Causes of Phase Transitions in Highway Traffic

C.F. Daganzo, M.J. Cassidy and R.L. Bertini

Department of Civil and Environmental Engineering
University of California, Berkeley, CA, 94720, U.S.A.
Phone: 510-642-3853
Email: daganzo@ce.berkeley.edu

Originally Submitted December 23, 1997
Revised January 26, 1998

Causes of Phase Transitions in Highway Traffic

C.F. Daganzo, M.J. Cassidy and R.L. Bertini
Dept. of Civil and Environmental Engineering, Univ. of California, Berkeley, CA, 94720, U.S.A.

(December 23, 1997)

Abstract

It is shown that all phase transitions in and out of freely flowing traffic reported earlier for a German site could be caused by bottlenecks, as are all transitions observed at two other sites examined here. Furthermore, the evidence indicates that bottlenecks cause these transitions in a predictable way, and no evidence is found that stoppages (jams) appear spontaneously in free flow traffic for no apparent reason.

This letter provides additional empirical observations and an alternative explanation of some traffic data from a German highway [1, 2, 3]. An expanded discussion including additional references, can be found in [6]. It will be shown that the additional (North American) observations are consistent with the German data. However, none of the data support the most striking conclusion in [1]: that free flowing traffic will spontaneously break down randomly, without obvious reasons, and then remain in that state due to traffic's tendency to self-maintain congestion. Rather, the evidence indicates that traffic *breaks down* (queues form) at locations of freeway inhomogeneities (bottlenecks) in a predictable way and that queues do not self-maintain away from a bottleneck in the absence of an incident. Highways appear to behave like crowds of people going through a series of queues at well defined locations; e.g., fans who wait in line to buy football tickets and then queue through one or more gates to enter the stadium. Some definitions derived from this analogy are introduced below.

Free flow (or *unqueued*) conditions refer to that state of traffic where small disturbances flow forward in space (e.g., as in supersonic gas flow). People walking away from the ticket window, for example, are in free flow because if one of them momentarily slows down just a little, the queue is not affected upstream of the server. Conditions that are not free flow will be called *queued*. In queued traffic, disturbances can and do flow upstream. *Bottlenecks* are those inhomogeneous locations such as a ticket window where queues can form and persist with free flow downstream. A bottleneck in this state will be said to be *active*. If the flow through an active bottleneck is nearly constant and reproducible this flow will be called the bottleneck *capacity*. The evidence available in German and North American traffic suggests that queues grow when bottlenecks become active and that instabilities may be the result and not the cause of queues.

Merges: Included here is evidence that traffic instabilities upstream of an active merge bottleneck appear not to influence the bottleneck flow significantly, i.e., one can define a *merge capacity*, that queues grow shortly after the merging flows exceed this capacity, and that queues dissipate predictably. This explains the breakdown described in Figs. 1, 2 and 3 of [1], as well as those observed on North American sites.

Figure 2(b) of [1] shows that at 7:16 a.m. a drop in the time series of measured speeds for the passing lanes was recorded at the first detector (D3) downstream of the merge, while the speeds upstream and downstream of it remained higher. This change in speeds was later detected upstream and downstream, suggesting that the bottleneck became active and a queue began to grow near detector D3 at around 7:16 a.m. The sharper changes in speed observed later at upstream locations are what one would expect if the back end of such a queue was growing in the upstream direction;

it passed over detector D2 at 7:22 a.m. and detector D1 later. It is further stated in [1] that the "phase change" (i.e., the onset of queueing) was brought about by an increase in the on-ramp flow and that the phase change was self-maintained for about two hours.

These observations are consistent with the simplest theories available [6]: when a queue is generated on the freeway, the bottleneck flow is expected to stabilize at a predictable level until the freeway queue dissipates. The data presented in [1] suggests that the fluctuations in ramp flow are absorbed by the mainline flow, as expected. Incomplete proof of this is contained in Fig. 2(e) of [1], which presents the rather stable time-series of flow in the left lane at the two detectors immediately downstream of the merge. The figure shows that these average flows did not change appreciably after the onset of queueing, suggesting that the cumulative number of vehicles entering the merge after this event did not change significantly either.

Far from exceptional, the behavior of the German bottleneck was reproduced qualitatively at a North American site analyzed in [4]. The main results are summarized below. The analysis is based on cumulative flow curves because they are more informative than the time series data of counts and speed used in [1]. The results are also presented in a form suitable for comparison with [1].

The analyzed site is a segment of the Queen Elizabeth Way in Toronto, Canada, with the geometry of Fig. 1(a). This freeway is instrumented with detectors that record vehicle counts, occupancies (i.e., the dimensionless measure of density obtained by loop detectors [5]) and time averaged speeds over 30-second intervals. (The detector station labels shown are those used by the Ontario Ministry of Transport.) Typical of several other days that were analyzed at this location, the data shown here were gathered on the morning of May 3, 1995.

Figure 1(b) presents transformed *N*-curves of cumulative flow (in all lanes) versus time for the

4 detector locations of our site during the onset of queueing. Note that an untransformed *N*-curve gives the cumulative number of vehicles to have passed detector station *I* by time *t*, starting the counts ($N = 0$) with the passage of a reference vehicle. Thus, horizontal separations between *N*-curves are trip times and vertical separations the accumulations between detectors [5]. In Figure 1(b), the curves (along with their respective time axes) have been shifted to the right by the average free flow trip time between the respective detector and detector 25, so that the vertical separation between curves now represents the excess vehicular accumulation between detectors due to vehicular delays. Such a shift is advantageous because two superimposed curves indicate that traffic in the intervening segment is freely flowing—since every feature (disturbance) of the upstream *N*-curve is passed to its downstream neighbor later. In addition, the figure only shows the difference between each curve and the line $N = q_o t$ because this background flow reduction magnifies details without changing the excess accumulations. The superimposed curve portions in this figure indicate that traffic was initially in free flow and remained in free flow between detectors 24 and 25. The marked separation of curves 24/25 from curve 23 from 6:29 a.m. onward (as shown by the arrow) indicates that a bottleneck was activated a little earlier between detectors 24 and 23. The subsequent separation of curve 23 from curve 22 indicates when the queue arrived to detector 23.

Figure 1© presents the transformed *N*-curves for stations 22 and 25 for the entire rush. (Additional tests explained in [4] confirmed that the bottleneck was active during this time.) Note the persistent displacement between these curves (i.e., the queue) while the bottleneck was active. As shown by the dashed line, the queue discharge rate measured at detector 25 varied slowly in time, but *N* never deviated by more than about 50 vehicles from this trend line with average rate 6,470 vph. Also evident in the figure is the maximum flow of 6,970 vph that persisted for 12.5 minutes during the onset of queueing.

Repeated observations at this and at another freeway merge in Toronto indicated that queue formations were always accompanied by brief periods of excessively high (combined) arrival rates and that the average discharge flows were reproducible from day to day [4]. Having demonstrated that the Toronto queue formed due to a bottleneck, we next show that the German data are qualitatively similar to Toronto's and thus, the former likely describes the activation of a merge bottleneck as well.

Figure 2 presents time-series plots of the average vehicle speeds at the Toronto site measured over 1-minute intervals in each travel lane and at each of the four detectors, as was done in [1]. As on the German site, speeds upstream of the bottleneck at detectors 23 and 22 (Figs. 2(b) and 2© respectively) drop markedly during the “phase change” to the queued state. Also reported in [1], speeds dropped downstream of the merge when the queue formed (as occurred here at station 24 by 6:29 a.m.) with smaller reductions exhibited a little later further downstream (as occurred here at station 25). This indicates that vehicles gradually accelerate after discharging from the bottleneck queue, and they discharge at an average rate equal to the bottleneck capacity, as shown in Fig. 1(c) of this letter and in Fig. 3(a) of [1].

Figure 3 presents scatter plots of 1-minute occupancies (densities) versus flows in the left-most travel lane of the Toronto site. Unshaded, smaller circles denote observations subsequent to the arrival of the queue formation waves at each location. Reference [1] demonstrates that the 1-minute flow-density data fluctuate markedly upstream of the merge, and this can also be seen here at station 22 (Fig. 3(a)). The reference also contains plots similar to Figs. 3(b) and 3(c), again indicating that vehicles gradually accelerated downstream of the bottleneck. In fact, Figs. 2 and 3 are so similar to

Figs. 2(a-d) and 3© of [1] there is little doubt that the German data describe a bottleneck formed by merging vehicles.

Diverges: It is further shown in [1] (Fig. 4) that a phase transition into the queued state also occurred at a different location and time. This disturbance was characterized as spontaneous and due to a random occurrence because it formed away from a merge. The disturbance began adjacent to an off-ramp with much lower speeds in the shoulder lanes at locations near the exit (see detector D9 in Fig. 4 of [1]) which is a *signature* of a brief interruption caused by an oversaturated (i.e. queued) diverge.

A diverge on a narrow freeway is not qualitatively different from a single line of cars at an uncontrolled T-junction were a left-turning car may force everyone behind to wait (e.g. because of opposing traffic) even if it is safe for other vehicles to proceed. Likewise, if a freeway off-ramp cannot accept the traffic wishing to exit, a queue entrapping some through-vehicles will form on the freeway and reduce the flow, perhaps even blocking all the lanes further upstream. According to current theory, these effects are more pronounced if flows are close to saturation and the freeway is narrow as in [1] (see [6] for more background).

The onset of queueing next to the off-ramp can be triggered in at least two ways: 1) a queue from the off-ramp spills-over and blocks the freeway traffic and 2) the off-ramp is unqueued but an increase in the (desired) exit flow greater than the off-ramp capacity creates a freeway queue. According to prevailing theory a recovery wave like that of Fig. 4 of [1] should be issued from the off-ramp in the latter case if the flow of exiting vehicles approaching the off-ramp drops below the saturation level of the ramp. Note that the effects due to 2) would occur *even if the rate at which vehicles arrived to the queue and the exit ramp flow remained nearly constant*; i.e., the effects could be interpreted as occurring for no apparent reason. The information in Fig. 4 of [1] does not include the ramp flows, but from the node conservation law it appears that both the freeway flows (at station D9) and the ramp flows (the difference between the flows of D9 and D10) were close to saturation prior to the genesis of the disturbance. Its short duration suggests a disturbance of type 2, although it could also be due to exiting traffic (perhaps involving trucks) that required some lane changing. The development of this disturbance into a stoppage that propagates upstream is consistent with the simplest theory of traffic instability, as explained in [6].

The situation observed in [1] does not appear to be exceptional. Evidence that off-ramp queues can block adjacent freeway lanes can be found even when freeways are wide and upstream flows well below saturation; i.e., when a disturbance would seem even less likely to affect adjacent lanes. As an illustration of this, Figs. 4(a)-(b) describe the evolution of an off-ramp queue on a segment of Interstate 880 in Hayward, California (U.S.A.). At its upstream end, the segment has four regular lanes and a median lane for car pools and buses, numbered as shown. (Also shown in the figure are the detector numbers used by the California Department of Transportation.) The shadings drawn here correspond to the 2-minute occupancies (densities) measured by the detectors at two different times during the afternoon rush on March 8, 1993. An occupancy of about 25% or more corresponds to queued traffic on this freeway [7] so that unshaded portions denote free flow. Figure 4(a) shows that the disturbance originates at the off-ramp while 4(b) shows that as the queue propagated upstream, it moved in the transverse direction and blocked all regular lanes. The queue eventually passed detector 20 [6].

Figure 5 presents the same data in the form used in [1]; data from the car pool and bus lane are not shown. Figure 5(a) shows that the flows at the upstream end of the section were always well below saturation in most of the regular lanes. This explains why the queue did not propagate upstream of detector 20 (Fig. 5(h)) or appreciably reduce the output flow measured at detector 8 (Fig. 5(b)). Figure 5 also reveals the speed reductions brought by the queue. Three similarities with [1] stand out: (1) larger speed reductions in the shoulder lanes at locations near the exit, (2) upstream propagation of the speed reductions with a delay and (3) equally low speeds in all lanes at locations sufficiently far upstream of the exit.

In summary, the attributes of the speeds measured in California are similar to those in [1], although the drop in the former persists for over 1 hour. Furthermore, although the flow drop on the California site was not severe, the occupancy shading on the off-ramp leaves no doubt that this ramp spilled over onto the freeway (reason 1) for more than 1 hour. Thus, the flow and speed data of Fig. 5, which might otherwise be even more puzzling than those of the German site, can be explained in terms of a clear cause.

Final remarks: In our view, and despite the existence of traffic instabilities, one cannot claim that traffic queues form spontaneously on homogeneous roadways and then self-maintain themselves based on the available evidence. Reference [6] shows that disturbances can grow and propagate in queues without affecting the bottleneck flow, and that the most unusual phenomena in [1, 2, 3] and [8] can be explained in terms of simple theories specific to traffic. We thank G.F. Newell for his comments.

----

[1] B.S. Kerner and H. Rehborn, Phys. Rev. Let. **79,** 4030 (1997).

[2] B.S. Kerner and H. Rehborn, Phys. Rev. E **53**, 4275 (1996).

[3] B.S. Kerner and H. Rehborn, Phys. Rev. E **53**, 1297 (1996).

[4] M. Cassidy and R. Bertini, U.C. Berkeley Institute of Transportation Studies Research Report, ITS-RR-97-07 (1997).

[5] C.F. Daganzo, *Fundamentals of Transportation and Traffic Operations* (Pergamon, 1997).

[6] C.F. Daganzo, M. Cassidy and R. Bertini, University of California at Berkeley Institute of Transportation Studies Research Report, ITS-RR-97-08 (1997).

[7] M. Cassidy, Transp. Res. **32B** (1998).

[8] J. Treiterer and J.A. Myers, in *Proc. 6th Int. Symp. On the Theory of Transportation and Traffic Theory*, edited by D.J. Buckley (Elsevier, 1974), pp. 13-38.

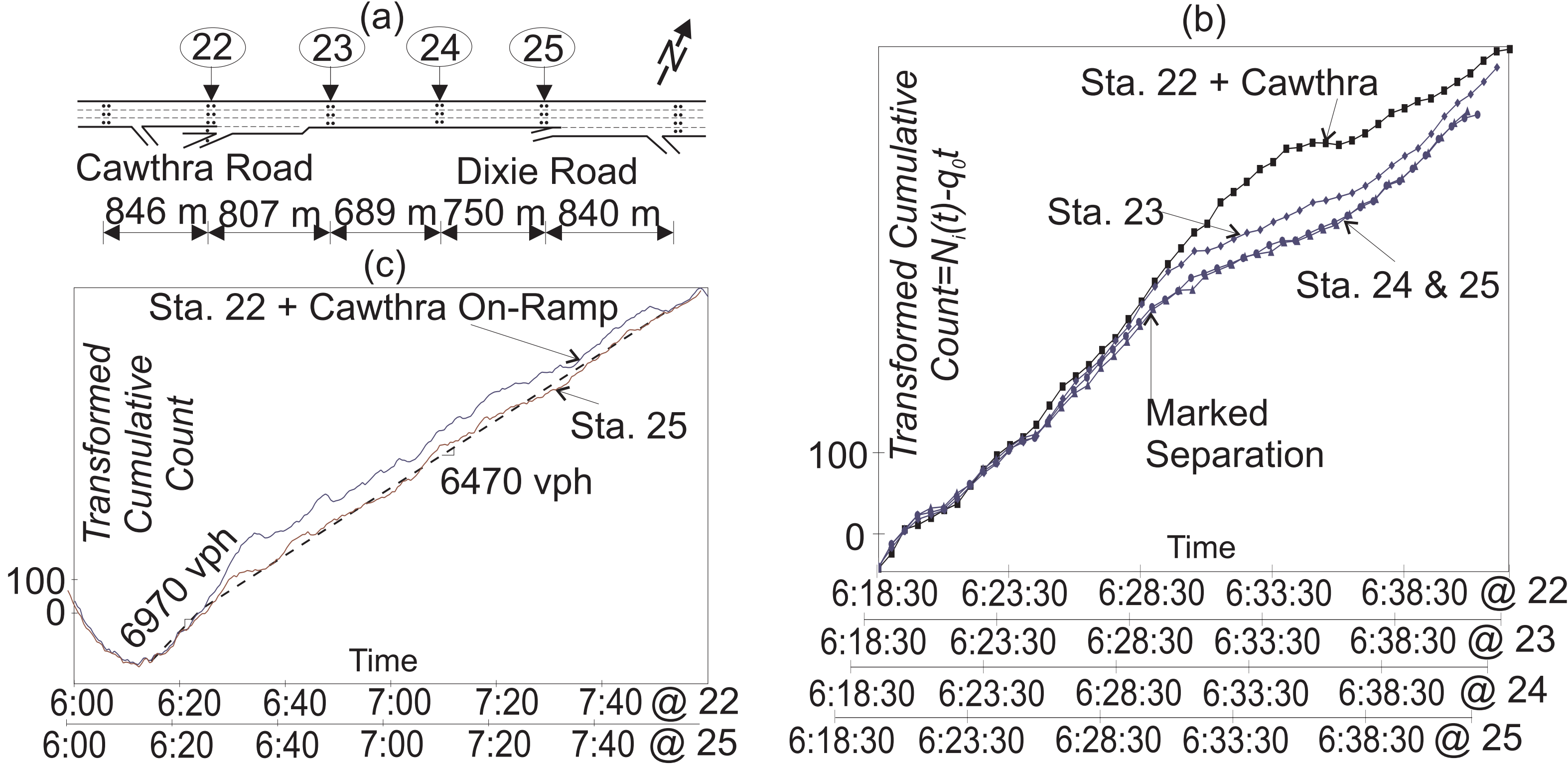

FIG. 1. Queue evolution at a merge bottleneck. (a) Segment of Queen Elizabeth Way, Toronto, Canada. (b) Cumulative curves showing queue formation at bottleneck. (c) Cumulative curves showing features of the active bottleneck. (d) Cumulative curve of on-ramp.

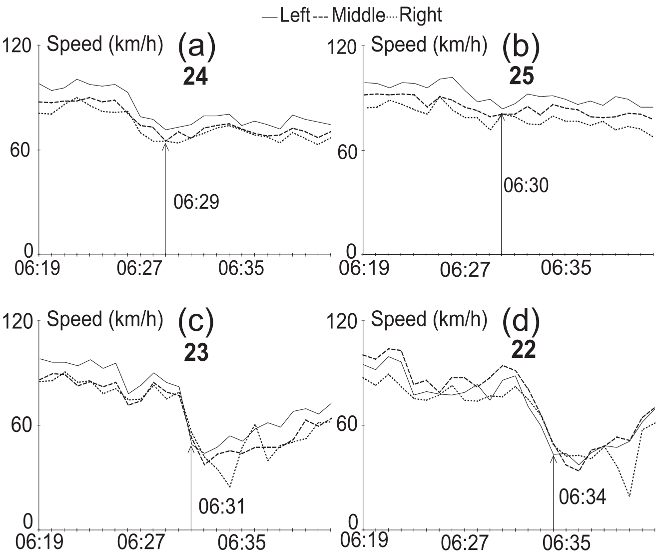

FIG. 2. Time series of vehicle speeds at merge bottleneck. (a) Station 24. (b) Station 25. (c) Station 23. (d) Station 22.

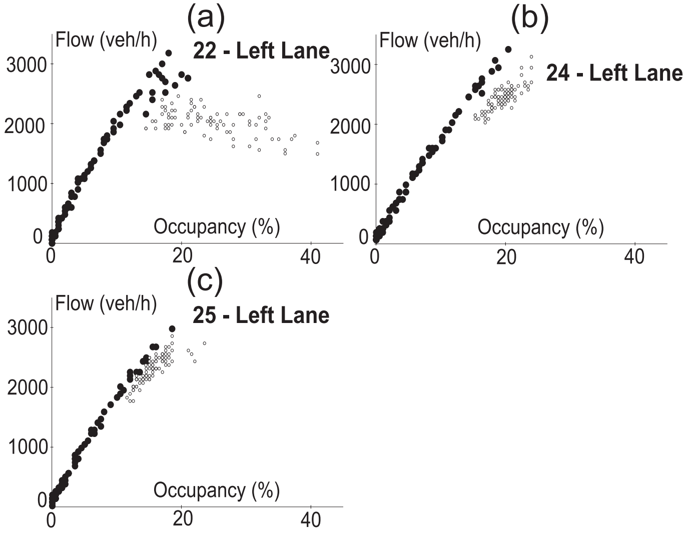

FIG. 3. Occupancy versus flow at merge bottleneck, 6:00 a.m. to 7:54 a.m. (a) Station 22. (b) Station 24. (c) Station 25.

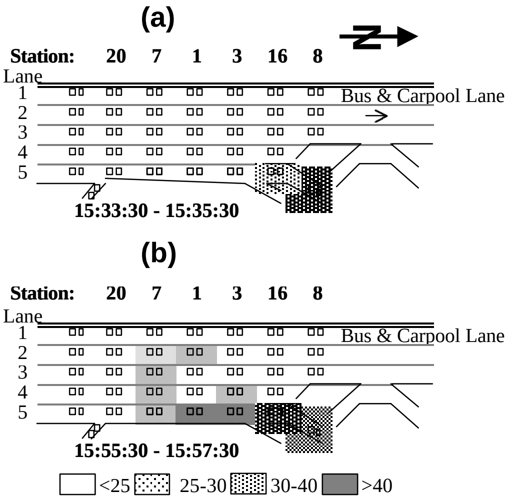

FIG. 4. Queue Evolution at a Diverge Bottleneck

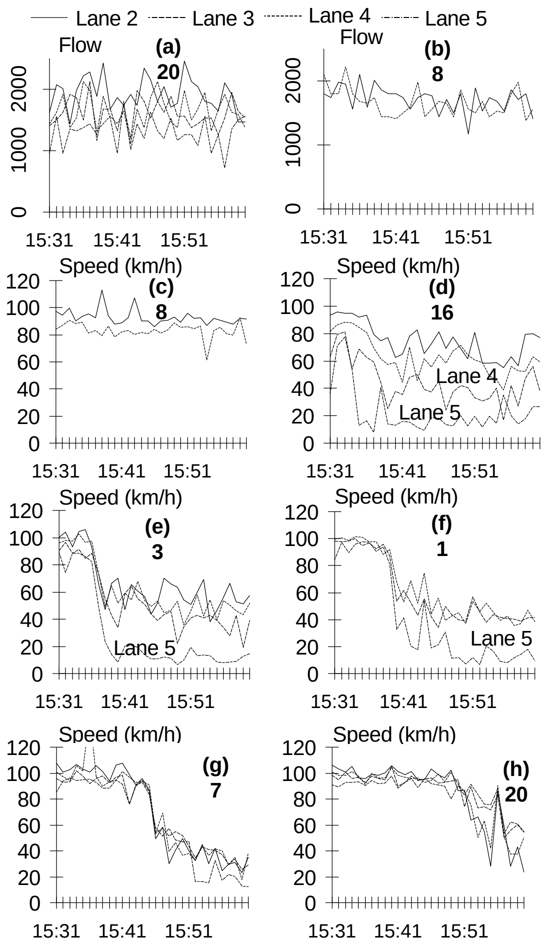

FIG.5. Time series of flows and vehicle speeds at diverge bottleneck. (a) Flow at Station 20. (b) Flow at Station 8. (c) Speed at Station 8. (d) Speed at Station 16. (e) Speed at Station 3. (f) Speed at Station 1. (g) Speed at Station 7. (h) Speed at Station 20.