Transportation Research Part B 45 (2011) 782–788

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

On the macroscopic stability of freeway traffic

Carlos F. Daganzo

Institute of Transportation Studies, University of California, Berkeley, CA 94720, United States

ARTICLE INFO

Article history:
Received 20 October 2010
Received in revised form 7 February 2011
Accepted 8 February 2011

Keywords:
Freeway traffic
Traffic instability
Macroscopic fundamental diagram
Hysteresis

ABSTRACT

A simple model of traffic flow is used to analyze the spatio-temporal distribution of flow and density on closed-loop homogeneous freeways with many ramps, which produce inflows and allow outflows. As we would expect, if the on-ramp demand is space-independent then this distribution tends toward uniformity in space if the freeway is either: (i) uncongested; or (ii) congested with queues on its on-ramps and enough inflow to cause the average freeway density to increase with time. In all other cases, however, including any recovery phase of a rush hour where the freeway's average density declines, the distribution of flow and density quickly becomes uneven. This happens even under conditions of perfect symmetry, where the percentage of vehicles exiting at every off ramp is the same. The flow-density deviations from the average are shown to grow exponentially in time and propagate backwards in space with a fixed wave speed. A consequence of this type of instability is that, during recovery, gaps of uncongested traffic will quickly appear in the unevenly congested stream, reducing average flow. This extends the duration of recovery and invariably creates clockwise hysteresis loops on scatter-plots of average system flow vs. density during any rush hour that oversaturates the freeway. All these effects are quantified with formulas and verified with simulations. Some have been observed in real networks. In a more practical vein, it is also shown that the negative effects of instability diminish (i.e., freeway flows increase) if (a) some drivers choose to exit the freeway prematurely when it is too congested and/or (b) freeway access is regulated in a certain traffic-responsive way. These two findings could be used to improve the algorithms behind VMS displays for driver guidance (finding a), and on-ramp metering rates (finding b).

1. Introduction

It has been proposed (Godfrey, 1969; Herman and Ardekani, 1984; Ardekani and Herman, 1987; Mahmassani and Peeta, 1993; Olszewski et al., 1995; Daganzo, 2007) that plots of average flow vs. average density observed in a city network should follow unimodal curves. It has also been suggested (Daganzo, 2007) that if these curves are reproducible they can be used to build aggregate models of network flow dynamics and control, treating the network as a single reservoir with evenly distributed vehicles. The latter reference also conjectures that curves of this type should be consistently reproduced if the average speed on different parts of the network they describe is evened out by drivers' navigation habits. Geroliminis and Daganzo (2008) shows that indeed, plots of average flow vs. average density for all the streets in Yokohama (Japan) are reproducible and well organized, and called these curves "macroscopic fundamental diagrams" (or MFD's). More recently, Buisson and Ladier (2009) shows that multi-day data from those intersections in Toulouse (France) that have detectors are also well organized .

There are exceptions, however, Buisson and Ladier (2009) also shows that the average network flows were considerably lower on a day when the network was subject to a large non-recurrent disturbance. On that day the data were more scattered

---

E-mail address: daganzo@ce.berkeley.edu

0191-2615/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2011.02.001

and formed a clockwise hysteresis loop. The lower flows and the scatter are attributed in Buisson and Ladier (2009) to the disturbance itself, which misdistributed congestion. But that reference does not explain the timing of the disturbance or the direction of the hysteresis loop.[1] Daganzo et al. (2011) shows, using a 2-bin model of network dynamics, that reduced driver adaptation on the day of the disturbance could also contribute to the lower flows. And, using a generalization of the same model, Gayah and Daganzo (2010) explains both, the scatter and the direction of the hysteresis loop. More specifically, Gayah and Daganzo (2010) shows that if drivers are insufficiently adaptive then congestion must become more uneven as the rush progresses and that as a result, flows for any given density are naturally lower when congestion is declining than when it is increasing. This insight strongly suggests that on any network where drivers do not, or cannot, adapt (the latter will happen if the network lacks redundant routes) clockwise hysteresis loops must arise on MFD scatter-plots.

Although most freeways systems lack sufficient route redundancy to guarantee an even distribution of congestion, freeway MFDs have been studied anyway. On the theoretical side, Cassidy et al. (2010) argues that if all the lanes of a freeway system are in the same congested or uncongested regime everywhere along their lengths, then the freeway system must exhibit a roughly triangular MFD. These single-regime conditions are stringent, however, and cannot be expected to hold for large systems. Indeed, well defined freeway MFDs have only been observed so far on systems with up to four ramps (Cassidy et al., 2010). Simulations of larger systems (Ji et al., 2010) show that portions of a freeway system often are congested while others are not, and the MFD is not well defined. Empirical data for even larger systems also reveal similar fragmentation and scatter (Endo et al., 2010), as well as large clockwise hysteresis loops as in the Toulouse street network (Geroliminis and Sun, 2010).

It is therefore valid to ask whether the forces that drive the distribution of congestion in street networks are also at work in freeway networks. The answer is obvious in some cases. For example, if all the traffic on a long homogeneous freeway flows through a single bottleneck, then this freeway's aggregate MFD data will exhibit clockwise hysteresis during any demand-driven episode of bottleneck queuing.[2] However, for freeways with distributed destinations the answer is not as obvious, especially when these systems are translationally symmetric.

Therefore, to look into this issue, the spatio-temporal distribution of congestion on a freeway of this type is studied below. The freeway is modeled with the kinematic wave theory of traffic flow (Lighthill and Whitham, 1955; Richards, 1956) using the boundary conditions for entrances and exits of the cell transmission model (Daganzo, 1995). This setup describes a freeway system more realistically than the 2-bin model.[3] In order to isolate the issue at hand the freeway will be a rotationally symmetric ring with even demand all around. This is the same system analyzed in a freeway gridlock study (Daganzo, 1996) which assumed congestion remains evenly distributed along the ring as gridlock develops.

It is shown below that while congestion does remain evenly distributed as it builds, it cannot stay even as it dissipates. Congestion must invariably fragment during this recovery phase. This fragmentation process reduces flow, retards full recovery and produces clockwise hysteresis loops, explaining their persistence in empirical data. More importantly, congestion's natural tendency toward fragmentation means that allowing a freeway beltway to become congested all around is more problematic than previously thought. Section 2 of the paper describes the behavior of the ring when it is in a single regime; Section 2 the transitions between regimes and the fragmentation process; Section 3 the effect of adaptive driver diversion to parallel routes; and Section 4 some practical implications and conclusions.

2. Ring performance in a single regime

Considered is a ring of length L on which traffic follows the kinematic wave (KW) theory with a triangular fundamental diagram as proposed in Newell (1993). The parameters of this model are labeled: v_f (free-flow speed), q_o (capacity), k_o (optimum density), κ (jam density) and $w > 0$ (backward wave speed). Ramps are uniformly distributed along the ring. Their flows are modeled as continuously distributed fluxes measured in units of flow per unit length of freeway (veh/h km). This continuum approximation is most accurate when individual ramps carry small flows compared with the freeway. The exit fluxes and influxes at particular locations are determined with the boundary conditions of the CTM (cell transmission model); see Daganzo (1995, 1996).

The demand influx from on-ramps, r (veh/h km), is assumed to be independent of location, x. The boundary conditions for on-ramps stipulate however that the actual influx is a fixed fraction of the freeway flow if the on-ramps are queued and the freeway is congested. This fixed fraction is denoted a (km^{-1}). Clearly then, although the demand is location-independent, the actual input flow will depend on x if the on-ramps are queued and the freeway is unevenly congested.

Off-ramps are assumed to have enough capacity to discharge everyone who wants to leave. It is also assumed that vehicles take exponentially distributed trips with average length l so that the fraction of traffic exiting per unit length of freeway,

---

[1] Clockwise hysteresis loops have been observed upstream of active bottlenecks due to a 'capacity drop' phenomeonon (Cassidy and Bertini, 1999; Bertini and Leal, 2005), which has been theoretically explained in Laval and Daganzo (2006).

[2] The reason is that the queue that forms upstream of the bottleneck will be in a fixed congested state with flow equal to the bottleneck's capacity, while the freeway's uncongested upstream portion will have higher flows when the queue is growing (demand greater than capacity) than when it is receding (demand less than capacity).

[3] Ahn et al. (2011) theoretically show using the same assumptions as our setup (and verify with real data) that off-ramps amplify disturbances in congested flow, while on-ramps dampen them. This real-world phenomenon is at the heart of the macroscopic analysis about to be described.