The Statistical Physics of Traffic Flows

James Cook

May 7, 2004

Abstract

The growth of cities in industrialized nations has created a pressing need to understand traffic and the problems inherent in many-particle self-driven systems. The methods of statistical physics and nonlinear dynamics have been used to understand many puzzling phenomna that have been observed in traffic flows. Specifically, the phase diagram of traffic flow has been predicted, and all of these phases have been observed in real world traffic flows. I will review the progress made on particle based "microscopic" models and gas-kinetic based "macroscopic" models of traffic flow. I will review how these models give qualitative explanations for the observed states of traffic flow.

1 Introduction

The flow of traffic is a complex system with rich dynamics that many people face on a daily basis. Real world congested traffic flows are non-equillibrium systems with self-driven particles that have been observed to have long-range order due to short range interactions. In fact, traffic jams are prototypical examples of long-range order emerging from short-range interactions. A driver is only influenced by the cars that he can see, and most heavily by the one directly in front of him. However, traffic jams in major cities often extend over several kilometers, well outside the view of any one driver, and contain rich dynamics. Longe-range velocity correlations have also been observed in congested traffic flows. While a mathematical description of congested traffic might appear difficult to obtain due to the uncertainty associated with human behavior, an understanding of human motivations may allow us to quantify a majority of human psychological factors in such a way that the fluctuations are small compared to the predictable responses to the narrow range of situations that arrise while driving on a highway. The study of this system may also help develope the techniques necessary to study similar more compicated systems where motivations may not be as well understood, such as, pedestrian and animal traffic in 2-dimensions and the motion of flocks of birds and schools of fish in 3-dimensions.

As early as 1959 Greenberg wrote "The volume of vehicular traffic in the past several years has rapidly outstripped the capacities of the nation's highways. It has become increasingly necessary to understand the dynamics of traffic flow and obtain a mathematical description of the process." [1] The situation has only gotten worse in the last 45 years. Most major cities in the United States and Europe have congested traffic around the clock. The amount of time that the average driver spends in traffic jams is several days a year [7]. There have been jams with lengths more than 100 km in Europe during holiday seasons [7].

Congested traffic also has serious economic and environmental costs. Congested traffic costs an estimated $100 billion each year in econimic loses due to lost time, but, perhaps even more important, is that it causes a similar cost due to accidents and pollution. Vehicle emissions and noise pollution are now at levels comparable to industrial production. Automobile manufacturers worry about how congested traffic will affect their future market, and have invested considerably in research on traffic dynamics. For these economic and environmental reasons, it is important to have an understanding of how traffic jams and congested traffic form, so that measures can be taken to reduce their effects.

The goal of this paper is to review the progress made in modeling traffic and to discuss how detailed mathematical models can explain many effects observed in congested traffic. Specifically, I present a description of how the techniques of statistical physics and nonlinear dynamics have been used to provide mathematical models of traffic flows, and discuss how well these methods have been able to reproduce empirically observed phenomena in traffic flows. The phenomena that I discuss include the current observed states of traffic flows and transitions between these states. The empirical observations are made with several types of detector, but the data that is most plentiful, and therefore, most analyzed is taken from induction-loop detectors. Currently, these detectors can measure the number of vehicles crossing in a given time interval, the time a vehicle spends in the detector, and some even can measure the vehicle velocities.

This paper reviews microscopic (modeling individual car motions) and macroscopic (modeling bulk flow quantities) models of traffic and how well they predict empirically observed phenomenon in congested traffic. In section II, I discuss the states of traffic seen in empirical data. In section III, I discuss the basics of constructing microscopic models of traffic flows, and the basics of constructing macroscopic models of traffic flows. In section IV, I summarize and discuss the results presented in the previous two sections.

2 Phenomena Observed in Real World Traffic Flows

In this section, I review some of the complex dynamics empirically observed in real world traffic flows. First, I introduce some common terminology in traffic flows and discuss how data is collected for these observations. Next, I review the generic properties of a typical data set taken on Friday, August 25, 1995 on a section of highway in Germany [2]. Finally, I discuss even more states of congested traffic that have been observed in real world flows.

2.1 Terminology and Data Collection

This discussion of termilonlogy and detectors closely follows that found in [7]. Most data on traffic flows are obtained by induction-loop detectors. These detectors measure the number of crossing vehicles N in a time T which can be changed, the times t^0_α and t^1_α which are the times a vehicle enters and leaves the detector respectively, and some measure individual vehicle velocities v_α and the vehicle lengths l_α (in this brief description I will label individual vehicle properties with a subscript α and leave bulk flow properties without a subscript). From these measurements people can easily construct the following quantities: the time headways (also called the gross or brutto time separations)

$$\Delta t_\alpha = t^0_\alpha - t^0_{\alpha-1}, \tag{1}$$

the time clearances (also called the netto time separations)

$$\Delta t_\alpha = t^0_\alpha - t^1_{\alpha-1}, \tag{2}$$

the headways (also called the brutto distances)

$$d_\alpha = v_\alpha \Delta t_\alpha, \tag{3}$$

and the clearances (also called the netto distances)

$$s_\alpha = d_\alpha - l_{\alpha-1}. \tag{4}$$

These quantities all have pretty simple meanings. The time headway is the amount of time between the time that vehicle $\alpha - 1$ enters the detector and vehicle α enters the detector. The time clearance is the amount of time between the time that vehicle $\alpha - 1$ leaves the detector and vehicle α enters the detector. The headway is the distance between the front of vehicle $\alpha - 1$ and the front of vehicle α when vehicle $\alpha - 1$ enters the detector (assuming vehicle α travels at a constant velocity). The clearence is the distance between the rear of vehicle $\alpha - 1$ and the front of vehicle α when vehicle $\alpha - 1$ enters the detector (assuming vehicle α travels at a constant velocity). These values frequently enter into microscopic descriptions of traffic flows. For macroscopic descriptions, new quantities must also be defined from the data sets.

The macroscopic quantities are all fairly straight forward except the density. The macroscopic quantities frequently used are vehicle flow, time occupancy, average velocity, and velocity variance, and they are defined as

$$\begin{aligned} Q(x,t) &= N/T, & (5)\\ O(x,t) &= \Sigma_\alpha (t^1_\alpha - t^0_\alpha)/T, & (6)\\ V(x,t) &= \langle v_\alpha \rangle = 1/N \Sigma_\alpha v_\alpha, & (7)\\ \theta(x,t) &= \langle [v_\alpha - \langle v_\alpha \rangle]^2 \rangle. & (8) \end{aligned}$$

One would also like to specify a vehicle density, but for various reasons (see [7]) there a few different ways that are common in the literature. The ways

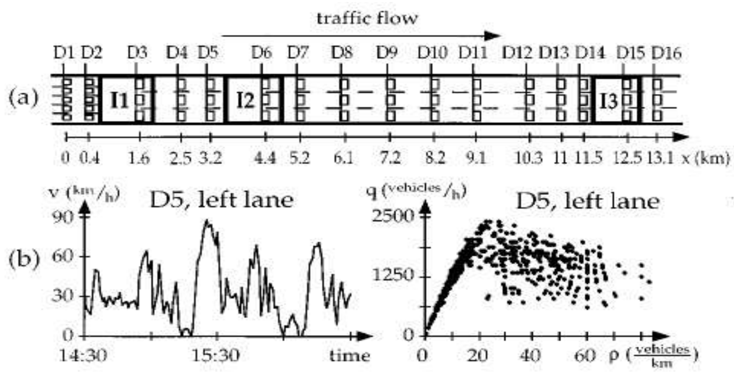

Figure 1: (a) Schematic of the detector configuration on the highway chosen. (b) the average velocity of vehicles for a time sample (left) and the measurement points on the flux-density plane for a large time sample (7:00-22:00) (right) at the detector D5 (left lane). Figure reproduced from [2].

macroscopic vehicle density is commonly defined are

$$\rho(x,t) = Q(x,t)/V(x,t), \tag{9}$$
$$\rho(x,t) = O(x,T)/[L(x,t)+L_D], \tag{10}$$

where $L(x,t)$ is the average vehicle length during the measurement interval and L_D is the detector length. Helbing [7] suggests that the best way to construct a vehicle density is to use equation (9) with the definition

$$\frac{1}{V(x,t)} = \left\langle \frac{1}{v_\alpha} \right\rangle. \tag{11}$$

When discussing macroscopic empirical results, I will specify which definition of vehicle density is being used, when discussing macroscopic theoretical models, it is not always possible to distinguish the different density definitions, but where it is possible, I will specify the definition used.

2.2 Emperical Properties of Complexity in Traffic

First, I will explore the typical states of traffic flows by examining the data from Friday, August 25, 1995 following the discussion in [2]. Then proceed to a more thorough description of congested traffic drawn from [6]. All vehicle densities discussed in this section were obtained using the density method recommended in [7].

Figures 1, 2, and 3 displays the broad states generally found in traffic flows. This data was taken with induction loop detectors which give a measurement of the average velocity of vehicles crossing the detector and the flux of vehicles crossing the detector for each of three lanes [2]. The road on which this data

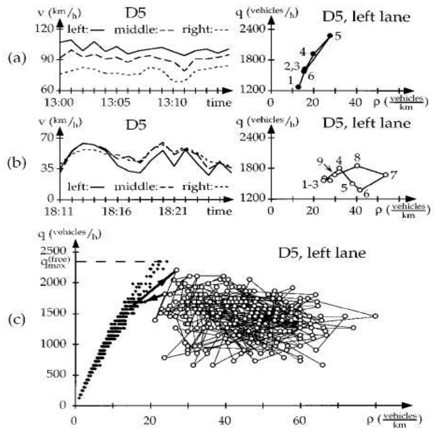

Figure 2: Distinguishing two phases of traffic: (a) and (b) are the average velocities of vehicles by lane (left) and experimental points on the flux-density plane (right) for free and sychronized traffic flow respectively. Figure reproduced from [2].

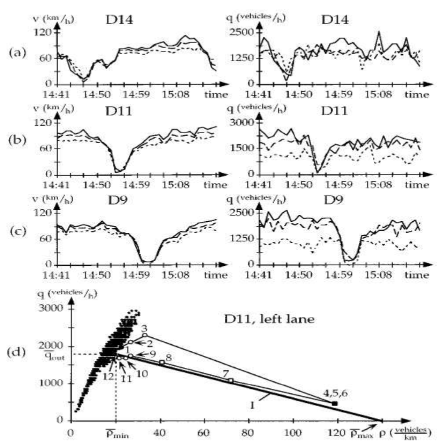

Figure 3: Exploring a typical traffic jam: (a)-(c) show the kinematics of the jam as it propagates upstream from detector D14 to detector D9. (d) shows the difference between a jam and the other phases of traffic in the flux-density plane. The solid thick line I represents the downstream front of wide jams for the cases when no hindrance exists in the outflow from the jam [3]. Figure reproduced from [2].

was taken was a section of highway A5 between Bad Homburg and Frankfurt in Germany [2]. The layout of the highway can be seen in figure 1 (a) where D1-D16 are the detectors and I1-I3 are intersections.

Figure (1) (b) shows a great deal of complexity in both the average velocity and scattering of data in the flux-density plane. There is one striking feature in the flux density plane and that is the very linear relationship up to a density of 20vehicles/km and reaching the maximum flux of about 2500vehicles/h. This certainly suggests some sort of phase transition.

Figure (2) (a) and (b) separates the data from these two regimes. The regime corresponding to the linear relation in the flux-density plane (see Figure [2] (a)) is frequently called free traffic flow, because it is associated with drivers being able to change lanes and reach their desired speed due to low desnities. One key feature of free traffic flow is that each lane has a different average speed, because faster drivers tend to stay on the left and slower drivers tend to stay on the right. Now notice the striking differences in the phase of traffic flow known as synchronized traffic flow (see Figure [2] (b)). In this phase, there appears to be essentially random motion in the flux-density plane. The most interesting feature is the fact that the average vehicle velocity is now the same in all three lanes (hence the name synchronized traffic flow). The interptretation of this effect is that the density is so high, that lane changes now tend regulate the flow between the lanes instead of allowing vehicles to freely increase to their desired speed. Also of note is that the increased density has caused a marked decrease in the average speed from free traffic flow (compare Figure [2] (a) with (b)). Finally, (c) shows all the data from 7:00 to 22:00 and the transitions from free to synchronized flow and back marked. All traffic jam data has been removed from the above data, because I wanted to focus on the transition from free traffic flow to congested state of traffic flow.

Now I will show the characteristic of traffic jams by discussing a particular jam occuring in the data set. A traffic jam is place where the vehicle velocity goes very close to zero. Figure (3) (a)-(c) shows such a jam progressing upstream from detector D14-D9. This upstream motion of the jam is typical. The speed of the jam front is about -15km/h [2]. Even though this jam is surrounded by synchronized flow (as can be seen from the lane specific vehicle velocity) in general jams can have either types of flow upstream or downstream (i.e. jams caused by narrowing of the road due to accidents frequently have free flow at both ends) [2]. Figure (3) (d) shows the jam on the flux-density plane. This is qualitatively very different from either free or synchronized traffic flows having a downard sloping relatively linear shape in the high density, low flow region of the diagram. This jam originated in intersection I3 (see Figure [1]). Notice that the jam starts very small and grows as it travels upstream, which suggests a phase transition that occurs through nucleation. This covers the typical states found in traffic flows.

A more detailed analysis of the congested traffic regime has found a number of different types of congested traffic [6]. They are usually excited by different situations, so here I will briefly catalogue their characteristics and the conditions under which they form (note that all these flows occur in the high density range and so typically emerge from synchronized flows discussed above). Homogeneous congested traffic is characterized by vehicle velocites that are essentially constant over large regions of space and is usually caused by a sudden decrease in the flow capacity (such as a vehicle accident blocking a lane). Oscillating congested traffic, usually occuring at bottleneck inhomogeneities in the vehicle capacity (i.e. a fixed reduction in lane number), is characterized by average vehicle velocities that are periodic in time and nearly constant in space, however, the nonlinear nature of traffic flows causes a number of modes to be visible in the fourier spectrum. Just as a note, the two previously discussed states have been observed in coexistence where a sudden lane closure due to an accident occured at a steep uphill gradient, which reduces the flow capacity at all times. Pinned and moving localized clusters in free traffic flows are states that look very similar to synchronized traffic in the flux-density diagram, but do display the drop in average velocity associated with the transition from free to synchronized flow. They're causes can be varied, and often moving localized clusters appear to be cause by nothing more than density fluctuations. Indeed, these states lie very near the transition between free and synchronized flow in the flux-density plane, signaling that the phase transition occurs through a nucleation process.

While many interesting phenomena have been observed empirically using the loop induction detectors, there are better ways to record traffic flows. The time interval of measurements taken with loop induction detectors acts like a coarse graining in time, and can give spurious behavior if the time scale is too short and miss critical behavior if the measurement times are too long. Furthermore, they provide no real information on the path of specific vehicles or where vehicles are changing lanes. Data based on aerial photography and video recordings allow many interacting vehicles to be tracked over long distances. Specifically lane-changing maneuvers and individual vehicle velocities may be tracked for a specific group of vehicles rather than obtaining just bulk flow properties [7]. Another way to get interesting data is to actually put a car with instruments to measure the velocity, acceleration, and distance to the nearest neighbors [7]. However, this data is not as available for study.

3 Theoretical Models of Traffic Flows

In this section I review the techniques used to construct microscopic and macroscopic models of traffic flows. First, I briefly go over what a good model of traffic flow should contain. A good theory of traffic flow should only contain a few pararmeters that can be interpreted intuitively. Ideally, these parameters could be measured independently of the flow (i.e. not used to fit the flow, but general properties of known human reactions). Finally, the model should not cause accidents (while they do occur in real life, there shouldn't be common situations that cause massive accidents in a believable model). Finally, the model should predict the observed state of traffic at roughly the correct flux and density (This is discussed in the Results and Conclusions section).

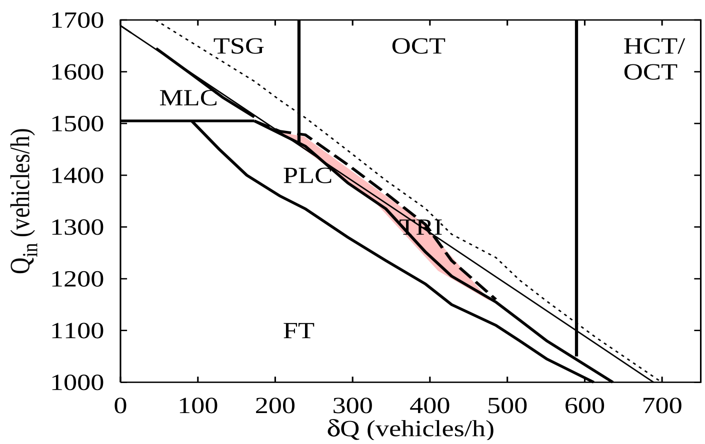

Figure 4: Typical phase diagram of traffic flow states resulting from an intelligent driver method calculation. FT stands for free traffic. PLC stands for pinned local clusters. MLC stands for moving local clusters. TSG stands for triggered stop-and-go waves. TRI is a state where all states on the boundary are metastable and seen in patches. OCT is oscillating congested traffic. HCT is homogeneous congested traffic. Figure reproduced from [6].

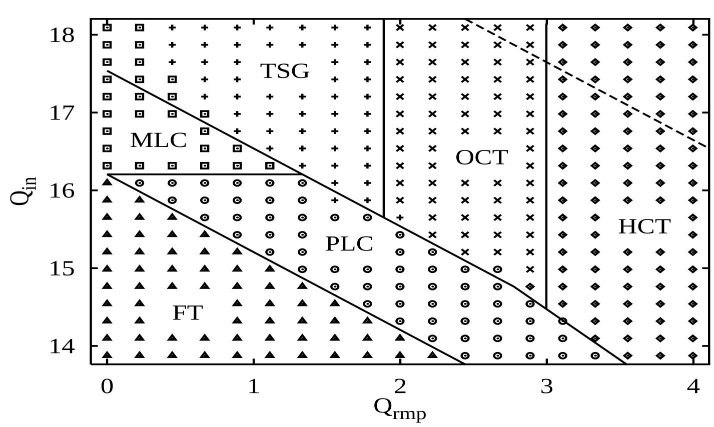

Figure 5: Typical phase diagram of traffic flow states resulting from an non-local gas-kinetic-based traffic model. All region names have the same meaning as in figure (4). Figure reproduced from [4].

3.1 Microscopic Models of Traffic Flows

Generally, the equation of motion for self-driven one-dimensional system is

$$\frac{dv_\alpha(t)}{dt} = \frac{v_\alpha^0 - \xi_\alpha(t) - v_\alpha(t)}{\tau_\alpha} + \sum_{\beta \neq \alpha} f_{\alpha,\beta}(t), \tag{12}$$

where v_α^0/τ_α is a self-driving term, $\xi_\alpha(t)/\tau_\alpha$ includes fluctuations, v_α/τ_α is a friction term, and $f_{\alpha,\beta}(t)$ is the interaction of particle β on particle α [7]. In the absence of the other terms, the driving and friction terms lead to an exponential adaption to the desired velocity v_α^0. Obviously, the interaction term does not obey Newton's third law, because people do not behave in the same way when approaching a vehicle from the rear as when a vehicle approaches them from the rear. In traffic flow problems, a follow the leader approach is taken where $f_{\alpha,\beta}(t) = 0$ for $\beta \neq \alpha - 1$. This assumption will be made throughout this discussion.

The models based on this technique are too numerous to catalogue, so I will now concentrate on one specific model called the intelligent driver model [5]. In this model, the drivers acceleration tendancy in taken as $a_\alpha[1 - (v_\alpha/v_\alpha^0)^\delta]$ and $f_{\alpha,\alpha-1} = -a_\alpha[s_\alpha^*(v_\alpha, \Delta v_\alpha)/s_\alpha]^2$, where *delta* is a model parameter, s_α^* is a desired clearence, s_α is the actual clearance, and $\Delta v_\alpha = v_\alpha - v_{\alpha-1}$ is the approaching rate. This gives

$$\frac{dv_\alpha(t)}{dt} = a_\alpha \left[1 - \left(\frac{v_\alpha}{v_\alpha^0}\right)^\delta - \left(\frac{s_\alpha^*(v_\alpha, \Delta v_\alpha)}{s_\alpha}\right)^2\right], \tag{13}$$

as the equation of motion. The form of $s_\alpha^*(v_\alpha, \Delta v_\alpha)$ still needs to be specified. Treiber and Helbing take

$$s_\alpha^*(v_\alpha, \Delta v_\alpha) = s'_\alpha + s''_\alpha \sqrt{\frac{v_\alpha}{v_\alpha^0}} + T_\alpha v_\alpha + \frac{v_\alpha \Delta v_\alpha}{2\sqrt{a_\alpha b_\alpha}}, \tag{14}$$

where T_α is the safe time clearance, a_α is the maximum acceleration, b_α is the comfortable deceleration, and s' and s'' are the jam distances. For simplicity they take the parameters to be the same for all vehicles and will always consider $s'' = 0$. A possible improvement would be to give each driver values for b_α, T_α, a_α, v_α^0, and s' chosen with a gaussian probability distribution (or even an empirically determined distribution). I will note that for some choices of parameters, analytic equillibrium solutions to this equation have been found, but they are not particularly enlightening.

In [6], this model is used to construct a phase diagram of traffic states (see Figure 4). The state was determined by letting the model run for an amount of time where a steady state was achieved (except for the TRI models which never achieved a steady state due to the believed metastability of three states in that region). The congested states appeared to form via nucleation where a region of that state would form due to fluctuations and then grow.

This method for modeling traffic is computationally slow. There have been efforts to use cellular automata to model traffic flows. These efforts give similar qualitative results and are much quicker to run. However, the above method is a much more accurate representation of traffic flow, because in the cellular automata calculations thus far space, time, and velocity must all be discretized. However, cellular automata include finite size effects that are neglected in the above model.

3.2 Macroscopic Models of Traffic Flows

The developement of a macroscopic traffic flow is a much more technical and less intuitave process, so I am going to give only a sketch of the proceedure and present the typical resulting equations. There is an obvious equation for the vehicle density that is simply a statement of conservation of vehicle number in the form of a continuity equation

$$\frac{\partial\rho}{\partial t}+\frac{\partial\rho V}{\partial x}=Q_{\mathrm{source}}, \tag{15}$$

where Q_{source} is a sink/source term that might be used to simulate particles leaving/entering at on ramps. This must be supplemented with an equation analogous to the Navier-Stokes equation for the average velocity V. This equation is usually derived from a Boltzmann-like distribution, however, the phase-space density is not conserved due to self-energy production and asymmetric interactions. Functional forms are then picked for these two effects, and one then derives an analogue to a Fokker-Planck equation. For a more detailed discussion of the construction of these theories, see [7]. The specific model I am going to present the results for is called the nonlocal, gas-kinetic-based traffic model [4]. The non-dimensionalized Navier-Stokes like equation in this model is

$$\frac{\partial V}{\partial t}+V\frac{\partial V}{\partial x}=-\frac{1}{\rho}\frac{\partial\rho\theta}{\partial x}+(V_0-V)-PA(\rho_a)\frac{(\rho_a V_a)^2}{(1-\rho_a)^2}B(\delta_V), \tag{16}$$

where V_0 is a bulk desired velocity, P is a cross section, θ is the velocity variance, $A(\rho)=0.171+0.417\{\tanh[10(\rho-0.27)+1\}$ represents a structure factor, and

$$B(\delta_V)=2\left[\delta_v\frac{e^{-\delta_V^2/2}}{\sqrt{2\pi}}+(1+\delta_V^2)\int_{-\infty}^{delta_V}dy\frac{e^{-y^2/2}}{\sqrt{2\pi}}\right] \tag{17}$$

is a Boltzmann factor arising from vehicle interactions. The index "a" indicates the the quantity is evaluated at an advanced "interaction point" $x_a=\gamma(1+TV)$, where γ is an anticipation factor and T in the safe time headway.

Helbing *et. al.* used the nonlocal gas-kinetic-based traffic model to construct a phase diagram of traffic states (see Figure 5). The state was determined by introducing a perturbation on top of a uniform flow and waiting to see the final steady state. The congested states appeared to grow from this initial seed perturbation, suggesting that the phase transition occurs through nucleation.

The lines in the diagram are analytic determinations of the phase boundaries made with this model.

This method for modeling traffic is computationally slow and conceptually further removed from the simple analysis of the microscopic traffic models. However, the macroscopic models' advantage is that they are typically better suited for analytically probing the system, which is evident from the analytic calculation of the phase boundaries in the diagram.

4 Discussion

Both the microscopic and macroscopic models discussed seem to qualitatively reproduce the types of traffic seen in empirical traffic flows. The same states of congested traffic appear in both theoretical models, and phase space locations of the boundaries are in good agreement (see Figures [4]and [5]). However, while all of the states predicted theoritically have been observed empirically, the phase space has not yet been fully explored meaning, that the phase boundaries have not all been mapped out with empirical observations. There is still alot of empirical data taken with induction loop detectors that can be explored, as well as, better techniques to collect data, such as the aerial video data and car following methods discussed earlier.

The theoretical models can also be significantly advanced. First, as mentioned briefly in the section III, the current microscopic and macroscopic models allow for no variation in driver. A sensible next step is to empirically collect data on reaction times, comfortable accelerations/decelerations, prefered velocity, safe time headways, and reactions to velocity gradients. This data could then be used to construct a model where the distribution of driver properties obey the distributions found empirically. Furthermore, one could correlate the properties appropriately (i. e. a driver that likes to drive faster might also be more likely to have a lower safe time headway). Another step would possibly be to try and determine what really affects driver behavior through surveys. This might reveal that local density plays an important role in driver responses. These ideas could be applied to both macroscopic and microscopic models in a fairly straight forward way. Finally, different vehicles need to be included in the simulations. In the models discussed thus far, the vehicle size was taken to be zero. It will be interesting to see how finite vehicle size and including different size vehicles influence the phase diagram of these models. These vehicle size features are also lacking from the macroscopic models, but it is far less obvious how to include these effects than it is to include them in the microscopic models.

Traffic flow models have been successful thus far in qualitatively describing the different phases present in real world traffic flows. Still, more data needs to be examined and collected to determine quantitative success of the theoretically predicted phase diagrams of traffic flows. The models themselves are still at a fairly early stage and could probably benefit from some simple additional features.

References

[1] Greenberg, H., Oper. Res. **7** 79 (1959).

[2] Kerner, B. S., and H. Rehborn, Phys. Rev. E **53**, R4275 (1996).

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

[4] Helbing, D., A. Hennecke, and M. Treiber, Phys. Rev. Lett. **82**, 4360 (1999).

[5] Treiber, M., and D. Helbing, e-print cond-mat/9901239, (1999).

[6] Treiber, M., A. Hennecke, and D. Helbing, Phys. Rev. E **62**, 1805 (2000).

[7] Helbing, D., Rev. Mod. Phys. **73**, 1067 (2001).