Theoretical Analysis of
Catastrophes in Traffic

UCI-ITS-TS-WP-98-9

Seungmin Kang
R. Jayakrishnan

Department of Civil and Environmental Engineering and
Institute of Transportation Studies
University of California, Irvine
rjayakri@uci.edu

July 1998

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu
Abstract

Catastrophe theory has been applied to traffic analysis to account for discontinuous behavior of freeway operations. However, the reason for catastrophe have not been developed, and there have not been any fundamental examination of whether catastrophe surfaces of the kinds proposed in the past are reasonable or even possible in traffic. This paper examines the short-time-period (30 second or less) dynamics of traffic flow based on fundamental considerations without making assumptions on the traffic relationships. The results show that the effect of vehicle length variations on occupancy calculations is one that could be leading to erroneous conclusions and that properly calculated occupancies for double loop detector stations show some interesting conclusions, that independent variation of the three variables (flow, occupancy and speed) is not possible, and that catastrophe folds showing a vertical drop of speed at constant flow and occupancy can not and does not exist. Finally observed traffic data for freeway operation does not confirm to a catastrophe surface, but a continuous fold-less surface.

Key words: traffic discontinuity, catastrophe theory, continuous traffic surface
Introduction

Over the years, a number of traffic stream models have been proposed for modeling freeway operations based on car-following theory and hydrodynamic principles. These models failed to describe the occurrence of discontinuous phenomena in the three basic traffic variables, speed, flow and density. This discontinuity of freeway operations has been identified by Edie [1] and Ceder [2] and their approach, called the two-regime model, proposed the use of different models for free-flow conditions and for congested flow conditions. The critical difficulty in the two-regime model is determining the breakpoint between regimes. On the other hand, research has been attempted to estimate continuous function accounting for this discontinuity, based on catastrophe theory, a theory about singularities which deals with the properties of discontinuities directly. Navin [3] proposed a cusp catastrophe traffic model to explain sudden changes in traffic stream behavior and later Hall and others [4, 5, 6] claimed that traffic stream behavior fits a cusp catastrophe surface.

When a cusp catastrophe is applied, a three-dimensional surface is drawn with speed, flow and occupancy on the coordinate axes. While the fundamental diagram shows a smooth three dimensional curve for the three variables, the three dimensional catastrophe surface shows that speed can change drastically under certain conditions even when the flow and occupancy show smooth variation. This drastic change of speed at unchanged flow and occupancy is considered to be of significance in automatic incident detection [6]. The previous analyses have attempted to calibrate such catastrophe surface models using real traffic flow data, utilizing data transformations [7], and catastrophe ideas have been used for the estimation of speeds from
inductance loop freeway flow and occupancy data [8]. However, the reason for catastrophe have not been developed, and there have not been any fundamental examination of whether catastrophe surfaces of the kinds proposed in the past are reasonable or even possible in traffic. This has possibly caused some confusion and misconceptions on what kind of catastrophe is reasonable for traffic.

This paper examines the short time-period dynamics of traffic flow based on fundamental considerations without making assumptions on the traffic relationships, and develops certain results on the traffic variations. The analysis is simple and uses intuitive arguments, but develops several significant insights. The results show that the effect of vehicle length variations on occupancy calculations is one that could be leading to erroneous conclusions and that properly calculated occupancies for double loop detector stations show some interesting conclusions, that independent variation of the three variables including flow, occupancy and speed is not possible, and that catastrophe folds showing a vertical drop of speed at constant flow and occupancy (as often shown by catastrophes in other dynamic systems) can not and does not exist. 30-second data from a real-world freeway is used to verify such results. The results may perhaps even lead to the conclusion that catastrophe theory may not be fully appropriate for traffic, a sentiment that some other traffic researchers may already hold.

**Effect of Point Measurement**
The most commonly used traffic detector is a presence-type inductance loop. Traffic variables such as flow, occupancy and speed are collected through the traffic detectors and used for measuring the operational conditions in traffic. Flow is the number of vehicles crossing a point during the aggregation interval, occupancy is the fraction of time the detector is covered by vehicles (usually used as a surrogate variable for density), and speed is an average of the speeds of vehicles crossing the detector. Space mean speed is defined based on vehicle passage times over a fixed detection length and a harmonic averaging, as:

$$u_s = \frac{n}{\sum_{i=1}^{n} 1/u_i} = \frac{n}{\sum_{i=1}^{n} t_i/d_i}$$

(1)

where, $u_s$ is a space mean speed, $u_i$ is the speed for a vehicle $i$, $n$ is the number of vehicles crossing during the aggregation interval, and $t_i$ is the time taken for vehicle $i$ to cross a fixed detection length $d_i$. A single loop detector, however, measures only the time period over which an inductance pulse exists, which is the time when any part of the vehicle is over the detection zone. This $t_i$ is the time for the vehicle to move a distance equal to the detection zone length plus the vehicle length. If an average vehicle length is assumed, for a single loop detector, Eq. (1) becomes:

$$u_s = \frac{n}{\sum_{i=1}^{n} 1/u_i} = \frac{n}{\sum_{i=1}^{n} t_i/d_i} = \frac{nd_i}{\sum_{i=1}^{n} t_i}$$

(2)
where, $d_e$ is the assumed effective detection length which is equal to detection zone length plus the average vehicle length. Occupancy is defined as:

$$o = \frac{\sum t_i}{T}$$

(3)

where, $o$ is occupancy and $T$ is aggregation interval. Since $n/T$ is the measured flow over the aggregation interval $T$, the Eq. (2) can be rewritten as:

$$u_s = qd_e / o$$

(4)

However, space mean speeds computed from Eq. (2) or Eq. (4) are not same as the theoretical speeds from Eq. (1), essentially unless all the vehicles have identical lengths.

One of the most obvious examples of occupancy leading to an incorrect interpretation of traffic is in the application of catastrophe theory to traffic analysis. The previous papers on catastrophe theory in traffic flow describe the existence of a vertical drop of speed at a given flow and occupancy [9], the vertical fold being in accordance with the so-called Maxwell convention of catastrophe theory. If all the vehicle lengths were the same, we can see from Eq. (4) that such a vertical drop of speed is not possible (due to the strict relationship between occupancy, flow and speed). This implies that if a vertical drop is
observed, it is not because of variation of traffic conditions but because of variation of vehicle length. The Maxwell convention has not been proved with real traffic data, either.

While there is no easy way to get around the problem of varying vehicle lengths at a single detector station, it is indeed possible to find occupancies not affected by vehicle lengths at a double-loop detector station (called a "speed-trap"). This is based on the time it takes for any given vehicle to move over a fixed distance between the upstream edges of the two detection zones. Space mean speeds can then be found as:

\[
    u_s = \frac{n}{\sum_{i=1}^{n} 1/u_i} = \frac{n}{\sum_{i=1}^{n} (t_{on})_{i}} / d_i = \frac{nd_i}{\sum_{i=1}^{n} (t_{on})_{i}} = \frac{qd_i}{\sum_{i=1}^{n} (t_{on})_{i} / T} = \frac{qd_i}{o} \tag{5}
\]

where \((t_{on})_{i}\) is the difference in turn-on times of the inductance pulses which is available (i.e., times when a vehicle \(i\) reaches the edges of the detection zones of two closely spaced detectors), \(d_i\) is the fixed distance from the upstream edge of the upstream detection zone to the upstream edge of the downstream detection zone, and \(o\) is the occupancy. We are proposing an occupancy here for a double-loop detector station, defined as the fraction of time when vehicles are over the double-loop detector station, as in Eq. (5). Note that no assumptions on vehicle lengths have been made and the detection length \(d_i\) essentially has a fixed value at a detection location. We can easily see that, again, there exists a strict relationship among flow, occupancy and speed. In other words, given flow and occupancy, a unique space mean speed is automatically determined by Eq. (5). This strict
relationship between these three variables also indicates that they form a continuous plane on a three dimensional space of these variables (see Figure 1).

Comparison of Characteristics of the Continuous and Catastrophe Surfaces

The continuous traffic surface in Figure 1 shows a few contradictions with catastrophe surfaces. Figure 1 shows exactly identical behavior between flow, occupancy and speed as in the three-dimensional surface shown in [10], on which the fundamental diagram (three dimensional curves) is placed. Figure 1 does not contain any kind of folds unlike cusp catastrophes. From the collected traffic data we examined, we do not see the traffic variables to follow the proposed cusp catastrophe function, either. While we show no restriction that the detected traffic variables should be on a fixed curve, we do see that they have to be on a surface as in Figure 1.

In fact, it is possible that the data in the vicinity of the assumed “folds” in previous research were scant, and the assumption was never examined closely with sufficient data confirming a fold. This is perhaps a case where observed data fell close to the assumed surface function but were mostly all away from the fold.

The traffic surface in Figure 1 is a unique surface, and does not depend on the length of time interval and location of traffic data collection. The functional form of the surface is solely dependent on the value of a conversion factor, \( d \) in Figure 1, which is a factor that combines effective detection zone, vehicle length, and a unit conversion. For longer time
intervals, say 5 minutes interval, or different locations, catastrophe theory model does not always hold consistent and acceptable goodness of fit measures in understanding gaps in variation of freeway traffic behavior that is frequently observed [6]. Traffic variables are required to be transformed for catastrophe theory models. So depending on the value of parameters related to the transformation, the performance of the catastrophe theory model varies [7].

Figure 1 also shows that vertical change of single variable at a given values of flow and occupancy is not possible because there exists a restrict relationship among traffic variables, and given values of two variables, the value of the other variable is automatically determined by Eq. (5). However, a continuous change in one variable is possible with the change of another variable while the third is at a constant value. Though fundamental does not allow this behavior, it does occur in traffic dynamics over short periods, but it is still not a change as in a vertical catastrophe. From Figure 1 and Eq. (5), we can see that the degree of vertical change of speed can be steeper as flow rates decrease at high speed and low occupancy conditions. This phenomena support Hall’s model displaying discontinuous behavior at low flow rate while Navin’s model displays discontinuous behavior at or near capacity. Note that even if Hall’s description of discontinuous behavior is supported by the continuous traffic surface, the discontinuous behavior occur on a fold-less continuous surface, not on the catastrophe surface.

**Empirical Analysis of Variations of Flow, Occupancy and Speed**
In this section, empirical analysis is performed to support our claims with observed traffic data through two closely spaced loop detectors. For our analysis, we use traffic data aggregated every 30 seconds from I-880 south of Oakland, California, on March 5, 1993 [11]. Time series of occupancy, flow, and speed from 5:30 - 10:00 AM are plotted in Figure 2. Traffic was congested for about an hour from 7:30 AM to 8:30 AM, and was free-flow for the rest of the period.

Our definition of occupancy is different than in the normal traffic engineering practice, where even at double loop locations, the occupancy is calculated from one of the loops which we believe is improper. Some comments are warranted on how much of effect vehicle lengths have in giving incorrect estimates of occupancy. Vehicle lengths are often much longer than the detector length itself. In the USA, the most common traffic detector is about 1.8 m (6ft) long and the length of common passenger car is usually longer than 3.7m (12ft). For periods as low as 10 to 30 seconds, when only a handful of vehicles are part of the aggregation, the occupancy calculated could have significant errors. It is unclear whether any of the earlier catastrophe model research conclusions were caused by these errors, but we have eliminated the errors using the proper definition of occupancy.

On I-880, detectors are about 4.3m (14ft) apart and the length of each detector is about 1.8m (6ft). Therefore $d_i$ in Eq. (5) is about 6.1m (20ft) long. By Eq. (5), we can compute the occupancy from flow and speed through double loop detectors. The computed occupancy is the fraction of time taken by vehicles crossing the fixed distance between
upstream edges of double loop detectors which is about 6.1m (20ft). By comparing the computed occupancy with the observed occupancy, we can see the effect of vehicle length because the observed occupancy varies with the change of vehicle length. The difference between the observed and computed occupancies is plotted in Figure 3 and Figure 4. Lane 1 is the one assigned exclusively for high occupancy vehicles (HOV), and lane 5 is the rightmost lane. The magnitude and ratio of the difference between the observed and computed occupancies on lane 1 is smaller than lane 5. Especially the difference in lane 5 is mostly positive which means that the observed occupancy is overestimated by longer vehicles. This is to perhaps be expected because fast-moving passenger cars travel on the lane close to median, and trucks usually do not travel on the HOV lane. These figures indicate that the conventionally observed occupancy is possibly overestimated due to long vehicles.

For the above reason, traffic analysis based on observed occupancy could be misleading while estimating the traffic condition. For instance, Persaud and Hall [6] showed a plot of flow-occupancy data under an incident condition [Figure 6 in 6]. From the flow-occupancy plot if the occupancy is the one in Eq. (5), we can estimate speed at given flow and occupancy using a conversion factor. And we should be able to point the location of data showing higher speed by finding the points with higher slope for the chord connecting them to the origin, in the flow-occupancy plot. However, they utilized the conventional occupancy for analyzing changes in flow, occupancy and speed under incident conditions, and Fig. 6 in Persaud and Hall [6] included several points showing smaller chord-slope
than the other slower speed points, which is a contradicting behavior indicating a problem possibly resulting from the occupancy definition used.

Occupancy as currently observed is not consistent with the condition of the traffic, causing erroneous observations of "catastrophe drops". Nevertheless occupancy is the most common traffic variable used for traffic management including incident detection and freeway ramp control, and a correct definition of occupancy and some insights into what kind of reasonable short-term variations are possible, are important.

As we expected from Eq. (5), the plots of flow, occupancy and speed our analysis using the I-880 does not show the Maxwell convention, the vertical drops of speed in the speed-flow-occupancy space (see Figure 5). We do, however, see inclined drops of speed (not vertical) with increase of occupancy and the decrease of flow at the onset of congestion.

Conclusion

Occupancy is one of the common variables used for traffic analysis. However, due to the basic characteristics of the definition used, sometimes the occupancy at a detector does not indicate the appropriate traffic information, and the misuse of occupancy may lead to improper traffic analysis. A case in point may be the application of catastrophe theory in traffic analysis.
Catastrophe theory model for traffic analysis is not based on theoretical reasons, but solely based on empirical observation. The justification for application of catastrophe theory to traffic analysis depends on researchers’ subjective conclusion that a relation exists for the traffic variables, flow, occupancy and speed as in cusp catastrophe models. Such a relation does not hold when we examine the definitions of these traffic variables.

From the theoretical equation (Eq. (5)) used to compute speed from flow and occupancy through double loop detectors, we can see that traffic variable form a unique and continuous surface, and catastrophe theory models do not appear applicable here.
References


Figure 1 The Continuous Traffic Surface

\[ u = q^* d / c \]
Figure 2 Plots of Occupancy, Flow, and Speed
Figure 3 Difference Between the Observed and Computed Occupancies in Magnitude

Lane 1

Lane 5
Figure 4 Difference Between the Observed and Computed Occupancies in Ratio

Lane 1

Lane 5
Figure 5 Plots of Traffic Variables on the Continuous Traffic Surface

\[ u = q^* d/\alpha \]