

\[ \int_{T}^{\infty} \frac{q^a t^{a-1}}{(a-1)!} e^{-aq} dt \] WHERE \( q = \text{FLOW (VPH)} \)

FIGURE 12
where

\[ p = P(t < T) = \int_0^T f(t) \, dt \quad (2) \]

and \( f(t) \) is the distribution of gaps in the major stream. The expected number of intervals for which a driver has to wait is given by

\[ E(n) = \frac{p}{1-p} = \frac{\int_0^T f(t) \, dt}{\int_{\infty}^T f(t) \, dt} \quad (3) \]

The average time for a ramp vehicle which is in position to merge to find an acceptable gap in the freeway traffic stream will be the product of the expected number of intervals less than \( T \), \( E(n) \), and the average length of interval. The average length of interval less than \( T \) is, in turn, equal to the total time less than \( T \) seconds divided by the number of intervals less than \( T \) seconds:

\[ \text{Aver. length of intervals } < T = \frac{\int_0^T t f(t) \, dt}{\int_0^T f(t) \, dt} \quad (4) \]

where \( q \) is the rate of flow. Multiplying equation (3) and (4) yields the average waiting time for a ramp vehicle in position to merge:

\[ d = \frac{\int_0^T t f(t) \, dt}{\int_{\infty}^T f(t) \, dt} \quad (5) \]

Recalling that the proportion of ramp vehicles actually delayed is given by (2), it is apparent that the average waiting time of those who suffer delay is

\[ d' = d/\int_0^T f(t) \, dt \quad (6) \]
It should be remembered that the delays expressed in equation (5) and (6) are for single vehicles approaching the merging area.

Theoretical solutions\textsuperscript{13,14} to the determination of delays for vehicles attempting crossing or merging maneuvers have been based on the assumption that the main stream of traffic is exponentially distributed,

\[ f(t) = qe^{-qt} \]  
(7)

Substituting (7) in (5), the mean delay \( d \) for a critical gap \( T \) and freeway lane flow \( q \) is

\[ d = q^{-1} (eqT - qT - 1) \]  
(8)

If the more general Erlang distribution is used to describe the distribution of main stream gaps, then --

\[ f(t) = (aq)^a t^{a-1} e^{-aqT}, \quad (a = 1, 2, \ldots) \]  
(9)

and the mean delay for \( a = 2, 3 \) and 4 is

\[ d_{a=2} = \frac{e^{2qT} - 2(qT)^2 - 2qT - 1}{q(2qT + 1)} \]  
(10)

\[ d_{a=3} = \frac{e^{3qT} - 4.5(qT)^3 - 4.5(qT)^2 - 3qT - 1}{q[4.5(qT)^2 + 3qT + 1]} \]  
(11)

\[ d_{a=4} = \frac{e^{4qT} - 10.67(qT)^4 - 10.67(qT)^3 - 8(qT)^2 - 4qT - 1}{q[10.67(qT)^3 + 8(qT)^2 + 4qT + 1]} \]  
(12)

Since the negative exponential is a special case of the Erlang distribution, \( d_{a=1} \) is given by Equation 8. Equation 8, 10, 11 and 12 are plotted in Figure 13.

It must be remembered that the theoretical delay values are based on a fixed critical gap for all drivers. Blunden\textsuperscript{11} correctly suggests that a more complete description of delays to traffic would be given if the fixed critical gap were replaced with a distribution of critical gaps, \( g(T) \). If it is assumed that the critical gap distribution is of the gamma type

\[ g(T) = \frac{b^a}{(a-1)!} T^{a-1} e^{-bT} \]  
(13)

a reasonable assumption as seen by Figure 11, then the mean delay \( D \) is readily obtainable from
MERGING DELAY IN TERMS OF THE FREEWAY FLOW $q$, CRITICAL GAP $T$, AND ERLANG CONSTANT $\alpha$.

**FIGURE 13**
\[ D = \int_{0}^{\infty} d(T) g(T) dT \]  

(14)

Substituting the delay function given by (8) and the critical gap distribution given by (13) in (14) yields

\[ D = q^{-1} \int_{0}^{\infty} \frac{e^{qT} T g(T) dT}{q} \int_{0}^{\infty} T g(T) dT - q^{-1} \int_{0}^{\infty} g(T) dT \]  

(15)

Realizing that the second term is the form for the mean of a distribution and the integral in the last term must equal unity, gives

\[ D = q^{-1} \int_{0}^{\infty} e^{qT} g(T) dT - T - q^{-1} \]  

(16)

Letting \( b = a/T \), the first term may be put in the form

\[ \frac{(a/T)^a}{q(a-qT)^a} \int_{0}^{\infty} \frac{a^{-qT a}}{T} \frac{T^{a-1}}{(a-1)!} e^{-(a-q\bar{T}) T/\bar{T}} dT \]  

(17)

\[ = \frac{1}{q} \left( \frac{a}{a-q\bar{T}} \right)^a \]  

(18)

provided that \( a>q\bar{T} \), since the integral represents the area under a gamma distribution with parameters \( a \) and \( (a-q\bar{T}) / \bar{T} \). Gathering terms, the general expression for delay to a merging vehicle is

\[ D = \frac{1}{q} \left( \frac{a}{a-q\bar{T}} \right)^a - \bar{T} - \frac{1}{q} \]  

(19)

It is interesting to note that (8) can be verified as the special case of (19) in which the critical gap is taken as a fixed value equal to the mean critical gap. For \( a = \infty \), we know that the variance of the gamma distribution is zero; so we may interpret \( g(T; u, \infty) \) as the family of situations for which the critical gap is constant and has the value \( \bar{T} \). From the definition of \( e \) it is apparent that

\[ \lim_{a \to \infty \left( 1 - q\bar{T} \right)}^{-a} = e^{q\bar{T}} \]  

(20)

Substitution of (20) in (19) gives (8).

Delay values calculated from (19) using finite values of "a" will always be higher than those obtained from (8) assuming a fixed critical gap. Moreover, the fact that the delay obtained using the fixed value is the special case of the delay obtained using the distribution for a variance of zero proves that the mean of the critical gap distribution, not
the median, should be used in delay computations.

Observed average merging delays were compared to the average merging delays predicted by the equations (using the mean critical gap) and the results are summarized in Figure 14. The value of "a" for the Erlang distribution used to enter the graphs in Figure 13 was calculated from the time headways in the outside lane using the method of moments in which "a" is the integer closest to the quotient obtained by dividing the mean headway squared by the variance in the headways. The merging delay concepts discussed in this section become important as inputs in a queueing model developed in the next section.

Queueing Considerations

A queueing system is composed of three elements: (1) a demand or flow of arrival requiring service, (2) some restriction on the availability of service, and (3) irregularity in either the demand or in the servicing operation or in both.

Consider vehicles on an entrance ramp arriving at the merging area at a rate described by \( f(q_r) \) where \( q_r \) is the average rate of arrivals. These vehicles are obliged to yield to the freeway traffic, thus forming a single lane waiting for successive vehicles at the head of the line to merge. If the distribution of time spent by vehicles at the head of the queue is \( f(\mu) \), then it is apparent that an entrance ramp merging operation is within the realm of a classical queueing system. In this section using Kendall's approach for a queueing system with random inputs and arbitrary service times, formulae are developed for the mean queue length and mean waiting time for a queue of ramp vehicles waiting to merge.

Let \( n_0, n_1 \) be the ramp queue lengths immediately after two successive ramp vehicles \( C_0, C_1 \) have merged, \( t \) be the service time of \( C_1 \), \( r \) be the number of ramp vehicles arriving while \( C_1 \) is being served. If a random variable \( \delta \) is introduced such that \( \delta = 1 \) if \( n_0 = 0 \) and \( \delta = 0 \) if \( n_0 \neq 0 \), then it follows that

\[
  n_1 = n_0 + r - 1 + \delta \quad (21)
\]

It is to be noted from the definition of \( \delta \) that

\[
  \delta^2 = \delta
\]

and

\[
  n_0 \delta = 0
\]

30
OBSERVED MERGING DELAYS AT VARIOUS RAMPS (SECONDS)

THEORETICAL MERGING DELAYS USING GRAPHS IN FIGURE 13

COMPARISON OF OBSERVED AND THEORETICAL MERGING DELAYS AT STUDY RAMPS

FIGURE 14
and hence, from (21), on taking expected values, we obtain

\[ E(n_1) = E(n_0) + E(r) - 1 + E(\delta). \]  

(22)

If the system is assumed to be in a state of statistical equilibrium, \( E(n_1) = E(n_0) \) and

\[ E(r) = \frac{q_r}{\mu} = \rho. \]

Thus substituting in (22)

\[ E(\delta) = 1 - \rho. \]

Squaring both sides of (21) and taking expected values as before,

\[ E(r-1)^2 + E(\delta^2) + 2E[n_0(r-1)] + 2E[\delta(r-1)] = 0, \]

which reduces to

\[ E(n_0) = \rho + \frac{E(r^2) - \rho^2}{2(1-\rho)}. \]  

(23)

It is now necessary to calculate \( E(r^2) \), the second moment of the number of arrivals in the service time \( T \), making use of its relationship to the mean and variance in arrivals. Assuming that ramp arrivals are Poisson and remembering that "averaging" here must be carried out with respect to both \( r \) and the service time \( t \), we have

\[ E(r^2) = q_r E(t) + q_r^2 E(t^2) \]

Since \( E(t) = \mu^{-1} \) and

\[ E(t^2) = \sigma^2 + E(t)^2 \]

then

\[ E(r^2) = \rho + \rho^2 + q_r^2 \sigma^2. \]  

(24)

Substituting (24) in (23) gives the expected queue length on the ramp as

\[ E(n_1) = \frac{q_r}{\mu} + \frac{\mu q_r^2 (\mu^{-2} + \sigma^2)}{2 (\mu - q_r)}. \]  

(25)

If \( w \) is the waiting time (before merging) of \( C_1 \), then \( n_1 \) ramp vehicles arrive in time \( t + w \). Thus since the mean arrival rate is \( q_r \),
\[ E(n)_r = q_r \cdot E(t + w) \]

It follows that the average waiting time for a ramp vehicle before merging is

\[ E(w) = \left[ E(n)_r / q_r \right] - \mu^{-1} \tag{26} \]

and the mean wait in the system for a ramp vehicle is

\[ E(v) = E(n)_r / q_r \tag{27} \]

Major and Buckley\textsuperscript{16} have interpreted the service time for the queue \( \mu^{-1} \) as identical to the summation of the rejected gaps for a ramp vehicle in position to merge (what we have referred to as the merging delay \( d \) in this paper). The Pearson Type III distribution was fitted to the observed distribution of delays at each ramp and the results are summarized in Table 4 for all ramps in the study. Based on the assumptions that the average service time is equal to the average merging delay and the distribution of service times is of the form of the Pearson Type III distribution, we obtain from (25)

\[ E(n)_r = dq_r + \frac{(dq_r)^2 (1 + a^{-1})}{2(1 - dq_r)} \tag{28} \]

where \( d \) is given by equation 8, 10, 11 or 12 depending upon the distribution of freeway headways and \( a \) is the parameter describing the distribution of service times.

If the \( a \) value in (28) calculated in Table 4 is rounded off to unity, equations 25-28 take the form of the conventional, Poisson - negative exponential, queueing formulas.
TABLE 4

SUMMARY OF SERVICE TIME MOMENTS AND PEARSON TYPE III PARAMETER "a"

<table>
<thead>
<tr>
<th>Ramp</th>
<th>Service Time Moments</th>
<th>Type III &quot;a&quot; Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td>Cullen</td>
<td>4.0</td>
<td>5.2</td>
</tr>
<tr>
<td>Dumble</td>
<td>3.5</td>
<td>5.5</td>
</tr>
<tr>
<td>Wayside</td>
<td>6.2</td>
<td>8.8</td>
</tr>
<tr>
<td>Griggs</td>
<td>1.5</td>
<td>2.8</td>
</tr>
<tr>
<td>Mossrose</td>
<td>4.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Woodridge</td>
<td>4.6</td>
<td>6.0</td>
</tr>
</tbody>
</table>
APPLICATIONS

Ramp Control

Almost any engineering problem may be described as a systematic attempt to resolve a capacity-demand relationship at an acceptable level of service. We try to build enough strength into the materials of a layered pavement system, for example, to withstand shear stresses due to anticipated loads. However, the mere fact that the strength (capacity) exceeds the load stresses (demand) does not guarantee an acceptable level of service. The deflection, smoothness, texture and color contrast also affect the driver's ride and, as such, are level of service factors that must be considered.

Traffic engineering is the science of measuring traffic characteristics and the application of this information to the design and operation of traffic systems. The traffic engineer's basic problem of resolving a capacity-demand relationship is similar to that of any other engineer: he must be able to either measure the parameters defining capacity and demand very accurately, or he must be able to control them. Returning to the pavement design analogy: although the strength of the materials in a pavement can probably not be estimated as accurately as the capacity of a freeway lane the pavement designer knows that the loads (demand) on the facility are controlled, and in most states limited by law. If urban freeways are to operate at the levels of service for which they were designed, the demand on these facilities must also be regulated.

Because the control of vehicles entering the freeway, as against the control of vehicles already on the freeway, offers a more positive means of preventing congestion, considerable emphasis is being placed on the technique of ramp metering. Any scheme for controlling the rate of flow of vehicles from an entrance ramp onto the freeway should be based on, and be capable of, reacting to operation on the freeway lanes. Chicago\textsuperscript{17} utilizes an occupancy (density) measurement in the middle freeway lane in advance of the merging area. Wattleworth\textsuperscript{18} proposes a systems approach in which the capacity of a freeway system is limited to the section of smallest capacity (bottleneck). If the bottleneck capacity is known, all ramps in the system can be metered according to the differential between the upstream freeway demand and bottleneck capacity.

The key to entrance ramp operation is, however, the synthesis of two lanes of traffic into a single lane. For this reason, it is believed that a microscopic approach - one sensitive to the operating conditions on the outside freeway lane - should be considered. Drew\textsuperscript{7} describes a "moving queues" model based on coordinating ramp metering with the detection of acceptable gaps in the outside freeway lane. Figure 15 illustrates
ENTRANCE RAMP

METERING LOCATION

TF

OUTPUT

IS

NO

T = Tc

YES

N = N + 1

Q = Q + 1

NO

t - Tq > 0

YES

qR = Q / Tc

q = N / Tc

E(n) = N / Q

INPUT

N

t

COMPUTER

OUTPUT

q

E(n)

APPLICATION OF QUEUEING MODEL TO FREEWAY CONTROL

FIGURE 15

NOTATION

N = VEHICLE COUNTER

t = TIME BETWEEN VEHICLE COUNTS

TF = FREEWAY TRAVEL TIME

TR = RAMP TRAVEL TIME

Tc = TF - TR

qR = METERING RATE

q = FLOW

E(n) = AVERAGE QUEUE LENGTH

Q = QUEUE COUNTER

T = REAL TIME LAPSED
the application of this model to the control of a merging situation at an entrance ramp. Thus, a vehicle in the outside lane of the freeway with a time headway less than the arbitrary queueing headway $T_Q$ is considered to be queued to the preceding vehicle. The control system illustrated consists of the flow of information from a detector located on the outside freeway lane to a computer and then to the metering signal on the ramp. For the "closed loop" system pictured, either a digital or analog computing device could be utilized. However, use of the former would necessitate a reduction in the "time constant" (time over which traffic conditions are averaged) by an interval equal to the time necessary for computation.

In considering an example, suppose the travel time during the peak period from detector to merging area is $T_F = 35$ sec. and from the metering station to the merging area is $T_R = 5$ sec. The critical gap (that headway in the outside freeway lane for which an equal percentage of ramp traffic will accept a smaller headway as will reject a larger one) is assumed to be 2.5 seconds. It is apparent that if control adjustment is to be made during the same period as detection, the time constant $T_c$ cannot be greater than $T_F - T_R$. Moreover, if the arbitrary queueing headway $T_Q$ is equated to the critical gap, the number of platoons or moving queues $Q$ will equal the number of critical gaps. The latter determines $q_r$, the number of ramp vehicles that can merge during $T_c$. Thus, in the example the dials on the controller would be set to $T = 30$ sec. and $T_Q = 2.5$ sec. If during $T_c$, $N = 10$ vehicles were detected and $Q = 5$ of the headways were greater than the queueing headway on the dial ($T > T_Q$), the metering rate during $T_c$ would be:

$$q_r = \frac{Q}{T_c} = \frac{5 \text{ veh.}}{30 \text{ sec.}} = 1 \text{ veh. every 6 sec.}$$

The rate of flow and congestion index at the detection station during $T_c$ would be:

$$q = \frac{N}{T_c} = \frac{10 \text{ veh.}}{30 \text{ sec.}} = 1 \text{ veh. every 3 sec.}$$

$$E(n) = \frac{N}{Q} = \frac{10 \text{ vehicles}}{5 \text{ queues}} = 2 \text{ vehs./queue.}$$

It is apparent from the variability of the critical gap among entrance ramps (See Figure 7) that controller settings would vary from entrance ramp to entrance ramp. The geometrics of the freeway would also affect the detector location and hence the time constant.
The significance of the control parameter $E(n)$ is twofold (1) by definition, it is the average number of vehicles in a freeway platoon and as such is a rational congestion index (a value of unity would be free flow, whereas at jam concentration $E(n)$ would approach infinity), (2) the moving queue length, $E(n)$, was formulated in such a way that it is the reciprocal of the probability of getting a gap larger than the queueing criteria, and therefore, is actually a measure of gap availability.

**Comparison of Ramp Metering Criteria**

An important contribution of this microscopic approach to ramp control is that the individuality of entrance ramps is taken into consideration. Figure 16 illustrates the relationship between the freeway outside lane volume, $q$, the headway distribution of the freeway volume $a$, the critical gap $T$, and the ramp volume $q_r$. Consider an entrance ramp operating with a critical gap of 4.0 seconds and with the distribution of freeway traffic conforming to an Erlang distribution with $a = 2$. It is apparent that the sum of the coordinates of any point on the line $T = 4$ in the graph $a = 2$ describes the merging service volume for that ramp. For example, the point described by $q = 1500$ and $q_r = 120$ tells us that the merging service volume is 1620, and that under these operating characteristics, a ramp arrival has a 67% chance of finding the ramp empty or a 33% of finding a vehicle ahead of it trying to merge.

The effect of poor ramp geometries is evident. If in the previous example the critical gap was 5.0 seconds, the maximum ramp service volume $q_r$ for the same freeway volume ($q = 1500$) would be 50 vph or a total merging service volume for the ramp of 1550 vph.

The graphs also point out the differences in the macroscopic philosophies of ramp metering and the microscopic approach explained in this paper. In the macroscopic approach, metering would be based on one of the curved lines (one representing the boundary between stable and unstable flow and the other representing possible capacity) regardless of the ramp geometrics or critical gap. This means that for all conditions except those described as unstable flow on the graph vehicles would be metered at a faster rate than the service rate at the merging area (available critical gaps) encouraging drivers to either accept smaller gaps than the critical gap or become part of a steadily growing queue at the merging area.

Figure 16 clearly illustrates the need for a ramp control technique combining both the macroscopic and microscopic approach. For conditions described on the graph by "unstable flow," the ramp geometrics do not govern and hence the macroscopic approach based on the downstream bottleneck service volume applies. However, to the left of the 1800 vph line
WHERE "a" REFERS TO THE ERLANG PARAMETER DESCRIBING THE DISTRIBUTION OF FREEWAY ARRIVALS AND T IS THE CRITICAL GAP

MAXIMUM SERVICE VOLUME \( q_r \) FOR A RAMP ARRIVAL TO HAVE A PROBABILITY OF .67 OF FINDING NO RAMP VEHICLES IN THE MERGING AREA

FIGURE 16
dividing stable and unstable flow, the critical gap governs since the merging service volume is less than the bottleneck service volume. (It should be explained that the term "service volume" is used here consistent with the level of service concept proposed in the 1965 Highway Capacity Manual, although the word capacity could be substituted in this particular comparison with only the risk in a loss of generality.)

Ramp Design

After the determination of the number of freeway lanes in a rational design procedure, the operating conditions at critical locations of the freeway must be investigated for the effect on capacity and level of service. Unless some designated level of service is met at every point on the freeway, bottlenecks will occur and traffic operation will break down. Critical locations on a freeway are manifest by either sudden increases in traffic demand, the creation of inter-vehicular conflicts within the traffic stream, or a combination of both.

Entrance ramps represent the third and most serious case since they create two potential conflicts with the maintenance of the adopted level of service of a roadway section. First, the additional ramp traffic may cause operational changes in the outside lane at the merge. This condition, of course, will be aggravated by any adverse geometrics, such as high angle of entry, steep grades, and poor sight distance. Second, the additional ramp volume may change the operating conditions across the entire roadway downstream from the on-ramp. This is particularly true where there is a downstream bottleneck.

The three basic procedures employed in checking capacity for the design of entrance ramps are based on the same philosophies discussed for metering ramps. One method is based on preventing the total freeway volume upstream from the ramp plus the entrance ramp volume from exceeding the capacity of a downstream bottleneck. A second method takes into consideration the distribution of freeway volumes per lane and then limits the ramp volume to the merging capacity less the upstream volume in the outside lane. The third method states that the ramp capacity is limited by the number of gaps in the shoulder lane which are greater than the critical gap for acceptance.

Figure 16 can be useful in the implementation of all three approaches. Thus, if a ramp on a new facility is of a high-type geometric design guaranteeing a low critical gap, methods 1 and 2 are applicable since the merging service volume will exceed any bottleneck service volume. However, due to the terrain, spacing of interchanges or ramp configuration, some compromise in the geometric design of the entrance ramp-freeway merging area is necessary, then the third method should be employed.
Freeway design is, as are most real world phenomena, a series of compromises. Because of the spacing of interchanges on many urban freeways the fulfillment of desirable entrance ramp design, desirable exit ramp design, and the provision for an adequate weaving section between them offer a dilemma. The alternatives are: (1) reduction in the standards of one or more of the features, (2) elimination of one of the features (such as one of the ramps) or (3) transferring the weaving from the freeway to the frontage road. These alternatives should be evaluated in terms of their cost, their effect on the freeway and ramp operation at that location, and their effect on adjacent facilities such as adjacent interchanges, cross street signalization, etc. The procedure exemplified by Figure 16 enables to some extent a designer to evaluate alternatives more rationally and if compromise is needed, to select the element or location where it will be the least objectionable.
SUMMARY AND CONCLUSIONS

Many researchers have defined a parameter called the critical gap in order to describe the operation of two interacting traffic streams. The critical gap, as a time gap that is just as likely to be accepted as it is to be rejected, is by implication a median gap. The advantages of this parameter for describing entrance ramp operation are: (1) it provides a single typical figure rather than the whole range of observed gap acceptance characteristics, (2) it gives simplicity in such advanced stages of ramp behavior as delay calculations, (3) it is sensitive to variations in geometrics and therefore provides a rational means of evaluating ramp design, and (4) like such parameters as capacity, optimum speed and optimum concentration, it is independent of traffic volumes and is therefore invariant for a given location under a given set of environmental and operating conditions. These concepts are verified in this report.

Of particular significance is the development of a distribution of critical gaps which enables a researcher to predict the probability of a driver having a certain critical gap. Delay values calculated using this distribution (Equation 19) will always be higher than those calculated assuming that all drivers have the same fixed critical gap (Equation 8). The fact that the delay obtained using the fixed value is the special case of the delay obtained using the distribution for a variance of zero proves that the mean of the critical gap distribution not the median should be used in delay computations.

The delay equations (8, 10, 11, 12) developed in this study are based on the assumption that the distribution of freeway headways is of the form of the Erlang frequency distribution. This represents a generalization of past work in which the distribution of freeway headways was assumed to be of that particular form of the Erlang distribution called the negative exponential distribution.

The distribution of time spent by merging ramp vehicles at the head of the queue may be approximated by a gamma distribution suggesting that an entrance ramp merging operation may be considered within the context of classical queueing theory. Based on this queueing model the ramp metering curves of Figure 16 were conceived. The metering technique suggested combines a macroscopic and a microscopic approach, the former based on a controlling downstream bottleneck, the latter on the merging area.

In conclusion, it is apparent that additional theoretical research is needed in order to provide a more complete relationship between the many variables associated with the interaction of vehicles traversing a ramp and merging into a freeway so as to determine the effects of traffic char-
acteristics, merging area geometrics, and environmental elements on merging capacity and operation. This information would be of profound significance in the operation and control of existing facilities, the design of future facilities, and the development of usable distributions of traffic variables for simulation programs.
APPENDIX

The Pearson Type III Distribution

A random variable $y$ is said to be distributed as the Type III distribution if its density is

$$f(y) = \frac{b^a y^{a-1} e^{-by}}{\gamma(a)} \quad 0 < y < \infty$$

(1)

where $\gamma(a)$ is called the gamma function and is defined by the formula

$$\gamma(a) = \int_0^\infty z^{a-1} e^{-z} \, dz$$

In order to apply the Type III distribution to traffic phenomena such as space headways in which the distribution curve does not go through the origin (the space headway between successive vehicles in the same lane can never be zero because vehicles possess length), it is necessary to translate the distribution $c$ units from the origin.

Recognizing that the area under $f(y)$ is unity, we obtain the desired generalized distribution by substituting $x = y + c$ for $y$ in (1);

$$\int_0^\infty f(y)dy = \int_0^\infty \frac{b^a}{c \gamma(a)} (x-c)^{a-1} e^{-b(x-c)} \, dx$$

giving

$$f(x) = \frac{b^a}{\gamma(a)} (x-c)^{a-1} e^{-b(x-c)}; \quad c < x < \infty$$

(2)

Taking moments about $c$,

$$\mu_k = \frac{b^a}{\gamma(a)} \int_c^\infty (x-c)^{k+a-1} e^{-b(x-c)} \, dx$$

$$= \frac{b^a}{\gamma(a)} \frac{\gamma(k+a)}{b^{k+a}} \int_c^\infty (x-c)^{k+a-1} e^{-b(x-c)} \, dx$$

$$= \frac{b^a}{\gamma(a)} \frac{\gamma(k+a)}{b^{k+a}}$$

(3)

since the integral represents the area under a Type III curve and must equal unity. The first moment about $c_1$ is obtained by letting $k = 1$ in (3),

44
\[
\mu_1^c = \frac{a}{b}
\]

(4)

Adding \( c \) to (4) gives the mean of this distribution

\[
\mu = \frac{a}{b} + c
\]

(5)

The variance is given by

\[
\sigma^2 = \mu_2^c - (\mu_1^c)^2 = \frac{a}{b^2}
\]

(6)
BIBLIOGRAPHY


BIBLIOGRAPHY continued


