Traffic Jam Dynamics in Traffic Flow Models

Dominic Jost, Dept. of Computer Science, ETH Zürich
Kai Nagel, Dept. of Computer Science, ETH Zürich

STRC 03 Conference paper
Session Model & Scheduling

STRC | **3rd Swiss Transport Research Conference**
Monte Verità / Ascona, March 19-21, 2003

Traffic Jam Dynamics in Traffic Flow Models

Dominic Jost
Department of Computer Science
ETH Zürich
CH-8092 Zürich, Switzerland

Phone: 01-632 31 26
Fax: 01-632 11 72
eMail: jost@inf.ethz.ch

Kai Nagel
Department of Computer Science
ETH Zürich
CH-8092 Zürich, Switzerland

Phone: 01-632 54 27
Fax: 01-632 13 74
eMail: nagel@inf.ethz.ch

Abstract

There is discussion if traffic displays multiple phases (e.g. laminar, jammed, synchronized) or not. This paper presents computational evidence that a stochastic car following model, by changing one of its parameters, can be moved from having two phases (laminar and jammed) to having only one phase. Models with two phases show three states: two being homogeneous states corresponding to each phase, and a third state which consists of a mix between the two phases (phase coexistence).

Although the gas-liquid analogy to traffic models has been widely discussed, no traffic-related model ever displayed a completely understood *stochastic* version of that transition. Having a stochastic model is important to understand the potentially probabilistic nature of the transition. Most importantly, if indeed 2-phase models describe certain aspects correctly, then this leads to predictions for breakdown probabilities. Alternatively, if 1-phase models describe these aspects better, then there is no breakdown. Interestingly, such 1-phase models can still allow for jam formation on small scales, which may give the impression of having a 2-phase dynamics.

Besides going into the details with the above arguments, we will also provide some more general overview about traffic jam dynamics and traffic jam modelling by microscopic and fluid-dynamical models.

Keywords

traffic flow theory – traffic breakdown – phase transition – 3rd Swiss Transport Research Conference – STRC 03 – Monte Verità

1. Introduction

Both from an operations and from a design perspective, the capacity of a road is an important quantity. If demand exceeds capacity, queues will form, which represent a cost to the driver and thus to the economic system. In addition, such queues may impact other parts of the system, for example by spilling back into links used by drivers who are on a path that is not overloaded.

This paper discusses freeway capacity. The question concerns the maximum flows that freeways can reach, and if the maximum flows sometimes observed (> 2500 vehicles per hour and lane) are sustainable flows or short-term fluctuations. Let us assume that there is traffic with a fairly high density ρ on a freeway, but vehicles are still able to drive at some fast velocity v. Throughput is $q = \rho\, v$. The question is what will happen if density is further increased: Can q further increase because ρ increases more than v decreases? Will q gradually decrease because ρ increases but v decreases faster? Or is there a possibility that traffic will break down, leading to stop-and-go traffic?

More technically, the question is if there is, for each density ρ, a velocity $V(\rho)$ and corresponding throughput $Q(\rho) = \rho\, V(\rho)$ at which traffic flow is smooth and homogeneous. Or is there a density range where that homogeneous traffic flow is unstable, and traffic has a tendency to reorganize into a stop-and-go pattern, with possibly lower throughput?

There is in fact a long history of publications about breakdown behavior in freeway traffic, sometimes called "reverse lambda shape of the fundamental diagram" (Koshi *et al.*, 1983; Kerner, 1999), "hysteresis" (Treiterer and Myers, 1974), "capacity drop" (Kerner and Rehborn, 1996a), "catastrophe theory" (Acha-Daza and Hall, 1993), and the like. From the modeling side, there has since long been discussions about an analogy to a gas-liquid transition (Prigogine and Herman, 1971; Reiss *et al.*, 1986), and recent work has established traffic models which display deterministic versions of a liquid-gas-like transition (Kerner and Konhäuser, 1994; Bando *et al.*, 1995).

On the other hand, measurements by Cassidy (1998) indicate that there can be stable homogeneous flow at all densities. Muñoz and Daganzo (in press) point out correctly that many of the "inverse lambda" observations could also be explained by geometrical constraints, in the following way. A bottleneck downstream of a measurement location can cause the following temporal sequence of measurements:

1. The system starts with low flow at low densities.
2. Both flow and density keep increasing, along the "free flow" branch of the fundamental diagram.
3. This flow can be larger than what can flow through the bottleneck. Then, a queue starts forming at the bottleneck, but that does not immediately influence the measurement.
4. Eventually, the queue will have spilled back to the measurement location. At that point in

time, data points will move to a much higher density, while the flow value will drop to the bottleneck capacity.

It can take up to 20 minutes for the transition zone (transition from free flow to queue) to traverse a fixed detector location, leading to fundamental diagram data points that lie between the free flow and the queue state (Muñoz and Daganzo, in press).

This mechanism generates data that looks similar to data shown in support of the breakdown hypothesis. Unfortunately, many of the published data sets do not provide enough information about the geometrical layout and the full spatio-temporal picture of the dynamics in order to resolve this question.

This question is not just academic. The correct use of technical devices such as ramp metering or adaptive speed limits (Zackor *et al.*, 1988) depends on the answer. For example, let us assume that the homogeneous solution is unstable in a certain density range, and that the alternative stop-and-go solution has a lower throughput than homogeneous traffic at the same density. In this case, the task of ramp metering might be to keep the density away from the unstable density range. If density approaches this value, on-ramp traffic should be reduced.

If, on the other hand, the homogeneous solution is stable everywhere, then ramp metering shifts capacity from the on-ramp to the through lanes, and it avoids slowdown on the freeway and its emission consequences. There would however be no net capacity effect, in the sense that –in the absence of additional obstructions– throughput downstream from the metered ramp would be the same no matter if ramp metering was switched on or not.

If, in addition, breakdown is probabilistic, that is, the homogeneous solution can survive for certain amounts of time, then the question becomes which risk of breakdown one would be willing to accept. Accepting higher flow rates in the ramp metering algorithm might increase *average* throughput, but it might also increase the probability of breakdown.

There is even discussion to include aspects of stochastic transitions into the Highway Capacity Manual (L. Elefteriadou, personal communcation). This could for example mean that, for certain flow levels, one would include a curve describing the probability that traffic flow has not broken down as a function of time. From such a curve, one could for example look up the maximum density and flow levels if one accepts a, say, 1% probability of breakdown.

Before continuing, let us make this more precise. Let us assume there is a density range where the homogeneous solution is unstable. The way this could (in principle) be tested is to have homogeneous traffic operating at a certain density, and to introduce a strong disturbance, say by stopping one car for several seconds. If the introduced disturbance heals out over time, then homogeneous traffic at that density is stable; if the disturbance grows over time, then the homogeneous solution is unstable at this density. The unstable solution needs two ingredients:

- Outflow from the jam is less than the maximally possible homogeneous flow.
- The jam, once it is there, remains compact; in other words, density inside the jam is a fixed quantity that should essentially be the same from one location to the next.

There are at least three references (Figs. 2 and 3 in Kerner and Rehborn (1996a); Fig. 3 in Kerner and Rehborn (1996b); Fig. 4 in Kerner and Rehborn (1997)) where the data points to the existence of a stable jam, embedded *both upstream and downstream* in free traffic, and where the outflow from the jam is lower than the inflow. In the 2nd and the 3rd of these references, one can in addition see that the jam is growing in width, as it should in such a situation, while remaining compact. In the 1st of these references, the data to decide this question is not sufficient.

Given this state of affairs, it makes sense to look at modeling. The task is to understand which model solutions are possible at all. This understanding will lead to the predictions of additional features that will go along with one mechanism or the other, and it might be possible to measure them, and so the issue will hopefully be eventually resolved. Until then, however, there is no agreement on the issue of breakdown in freeway traffic, and in consequence all engineering relying on one or the other assumption may not work as intended.

The starting point for our work are single-lane car following models. These models are typically either of the type $v(t+\tau_v) = f(g, \Delta v, ...)$ or of the type $a(t+\tau_a) = h(g, \Delta v, ...)$, where $v(t)$ is the velocity of a car at time t, Δv is the velocity difference to the car ahead, and a is the acceleration. g is the gap to the car ahead, where $g = \Delta x - l_c$, with Δx the front-buffer-to-front-buffer distance, and l_c is the space the car occupies in a jam. These models can for example be found in Newell (1961)

$$v(t{+}\tau_v) = V(g(t)) \ , \ \text{ with } \ V(g) = v_f - v_f \ \exp(-\lambda g/v_f) \ , \tag{1}$$

in Bando *et al.* (1994)

$$a(t) = \alpha \cdot \Big(V(g(t)) - v(t) \Big) \ , \ \text{ with } \ V(g) = v_f \cdot (\tanh(g+\ell) - \tanh(\ell)) \ , \tag{2}$$

or in Herman *et al.* (1959)

$$a(t{+}\tau_a) \propto \frac{[v(t{+}\tau_a)]^l}{[\Delta x(t)]^m} \Delta v(t) \ . \tag{3}$$

Additional parameters here are v_f (the free speed), λ, α, l, and m.

When these models are implemented on a computer, they need to be discretized in time, and one has to concern oneself with the size of the integration time step, Δt. A typical discretization is

$$a(t) = \text{given by the model} \tag{4}$$

$$v(t + \frac{\Delta t}{2}) = v(t - \frac{\Delta t}{2}) + \Delta t \, a(t) \ , \tag{5}$$

and

$$x(t + \Delta t) = x(t) + \Delta t \, v(t + \frac{\Delta t}{2}) \ . \tag{6}$$

These discretizations are meant to approach the original coupled differential equations for $\Delta t \to 0$, and there is a whole body of literature available for this (see, e.g., Press *et al.* (different years), and references therein). Once time delays (via $\tau > 0$) are introduced into such equations, numerical treatment becomes more difficult, because the dynamical history between t and $t - \tau$ needs to be memorized in increments of Δt.

In this situation, it makes sense to look for computational models which are not based on the limit $\Delta t \to 0$, but which generate useful results also for relatively large time steps of, say, one second. The model that we will use in this paper has been introduced by Krauß (1997a); it is a variant of a model used by Gipps (1981). The Krauß model has been shown to be free of collisions, i.e. that $g < 0$ never occurs (Krauß, 1997a; Nagel *et al.*, forthcoming).

In addition to being crash free at large time steps, the Krauß model is also stochastic. The important parameter for our study is a noise amplitude ϵ, which we will vary from 0.5 to 2. For $\epsilon < 0$ or $\epsilon \geq 2$ the model leaves the range of where it is plausible for traffic.

Our main results are the following:

- For medium ϵ, there are three states of traffic, which we will call laminar, coexistence, and jammed. The state depends on the density. Laminar, occurring at low density, means that nearly all vehicles have large spacing, and are driving at or near free speeds. Occurrence of some mini-jams is possible, but these mini-jams are not sustained and far apart. Jammed, occurring at high density, means that nearly all vehicles have small spacing, and are driving at low speeds or are stopped. Coexistence, occurring at intermediate density, means that the system is a mix of laminar and jammed traffic. In the coexistence state, traffic is strongly inhomogeneous.

  It is important to note that there are three states (laminar, jammed, coexistence) but only two phases (laminar, jammed). The phases refer to homogeneous sections of the system; the state refers to the system as a whole.

- For large ϵ, there is only one phase of traffic and therefore only one state. When going from low to high density, cars move closer and closer together, but traffic remains homogeneous at all times.

- At some ϵ in between, there is a transition from the 2-phase to the 1-phase regime.

- In the Krauß model, changes of ϵ also change the average acceleration. This is an unfortunate coincidence, and we believe that our general results regarding the number of phases are not related to this effect.

- Deterministic models, formulated either as car following models or as fluid-dynamical models, can display 1-phase or 2-phase behavior. They can however *not* display stochastic transitions between the phases.

The results are important for model building as well as for understanding field measurements. In a 2-phase model, theory predicts that there can be a hysteretic transition from the laminar to the coexistence state *without a change in density*. This means that, at a given density, traffic can operate in the laminar flow state for long times, until it will eventually “break down” and switch to the coexistence state. In a 1-phase model, this is impossible, and there is only one phase for any given density.

A direct consequence of this is that, if traffic follows a 1-phase model, any initial jam will "smear out" and thus eventually go away, *even with unchanged traffic conditions*. Conversely, in a 2-phase model with density in the coexistence range, jams have a typical density and a typical shape of their upstream and downstream front. These shapes are stable under disturbances, that is, the system will restore these densities and shapes after disturbances.

This paper starts with Sec. 2 which describes the general idea of a gas-liquid transition. Sec. 3 describes the general simulation setup including the car following model that is used, discusses space-time plots of the resulting dynamics, and investigates transients vs. the steady state. Sec. 4 then establishes how a coexistence state can be numerically detected for a given model. Sec. 5 reports similar results for cellular automata (CA) models. Sec. 6 discusses how these results relate to deterministic models; the paper is concluded by a discussion and a summary.

This paper is nearly identical to the TRB'03 Annual Meeting paper 03-4266.

2. Phases in Traffic

The analogy between a gas-liquid transition and the laminar-jammed transition of traffic was pointed out many times (e.g. Reiss *et al.* (1986); Bando *et al.* (1994)). The description of traffic in the well-known 2-fluid-model (Herman and Prigogine, 1979) assumes the existence of two phases; and all simulation models which use spatial queues (e.g. DYNAMIT, DYNASMART (for both see www.dynamictrafficassignment.org), Gawron (1998)) will display two phases because of the definition of the dynamics. The two phases in models with queues are however much easier to understand than the phases in more realistic models.

In a gas-liquid transition, one observes the following (see also Fig. 1(a) left):

- In the gas phase, at low densities, particles are spread out throughout the system. Distances between particles vary, but the probability of having two particles close to each other is very small.
- In the liquid phase, at high densities, particles are close to each other. There is no crystalline structure as in solids, but the density is similar and in some cases (e.g. in water) even higher in the liquid than in the gas phase. Because of the fact that the particles are so close to each other, it is difficult to compress the fluid any further.
- In between, there is the so-called coexistence state, where gas and liquid coexist. In typical experiments in gravity, the liquid will be at the bottom and the gas will be above it. Without gravity, as well as for example within clouds, droplets form within the gas and remain interdispersed. In clouds, small droplets will eventually merge together into bigger droplets (**coagulation**), which will fall out of the cloud as rain. Without gravity, the droplets will just merge but never fall out. The final state of the system is having one big droplet of liquid, surrounded by gas.

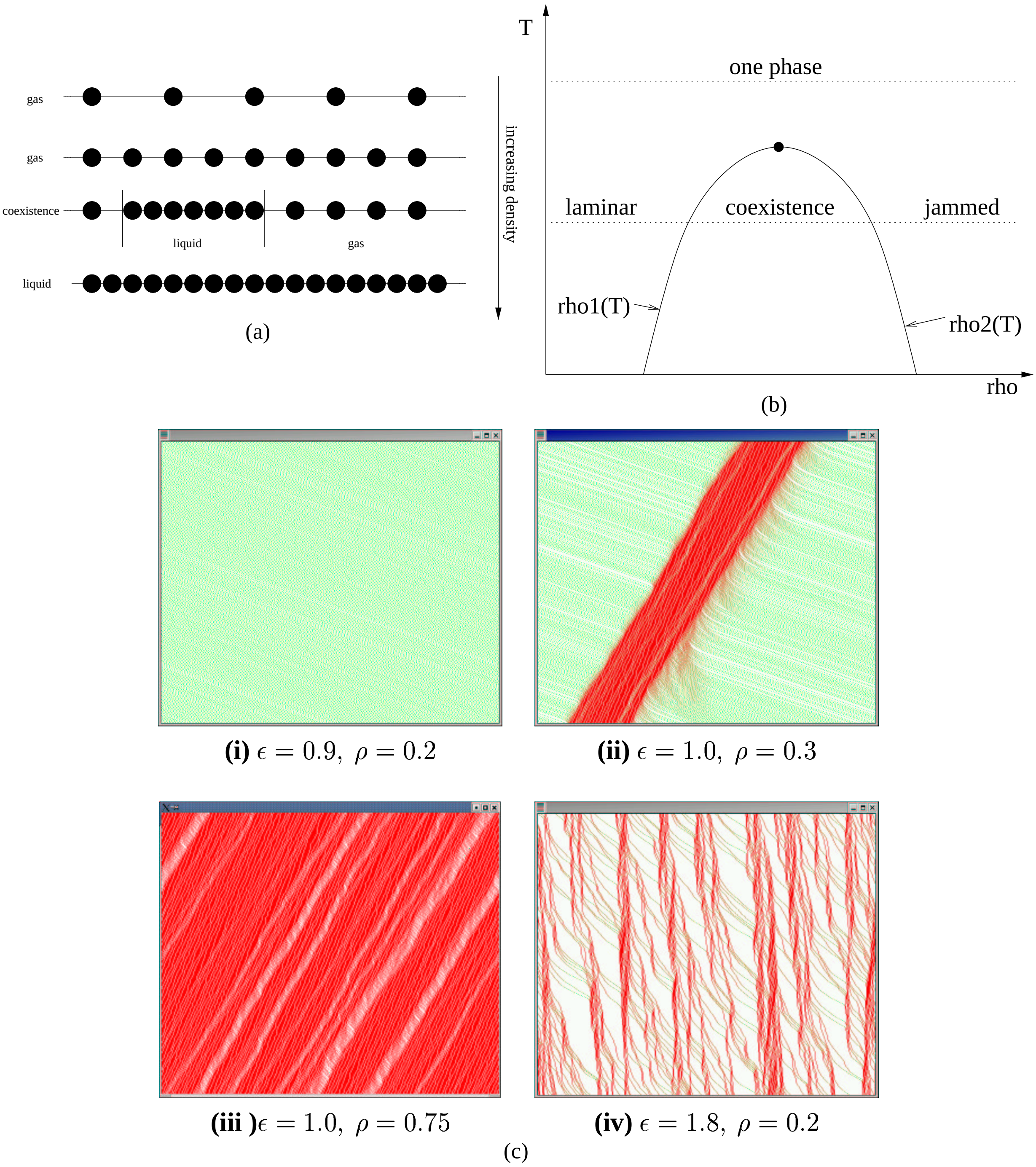


Figure 1: (a) Schematic representation of the gas-liquid transition in one dimension. (b) State of the gas-fluid model as a function of the density and the temperature T. (c) Space-time plots for different parameters. Space is horizontal; time increases downward; each line is a snapshot; vehicles move from left to right; fast cars are green, slow cars red. $L = 600$ for all plots.

> If a system in the coexistence state is compressed, more droplets form and/or existing ones grow, but the density both inside and outside the droplets remains constant. That is, the system reacts by allocating more space to the high density phase, but *not* by changing the density either of the gas or the liquid phase. Let us call those two densities ρ_1 and ρ_2. Eventually, all the space is used up by the liquid. At this point, the system will be homogeneous again and remain so if density is increased further.

The kinetics of the droplet formation (e.g. Lifschitz and Pitajewski (1987)) is ruled by a balance between surface tension and vapor pressure. Since surface tension pulls the droplet together, it increases the pressure inside the droplet. This interior pressure pushes water molecules out of the droplet. Vapor pressure outside the droplet is the balancing force – it pushes particles into the droplet.

Surface tension and thus interior pressure depend on the droplet radius – the smaller the droplet, the larger the surface tension and thus the interior pressure. The result is that, when coming from small densities, there is a regime, starting at ρ_1, where large droplets would already be stable, but small droplets are not. That is, if the system were in equilibrium, there would be a coexistence between gas and droplets. But when coming from low density, the homogeneous gaseous phase can survive for some time. This super-critical gas is thus **meta-stable**. A direct consequence of meta-stability is **hysteresis**: When coming from low densities, it is possible to go beyond ρ_1 and still remain in the gaseous phase. Only after some waiting time will, by a fluctuation, some particles get close enough to each other to start the formation of a droplet.

When increasing temperature T in a gas/liquid system, the 2-phase structure will eventually go away. This happens via ρ_1 and ρ_2 approaching each other and eventually meeting (see Fig. 1(b)). That is, depending on the temperature T, a fluid system will either display transitions from gas to coexistence and from coexistence to liquid, or there will be *no transition at all*.

We will now move on to describe the supporting evidence for our claims. As is typical in computational science, our evidence is based on computer simulations. It is however backed up by generic knowledge about phase transitions as they are well understood in physics.

3. The Simulations

3.1 Krauß Model

The velocity update of the Krauß model (Krauß *et al.*, 1996; Krauß, 1997b) reads as follows:

$$v_{\text{safe}} = \tilde{v}(t) + \frac{g(t) - \tilde{v}(t)\tau}{\overline{v}/b + \tau} \quad (7)$$

$$v_{\text{des}} = \min\{v(t) + a\,\Delta t, v_{\text{safe}}, v_{\max}\} \quad (8)$$

$$v(t+\Delta t) = \max\{0, v_{\text{des}} - \epsilon\, a\, \eta\}\ . \quad (9)$$

$\tilde{v}$ is the speed of the car in front, $\overline{v} = (v + \tilde{v})/2$ is the average velocity of the two cars involved, v_{max} is the maximum allowed velocity, a is the maximum acceleration of the vehicles, b their maximum deceleration for $\epsilon = 0$, ϵ is the noise amplitude, and η is a random number in $[0, 1]$. The meaning of the terms is as follows:

- Eq. (7): Calculation of a "safe" velocity. This is the maximum velocity that the follower can drive when she wants to be sure to avoid a crash (Krauß *et al.*, 1996; Krauß, 1997b). The main assumption is that the car ahead will never decelerate faster than b, and that the car of the follower can also decelerate with up to b.
- Eq. (8): The desired velocity is the minimum of: (a) current velocity plus acceleration, (b) safe velocity, (c) maximum velocity (e.g. speed limit).
- Eq. (9): Some randomness is added to the desired velocity.

After the velocities of all vehicles are updated, all vehicles are moved.

The Krauß model has been proven to be free of crashes for numerical time steps Δt smaller than or equal to the reaction time, τ (Krauß, 1997b; Nagel *et al.*, forthcoming). We will use $\Delta t = \tau = 1$ as has conventionally been used for the Krauß model. We further use $a = 0.2$, $b = 0.6$, $v_{max} = 3$ for all simulations.

The model is free of units; this is a property that it has inherited from the cell-based cellular automata models. A reasonable calibration is: time steps correspond to seconds and cells correspond to 7.5 meters. The reaction time then is assumed to be 1 second, and $v_{max} = 3$ corresponds to 22.5 m/s or 81 km/h. $a = 0.2$ corresponds to a maximum acceleration of 1.5 m/s per second or 5.4 km/h per second. $b = 0.6$ corresponds to a maximum deceleration of 16.2 km/h per second.

All simulations are done in a 1-lane system of length L with periodic boundary conditions (i.e. the road is bent into a ring). Let N be the number of cars on the road. The (global) density is $\rho = \frac{L}{N}$.

3.2 Pictures

Before analysing the Krauß-model numerically, it is instructive to look at the space-time plots in Fig. 1(c). Space-time plots are pictures of the time evolution of the system. In Fig. 1(c), vehicles drive to the right and time points down. Each row of pixels is a "snapshot" of the state of the road. In principle, one could reconstruct the trajectory of a particular car by connecting the corresponding pixels. In practice, at the displayed resolution this is close to impossible and one mostly observes the larger scale traffic jam structure. Traffic jams move *against* the direction of driving. The following refers to each individual case (i)–(iv) of Fig. 1(c):

(i) The laminar state: All cars drive at high speed. The available space is shared evenly among the cars. The traffic is homogeneous.

(ii) The coexistence state: The slow cars are all together in one big jam. On the rest of the road, the cars drive at high speed. In consequence, the traffic is very inhomogeneous.

(iii) The jammed state: The density is so high that no single car can drive fast. As in (i), the traffic is homogeneous.

(iv) The single phase at high ϵ: Many small jams are distributed over the whole system. There is neither a larger area of free flow, nor a major jam. The traffic is homogeneous.

Note that "homogeneous" here means "homogeneous on large scales". This means that there is a spatial measurement length ℓ above which all density measurements return the same value. If a system goes from a 2-phase to a 1-phase model, then even in the regime which technically has only one phase, structure formation on small scales is still possible. Fig. 1(c) bottom right is indeed an example for this. With larger distance from the 2-phase model, i.e. larger ϵ, the scale of these structures becomes smaller and smaller, which means that the system is homogeneous already on smaller scales. This statement can be quantified, for example via the gap distribution, i.e. the distribution of the distances between jams (Jost, 2002).

3.3 Defining a Jam

In order to make progress, one needs to define where a jam starts and where it ends. Our definition of homogeneity (see later) will not depend on this, however. A *jam* is a sequence of adjacent cars driving with speed less or equal $v_{max}/2$. The cars between two neighbouring jams are in *laminar* flow.

This definition is very simple, but will not always correspond to our natural understanding of the word jam. Thus, whether a car is jammed or not according to this definition is just a starting point and not the final answer.

3.4 Initial Condition and Relaxation

For many parameters of the Krauß model, there is a unique equilibrium state, which the system will attain after a finite time t_{relax}, no matter how it was started. Deciding when the equilibrium is reached is not trivial.

Let r_t be the state of the road at time t. To find the equilibrium value of some property, $E[f(r_{t_{relax}})]$, we use the following idea: For small t, $E[f(r_t)]$ will depend on the initial condition. With increasing time, $E[f(r_t)]$ converges towards the equilibrium value. Assume the convergence is from above. Now we need another initial condition that approaches the equilibrium value from below. Once these two sequences are close enough together, an estimate for the equilibrium value is found. Unfortunately, it cannot be guaranteed that the value thus obtained really is the equilibrium value. We use the following two initial conditions:

- Laminar start: The cars are positioned equidistant over the road with speed zero.
- Jammed start: All cars are cramped together in a big jam without any gap. Their speed is zero.

An example of this method is shown in Fig. 2(a). For $f(.)$ the number of jams was used. Since both initial conditions start with $v = 0$, the criterion of Sec. 3.3 finds one large jam at $t = 1$. After this, the following happens:

- Laminar start: Vehicles accelerate, but because of interaction, many small jams form, and the number of jams increases rapidly. These jams then slowly coagulate, which slowly reduces the number jams in the system, until the equilibrium value is reached.
- Jammed start: Vehicles accelerate out of the jam, but no or very few jams form in that outflow. Only very slowly does the number of jams increase, either via new jams in the outflow, or because of a “breaking apart” of the initial jam.

In Fig. 2(a), one sees that for both initial conditions the system eventually reaches the same number of jams. With $\epsilon = 1.0$, the system in equilibrium has, in the average, about 1.8 jams. In contrast, with $\epsilon = 1.5$, the system converges to an average of more than 20 jams.

4. Establishment of a Phase Diagram Via a Measure of Inhomogeneity

In this section, a criterion is established that distinguishes homogeneous from coexistence states. As pointed out before, coexistence states, for example at $\epsilon = 1.0$ and $\rho = 0.3$ in our model, see Figs. 1(c) (ii) and 2(a), are characterized by the coexistence of laminar and jammed traffic. Deep inside the coexistence regime, one would expect that the phases coagulate, leading to one large laminar and one large jammed section in the system. As one has seen in Fig. 2(a) for $\epsilon = 1.0$, this is essentially correct, except that small additional mini-jams always lead to the detection of a small number of additional jams. When approaching the boundaries of the coexistence regime, this characterization will become less clear-cut, and it may be possible to have more than one jam. Typically, there will be one major jam and many small ones, and for many measurement criteria this will cause enough problems to no longer be able to differentiate between the coexistence and a homogeneous state. This is particularly true for criteria that attempt a binary classification into homogeneous or not. In contrast, our criterion will show a gradual decrease in differentiating power.

The criterion is defined as follows: Partition the road into segments of length ℓ (for simplicity let ℓ divide L without remainder). For each segment the local density ρ_ℓ can be computed as the number of cars in that segment divided by ℓ. An interesting value is the variance of the local density:

$$\mathrm{Var}(\rho_\ell) = \frac{1}{L/\ell} \sum_{i=1}^{L/\ell} \Big(\rho_\ell(i) - \mathrm{E}(\rho_\ell) \Big)^2 \,, \tag{10}$$

where $E(.)$ is the expected value, which in our case is the same as the systemwide density. Note that since the density lies within $[0, 1]$, the variance cannot exceed $1/4$.

What this value picks up is how much each individual measurement segment of length ℓ deviates, in terms of its density, from the average density. Assume a system consisting of jammed and laminar traffic. If there is a jam in one segment, then the segment's density will be much higher than the average density. Conversely, if there is only laminar traffic in a segment, then the segment's density will be much lower than the average density. Var(ρ_ℓ) takes the average over the square of these deviations.

Fig. 2(b) shows this value as a function of the global density ρ and the noise parameter ϵ. Each gridpoint is the result of a computer simulation. The simulations run until the average number of jams over the last 100'000 time steps is (almost) equal for a system started with a big jam and a system started with laminar flow (see Section 3.4). Over these last times the variance of the local density is averaged.

Look at Fig. 2(b) for fixed ϵ, say $\epsilon = 1$. One sees that at densities up to $\rho \approx 0.2$, the value of Var(ρ_ℓ) is close to zero, indicating a homogeneous state, which is in this case the laminar state. Similarly, for densities higher than 0.8, Var(ρ_ℓ) is again close to zero, indicating a homogeneous state, which is in this case the jammed state. In between, for $0.2 \lesssim \rho \lesssim 0.8$, the value of Var$(\rho_\ell)$ is significantly larger than zero, indicating a coexistence state.

Now slowly increase ϵ. We see that the laminar regime ends at smaller and smaller densities, while the jammed regime starts at smaller and smaller densities (see Fig. 2(b) bottom). The latter means that for large ϵ, the jammed phase has many relatively small holes, which reduce the density, but do not break the jam. At $\epsilon \approx 1.7$, the coexistence phase completely goes away; for larger ϵ, we do not pick up any inhomogeneity at *any* density (look at the bottom plot in order to get information about behavior not visible in the 3d plot). Compare this to the theoretical expectation in Fig. 1(b), where for increasing T the two densities eventually merge and thus the different phases go away. Note that close to the transition the system still looks like it possesses different phases (see Fig. 1(c)(iv) and locate the corresponding $\epsilon = 1.8$ and $\rho = 0.2$ in Fig. 2(b)). These structures do however exist *on small scales only*. This means that for system size $L \to \infty$ and measurement interval $\ell \to \infty$ (but $\ell \ll L$), all intervals of size ℓ will eventually return the same density value. A segment length of $\ell = 62.5$, as used for Fig. 2(b), is already sufficient in order to not measure any inhomogeneity for the state in Fig. 1(c)(iv). This will not be the case for coexistence states: In coexistence state, there will always be segments with different densities, unless $\ell \approx L$. This is because droplets will coagulate so that they will eventually show up on all possible length scales ℓ.

Remember again that ϵ is a model parameter while ρ is a traffic observable. That is, once one has settled for an ϵ, the model behavior is fixed, and one has decided if one can encounter a second phase or not. *If* one can encounter a second phase, it will come into existence through changing traffic demand throughout the day – traffic can move from the laminar into the coexistence and potentially into the jammed state and back.

As a side remark, let us note that there is also another 1-phase regime for $\epsilon \to 0$. Albeit potentially interesting, this is outside the scope of this paper.

In summary, one obtains, for the traffic model, a phase diagram as in Fig. 1(b), which is the schematic phase diagram for a gas-liquid transition in fluids. Again, the important feature of this phase diagram is that there are three states for low temperatures (small T or smallϵ): gas/laminar; coexistence; liquid/jammed. For higher temperatures, the coexistence range becomes more and more narrow, while the density of the gas phase and the density of the liquid phase in the coexistence state approach each other. Eventually, these densities become equal, and the coexistence state dies out. The only important difference is that for our traffic model the phase diagram is bent to the left with increasing ϵ.

There are other criteria which can be used to understand these types of phase transitions. In particular, one can look at the gap distribution between jams, and one would expect a fractal structure at the point where the 2-phase and the 1-phase model meet, i.e. at $\rho \approx 0.2$ and $\epsilon \approx 1.7$. This is indeed the case but goes beyond the scope of this paper; see Jost (2002) for further information.

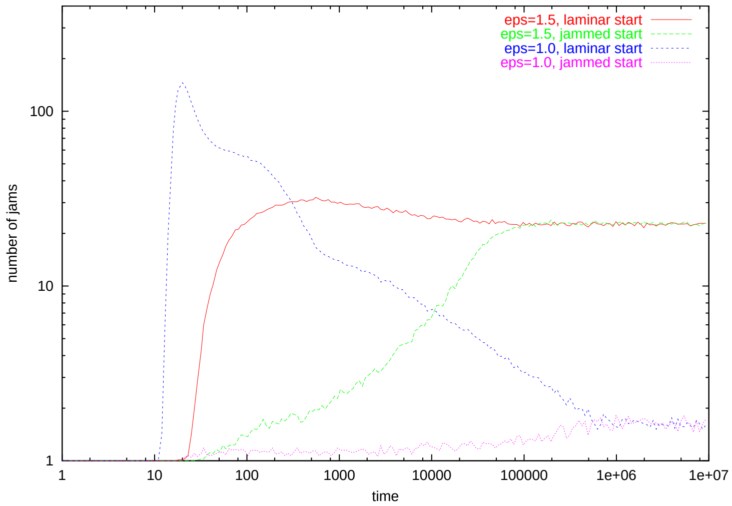
5. Cellular Automata Models

Many of the arguments regarding the nature of a stochastic and possibly critical phase transition (Nagel, 1994; Nagel and Paczuski, 1995; Sasvari and Kertesz, 1997; Roters *et al.*, 1999; Chowdhury et al, 2000) have been made using so-called cellular automata (CA) models. CA models use coarse spatial, temporal, and state space resolution. For traffic, a standard way is to segment a 1-lane road into cells of length l_c, where l_c is the length a vehicle occupies in the average in a jam, i.e. $l_c = 1/\rho_{jam} \approx 7.5$ m.

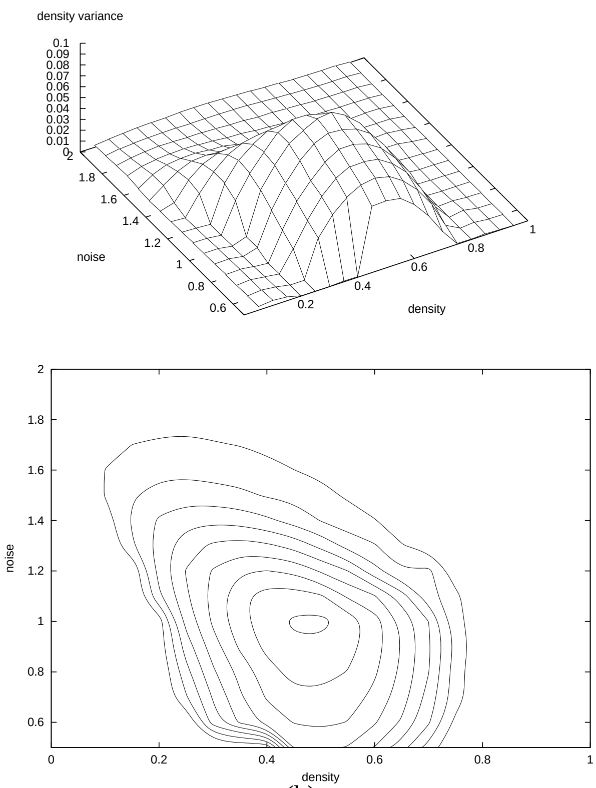
As with the Krauß model, the time step for the CA models is best selected similar to the reaction time; a time step of 1 second works well in practice. Taking the time step together with l_c, one finds that a speed of 135 km/h corresponds to five cells per time step; this is often taken as maximum velocity v_{max}.

Stochastic CA models contain a noise parameter p that introduces randomness into the driving rules: With a probability p, the deterministically calculated velocity gets reduced by one. This is the same idea as Eq. (9) in the model by Krauß. One can make p dependent on the velocity (Barlovic *et al.*, 1998); the resulting models are sometimes called models with "velocity-dependent randomization (VDR)". Often one uses just two probablilities: p_0 when the car is standing and $p_{>0}$ when the car is driving. Standard values are $p_0 = 0.5$ and $p_{>0} = 0.01$. This models that drivers, once stopped, are a bit sloppy in restarting again.

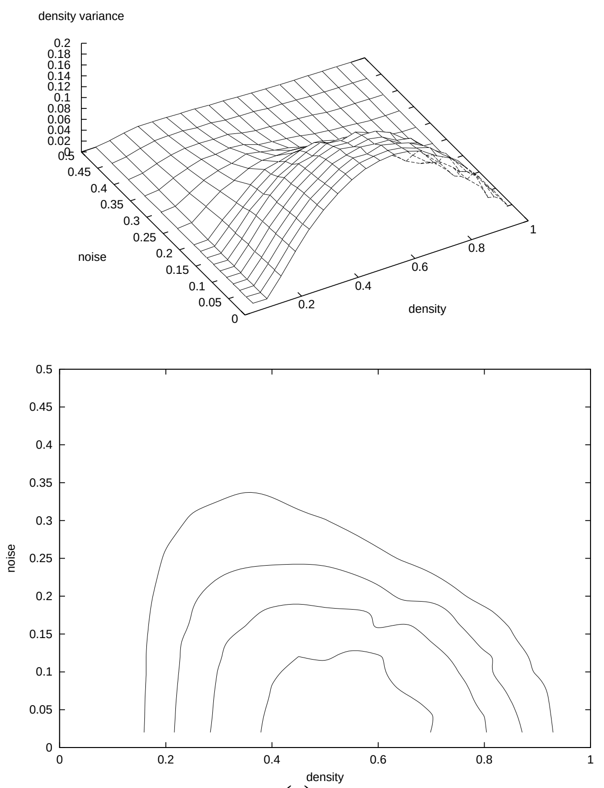
With this family of models, one can again plot the density variance (Fig. 2(c)). Instead of the



(a)



(b)



(c)

Figure 2: (a) Time evolution of the number of jams. All four curves are for 1000 cars and $\rho = 0.3$. Each curve is an average over at least 80 realizations, each with a different random seed. (b) 3d-plot and isolines of the density variance in the Krauß model. The outermost isoline is $\text{Var}(\rho_\ell) = 0.01$, the innermost $\text{Var}(\rho_\ell) = 0.09$. $L = 4000$ and $\ell = 62.5$ (c) 3d-plot and isolines of the density variance in the cellular automata model with velocity dependent randomization. The outermost isoline is $\text{Var}(\rho_\ell) = 0.04$, the innermost $\text{Var}(\rho_\ell) = 0.16$. $L = 2000$ and $\ell = 62.5$

noise amplitude ϵ, the parameter $p_{>0}$ is used. $p_{>0} = 0$ means deterministic driving except when accelerating from zero; increasing $p_{>0}$ means increasingly more randomness when moving. In this plot, one finds a behavior similar to Fig. 2(b): For small $p_{>0}$, the system displays three states (laminar, coexistence, and jammed). For $p_{>0} > 0.4$, the system is a 1-phase system. Close to $p_{>0} = 0.4$ and still at $p_{>0} = 0.5$, the system displays a lot of jam formation and structure which vanishes only when observed at very large scales (very large ℓ). In consequence, it is now clear why there was so much discussion about possible fractals for the original model (Nagel and Schreckenberg, 1992) where $p_0 = p_{>0} = 0.5$: It *is* indeed close to a critical point, and therefore fractal behavior up to a certain cut-off length scale should be expected.

6. Phase Transitions in Deterministic Models

Only stochastic models can display *spontaneous* transitions between homogeneous and coexistence states. The nature of the transition can however also become clear in deterministic models. We will discuss these similarities first for a deterministic car following model and then for deterministic fluid-dynamical models.

6.1 Car Following Models

For the model of Eq. (2), it has been shown (Bando *et al.*, 1995) that the homogeneous solution of the model is linearly unstable for densities where $V'(g) > \alpha/2$, where V' is the first derivative of the function $V(g)$, and $g = \frac{1}{\rho} - l_c$ is the gap. The instability sets in for intermediate densities; for low and high densities *all* models are stable in the homogeneous (laminar or jammed) state. For intermediate densities, one can select the curve $V(g)$ and the parameter α such that the model either has unstable ranges, or not.

If all parameters including the density are such that the homogeneous solution is not stable, then the system rearranges itself into a pattern of stop-and-go traffic, corresponding to the coexistence state. The density of the laminar and the jammed phase in the coexistence state are independent from the average system density, that is, if in that state system density goes up, it is reflected in the jammed phase using up a larger fraction of space.

The type of the instability is similar to the better-known instability of Eq. (3). However, once the instability is triggered in Eq. (3), it will just grow exponentially, and no stable 2-phase solution is found (e.g. Gerlough and Huber (1975)).

6.2 Fluid-Dynamical Models

Standard Lighthill-Whitham theory, of the type

$$\partial_t \rho + \partial_x Q(\rho) = 0 \qquad (11)$$

with a strictly convex flow-density-curve $Q(\rho)$, results in a 1-phase model, meaning that shocks smear out over time. When $Q(\rho)$ has linear sections, then in those sections shock waves are marginally stable, in the sense that disturbances to those shocks are neither amplified nor dissipated away.

Fluid-dynamical theory, of the type

$$\partial_t \rho + \partial_x(\rho\, v) = 0 \tag{12}$$

and

$$\partial_t v + v\,\partial_x v = \frac{1}{\tau}\Big(V(\rho) - v\Big) + \alpha(\rho)\,\partial_x \rho + \nu(\rho)\,\partial_x^2 v \tag{13}$$

can, depending on the choice of parameters including the $V(\rho)$-curve, either be a 1-phase/1-state or a 2-phase/3-state model (Kühne and Beckschulte, 1993). For example, the homogeneous solution of the model with $\alpha(\rho) = \frac{c_0^2}{\rho}$ and $\nu(\rho) = \nu_0$ is linearly unstable at densities where $|V'(\rho)| > \frac{c_0}{\rho}$, where $V'(\rho)$ is the first derivative of V with respect to ρ (Kühne and Beckschulte, 1993). This is similar to the instability condition in Sec. 6.1; note that $V'(\rho)$ and $V'(g)$ are, albeit related, not the same.

As pointed out before, these models are deterministic, so in no situation will these models display stochastic transitions.

7. Discussion

This paper establishes that stochastic models either possess one homogeneous phase of traffic across the whole density range (1-phase behavior), or they possess two disjoint homogeneous phases, "laminar" and "jammed", which are separated by a density regime where the two phases coexist (2-phase behavior). Speculations about this have been around for a rather long time (e.g. Prigogine and Herman (1971); Reiss *et al.* (1986); Treiterer and Myers (1974)); corresponding deterministic models have been established more recently (e.g. Kerner and Konhäuser (1994); Bando *et al.* (1995)). However, despite much discussion (e.g. Nagel (1994); Nagel and Paczuski (1995); Sasvari and Kertesz (1997); Roters *et al.* (1999); Chowdhury et al (2000)) no clear picture for stochastic models was established. Only stochastic models allow to look at meta-stable states, spontaneous transitions, and fractal-like structure, all of which are important for real world traffic. Importantly, 1-phase and 2-phase behavior can be obtained from the same model by just changing one parameter.

With respect to reality, there is no general agreement if measurements show 1-phase/1-state or 2-phase/3-state traffic. As discussed in the introduction, there is some evidence for 2-phase behavior in German data (Kerner and Rehborn, 1996a,b, 1997). Measurements in Northern America (Cassidy, 1998) point towards 1-phase behavior. In addition, many of the earlier measurements that point towards 2-phase behavior can in fact be explained by 1-phase models together with geometric constraints (Muñoz and Daganzo, in press). To make matters worse, newer publications claim the existence of three (e.g. Kerner and Rehborn (1996b)) or even more (e.g. Helbing *et al.* (1999)) phases, while other publications (e.g. Daganzo *et al.* (1999)) claim that these different phases are just queues.

Since there is discussion of entering the notion of stochastic breakdown into the Highway Capacity Manual, and since, as discussed in the introduction, the correct operation of devices, such as ramp metering and adaptive speed limit, depends on the answer of the breakdown question, it seems critical to fully understand these issues. It also seems critical to consider stochastic models, in order to not base the notion of stochastic breakdown on deterministic models. This paper's contribution is a solid step towards understanding the consequences of *stochastic* traffic breakdown, if it exists. In other words, this model will allow the development of further predictions, which are impossible to make by deterministic models, and these predictions could be tested against field data. For example, a stochastic model would predict a certain wave structure inside a queue caused by a downstream bottleneck, similar to Windover and Cassidy (2001), although a bottleneck with fixed capacity would be better suited to test the theory.

The basic theory of phase transitions, which is behind the much of this modeling work, applies in the so-called thermodynamic limit, which refers to infinitely large systems. Since traffic systems are small when compared to thermodynamic systems, the theory needs to be modified for those smaller-scale systems. Both the theory and computer modeling provide the tools for this, but great care has to be taken to find predictions which could actually be tested in the real world with finite queue lengths and finite durations. In consequence, such comparisons are highly desirable, but outside the scope of this paper.

8. Summary

This paper shows, via computational evidence, that two specific stochastic car following models can either display 1-phase/1-state or 2-phase/3-state traffic, depending on the choice of parameters. With 2-phase parameters, the two phases are: “laminar”, and “jammed”. These phases also correspond to two of the three states. Those states are homogeneous. The third state, at intermediate densities, is a coexistence state, consisting of sections with jammed and sections with laminar traffic.

The transition to a 1-phase/1-state model happens via the densities of the laminar and of the jammed phase approaching each other until they become the same. Beyond this point, there is only one homogeneous phase of traffic.

Some of these findings can be understood by looking at deterministic models for traffic, either car-following or fluid-dynamical. However, the stochastic elements of the transition cannot be explained by deterministic models. An important stochastic element is meta-stability, which means that a “super-critical” homogeneous state can survive for long times before it “breaks down” and reorganizes into stop-and-go traffic. Another important stochastic element is that structure formation and strong variability can also happen in a 1-phase model as long as the parameters are close to the 2-phase model – a deterministic model would converge to a homogeneous solution here.

It is important to understand this possibility of stochastic models to be in different regimes if one considers to enter discussions of traffic breakdown probabilities into the Highway Capacity Manual. If traffic is best described by a 1-phase model, then there is, in our view, no theoretical justification for such probabilities. If, however, traffic is best described by a 2-phase model, then the 2-phase model could give theoretical predictions for breakdown probabilities. A discussion of breakdown probabilities in 2-phase models can be found in Nagel *et al.* (In press).

Acknowledgments

We thank F. Wagner for intensive discussions. Also D. Stauffer and M. Droz have significantly helped in defining good computational criteria for the detection of multiple phases. P. Wagner, C. Daganzo, and M. Cassidy have helped indirectly (and maybe inadvertently), via the discussion of a separate but related paper. Several weeks of computing time on 192-CPU cluster Xibalba and on 44-CPU cluster Linneus was used for some of the computational results.

References

Acha-Daza, J. and F. Hall (1993) Graphical comparison of predictions for speed given by catastrophe theory and some classic models, *Transportation Research Record*, **1398** 119.

Bando, M., K. Hasebe, A. Nakayama, A. Shibata and Y. Sugiyama (1994) Structure stability of congestion in traffic dynamics, *Japan Journal of Industrial and Applied Mathematics*, **11(2)** 203–223.

Bando, M., K. Hasebe, A. Nakayama, A. Shibata and Y. Sugiyama (1995) Dynamical model of traffic congestion and numerical simulation, *Physical Review E*, **51(2)** 1035.

Barlovic, R., L. Santen, A. Schadschneider and M. Schreckenberg (1998) Metastable states in cellular automata, *European Physical Journal B*, **5** (3) 793–800, 10 1998.

Cassidy, M. (1998) Bivariate relations in nearly stationary highway traffic, *Transportation Research B*, **32B** 49–59.

Chowdhury et al, D. (2000) Comment on: “Critical behavior of a traffic flow model”, *Physical Review E*, **61** (3) 3270–3271.

Daganzo, C. F., M. J. Cassidy and R. L. Bertini (1999) Possible explanations of phase transitions in highway traffic, *Transportation Research A*, **33** 365–379.

Gawron, C. (1998) An iterative algorithm to determine the dynamic user equilibrium in a traffic simulation model, *International Journal of Modern Physics C*, **9** (3) 393–407.

Gerlough, D. L. and M. J. Huber (1975) *Traffic Flow Theory*, Special Report No. 165, Transportation Research Board, National Research Council, Washington, D.C.

Gipps, P. G. (1981) A behavioural car-following model for computer simulation, *Transportation Research B*, **15** 105–111.

Helbing, D., A. Hennecke and M. Treiber (1999) Phase diagram of traffic states in the presence of inhomogeneities, *PRL*, **82** (21) 4360–4363.

Herman, R., E. Montroll, R. Potts and R. Rothery (1959) Traffic dynamics: Analysis of stability in car following, *Operations Research*, **7** 86–106.

Herman, R. and I. Prigogine (1979) A two-fluid approach to town traffic, *Science*, **204** 148–151.

Jost, D. (2002) Breakdown and recovery in traffic flow models, Master's thesis, ETH Zurich, Switzerland. See sim.inf.ethz.ch/papers.

Kerner, B. S. (1999) Phase transitions in traffic flow, in D. Helbing, H. J. Hermann, M. Schreckenberg and D. E. Wolf (Eds.), *Traffic And Granular Flow '99*, 253–284, Springer.

Kerner, B. S. and P. Konhäuser (1994) Structure and parameters of clusters in traffic flow, *Physical Review E*, **50** (1) 54.

Kerner, B. S. and H. Rehborn (1996a) Experimental features and characteristics of traffic jams, *Physical Review E*, **53** (2) R1297–R1300.

Kerner, B. S. and H. Rehborn (1996b) Experimental properties of complexity in traffic flow, *Physical Review E*, **53** (5) R4275–R4278.

Kerner, B. S. and H. Rehborn (1997) Experimental properties of phase transitions in traffic flow, *Physical Review Letters*, **79** (20) 4030–4033.

Koshi, M., M. Iwasaki and I. Ohkura (1983) Some findings and an overview of vehicular flow characteristics, in *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*, 403–426, Elsevier, Amsterdam, Toronto.

Krauß, S. (1997a) Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics, Ph.D. thesis, University of Cologne, Germany. See www.zpr.uni-koeln.de.

Krauß, S. (1997b) Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics, Ph.D. thesis, University of Cologne, Germany. See www.zpr.uni-koeln.de.

Krauß, S., P. Wagner and C. Gawron (1996) Continuous limit of the Nagel-Schreckenberg model, *Physical Review E*, **54** (4) 3707–3712.

Kühne, R. and R. Beckschulte (1993) Non-linearity stochastics of unstable traffic flow, in C. Daganzo (Ed.), *Proc. 12th Int. Symposium on Theory of Traffic Flow and Transportation*, 367, Elsevier, Amsterdam, The Netherlands.

Lifschitz, E. and L. Pitajewski (1987) Statistische Physik, Teil 1, Lehrbuch der Theoretischen Physik, Akademie-Verlag.

Muñoz, J. C. and C. F. Daganzo (in press) Structure of the transition zone behind freeway queues, *Transportation Science*. PATH Working paper UCB-ITS-PWP-2000-24.

Nagel, K. (1994) Life-times of simulated traffic jams, *International Journal of Modern Physics C*, **5(3)** 567.

Nagel, K., C. Kayatz and P. Wagner (In press) Breakdown and recovery in traffic flow models, in Y. Sugiyama et al (Ed.), *Traffic and granular flow '01*, Springer, Heidelberg.

Nagel, K. and M. Paczuski (1995) Emergent traffic jams, *Physical Review E*, **51** 2909.

Nagel, K. and M. Schreckenberg (1992) A cellular automaton model for freeway traffic, *Journal de Physique I France*, **2** 2221.

Nagel, K., P. Wagner and R. Woesler (forthcoming) Still flowing: Approaches to traffic flow and traffic jam modeling, *Operations Research*. See sim.inf.ethz.ch/papers.

Newell, G. (1961) Nonlinear effects in the dynamics of car following, *Operations Research*, **9** (2) 209–229.

Press, W., S. Teukolsky, W. Vetterling and B. Flannery (different years) *Numerical Recipes: The art of scientific computing*, Cambridge University Press.

Prigogine, I. and R. Herman (1971) *Kinetic theory of vehicular traffic*, Elsevier, New York.

Reiss, H., A. Hammerich and E. Montroll (1986) Thermodynamic treatment of nonphysical systems: Formalism and an example (single-lane traffic), *Journal of Statistical Physics*, **42** (3/4) 647–687.

Roters, L., S. Lübeck and K. Usadel (1999) Critical behavior of a traffic flow model, *Physical Review E*, **59** 2672.

Sasvari, M. and J. Kertesz (1997) Cellular automata models of single lane traffic, *Physical Review E*, **56** (4) 4104–4110.

Treiterer, J. and J. Myers (1974) The hysteresis phenomenon in traffic flow, in D. Buckley (Ed.), *Proc. 6th ISTT*, 13, A.H. & A.W. Reed Pty Ltd., Artarmon, New South Wales.

Windover, J. and M. Cassidy (2001) Some observed details for freeway traffic evolution, *Transportation Research A*, **A35** 881–894.

Zackor, H., R. Kühne and W. Balz (1988) Untersuchungen des Verkehrsablaufs im Bereich der Leistungsfähigkeit und bei instabilem Fluß, *Forschung Straßenbau und Straßenverkehrstechnik*, **524**, Bundesminister für Verkehr, Bonn–Bad Godesberg.