Cooper@tive Tr@nsport@tion Dyn@mics **1**, 2.1–2.24 (2002)

Critical Discussion of "Synchronized Flow"

Dirk Helbing[a,b,1] **and Martin Treiber**[a,2]

[a] Institute for Economics and Traffic, Dresden University of Technology, Andreas-Schubert-Str. 23, D-01062 Dresden, Germany

[b] Collegium Budapest – Institute for Advanced Study, Szentháromság u. 2, H-1014 Budapest, Hungary

(Submitted: March 13, 2002; Revised: May 3, 2002; Published: May 7, 2002)

Abstract: We critically discuss the concept of "synchronized flow" from a historical, empirical, and theoretical perspective. Problems related to the measurement of vehicle data are highlighted, and questionable interpretations are identified. Moreover, we propose a quantitative and consistent theory of the empirical findings based on a phase diagram of congested traffic states, which is universal for all conventional traffic models having the same instability diagram and a fundamental diagram, in particular deterministic ones. New empirical and simulation data supporting this approach are presented as well.

Key Words: Synchronized flow; Measurement problems; Wide scattering and interpretation of aggregate traffic data; Correlations of macroscopic variables; Phase diagram of congested traffic states.

[1]E-mail: `helbing@trafficforum.org`, URL: `www.helbing.org`

[2]E-mail: `martin@mtreiber.de`, URL: `www.mtreiber.de`

1 Freeway Traffic: "Synchronized flow", "Pinch Effect", and Measurement Problems

1.1 What is new about "synchronized flow"?

Congested traffic shows subtle effects and has, therefore, been investigated for many decades. To identify the reasons for its complex phenomenology, Kerner and Rehborn have removed the data belonging to wide moving jams (see MLC and TSG in Figs. 9 and 10) and found that the remaining data of congested traffic still displayed a wide and two-dimensional scattering [42], see Fig. 4(c). By mistake (see Figs. 1, 4(c), and Sec. 1.2), they concluded that all models assuming a fundamental diagram were wrong and defined a new state called *"synchronized flow"* ("synchronized" because of the typical synchronization among lanes in congested traffic, see Fig. 2(a), and "flow" because of flowing in contrast to standing traffic in fully developed jams). Since then, Kerner suggests to classify *three phases*:

1. *free flow,*
2. *"synchronized flow"*, and
3. *wide moving jams* (i.e. *moving localized clusters* whose width in longitudinal direction is considerably higher than the width of the jam fronts).

In some applied empirical studies, however, Kerner *et al.* additionally distinguished a fourth state of stop-and-go traffic [45, 46].

Free flow is characterized by the average desired velocity V_0 and, therefore, by a strong correlation and quasi-linear relation $Q \approx \rho V_0$ between the local flow Q and the *local* density ρ [64]. It is also well-known that *wide moving jams* propagate with constant form and (phase) velocity $C \approx -15$km/h [7, 43, 61]. Kerner found that this propagation is not affected by bottlenecks or the presence of "synchronized flow". Moreover, he showed that the outflow Q_{out} from wide jams is a self-organized traffic constant as well [39, 43]. In contrast to wide moving jams, the flow inside of *"synchronized flow"* remains finite, and its downstream front is normally fixed at the location of some bottleneck, e.g. an on-ramp. Therefore, "synchronized flow" basically agrees with previous observations of queued or congested traffic (see, e.g., Refs. [2, 3, 66] and the references therein). In his patent [37], Kerner applies the queuing theory himself, which goes back to the fluid-dynamic traffic model by Lighthill and Whitham [56].

The synchronization of the average velocities among neighboring lanes has been already described by Koshi *et al.* [52] (but see also Refs. [15, 18, 61]). It is true on a *macroscopic*

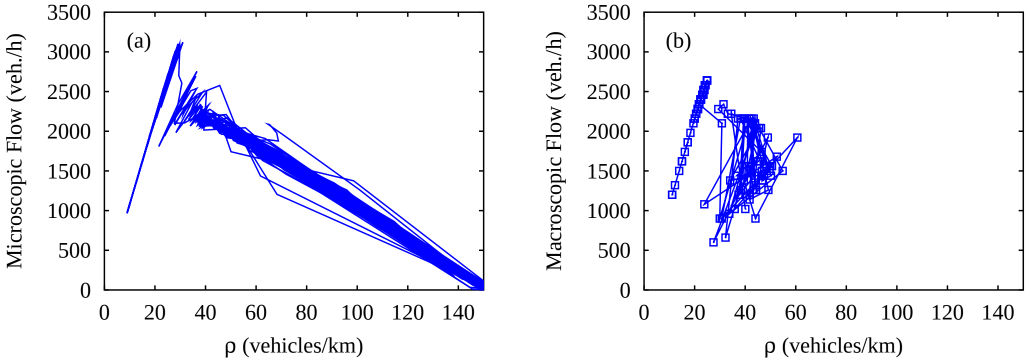

Figure 1: (a) Flow-density relation for narrow moving jams simulated with a microscopic traffic model. (b) The aggregated (1-minute) data corresponding to the narrow moving jams displayed in (a) show a wide scattering and erratic movement in the flow-density plane. By mistake, this is used to characterize "synchronized flow".

level for *all* forms of congested traffic including wide moving jams. Simulations have shown that this is a result of lane changes [53], while the assumed reduction in the lane changing rate [44] occurs only after the speeds in neighboring lanes have been successfully balanced [69]. On a *microscopic* scale, over-taking maneuvers continue to exist almost at all densities [27], see Figs. 2(b), (c). Nevertheless, the probability of lane changes drops considerably with increasing density, when most of the road is used up by the vehicles' safety headways [27], but not in the postulated Z-shaped way [35]. Due to the reduced opportunities for overtaking and the related bunching of vehicles, the velocity variance goes down with increasing density as well (also because of the reduced velocities) [22, 23, 44].

The transition between free and congested traffic is of hysteretic nature, i.e. the density immediately before the transition is higher and the average velocity lower than immediately after the inverse transition. This has been known for a long time [19, 65, 76]. Kerner specifies that the transition is usually caused by a localized and short perturbation of traffic flow that starts downstream of the on-ramp and propagates upstream with a velocity of about $-15\,\mathrm{km/h}$. As soon as the perturbation passes the on-ramp, it triggers the breakdown which spreads upstream in the course of time. The congested state can then persist for several hours [44].

Moreover, Kerner and Rehborn distinguish three types of "synchronized flow" [42], which may be short-lived:

(i) Stationary and homogeneous states where both the average speed and the flow rate

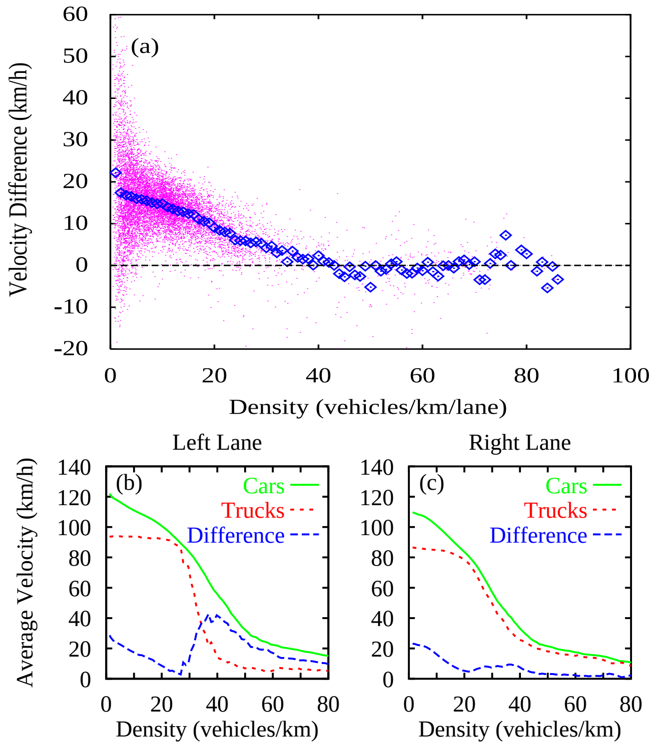

Figure 2: (a) The difference of the average velocity in the left and the right lane vanishes at densities above 30 vehicles per kilometer, corresponding to a macroscopic synchronization of the speeds [22]. (b), (c) The difference in the empirically determined velocities of cars and trucks, however, show that overtaking maneuvers continue to exist even at higher densities [27].

are stationary (see, e.g., also Refs. [20, 67, 77]. Later on, we will call these *"homogeneous congested traffic"* (HCT) [26].

(ii) States where only the average vehicle speed is stationary, named *"homogeneous-in-speed states"* (see also Refs. [32, 55]). We interpret this state as *"recovering traffic"* [25], as it bears several signatures of free traffic and mostly appears downstream of bottlenecks with congested traffic.

(iii) Non-stationary and non-homogeneous states (see also Refs. [6, 32, 75]). For these, we will use the term *"oscillating congested traffic"* (OCT) [26].

At least types (i) and (iii) are characterized by a considerably scattering and erratic change of the flow-density data, the various sources of which will be addressed in the following subsection. Continuous (spatial) transitions between these types are probably the reason for the so-called *"pinch effect"* [33], see Fig. 3(a):

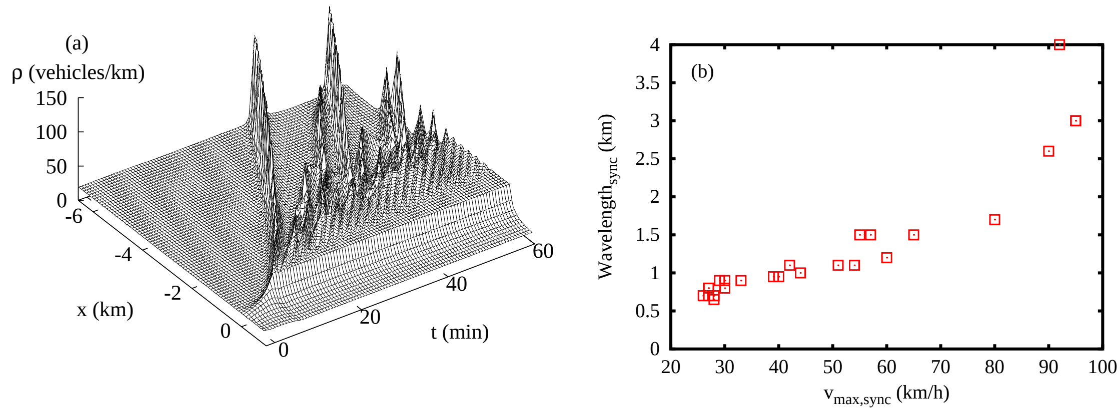

Figure 3: (a) Simulation of the pinch effect with a deterministic microscopic model showing stable traffic at low and high densities, linearly unstable traffic at medium densities, and metastable traffic in between. The spatio-temporal density plot illustrates the breakdown to homogeneous congested traffic (HCT) upstream of a bottleneck, emerging oscillating congested traffic (OCT) further upstream, and a few stop-and-go waves (TSG) emanating from this region. The conditions for this spatial coexistence of congested traffic states are as follows [25, 72]: The density in the congested region immediately upstream of the bottleneck should be in the linearly unstable, but convectively stable range, where perturbations are convected away in upstream direction [10, 59]. In this case, traffic flow will appear stationary and homogeneous close to the bottleneck, but small perturbations will grow as they propagate upstream in the congested region starting at the bottleneck. If the perturbations propagate faster than the congested region expands, they will reach the area of free traffic upstream of the bottleneck. During rush hours, it is quite likely that this free flow is in the metastable range between the free and linearly unstable density region. Consequently, sufficiently large perturbations can trigger the formation of jams, which continue travelling upstream, while small perturbations are absorbed. (b) The wavelength (average distance between density maxima) determined from (a) qualitatively displays the empirical increase with the vehicle velocity observed by Kerner [33].

Upstream of a section with homogeneous congested traffic close to a bottleneck, there is a so-called *"pinch region"* characterized by the spontaneous birth of small and narrow density clusters, which are growing while they travel further upstream. Wide moving jams are eventually formed by the merging or disappearance of narrow jams, in which the cars move slower than in wide jams [33]. Once formed, the wide jams seem to suppress the occurence of new narrow jams in between. Similar findings were already reported by Koshi *et al.* [52], who observed that "ripples of speed grow larger in terms of both height and length of the waves as they propagate upstream".

1.2 Wide scattering of congested flow-density data

The collection and evaluation of empirical freeway data is a subject with often underestimated problems. To make reliable conclusions, in original investigations one should specify

1. the measurement site and conditions (including applied control measures),
2. the sampling interval,
3. the aggregation method,
4. the statistical properties (variances, frequency distributions, correlations, survival times of traffic states, etc.),
5. data transformations,
6. smoothing procedures,

and the respective dependencies on them.

The measurement conditions include ramps and road sections with their respective in- and outflows, speed limits, gradients, and curves with the respectively related capacities, furthermore weather conditions (like rain, ice, blinding sun, etc.), presence of incidents (including the opposite driving direction), and other irregularities such as road works, which may trigger a breakdown of traffic flow. Moreover, one should evaluate the long vehicles ("trucks") separately, as their proportion varies significantly, see Fig. 4(a). This can explain a considerable part of the wide scattering of congested traffic [73], see Figs. 4(c), (d). Presently, this explanation is still the only one for this observation that has been quantitatively checked with empirical data. Note that a considerable variation of the time headways is also observed among cars, see Fig. 4(b). This is partly due to different driver preferences and partly due to the instability of traffic flow, see Fig. 5(a). While vehicle platoons with reduced time headways imply an increase of the flow with growing density, a reduction in vehicle speeds is usually related with a decrease. According to Banks [5],

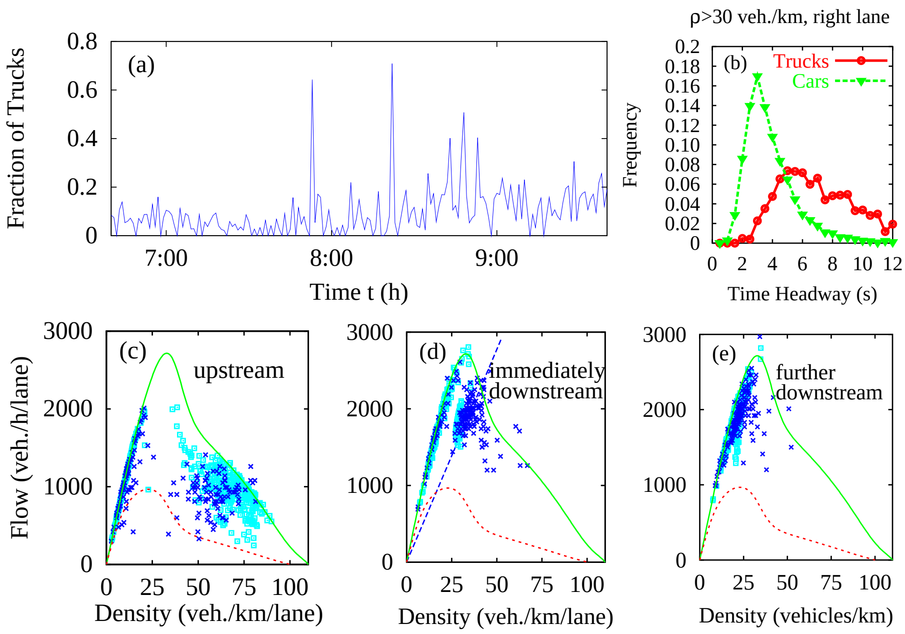

Figure 4: (a) The empirical truck fraction varies considerably in the course of time. (b) The time headways of long vehicles ("trucks") are on average much higher than those of short vehicles ("cars"). (c)-(e) Assuming a fundamental diagram for cars (solid line), a separate one for trucks (dashed line), weighting them according to the measured truck fraction, and using empirical boundary conditions allows to reproduce the observations in a (semi-)quantitatively way [73]: Free traffic (at low densities) is characterized by a (quasi-)one-dimensional curve. (c) Data of congested traffic *upstream* of a bottleneck are widely scattered in a two-dimensional area. (d) *Immediately downstream* of the bottleneck, one observes homogeneous-in-speed states reflecting recovering traffic. (e) Further downstream the data points approach the curve describing free traffic. Dark symbols correspond to empirical one-minute data, light ones to corresponding simulation results.

this can account for the observed erratic changes of the flow-density data. We support this interpretation.

The strong variations of traffic flows imply that all measurements of macroscopic quantities should be complemented by error bars (see, e.g., Ref. [21]). Due to the relatively small "particle" numbers behind the determination of macroscopic quantities, the error bars are actually quite large. Hence, many temporal variations are within one error bar, when traffic flow is unstable, see Fig. 5(b). It is, therefore, very questionable whether it is possible to empirically prove the existence of small perturbations triggering a breakdown

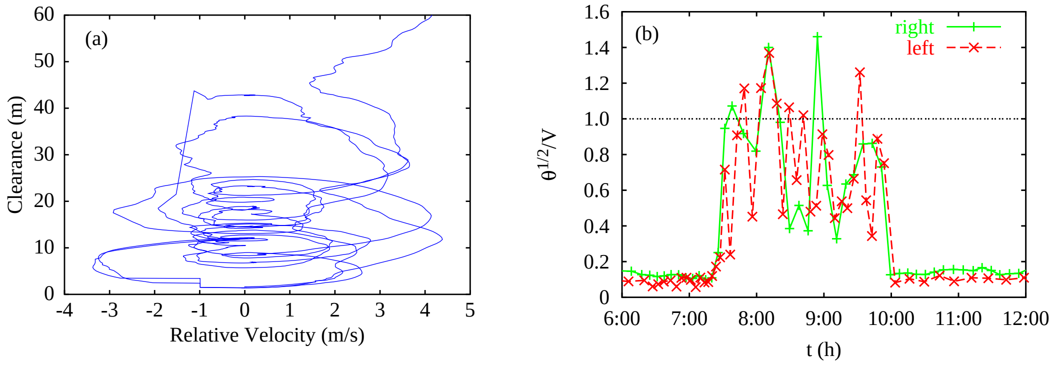

Figure 5: (a) The measured oscillations of the clearance and the relative velocity [31] indicate an instability in the car-following behavior [28]. (b) The empirical standard deviation $\sqrt{\theta(t)}$ of vehicle velocities divided by the average velocity $V(t)$ is particularly large during the rush hour, where traffic flow is congested and unstable [25, 70].

of traffic flow [44] or of the “birth” and merging of narrow density clusters in the “pinch region” [33]. At least, this would require a thorough statistical support. Consequently, we deny that such kind of data are presently suited as starting point for the development of new models [38] or traffic theories [35]. There is a considerable risk of overinterpreting particular (possibly statistical) features of the data recorded at special freeway sections and to construct new models that merely simulate what has been incorporated by means of the model assumptions. In fact, the only reason why we personally believe in the correctness of these interpretations is the existence of plausible deterministic traffic models reproducing these hard-to-see effects without any special assumptions or extensions (see Fig. 3 and Refs. [29, 53, 72]). That is, models that have been developed to describe other effects, have been found to produce also phenomena corresponding to the claimed ones, but this is no proof.

Because of the above mentioned problems, we call for more refined measurement and evaluation techniques, which are required for more reliable data analyses. These must take into account *correlations* between different quantities, as is pointed out by Banks [4].

For example, approximating the vehicle headways by $d_\alpha = v_\alpha \Delta t_\alpha$ (where v_α is the velocity and Δt_α the time headway of vehicle α) and determining arithmetic multi-vehicle averages $\langle \dots \rangle$ at a fixed location, one obtains for the inverse vehicle flow

$$\frac{1}{Q} = \langle \Delta t_\alpha \rangle = \left\langle \frac{d_\alpha}{v_\alpha} \right\rangle = \langle d_\alpha \rangle \left\langle \frac{1}{v_\alpha} \right\rangle + \mathrm{cov}\left(d_\alpha, \frac{1}{v_\alpha} \right) . \quad (1)$$

Herein, $\mathrm{cov}(d_\alpha, 1/v_\alpha)$ denotes the covariance between the headways d_α and the inverse velocities $1/v_\alpha$, which is negative and particularly relevant at large vehicle densities, as

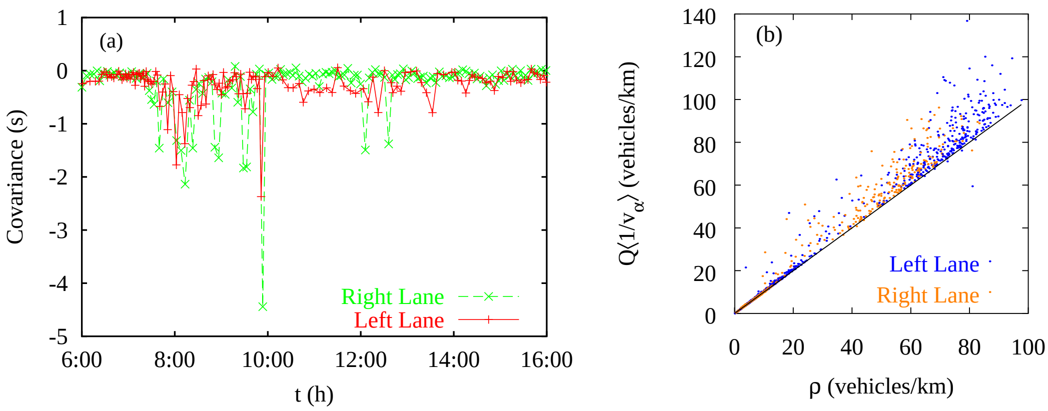

Figure 6: (a) The covariance between headways d_α and inverse velocities $1/v_\alpha$ shows large variations and significant deviations from zero in congested traffic, while it approximately vanishes in free flow, compare Fig. 5(b). Even after traffic has recovered, there seem to remain weak correlations between headways and vehicle speeds for a considerable time. These are probably a reminiscence of congestion due to platoons which have not fully dissolved [25, 70]. (b) 50-vehicle averages of the density $Q\langle 1/v_\alpha \rangle$ as a function of the actual density $\rho = 1/\langle d_\alpha \rangle$. If both measurement techniques were equivalent, as is usually assumed, all data points should lie on the solid line. Instead, the neglect of the covariance leads to an overestimation of the actual density by the formula $Q\langle 1/v_\alpha \rangle$. In contrast, the formula $Q/\langle v_\alpha \rangle$ underestimates the actual density [25, 71].

expected (see Fig. 6). Defining the local density ρ by

$$\rho = 1/\left\langle d_\alpha \right\rangle \tag{2}$$

and the average velocity V via

$$\frac{1}{V} = \left\langle \frac{1}{v_\alpha} \right\rangle , \tag{3}$$

we obtain the fluid-dynamic flow relation

$$Q = \rho V \tag{4}$$

by the conventional assumption $\text{cov}(d_\alpha, 1/v_\alpha) = 0$. According to Fig. 6, however, this overestimates the density systematically, since the covariance tends to be negative due to the speed-dependent safety distance of vehicles. In contrast, the common method of determining the density via $Q/\langle v_\alpha \rangle$ systematically underestimates the density [25, 71].

Moreover, because of the large variation of the covariance in time, errors in the measurement of the density due to a neglect of correlations partly account for the observed scattering of flow-density data in the congested regime.

A similar problem occurs when the density is determined via the time occupancy of a certain cross section of the road. Considering that $\Delta t_\alpha = T_\alpha + l_\alpha/v_\alpha$, where T_α is the (netto) time clearance and l_α the length of vehicle α, we have the relation

$$\rho = \rho_{\max}\frac{\langle l_\alpha/v_\alpha\rangle}{\langle \Delta t_\alpha\rangle} = \rho_{\max} Q\langle l_\alpha/v_\alpha\rangle = \frac{Q}{\langle l_\alpha\rangle}\langle l_\alpha/v_\alpha\rangle\,, \tag{5}$$

where $\rho_{\max} = 1/\langle l_\alpha\rangle$ is the maximum density and $\langle l_\alpha\rangle$ the average vehicle length. For a finite detector length L_D, we have to replace l_α by $l_\alpha + L_D$ [25, 60]. The formula $1/V = \langle l_\alpha/v_\alpha\rangle/\langle l_\alpha\rangle$ for the average velocity results in the correct expression $1/V = \langle 1/v_\alpha\rangle$ only, if the individual vehicle lengths and velocities are not correlated, which is usually not the case.

2 A Quantitative Theory of Congested Traffic States

When Kerner started to question all traffic models with a fundamental diagram, physicists were used to simulate traffic in closed systems with periodic boundary conditions. With the Kerner-Konhäuser model, it was possible to produce free traffic, emergent stop-and-go waves, and triggered wide moving jams [40, 50, 51]. However, attempts to simulate "synchronized flow" failed even when small ramp flows were added to the system. They resulted in what we call a *pinned localized cluster* (PLC) located at the on-ramp [41] (see Figs. 9 and 10). Because of the sensitivity of the model and problems with the treatment of open systems, it was not possible to simulate open systems with considerable ramp flows. Other independent studies for periodic systems with localized bottlenecks produced either homogeneous vehicle queues (HCT) or *oscillating congested traffic* (OCT) [8, 9, 11, 16, 30, 48, 53, 58, 62, 63, 68], but at that time nobody could make sense of these apparently inconsistent findings. This situation changed, when Helbing *et al.* derived a phase diagram of congested traffic states. They managed to simulate a macroscopic traffic model with open boundary conditions even in extreme situations and investigated a freeway stretch with a single ramp [26]. Instead of the densities, they identified the main flow on the freeway and the on-ramp flow as the suitable control parameters for on open system and varied them systematically. In this way, they found that a perturbation could trigger different kinds of congested traffic states. Moreover, the boundaries separating different states could be related to the *instability diagram* for homogeneous freeways and other characteristic quantities [26]. For this reason, they concluded that the phase diagram should be qualitatively the same, i.e. *universal*, for all microscopic, mesoscopic, or macroscopic traffic models having the same instability diagram. This has been supported in the meantime [26, 54, 75]. Apart from this, the phase diagram of traffic models with different instability diagrams can be directly derived [25]. Generalizations to other kinds of bottlenecks (e.g. gradients) have been developed as well [75].

In the following, we will sketch the basic ideas behind the phase diagram of congested traffic states (for a more detailed discussion see Ref. [25]). Let us assume our traffic model has a fundamental diagram

$$Q_{\rm f}(\varrho) = \varrho V_{\rm f}(\varrho) \tag{6}$$

describing the relation between the vehicle flow Q, the *spatially averaged* density ϱ, and the average velocity V in homogeneous and stationary traffic. (The flow-density relation of emergent stop-and-go waves is characterized by a linear relation, i.e. it looks differently [40].) Moreover, let us assume the model has ranges of stable traffic flow at small and high densities, a range of linearly unstable traffic flow at medium densities, and ranges of metastable traffic flow in between. This kind of instability diagram is, for example, found for the macroscopic model used by Kühne, Kerner and Konhäuser, or Lee *et al.* [40, 50, 53], for the microscopic optimal velocity model [1], for the non-local gas-kinetic-based traffic model [74], or the microscopic intelligent driver model [75] (among which the first two models are rather sensitive to parameter variations, but the latter two are quite robust).

In contrast to circular freeways, emergent "phantom traffic jams" are not typical for open homogeneous freeway stretches, as it is normally impossible to reach the linearly unstable density regime by feeding in vehicles at the upstream boundary. This is in agreement with empirical observations [34]. Most cases of traffic congestion on an n-lane freeway are observed upstream of on-ramps or other bottlenecks. They can be triggered by perturbations significantly below the theoretical capacity, as soon as the sum of the upstream freeway flow $Q_{\rm up}$ and the on-ramp flow

$$\Delta Q = Q_{\rm rmp}/n \tag{7}$$

per lane exceeds the outflow $Q_{\rm out}$ from congested traffic: If a disturbance leads to temporary congestion, the drivers must accelerate again and suffer some time delay, which reduces the capacity to $Q_{\rm out}$. Therefore, the following vehicles will queue up, and the temporary perturbation grows to form a persistent kind of congestion. The initial perturbation can even be a temporary *reduction* of the traffic flow and/or vehicle density, which can be caused by temporal variations of the traffic volume or even by vehicles leaving the freeway at some off-ramp [14, 24], see Fig. 7.

If the total traffic volume

$$Q_{\rm tot} = (Q_{\rm up} + \Delta Q) \tag{8}$$

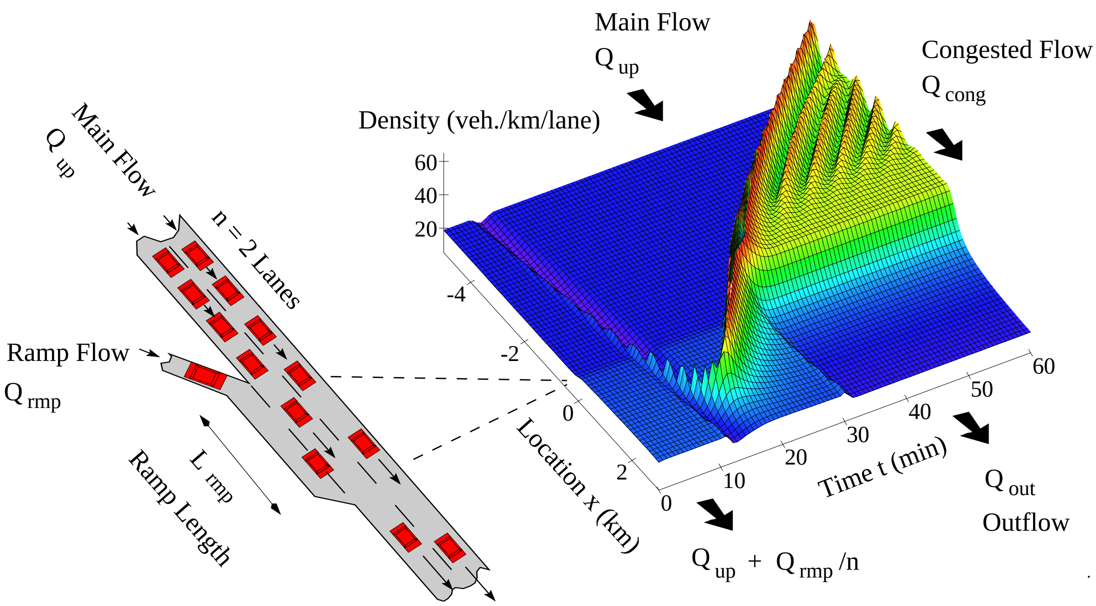

Figure 7: Negative perturbation triggering oscillating congested traffic. When the traffic density has sufficiently increased to reach the metastable regime, the "negative" perturbation will be amplified if it only exceeds the critical amplitude. While it is small, it will move downstream with the vehicles, so one could hope it would pass the bottleneck and leave the system. However, when the density wave grows larger, it will reduce its speed and even change its propagation direction. Once it is fully developed, it moves upstream with constant velocity, since vehicles leave the jam at the downstream front, while new ones join it at the upstream front. Hence, the perturbation returns to the bottleneck (see Figs. 8 and 10 for empirical signs of this *"boomerang effect"*), and it triggers a breakdown of traffic, when it passes the bottleneck in upstream direction, as it thereby reduces the effective capacity to Q_{out} [24, 25, 29].

is greater than the dynamic capacity Q_{out}, we will automatically end up with a growing vehicle queue upstream of the on-ramp. The traffic flow Q_{cong} resulting in the congested area normally gives, together with the inflow ΔQ, the outflow Q_{out}, i.e.

$$Q_{\text{cong}} = (Q_{\text{out}} - \Delta Q) \tag{9}$$

(but see the footnote on p. 1111 of Ref. [25] for exceptions). One can distinguish the following cases [25, 26, 75] (see Fig. 9): If the density ϱ_{cong} associated with the flow

$$Q_{\text{cong}} = Q_{\text{f}}(\varrho_{\text{cong}}) \tag{10}$$

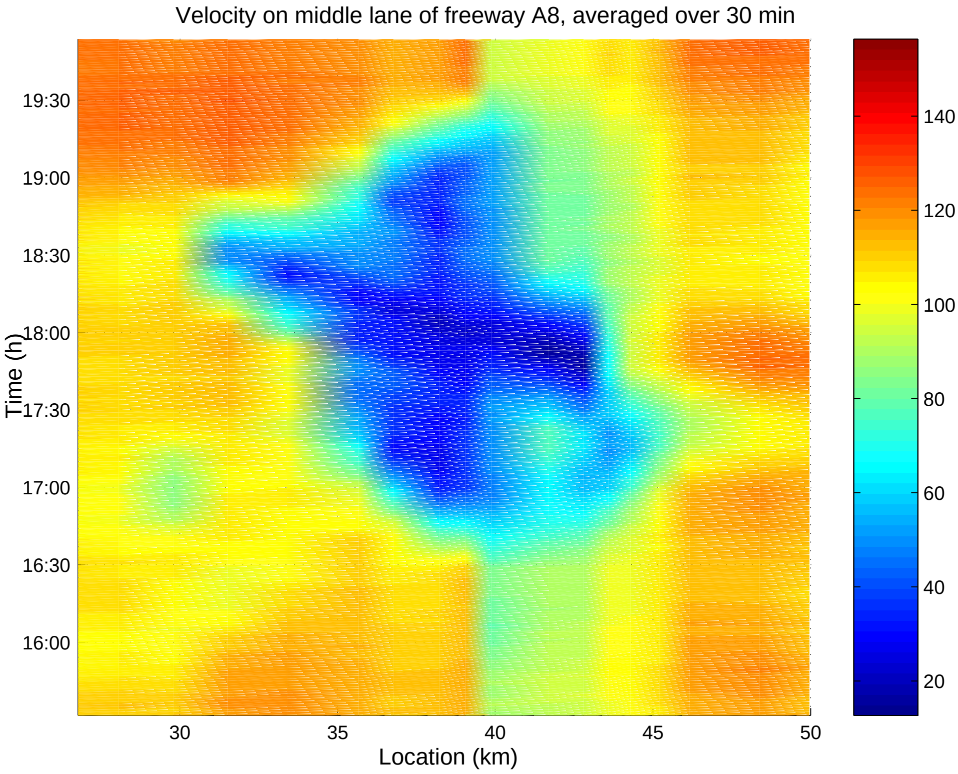

Figure 8: The wide moving jam upstream (left) of kilometer 43 starts with a "boomerang effect" and travels through the "synchronized" congested traffic flow left of kilometer 41 (dark area). A careful data analysis shows that the boomerang-like perturbation caused an accident on the right lane which was responsible for the bottleneck at kilometer 43 [75]. In this contour plot, the displayed data are moving time averages over intervals of 30 minutes, so that short events of low speed are illustrated with a color representing higher speeds. (Reproduction with kind permission of Rudolf Sollacher, Siemens AG, Munich.) Note that several more representative examples of boomerang effects have been found in a current analysis of new data. These will be presented and discussed in a forthcoming publication.

is stable, we find *homogeneous congested traffic* (HCT) such as typical traffic jams during holiday seasons. For a smaller on-ramp flow ΔQ, the congested flow Q_{cong} is linearly unstable, and we either find *oscillating congested traffic* (OCT) or *triggered stop-and-go traffic* (TSG).

In contrast to OCT, stop-and-go traffic is characterized by a sequence of moving jams,

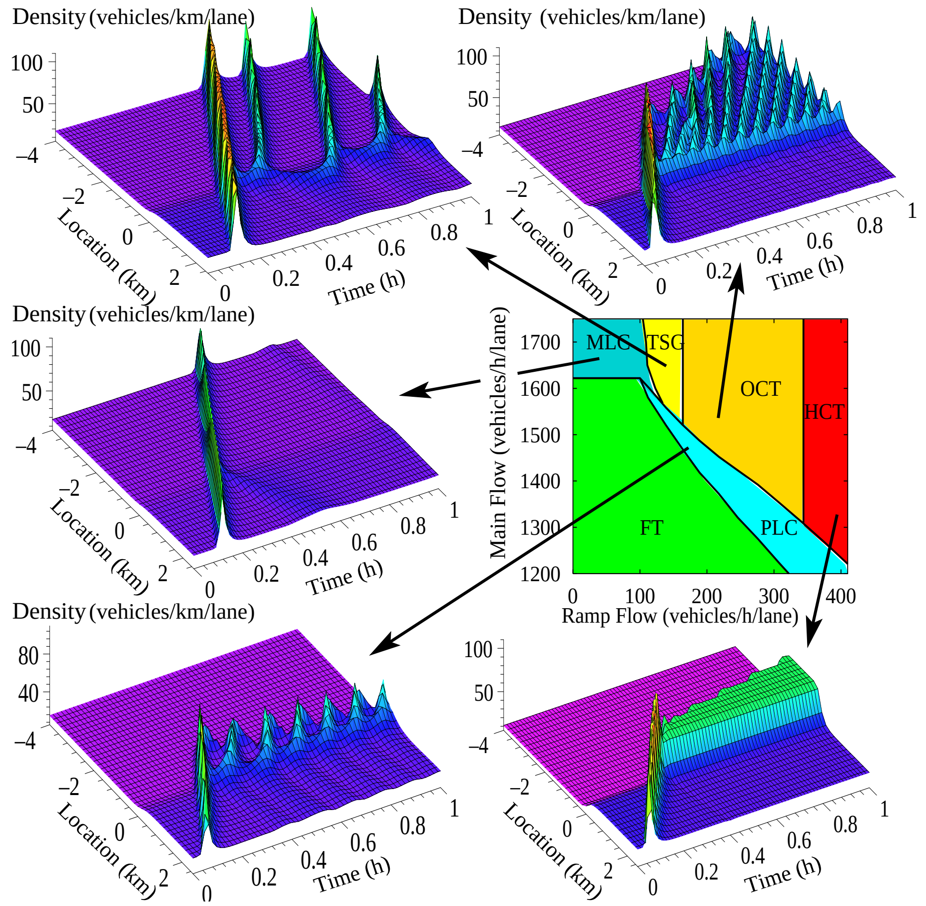

Figure 9: Representatives of different congested traffic states triggered by a large perturbation travelling upstream. The simulation results were obtained with the intelligent driver model (IDM) for identical vehicles. Heterogeneous driver-vehicle units would lead to less regular structures [75]. Center right: Phase diagram of the traffic states as a function of the (upstream) traffic volume $Q_{\rm up}$ on the freeway and the strength ΔQ of a bottleneck at location $x = 0$ km, e.g. an on-ramp with inflow ΔQ per freeway lane.

between which traffic flows freely. This state can either emerge from a spatial sequence of homogeneous and oscillating congested traffic (so-called *"pinch effect"* [33]), or it can be caused by the inhomogeneity at the ramp. In the latter case, each traffic jam triggers another one by inducing a small perturbation in the inhomogeneous freeway section (see Fig. 9), which propagates downstream as long as it is small, but turns back when it has grown large enough (*"boomerang effect"*, cf. Figs. 7 to 10). This, however, requires

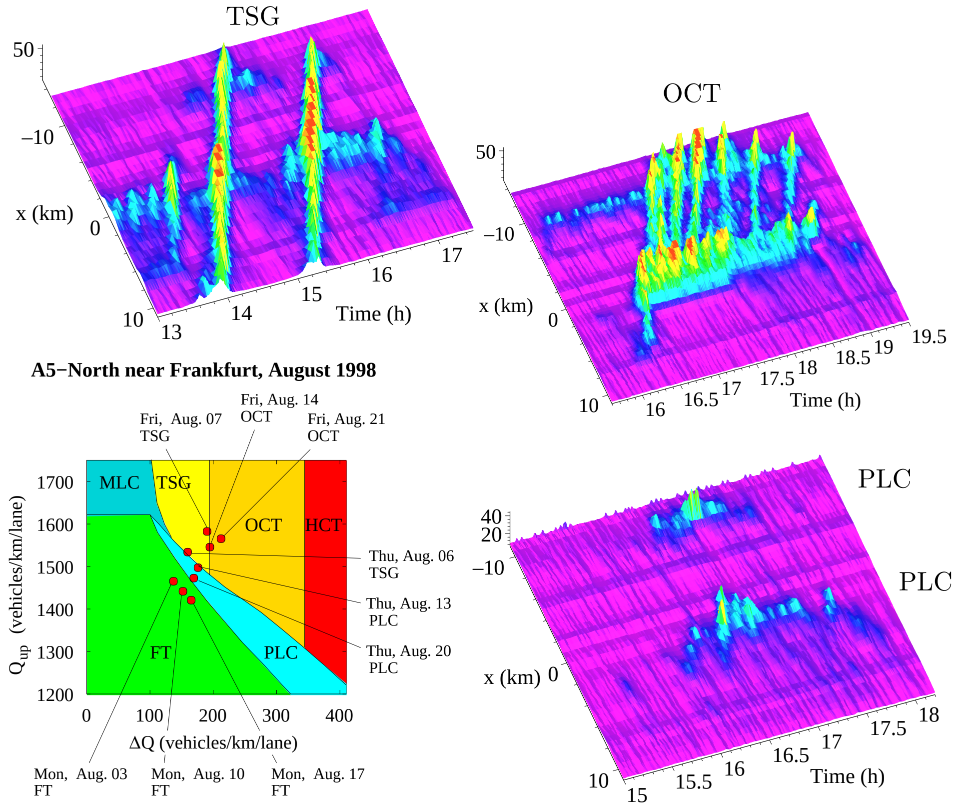

Figure 10: Empirical examples of triggered stop-and-go traffic (TSG), oscillating congested traffic (OCT), and pinned localized clusters (PLC), and the location of the empirical data points in the phase diagram for the German freeway A5 near Frankfurt. The full data set contained data of two weeks only. A more detailed study is currently carried out with new traffic data of 180 days, containing examples of the less frequent moving localized cluster (MLC) and homogeneous congested traffic (HCT) states.

the downstream traffic flow to be linearly unstable. If it is metastable instead (when the traffic volume is further reduced), a traffic jam will usually not trigger a growing perturbation. In that case, one finds either a single *moving localized cluster* (MLC), or a *pinned localized cluster* (PLC) at the location of the ramp. The latter requires the traffic flow in the upstream section to be stable, so that no traffic jam can survive there. Finally, for sufficiently small traffic volumes, we find *free traffic* (FT), as expected.

The different congested traffic states found in the microsimulations (as displayed in Fig. 9) could all be identified in real traffic data (see Fig. 10 for some examples). Moreover, according to our first investigation results, the traffic patterns observed on the German freeway A5 near Frankfurt have a typical dependence on the respective weekday and are even quantitatively consistent with the phase diagram (see Fig. 10). Of course, the empirically measured patterns look less regular, as the simulation results displayed in Fig. 9 are for a deterministic model with identical vehicle parameters. Presently, we are carrying out a more detailed analysis based on data of 180 different days. In particular, we are searching for traffic patterns not consistent with the above phase diagram, but until now, we have not discovered any phenomena contradicting the above theory of congested traffic.

2.1 Reply to criticisms of the phase diagram

In the following, we will try to clear up the misunderstandings behind the italicized criticisms of the phase diagram of congested traffic states:

1. *On freeways, some observed traffic patterns are more complex than the five congested traffic states mentioned above.* This is usually an effect of the simultaneous presence of several bottlenecks interfering with each other. For example, we find the spatial *coexistence* of states such as OCT and PLC, temporal *transitions* between different states, extended congested traffic states (HCT, OCT, or TSG) with a *fixed upstream front,* and other phenomena. This is, however, still compatible with the above theory, as similar patterns are found in the simulation of these more complex freeway geometries with several bottlenecks. The phenomenon of multistability and coexisting states is, by the way, already found for the case of one single bottleneck (see Ref. [54] and Fig. 8 in Ref. [75]).
2. *Sometimes, moving localized clusters or wide moving jams propagate through areas of stable traffic associated with a pinned localized cluster.* This "tunnel effect" is due to a certain *"penetration depth"* of unstable patterns into regions of stable traffic [75], which is similar to other phenomena found in physics. In *meta*stable traffic, by the way, perturbations have always a chance to survive, if they are large enough.
3. *The variation and scattering of the flow-density data is much higher in empirical measurements than in the above simulations.* It can, however, be well reproduced, if different driver-vehicle types are distinguished (and overtaking maneuvers are taken into account).
4. *A transition of traffic patterns containing moving jams to "synchronized flow" is sometimes found with* decreasing *on-ramp flows* [36]. In these observations, the term "synchronized flow" is usually used for patterns we would classify as pinned localized

clusters (PLC) or oscillating pinned localized clusters (OPLC), which are a subtype of the former ones (see Fig. 9). This judgement is based on traffic data from the same freeway stretch of the German freeway A5 that Kerner uses to analyze. As a consequence, the above mentioned transition observed by Kerner corresponds to a transition from TSG to PLC patterns with decreasing ramp flow, which is fully consistent with the phase diagram of congested traffic states.

Nevertheless, we note that, in congested traffic, drivers seem to change their behavior after some time. Due to a "frustration effect", the average time headway appears to increase and the acceleration strength (i.e. the desired velocity) to decrease in the course of time [25, 47]. This should influence the stability of traffic flow and spatio-temporal transitions between different traffic patterns.

5. *Transitions from "synchronized flow" to stop-and-go patterns can occur for increasing ramp flows.* If the "synchronized flow" corresponds to PLC or OPLC states, this is just the opposite of the last point. If the "synchronized flow" is of HCT type, such observations can be understood as a "pinch effect" with time-dependent ramp flows. The "pinch effect" corresponds to a spatio-temporal transition between HCT, OCT, and TSG states, occuring in a certain area of the phase diagram, i.e. for certain combinations of ramp and freeway flows (which may vary in time, although the "pinch effect" does not require that). This particular area, which is not separately marked in the phase diagram of Fig. 9, is expected to agree with subregions of the OCT and PLC phases, as the congested flow must be linearly unstable like in OCT (but convectively stable) [72, 25], and a (meta-)stable upstream flow like in localized cluster states is one possible mechanism for the disappearance of narrow jams.

6. *According to measurements comparable to Fig. 3(b), homogeneous congested traffic appears at relatively high vehicle speeds, namely in the limit of large wavelengths.* In this limit, we find a localized cluster state (MLC or PLC) with the empirically required properties rather than HCT associated with low vehicle speeds. Homogeneous congested traffic does not correspond to the limit of infinite wavelengths, but to the limit of vanishing oscillation amplitudes. If we start with HCT and decrease the ramp flow in the course of time, our theory of congested traffic states predicts an increasing oscillation amplitude below some critical ramp flow, which gives rise to oscillating congested traffic (OCT). When the oscillation amplitude becomes so large that we have an alternation between free flow and congested traffic, it cannot grow further and the triggered stop-and-go regime (TSG) is entered. A further reduction of the ramp flow will increase the average *wavelength* of the stop-and-go patterns [25]. The empirically observed increase of the vehicle speeds in "synchronized flow" with the average oscillation wavelength [33] is qualitatively well reproduced, see Fig. 3(b).

7. *It has been observed that the flow downstream of congested traffic can exceed the dynamical capacity* $Q_{\rm out}$ [35]. This finding is reproduced by our theory, if vehicles can enter the freeway via the ramp downstream of the congestion front. This is practically relevant for *long* on-ramps like freeway intersections, in particular for the cases studied by Kerner (see the footnote on p. 1111 of Ref. [25]).

Finally, note that the classification of congested traffic states according to the phase diagram in Fig. 9 and Kerner's three-phase traffic theory in Sec. 1.1 do not necessarily contradict each other. In the phase diagram, the states are distinguished by their different spatio-temporal behavior, whereas the classification of the three-phase theory is mainly based on local, time-dependent features like the properties of flow-density data measured at some freeway cross section. Thus, "synchronized flow" can correspond to PLC, HCT, or OCT states, while "wide moving jams" can correspond to MLC or TSG states. The "pinch effect" denotes a spatio-temporal transition between these two types of states.

3 Summary

Nowadays, many aspects of traffic dynamics have been understood on the basis of self-driven many-particle models [25]. The observed phenomena can be (semi-)quantitatively reproduced by simulations using measured boundary conditions [75], even with deterministic models. Moreover, a universal phase diagram of congested traffic states for freeway sections with one bottleneck has been found, and the generalization to more complex situations is straightforward. The reproduction of fine details, however, will require a more refined measurement of the interactive vehicle dynamics and possibly the consideration of psychological aspects [12, 13]. Although these may be described in a mathematical way as well [17, 38, 49, 57, 78], in most cases it will be hardly possible to prove or disprove the corresponding models. There are certainly phenomena that can *only* be satisfactorily understood with psychological concepts, but fewer than expected some years ago. The scientific progress during the last decades has shown that many observations have simpler explanations, which are easier to verify.

Acknowledgments

The authors would like to thank the German Research Foundation (DFG) for financial support through the grant He 2789/2-1. They are also grateful to the *Hessisches Landesamt für Straßen und Verkehrswesen* for providing the traffic data.

References

[1] M. Bando, K. Hasebe, K. Nakanishi, A. Nakayama, A. Shibata & Y. Sugiyama, 1995, Phenomenological study of dynamical model of traffic flow, *J. Phys. I France* **5**, 1389–1399.

[2] J. H. Banks, 1990, Flow processes at a freeway bottleneck, *Transpn. Res. Rec.* **1287**, 20–28.

[3] J. H. Banks, 1991, Two-capacity phenomenon at freeway bottlenecks: A basis for ramp metering? *Transpn. Res. Rec.* **1320**, 83–90.

[4] J. H. Banks, 1995, Another look at a priori relationships among traffic flow characteristics, *Transpn. Res. Rec.* **1510**, 1–10.

[5] J. H. Banks, 1999, An investigation of some characteristics of congested flow, *Transpn. Res. Rec.* **1678**, 128–134.

[6] M. J. Cassidy, R. L. Bertini, 1999, Some traffic features at freeway bottlenecks, *Transpn. Res. B* **33**, 25–42.

[7] M. J. Cassidy & M. Mauch, 2001, An observed traffic pattern in long freeway queues, *Transpn. Res. A* **35**, 143–156.

[8] K. H. Chung & P. M. Hui, 1994, Traffic flow problems in one-dimensional inhomogeneous media, *J. Phys. Soc. Jpn.* **63**, 4338–4341.

[9] M. Cremer, 1979, *Der Verkehrsfluß auf Schnellstraßen* (Springer, Berlin).

[10] M. C. Cross & P. C. Hohenberg, 1993, Pattern formation outside of equilibrium, *Rev. Mod. Phys.* **65**, 851–1112.

[11] Z. Csahók & T. Vicsek, 1994, Traffic models with disorder, *J. Phys. A: Math. Gen.* **27** L591–L596.

[12] C. F. Daganzo, 1999, A behavioral theory of multi-lane traffic flow, Part I: Long homogeneous freeway sections (ITS Working Paper, UCB-ITS-RR-99-5, revised June 20, 2000).

[13] C. F. Daganzo, 1999, A behavioral theory of multi-lane traffic flow, Part II: Merges and the onset of congestion (ITS Working Paper, UCB-ITS-RR-99-6).

[14] C. F. Daganzo, M. J. Cassidy & R. L. Bertini, 1999, Possible explanations of phase transitions in highway traffic, *Transpn. Res. A* **33**, 365–379.

[15] L. C. Edie & R. S. Foote, 1958, Traffic flow in tunnels, *High. Res. Board Proc.* **37**, 334–344.

[16] H. Emmerich & E. Rank, 1995, Investigating traffic flow in the presence of hindrances by cellular automata, *Physica A* **216**, 435–444.

[17] M. Fellendorf, 1996, VISSIM for traffic signal optimisation, in *Traffic Technology International '96* (UK & International, Dorking), pp. 190–192.

[18] T. W. Forbes, J. J. Mullin & M. E. Simpson, 1967, Interchange spacings and driver behavior effects on freeway operations, in *Proceedings of the 3rd International Symposium on the Theory of Traffic Flow*, edited by L. C. Edie (Elsevier, New York, N. Y.), pp. 97–108.

[19] F. L. Hall, 1987, An interpretation of speed-flow-concentration relationships using catastrophy theory, *Transpn. Res. A* **21**, 191–201.

[20] F. L. Hall & K. Agyemang-Duah, 1991, Freeway capacity drop and the definition of capacity, *Transpn. Res. Rec.* **1320**, 91–108.

[21] F. L. Hall, B. L. Allen & M. A. Gunter, 1986, Empirical analysis of freeway flow-density relationships, *Transpn. Res. A* **20**, 197–210.

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

[23] D. Helbing, 1997, Fundamentals of traffic flow, *Phys. Rev. E* **55**, 3735–3738.

[24] D. Helbing, 2001, Die wundervolle Welt aktiver Vielteilchensysteme, *Phys. Bl.* **57**, 27–33.

[25] D. Helbing, 2001, Traffic and related self-driven many-particle systems, *Reviews of Modern Physics* **1067**, 1067–1141.

[26] D. Helbing, A. Hennecke & M. Treiber, 1999, Phase diagram of traffic states in the presence of inhomogeneities, *Phys. Rev. Lett.* **82**, 4360–4363.

[27] D. Helbing & B. A. Huberman, 1998, Coherent moving states in highway traffic, *Nature* **396**, 738–740.

[28] D. Helbing & B. Tilch, 1998, Generalized force model of traffic dynamics, *Phys. Rev. E* **58**, 133–138.

[29] D. Helbing & M. Treiber, 1998, Gas-kinetic-based traffic model explaining observed hysteretic phase transition, *Phys. Rev. Lett.* **81**, 3042–3045.

[30] M. Hilliges, 1995, *Ein phänomenologisches Modell des dynamischen Verkehrsflusses in Schnellstraßennetzen* (Shaker, Aachen).

[31] D. H. Hoefs, 1972, *Untersuchung des Fahrverhaltens in Fahrzeugkolonnen* (Bundesministerium für Verkehr, Abt. Straßenbau, Bonn-Bad Godesberg).

[32] B. S. Kerner, 1998, A theory of congested traffic flow, in *Proceedings of the 3rd International Symposium on Highway Capacity*, edited by R. Rysgaard (Road Directorate, Denmark 1998) Vol. 2, pp. 621–642.

[33] B. S. Kerner, 1998, Experimental features of self-organization in traffic flow, *Phys. Rev. Lett.* **81**, 3797–3800.

[34] B. S. Kerner, 2000, Experimental features of the emergence of moving jams in free traffic flow, *J. Phys. A.: Meth. Gen.* **33**, L221–L228.

[35] B. S. Kerner, 2001, Complexity of synchronized flow and related problems for basic assumptions of traffic flow theories, *Networks and Spatial Economics* **1**, 35–76.

[36] B. S. Kerner, 2002, Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks, *Physical Review E* **65**, 046138.

[37] B. S. Kerner, H. Kirschfink & H. Rehborn, 1999, *Method for the automatic monitoring of traffic including the analysis of back-up dynamics* (German Patent DE 196 47 127.3; US Patent US 5,861,820).

[38] B. S. Kerner & S. L. Klenov, 2002, A micropscopic model for phase transitions in traffic flow, *J. Phys. A: Math. Gen.* **35**, L31–L43.

[39] B. S. Kerner, S. L. Klenov & P. Konhäuser, 1997, Asymptotic theory of traffic jams, *Phys. Rev. E* **56**, 4200–4216.

[40] B. S. Kerner & P. Konhäuser, 1994, Structure and parameters of clusters in traffic flow, *Phys. Rev. E* **50**, 54–83.

[41] B. S. Kerner, P. Konhäuser & M. Schilke, 1995, Deterministic spontaneous appearance of traffic jams in slightly inhomogeneous traffic flow, *Phys. Rev. E* **51**, R6243–R6246.

[42] B. S. Kerner & H. Rehborn, 1996, Experimental properties of complexity in traffic flow, *Phys. Rev. E* **53**, R4275–R4278.

[43] B. S. Kerner & H. Rehborn, 1996, Experimental features and characteristics of traffic jams, *Phys. Rev. E* **53**, R1297–R1300.

[44] B. S. Kerner & H. Rehborn, 1997, Experimental properties of phase transitions in traffic flow, *Phys. Rev. Lett.* **79**, 4030–4033.

[45] B. S. Kerner, H. Rehborn, M. Aleksic, A.Haug & R. Lange, 2000, Verfolgung und Vorhersage von Verkehrsstörungen auf Autobahnen mit 'ASDA' und 'FOTO' im online-Betrieb in der Verkehrsrechnerzentrale Rüsselsheim, *Straßenverkehrstechn.* **10**, 521–527.

[46] B. S. Kerner, H. Rehborn & M. Aleksic, 2000, Forecasting of traffic congestion, in *Traffic and Granular Flow '99*, edited by D. Helbing, H. J. Herrmann, M. Schreckenberg & D. E. Wolf (Springer, Berlin), pp. 339–344

[47] Contributions by M. Kikuchi, M. Koshi, H. Ozaki, and S. Yukawa, 2002, in M. Fukui, Y. Sugiyama & M. Schreckenberg, Eds., *Traffic and Granular Flow '01* (Springer, Berlin), in print.

[48] A. Klar, R. D. Kühne & R. Wegener, 1996, Mathematical models for vehicular traffic, *Surv. Math. Ind.* **6**, 215–239.

[49] W. Knospe, L. Santen, A. Schadschneider & M. Schreckenberg, 2002, Human behavior as origin of traffic phases, *Phys. Rev. E* **65**, 015101 (R).

[50] R. D. Kühne, 1984, Macroscopic freeway model for dense traffic—Stop-start waves and incident detection, in *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, edited by I. Volmuller & R. Hamerslag (VNU Science, Utrecht), pp. 21–42.

[51] R. D. Kühne, 1991, Traffic patterns in unstable traffic flow on freeways, in *Highway Capacity and Level of Service*, Proceedings of the International Symposium on Highway Capacity, edited by U. Brannolte (Balkema, Rotterdam), pp. 211–223.

[52] M. Koshi, M. Iwasaki & I. Ohkura, 1983, Some findings and an overview on vehicular flow characteristics, in *Proceedings of the 8th International Symposium on Transportation and Traffic Flow Theory*, edited by V. F. Hurdle, E. Hauer & G. N. Stewart (University of Toronto, Toronto, Ontario), pp. 403–426.

[53] H. Y. Lee, H.-W. Lee & D. Kim, 1998, Origin of synchronized traffic flow on highways and its dynamic phase transitions, *Phys. Rev. Lett.* **81**, 1130–1133.

[54] H. Y. Lee, H.-W. Lee & D. Kim, 1999, Dynamic states of a continuum traffic equation with on-ramp, *Phys. Rev. E* **59**, 5101–5111.

[55] H. Y. Lee, H.-W. Lee & D. Kim, 2000, Phase diagram of congested traffic flow: An empirical study, *Phys. Rev. E* **62**, 4737–4741.

[56] M. J. Lighthill & G. B. Whitham, 1955, On kinematic waves: II. A theory of traffic on long crowded roads, *Proc. Roy. Soc. London, Ser. A* **229**, 317–345.

[57] J. Ludmann, D. Neunzig & M. Weilkes, 1997, Traffic simulation with consideration of driver models, theory and examples, *Veh. Syst. Dyn.* **27**, 491–516.

[58] Y. Makigami, T. Nakanishi, M. Toyama & R. Mizote, 1983, On a simulation model for the traffic stream in a freeway merging area, in *Proceedings of the 8th International Symposium on Transportation and Traffic Flow Theory*, edited by V. F. Hurdle, E. Hauer & G. N. Stewart (University of Toronto, Toronto, Ontario), pp. 427-451.

[59] P. Manneville, 1990, *Dissipative Structures and Weak Turbulence* (Academic, New York).

[60] A. D. May, 1990, *Traffic Flow Fundamentals* (Prentice Hall, Englewood Cliffs, NJ).

[61] H. S. Mika, J. B. Kreer & L. S. Yuan, 1969, Dual mode behavior of freeway traffic, *Highw. Res. Rec.* **279**, 1–13.

[62] P. K. Munjal, Y.-S. Hsu & R. L. Lawrence, 1971, Analysis and validation of lane-drop effects on multi-lane freeways, *Transpn. Res.* **5**, 257–266.

[63] T. Nagatani, 1997, Instability of a traffic jam induced by slowing down, *J. Phys. Soc. Jpn.* **66**, L1928–L1931.

[64] L. Neubert, L. Santen, A. Schadschneider & M. Schreckenberg, 1999, Single-vehicle data of highway traffic: A statistical analysis, *Phys. Rev. E* **60**, 6480–6490.

[65] H. Payne, 1984, Discontinuity in equilibrium traffic flow, *Transpn. Res. Rec.* **971**, 140–146.

[66] B. N. Persaud, 1986, *Study of a freeway bottleneck to explore some unresolved traffic flow issues* (Ph.D. thesis, University of Toronto, Canada).

[67] B. Persaud, S. Yagar & R. Brownlee, 1998, Exploration of the breakdown phenomenon in freeway traffic, *Transpn. Res. Rec.* **1634**, 64–69.

[68] W. F. Phillips, 1977, *Kinetic Model for Traffic Flow* (Technical Report DOT/RSPD/DPB/50-77/17, National Technical Information Service, Springfield, VA 22161).

[69] V. Shvetsov & D. Helbing, 1999, Macroscopic dynamics of multilane traffic, *Phys. Rev. E* **59**, 6328–6339.

[70] B. Tilch, 2001, *Modellierung und Simulation selbstgetriebener Vielteilchensysteme mit Anwendung auf den Straßenverkehr* (Ph.D. thesis, University of Stuttgart, in preparation).

[71] B. Tilch & D. Helbing, 2000, Evaluation of single vehicle data in dependence of the vehicle-type, lane, and site, in *Traffic and Granular Flow '99*, edited by D. Helbing, H. J. Herrmann, M. Schreckenberg & D. E. Wolf (Springer, Berlin), pp. 333–338.

[72] M. Treiber & D. Helbing, 1999, Explanation of observed features of self-organization in traffic flow (e-print `cond-mat/9901239`).

[73] M. Treiber & D. Helbing, 1999, Macroscopic simulation of widely scattered synchronized traffic states, *J. Phys. A: Math. Gen.* **32**, L17–L23.

[74] M. Treiber, A. Hennecke & D. Helbing, 1999, Derivation, properties, and simulation of a gas-kinetic-based, non-local traffic model, *Phys. Rev. E* **59**, 239–253.

[75] M. Treiber, A. Hennecke & D. Helbing, 2000, Congested traffic states in empirical observations and microscopic simulations, *Phys. Rev. E* **62**, 1805–1824.

[76] J. Treiterer & J. A. Myers, 1974, The hysteresis phenomenon in traffic flow, in *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, edited by D. Buckley (Reed, London), pp. 13–38.

[77] D. Westland, 1998, Effect of highway geometry on freeway queuing at merge sections, in *Proceedings of the 3rd International Symposium on Highway Capacity*, edited by R. Rysgaard (Road Directorate, Denmark), pp. 1095–1116.

[78] R. Wiedemann, 1974, *Simulation des Straßenverkehrsflusses* (Institut für Verkehrswesen, Universität Karlsruhe).

