Freeway Level of Service as Influenced by Volume and Capacity Characteristics

DONALD R. DREW and CHARLES J. KEESE,
Texas Transportation Institute, Texas A & M University, College Station

● AS APPLIED TO the traffic operation on a particular roadway, level of service refers to the quality of the driving conditions afforded a motorist by a particular facility. Factors which are involved in the level of service are (a) speed and travel time, (b) traffic interruption, (c) freedom to maneuver, (d) safety, (e) driving comfort and convenience, and (f) vehicular operational costs.

Each of the foregoing factors is somewhat related to all the others. The volume of traffic using a facility affects all of the factors and, in general, the greater the volume, the more adverse are the effects. As the ratio of the volume of traffic on a facility to the volume of traffic the facility can accommodate approaches unity, congestion increases. Congestion is a qualitative term, long used by the general public as well as traffic engineers, which refers to what can quantitatively be defined as vehicular density. The end result of an oversupply of vehicles is the formation of a queue of stopped (or "crawling") vehicles at bottleneck locations (a "breakdown" of the operation) such that volumes momentarily drop to zero, leaving only congestion on the facility until a clearout can be effected.

Traffic volumes are known to be continuously variable; even at very low hourly volumes there will be infrequent, short-term occasions when a relatively large number of vehicles will pass a given point. There also are regions on a facility which, due to the geometry, inherently will tend to accommodate fewer vehicles. This implies that bottlenecks do exist and thus the level of service on a given facility may vary even with a "constant" hourly volume. Bottlenecks may be fixed in space due to the aforementioned geometrical considerations of the facility and thus may be studied at the particular location. Such geometrical aspects as entrance and exit ramps have been studied and evaluated as bottlenecks. It is possible also, however, that the random "bunching" of vehicles at any point in space may result in "bottlenecking" due to the statistically variable nature of streams of vehicles, in which case, the designers should be able to predict such peaking characteristics in order to assure acceptable levels of service.

Basically, congestion will be the direct result of the nature of the "supply and demand" on a facility. The supply, in terms of traffic engineering, has been referred to as capacity; the demand placed on the facility is, as it implies, the number of motorists who would seek to use the facility, and can be estimated by origin-and-destination surveys if the times of the desired trips are obtained. It is often futile to measure the flow of traffic on an existing facility with the objective of determining the demand on that facility. About the only relationship between existing volumes and actual demand which can be determined from such measurements is whether the demand is, in fact, as large as the capacity during any significant length of time during a day. In borderline instances, peaking occurring within a peak hour might show where capacity is exceeded by demand for intervals of time less than one hour. Even this feature cannot be exhibited by a traffic system which is so inadequate that it limits (or "meters") the input of vehicles such that the volumes are less than the capacity of the particular facility being studied.

If it is possible, in a given system, for more vehicles to enter than the facility can handle, then congestion will result whenever the demand exceeds capacity and the accompanying inefficiency results in fewer numbers of vehicles being accommodated

by the facility in a given period of time. It is theoretically true that there is some maximum number of vehicles which can use a facility (1). This "possible capacity" is the volume of traffic during the peak rate of flow that cannot be exceeded without changing one or more of the conditions that prevail. From this value more restrictive conditions of roadway and traffic conditions are imposed to describe the measure of "level of service" that a given lane or roadway should provide. If the conditions are associated with highways or streets to be constructed at a future date, it is defined as design level of service. If the conditions express prevailing traffic flow conditions, it is designated as operational level of service.

Various volume levels can cause various levels of operating conditions, or levels of service. For any volume of vehicles using a particular facility, there is an associated level of service afforded these vehicles. It is possible that the input of vehicles into a particular facility will be regulated such that traffic volumes will not exceed a predetermined, suitable level of service volume. Such operational control procedures are being investigated and seem to offer considerable promise.

Although the use of a design level of service volume has considerable appeal in that it conforms to traditional engineering practice, the determination of such a volume, relative to various levels of service, is complex. There are regions on a freeway which are subject to more restrictive vehicular operation, such as in the vicinity of an entrance ramp or exit ramp. Such regions should be considered when determining the design volume of a facility and a knowledge of the operational characteristics and traffic requirements at such locations is necessary for proper planning in order to avoid future bottlenecks.

SCOPE

This report deals with two main topics which affect traffic operation, and thus the level of service, on freeways. They are volume characteristics and capacity characteristics. The capacity of a highway facility is a measure of its ability to accommodate vehicular traffic. This ability depends, not only on the physical features of the road itself, but also on the traffic demand and the interaction of vehicles in the traffic

should be apparent then, that an appreciation of freeway volume characteristics is important in the planning, design and operation of freeway facilities.

Freeway volume characteristics utilize some of the same parameters defined in the more general subject of volume characteristics. First, the maximum observed volumes on different types of highway facilities, not only provide an indication of the magnitude of traffic demand, but also establish a lower limit of possible highway capacity. The variations in volume for different time periods are explored for their effect in the selection of representative design hourly volumes. These cyclic patterns include monthly, daily, hourly and peak period variations. The distribution of vehicles by direction and lane and the composition of traffic are important design considerations, whereas the longitudinal distribution of vehicles has considerable practical significance as our emphasis gradually shifts from freeway design to freeway operations and control.

Two main topics under volume characteristics are considered and developed. The first is an analysis of such peak flow characteristics as peak rates of flow, correlation with various parameters, and a comparison of peak 2-hr volumes with peak 1-hr volumes. The second main topic under volume characteristics deals with studies of lane distribution of vehicles on freeways with particular attention given to the effect of entering and exiting vehicles on the lane use distributions.

The second section deals with freeway capacity characteristics, including a theoretical approach to providing a rational relationship between capacity and level of service. It is significant that the first section deals with freeway demand and the second with capacity. In the third section, applications are made of these demand and capacity characteristics to freeway design and operations.

VOLUME CHARACTERISTICS

Peaking Characteristics

The need to consider the peak rates of flow within the peak hour has been recognized for several years (2, 3, 4, 5) and has been given considerable attention by the HRB Committee on Highway Capacity.

No matter what criteria are used for the design and operation of a freeway, it is necessary to know what the traffic demand will be. Although no definite limits have been established as yet for the factors affecting the level of service, there will have to be some correlation between peak volume and level of service. An origin-destination survey which denoted, to the nearest five minutes, when motorists would desire to begin their trips, could yield valuable information about the true nature of the existing demand on an urban transportation system. The peak demand periods, however, would likely exceed the economical limits of any system which could be provided. In some instances, depending on the data available, reasonable estimates can be made of future peak-hour volumes. A peak-hour volume does not, however, necessarily imply that a high rate of flow will exist for less than a full hour, more than an hour, or approximately one hour; it is simply an estimate of the maximum number of vehicles expected on a facility during a full 60-min period. Due to the nature of the peak-hour demand and the statistically variable nature of traffic, it is known that short-term rates of flow within the peak hour are often quite variable.

The statistical variability of volumes of traffic is affected by the time period involved. As the time period is reduced, the average number of vehicles for that time period will reduce accordingly. For example, if the average hourly volume were 1,800 vph, the average minute volume would be 30 vpm and the average second volume would be $\frac{1}{2}$ vps based on the hourly volume. The variability of smaller mean values is greater than that of larger mean values when expressed as a percentage of the mean. The narrowing of the confidence interval band for increasing mean values is characteristic of not only the Poisson distribution, which closely approximates the distribution of light volumes of traffic, but also the many other distributions which have been used to approximate various actual traffic distributions. Even the normal distribution exhibits these same characteristics although it is rarely used as an example of existing traffic distributions. Thus, even without the occurrence of a change of volume within a given peak hour, a short-term period within this hour has increased probability of exceeding its mean by a given percent than does the whole hour.

For planning purposes, future volumes are presently estimated for the peak hour or 2-hr period. To relate such volumes into a design peak rate of flow, the factors which affect this relationship must be established and evaluated.

Data from more than 200 freeway traffic studies were obtained from the Texas Highway Department, Bureau of Public Roads (Ramp Capacity Studies), previous studies conducted by the Texas Transportation Institute and specific studies conducted on this project. The relationship between short-period traffic flow (5-min flow) to total hourly flow was determined from these data.

Because the "loading" of freeway in some instances is controlled by the capacity and operation of the supporting street system and inadequate capacity also sometimes limits "unloading" which results in impaired freeway operation, there was not sufficient knowledge of each of the freeways, except those in Texas and a few other specific sites, to permit consideration of these characteristics. It is possible that much better correlations of the results would have been possible had all conditions been known. Those freeways known to have good loading and unloading characteristics showed very good correlation of the data.

Although many characteristics related to trip generation such as geographical and time concentrations of trips, character of the freeway (radial, circumferential, etc.), character of supporting street system, population, area served, and others perhaps have marked effects on the peaking characteristics, it was possible from the data available to study only the relationship of peaking to the population of the city or urban area. The results are shown in Figure 1. These curves are based on the data for 132 peak periods from studies in 31 cities in 18 states. Congestion was not apparent in the
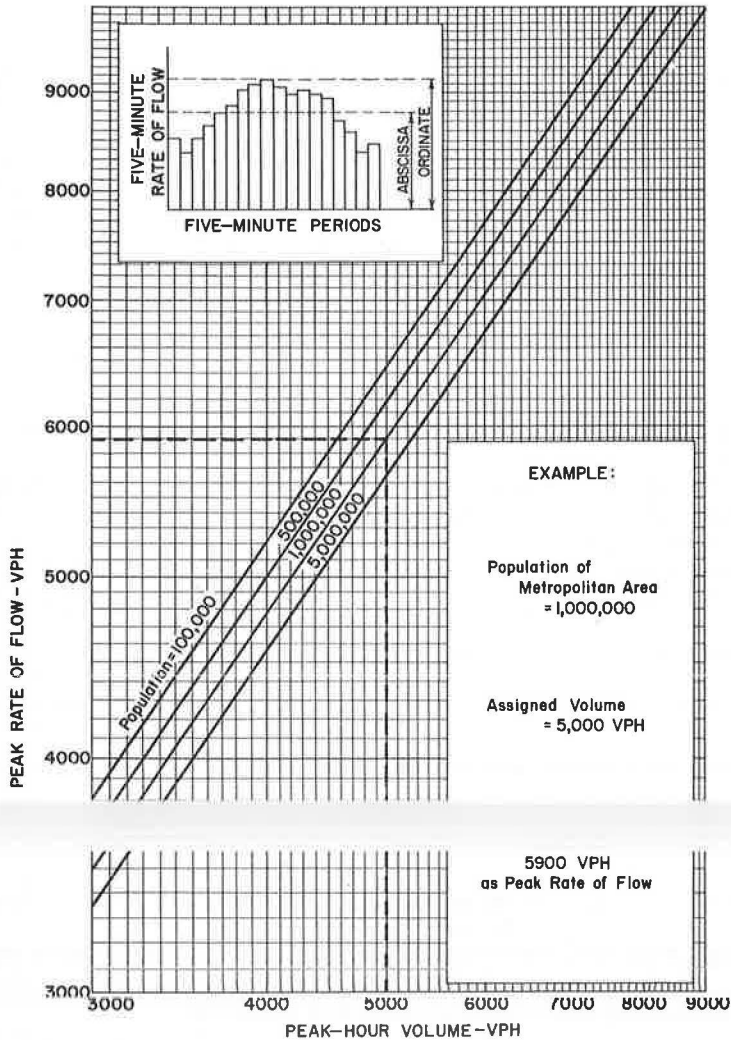
Figure 1.   Determination of rate of flow for highest 5-min interval from rate of flow for whole peak hour.

immediate vicinity of any of the study sites. The variables are statistically significant and the curves fit the available data with a standard deviation of 5 percent.

Figure 2 shows the relationship between estimated rates of flow and observed rates of flow and includes a 10 percent error band within which most points were included. Figure 3 shows the frequency distribution of the percent error involved in using Figure 1 to estimate the peak rates of flow. As can be noted, the errors are somewhat normally distributed.

A series of multiple regressions were run on available data in an effort to determine the factors affecting the maximum rate of flow occuring in a peak hour. Consideration was given to the following items: the physical size of the metropolitan area, the concentration of the central business district, the distance of the study site from the main destination or generator (in general the central business district), the population of the metropolitan area as measure of the complexity of the traffic system, and the actual size of the peak period itself, and whether the peak occurred in the morning or afternoon. A few of the cities with multiple study sites indicated a definite relationship with
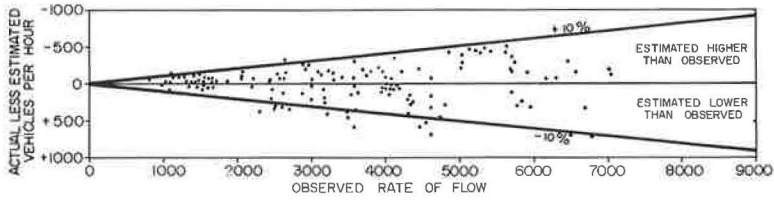
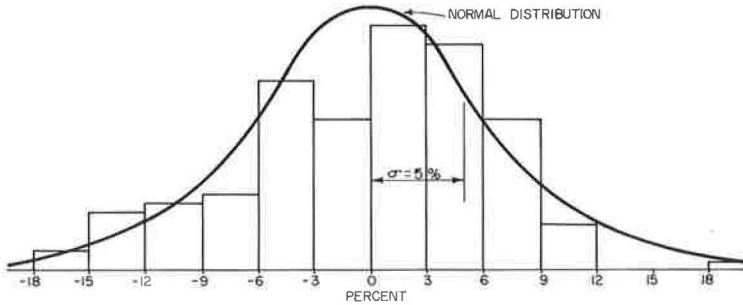Figure 2. Relationship between error, in vph, and observed rate of flow.

Figure 3. Frequency distribution of percent difference between estimated and observed rates of flow.

the distance from the major generator, while others did not. It was disconcerting to note that this relationship was often contradictory between studies. In the light of such varying results, no relationship could be positively identified. The same was true of most of the other factors.

Improved traffic assignment methods, involving comprehensive programs utilizing large digital computers, are being used to develop predictions of urban traffic volumes for a peak 2-hr period. To establish a relationship between 2-hr peak periods and 1-hr peak periods, the data from the studies mentioned earlier were analyzed. Figure 4 shows the results obtained from 95 studies for which 2-hr peak volumes were available. It can be noted that the peak-hour volume can be expected to lie between 55 and 60 percent of the peak 2-hr volume.

By using the relationships shown in Figures 4 and 1, the design volume can be obtained from the peak 2-hr volume. These design volumes will take into account the peaking effect.

Lane Distribution Characteristics

Critical sections on the freeway often exist adjacent to ramps and, if a certain level of service is to be assured the motorists, it is necessary to give close consideration to such areas in the design of freeways. Because the merging problem directly involves traffic in the outside lane and the entering ramp traffic, a study was made of the percent of total freeway traffic using the outside lane. Only six-lane freeways were considered in this particular project.

A three-part research project was made involving data obtained from 49 study sites located on 14 different six-lane freeways in 10 different states. First, empirical relationships were developed which best fit the data of eight representative study sites. Parameters which were found to be significant were freeway volume, entrance ramp volume, upstream ramp volume, distance to upstream ramp, downstream exit ramp volume, and distance to downstream exit ramp. Second, the relationships which were developed were then used to test against all study sites for validity. The third step was the development of a design procedure which was used for practical applications.
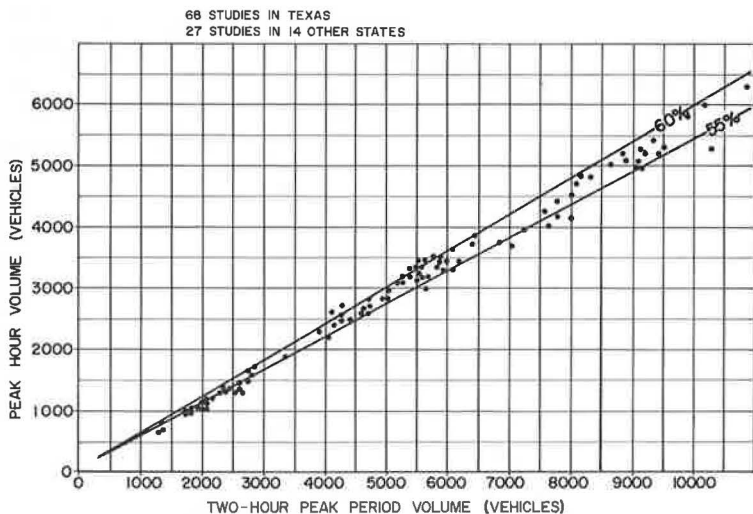
Figure 4.   Determination of rate of flow for peak hour from rate of flow for a 2-hr period.
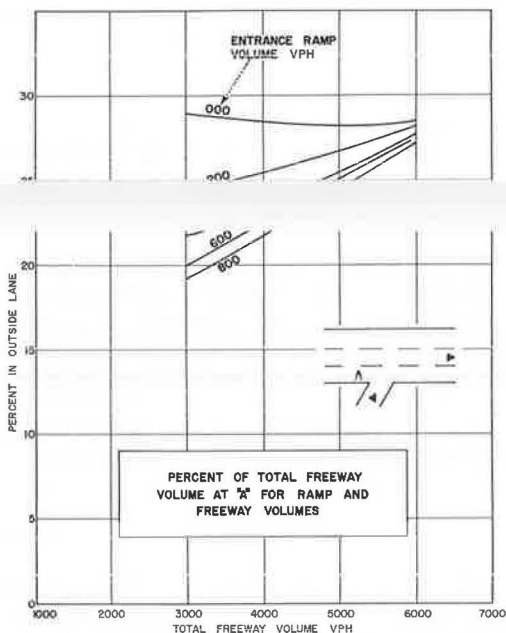

Figure 5.   Relationship between percent of total freeway volume in outside lane and entrance ramp volume.

Figure 5 shows the relationship between the percent of the total freeway volume in the outside lane and two of the parameters—total freeway volume and entrance ramp volume. The monographs shown in Figures 6, 7 and 8 represent the relationship between the percent of the total traffic and distance to, an upstream exit ramp, a downstream exit ramp, and an upstream entrance ramp, respectively. Distances are referenced to the ramp nose in each case. A downstream entrance ramp was considered to have no effect on the percent of traffic in the outside lane.

The predicted percent in the outside lane can have no less variability than the ordinary or natural variability of the data in general. The 90 percent confidence limits for each of the freeway and ramp volume groups were calculated as follows:

Freeway Volume (vph)	Natural Variability in Outside Lane (%)
2,000-3,000	±8
3,000-4,000	±5
4,000-5,000	±3
5,000-6,000	±2

The "natural" variability of the data was arbitrarily defined as the range within the 90 percent confidence interval. The large natural variability at freeway volumes less
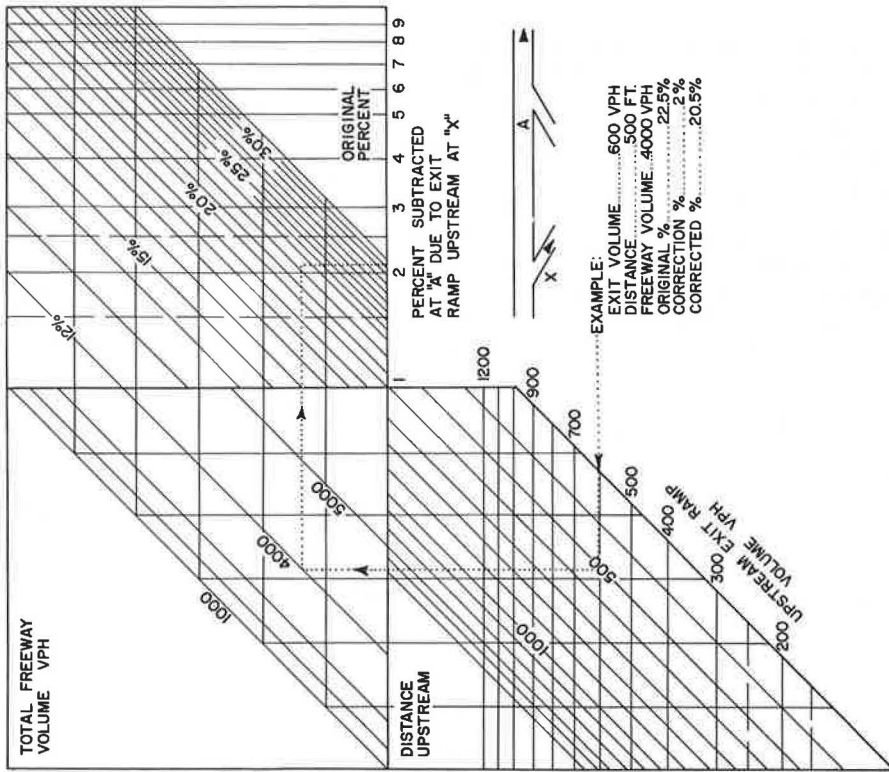
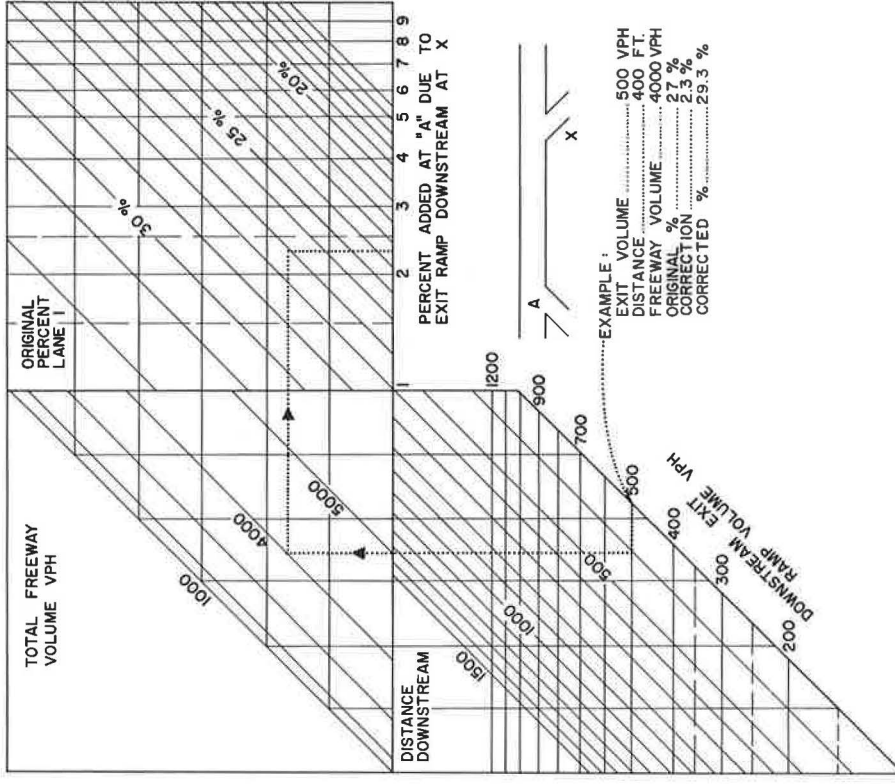Figure 7. Correction to percent in lane 1 due to exit ramp downstream.

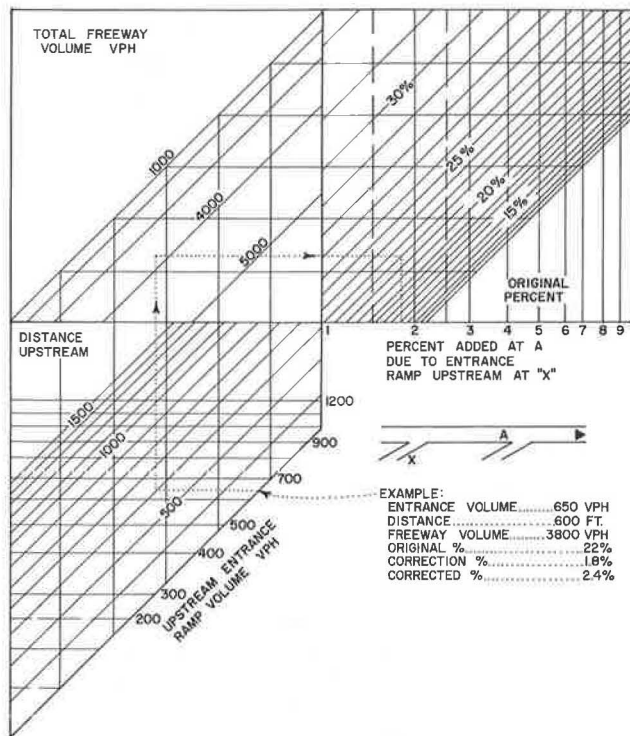Figure 6. Correction to percent in lane 1 due to exit ramp upstream.

Figure 8. Correction to percent in lane 1 due to entrance ramp
upstream.

the predicted percentages in the outside lane was made for twelve 5-min volumes taken
from the 49 study sites (Table 1).

It is intended that the method presented here for six-lane freeways be extended to
include four- and eight-lane freeways. In an effort to determine general relationships
for four-, six- and eight-lane freeways similar to those shown for only six lanes in
Figure 8, data from 132 peak periods were analyzed. These data were the same as
those used in the peaking characteristics study and included more than 2,000 five-
minute volumes from 31 cities located in 18 states. Multiple regression techniques
were used to develop equations from these data and Figure 9 shows the families of
curves which were developed from the equations. A multiple correlation coefficient
squared, R^2, is generally understood to represent that portion of the variability, or
variation, which is accounted for by the derived equations. The obtained R^2s are as
follows: four-lane freeways, 0.51; six-lane freeways, 0.70; and eight-lane freeways,
0.48. It is interesting to note that, for the entire range of values shown in Figure 9,
the standard deviation was approximately 200 vph. Expressed as a percentage, how-
ever, at lower outside lane volumes of 1,000 vph, there would be a standard deviation
of 20 percent.

It is known that factors other than freeway volume and ramp volume affect the per-
cent of total freeway traffic in the outside lane. Further studies of those factors are
being made so that nomographs similar to those shown in Figures 6 to 8 can be pre-
pared.

As previously mentioned, there is an important relationship between trip length and
outside lane utilization. Drivers entering the freeway for only relatively short trips
could reasonably be expected to remain in the outside lane. In an attempt to evaluate
such relationships, a single study was conducted on the North Central Expressway in
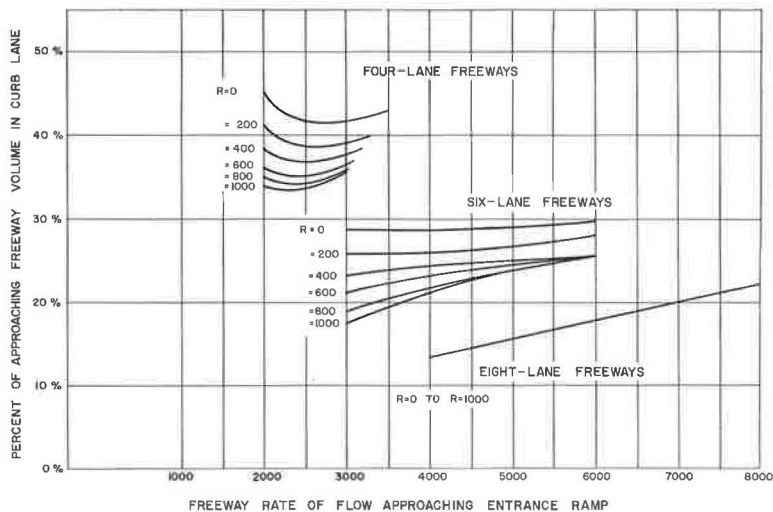
Figure 9. Relationship between percent of total freeway volume in outside lane and freeway and ramp volumes (R = ramp volume).
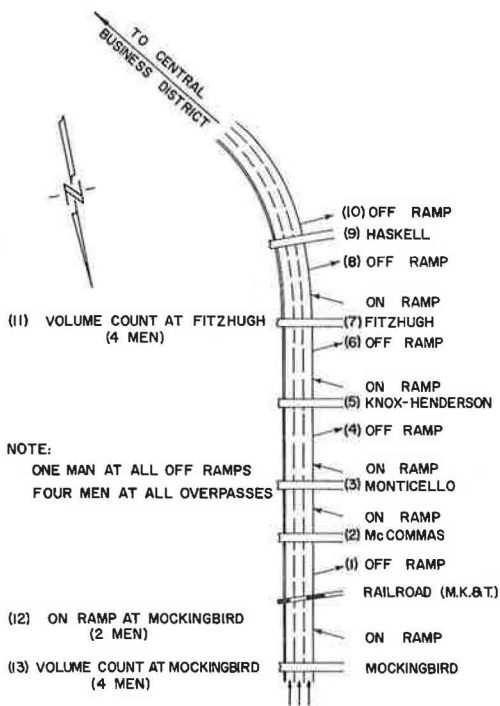

Figure 10. Strip map for observation locations.

Dallas using the "lights-on" study technique previously used by the New York State Department of Public Works and the Port of New York Authority (6).

Figure 10 shows a scheme of entrances, exits, and observation points along the 2½ miles of the six-lane freeway study section. Vehicles entering at the Mockingbird entrance ramp were advised by signs to drive with their lights on for 20 minutes. Policemen were standing near the signs and would call motorists' attention to the message by simply pointing to the signs. Observers were located at strategic points along the freeway, generally on the cross-street overpass structures and at each off-ramp.

As the vehicle entered the freeway, the last four digits of the license number and the vehicle type were recorded. Failure to turn on headlights was noted to determine compliance ratios.

As the vehicle with headlights on moved through the study area, the observer at each point would record the license number, vehicle type and the lane in which the vehicle was traveling. Observers located at each exit ramp recorded the license number of each vehicle with headlights on. Each observer noted the end of each 5-min period on the data sheet.

Freeway volumes were counted, by lanes, at two locations during the entire study period of four hours. One volume count station was located at the beginning of the study section and one at the end of the section. Volumes were recorded in 5-min intervals.

TABLE 1

TABULATION OF THE NUMBER AN... [GNITUDE OF THE TIMES THE DIFFERENCE BETWEEN]
THE OBSERVED PERCENTA... [ESS THE PREDICTED PERCENTAGE EXCEEDED]
THE ... [RAL DATA VARIABILITY]

NUMBER	STATE	FREEWAY	NUMBE 5 Min. PERIOD	...UMBER OF TIMES THE NATURAL DATA VARIABILITY IS EXCEEDED BY:								
				1%	1-2%	2-3%	3-4%	4-5%	5-6%	6-7%	7-8%	8-10%	10%
05	Texas	Gulf	24		0	0	0	0	0	0	0	0	0
07	Texas	Dallas	32		2	2	1	0	0	0	0	0	0
12	Texas	Gulf	32		0	0	0	0	0	0	0	0	0
13	Texas	Gulf	32		1	0	0	1	0	0	0	0	0
31	Texas	Gulf	30		0	0	0	0	0	0	0	0	0
33	Texas	Gulf	30		1	2	0	0	0	0	0	0	0
41	Texas	Gulf	36		2	0	0	0	0	0	0	0	0
44	Texas	Gulf	36		0	0	1	0	0	0	0	0	0
0501	Calif.	San Bernardino	35		3	2	1	1	1	2	1	0	0
0503*	Calif.	Harbor	24		2	1	1	3	2	0	0	0	0
0505*	Calif.	Hollywood	30		3	5	5	3	3	1	2	1	1
0609(1)	Colo.	Valley Highway	24		1	0	0	0	0	0	0	0	0
0610*	Colo.	Valley Highway	24		0	2	4	3	2	2	3	1	3
1101	Ga.	Atlanta Expressway	24		0	1	0	1	1	0	0	0	0
1102	Ga.	Atlanta Expressway	24		0	0	0	1	0	0	0	0	0
1103	Ga.	Atlanta Expressway	18		1	0	0	0	0	0	0	0	0
1104	Ga.	Atlanta Expressway	24		4	0	0	1	0	0	0	0	0
1105	Ga.	Atlanta Expressway	24		2	1	1	0	2	0	0	1	0
1106	Ga.	Atlanta Expressway	18		3	2	0	0	1	0	0	1	0
1107	Ga.	Atlanta Expressway	24		3	2	2	1	0	0	0	0	0
1108	Ga.	Atlanta Expressway	26		1	1	0	0	1	1	2	0	0
1109	Ga.	Atlanta Expressway	18		2	1	2	1	0	0	0	0	0
1430*	Ill.	Congress Street	24		5	4	4	2	3	0	0	1	0
1503**	Ind.	Tri-State	30		0	0	0	0	0	0	0	0	0
2313	Mich.	Edsel Ford	28		3	1	0	1	0	0	0	0	0

ID	State	Freeway		0%	1%	2%	3%	4%	5%	6%	7%	8%	10%
2317(1)	Mich.	Edsel Ford	35	2	3	2	4	1	2	0	0	0	0
2318(1)	Mich.	Edsel Ford	19	2	3	0	0	1	0	1	0	0	0
2321	Mich.	Edsel Ford	14	1	0	1	1	0	0	0	0	0	0
2322	Mich.	Edsel Ford	15	0	2	3	1	0	0	0	0	0	0
2323(1)	Mich.	Edsel Ford	26	1	0	0	0	0	0	0	0	0	0
2503(1)	Mo.	Intercity	24	0	2	0	0	0	0	0	0	0	0
3101	N.J.	Garden State	30	2	6	1	1	1	0	0	0	0	0
3315	N.Y.	N.Y. State	24	0	0	0	0	0	0	0	0	0	1
222	Texas	Eastex	26	2	1	0	0	0	0	0	0	0	0
131	Texas	Eastex	18	3	2	2	1	0	0	0	0	0	0
221	Texas	Eastex	25	0	4	1	1	0	0	0	0	0	0
130	Texas	Eastex	24	1	1	2	0	0	0	0	1	0	0
3010	Texas	Gulf Freeway	16	0	2	1	0	0	0	0	0	0	0
4020	Texas	Gulf Freeway	15	2	0	2	0	0	0	0	0	0	0
5030	Texas	Gulf Freeway	13	2	2	4	0	0	0	0	0	0	0
5930	Texas	Gulf Freeway	10	0	2	1	1	0	0	0	0	0	0
6050	Texas	Gulf Freeway	16	0	1	1	1	0	0	0	0	0	0
7060	Texas	Central Expressway	13	2	1	1	0	0	0	0	0	0	0
4413*	Texas	Central Expressway	36	2	4	4	0	2	0	0	0	0	0
4414*	Texas	Central Expr.	36	3	1	3	4	1	1	1	2	8	1
R12(1)	Texas	Gulf Freeway	36	3	2	2	1	0	1	0	0	0	0
R13(1)	Texas	Gulf Freeway	36	2	2	1	0	0	0	0	0	0	0
GRAND SUM			1212	83	86	61	38	26	20	8	11	13	6
				0%	1%	2%	3%	4%	5%	6%	7%	8%	10%
				29.1%	22.2%	15.1%	10.1%	6.9%	4.8%	3.1%	2.5%	1.6%	0.5%

Percent of the time that the difference exceeds the Natural Data Variability by a percent more than:

* Unusual Geometrics (lane dropped or added)
** Freeway Volume is less than 2000 vph
(1) Study with complete adjacent ramp data
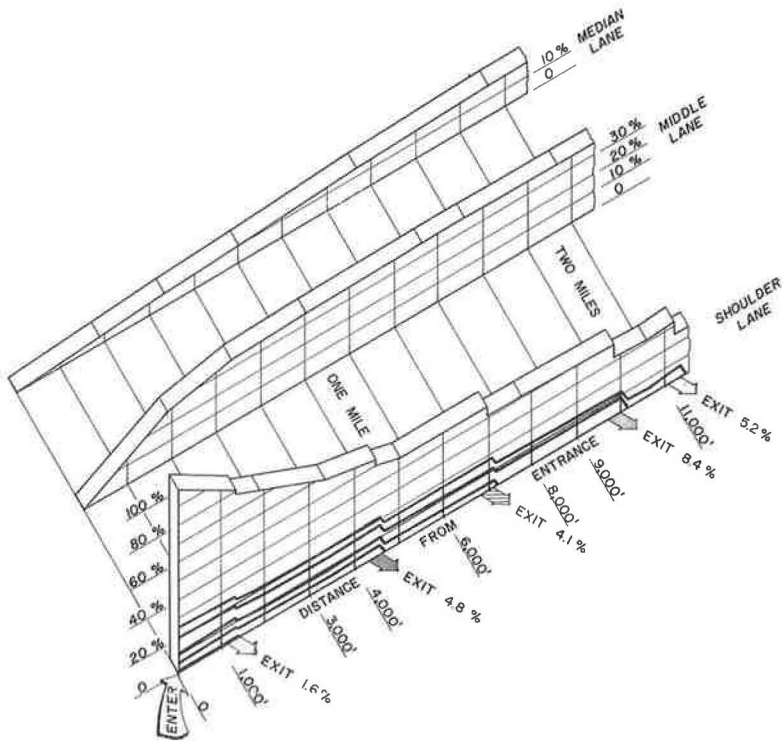
Figure 11. Lane use distribution by vehicles entering at Mockingbird ramp, distribution by percent in each lane, 1,444 vehicles in sample.

During the 4-hr study period more than 1,800 vehicles entered on the Mockingbird ramp and more than 1,500 complied by turning on their headlights. About 1,100 of the complying vehicles did not exit within the study area and about 400 vehicles were noted to exit at one of the five exit points within the study section.

Regarding the study method, the following can be summarized:

1. The "lights-on" study method was an effective means of studying the lane use related to trip length on this particular freeway which had numerous major street overpasses.

2. The presence of the signs, policemen, or observers had no noticeable effect on the traffic flow.

3. Compliance ratios in the 75-85 percent range were obtained by using this method. The sample size was much higher and less expensive to obtain than by other methods considered.

4. Proper design and location of signs is a big factor in obtaining high compliance ratios.

5. A loudspeaker was not essential.

6. A policeman can be used to advantage to encourage motorists to observe the signs.

7. Newspaper, radio, and television publicity can be beneficial, but is not necessary for adequate compliance ratios.

Figure 11 depicts three-dimensionally how the vehicles entering at the Mockingbird ramp were distributed over the three lanes as they traveled toward the central business district. The dark portion at the bottom of the outside lane represents that portion of those entering vehicles which exited within the study area. A large percentage of all such exiting vehicles used only the outside lane and, as a result, those vehicles cannot be noticed in the two inside lanes.
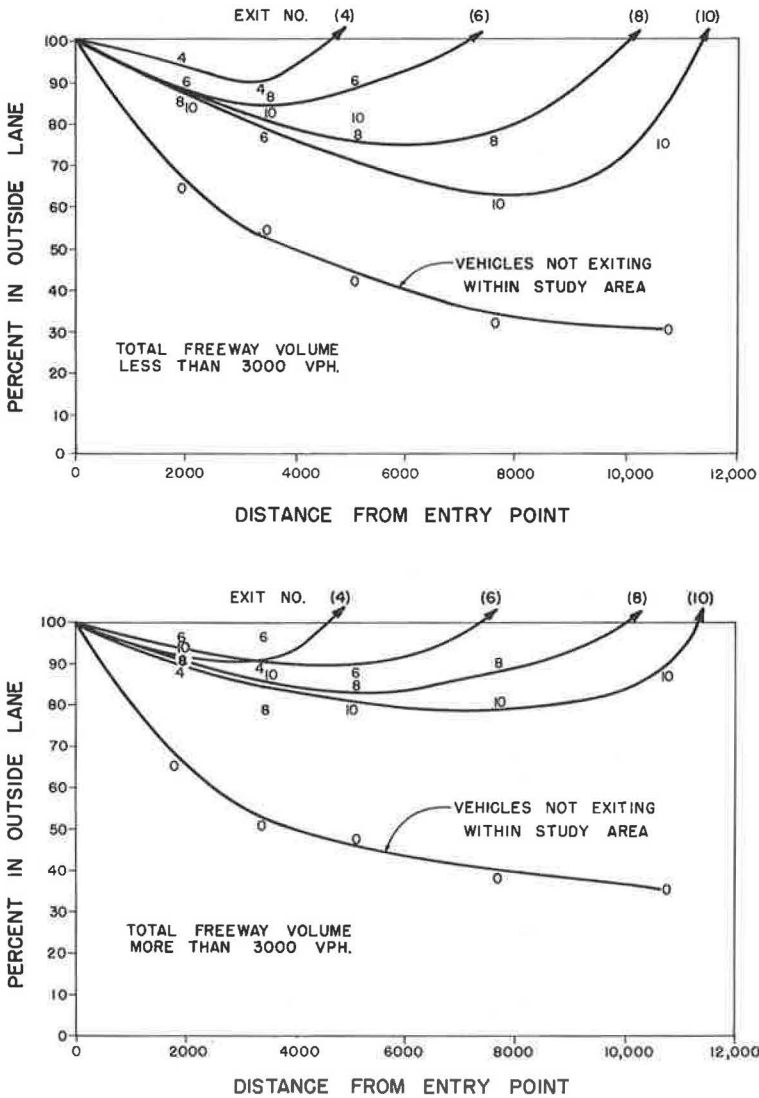
Figure 12. Outside lane use relationship with trip lengths.

The total volume of traffic on the freeway has an effect on the lane usage of entering traffic. As the total volume increases, entering vehicles are more restricted to the outside lane and motorists view a temporary lane change less worthwhile in view of the fact that another lane-change opportunity must be found to return to the outside lane prior to exiting. Figure 12 shows the relationship between trip length and the percent of the entering traffic in the outside lane. The upper portion of Figure 12 pertains to only light total freeway traffic volumes of less than 3,000 vehicles per hour, one way. The lower portion of Figure 12 corresponds to conditions of moderate to heavy total freeway traffic volumes of more than 3,000 vehicles per hour, one way. Although the observation points are widely spaced, it would seem that the restrictive effect of the higher volumes is noticeable.

Figure 12 indicates that "ordinary" or "equitable" lane use distribution is only attained by those vehicles traveling several miles on the freeway. Those vehicles traveling less than three miles cannot be expected to reach that "steady-state" lane distribution which is characteristic of "through" vehicles.

It must be pointed out that this is the result of a single study. It is expected that this study will provide a basis for further work of a similar nature.

CAPACITY CHARACTERISTICS

<u>Theoretical Approach to the Capacity-Level of Service Concept</u>

Both capacity and level of service are functions of the physical features of the highway facility and the interaction of vehicles in the traffic stream. The distinction is this: A given lane or roadway may provide a wide range of levels of service, but only one possible capacity. The various levels for any specific roadway are a function of the volume and composition of traffic. A given lane or roadway designed for a given level of service as a specified volume will operate at many different levels of service as the flow varies during an hour, and as the volume varies during different hours of the day, days of the week, periods of the year, and during different years with traffic growth. In other words, fluctuations in demand do not cause fluctuations in capacity, but do effect changes in the quality of operation afforded the motorist. In a very general way then highway planning, design and operational problems become a case of whether a certain roadway (capacity) can handle the projected or measured demand (volume) at an acceptable level of service (speed, etc.). Because of both observed and theoretical speed-volume relationships on freeway facilities, which are considered later, it is possible to anticipate to some degree just what level of service can be expected for a given demand-capacity ratio. The obvious weakness lies in the fact that most of the qualitative factors affecting level of service cannot be related directly to traffic volume.

Greater dependency on motor vehicle transportation has brought about a need for greater efficiency in traffic facilities. The motorist is no longer satisfied to be "out of the mud." In fact, fewer and fewer folks remember the days of unpaved roads. The freeway is an outgrowth of the demand for highways providing higher levels of service. The place that motor vehicle transportation plays in our society demands dependable service be provided by traffic facilities and the popularity or attraction to the freeway illustrates this point. It is very important that the engineer clearly understands the

terms of volume accommodated. He evaluates efficiency in terms of his trip—the service to him. He evaluates the operating conditions of speed, travel time, traffic interruptions, freedom to maneuver, safety, driving comfort and convenience, operating costs, etc. The level of service is a term which denotes the different operating conditions that occur on a given lane or roadway when accommodating various traffic volumes.

Recent contributions (7) to traffic flow theory regard traffic as a one-dimensional compressible fluid with a concentration, k, and a fluid velocity, u. The conservation of vehicles is explained by the following equation of continuity:

$$\frac{\partial k}{\partial t} + \frac{\partial (ku)}{\partial x} = 0 \tag{1}$$

If it is assumed that drivers adjust their speed in accordance with the traffic conditions about them as expressed by the general expression $k^n \partial k/\partial x$, the acceleration of the traffic stream becomes

$$\frac{du}{dt} = - c^2 k^n \frac{\partial k}{\partial x} \tag{2}$$

Solving Eqs. 1 and 2 for $u = f(k)$, and making use of $q = ku$ yields the following generalized equation of state for a traffic stream

$$q = k u_f \left[1 - \left(\frac{k}{k_j} \right)^{(n+1)/2} \right], \quad n > -1 \tag{3}$$

where u_f is the free speed and k_j is the jam concentration. The exponent, n, provides some flexibility in fitting a theoretical flow concentration curve to a particular highway (8).

The speed of waves carrying continuous changes of flow through the stream of vehicles is given by the derivative of the q - k equation in Eq. 3

$$q' = u_f \left[1 - \frac{(n+3)}{2} \left(\frac{k}{k_j} \right)^{(n+1)/2} \right], \quad n > -1 \tag{4}$$

The concentration, k_m, at which flow is a maximum is obtained by setting Eq. 4 equal to zero and solving for k:

$$k_m = \left[(n+3)/2 \right]^{-2/(n+1)} k_j, \quad n > -1 \tag{5}$$

Repeating for $dq/du = 0$ gives the speed, u_m, at which flow is a maximum

$$u_m = \left[(n+1)/(n+3) \right] u_f, \quad n > -1 \tag{6}$$

It therefore follows that the maximum traffic flow obtainable on a roadbed (capacity) is

$$q_m = k_m u_m \tag{7}$$

It has been hypothesized (9) that discontinuities in traffic flow are propagated in a manner similar to "shock waves" in the theory of compressible fluids. The speed of a shock wave, U, is given by the slope of the chord joining the two points of the flow-concentration curve which represent the conditions ahead of and behind the shock wave

$$U = (q_2 - q_1)/(k_2 - k_1) \tag{8}$$

Application of the mean value theorem suggests that the speed of the shock wave is approximately the mean of the speeds of the waves running into it from either side

$$U = 1/2 (q'_1 + q'_2) \tag{9}$$

where q'_1 is given in Eq. 4.

The very strong analogy between traffic flow and fluid flow suggests that the conditions of continuity of momentum and energy should be fulfilled at the surface of a traffic shock wave, just as the equations of dynamic compatibility must be fulfilled in fluid dynamics. Multiplying Eq. 1 by u and Eq. 2 by k, then adding the two equations, we obtain

$$\frac{\partial (ku)}{\partial t} = - \frac{\partial \left(ku^2 + k^{n+2} \frac{c^2}{n+2} \right)}{\partial x} \tag{10}$$

Eq. 10 is the law of conservation of momentum in the differential form as applied to traffic flow. Comparing Eqs. 1 and 10 with the classical forms in hydrodynamics, we can complete the analogy between the fluid and traffic quantities. This correspondence is illustrated in Table 2.

Kinetic energy, ku^2, is the energy of motion of the traffic stream. The measure of the jerkiness of the driving in this stream is given by the standard deviation of the ac-

TABLE 2

CORRESPONDENCE BETWEEN PHYSICAL SYSTEMS

Factor	Hydrodynamic System	Traffic System
Variables	Mass density, p Velocity, v	Concentration, k Speed, u
Parameters	Momentum, pv Shock wave velocity, U Kinetic energy, pv 2/2 Internal energy, ϵ	Flow, ku Shock wave velocity, U Kinetic energy, ku 2/2 Acceleration noise, σ

celeration or acceleration noise, σ. The units of both parameters are those of acceleration. Energy, as expressed in these two quantities, is consistent with the level of service concept previously defined. Thus, the kinetic energy of the stream fulfills the first level of service factor (speed and travel time), whereas internal energy (acceleration noise) measures such level of service factors as traffic interruption and freedom to maneuver, and to some degree safety, comfort and operation costs.

Utilizing energy, rather than momentum, as the criteria for optimization depends on finding those values of k and u that maximize the kinetic energy of the traffic stream, E, and minimize the internal energy or lost energy, σ. Division of Eq. 3 by k, squaring the term and then multiplying by k yields

$$E = ku_f^2 \left[1 - 2 \left(\frac{k}{k_j} \right)^{(n+1)/2} + \left(\frac{k}{k_j} \right)^{(n+1)} \right], \quad n > -1 \tag{11}$$

$$k'_m = (n+2)^{-2/(n+1)} k_j, \quad n > -1 \tag{12}$$

$$u'_m = \left[(n+1)/(n+2) \right] u_f, \quad n > -1 \tag{13}$$

and

$$q'_m = k'_m u'_m \tag{14}$$

Dividing Eqs. 5, 6 and 7 by Eqs. 12, 13, and 14, respectively, we get

$$\frac{k_m}{k'_m} = \left[\frac{2(n+2)}{(n+3)} \right]^{2/(n+1)}, \quad n > -1 \tag{15}$$

$$\frac{u_m}{u'_m} = \frac{n+2}{n+3}, \quad n > -1 \tag{16}$$

and

$$\frac{q_m}{q'_m} = 2^{2/(n+1)} \frac{(n+2)^{(n+3)/(n+1)}}{(n+3)}, \quad n > -1 \tag{17}$$

It is apparent that $k'_m < k_m$ and $u'_m > u_m$ which, from the point of view of the motorist, suggests that energy is a better criteria for defining optimum operation. Of course, this is accomplished by a sacrifice in traffic flow, since $q'_m < q_m$.

Traffic engineers have long been faced with the dilemma of relating possible capacity to level of service quantitatively. Much of the difficulty can be attributed to the fact that capacity is expressed in the units of volume, whereas the term level of service is highly subjective in nature. Volume is a logical measure of efficiency from the point of view of the engineer, whereas motion in the form of speed and the magnitude and frequency of speed changes is an important measure of level of service from the point of view of the individual driver.

The momentum-kinetic energy analogy seems to apply. If a moving mass strikes a stationary object without rebounding, as when the descending block of a pile driver strikes the pile, the resulting motion of the latter depends on the momentum of the block and not upon its kinetic energy. Therefore, if a fixed amount of energy is available, it is more effective to use a heavy mass moving at a relatively low speed than a lighter mass moving at a high speed. The momentum of the slow-moving heavy mass after falling a short distance is greater than that of a small rapidly moving mass which has been lifted higher by the expenditure of the same amount of energy.

Replacing mass in the previous discussion with traffic density, it is seen that efforts to measure efficiency in terms of momentum (traffic "throughput") must necessarily be achieved with a high traffic stream density and a low traffic stream speed. This is obviously not consistent with the level of service concept. However, because energy is a scalar quantity, the energy of a system such as a traffic stream is equal to the sum of the energies of its constituent particles and will be a maximum when internal friction caused by vehicular interaction is a minimum. Inasmuch as this internal friction reflects some compromise on the individual driver's freedom to maneuver, his comfort, and his safety, the energy concept includes most of the qualitative ingredients defined in level of service. Moreover, the energy concept affords the engineer the opportunity to treat level of service quantitatively.

If Eqs. 3 and 11 are expressed in terms of speed only, and then normalized, they become (for the special case of $n = 1$)

$$\frac{q}{q_m} = 4\left[\left(\frac{u}{u_f}\right) - \left(\frac{u}{u_f}\right)^2\right] \tag{18}$$

and

$$\frac{qu}{q_m u_f} = 4\left[\left(\frac{u}{u_f}\right)^2 - \left(\frac{u}{u_f}\right)^3\right] \tag{19}$$

The curves of Eqs. 18 and 19 are plotted in Figure 13. The right side of the graph is the well known volume-speed relationship normalized so that the abscissa is the ratio of flow to capacity and the ordinate the ratio of speed to free speed. Because division of the abscissa by the ordinate

$$\frac{q}{q_m} \div \frac{u}{u_f} = \frac{ku}{4k_j u_f} \div \frac{u}{u_f} = \frac{k}{4k_j} \tag{20}$$

it is apparent that the slope of any ray is one-fourth the normalized traffic density. The optimum density rays, using both the momentum and energy criteria, are plotted.

The left side of the graph shows the relationship between the kinetic energy and speed of the traffic stream. Division of the abscissa by the ordinate

$$\frac{qu}{q_m u_f} \div \frac{u}{u_f} = \frac{q}{q_m} \tag{21}$$

gives the flow-capacity ratio. Thus, the maximum abscissa gives maximum kinetic energy (located at $u'_m = 2/3 \, u_f$), while the maximum slope gives the maximum momentum or flow (located at $u_m = 1/2 \, u_f$).

The relationship between capacity and level of service is so fundamental to such practical aspects of traffic engineering as planning, design and operations, it is important that the distinction between these terms be appreciated. This can best be accomplished by a quantitative relationship, based on the energy-momentum analogy, as expressed in the following definitions:

Possible capacity is the maximum number of vehicles that can be handled by a particular roadway component under prevailing conditions. It is that product of the density and speed that maximizes the momentum of the traffic stream.

Level of Service refers to the quality of driving conditions afforded a motorist by a particular facility as reflected by (a) speed and travel time, (b) traffic interruption, (c) freedom to maneuver, (d) safety, (e) driving comfort and convenience, and (f) vehicular operating costs. Seven levels of service are described. (See Fig. 13.)

Level of Service A describes a free flow accompanied by low volumes, low densities, and high speeds which are controlled by the driver desires and physical roadway conditions (free speed). Although the variance in speeds is high, there is no restriction in maneuverability due to the presence of other vehicles, and drivers can maintain their desired speeds with little or no delay. This is the service expected in rural locations.

Level of Service B, C, and D describes the zone of stable flow. The upper limit is set by the zone of free flow, whereas the lower limit is defined by the optimum density, k'_m, and optimum speed, u'_m, based on maximizing the kinetic energy of the traffic stream. The conditions at u'_m and k'_m are acceptable for urban design practice. The divisions associated between levels B-C and C-D are arbitrary.

Level of Service E_1 & E_2 describes the zones of unstable flow. Zone E_1 is set between the optimum conditions described by the energy and momentum criteria. In this zone, small increase in volume is accompanied by both a large decrease in

a traffic density greater than k_m, yet with a vehicular flow greater than q_m. This type of operation cannot persist and leads inevitably to congestion.

Level of Service F describes a forced flow condition at low speeds and very high internal friction. Volumes are below capacity and storage areas consisting of queues of vehicles form. Normal operation is not achieved until the storage queue is dissipated.

Measurements and Relationships

The primary characteristics of traffic movement are concerned with speed, density and volume. These three fundamental characteristics are dependent on the geometric design of the roadway and the operational requirements of the traffic stream. Interest in these characteristics is manifest in the need for establishing representative possible and practical capacities for freeway sections.

While any set of speed observations may be influenced by such items as demand, capacity, design, weather and controls, it is not generally appreciated that the location at which measurements are made has a great deal to do with the adequate description of operating conditions. For example, traffic data taken just beyond an entrance ramp may reflect smooth and uniform operation when actually the traffic behind the ramp may be operating under stop-and-go conditions. Because congestion at one point may cause congestion for a great distance back along the freeway, a survey made at a "point" behind this critical ramp area will reflect poor conditions (low speed and relatively low volume) without a direct association with the cause of the congestion being possible.

Motion picture study procedures provide the advantage of having a view of a reasonably long section of freeway and reasonably precise measurements of numerous traffic characteristics such as speed, volume and density. However, even with the view of a section of freeway 1,000 to 2,000 feet in length, it is often difficult to determine accu-
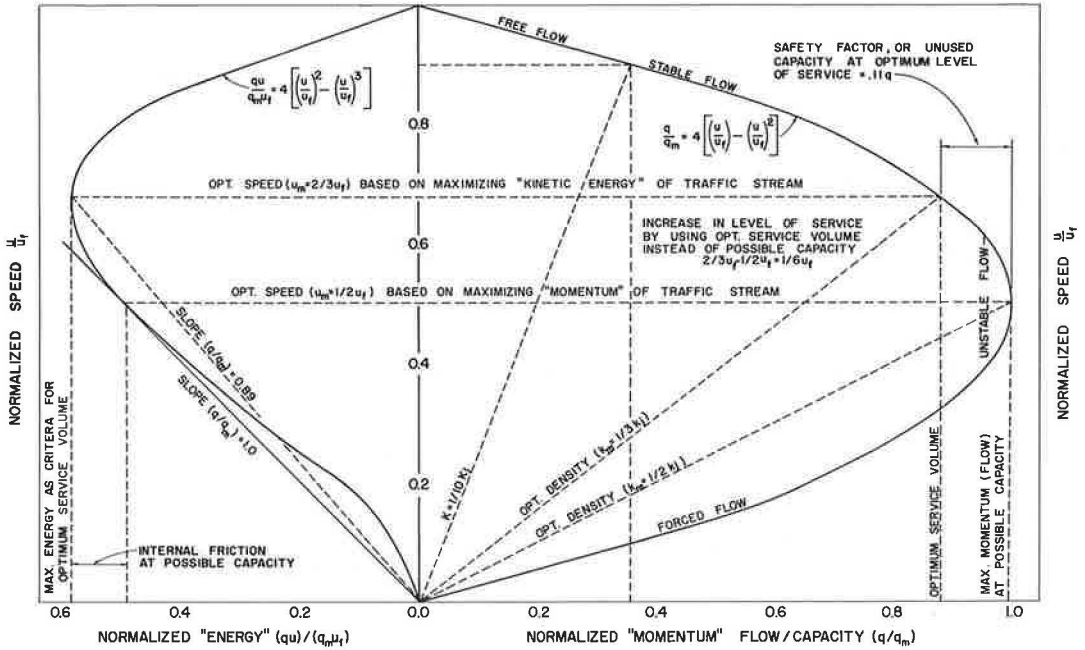
Figure 13. Quantitative approach to level of service-capacity relationship using energy-momentum analogy.

rately the cause of congestion. Congestion at one study area may actually be caused by conditions existing at a point farther along the freeway. Although the motion picture provides some possibilities of continually examining conditions throughout a section for possible influencing factors, the section studied from a single camera location is not always long enough to reveal whether congestion and "stoppages" are caused by conditions within the study area or by conditions ahead.

Two methods which have been successfully utilized and that do not rely on point survey data for describing flow characteristics are television camera surveillance and aerial photography. Pursuant to this research, two types of aerial photography have been studied (10): (1) strip photography where two continuous pictures are taken simultaneously over the entire study section, and (2) time-lapse photography where individual overlapping pictures are taken at short intervals of time. Time-lapse photography seems to be more suited for speed and density measurements, and can provide multiple speeds for each vehicle from which acceleration can be calculated. Of course, the shortcomings of the aerial approach, though distinct from the point survey approach, are nevertheless significant in that it is essentially an "instant" survey.

By careful design of a freeway study, either "point" or "instant" surveys can be used to give continuous coverage in both time and space. This two-dimensional interpretation is indispensable because all traffic characteristics vary in both time and space. "Contour maps" provide a means of illustrating these two-dimensional variations in the characteristics. These maps are drawn by using time as the ordinate, and distance along the freeway as the abscissa. If aerial photography is used (the "instant" survey approach), flight runs must be made at periodic intervals (about 10 minutes), and the speeds are averaged at about 600-ft intervals giving about 8.5 points per mile per 10 minutes. Interpolating in both time and space and connecting points of equal speed yield a speed contour map. Obviously, the same contour map could be obtained by interpolating between "point" surveys spaced at 600-ft intervals. Figure 14 illustrates speed contours for total inbound traffic on the Gulf Freeway, Houston, obtained from the aerial photographic survey method.
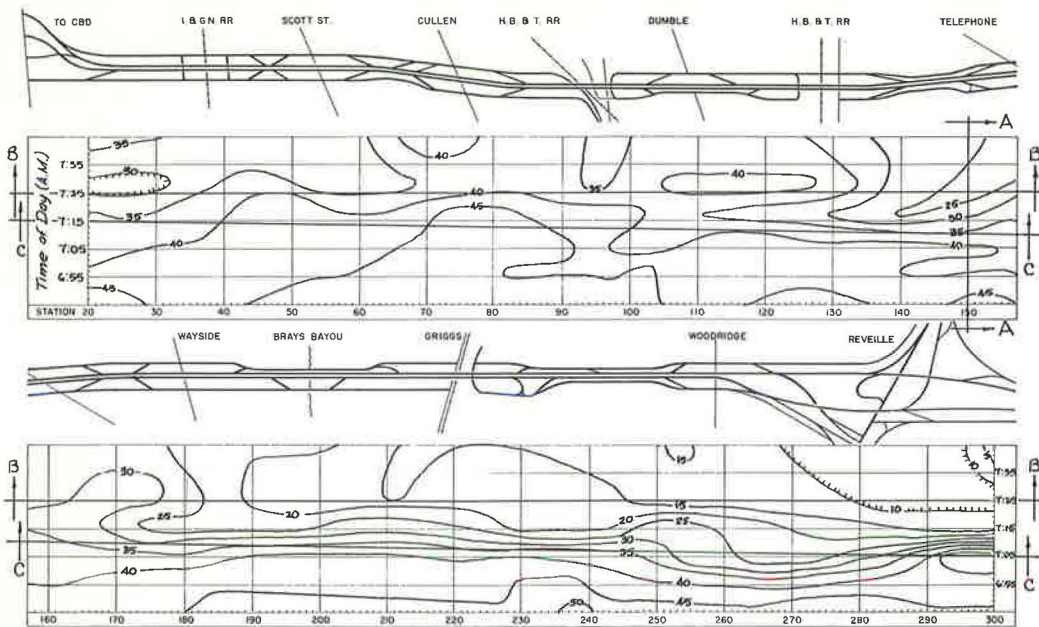
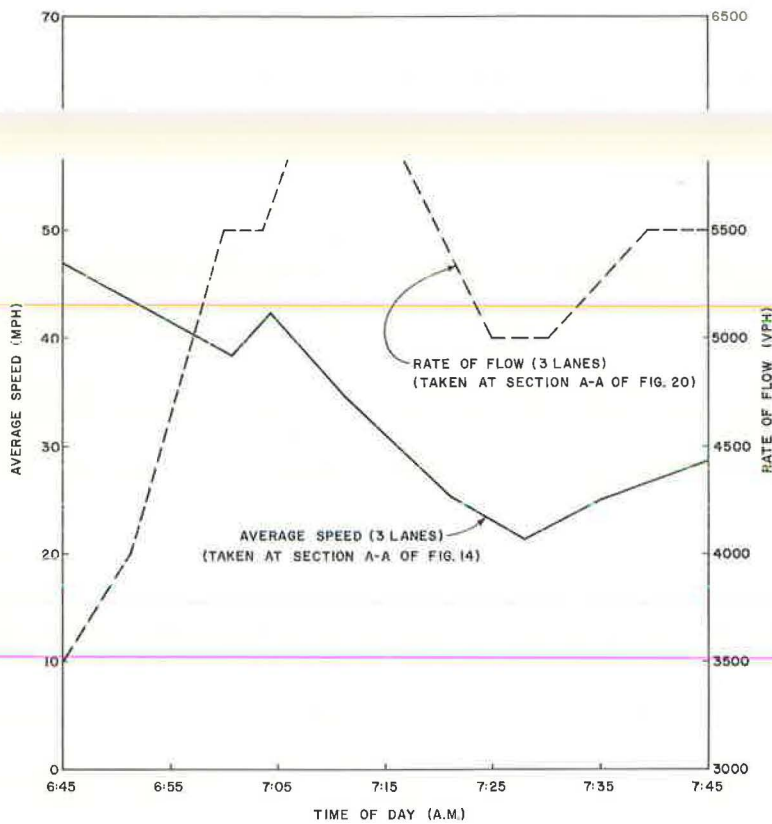Figure 14. Speed contours (total inbound traffic).

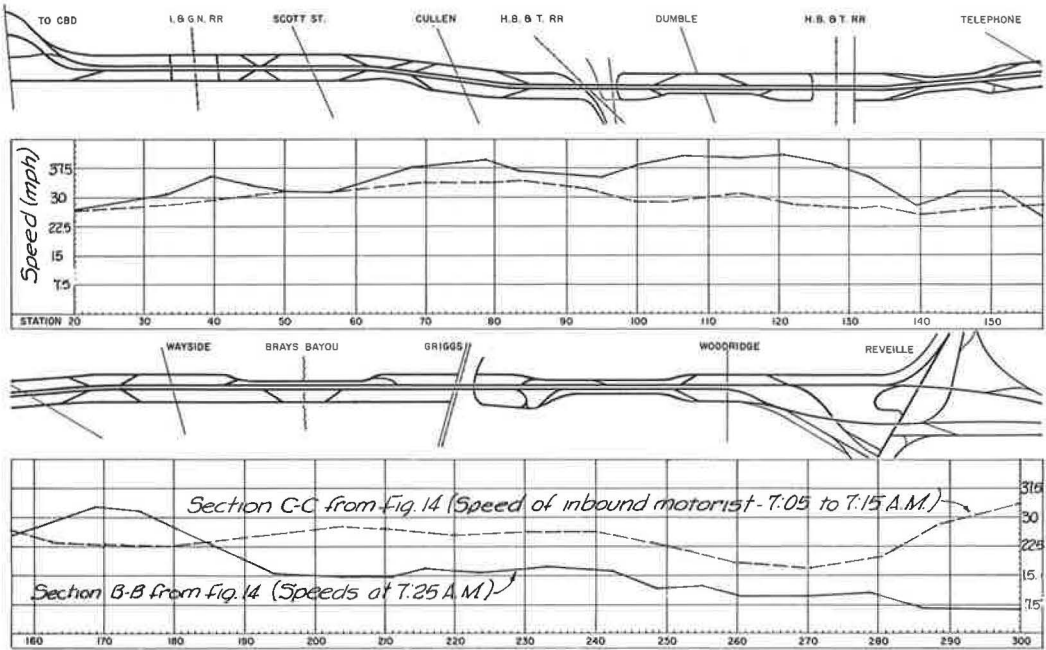Figure 15. Speed-flow relationship as obtained from contour maps.

Figure 16. Speed profiles obtained from "contour map."

In addition to providing a continuous record of speed for an entire facility for a sustained period such as the peak hour, the contour map affords the opportunity to isolate critical locations and periods and then to study them in more detail. Consider, for example, the Telephone Interchange Entrance Ramp located at Station 150 + 00 (Fig. 14). The profile of section A-A is plotted in Figure 15. A speed-flow relation, obtained solely from contour maps, is illustrated. A profile plotted from section B-B' in Figure 14 is plotted in Figure 16. This profile illustrates the performance of the entire facility at 7:25 a.m., as reflected by speeds.

Although profiles at either a point or at an instance may have some conceptual appeal to the engineer in evaluating freeway operations, the level of service concept previously discussed is based on the driving conditions afforded an individual motorist as reflected, for example, by his speed. Section C-C in Figure 14 indicates the path of a hypothetical motorist traversing the freeway starting from the Reveille Interchange at 7:05 and arriving at the downtown distribution system at 7:15. This profile is also plotted in Figure 16.

It is generally accepted that the largest number of vehicles that can pass a given point in one lane of a multilane highway, under ideal conditions, is between 1,900 and 2,200 veh/hr. This represents an average maximum volume per lane sustained during the period of one hour. Figures 17, 18, 19 and 20 illustrate volume contours for the three inbound lanes on the Gulf Freeway, Houston, as well as the three-lane total for inbound traffic. It is evident that higher rates of flow exist for specific lanes or for short periods of time. Although the capacity of a freeway under ideal conditions is considered to be 2,000 veh/lane/hr, only in situations where the peak period demand extends very nearly for the entire hour will this capacity value be realized.

To illustrate the speed-density-volume relationship over shorter sections of freeway, the 14 graphs of Figure 21 were plotted for the inbound shoulder lane. Each graph represents the average conditions encountered throughout the study hour for the 2,000-ft length of freeway directly above the graph. The graphs' ordinate is speed in miles per hour and the upper abscissa is volume in vehicles per hour, while the lower abscissa is density in vehicles per mile. The numbers from 1 to 6 refer to the times of the 6
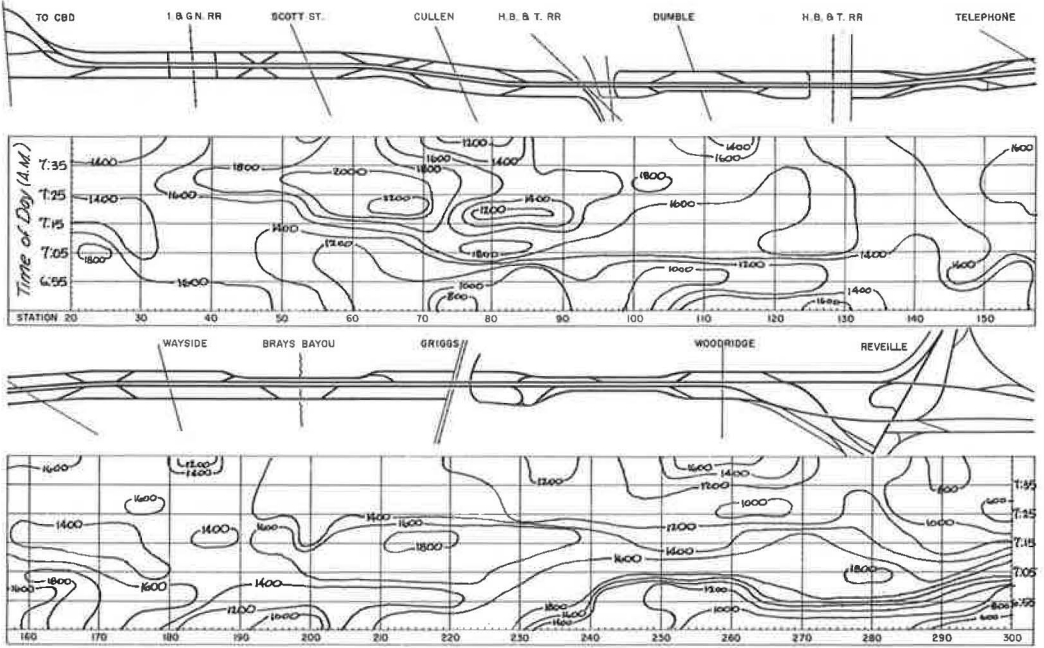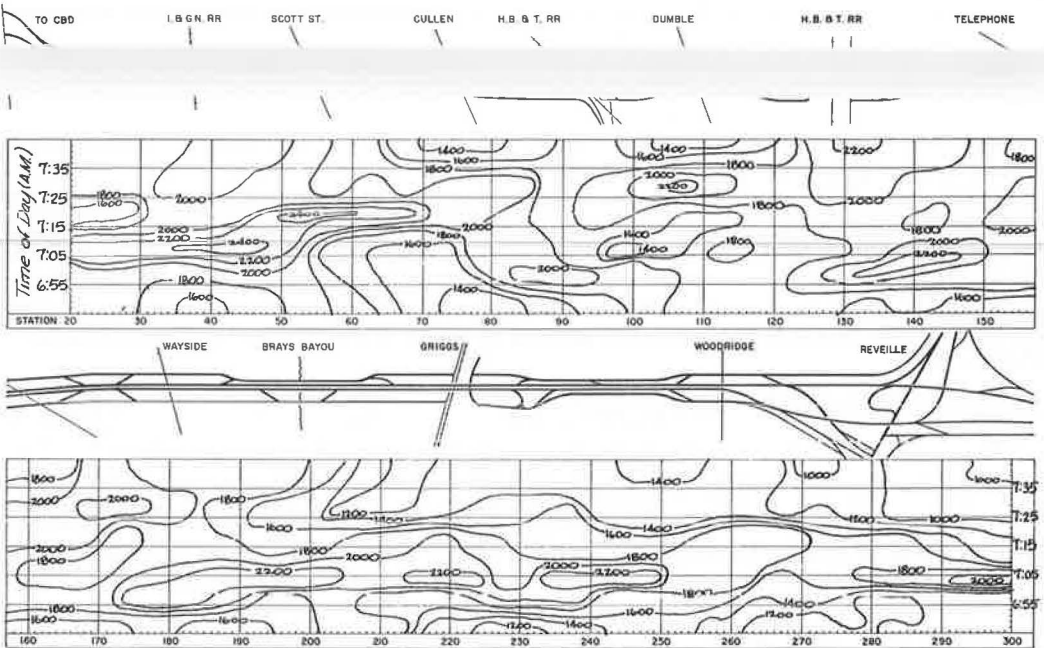
Figure 17. Volume contours (shoulder lane).

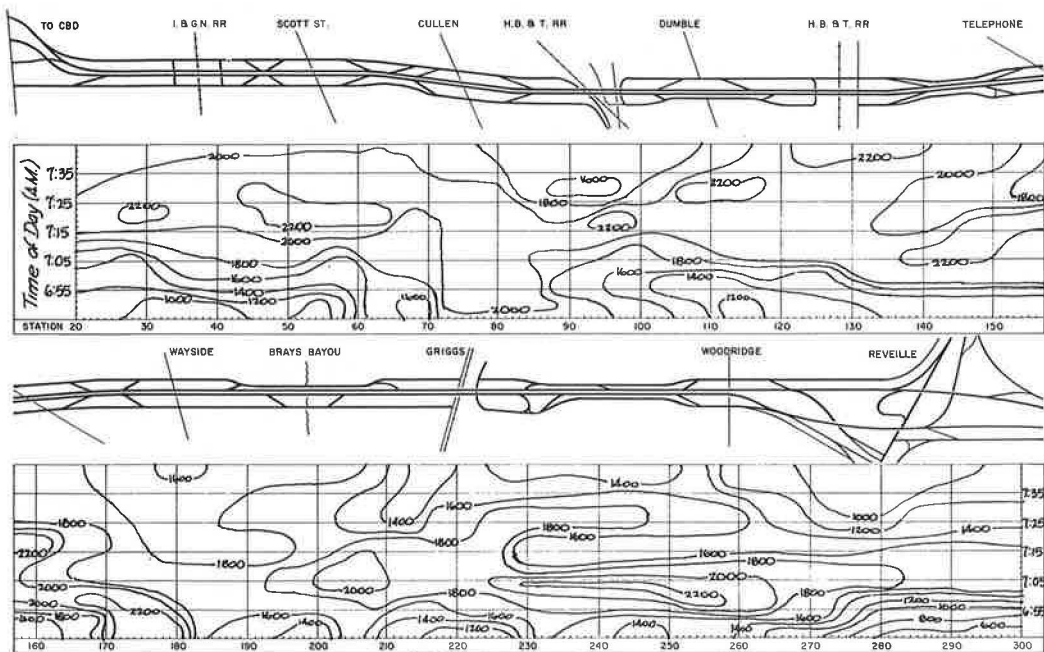Figure 18. Volume contours (center lane).
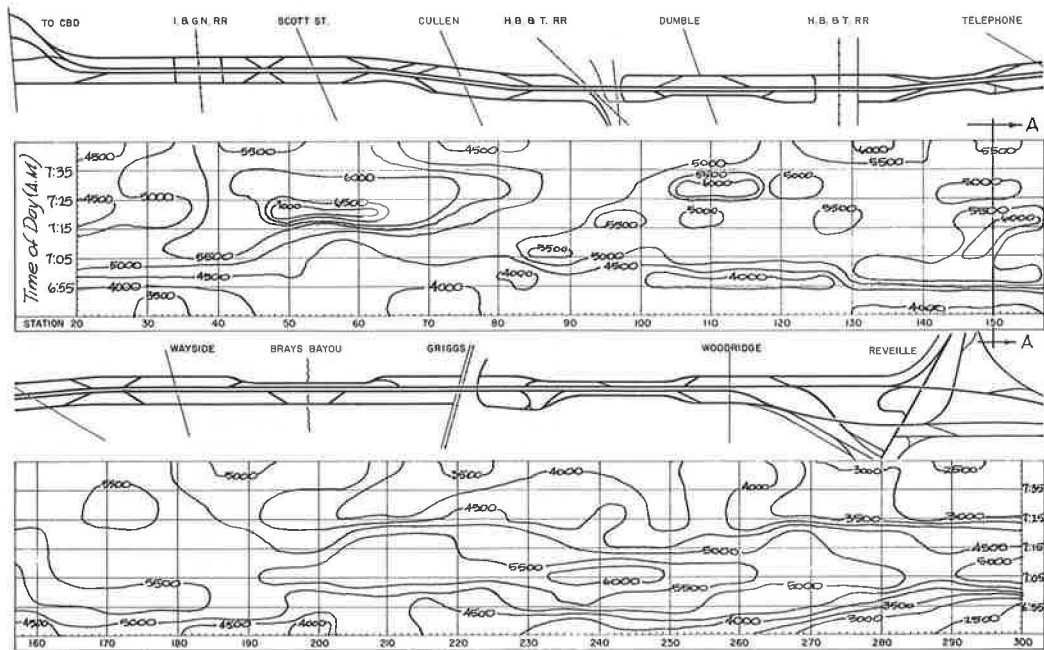
Figure 19. Volume contours (median lane).

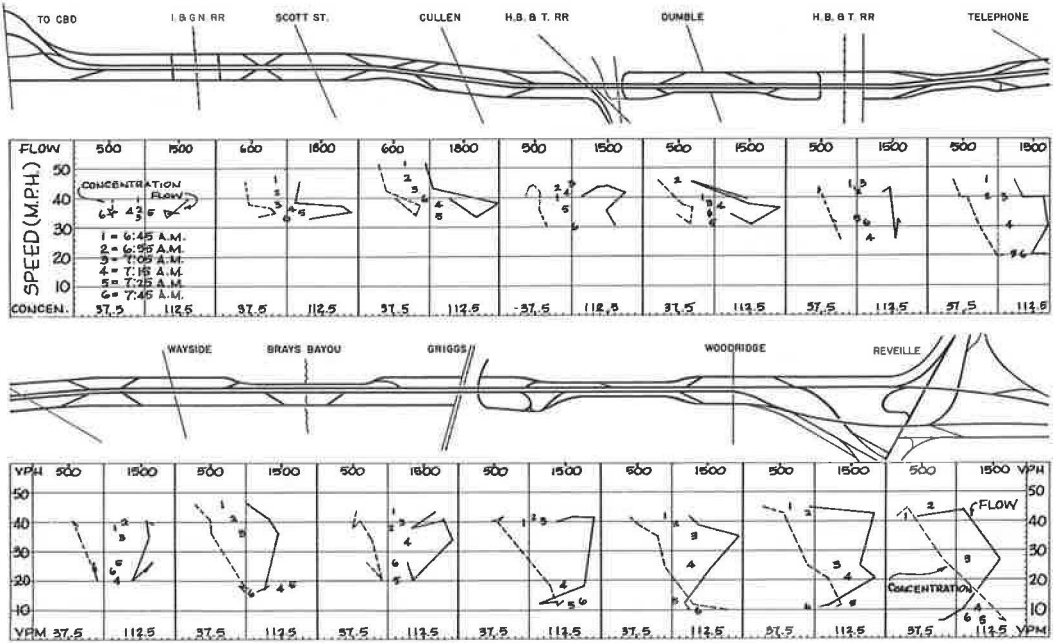Figure 20. Volume contours (total inbound traffic).

Figure 21. Observed speed-volume-density relationships (lane 1) in time and space.

flight runs, thus giving a chronological plot of data. The speed-density relationship is shown by a dotted line; the speed-volume relationship is shown by a solid line.

the interval within the Reveille Interchange, the range in average speeds varies from 45 to 5 mph over the study hour. At the other end of the freeway, however, average speeds remain at from 40 mph to 35 mph during the same period. The range of average speeds plainly decreases throughout the morning peak hour as one travels toward the CBD on the facility.

The volume-speed relationship is more difficult to explain. Decreases in speed do not always accompany increases in flow. However, several of the graphs exhibit a characteristic parabolic loop resulting from the decrease of speed at excessive flows. Keese, Pinnell and McCasland (2) explain that as peak flows build up, the average speed drops and generally does not recover to the original relationship with volume until the peak flow or demand has passed. Ryan and Breuning (11) utilize the concept of critical vs noncritical flow, with the dividing point being the maximum flow. They report that all three relations among speed, volume and density are linear within the noncritical flow region (before congestion sets in). May, et al. (12) define 3 zones which may be described as constant speed, constant volume and constant rate of change of volume with density. In zone 1, the speed of the vehicle is determined by the facility itself and the volume matches the demand. Zone 2 represents impending poor operations; average speed drops but the flow rates may be sustained at a high level. In zone 3 both speed and volume rates decrease, which in itself may serve as a definition of congestion.

The analysis of highway traffic is more than the making of measurements and collection of facts. Although this exploration of the true nature and characteristics of freeway traffic is the necessary beginning in providing new ways of improving performance, the freeway traffic phenomena are so complex that a collection of bare data tells us a little more than we already know. Earlier in this article, a hydrodynamic model was explained. An important aspect of this model is the momentum-energy analogy which is summarized in Figure 13. These concepts yield several parameters which provide the means of organizing and interpreting the traffic characteristics collected.
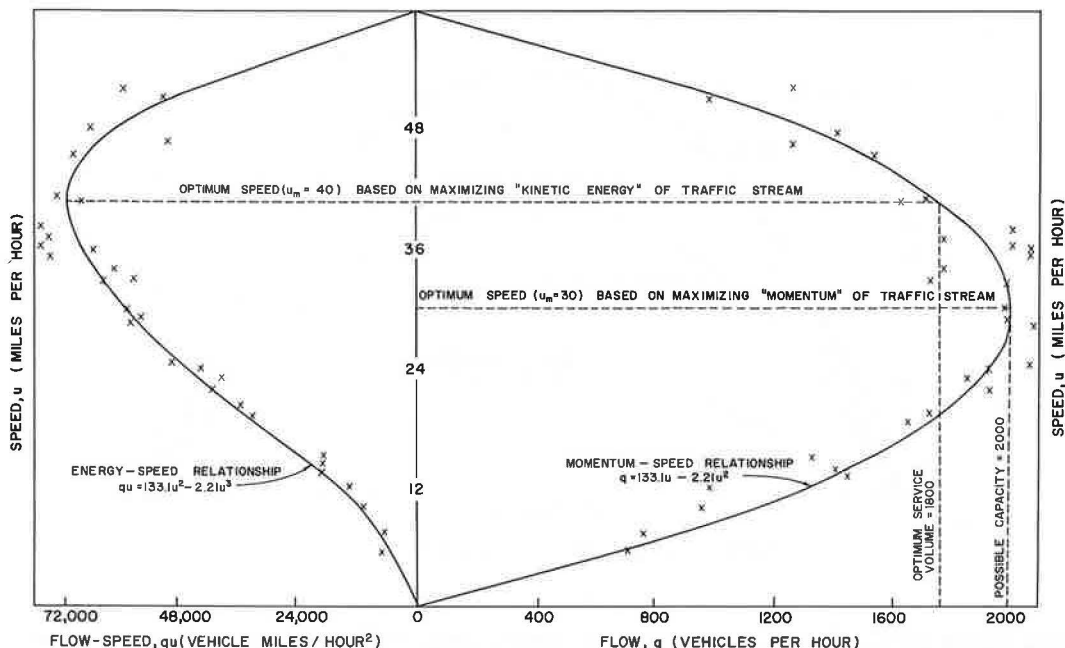
Figure 22. Verification of energy-momentum analogy.

Using Figure 22 as a model, volume vs speed and (volume × speed) vs volume data were plotted for lane 3 of the Gulf Freeway, inbound traffic, during the peak hour. The constants for a curve of the form

$$q = k_j u \left[1 - \left(\frac{u}{u_f} \right) \right] \tag{22}$$

was calculated by fitting a regression line to speed-density data taken from aerial photographs. Values of free speed, u_f, and jam density, k_j, are 60.3 mph and 133.1 vpm. Thus, the equation of the curve in the q - u plane becomes

$$q = 133.1 \, u - 2.21 \, u^2 \tag{23}$$

The equation for the "energy"-speed relationship is

$$qu = 133.1 \, u^2 - 2.21 \, u^3 \tag{24}$$

It can be observed that the curves provide excellent estimates of the points plotted, and verify the theoretical approach to the capacity-level of service concept suggested by the momentum-energy analogy of the hydrodynamic model of traffic flow.

The hydrodynamic model, so useful in describing the level of service-capacity concept, can be extended to aid in the explanation of a bottleneck (9). A bottleneck is a stretch of roadway with a flow capacity less than the road ahead (Fig. 23). The upper q - k curve in Figure 23 is for the roadway ahead of the bottleneck, and the lower one refers to the bottleneck itself. When the traffic volume reaches the capacity of the bottleneck, the velocity in the bottleneck, u_1, is less than ahead of the bottleneck, u_3. However, this difference in speed is not significant in urban area capacity problems
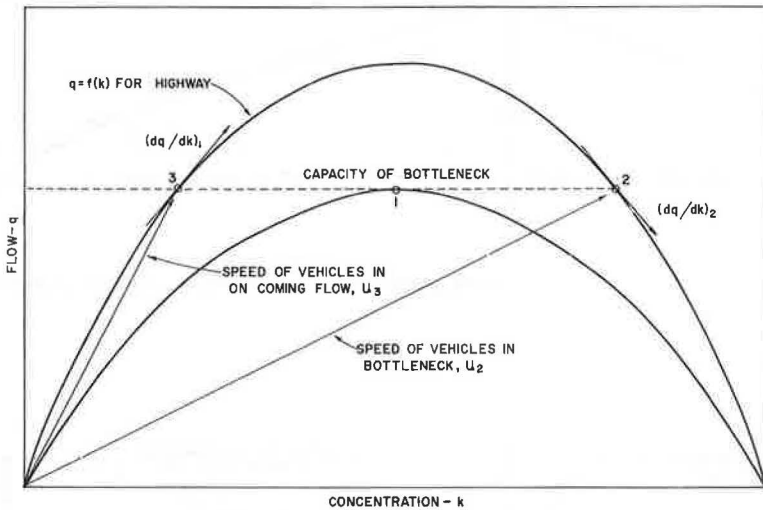
Figure 23. Traffic flow in a bottleneck.

and should not be used as a criterion for determining acceptable operation. But, any further increase in demand (volume) accumulates as a queue in advance of the bottleneck, and the traffic conditions in this region shift from those expressed by point 3 in Figure 23 to those expressed by point 2 (density changes from k_3 to k_2).

When a bottleneck is operating at capacity, the speed of traffic is independent of the geometric conditions in the upstream section. Because congestion may last much longer (Fig. 24) than that interval in which demand exceeds capacity, it is important that precautions be taken to prevent this. A stimulated rate-of-flow for a 5-min period

designer. This is illustrated by the design curves relating level of service to a 5-min rate of flow (Fig. 25). Thus, a service volume of 1,800 has a probability of 0.50 of guaranteeing "stable flow" during the peak 5-min period. On the other hand, there is a 50 percent chance of "unstable flow" occurring; and a 2.5 percent chance of "forced flow." Of course, forced flow is congestion, and unstable flow can lead to congestion, due to the statistical variability of the vehicle headways. It is interesting to note that a relatively small reduction of 100 in the service volume, to a flow of 1,700 vph, greatly increases the probability of maintaining stable flow. These curves represent an attempt to put such a decision in the hands of highway administrators and designers. After this choice is made, the service volume to be used for design for the peak hour would be obtained from Figure 1, depending on the population of the city. (See Table 3.)

Figure 25 was obtained by determining the probability of getting observed rates of flow greater than the predicted values utilizing the data shown in Figures 2 and 3 and assuming the errors are normally distributed. The first four columns in Table 3 are taken directly from Figure 25; the last four columns utilize the peaking relationships expressed in Figure 1.

APPLICATIONS

Freeway Design

Highway design is an engineering function—not a handbook problem. The engineer is faced with the problem of predicting traffic demands in future years and providing facilities that will accommodate that traffic under a selected set of operating conditions or levels of service. Too often highway design has been accomplished by adopting a set of handbook "standards" which when coupled with traffic "guestimates" have resulted in the construction of many seriously inadequate facilities.
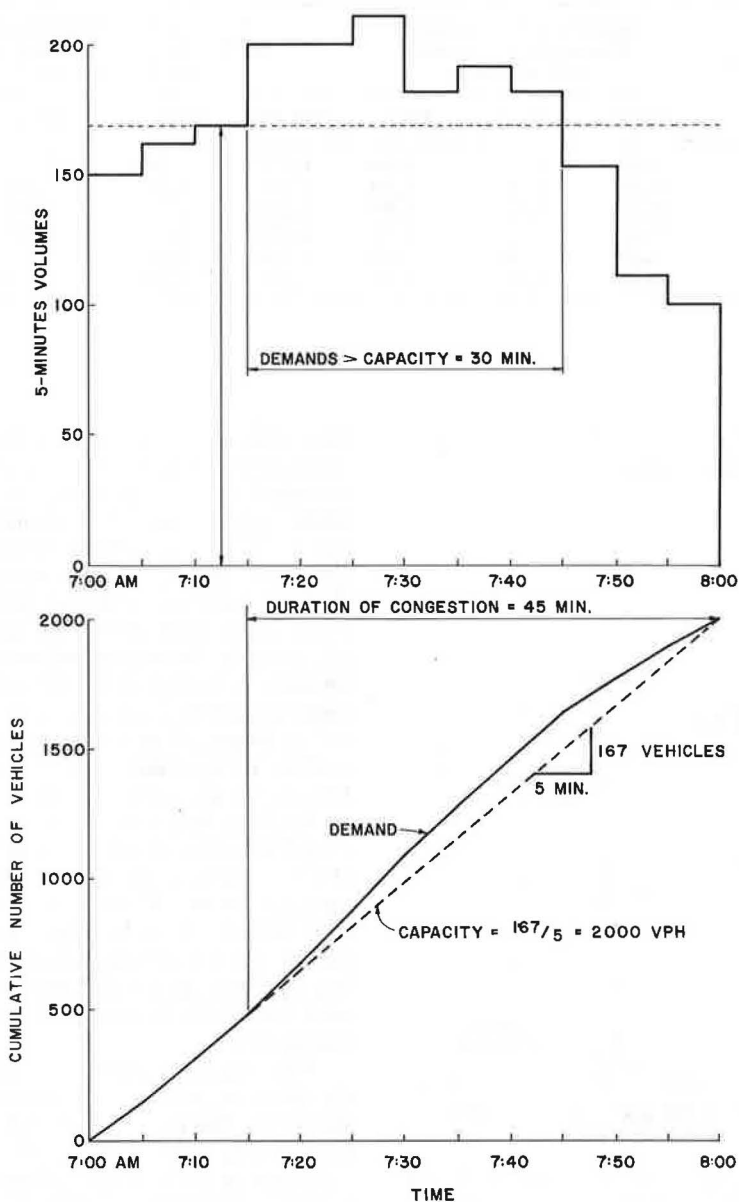
Figure 24. Relation between demand, capacity and congestion.

Traffic prediction, traffic operation and design have now developed to the point where it is possible for engineering (the application of science) to produce rather reliable results.

A freeway is _not_ built for some date 20 years in the future. It must go to work the first day and serve efficiently all through its expected life. And, if history is not changed, many will be serving for quite a number of years beyond the "design" year.

The freeway is only one facility in a network of system of streets and highways. It has its place, but the system as a whole must be made to function efficiently. The day

TABLE 3

FREEWAY CAPACITY WITH CONFIDENCE LIMITS

Peak 5-Min Flow (vph)	Approx. Probabilities of Various Types of Flow in Peak 5-Min			Freeway Design Service Volume by Population of Metropolitan Area (total hourly vol/lane)			
	Stable	Unstable	Forced	100,000	500,000	1,000,000	5,000,000
1,500	1.00	0.00	0.00	1,100	1,200	1,300	1,300
1,600	0.98	0.02	0.00	1,200	1,300	1,300	1,400
1,700	0.85	0.15	0.00	1,300	1,400	1,400	1,500
1,800	0.50	0.48	0.02	1,400	1,500	1,500	1,600
1,900	0.15	0.69	0.16	1,500	1,600	1,600	1,700
2,000	0.03	0.47	0.50	1,500	1,600	1,700	1,800
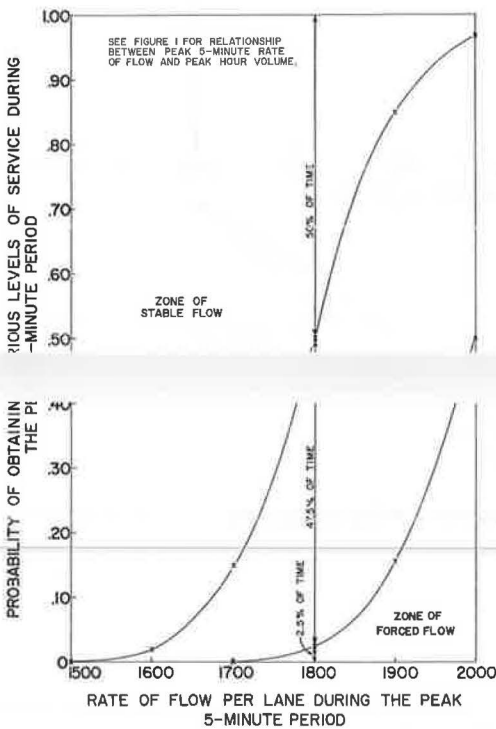
Figure 25. Design curves relating level of service to flows during the peak 5-min.

has gone when a freeway can be designed within the confines of two parallel right-of-way lines. Likewise, the day has gone when only the 20-yr "complete system" can be considered when designing a particular facility. Traffic projections and designs must be made on partial or incomplete systems if desirable service is to be obtained in the years before the whole system is completed. With the modern tools available, the designer should have at his disposal an accurate estimate of traffic demand for each stage of comple-

Engineering and management must be coupled in the selection of a level of service for design that is best adapted to the specific need. Economics and other factors will continue to play a major part in facility programming and even in design, but realistic projected service analysis will lead to more realistic priority programming.

The largest number of vehicles that can pass a given point in one lane of a multilane highway, under ideal conditions, is between 1,900 and 2,200 vph. This represents an average maximum volume per lane sustained during the period of one hour. Studies have found higher volumes for specific lanes or for short time periods. Where at least two lanes are provided for movement in one direction, and disregarding distribution by lanes, the capacity of a freeway under ideal conditions is considered to be 2,000 veh/lane/hr, as explained earlier in this report.

Where conditions are less than ideal because of reduced widths, sight distance, grades and commercial vehicles, etc., the capacity will be somewhat lower. Moskowitz and Newman (5) suggest some correction value to be used when these conditions are anticipated.

When the traffic volume equals the capacity of a freeway, operating conditions are poor. Speeds are low, with frequent stops and high delay. In order for the highway to

TABLE 4

LEVELS OF SERVICE AS ESTABLISHED BY ENERGY-MOMENTUM CONCEPT

Level of Service Zone	Description	Zone Limits[a]	
		Upper	Lower
A—free flow	Speeds are controlled by driver desires and physical roadway conditions; type of service expected in rural locations.	u_f	$0.9\,u_f,\ 0.35\,q_m$
B, C & D—stable flow	Flow concentration, speed $^{-1}$. Conditions at u'_m and k'_m are acceptable for urban design practice; divisions between B-C and C-D are arbitrary.	$0.9\,u_f,\ 0.35\,q_m$	$u'_m,\ q'_m$
E₁—unstable flow	A small increase in demand (flow) is accompanied by a large decrease in speed leading to high densities and internal friction which contribute to instability.	$u'_m,\ q'_m$	$u_m,\ q_m$
E₂—unstable flow	This type of high density operation cannot persist and leads inevitably to congestion.	$u_m,\ q_m$	$0.5\,u'_m,\ q'_m$
F—forced flow	Flows are below capacity and storage areas consisting of queues of vehicles form. Normal operation is not achieved until the storage queue is dissipated.	$0.5\,u'_m,\ q'_m$	0

[a] See Figure 13.

provide an acceptable level of service to the road user, it is necessary that the service volume be lower than the capacity of the roadway.

The level of service approach establishing levels of operation from free flow to capacity which is being considered by the HRB Committee on Highway Capacity is designed to allow the engineers and administrators to provide the highest level of service economically feasible. The momentum-energy analogy derived in the previous section is an effort to explain the capacity-level of service relationship rationally and quantitatively (Table 4). It must be recognized that highway traffic represents a stochastic phenomenon. Therefore, any highway facility, designed to accommodate traffic, must be designed with the realization that from time to time demand will exceed capacity. The organization of Table 3 is useful in that it provides the designer with confidence limits in determining the number of main lanes needed on a freeway.

After the determination of the number of freeway lanes, 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 intervehicular conflicts within the traffic stream, or a combination of both. An entrance ramp is an example of the first type of critical location, whereas exit ramps and grades can cause intervehicular conflicts.

It is interesting to note that a distinction can be made between the terms "critical location" and "bottleneck." A bottleneck is a section of roadway with a capacity lower than the adjacent upstream section. Thus, the volume input can exceed the capacity of

the bottleneck, and the roadway upstream becomes a storage area whose level of serv-
ice and rate of flow are governed by the capacity and operating conditions through the
bottleneck. If traffic backed up at a bottleneck is required to stop, the capacity of the
bottleneck becomes a function of vehicular departure headways from a stopped condi-
tion. Strictly speaking then, not all "critical locations" are "bottlenecks"; a bottleneck
is only one form of critical location. The most important consideration is that opera-
tion in critical sections never drops below the adopted level of service. Often, this
can be effected by not allowing the demand on the facility to exceed the bottleneck capac-
ities.

As traffic operations deteriorate, vehicles tend to form platoons. In general,
"platooning" is a function of the number of slow vehicles and the speed of slow vehicles.
When an upgrade is introduced, speeds are reduced and platoon lengths increase. Ac-
cident potential and capacity mitigating lane change maneuvers are a direct result.
The percentage of trucks in the stream, as well as the steepness and length of grade,
will determine just how adverse this effect may be.

One possible solution to maintaining a level of service on the grade equal to a level
grade, would be to add a climbing lane whenever the passenger car volume and speed
varies below the adopted level of service. This, of course, is not always feasible on
an urban freeway facility where lane reductions present severe operational problems.
The policy of restricting trucks to the outside lane (predicated on the theory that if all
trucks are traveling on the outside lane, then vehicles on the remaining lanes can main-
tain the operation levels achieved on level grade) ignores the fact that vehicles in the
outside lane will not accept the same level of service as trucks and will attempt to
change lanes. It is impossible for adjacent lanes to operate at drastically different
levels of service. This influence of the operation of one lane on the operation of another
is a well-established speed characteristic and is sometimes called "speed sympathy."
Thus, the most acceptable solution to the problem of operating on freeway upgrades is
to see that demand volumes never exceed the adopted level of service.

The traffic demand on a freeway can only change at entrance or exit ramps. Two of
the most critical points on a freeway will be upstream from an exit ramp and down-

described at an upgrade, but can be much more severe where there is a back-up from
the exit ramp onto the main roadway proper. Many exit ramp problems could be
avoided by providing for the speed reduction on the ramp rather than on the shoulder
lane of the freeway. Even where long parallel deceleration lanes are provided, they
are not used because of the unnatural maneuver involved. Unfortunately, the close
spacing of interchanges and use of frontage roads favor the use of short slip-type ramps.
Where a high exit volume slip ramp is used, definite consideration should be given to
placing yield signs on the frontage roads, thus preventing back-up from the exit ramp
onto the freeway.

Entrance ramps may 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.

There are three basic procedures employed in determining the capacity of entrance
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 vol-
umes per lane (discussed in the first section of this paper and also treated extensively
by Hess (13)) and then limits the ramp volume to the merging capacity (assumed here
to be equal to the service volume selected in Table 3) 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 (8).
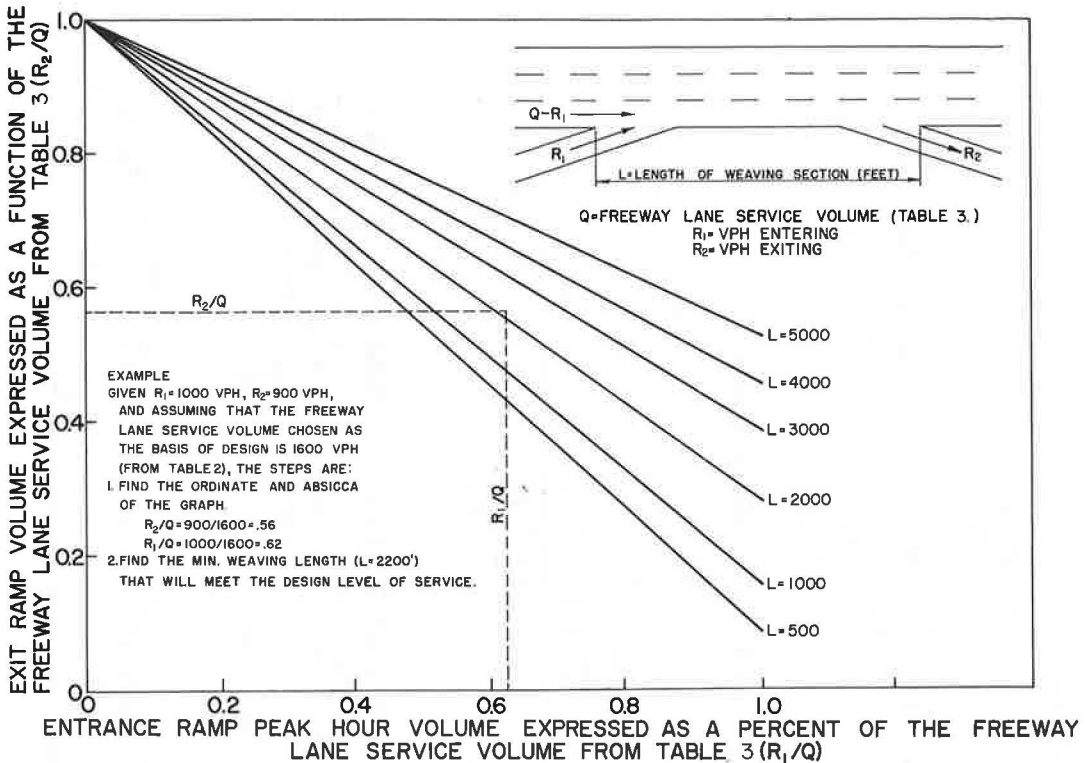It is believed that the second method (Figs. 5 to 9) is the most practical in designing a

EXIT RAMP VOLUME EXPRESSED AS A FUNCTION OF THE FREEWAY LANE SERVICE VOLUME FROM TABLE 3 (R_2/Q)

1.0

0.8

0.6

R_2/Q

0.4

0.2

0

Q-R_1

L=LENGTH OF WEAVING SECTION (FEET)

Q=FREEWAY LANE SERVICE VOLUME (TABLE 3.)
R_1= VPH ENTERING
R_2= VPH EXITING

L=5000
L=4000
L=3000
L=2000
L=1000
L=500

EXAMPLE
GIVEN R_1= 1000 VPH, R_2= 900 VPH,
AND ASSUMING THAT THE FREEWAY
LANE SERVICE VOLUME CHOSEN AS
THE BASIS OF DESIGN IS 1600 VPH
(FROM TABLE 2), THE STEPS ARE:
1. FIND THE ORDINATE AND ABSICCA
OF THE GRAPH
R_2/Q= 900/1600= .56
R_1/Q= 1000/1600= .62
2. FIND THE MIN. WEAVING LENGTH (L= 2200')
THAT WILL MEET THE DESIGN LEVEL OF SERVICE.

R_1/Q

0   0.2   0.4   0.6   0.8   1.0

ENTRANCE RAMP PEAK HOUR VOLUME EXPRESSED AS A PERCENT OF THE FREEWAY LANE SERVICE VOLUME FROM TABLE 3 (R_1/Q)

Figure 26. Determination of minimum length of weaving section to meet design level of service.

new facility. The first method is predicated on knowing the capacity of bottlenecks— something that is not known in the case of a new facility. Research concerning the third method is now underway; the advantage of this approach is that it recognizes that ramp capacity and operation must be affected by the geometrics of the ramps.

The last "critical location" to be considered is the weaving section. Weaving sections often simplify the layout of interchanges and result in right-of-way and constrution economy. The capacity of a weaving section is dependent on its length, number of lanes, running speed and relative volumes of individual movements. When large volume weaving movements occur during peak hours, approaching the possible capacity of the section, probably results are traffic stream friction, reduced speeds of operation, and a lower level of service. This can sometimes be avoided by the use of additional structures to separate ramps, reversing the order of ramps so as to place the critical weaving volumes on frontage roads, and the use of collector-distributor roads in conjunction with cloverleaf interchanges.

Weaving sections should be designed, checked and adjusted so that their capacity is greater than the service volume used as the basis for design. This is consistent with the level of service concept used in determining the number of main lanes and checking the merging capacities at entrance ramps. The determination of minimum length of weaving section to meet the controlling level of service is illustrated in Figure 26. These relationships were obtained by considering the outside lane use relation with trip length (Figs. 11 and 12). Referring to Figure 26, the maximum number of vehicles at an exit, R_2, cannot exceed Q - R_1, plus the number of entrance ramp vehicles that change lanes within the merging section.

Figure 27 illustrates four steps to be followed in the design of a freeway system, as follows:
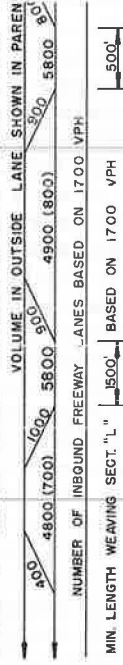
32

STEP 1 – PEAK HOUR VOLUMES (A.M. PEAK); DIRECTIONAL ...TRIBUTION = 2:1
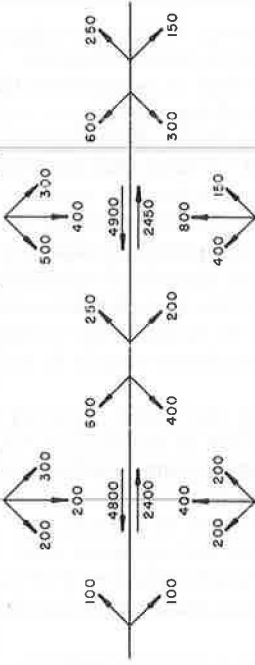
STEP 2 – INTERCHANGE REQUIREMENTS (A.M. PEAK) – SEE ...GURES 28 & 29 FOR PHASING & CAPACITY
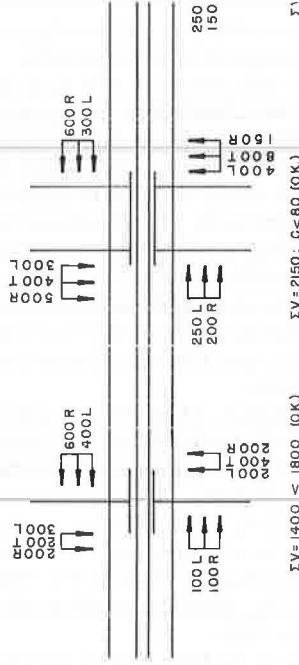
STEP 3 – MAIN LANE REQUIREMENTS – A.M. PEAK (...LE 3) & CHECK OF CRITICAL LOCATIONS (FIGS. 5 TO 9, & 26)
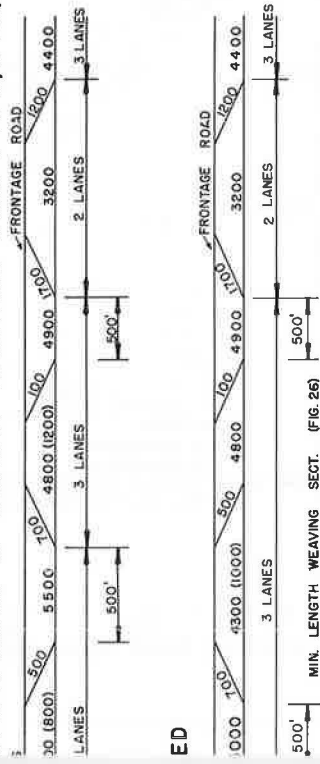
STEP 4 – ALTERNATE DESIGN WITH RAMPS REVERSED

Figure 27. ...way design procedure.
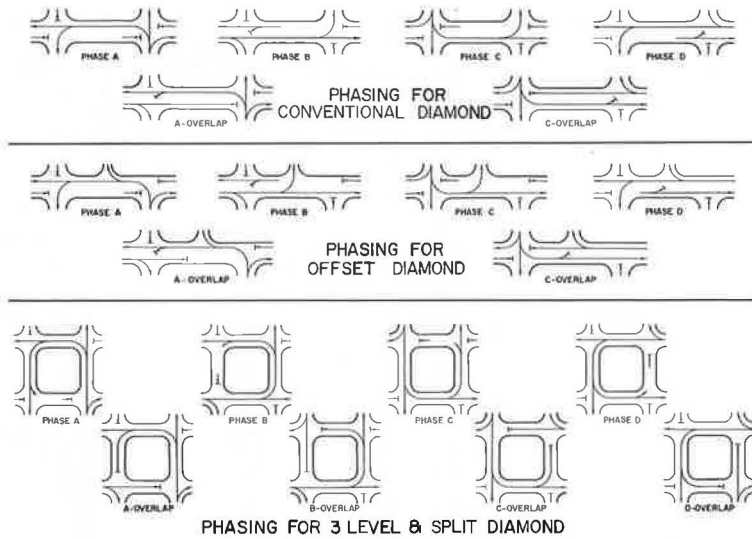
Figure 28.

Step 1—Determine the peak hour volumes through the application of the peak hour and directional distribution factors to the assigned daily traffic volumes. In an actual problem the PM peak would also be checked.

Step 2—Determine interchange requirements. It is important that this be done before freeway main lane requirements be investigated, because the number of ramps depends on the choice of interchange. Thus, a cloverleaf interchange and a directional interchange may have one or two entrance ramps and one or two exit ramps in each direction; whereas diamond interchanges have one entrance ramp and one exit ramp in each direction. If the interchange is to be signalized, a capacity check is made to see if the planned facilities will handle the traffic with reasonable cycle lengths (Figs. 28 and 29) (14). Should a facility be apparently underdesigned, additional approach lanes may be added or a higher type facility be substituted in its place.

Step 3—The number of main lanes depends on what service volume value is chosen as the design capacity. The freeway design service volumes in Table 3 enable the designer to judge what level of service can be expected for a given service volume based on the probability of obtaining various types of flow conditions during the peak 5-min period. For the purposes of this example a service volume of 1,700 vph is chosen. The operating conditions at critical locations must be checked to insure that the designated level of service is met at every point on the freeway. The critical sections considered in this paper are merging and weaving sections. Figures 5 to 9 provide the basis for determining if the merging capacities at entrance ramps are exceeded, where the merging capacity is defined as the service volume chosen in Table 3. Thus, since a total hourly volume of 1,700 vph is used as the basis for determining the number of lanes, then 1,700 vph would represent the merging capacity in this procedure. Figure 26 provides the basis for determining if weaving sections on the freeway meet the designated level of service.

Step 4—Alternate designs should always be considered. In Figure 27, one alternative is illustrated by merely reversing the order of entrance and exit ramps, resulting in 3 lanes in each direction instead of 4 lanes.

The level of service should be "in harmony" along the stretch of freeway being considered. Because operational problems at one point are reflected along the freeway for a distance depending on the volume-capacity relationship, it is not practical to consider a lower level of service at one or more critical points, rather the level of service selected for design should be met or exceeded at the critical or bottleneck points. This concept is referred to as balanced design and it is a must for freeways.
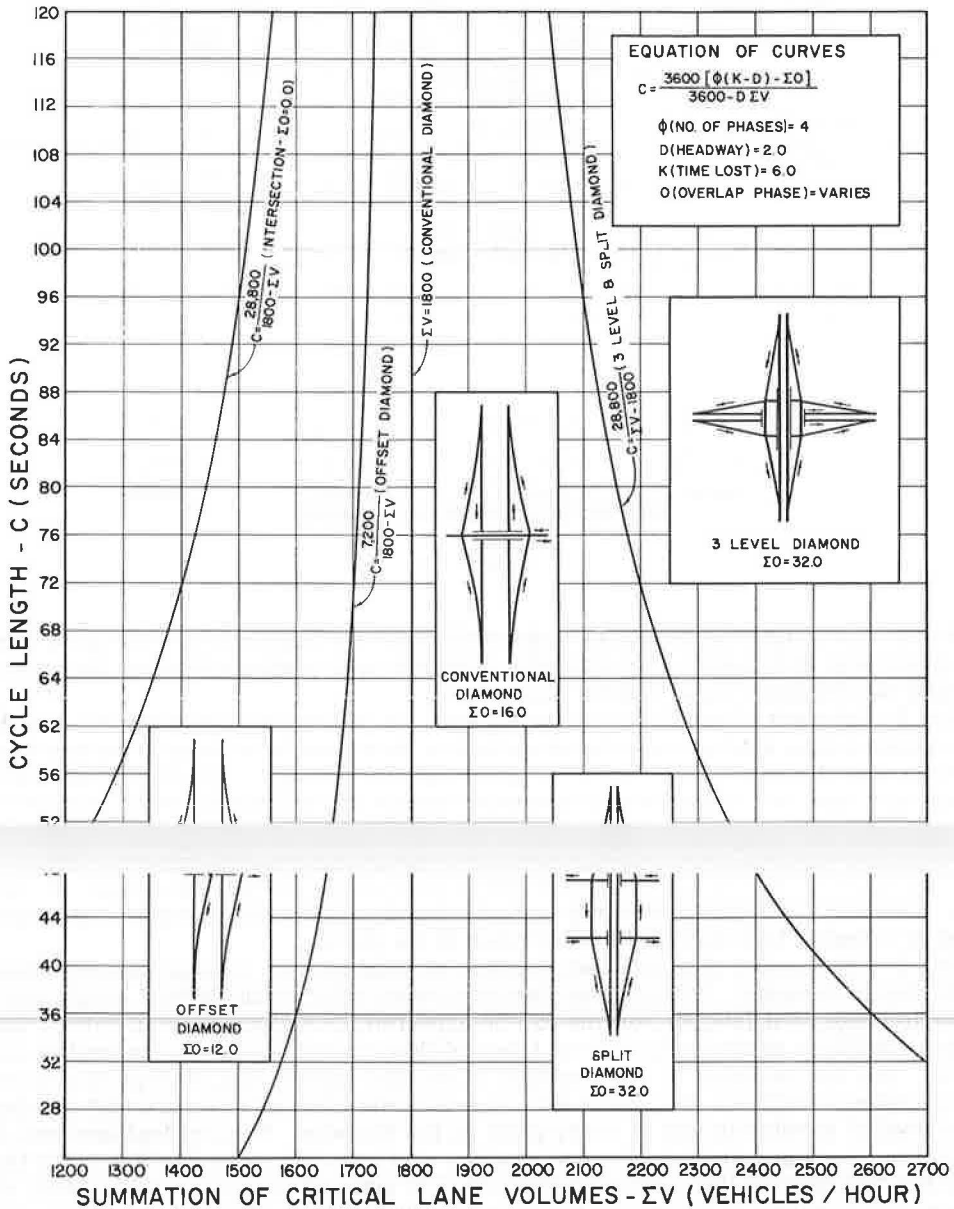
Figure 29. Design curves for 4-phase facilities.

Freeway Operations

The freeway motorist expects to have his needs anticipated and fulfilled to a much higher degree than on conventional roads. This expectation can sometimes be fulfilled by the application of capacity considerations to rational geometric design. More often than not, however, actual traffic and travel patterns differ from the projected values making constant freeway operational attention after construction a must.

Congestion occurs on a freeway section when the demand exceeds the capacity of that section for some period of time. Bottlenecks can be caused by changes in the freeway alignment (horizontal or vertical) or reductions in the freeway section (reduction of number of lanes, reduction in lane widths, the presence of an entrance ramp,

etc.). Accidents, disabled vehicles and maintenance or law enforcement operations can also cause temporary bottlenecks by reducing the effective capacity or level of service provided.

Freeway design does not always eliminate the need for sound traffic regulation. A reasonably homogeneous traffic stream, particularly with respect to speed, is essential for efficient freeway operations. Pedestrians, bicycles, animals and animal-drawn vehicles are excluded from freeways. Motor scooters, non-highway (farm and construction) vehicles and processions, such as funerals, are also generally prohibited from the freeway. Towed vehicles, wide loads or other vehicle combinations such as trailers drawn by passenger vehicles which impede the normal movement of traffic may be barred during the peak traffic hours or during inclement weather.

Minimum speed limits are being increasingly used and have been found of great benefit, particularly on high-volume sections. The effect of this type of control is to reduce the number of major accident potential lane change maneuvers. The effects of slow-moving vehicles on both capacity and accident experience are so pronounced that a greater use of minimum limits appears probable. There is a need to eliminate all vehicles incapable of compatible freeway operation.

Increasing attention has been given to the possibility of and need for using variable speed control on urban freeway sections as a means of easing the accordion effects in a traffic stream as congestion develops. Drastic speed variations might be dampened by automatically adjusted speed message signs in advance of bottlenecks.

A properly designed entrance ramp with provision for adequate acceleration should allow the entering driver adequate distance to select a gap and enter the outside lane of the freeway at the speed of traffic in the lane. These merging areas operate best when there is a mutual adjustment between vehicles from both approaches. "Yield" signs impose rather drastic speed restrictions under the laws of a number of states thus causing operational problems, and are no longer mandatory on the Interstate system. It is generally felt that any speed restriction or arbitrary assignment of right-of-way should be avoided unless inadequacies in the design make it imperative.

It is generally agreed that one key to significant progress in operation of urban freeways lies in improved surveillance technique. In its most basic form, urban freeway surveillance is limited to moving police patrols. Recently, helicopters have been used for freeway surveillance in Los Angeles and other communities. Efficient operation of high density freeways is, however, more than knowing the locations of stranded vehicles; it may require closing or metering entrance ramps, or excluding certain classes of vehicles during short peak periods. Therefore, what is needed is a reliable, all-weather source of surveillance information with no excessive time lag.

Experimentation with closed circuit television as a surveillance tool was initiated on the John C. Lodge Freeway in Detroit. This offers the possibility of seeing a long area of highway in a short or instantaneous period of time, made possible by spacing cameras along the freeway so that a complete picture can be obtained of the entire section of roadway. Evaluation of the freeway operation depends mostly on the visual interpretation of the observers. However, many traffic people believe that this is not enough. The Chicago Surveillance Research Project, for example, is predicated on the assumption that trained observers offered no uniform objectivity. In other words, if an expressway is operating well, this quality can be detected by observing operating characteristics. When the characteristics drop below a predetermined level, action may be taken.

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 control logic of a surveillance system, or any system, is that combination of techniques and devices employed to regulate the operation of that system. The analysis shows what information is needed and where it will be obtained. Then, and only then, can the conception and design of the processing and analyzing equipment necessary to convert data into operational decisions and design warrants be described.

In research conducted during the past year by the Texas Transportation Institute on the Gulf Freeway Surveillance Project, the application of many control parameters to
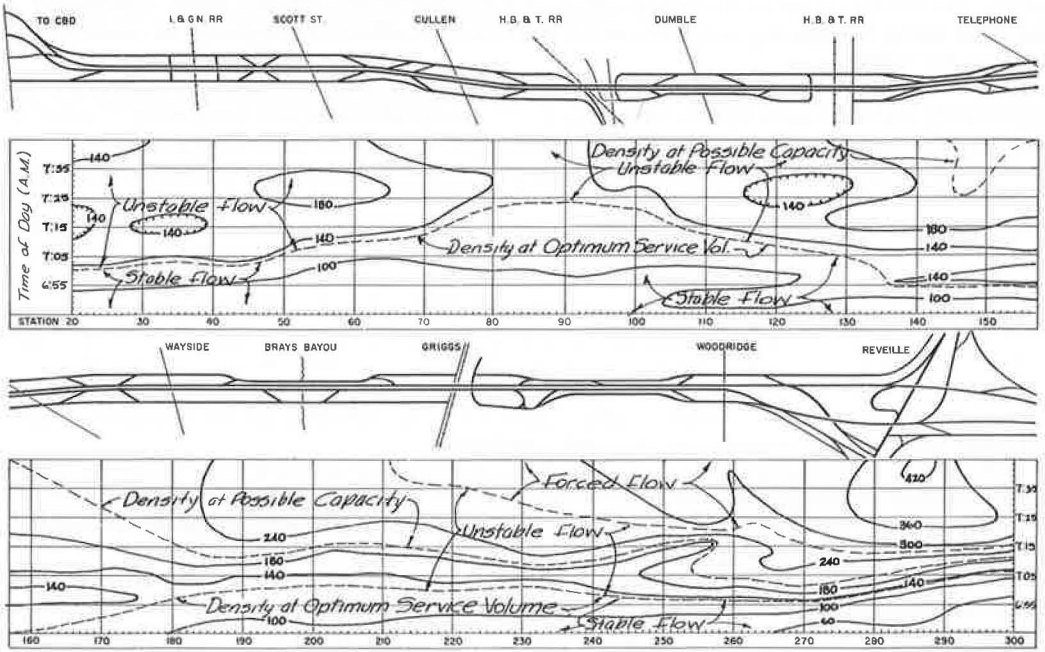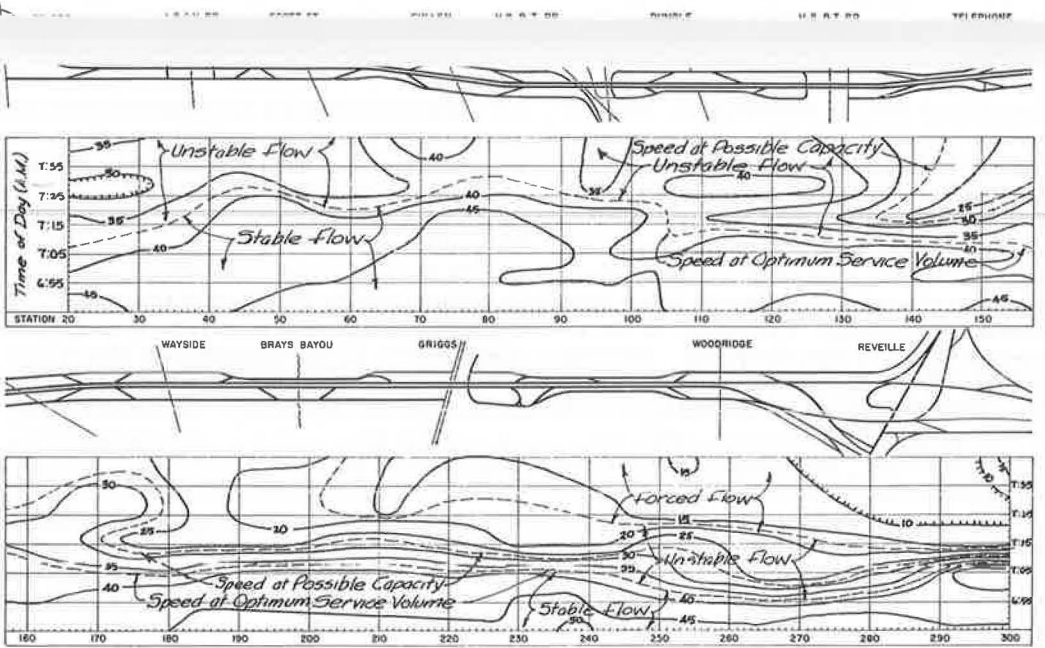
Figure 30. Density contours (3-lane total).

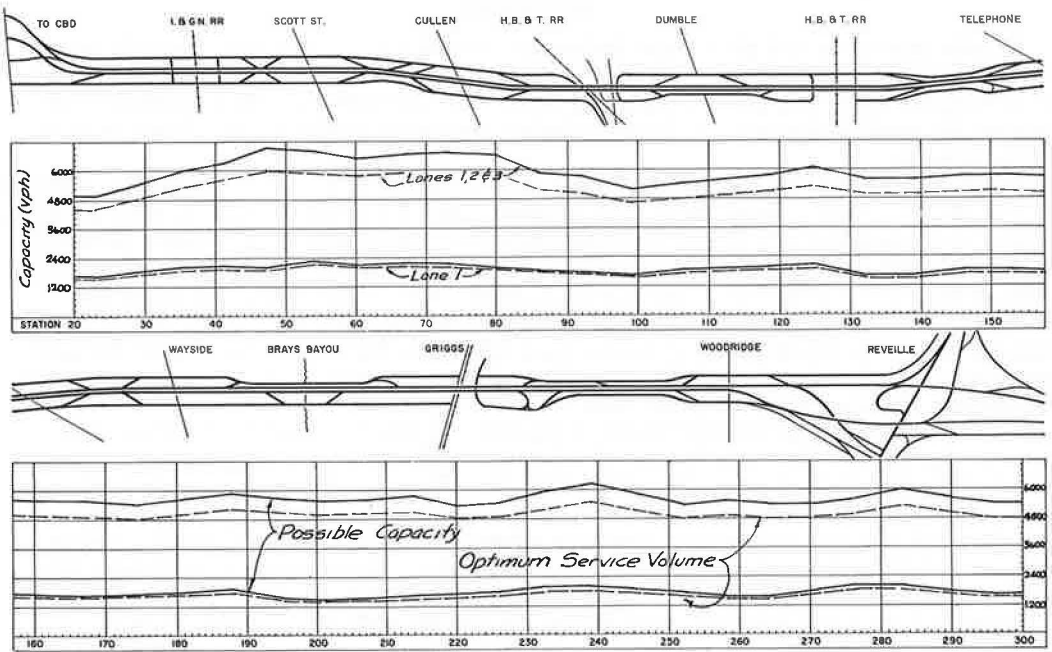Figure 31. Speed contours (total inbound traffic).

Figure 32. Capacity profiles (linear model).

the description and eventual control of freeway congestion was explored. Figures 30 and 31 illustrate the operation of the 3 inbound lanes of some six miles of the facility during the morning peak hour as obtained from time-lapse aerial photographic studies. Four control parameters, derived in the previous section, are superimposed on the contour maps: (1) the speed at possible capacity, u_m; (2) the density at possible capacity, k_m; (3) the speed at the optimum service volume, u'_m; and (4) the density at the optimum service volume, k'_m. These parameters afford a rational, quantitative means for describing the level of operation on the facility: stable flow, unstable flow and forced flow.

Figure 32 illustrates continuous profiles of the possible capacity, q_m, and the optimum service volume, q'_m, which were derived by applying the momentum-energy analogy to speed-density data taken from aerials of the facility. Thus, if stable flow is to be maintained on the facility, demand must be kept below the optimum service volume. Use of possible capacity as a basis for ramp metering or control places operation of the facility in the unstable zone of operation, and provides absolutely no safety factor against breakdowns due to statistical variability in demand.

Efforts to measure freeway operational efficiency in terms of traffic "throughput" (momentum) are obviously inconsistent with the level of service (energy) concept, because maximum throughput must necessarily be achieved with a high traffic stream density, a low traffic stream speed, and a level of operation typified by "unstable flow." On the other hand, the optimum service volume provides for speeds 33 percent higher, densities 33 percent lower, a level of operation typified by "stable flow," and with only a 10 percent reduction in flow. Actually, because there is less probability of attaining "forced flow" (congested flow inevitably accompanied by complete breakdown), the "throughput" from day to day might very well be higher because of less frequent breakdowns.

SUMMARY

Traffic operation and geometric design are essentially systematic attempts to re-solve a demand-capacity relationship for a given facility in a manner that will provide an acceptable level of service to the motorist. This report deals with urban freeway volume (demand) and capacity characteristics, and their application to freeway design and operation.

An important volume characteristic which is discussed is the distribution of demand during the peak hour. Peak rates of flow within the peak hour exceed the average hourly rate of flow on urban freeways. It is important that the peak hourly volume presently used as a criterion for design be expanded to accommodate the higher rates of flow which exist over shorter intervals within the peak hour. This is true because a condition in which the demand exceeds the capacity can extend congestion for a much longer time than just the duration of the peak flow period (see Fig. 24). In Figure 1 this variation has been related to the size of the city, thus affording a means of estimating the highest 5-min rate of flow during the peak hour on a facility.

Another vital volume characteristic considered is the lane use distribution of vehi-cles on freeways. The outside lane of a freeway will, on the average, have a lower volume than the other lanes. The most important factors influencing the percent of the total freeway traffic in the outside lane immediately upstream from an entrance ramp are the total freeway volume and entrance ramp volume (see Figs. 5 and 9). Other factors are the sequencing of entrance and exit ramps, their spacing, and their vol-umes. (See the nomographs illustrated in Figs. 6, 7 and 8.)

A second aspect of lane use distribution discussed is the relationship between trip length and outside lane utilization obtained through the use of a "lights-on" study tech-nique (Figs. 10, 11 and 12). This relationship was useful in establishing the minimum length of freeway weaving sections based on entrance and exit ramp volumes (see Fig. 26).

The second section of this report deals with freeway capacity characteristics. A theoretical approach to providing a rational relationship between capacity and level of service is formulated utilizing a hydrodynamic model and based on an energy-momentum description of level of service is suggested in the preparation of a table of Freeway Design Service Volumes (Table 3) to be used in determining freeway main lane require-ments. Fundamental volume-speed-density relationships (Fig. 21) and their measure-ments (Figs. 14, 15 and 16) on urban freeways are discussed. The importance of the contour format as a means of presenting a complete picture of performance over long sections of freeway and extended intervals of time is also described in this section.

It is significant that the first section of the body of the report deals with freeway demand and the second section with capacity. In the third section, applications are made of these demand and capacity characteristics to freeway operations (Figs. 30, 31 and 32) and to freeway design (Figs. 26, 27, 28 and 29), featuring a step-by-step design procedure.

REFERENCES

1. U. S. Bureau of Public Roads. Practical Applications of Research. Highway Capacity Manual, U. S. Govt. Print. Office, Washington, D. C., 1950.
2. Keese, Charles J., Pinnell, Charles, and McCasland, W. R. A Study of Freeway Traffic Operations. Highway Research Board Bull. 235, pp. 1-18, 1960.
3. Keese, Charles J. Improving Freeway Operations. Proc. Western Section, Inst. of Traf. Eng., 1960.
4. Design Manual for Controlled Access Highways. Texas Highway Dept., Austin, Jan. 1960. (Revised Dec. 1962.)
5. Moskowitz, Karl, and Newman, Leonard. Notes on Freeway Capacity. Highway Research Record No. 27, pp. 44-68, 1963.
6. Route 3 "Lights-On" Traffic Survey. Port of N. Y. Authority, Dev. Dept., Planning Div., New York, 1960.
7. Greensberg, H. An Analysis of Traffic Flow. Operations Research, Vol. 7, 1959.
8. Drew, Donald R. A Study of Freeway Traffic Congestion. Texas A & M Univ., Dissert. 1964.
9. Lighthill, M. J., and Whitham, G. B. On Kinematic Waves: A Theory of Traffic on Long, Crowded Roads. Proc. Roy. Soc. London, Vol. 229, 1955.
10. McCasland, W. R. Comparison of Two Techniques of Aerial Photographs for Application in Freeway Traffic Operations Studies. Highway Research Record No. 65, pp. 95-115, 1965.
11. Ryan, D. P., and Breuning, S. M. Some Fundamental Relationships of Traffic Flow on a Freeway. Highway Research Board Bull. 324, pp. 73-84, 1962.
12. May, A. D., Athol, P., Parker, W., and Rudden, J. B. Development and Evaluation of Congress Street Expressway Pilot Detection System. Highway Research Record No. 21, pp. 48-68, 1963.
13. Hess, Joseph W. Capacities and Characteristics of Ramp-Freeway-Connections. Highway Research Record No. 27, pp. 69-115, 1963.
14. Drew, Donald R. Design and Signalization of High-Type Facilities. Traf. Eng., Vol. 33, No. 10, July 1963.

Discussion

ADOLF D. MAY, JR., Director, Expressway Surveillance Project, Illinois Division of Highways—This paper contributes significantly to the existing knowledge in freeway capacity and operations. The following discussion concentrates on three aspects of the paper: (1) linear speed density relationships, (2) density as an indication of levels of service, and (3) effect of levels of service on fuel economy.

The concepts of maximizing kinetic energy or traffic stream momentum as presented in Figure 13 are based on linear speed-density relationships, and because of the importance of these concepts some confirmation of the linearity of the speed-density relationship seems appropriate to mention in this discussion. Data obtained by the Chicago Area Surveillance Project has been analyzed (15) in order to evaluate the linearity of 1-min average speed and density observations, and the results are shown in Figure 33. Table 5 is a summary comparing the characteristics of the linear speed-density relationship obtained by the authors and the discussor.

The only significant difference between the two sets of data was that Drew-Keese's jam concentration was higher, and therefore the maximum flow was higher. The r^2 value of 0.88 indicates a rather strong linear speed-density relationship in that 88 percent of the variation in speed can be explained by density using a linear relationship. The σ value of 4.4 indicates that 95 percent of the observed minute speed measurements will lie within ± 8.8 mph of the speeds obtained by the equation. The new data introduced support the use of the linear speed-density relationship, although there are indications that well-selected nonlinear equations might provide even a higher correlation.

TABLE 5

Characteristics	Drew-Keese Data	Discussor's Data
Equation:	$\mu = 60.3 \left[1 - \dfrac{k}{133.1}\right]$	$\mu = 61.1 \left[1 - \dfrac{k}{117.7}\right]$
Mean free-flow speed (u_f)	60.3	61.1
Jam concentration (k_j)	133.1	117.7
Optimum speed (U_m)	30.2	30.6
Optimum density (K_m)	66.6	58.8
Maximum flow (Q_m)	2,006	1,798
Correlation coefficient (r)	0.97	0.94
Coefficient of determination $(r)^2$	0.95	0.88
Std. dev. of regression (σ)	3.2	4.4

TABLE 6

Level of Service Zone	Zone Limits	
	Upper	Lower
A	$K_j/10$ (13, 12)	0
B, C, D	$K_j/3$ (44, 39)	$K_j/10$ (13, 12)
E_1	$K_j/2$ (67, 59)	$K_j/3$ (44, 39)
E_2	$2K_j/3$ (39, 73)	$K_j/2$ (67, 59)

Table 4 of the Drew-Keese paper presents the levels of service as established by the energy-momentum concept. The authors use the combination of speed and volume to establish limits for the various levels of service. The consideration of this concept as a guide for operating a freeway surveillance and control system, and the authors comment in their introduction that congestion can quantitatively be defined in terms of vehicular density, led the discussor to investigate the use of density as the parameter to establish limits for the various levels of service for operating sections of freeways (see Fig. 34). The resulting density values for the various levels of service are shown in Table 6. The normalized limits in terms of k_j are given as well as the actual density levels in vehicles per lane-mile based on Drew-Keese's data and the discussor's data.

For the two sets of data, a lane density of approximately 40 veh/mi is the limit between stable flow (levels of service B, C and D) and unstable flow (level of service E_1), and is the optimum density based on maximizing kinetic energy. A lane density of approximately 60 veh/mi is the limit between the two levels of service (E_1 and E_2) which represent unstable flow, and is the optimum density based on maximizing traffic stream momentum. A lane density of approximately 80 veh/mi is the limit between unstable flow (level of service E_2) and forced flow (level of service F). It is interesting to note that a density of 40 veh/mi (15 percent occupancy) is currently used as the control parameter level for initiating ramp metering on the Eisenhower Expressway. If density continues to increase, the metering rate is reduced until at a density level of 67 veh/mi (25 percent occupancy) the most restrictive ramp metering rate is selected.

In the introduction the authors indicated that vehicular operational cost is one of the factors influencing level of service. In a recent study (16) on a 3.13-mi section of the
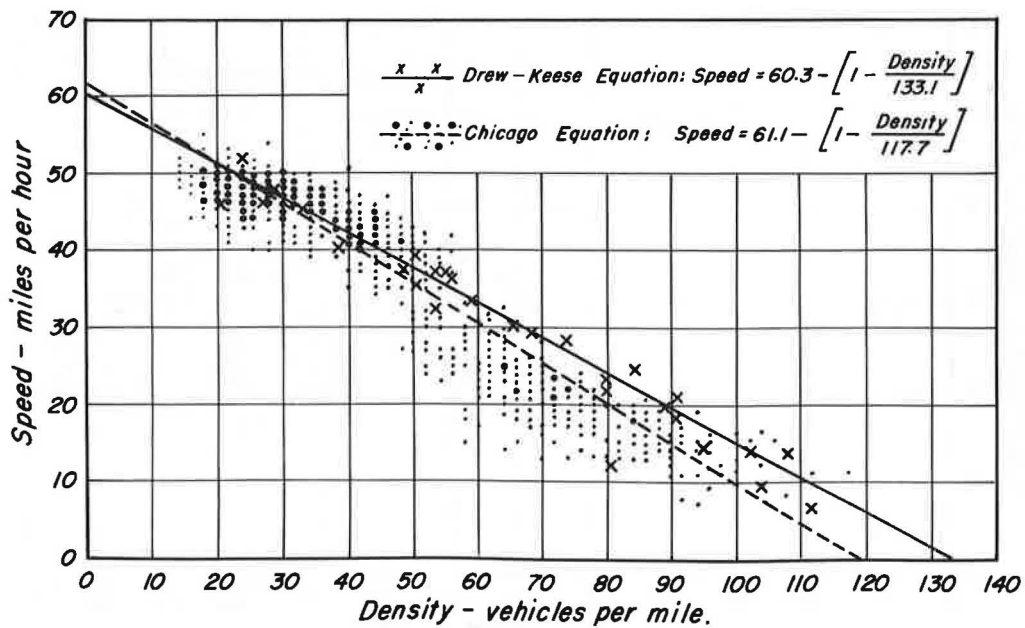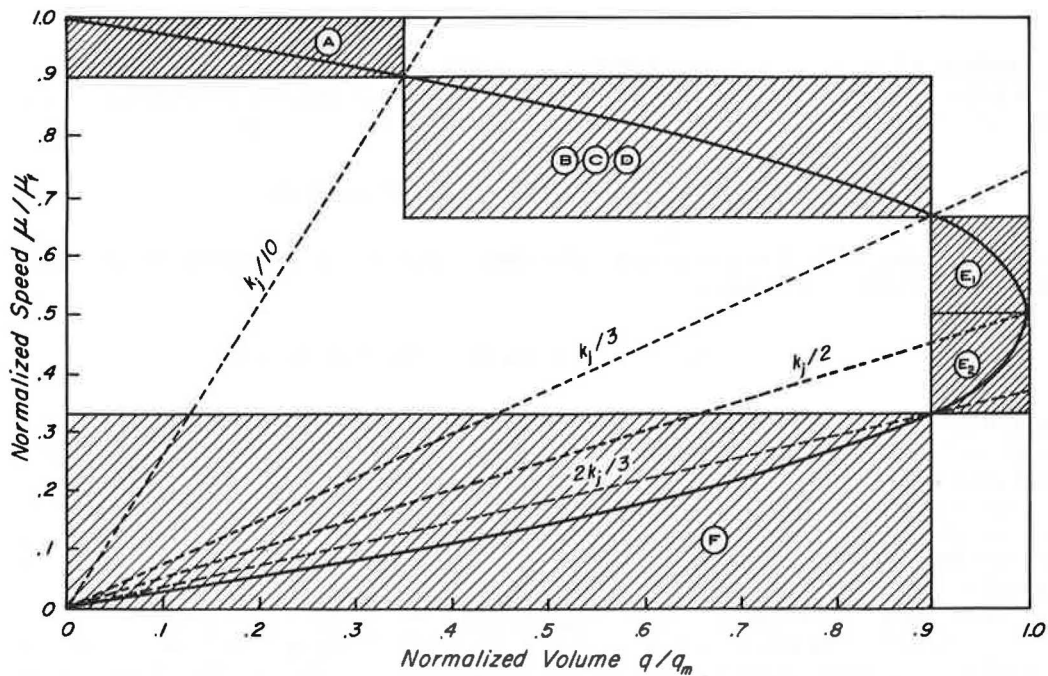
Figure 33. Speed-density relationship.

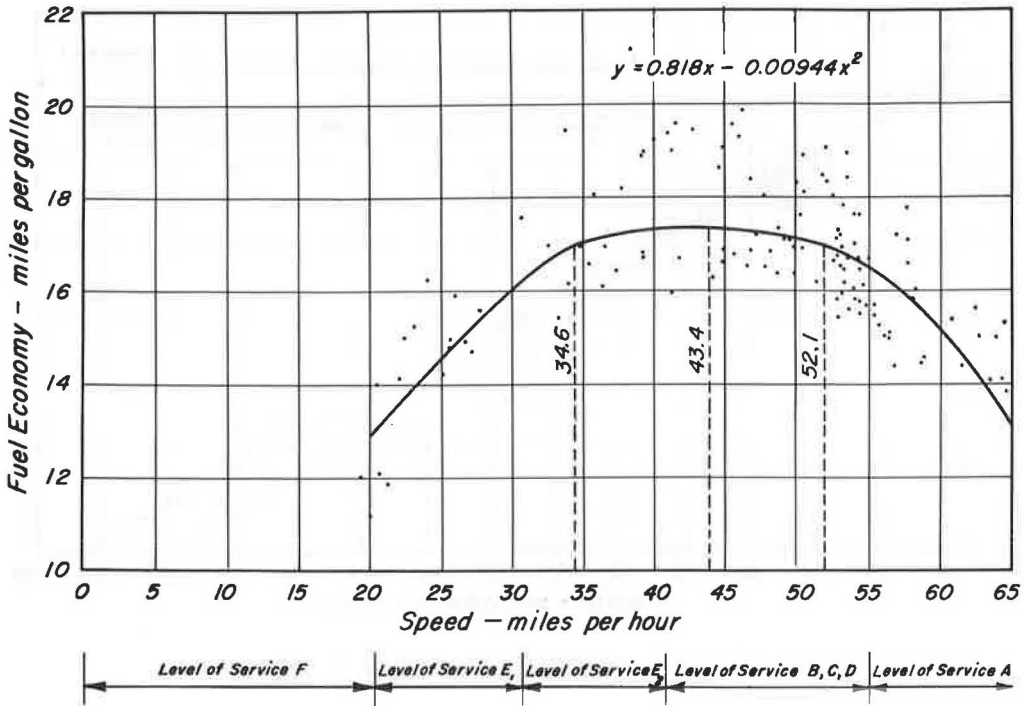Figure 34. Levels of service related to density.

Figure 35. Fuel economy vs speed.

of fuel economy to speed versus speed. The resulting equation was

$$\frac{\text{Fuel economy}}{\text{Speed}} = 0.818 - 0.00944 \,(\text{speed})$$

and an r^2 value of 0.92 was obtained. The above equation was transformed into a fuel economy versus speed equation

$$\text{Fuel economy} = 0.818 \,(\text{speed}) - 0.00944 \,(\text{speed})^2$$

and is shown in Figure 35. The highest fuel economy was obtained at a speed of 43.4 mph, and fuel economies greater than 17.0 mpg were obtained at speeds between 34.6 and 52.1 mph.

Using the density levels calculated in the earlier portion of this discussion, and the linear speed-density equation, the levels of service were determined and are shown in Figure 35. It can be seen that better fuel economy is obtained at the optimized speed based on the kinetic energy concept than the optimized speed based on the momentum concept. This would seem to further support the kinetic energy concept.

In summary, this discussion of the paper: (1) supports the authors' use of a linear speed-density relationship as a basis for the kinetic energy and momentum concepts, (2) indicates that the kinetic energy and momentum concepts can be used as guides for operating existing expressways, and (3) presents evidence that fuel economy is higher when kinetic energy is maximized than when momentum is maximized.

REFERENCES

15. Ellis, Raymond H. Analysis of Linear Relationships in Speed-Density and Speed-Occupancy Curves. Unpub. term paper, Northwestern Univ., Dec. 1964.
16. Oberle, George F. Travel Time vs Fuel Consumption on the Eisenhower Expressway. Unpub. term paper, Northwestern Univ., Dec. 1964.

KARL MOSKOWITZ, Assistant Traffic Engineer, California Division of Highways, Sacramento—The authors have made a significant contribution, not only by introducing a new concept, the "energy analogy," but in collecting and collating a great deal of previous work done by them and others.

Each mile of a freeway flowing at a rate of 5,400 vph (one direction) at 15 mph is providing 5,400 vehicle-miles of transportation per hour, and each mile of another freeway flowing at a rate of 5,400 vph at 50 mph is also providing 5,400 vehicle-miles of transportation per hour. The quantity of service seems to be the same, yet the quality is obviously different, and if a person wants to describe quality in terms of volume (or flow) he cannot do it.

The product of flow and speed suggested by the authors may be helpful in providing a number that would overcome this anomaly, but until people get used to it, the usefulness might be limited. Some of us might still find it aids our understanding to use more than one parameter to describe flow; i.e., to keep flow and speed separate.

Drew and Keese reproduce in Figure 24 (b) a graph that this reviewer has found very useful in understanding the relation of flow and delay at (and upstream of) a bottleneck, which is a very common occurrence on freeways and all other traffic facilities, including nonmotor vehicle facilities. Many investigators get involved with speed-volume curves (such as the right-hand side of Fig. 22), or flow-concentration curves (such as Fig. 23), which are based on observations that may or may not be upstream of a bottleneck (the points on the curves representing higher speeds are seldom upstream of bottlenecks), and they do not always realize that inferences drawn from the speed-volume curve or q - k curve may not be compatible with the facts as shown in Figure 24 (b).

One of the most useful things about Figure 24 (b) (reproduced herein with added notes and dimensions as Figure 36) is that it shows flow as a rate (i.e., a mathematical slope) which is time-related, instead of as a point which is not time-related. Since by definition flow is a rate (i.e., n/t where n is number of vehicles and t is time), Figure 36 is much easier for an engineer to understand. Engineers have great difficulty discerning the slope of a point, or the meaning of a slope represented by a point on a curve, or worse yet, a point which might be anywhere within an envelope represented by a curve.

Figure 36 represents cars going through a bottleneck as a function of time during a peak period. The slope of the upper curve at any point is the arrival rate and the slope of the lower curve is the service rate, or the actual flow. It can be seen that at t_a, the arrival rate begins to exceed the service rate. This might be 4:45 p.m. in a city like Sacramento, Calif., and many other cities in this country. At t_b, which might be 5:15 p.m., the arrival rate is equal to the service rate, and from t_b until t_d, the arrival rate is less than the service rate. At t_d, which might be 5:45 p.m., the congestion is over.
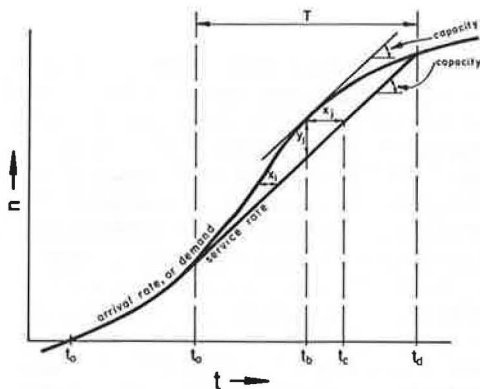

Figure 36. Relation between flow and delay.

Some of our colleagues in the traffic assignment field use speed-volume curves, such as Figure 22 in Drew and Keese, to predict journey times as a function of demand. Yet it can be seen that the delay to an individual vehicle, x_j in Figure 36, is much greater at t_c (when vehicle j goes through the bottleneck) than it is at t_a, or at any other time of day between t_a and t_b, although the predicted demand is less at t_c than it is at any time between t_a and t_b. Using a speed-volume curve, a person would imply that the speed was close to 0, at $t_a + \Delta t$, just after the queue begins to form, because the volume/capacity ratio at $t_a + \Delta t$ is very high and is considerably greater than 1.0, whereas in fact the delay is very small, as can be seen by looking at the graph. The same person might infer that subsequent to t_b, speed would be very high because the demand/capacity ratio is less than 1, and yet a glance at Figure 36 shows that the j th vehicle, arriving at t_b, suffers the greatest delay of any vehicle in the universe we are here concerned with.

The actual flow rate stays nearly constant from t_a to t_d, but the journey time of any individual vehicle varies considerably, and the queue length (y_j) also varies. The spot speed immediately upstream of the bottleneck will be variable, from 0 up, on a moment-to-moment basis, and this is what Drew and Keese as well as other investigators call "forced flow." The spot speed in the bottleneck or downstream of it will be more uniform during time and will vary depending on how far downstream the observation is made, because cars will generally be accelerating. If the bottleneck is a merging area on a freeway, the volume or flow upstream of the bottleneck is less than the volume in the bottleneck by the amount added at the ramp; i.e., the ramp volume plus the freeway volume is equal to the downstream volume.

A third axis, namely distance, can be added to Figure 36 and the resulting graph would be a surface that truly portrays traffic flow in a single file, because speed would show. The slope of this surface in one plane would be flow (this is the plane shown in Figure 36); in the second plane would be speed and in the third plane, density. An examination of this surface during periods of forced flow would show that speed-volume-density relationships are far too complex to describe on a rate vs rate graph (i.e., flow vs concentration or flow vs speed).

But the loss of time, or delay, owing to this speed reduction is so small in comparison with the delay that occurs when the arrival rate exceeds capacity, that it really is not very important to know much about in urban freeway operational problems. The reduction in level of service, or subjective reactions of drivers, in this range is related more to nervous tension and necessity for continuous concentration on the driving task than it is to speed or delay.

J. W. HESS, U. S. Bureau of Public Roads—This report deals with volume characteristics and capacity characteristics which affect traffic operation, and thus the level of service, on freeways. Applications of the developed design tools are given. I am very impressed by the comprehensiveness and thoroughness of this report. The Texas Transportation Institute engineers working on the HPR "level of service" project in Texas have turned out a remarkable quantity and quality of research in this past year. I understand the material in this report has been under development for several years.

The subject of peaking is studied as related to the population of the city or urban area. For instance, the level of service D for a 6-lane freeway has a maximum 5-min volume rate of 5,250 when the design is for 4,800 vph for a city of 5 million metropolitan area population. A higher maximum 5-min rate would be found in other cities of smaller population (for instance, a 5,600-vph rate for a city of 1,000,000). My own studies have shown the importance of population in high volume merging. The population of a large city represents a steadier demand, greater driver experience, and more aggressive driver behavior. The latter should not be misconstrued as meaning drivers in large cities are more rude. I feel that they are actually more polite in free-

way driving because they do not insist on their right-of-way and will help open a gap in the stream at merging areas.

Perhaps the indications are that smaller cities should design at a higher level of service because of greater peaking rates and poorer traffic performance. An alternative is to use lower-service volumes. Traffic performance should improve as the freeway systems are developed, but peaking will probably not change much.

The authors had trouble relating factors which might have an effect on the maximum rate of flow occurring in a peak hour. Such factors were physical size of the metropolitan area, distance of the study site from the CBD or principal traffic generator, the actual size of the peak period, itself, etc. Other engineers who have tackeled this subject have encountered the same trouble. Perhaps, the main reason for the lack of correlation is the incompleteness of traffic studies as far as a simultaneous study of traffic movement over the complete area of influence.

The design charts presented for use in determining lane 1 volumes at entrance ramp locations gave answers very similar to those obtained by my own nomographs. The only differences of any size were at successive on-ramps for 6-lane freeways where the authors' lane 1 volumes are 150-200 vph lower, and lane 1 volumes on 8-lane freeways where my volumes are lower. I have one specific question—does ± 8 percent connote ± 8 percentage points or ± 8 percent?

I was especially interested in the results of the "lights on" study because our weaving study crew made four such studies in Detroit in September 1963. The authors' study instructed the drivers to keep their lights on for 20 minutes, whereas our studies instructed drivers to keep their lights on for 2 miles. Does this mean that there is more congestion in Texas or do Texans tend to be "clockwatchers?" Seriously, we were studying the same thing—the lane usage of ramp vehicles downstream after merging. Our manpower allowed only a distance of 5,000-6,000 feet downstream for study purposes, whereas in the authors' study the distance was over 10,000 feet. The curves were very similar, with the authors' study curve falling along the same line as our uppermost curve. One difference in the respective studies is that the authors' results showed more vehicles remaining in lane 1 (after merging) at volumes above 3,000 vph, whereas our results showed less staying in lane 1 at the higher volume rates. Although it is more difficult to vacate lane 1 at higher volume rates, it also appears to be more desirable if the driver has the motivation of improving his position in the traffic stream. Northwestern University has recently done some similar work for left-hand ramps, both entrance and exit. The study method used by them utilized a series of time-lapse cameras mounted on overpasses. All of this information is very important in determining the operation to be expected for a given design, and it may well give us our best traffic evaluation for interchange spacing. Lane usage by prospective off-ramp vehicles upstream from their exit ramp is a vital cog still missing in this evaluation of interchange effects. I can also see an important use of this information in simulation of the traffic stream.

In the "Capacity Characteristics" section of the report, the following can be quoted— "Both capacity and level of service are functions of the physical features of the highway facility and the interaction of vehicles in the traffic stream. The distinction is this: A given lane or roadway may provide a wide range of levels of service, but only one possible capacity...." I wonder about the "one possible capacity." I believe that the interaction of vehicles can change, and with it, the possible capacity, for instance, at a ramp merging area where the ratio of lane 1 to the ramp might change considerably.

The energy-momentum analogy developed in this report to provide a quantitative approach to the level of service-capacity relationship should be studied by everyone interested in the developments behind the new Capacity Manual approach. This is not to say that the analogy led to the Capacity Manual approach—only that it helps explain some of the speed-volume-density relationships.

The last part of this section on "Capacity Characteristics" presents a very interesting "Probabilities of Types of Flow" table. It would be especially useful if eventually these probabilities could be related to the design standards or features of the freeway. The actual freeway design, especially at a bottleneck area, as well as the metropolitan population which is used should be a strong factor.

In the "Applications" section of this report, the authors discuss critical sections along the freeway such as merging and weaving areas. Though no procedures are given in this report, intervehicular conflict upstream from an exit ramp is also mentioned as a problem. Ramp vehicles tend to decelerate on the through freeway lane and this is a section of vehicle concentration. This problem is being considered in the new Highway Capacity Manual by the assignment of a service volume level to the lane 1 immediately upstream of the exit ramp. It is interesting to note that in the report's design problem of a length of freeway, application of the exit ramp equations points up at least three locations where the 1,700-vph service volume is exceeded, if only slightly. However, service volumes for diverges could logically be slightly higher than those for merges.

The "Graph for Determination of Minimum Length of Weaving Section to Meet the Design Level of Service" could be very useful. I note, by the slope of the weaving distance lines, that entrance ramp volumes are considered more critical than exit ramp volumes—this lends credence to my belief that lane 1 volume upstream from an exit ramp should have a higher service volume (perhaps 100 vph) than a merge.

Finally, the authors mention the gap acceptance method of evaluating entrance ramp capacity. I certainly agree that more research is needed in this area to cover, for instance, aside from geometrics, the effect of intensity of light, inclement weather, driver experience, etc. The measurement of gap arrivals and the metering of ramp vehicles to fill these gaps is to become of more importance. In this vein, we can no longer ignore freeway operation during more or less adverse conditions of weather. I guess we literally will have to get our feet wet in the near future if we are to take an all-conditions approach to freeway operation.

DONALD R. DREW and CHARLES J. KEESE, Closure—The authors are appreciative of the excellent reviews by Dr. A. D. May, Mr. Karl Moskowitz and Mr. Joseph W.

tics and operation (May), freeway capacity (Moskowitz), and ramp capacity (Hess), because these gentlemen have made particularly important contributions in these respective areas.

It is encouraging to note that Dr. May's data based on lane occupancy tend to substantiate the model developed in this paper, and that the optimum speed based on fuel consumption is in agreement with the energy-momentum concept for defining level of service. Vehicular operating cost is one of the factors considered in the level of service concept. The fuel data were collected by Donald G. Capelle for use in his Doctoral Dissertation at Texas A & M University. His studies were conducted on the Eisenhower Expressway in cooperation with Dr. May and will be reported in detail at a later date.

We concur with Mr. Moskowitz regarding the utility of the relationship between flow and delay illustrated in Figure 36 reproduced, in part, in our Figure 24. However, the addition of the top half of Figure 24 serves to dramatize the fact that the duration of congestion can be appreciably longer than the duration in which demand exceeds capacity.

It is felt that the ramp capacity analyses conducted by Mr. Hess substantiate the procedure developed in this paper which was based on several highly selective locations. Mr. Hess has analyzed a great deal more data covering numerous diverse types of facilities and conditions. The fact that his findings generally substantiate the theoretical approach and practical application developed in this paper indicates that this complex operational characteristic can be analyzed and used with reasonable confidence in the solution of freeway design and operation problems.

These reviews lend support to the philosophy that the level of service concept (or quality of traffic service) can be quantified for use in design of new facilities and in the application of control measures to improve operation on existing facilities.

The authors again express appreciation to Messrs. Ginn, Haynes, Jones and Lipscomb who conducted the individual studies and to the various members of the staff