MOVING BOTTLENECKS: A THEORY GROUNDED ON EXPERIMENTAL OBSERVATION

Juan Carlos Muñoz[1] and Carlos F. Daganzo, Inst. of Transportation. Studies, University of California, Berkeley, California, USA

ABSTRACT

This paper presents the most complete picture yet of moving bottlenecks on freeways, including experimental observations and a theory. The experimental observations include the "fingerprint" of a moving bottleneck on a series of loop detectors, and a set of controlled experiments in which moving bottlenecks were artificially introduced in the traffic stream. The paper also contrasts this evidence with current theories and describes a new one that is consistent with the data.

High-resolution oblique plots of loop detector data from freeway I-880 in Oakland (California) are used to analyze the aforementioned fingerprint. They clearly display the presence of the bottleneck and its evolution in time and space, including the precise location in space-time where it appeared. The data also reveal a fleeting but real change in the drivers' car-following attitude shortly after the bottleneck's appearance.

The controlled experiments reveal that the flow downstream of the bottleneck increases with the speed of the bottleneck when the bottleneck holds back a queue—in contradiction with two previous theories (Gazis and Herman, 1992, and Newell, 1993).

The new theory includes these as special cases. It treats the moving bottleneck as a boundary condition that can be integrated with kinematic wave (KW) theory and also with variants of this theory that account for multiple vehicle types and changes in driver psychology. The empirical evidence suggests that the lengths of queues upstream of moving bottlenecks and the ensuing vehicle delays can now be predicted with good accuracy.

---

[1] Instructor at the Pontificia Universidad Católica de Chile, Ph.D. student at U.C. Berkeley.

1. INTRODUCTION

This paper is concerned with the effects of slow-moving obstructions on a traffic stream. The main determinants of these effects are (i) the maximum rate at which a queue held back by the moving bottleneck discharges, and (ii) the density of the queue for a given discharge rate. The former is important because it determines whether a queue can form for a given flow and bottleneck speed and, if it does, whether the queue will grow or dissipate as traffic conditions change. The latter is important because it establishes the spatial length of the queue. It is found from field observations that the queue density behind a moving bottleneck closely matches the vehicular density of kinematic wave (KW) theory, in agreement with existing theories: Gazis and Herman (1992) and Newell (1993). Additionally, a set of controlled experiments reveal a reproducible relation between bottleneck speed and queue discharge rate, but this relation contradicts the theories. Fortunately, the experimental data strongly suggest the form of a complete and empirically sound theory of queue dynamics around moving obstructions.

The uncontrolled observations were made possible by a new way of processing loop detector data (oblique-plots) that can reveal information at the platoon level, and sometimes even at the individual vehicle level. For a tutorial on this processing technique see Muñoz and Daganzo (2000b). In our case, it was possible to see that drivers in the queue initially followed the obstruction very closely, exceeding "capacity" flow, and shortly thereafter at the equilibrium spacings. This change was observed *among the same set of drivers*. As such, it is the first direct evidence that drivers can change their following behavior in response to external stimuli. A similar effect has been proposed in Daganzo (1999a and 1999b) as an explanation of several puzzling traffic phenomena.

The paper has been organized as follows. Section 2 describes all moving bottleneck theories and the proposed one. The experimental basis for the new theory is then presented in Sections 3 and 4. Section 3 describes how a steadily moving bottleneck on Freeway I-880 in Oakland (California) was identified and characterized from loop detector data, including the behavior of the traffic stream around it. Section 4 presents two batteries of floating-vehicle experiments involving a slow test-vehicle that revealed a relation between maximum passing rate and bottleneck speed for one-lane obstructions in two- and three-lane freeways. Section 5 discusses the results. Finally, a possible extension of the theory is detailed in Appendix A.

2. THE MOVING BOTTLENECK THEORY REVISITED

Recall that the rate, q_r, at which cars pass an observer that moves with speed v_b when traffic is in a steady flow-density state (q,k) is given by the flow conservation formula:

$$q_r = q - kv_b \qquad (1)$$

If v_b and (q,k) are given then q_r is the vertical separation between the corresponding steady-state point on the (k,q)-plane and the dotted ray shown in Fig. 1a.

Equation (1) applies in particular to an observer that either trails or precedes a moving bottleneck by a substantial but fixed distance; i.e. to the steady traffic states on either side of a bottleneck.

Such a bottleneck is said to be "active" when, as a result of its presence, the steady states upstream and downstream of it are different. This occurs in practice when the bottleneck is holding back a queue; i.e. when a queue is detected behind it but no queue exists for a long stretch of road downstream. Equation (1) implies that if a stable passing rate q_r exists when an active bottleneck moves at speed v_b, then the two steady states on its sides must be somewhere on the dark straight line of Fig. 1a. From now on q_r will exclusively denote the passing rate when a bottleneck is active. Note that in general, q_r may depend on v_b.

If there exists a reproducible (q,k)-curve of equilibrium states (e.g., as in Fig. 1a) which is not affected by the presence of the bottleneck, then the two active bottleneck states must be on the intersection of this line with the (q,k)-curve. Points "U" (upstream) and "D" (downstream) of the figure illustrate this. The upstream state (U) should always
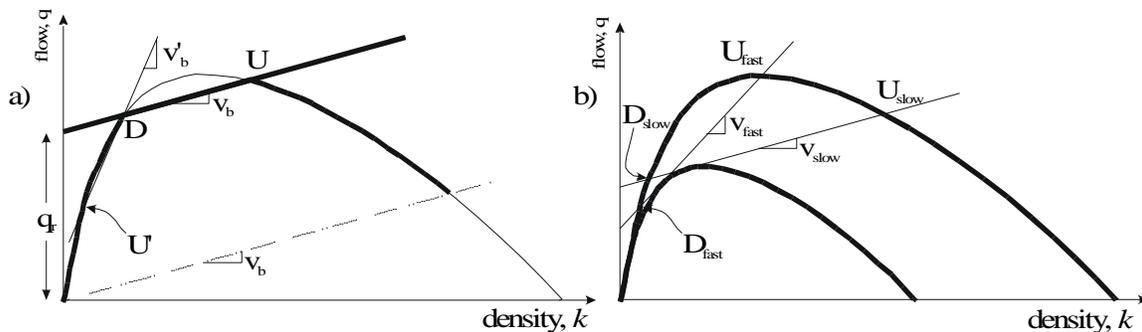

Figure 1. a) Possible stationary states on the flow-density (q, k) plane, upstream and downstream of a moving bottleneck. Effect of bottleneck speed on queue behavior in Gazis-Herman theory. b) Effect of bottleneck speed on bottleneck capacity in Newell's theory.

be the one with higher density and lower speed.[2]

If the bottleneck is not active then the upstream and downstream states must coincide and be on the (k,q)-diagram. If one now assumes that vehicles cannot pass the bottleneck on a sustained basis at a rate higher than q_r ; i.e. that the maximum passing rate is achieved when the bottleneck is holding back a queue, we see that the thin portion of the (q,k)-curve protruding above the thick straight line cannot arise. Furthermore, since a moving bottleneck cannot travel faster than traffic, the thin portion of the curve below the ray cannot arise either. Thus, only the solid portions of the (q,k)-curve are possible.

Note that the highest possible unqueued flow that can persist upstream and downstream of the bottleneck (i.e. when the bottleneck is not active) is achieved at point "D". Since this is also the downstream flow when the bottleneck is active, flows corresponding to points such as "D" will be generically called *bottleneck capacities*. In addition, the solid slanted line corresponding to (1) will be called the *capacity line*.

Two papers consistent with these ideas have been published in the last decade. Gazis and Herman (1992) introduced the moving bottleneck problem. Based on relatively few assumptions, they estimated several features of the system such as the (critical) speed at which a queue would form behind the bottleneck. The assumptions were that flow was conserved, that an equilibrium flow-density relation existed independent of location,[3] and most importantly that the capacity state (or escape state in their terminology) was given, presumably independently of the bottleneck speed. Unfortunately, fixing the capacity state (point "D" in Fig. 1a) for all feasible bottleneck speeds can lead to strange results. For example, the bottleneck could travel so fast that the capacity line could intersect the (q,k)-curve to the left of the capacity state, as shown by the thin straight line with slope v_b'of Fig. 1a and by point U'. Vehicles would then be moving faster in the queue upstream of the bottleneck than after passing it.[4] This strange result does not arise if the uncongested branch of the flow-density curve is a straight line.

Newell (1993) provided a more complete but still unsatisfactory picture of the problem with a slightly different and more complete set of assumptions. Instead of a fixed capacity-state he assumed that kinematic wave (KW) theory held, that the

---

[2] The reverse position for these points does not correspond to an active bottleneck. It physically means that the back of a queue caused by some downstream obstruction is moving forward and (by a remarkable coincidence) together with the bottleneck. The bottleneck is not said to be active because it is not the cause for the different states.

[3] Explicitly, they only assumed that equilibrium states existed and were reproducible. But, of course, the collection of all such states would define a curve. Gazis and Herman (1992) show such a curve toward the end of the paper.

[4] Gazis and Herman (1992) do not state specifically that the capacity state is on the (q,k)-curve, as in Fig. 1a. However, If the capacity-state were not to be on the curve one would have to develop the theory further to explain how downstream traffic returns to equilibrium without disrupting the bottleneck.

bottleneck was a long convoy, and that the traffic stream next to it would behave as in a scaled-down version of the freeway's flow-density curve (see Figure 1b). This implicitly assumes that people's form of driving is not affected by the speed of the convoy. These assumptions allow for variable capacity-states and can be used to construct a solution to any well-posed problem, but if the uncongested portion of the flow-density relation is curved, the model still produces strange results. It predicts that capacity would decline with bottleneck speed as shown in Fig. 1b. This means that a queue could form behind a moving obstruction when the speed of the obstruction *increases*. This does not seem reasonable[5]. It is particularly counter-intuitive for short bottlenecks, such as slow-moving cars and trucks.

Fortunately, the strange predictions of Newell's model disappear if the uncongested branch of the (q,k)-curve is a straight line. This is fortunate because the straight line appears to be a good approximation for uncongested traffic. In this special case the Newell model has a fixed capacity state and the family of capacity lines corresponding to different bottleneck speeds fans out from a single point on the uncongested branch of the (q,k)-diagram, as occurred in the Gazis and Herman case; i.e., both theories coincide. Although the strange predictions disappear in this simple case, the fixed-capacity assumption (i.e. same capacity for any bottleneck speed) is not necessarily correct. Only experiments can determine that. The rest of this section describes a more general theory, which will be tested experimentally.

The general model is specified by assuming the following two things:

(i)   There is a reproducible relation, $q_r(v_b)$, between the bottleneck speed, v_b, and the bottleneck passing rate, q_r , when the bottleneck is active. This is also the maximum rate at which vehicles can pass the bottleneck. The passing rate is assumed to be independent of the history of the system. It can be determined experimentally.

(ii)   KW theory can be used as a first approximation when conditions are changing dynamically, using equation (1) as a boundary condition for the bottleneck when it is active, and making sure that that the relation implied by Figure 1a, $0 \le q - kv_b \le q_r$, holds at all times.

The steps necessary to solve any problem with this general theory, e.g., a problem where the demand is variable and the moving bottleneck changes speeds, are ostensibly the same as those for the simpler case where the Gazis-Herman and Newell theories coincide; see Daganzo (1997) for details about the procedures.

To test the general theory, one should check if traffic around moving bottlenecks exhibits flow- density states on an equilibrium (q,k)-curve (assumption 2). Section 3, below, will furnish evidence suggesting that this is the case. One should also check if

---

[5] Newell (1993) states that the theory is not realistic for light traffic.

reproducible passing rates are obtained for different speeds (assumption 1). Section 4 will show that the experimentally determined $q_r(v_b)$ relations are reproducible and that (contrary to the simple model) they imply an increasing relation between bottleneck speed and capacity.

3. MOVING BOTTLENECK DETECTION AND DIAGNOSIS

3.1 Site description and data preparation

The *Freeway Service Patrol* data set (Skabardonis *et al*., 1994) was used for our study because of its fine level of detail. This data set includes 2-second counts and 2-minute occupancies at every station and lane for a long stretch of U.S. Interstate Freeway I-880 in the East Bay Area for many days of 1993. The sub-site examined in this paper is a 3 km northbound section of the freeway, directly upstream of the connecting off-ramp with freeway I-238; see Fig. 2. Boldface numbers on the bottom of the figure indicate detector stations. These labels express the approximate distance between each station and the I-238 off-ramp in multiples of 100 meters. For example, the first station from the left which is 2705 meters away from station 0 is labeled "27". The diamonds on the top lane designate an HOV lane, but this restriction was not in force during the study period (between 14:15 and 14:45 hrs.)

The data were slightly adjusted to correct for detector drift and bias, as usual, to ensure that all the stations counted the same total number of vehicles. Details of this procedure can be found in Muñoz and Daganzo (2000a). This reference also explains how the N-curves were synchronized, to ensure that all stations started the count with the same vehicle, and also displays the final result of this process.
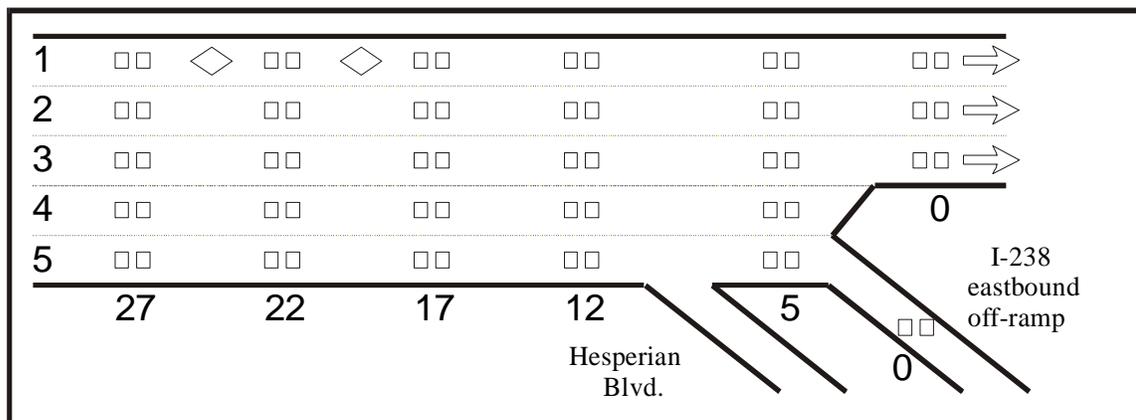

Figure 2. Freeway I-880, northbound; site geometry not to scale; diamonds denote an inactive HOV lane.

It was concluded in Muñoz and Daganzo (2000a) that vehicles were moving in free-flow conditions during the whole period at a speed of about 100 km/hr. This conclusion was reached in three different ways: by cross-correlating the cumulative count series at different detector stations, by cross-correlating the series of ordinary counts, and also from the time series of vehicular speeds at different locations. Figure 3 shows a flow-density scatter plot for Station 27, using a two-minute aggregation of the data from 14:00 to 14:45 hrs. The line shown in the figure is the result of least squares regression, without an intercept. It corresponds to a speed of 100.2 km/hr. The deviations from the line were caused by driver differences and platooning.
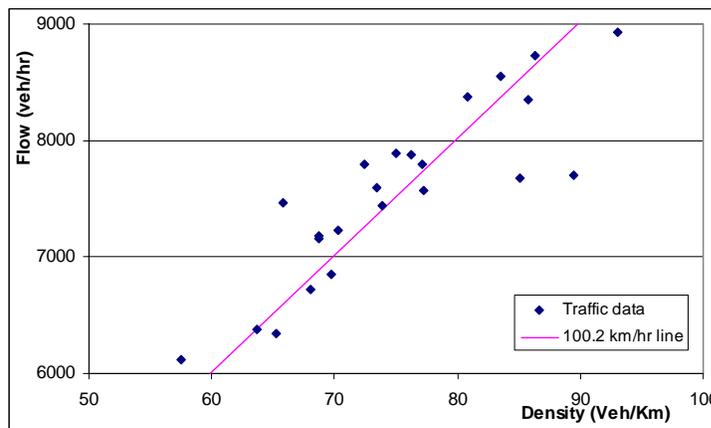

Figure 3. Flow-Density Scatter Plot at Station 27 between 14:00 and 14:45 hrs (2 minute counts).

3.2 A close look at the data

Examination of the data between 14:15 and 14:45 does not reveal anything that contradicts current understanding of free-flow traffic conditions; in particular, the presence of the I-238 diverge does not seem to disrupt uncongested traffic in a significant way. Figure 4a, for example displays the cummulative N-curves of vehicle count for stations 12 and 27 on an oblique coordinate system, defined by two non-orthogonal families of individually labeled parallel lines. In this paper, these lines will be either vertical or slanted, as shown in the figure. Coordinate labels will be shown in boldface for vertical (time) lines and unbolded for slanted (number) lines. The slope of these lines will be -6930 veh/hr; its absolute value is called the "background flow" of the diagram. Note that vehicle accumulation is given by the vertical separation between curves, as in an orthogonal plot, but that trip times are now given by curve separations in the direction parallel to the oblique axis. A tutorial on this technique can be found in Muñoz and Daganzo (2000a and 2000b). Figure 4a clearly shows that disturbances in count

propagate forward with the traffic stream and that trip times remain quite constant despite the fluctuations in count as noted in Cassidy and Windover (1995). The close match is shown more clearly in part b of the figure, where curve N_{27} has been translated toward N_{12} along the slanted lines by the trip time from station 27 to station 12 with a 100 km/hr speed.

(a)

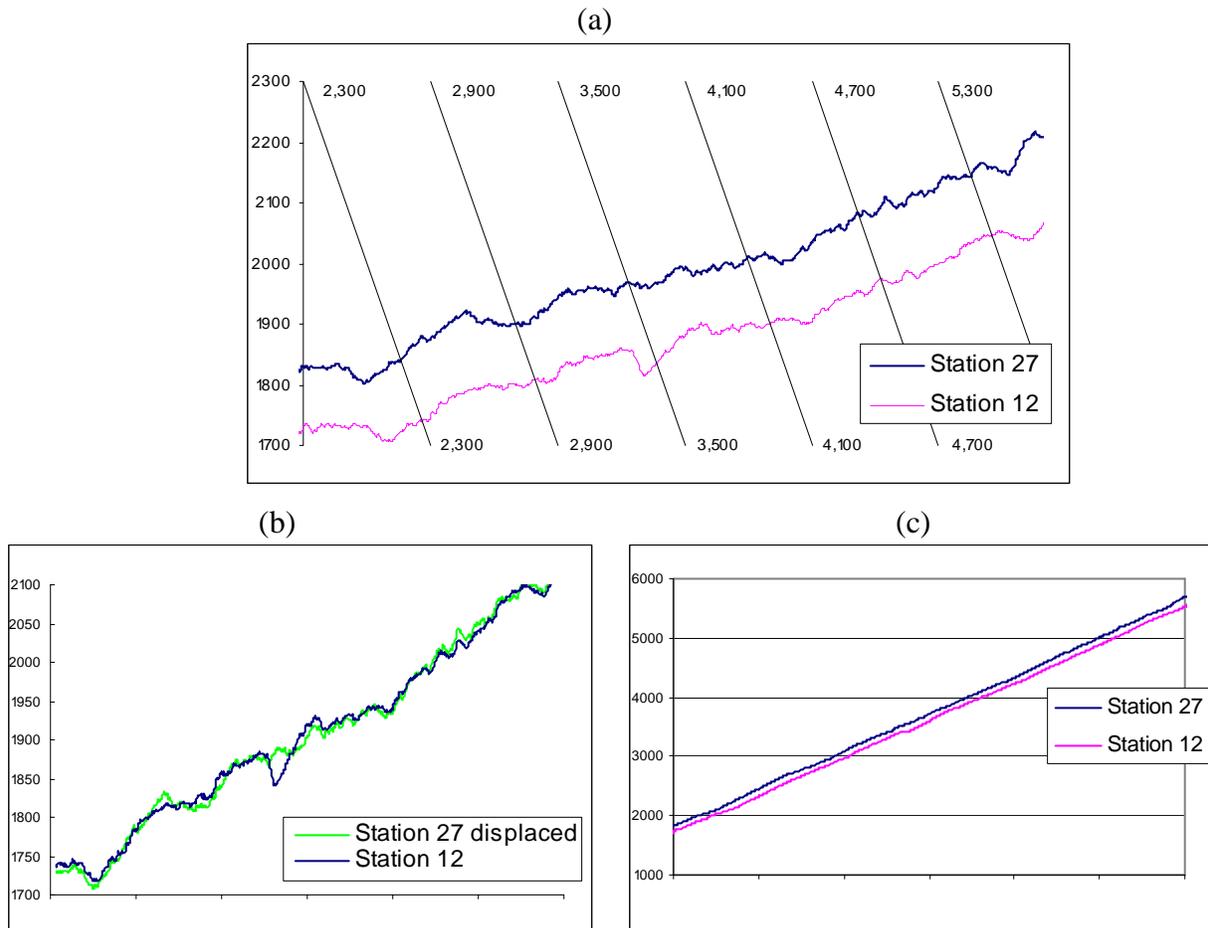(b)   (c)

Figure 4. Traffic conditions at stations 12 and 27 between 14:15 and 14:45 hrs. (a) Oblique N-plot (background flow = 6930 vph); (b) shifted N-plot with N_{27} displaced along the oblique coordinates toward N_{12} (shift corresponds to a speed of 100 km/hr); (c) Ordinary N-plot.

The two curves of Fig. 4b show a large discrepancy around 14:28 hrs, however, and this is examined below. The magnification capabilities of oblique plots allowed us to detect regularities within the discrepancy, which ruled out everything but a moving bottleneck. None of this would have been easy, or perhaps even possible, without this methodology. Note that the discrepancy would be considerably harder to identify with rectangular cumulative plots; see Fig. 4c.

The discrepancy between curves was first magnified by means of a smaller scale plot, as shown in Fig. 5a. This figure also displays the intermediate N-curve

corresponding to station 22. The new figure shows that the discrepancy at station 12 corresponds to a 30-sec period of very low flow. A detector malfunction was ruled out because a similar but shorter-lived effect was found at the intermediate detector, and because the cumulative count later recovered to the "correct" level. A sudden decrease in demand was also ruled out because demand fluctuations should propagate across detectors with little change, but this was not the case here.

Although there is no sure-fire way of finding what happened, the most likely explanation for the patterns in Fig. 5a is that one or more slow-moving vehicles entered the road somewhere upstream of station 27 (for $N \approx 3400$) and that these vehicles only
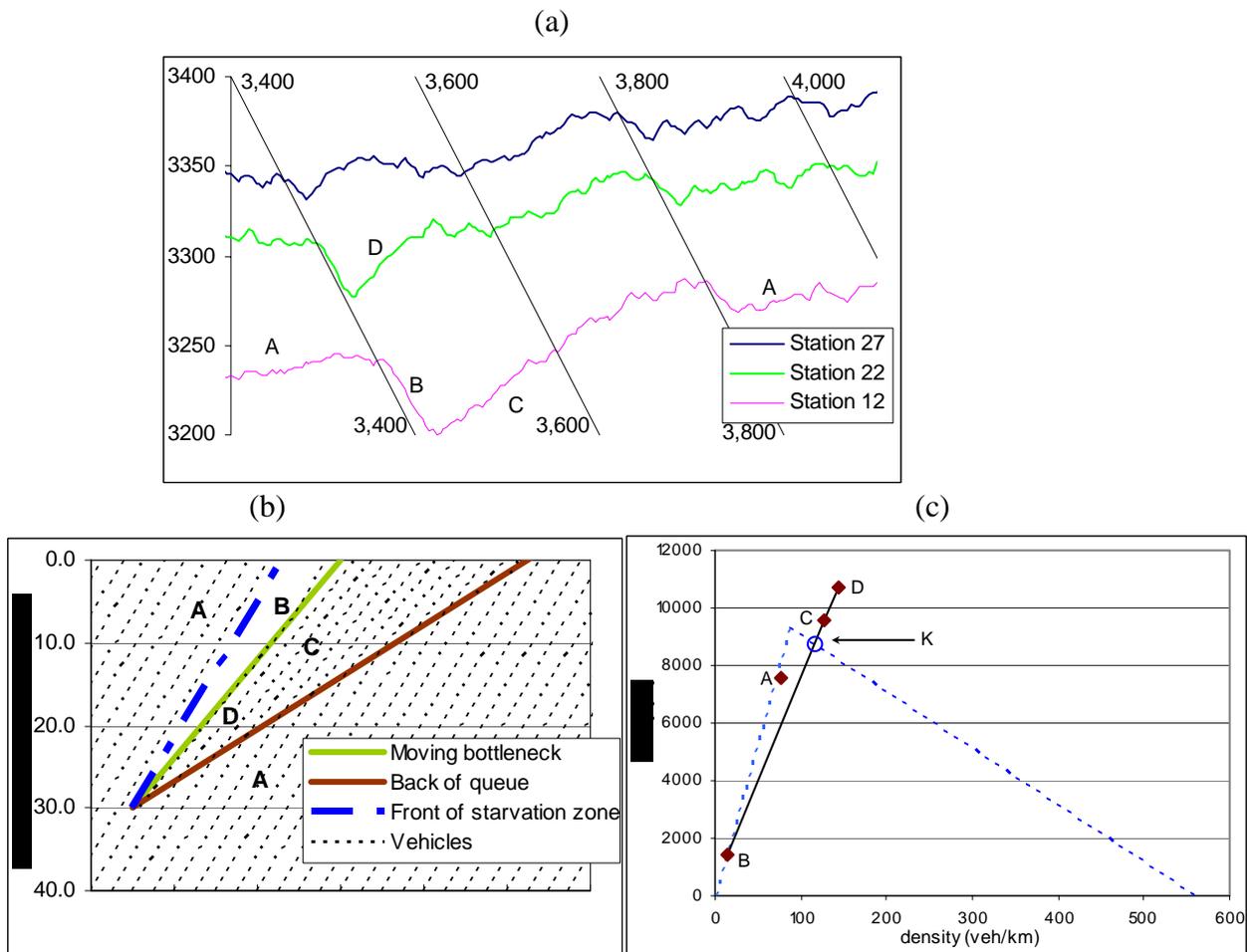
(a)

(b) (c)

Figure 5. Traffic patterns around a moving bottleneck: (a) N-curve signature detected at stations 27, 22 and 12; (b) evolution in time and space; (c) flow and density states.

allowed a small flow to pass. Given the smallness of the passing rate (see below), and the fact that the freeway had five lanes, we speculate that the obstruction may have been a police car traveling close to the speed limit (88 km/hr). This may have induced drivers

to pass hesitatingly, even on the wide freeway. (Both authors have observed this effect repeatedly in their daily lives.)

3.3 Moving bottleneck characteristics

The position of this "moving bottleneck" in each one of the curves is the bottom of the "V". Examination of the figure reveals that the bottleneck speed was 72 km/hr, and that its speed was the same between stations. We can also see by extrapolating the length of the lull in the upstream direction that the incident must have started approximately 3.0 km upstream of station 0.

The slope of left side of the "V" in the oblique coordinate system is the flow downstream of the bottleneck, and the slope of the right side is the flow inside the moving queue directly upstream of the bottleneck. The downstream flows are 1500 veh/hr at station 22 and 1800 veh/hr at station 12.

The rate at which vehicles pass through the bottleneck (in a frame of reference attached to the bottleneck) can also be estimated from the figure since it is the rate at which the N-coordinate of the bottom of the "V" increases with time. Note that "N" increases by about 10 veh (+/- 2 veh) between stations 27 and 12, and that it does so in 75 sec. Hence, the passing rate is about (10/75)(3600) = 480 veh/hr (+/- 100 veh/hr);[6] i.e., only about 2 veh/min/lane, assuming that 4 lanes were available to pass the obstruction.

The flow upstream of the incident is much larger: 10740 veh/hr at station 22, and 9540 veh/hr at station 12, where the queue is more fully developed. Note that the first value is very high (for a 5-lane freeway) and that the two flow levels are experienced by a single set of drivers, N ≈ (3430, 3500), which were in the queue at both locations. This suggests that the differences in flow are not caused by driver differences, but by a change in driver behavior. The data indicate that the first hundred drivers pack themselves closely upon joining the queue and later relax. The data also show that following drivers do not act in this way, since higher numbered drivers (up to N= 3750) never experience the high flows. This differential behavior is probably caused by proximity to the bottleneck when vehicles join the queue. We speculate that the first drivers may be adopting very short spacings in the hope of getting through the bottleneck quickly, and

---

[6] The passing rate could also be obtained without the figure, since vehicle conservation implies that the flow through the bottleneck (in a frame of reference moving with the bottleneck) only depends on the flow and speed prevailing on one of its sides and the speed of the bottleneck. Reassuringly, the result of this calculation is: 1800(100-72)/100 = 504 veh/hr.

that their motivation disappears once they realize that they will be queued for a while[7]. Later drivers act differently because they may not be able to see the bottleneck and do not see a benefit from following aggressively[8]. Therefore, the platoon flow seen at station 12 (about 9500 veh/hr) is more likely to be sustained at succeeding locations. This flow is close to the "capacity" of a 5-lane freeway.

The regime transition corresponding to the back of the moving queue can also be identified in Fig. 5a by the upper right side of the "V's". Although the queue grows in length, it does not grow fast enough to overcome the speed of the bottleneck and propagate upstream; thus, the back of the queue propagates forward and passes all the observers. Note that it passes station 27 at about 14:28 hrs and station 12 around14:30:30 hrs. Thus, it takes about 2.5 minutes to travel 1.5 km, yielding a forward speed of approximately 36 km/hr. Forward propagation is logical since the upstream flow is lower than the queued flow.

Figure 5 also displays the flow-density states that would be consistent with all this information, both, as part of a time-space diagram (in part b) and on an associated flow-density diagram (in part c). Figure 5b also shows the trajectories of several items: (i) the moving bottleneck, (ii) the back of the queue, (iii) the front of the starvation zone and (iv) some vehicles. This part of the figure also indicates the likely time and place where the bottleneck first emerged. Reassuringly, there is an uninstrumented on-ramp at that location. The diagram is consistent with the data of Fig. 5a, in the sense that the timing and magnitude of the regime transitions at the locations of the detectors coincide with the flow changes of Fig. 5a.

Figure 5c displays on the density-flow plane the states "A", "B" and "C" that were observed at detector 12, and also the compressed state, "D", observed at detector 22. The dotted line is an estimated flow-density relation for this freeway, and the open circle the queued state, "K", that kinematic wave theory would have predicted instead of "C" or "D". The flow-density relation was obtained from a broader analysis that included a 3-hour congested period. Complete details can be found in Muñoz and Daganzo (2000a).

Note how state "D" is only experienced at the outset and how the system quickly reaches a state "C", which is close to "K". Thus, except for the small difference between states "C" and "K", which is comparable with the ordinary 2-minute fluctuations during the 3-hour period that were found in Muñoz and Daganzo (2000a), the behavior of the

---

[7] It is not unusual to see very high flows at locations where drivers can be motivated to follow closely for brief periods of time; e.g., the short weaving areas of clover-leaf interchanges, where flows in excess of 3000 veh/hr per lane have been observed (Brilon, 1999)

[8] Changes in following behavior have recently been proposed as explanations for several puzzling phenomena (Daganzo, 1999a and 1999b).

queue behind the moving obstruction is similar to that of an ordinary obstruction. Therefore, despite what appear to be temporary changes in driver psychology, observations are roughly consistent with KW theory.

It should be clear that the detection of our "moving bottleneck", and the quantification of its interesting regularities was only made possible by the oblique plot technique.

4. MODEL OF PASSING RATES[9]

4.1 A two-lane freeway experiment

An experiment to measure the $q_r(v_b)$, relation was conducted on the Richmond-San Rafael (RSR) Bridge of the San Francisco Bay Area, late during the morning rush hour on June 14, 2000. The experiment consisted in creating an artificial bottleneck by introducing a test vehicle into the traffic stream. The time-of-day was chosen so that the disruptions to traffic would be minimized, while still ensuring that queues would develop behind the artificial moving bottleneck. Speeds below 30 mi/hr were not used, to prevent large speed differences from developing around the bottleneck. This important feature of the traffic stream was constantly monitored for the safety of all concerned.

The RSR Bridge has two decks, with two 1-directional lanes (and a shoulder) each. The bridge has no entrances or exits, and is 5.5 miles long. The westbound direction is more heavily traveled during the morning commute, but the eastbound direction also had sufficient traffic for our experiment. The prevailing space-mean speed during our experiment was about 62.5 mi/hr (+/- 1 mi/hr) in both directions, despite the different flows. This suggests that the free-flow speed, v_f , for this bridge is about 62.5 mi/hr, and that there is no significant dependence of speed on flow for traffic conditions ranging from light to heavy (free-flow). These estimates were obtained with the standard "moving-observer" method; i.e., by traveling a set distance while maintaining position in the traffic stream (passing as many vehicles as passed the test vehicle) and then recording the test vehicle's average speed over the distance.

For the bottleneck capacity measurement part of the experiment, the test vehicle first traveled at a reduced fixed speed until a queue developed and then its occupants started to record the number of vehicles that passed it in consecutive 10-sec intervals.

---

[9] This section does not use SI units because we did not use SI instruments, and conversion of their data to SI units would introduce awkward numerical values which could be distracting.

The test vehicle maintained the same speed with cruise control for a run of about 10 intervals (with some variation in this number), and then changed its speed. The data are provided in Muñoz and Daganzo (2000b), and summarized in Fig. 6a. Each point (diamond) in this diagram is the average of the ten or so values obtained for each run. The vertical line through each point denotes the estimated standard error in the average. Runs were done on both directions of the Bridge with two different passenger vehicles as test vehicles. They were done both, with and without the hazard lights on. No systematic effects were detected by these variations. No systematic effects were observed either when a large truck closely followed the test vehicle, essentially becoming the test vehicle. This would suggest (although not definitively) that the effects of slow trucks on the traffic stream are similar to those of slow passenger vehicles; i.e. that vehicle size does not change the properties of the bottleneck appreciably. This comment does not extend to "non-ordinary" vehicles that cause gawking, such as police cars and emergency vehicles; such vehicles should probably have quite different properties.

A hollow diamond is used with the data-line at $v = 0$ because these data were not the result of our experiment. The low end of the vertical line through the hollow diamond ($q = 1400$ veh/hr) is the capacity predicted by the most recent version of the Highway Capacity Manual (Bloomberg et al, 1999) for the case of a 1-lane (stationary) blockage of a 2-lane freeway. This value is probably affected by gawking (rubbernecking) and may not be perfectly representative of incidents moving at very low speeds, since in these cases gawking may not be a factor. Thus, we use it as a lower bound. The upper end of the line (1800 veh/hr) is the saturation flow per lane of a traffic signal. It is reasonable to use this value as an upper bound for $q_r(v_b)$ for $v_b \rightarrow 0$, because traffic signal flows are not disrupted by the merge of two lanes. Thus, the line of the figure represents the likely range of possible values for $q_r(0)$ (for cases without gawking). A reasonable guess for the actual value is the mid-range (1600 veh/hr), which is depicted by means of the open diamond.

The figure also contains a dotted vertical line at $v = 55$ mi/hr, to signify that bottleneck speeds above this level would probably have no effect. We believe that this upper bound, v_{max}, is the average speed on the shoulder lane when the freeway experiences its maximum sustainable flow. We speculate that the maximum relevant speed should be about 55 mi/hr since we observed an average speed of 60 mi/hr on the shoulder lane with sub-capacity conditions, and lower speeds should prevail on the shoulder lanes when the freeway is close to saturation.

Figure 6b displays the capacity lines on the density-flow plane obtained from these data. Only one capacity line was drawn for each speed, using all the information available for the given speed. The figure also contains a darker line, which is the
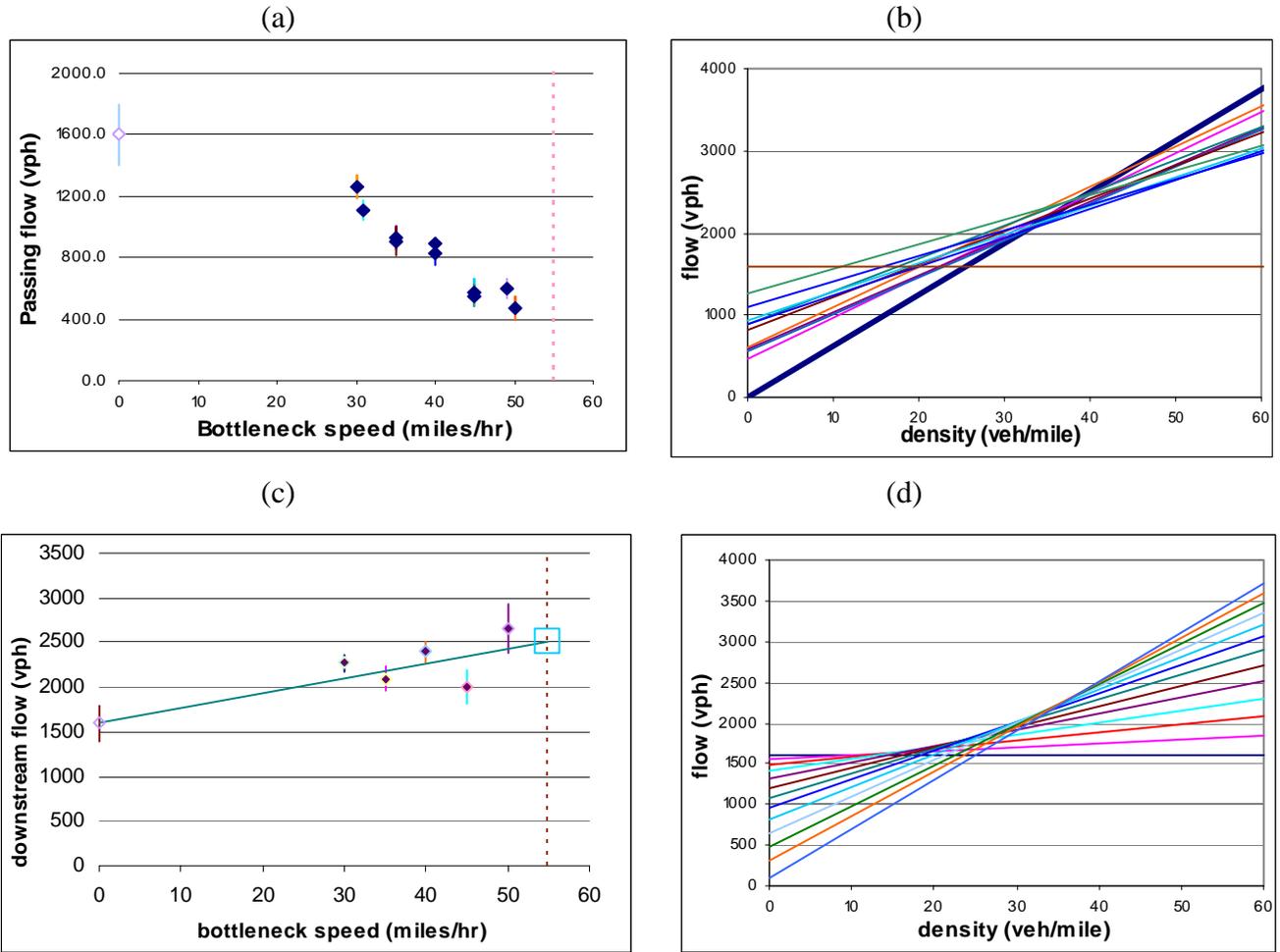
Figure 6: (a) Average passing rate and standard deviation estimated from measurements for different bottleneck speeds. (b) capacity lines on the density-flow plane obtained from the experimental data (c) Downstream (capacity) flow v/s bottleneck speed (d) Family of capacity lines generated by a linear relation between bottleneck speed and downstream flow

(straight) unqueued branch of the flow-density diagram that was estimated for the bridge. The intersections of the thin lines with the darker line give the (capacity) flows, q_D, downstream of the bottleneck. The formula linking q_r and q_D is:

$$q_D = q_r /(1 - v_b/v_f). \qquad (2)$$

Figure 6c plots these downstream (capacity) flows vs. the bottleneck speed, v_b, using the same convention as in part (a) to display standard errors.[10] The results clearly

---

[10] The error increases with speed because the factor that multiplies q_r (and its error) to obtain q_D (and its error) is larger for larger speeds.

establish that capacity flows are considerably greater for a moving bottleneck than for a stationary one. They also suggest that capacity may increase monotonically with speed.

Perhaps the simplest model to capture this effect approximately is one where capacity varies linearly with speed, in the range from $v_b = 0$ to $v_b = v_{max}$. When $v_b = v_{max}$, we would expect q_D to be equal or slightly greater[11] than the flow on the passing lane when the freeway is experiencing its maximum sustainable flow. If we use Q_D to denote this passing lane flow, then the formula for q_D in this crude approximation is:

$$q_D = q_r(0) + [Q_D - q_r(0)](v_b/v_f) \qquad (3)$$

Cassidy and Bertini (1999) have reported that flows of 2630 veh/hr on the median lane of a 3-lane freeway were sustained for a period of 1 hour. For a 2-lane freeway, the maximum passing lane flow should be lower since the left lane will in this case include a broader mix of vehicles. Perhaps, using $Q_D \approx 2500$ veh/hr when $v_b \approx 55$ mi/hr is reasonable for the RSR bridge. This data-point is shown by a large square on Fig. 6c. Thus, a plausible model for the Bridge is: $q_D = 1600 + 900 \, (v_b / 55)$, where flow is measured in veh/hr, speed in mi/hr and $v_b < 55$. This is the slanted straight line on Fig. 6c.

The family of capacity lines generated by this model is displayed on Fig. 6d. Note that capacity lines could be generated graphically without any calculation by specifying the upper envelope of all the lines (which is a convex increasing curve in our case), and then looking for the tangent with the desired slope.

4.2 A three-lane freeway experiment

Another experiment was also conducted on a section of I-80 (westbound) near Richmond, California, using the same procedure. This road is a 4-lane freeway where the median lane is reserved for HOVs. This lane was lightly traveled during our experiment because the HOV restriction was in force at the time. Thus, the experiment should be representative of three-lane freeways. We believe that the results of this experiment are less reliable than those of the RSR Bridge because: (1) the effect of a single-lane restriction is less noticeable on wide freeways; (2) on occasion, vehicles with one occupant used the HOV lane to pass the obstruction and created a void in the queue; and (3) vehicles would sometimes use the auxiliary lanes near on-ramps and off-ramps to pass the obstruction and also created a void.
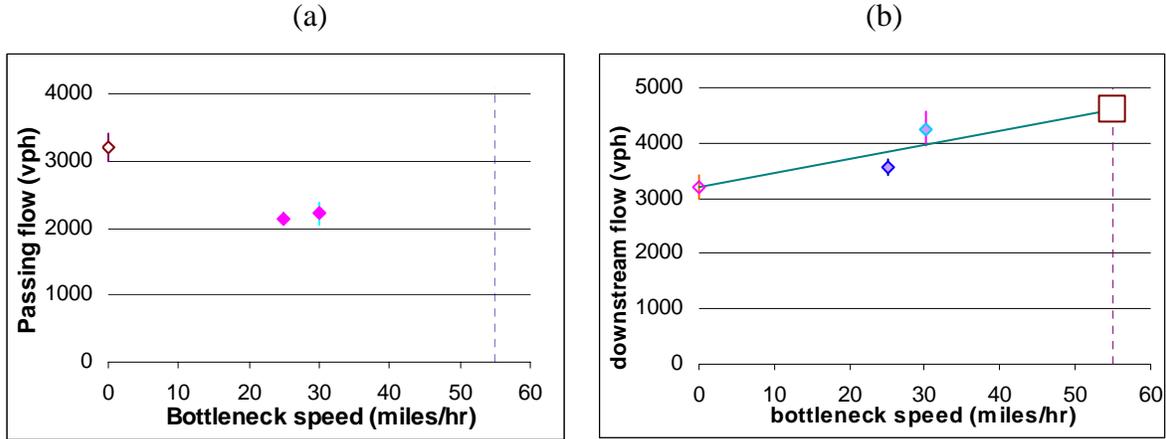
Figure 7: (a) Average passing rate and standard deviation estimated from measurements for different bottleneck speeds. (b) Downstream (capacity) flow v/s bottleneck speed.

Figure 7 presents the results of this experiment using the same conventions of Fig. 6. The raw data is provided in Muñoz and Daganzo (2000b). As in the previous case, there should be little doubt that downstream capacity increases with bottleneck speed. The lower bound for $v_b = 0$ is again taken from Bloomberg et al (1999), and the upper bound from saturation flows. Our best guess for $q_r(0^+)$ continues to be the mid-range between the bounds; i.e, $q_r(0^+) \approx 3300$ veh/hr. For $v_b = v_{max}$, the capacity flow ($Q_D \approx 3600$ veh/hr) was taken directly from Cassidy and Bertini (1999). The linear model is then as shown on the figure.

5. DISCUSSION

Police cars and linearity: It should be emphasized that there is no theoretical justification for assuming a linearly increasing relation between capacity and speed, except simplicity. Thus, the proposed model should be taken with a "grain of salt". It should also be noted that the linear model should not be applied to police cars, since vehicles should be quite hesitant to pass the police car when it is traveling close to the speed limit. In this case, the bottleneck capacity may actually decline with increasing speed, and may approach zero for the speed limit. In Sec. 3 we found a bottleneck that moved at 45 mi/hr and reduced downstream flow to 1800 veh/hr. An ordinary bottleneck would have to span 4 out of 5 lanes to have such a noticeable effect, but a humble police car traveling at such speed could have easily achieved the feat.

---

[11] Drivers only have to sustain the short headways for a short time while passing the moving bottleneck.

Long bottlenecks: The proposed capacity formulas should not be applied to long bottlenecks such as military convoys, because in these cases drivers may be less willing to sustain next to the convoy the short headways required to achieve the high passing rates observed in our experiments for high bottleneck speeds. As a result, one may find experimentally that bottleneck capacity increases less rapidly with bottleneck speed in this case. Perhaps, capacity is nearly independent of speed for sufficiently long bottlenecks. This, question is somewhat academic, though, since most bottlenecks are short.

High bottleneck speeds and driver differences: The general model proposed in Sec. 2, even with variable capacity states, is only a first approximation. The main weakness is that it relies on the KW model for light traffic when passing is possible; in these situations kinematic waves do not exist and driver differences matter. We are most doubtful of its performance for high bottleneck speeds, when not all vehicles in the queue may have the same incentive to pass. Predictions can perhaps be improved with extensions of the KW model that allow for passing, different vehicle classes and changes in driver psychology.[12] These models can better explain why a long queue behind a bottleneck may begin to dissipate before an accelerating bottleneck reaches the free-flow speed. As an illustration of this possibility, Appendix A explains the time-space solution of a problem where a bottleneck accelerates from a standstill until it reaches the desired speed of the slowest vehicles, using the two-vehicle class theory in Daganzo (1999a).

REFERENCES

Bloomberg, L, S Cohen, B Eads, F Hall, D May, N Rouphail. and T Urbanik. (1999) "Freeway systems," Highway Capacity Manual: First draft of Chapter 25, Presented at Trans. Res. Board, Washington, D.C.

Brilon, W (1999) Private communication.

Cassidy, M and R Bertini (1999) Observations at a freeway bottleneck. In *Transportation and Traffic Theory*, Proc. 14th Int. Symp. On Transportation and Traffic Theory, pp.107-146 (A. Ceder, editor), Pergamon, New York, N.Y.

Cassidy, M and J Windover (1995). Methodology for assessing dynamics of freeway traffic flow. *Trans. Res. Rec*. **1484**, 73-79.

Daganzo, C F (1997). *Fundamentals of Transportation and Traffic Operations,* Pergamon-Elsevier, Oxford, U.K.

---

[12] These models relax two basic assumptions: (a) the existence of a reproducible (q,k)-relation across all traffic lanes, and (b) the non-existence of passing rates greater than q_r when there are no queues.

Daganzo, C F (1999a). A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections. *Institute of Transportation Studies Research Report* UCB-ITS-RR-99-05, University of California, Berkeley

Daganzo, C F (1999b). A behavioral theory of multi-lane traffic flow. Part II: Merges and the onset of congestion. *Institute of Transportation Studies Research Report* UCB-ITS-RR-99-06, University of California, Berkeley.

Gazis, D C and R. Herman (1992). The moving and 'phantom' bottlenecks. *Trans. Sci.,* **26**, 223-229.

Lighthill, M J and G B Whitham (1955). On kinematic waves. I flow movement in long rivers. II A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. A*, **229**, 281-345.

Muñoz, J C and C F Daganzo (2000a). Experimental Characterization of Multi-Lane Freeway Traffic Upstream of an Off-Ramp Bottleneck. *Institute of Transportation Studies Research Report* UCB-ITS-PWP-2000-13, University of California, Berkeley.

Muñoz, J C and C F Daganzo (2000b). Experimental Results with Moving Bottlenecks. *Institute of Transportation Studies Research Report* UCB-ITS-RR-2000-6, University of California, Berkeley

Newell, G F (1982), *Applications of queueing theory* (2nd edition), Chapman Hall, London.

Newell, G F (1993). A moving bottleneck, *Institute of Transportation Studies Research Report* UCB-ITS-RR-93-3, University of California, Berkeley, California. (In abridged form, *Trans. Res.* 32B, 531-537, 1988.)

Richards P I (1956). Shockwaves on the highway. *Opns. Res.*, **4**, 42-51.

Skabardonis, A, H Noeimi, K Petty, D Rydzewski, and P Varaiya (1994). Freeway service patrol evaluation. *Institute of Transportation Studies Research Report* UCB-ITS-PRR-94, Univ. of California, Berkeley, CA.

Smilowitz, K, C F Daganzo, M J Cassidy and R. Bertini (1999). Some observations of highway traffic in long queues. *Trans. Res. Rec.* **1678**, 225-233.

Windover, J R (1998). Empirical studies of the dynamic features of freeway traffic. *PhD thesis*, Dept. of Civil and Environmental Engineering, Institute of Transportation Studies, University of California, Berkeley, CA.

Appendix A: An Accelerating Bottleneck

Figure A1 illustrates the stationary states we expect to find when an accelerating bottleneck is introduced in a traffic stream composed of two types of drivers, "slugs" and "rabbits", with different desired free-flow speeds (v_{fs} and v_{fr} respectively). The notational conventions in the figure are taken from Daganzo (1999a). We assume that rabbits outnumber slugs, and that rabbits are faster ($v_{fr} > v_{fs}$). As in that reference we also assume that in traffic conditions where average speed is faster than v_{fs} slugs use only the right lane, and rabbits the left lane. Otherwise, both types can be found in both lanes.

The thick lines of Figure A1a are the q-k diagrams for this model. The lower triangle is for the shoulder lane, which can only be populated by slugs in free-flow. The broken line directly above it is for the median lane, which can only be populated by rabbits when the speed on that lane exceeds v_{fs}. The upper down-sloping line is the congested (1-pipe) portion of the diagram for both lanes together. In this regime, all the vehicles travel at the same speed, and the data points representing the different classes line up on a ray from the origin; see for example the points for state "A". In the (2-pipe) passing regime, the passing-lane point contains only rabbits and the shoulder lane point only slugs; state "D" is an example. The position of the combined data point for a 2-pipe state is not on any curve since it depends on the particular mixture of rabbits and slugs.

Note that the maximum flow of rabbits on the median lane can be considerably higher than the maximum flow on the shoulder lane. This is because, in this theory, rabbits can become quite motivated to follow closely under certain stimuli. Under normal conditions, though, e.g., if a restriction holding back a queue is removed, the maximum rabbit flow is that of point "D". (Higher flows can only be achieved in the theory as the stream flows past an on-ramp and other rabbits squeeze into the passing lane; this effect does not play a role in Figure A1, however.)

Figure A1b is a time-space diagram of an accelerating bottleneck and the traffic states next to it. The thick line represents the bottleneck trajectory. The thin lines are vehicle trajectories and interfaces between stationary states. The states refer to those in the q-k diagram above. It is assumed that passing occurs as per the capacity lines of part (a) of the figure. Note that the downstream flow when $v_b = v_{fs}$ is that of point "D". The geometrical construction for the part of the diagram that is upstream of the bottleneck is exactly as in KW theory because traffic is in a 1-pipe regime. However, the downstream behavior is different because there is passing. Trajectories are not drawn downstream of the bottleneck because this would complicate the picture considerably. Therefore, this appendix explains how one would determine the downstream states precisely.

Consider first the initial part of the diagram, when the bottleneck is stopped. Here,

the bottleneck line is flat and the passing rate is also the total flow upstream and downstream of the bottleneck. Data points "A", represent the upstream state. Note that the combination point for both lanes is on the capacity curve and that the speed in the queue is the same for all vehicles. Once vehicles pass the bottleneck and reach free-flow conditions, however, slugs take the shoulder lane and rabbits stay on the passing lane; a two-pipe state arises. Therefore, because rabbits and slugs are conserved across the bottleneck, the downstream conditions on each lane are simply obtained by projecting the total-rabbits data point (dotted circle) and the total-slugs data point (empty square) horizontally, across to their respective free-flow branches of the q-k diagram. This ensures that the flow of rabbits and slugs is the same on both sides of the bottleneck. The downstream state is denoted "a".

When the bottleneck starts moving at the speed shown in Figure A1b the passing rate changes. The new bottleneck line will pass through a new combination point ("B"), which will still be queued as a 1-pipe. Since the percentage composition of the queue cannot change, the component points are obtained by projecting the "A" points onto the "B" ray in a direction parallel to the congested branch of the (q,k)-curves. As in case "A", traffic downstream of the moving bottleneck will be in a 2-pipe state. Furthermore, conservation of rabbits and slugs across the bottleneck continues to determine the downstream state. Now one would project the total-rabbit and total-slug states on the B-ray towards their respective free-flow branch of the q-k diagram, but would project them in a direction parallel to the bottleneck line (since slope = speed of the bottleneck). The downstream state is denoted "b". Note that that there would be a portion of time-space with rabbits from state "b" and slugs from state "a". This is denoted "ab".

Notice that if the bottleneck speeds up to a v_{fs} , something unusual would happen. Upon reaching this speed, slugs no longer wish to pass. Therefore, the bottleneck essentially becomes a slug. It would leave the queue behind, and the bottleneck would shift to the slugs upstream. This process is also shown on Figure A1.

Upstream of the bottleneck a first acceleration wave would travel shifting the 1-pipe state from "B" to "C" (the capacity, 1-pipe state). A second (slower) wave, depicted by a double line, later shifts the queue into a 2-pipe state. Conservation of rabbits, since their discharge state "D" is known, determine the speed of this transition. To understand the process intuitively, imagine that you are a rabbit in the right lane of the queue. First, you will notice that everybody increases its speed to v_{fr} (wave "BC"), and later you will notice that the left lane is moving faster than your lane and that vehicles in your lane don't want to speed up. If you are a rabbit, you will change lanes, and as you do this, you will be experiencing the regime transition ("CD"). Eventually, you will overtake the original bottleneck. If you are a slug the regime transition does not affect you, and you

will not catch up with the bottleneck.

It should be obvious by now, that in this theory there must be a critical bottleneck speed below v_f, above which queues cannot grow, no matter how high the initial flow.
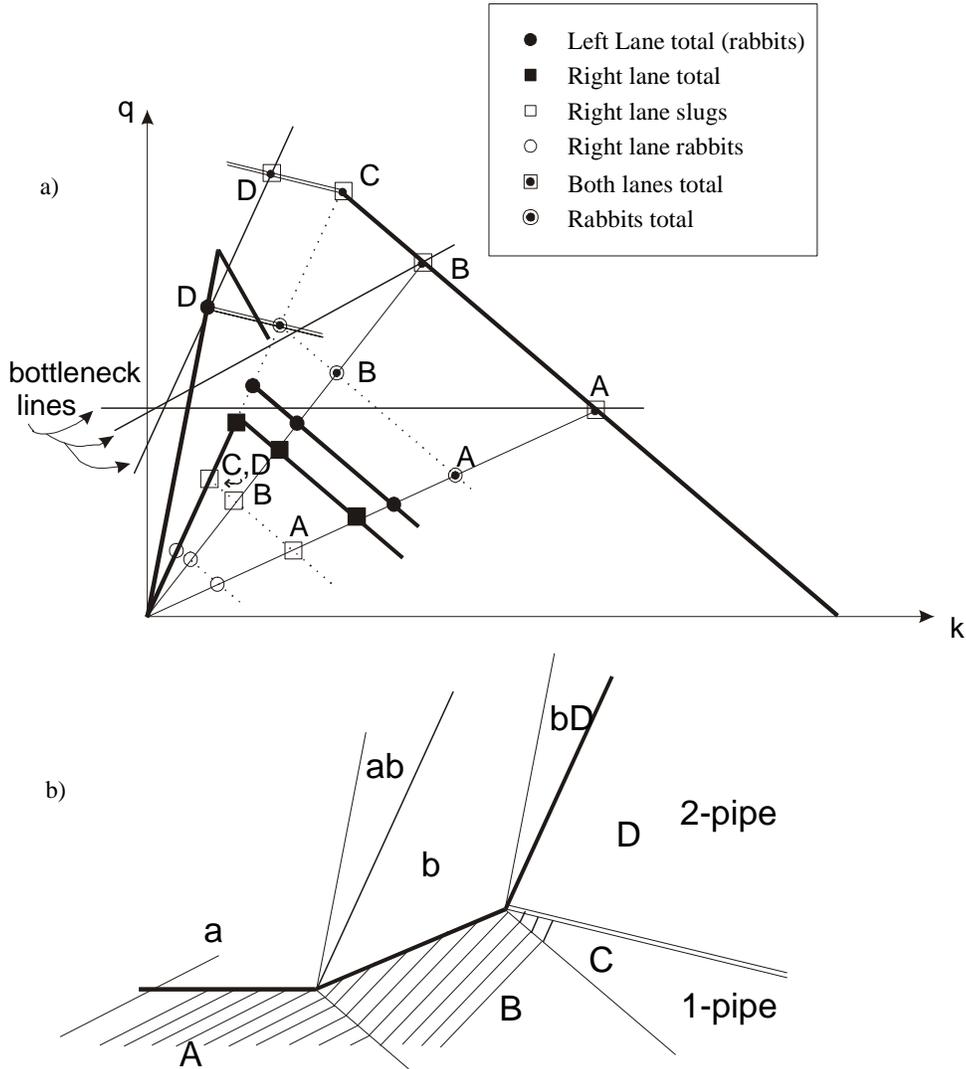

Figure A1: Solution of a problem where a bottleneck accelerates from a standstill until it reaches the desired speed of the slowest vehicles. a) Flow-density diagram, b) Time-space diagram.