The Emergent Behaviour of Traffic

Joseph B. Altepeter*
Department of Physics, University of Illinois,
1110 W Green St, Urbana, IL, USA 61801
(Dated: May 8, 2004)

The dynamics of freeway traffic, though fundamentally dependent on the complex interactions of many human agents, exhibit emergent behaviour which can be modelled using the machinery of statistical physics. This observed behaviour can be unintuitive, as in the case of "phantom traffic-jams", whereby fast-moving traffic suddenly congeals into a slow-moving jam for no apparent reason. Using models based both on the individual motivations of each driver and on macroscopic variables describing traffic flow as a whole physicists have been able to accurately reproduce the dynamics of real freeway traffic flow, including these surprising experimental observations.

*Electronic address: `altepete@uiuc.edu`

I. INTRODUCTION

The study of traffic is an area of active research almost 60 years old [1] the implications of which can be seen in virtually every aspect of modern life. Not only are freeways essential to modern intercity trade, intracity roadway traffic is ubiquitous in the developed world. While the widespread nature of traffic is certainly motivation for its study, there are significant economic gains to be made by understanding (and eliminating) inefficiencies due to congested traffic. In the United States alone the average annual loss due to traffic jams exceeds $65 billion for the country and 60 hours for each driver [2].

In addition to the direct applicability of this research to freeway systems, there are a number of related fields which can benefit from and influence these studies: pedestrian traffic, both on organized walkways and in emergency situations (e.g., many people trying to escape a burning building) [3]; flocking and herding of birds and other animals [4]; economic systems, such as stock market oscillations [5].

This paper attempts to summarize the basics of traffic flow. Section II discusses the empirical observations of freeway traffic, focusing on both interesting individual phenomena (e.g., phantom traffic-jams) and those characteristics addressed by modern traffic models. Section III discusses the modelling of traffic flow, beginning with the distinction between microscopic models (which treat each driver as an individual particle) and macroscopic models (which deal with system-wide variables and effects. This section will then examine attempts to link these two regimes and conclude with a discussion of Ginzburg-Landau theory as applied to traffic.

II. EMPIRICALLY OBSERVED DYNAMICS OF FREEWAY TRAFFIC

A. Measuring freeway traffic

Several measurement techniques are available for measuring traffic flow, vehicle speed and acceleration, and lane-changing maneuvers. The most accurate and complete method involves aerial photography or video-capture, a technique first introduced in the 1960's [6, 7]. This allows a complete analysis of each vehicle's behaviour, but requires active human involvement for both filming and analysis.

Another method which requires human action is car-following, by which a single car equiped with a variety of detectors monitors the behaviour of a vehicle by shadowing or following its movements. This gives less information that aerial methods and is more difficult that automated methods.

The easiest way to collect information on traffic patterns uses automated detectors at freeway cross-sections. By using single induction-loop detectors, one can measure a vehicle's arrival time t^0_α, a vehicle's departure time t^1_α, a vehicle's velocity v_α, and a vehicle's length l_α, for each vehicle α. By analyzing these results for multiple vehicles one can find the number of vehicles ΔN that cross the detector during a sampling interval ΔT.

These measured values allow us to derive a number of useful parameters. The *time headway*, $\Delta t_\alpha = \left(t^0_\alpha - t^0_{\alpha-1}\right)$, measures the arrival time difference between following vehicles. The *time clearance*, $t_c = \left(t^0_\alpha - t^1_{\alpha-1}\right)$, measures the temporal spacing between vehicles. The *headway*, $\Delta x_\alpha = v_\alpha \Delta t_\alpha$, measures the spatial distance between following vehicle arrivals while the *clearance* or *netto distance*, $s_\alpha = (d_\alpha - l_{\alpha-1})$, measures the space between vehicles. Of particular importance will be the *vehicle flow rate*,

$$Q(x,t) = \frac{\Delta N}{\Delta T}, \tag{1}$$

which corresponds to the rate that vehicles pass the detector, and the *arithmetic average velocity*,

$$V(x,t) = \langle v_\alpha \rangle = \frac{1}{\Delta N} \sum_{\alpha=\alpha_0+1}^{\alpha_0+\Delta N} v_\alpha. \tag{2}$$

Together, equations 1 and 2 can be used to define the vehicle density $\rho(x,t)$, which will be of central importance to later analysis:

$$\rho(x,t) = \frac{Q(x,t)}{V(x,t)}. \tag{3}$$

These definitions, while they may seem straightforward, lead to a subtle problem. Consider the difference between measuring an ensemble of vehicles that pass a specific point within interval ΔT as opposed to measuring an ensemble of vehicles that at a given instant lie within a stretch of freeway ΔX. Because faster vehicles will pass a single point more often they will more heavily influence the former ensemble average than the latter. This implies a fundamental difference between spatial and temporal averaging. Unfortunately, equation 3 simultaneously depends on both a spatial and a temporal average. In order to correct this

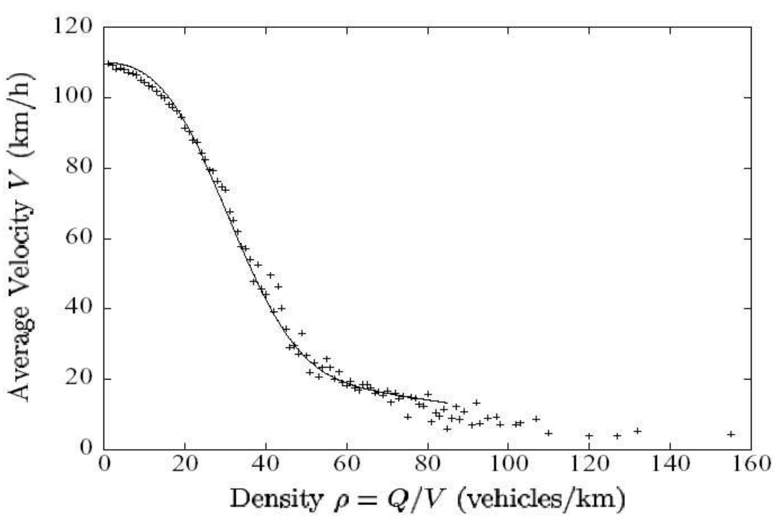

FIG. 1: Comparison of empirical velocity-density relations for different definitions of the average velocity. The 'plus' symbols represent one-minute averages determined using the harmonic velocity formula: $V = \frac{1}{\langle 1/v_\alpha \rangle}$. The solid line is a fit to the same data interpreted using the traditional arithmetic formula, $V = \langle v_\alpha \rangle$ [9].

problem, we replace the arithmetic mean (equation 2) with the *harmonic average velocity*,

$$\frac{1}{V(x,t)} = \left\langle \frac{1}{v_\alpha} \right\rangle, \tag{4}$$

which automatically grants additional weight to low velocities. Figure 1[1] shows a comparison between the arithmetic and harmonic mean velocities. They show good agreement, with the harmonic mean extending to much higher densities. The disadvantage of using the harmonic mean is its sensitivity to errors for small velocities.

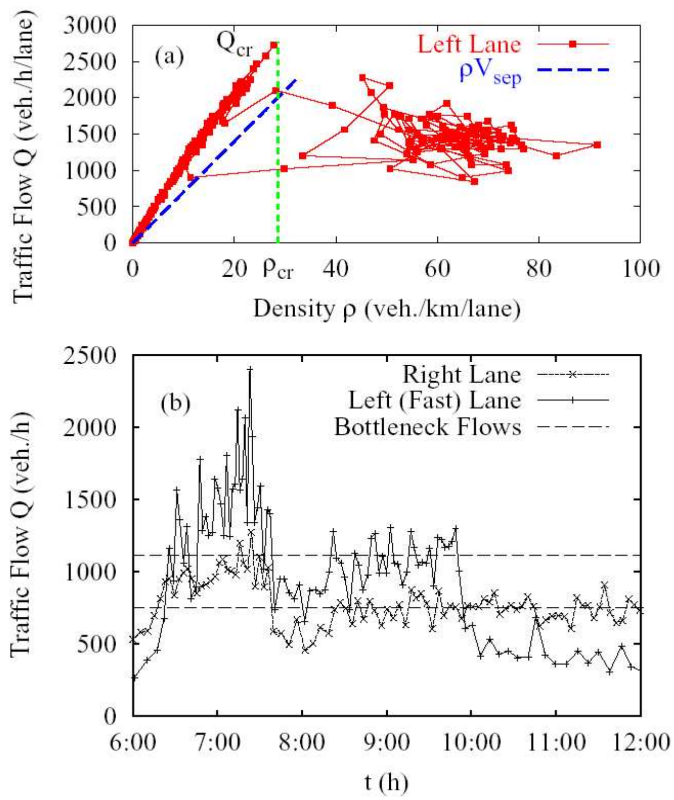

FIG. 2: The fundamental diagram: hysteresis in traffic flow and free versus congested traffic. (a) Traffic-flow time series as a function of density. Notice that in addition to hysteresis this plot shows two distinct regimes corresponding to free traffic flow (linear realationship between flow and density) and congested traffic flow (erratic relationship). These regimes can be separated by the line ρV_{sep}, with $V_{sep} = 70 km/h$. (b) Time series of traffic flow as a function of time. Immediately after breakdown from an unstable high-flow state to a congested state, the flow falls below typical bottleneck levels [8, 10, 11].

[1] Except where noted, figures are taken from [8], Dirk Helbing's article from *Rev. of Mod. Phys.*. The cite after each figure caption refers to the source of the original data.

B. Types of traffic flow

The empirical methods and definitions outlined in section II A provide the methods to gather data and the vocabulary to discuss them. These methods immediately lead to the analysis of the relationship between flow (the primary goal of traffic systems) and density (the primary variable upon which jams are dependent). Empirical results provide us with a *fundamental diagram* of flow as a function of density (see figure 2).

This time series of traffic-flow as a function of density reveals two distinct regimes: free traffic and congested traffic. Free traffic is characterized by a linear relationship between flow and density. This is the behaviour one would expect when the interactions between vehicles are neglected, where doubling the number of vehicles doesn't affect the speed of those vehicles and therefore doubles the flow.

The second regime, congested (also called synchronized) traffic, is far more erratic and less predictable. Real data exhibits hysteretic results widely spread in two dimensions on the flow/density diagram. In addition, it is possible for the density to be reduced under roughly constant flow, implying that the breakdown from free traffic to congested traffic is not reversible. As can be seen in figure 2b, the flow rate after a breakdown can temporarily fall below the bottleneck flow rate (the typical flow-rate for maximum density).

There is no clear division where free traffic breaks down. Instead, high-flow high-density traffic is unstable, becoming more likely to breakdown the longer the state persists and the higher the flow becomes. A plot of breakdown probability as a function of time and flow is shown in figure 3. In addition, the transition from free to congested traffic is discontinuous and exists for a wide range of densities, so that for many traffic densities it is possible to have either free or congested traffic flow.

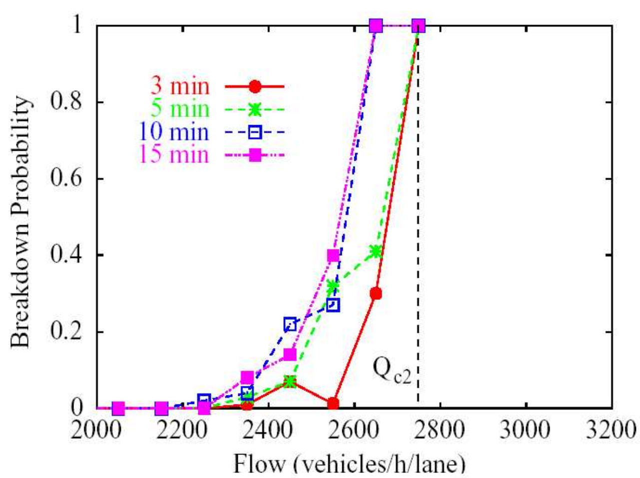

FIG. 3: Probability of free-traffic breakdown as a function of traffic flow. Several waiting times are shown, showing the instability of free-traffic flow at high-flow rates [10].

C. Characteristics of congested traffic

Even though the transition from free to congested traffic looks chaotic when its flow is plotted as a function of density, velocity distributions very closely follow normal distributions, even for densities within the congested-traffic regime (see figure 4). The regularity of these distributions, however, belies the range of phenomena which can occur within congested traffic.

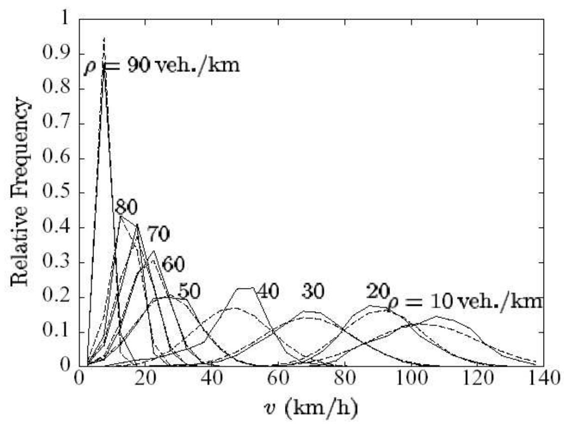

FIG. 4: Empirical velocity distributions for vehicle densities ranging from 10 to 90 vehicles per kilometer. The solid lines represent data while the dashed lines are normal distributions having the same mean and standard deviation [9, 12].

Rather than attempt to describe the full gamut of congested traffic, here we will describe two interesting characteristics: stop-and-go waves, familiar to most drivers, and "phantom traffic jams", wherein steadily moving traffic jams up suddenly and without apparent cause.

1. Stop-and-go waves

The flow of congested traffic can often manifest as periodic waves, resulting in the individual driver alternatingly spending time stationary and in motion. These traffic waves have a period of between 4 and 20 minutes [9, 12] and a wavelength of between 2.5 and 5 km [13]. These waves have no characteristic frequency, but instead are composed of several frequencies. When analyzing the power spectrum of these frequencies one finds white noise at high frequencies and a power law at low frequencies. The strength of a given frequency is approximately proportional to $\omega^{-1.4}$ [14].

It is possible that this power law is indicative of the emergent behaviour of self-organizing criticality [15]. Reference [15] argues that vehicles leaving a traffic jam self-organize by acheiving maximum throughput (all drivers want to leave the jam as quickly as possible). This naturally pushes flow towards its maximum, but allows small fluctuations to induce a breakdown of this high-flow rate, far downstream. In effect, the natural tendency of drivers to reach maximum throughput makes it far more likely that traffic jams will develop.

2. Phantom traffic jams

Empirical evidence (see figure 5) has shown that it is possible for free-traffic to spontaneously breakdown into a "phantom traffic jam", congealing into slow movement with no apparent cause. This is in fact the direct result of the discontinuity in the fundamental flow/density diagram, and represents the metastable high-flow high-density state devolving into the congested low-flow high-density state which coexists on the fundamental diagram at the same ρ. Nevertheless, the lack of an apparent cause for the spontaneous jam (such as an accident or a bottleneck) has been the subject of conjecture in both popular and scientific literature. Some research [16] suggests that these phantom jams are in fact caused by a lane change directly in front of a high density section of vehicles. Even if this is true, it can be unintuitive that such a small disturbance can have such a large effect.

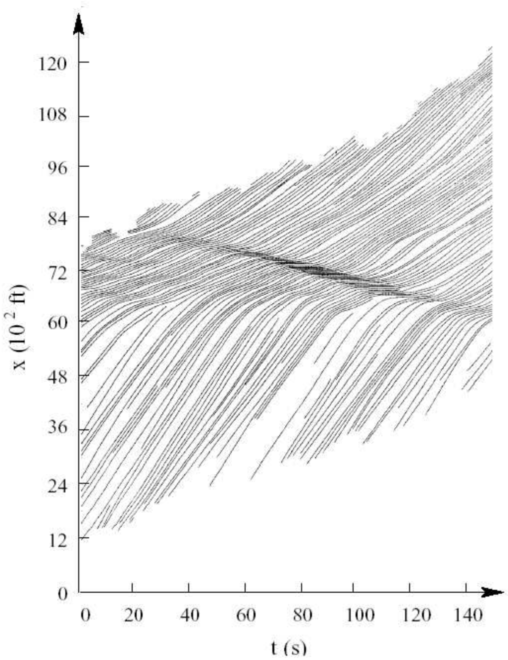

FIG. 5: Data for a "phantom traffic jam," taken using aerial photography. Broken lines indicate lane changes. As the line slopes (vehicle velocities) decrease, the line density (spatial vehicle density) increases, corresponding to spontaneous traffic jam. The jam front then propagates upstream with constant velocity [7, 17].

III. MODELLING TRAFFIC FLOW

A. Microscopic traffic models

A natural first step towards modelling traffic flow is to model the dynamics by which each vehicle is governed. The most basic of these microscopic models are "follow-the-leader" models, where the primary influence on a vehicle's behavior is the vehicle directly in front of it. Each driver keeps a minimum, velocity-dependent safe distance between itself and the vehicle in front of it. In addition, each vehicle accelerates to close a velocity differential and deccelerates quickly to avoid a collision.

This is a good first step, but needs to be immediately adapted to account for delays due to driver reaction time. In this way the initial theory can reproduce empirically observed density waves which are otherwise absent. The next unsettling characteristic of this theory is that in the absense of a leading car, traffic does not move. In other words, there is no way to predict the motion of a lone vehicle. By adapting the above follow-the-leader model to an optimal-velocity model, we assume that there is a preferable velocity that drivers strive for (such as the speed limit), which in the absense of strong outside influence (nearby vehicles) governs acceleration and decceleration. This optimal velocity model can predict excellent results, including the existence of phantom traffic jams (see figure 6).

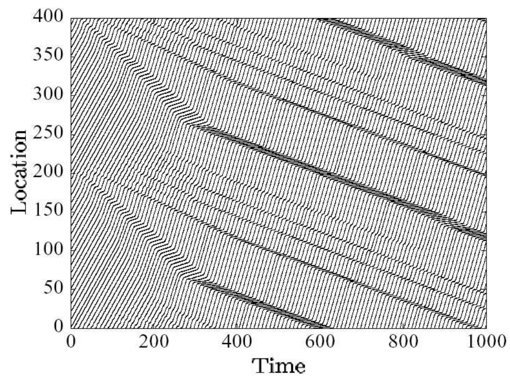

FIG. 6: Simulated trajectories calculated using the optimal velocity model [18] for every fifth vehicle. As shown, this model is able to reproduce the phenomenon of "phantom traffic jams", shown experimentally in figure 5 [9].

The optimal-velocity model tends to produce accidents when fast cars approach slow ones (among other problems). An "intelligent driver model" can be constructed which attempts to take into account how actual drivers behave. This model balances the tendency of drivers to accelerate on a free road and deccelerate when confronted with other drivers. In addition, recent modifications [19] include memory dependent effects whereby a driver's behaviour is directly influenced by the driving conditions experienced over the last few minutes. This type of model is able to reproduce virtually every important feature of the fundamental diagram (see figure 7).

Another noteworthy school of thought within the realm of microscopic models is cellular automata theory. Much more recent than other theories, traffic models based on cellular automata divide both space and time into discrete rather than continuous regions and calculate the next position in time and space based on the current configuration. These models have the advantage of being computationally very fast due to their discrete nature.

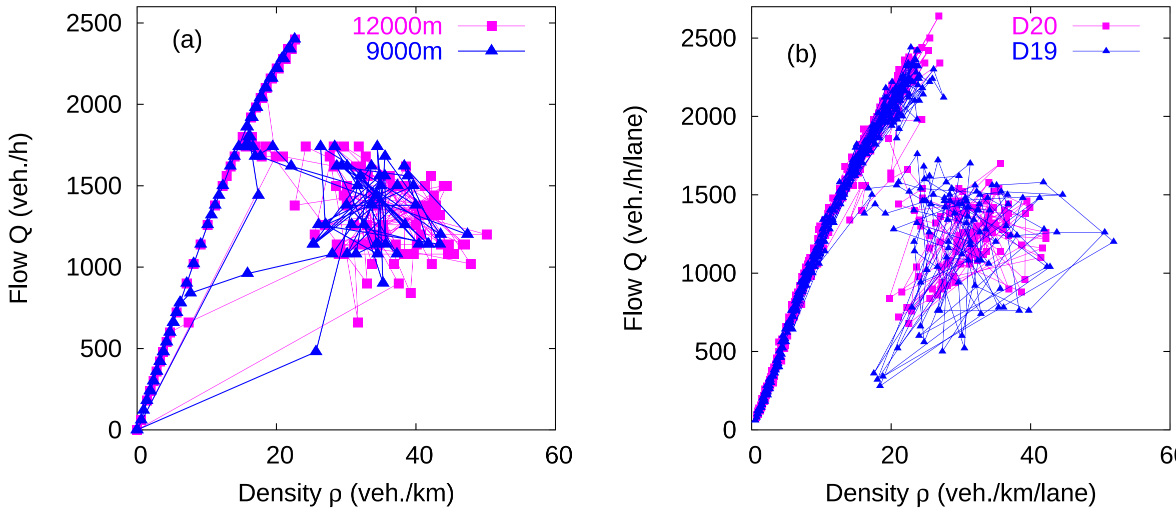

FIG. 7: Graphs of traffic flow as a function of traffic density. Here we compare simulated data for an intelligent-driver model that incorporates short-term driver memory (a) to empirical data taken on the German freeway A9-South near Frankfurt on July 31, 2001 (b). There is excellent agreement between theory and experiment, showing the characteristic free and congested regimes in addition to hysteretic effects. Figures and data from [19].

B. Macroscopic traffic models

Rather than consider the motion of each vehicle, the emergent nature of traffic allows us to consider modelling *only* the macroscopic quantities involved. A number of these theories have been suggested since the 1950's, but have since been shown [9] to follow a general pair of equations for density

$$\frac{\partial \rho}{\partial t} + V\frac{\partial \rho}{\partial x} = -\rho\frac{\partial V}{\partial x} + D(\rho)\frac{\partial^2 \rho}{\partial x^2} + \xi_1(x,t), \tag{5}$$

and velocity

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = -\frac{1}{\rho}\frac{dP}{d\rho}\frac{\partial \rho}{\partial x} + \nu\frac{\partial^2 V}{\partial x^2} + \xi_2(x,t). \tag{6}$$

Here $D(\rho)$ is the diffusion, $\xi_{1,2}(x,t)$ are fluctuations, $P(\rho)$ is the traffic pressure, $\nu(\rho)$ is a quantity similar to viscosity, and $V_e(\rho)$ is the equilibrium velocity.

In general, the individual interpretations which are special cases of these equations require fundamental vehicle conservation relations and involve macroscopic motional terms representing the convection of vehicles given their current velocity, the anticipation of drivers to conditions in front of them, and the tendency of a system to relax over some characteristic time into an equilibrium configuration.

C. Gas-kinetic models

The next step is to attempt to create a link between microscopic and macroscopic models, deriving the latter from the former. The mathematics for this process can be quite cumbersome and in general depend on which formulations one tries to use. Though some success has been attained (see figure 8), a completely satisfactory derivation has yet to be achieved. These models look structurally very similar to those of gases and fluids (e.g., one such model is based on the Navier-Stokes equations).

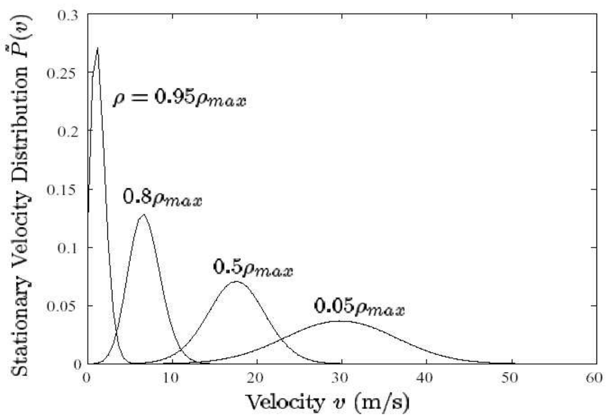

FIG. 8: Velocity distributions shown for different fractions of maximum vehicle density. These distributions were obtained through numerical solution of the gas-kinetic traffic model [20].

D. Emergent order and the Ginzburg-Landau equations

Recently, Nagatani [21] has applied statistical mechanics to the traffic problem by deriving the appropriate time-dependent Ginzburg-Landau equations from other models already discussed. He begins with a simple car-following model given by,

$$v_\alpha(t+\tau) = V\left(\Delta x_\alpha\left(t\right)\right), \tag{7}$$

where Δx_α is the headway and τ is a delay time. Here, a driver adjusts his velocity based on the observed distance to the next vehicle. Nagatani next uses an optimal-velocity given by

$$V(\Delta x_\alpha) = \frac{v_{max}}{2}\left\{\tanh(\Delta x_\alpha - h_c) + \tanh(h_c)\right\}, \tag{8}$$

where h_c is the "safety distance" at which a driver starts to deccelerate because he is too close to the car in front of him. This decellerating behaviour indicates a turning point in the above equation at h_c – a turning point which is necessary to derive the Ginzburg-Landau equations.

While the full derivation is too long to be included here, it can be found in reference [21]. Here we will list only the results. The order parameter in this model corresponds to the headway Δx_α and the temperature corresponds to $\frac{1}{\tau}$. Using the two equations above one can derive

$$\partial_{t_1} S = -[\partial_{x_1} - \frac{1}{2}\partial_{x_1}^2]\frac{\partial \Phi(S)}{\partial S}, \tag{9}$$

where $S = \Delta x_\alpha - h_c$ is the normalized headway (the order parameter) and

$$\Phi(S) \equiv \int dx \left[\frac{V'}{48}\left(\partial_{x_1} S\right)^2 + \phi(S)\right]. \tag{10}$$

Here $\phi(S)$ is the thermodynamic potential and

$$V' = \frac{dV(\Delta x_\alpha)}{d\Delta x_\alpha}\|_{\Delta x_\alpha = h_c}. \tag{11}$$

IV. CONCLUSION

Traffic, while fundamentally based on human social interactions and perceptions, can be effectively modelled using techniques originally developed for the solution of unrelated problems in physics. Like other emergent systems, the study of traffic has simultaneously progressed on both the microscopic and the macroscopic front, with the most effective research ultimately considering both regimes.

The success of traffic models, particularly their ability to accurately reproduce seemingly chaotic 'phantom' behaviour, refutes the initial assumptions that these strange behaviours are essentially random. This research can be extended to related problems such as pedestrian or avian traffic, extending this fundamentally one-dimensional problem into the second and third dimension, respectively.

---

[1] B. D. Greenshields, in *Proceedings of the Highway Research Board* (1935), vol. 14, p. 448.

[2] *Texas Transportation Institute Urban Mobility Report* (2002).

[3] D. Helbing, I. J. Farkas, and T. Vicsek, Phys. Rev. Lett. **84**, 1240 (2000).

[4] C. W. Reynolds, Comput. Graph. **21**, 25 (1987).

[5] T. Lux and M. Marchesi, Nature **397**, 498 (1999).

[6] J. Treiterer and J. I. Taylor, Highw. Res. Rec. **142**, 1 (1966).

[7] J. Treiterer and J. A. Myers, in *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, edited by D. Buckley (Reed, London, 1974), p. 13.

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

[9] D. Helbing, *Verkehrsdynamik* (Springer, Berlin, 1997).

[10] B. N. Persaud, S. Yagar, and R. Brownlee, Trans. Res. Rec. **1634**, 64 (1998).

[11] M. J. Cassidy and R. L. Bertini, Trans. Res., Part B: Methodol. **33B**, 25 (1999).

[12] D. Helbing, Phys. Rev. E **55**, 3735 (1997).

[13] B. S. Kerner, Phys. Rev. Lett. **81**, 3797 (1998).

[14] T. Musha and H. Higuchi, Jpn. J. Appl. Phys. **17**, 811 (1978).

[15] K. Nagel and M. Paczuski, Phys. Rev. E **51**, 2909 (1995).

[16] C. F. Daganzo, Tech. Rep. UCB-ITS-RR-99-5, Institute of Transportation Studies. University of California, Berkeley, CA (1999).

[17] W. Leutzbach, *Introduction to the Theory of Traffic Flow* (Springer, Berlin, 1988).

[18] M. Bando, K. Hasebe, K. Nakanishi, and A. Nakayama, J. Phys. I **5**, 1389 (1995).

[19] M. Treiber and D. Helbing, Phys. Rev. E **68**, 046119 (2003), cond-mat/0304337.

[20] C. K. J. Wagner, Phd thesis, Technische Universität (1997).

[21] T. Nagatani, Phys. Rev. E **58**, 4271 (1998).