

This document is a preprint version of the following article:

Title: Evidence of Convective Instability in Congested Traffic Flow: A Systematic Empirical and Theoretical Investigation

By: Treiber, Martin; Kesting, Arne

Published in: Transportation Research Part B: Methodological - Volume 45, Issue 9, November 2011, Pages 1362-1377

Publisher: Elsevier | Year: 2011

DOI: 10.1016/j.trb.2011.05.011



Brought to you by the Scientific
Information Service for Mobility
and Transport Research

Fachinformationsdienst Mobilitäts- und Verkehrsforschung

Website: www.fid-move.de

Repository: publish.fid-move.de

Contact: publish@fid-move.de

© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license.

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Evidence of Convective Instability in Congested Traffic Flow: A Systematic Empirical and Theoretical Investigation

Martin Treiber^{a,*}, Arne Kesting^a

^a*Technische Universität Dresden, Institute for Transport & Economics,
Würzburger Str. 35, 01062 Dresden, Germany*

Abstract

An extended open system such as traffic flow is said to be convectively unstable if perturbations of the stationary state grow but propagate in only one direction, so they eventually leave the system. By means of data analysis, simulations, and analytical calculations, we give evidence that this concept is relevant for instabilities of congested traffic flow. We analyze detector data from several hundred traffic jams and propose estimates for the linear growth rate, the wavelength, the propagation velocity, and the severity of the associated bottleneck that can be evaluated semi-automatically. Scatter plots of these quantities reveal systematic dependencies. On the theoretical side, we derive, for a wide class of microscopic and macroscopic traffic models, analytical criteria for convective and absolute linear instabilities. Based on the relative positions of the stability limits in the fundamental diagram, we divide these models into five stability classes which uniquely determine the set of possible elementary spatiotemporal patterns in open systems with a bottleneck. Only two classes, both dominated by convective instabilities, are compatible with observations. By means of approximate solutions of convectively unstable systems with sustained localized noise, we show that the observed spatiotemporal phenomena can also be described analytically. The parameters of the analytical expressions can be inferred from observations, and also (analytically) derived from the model equations.

Key words: Spatiotemporal dynamics, traffic database, traffic flow oscillations, growth rate, convective instability, analytical results

©2011. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

1 Introduction

Traffic flow dynamics and congestion phenomena have been investigated for decades (Reuschel, 1950; Pipes, 1953) (see also the reviews (Helbing, 2001; Hoogendoorn and Bovy, 2001)). Observations of instabilities date back several decades as well (Treiterer and Myers, 1974), and there is an ongoing flow of new empirical results (Kerner and Rehborn, 1996; Zielke et al., 2008; Schönhof and Helbing, 2007; Treiber and Kesting, 2010a). Some fundamental questions, however, are not yet settled. Can the multitude of observed spatiotemporal patterns of congested traffic be decomposed into precisely defined elementary patterns (Kerner and Rehborn, 1996; Schönhof and Helbing, 2009; Treiber et al., 2010a)? If so, into how many patterns, and what are their defining properties (Schönhof and Helbing, 2007; Kerner, 2002)? What is the relation between traffic patterns and (dynamic) traffic phases (Treiber et al., 2010a)? Particularly, are there two or three phases (Wagner and Nagel, 2008)?

All this is related to oscillations in extended regions of congested traffic, i.e., congested traffic spreading over several detector locations. Such oscillations have been observed frequently, e.g., in Germany, Holland, England, and the USA (Zielke et al., 2008; Bertini and Leal, 2005; Ahn and Cassidy, 2007; Helbing et al., 2009; Wilson, 2008; Treiber et al., 2010b). Their common properties have been summarized by the following qualitative facts (Treiber et al., 2010a).

- (1) Congestion patterns are typically caused by bottlenecks. The downstream front is either stationary at a bottleneck, or moves at a constant velocity c_{cong} . Analyzing about 400 congestion patterns on the German freeways A5-North and A5-South did not bring conclusive evidence of a single breakdown without a bottleneck (Schönhof and Helbing, 2007).
- (2) Most extended traffic patterns on freeways exhibit distinct internal oscillations (Daganzo, 2002a; Mauch and Cassidy, 2002). Only minor or severe bottlenecks may cause nearly homogeneous congested traffic where possible oscillations cannot be distinguished unambiguously from noise (Treiber et al., 2010a).
- (3) The oscillations propagate in upstream direction at the characteristic velocity c_{cong} of the moving downstream fronts (cf. Fact 1) (Li et al., 2010; Orosz et al., 2010; Windover and Cassidy, 2001). This velocity depends weakly on the country and traffic composition (Zielke et al., 2008; Bertini and Leal, 2005; Ahn and Cassidy, 2007), but not on the type of congestion.

* Corresponding author. Tel.: +49 351 463 36794; fax: +49 351 463 36809

Email addresses: treiber@vwi.tu-dresden.de (Martin Treiber),

kesting@vwi.tu-dresden.de (Arne Kesting).

URL: <http://www.mtreiber.de> (Martin Treiber).

- (4) The amplitude of the oscillations increases while propagating upstream (Zielke et al., 2008). As a consequence, one observes stationary traffic near the bottleneck and growing stop-and-go waves further upstream.
- (5) If oscillations are present, their frequency increases with the severity of the bottleneck.

We emphasize that, for the sake of simplicity, we do not consider details of the bottleneck which may be caused by onramps, offramps, gradients, road works, accidents, or combinations thereof. Nor do we consider the immediate region of a bottleneck itself where the dynamics generally depends on the details of the activation mechanism. For example, on US freeways with many lanes, lane separation effects such as the “rabbits and slugs” phenomenon (Daganzo, 2002b) may be relevant in the vicinity of the bottleneck. In contrast, in our data base, we always have observed nearly perfect speed synchronization across all lanes.

Of course, any “theory” must be consistent with the above observations. However, these observations do not settle unambiguously some of the most intriguing questions about the nature of traffic instabilities: Are they of a nonlinear type requiring a finite perturbation for activation, or are the observations consistent with linear instabilities? Can insights into stability properties of homogeneous systems (or closed ring roads) be transferred to real open systems (Helbing et al., 1999, 2009; Treiber et al., 2010a)? Can the above observations be “explained” quantitatively and analytically in terms of the convective instability? These questions are clearly fundamental since instabilities are the main building block for dynamic traffic phases and spatiotemporal patterns (Wagner and Nagel, 2008; Treiber et al., 2010a).

In this paper we investigate the problem from two perspectives. On the empirical side, we analyze detector data of several hundred traffic jams and quantify the above qualitative facts with estimates that can be evaluated semi-automatically. The main part of the paper is dedicated to a theoretical analysis of these results. After providing analytic criteria for the thresholds of absolute and convective linear instabilities for a wide spectrum of microscopic and macroscopic traffic flow models, we divide the models into five classes which uniquely determine the set of observable spatiotemporal patterns in real systems. Both simulations and analytical solutions show that only two of them are consistent with observations. Our main result is that the largely neglected concept of convective instability (Huerre and Monkewitz, 1990) which only very recently has attracted more attention in the traffic flow community (Ward and Wilson, 2011; Wilson and Ward, 2011) has a high explanative power in describing the observed spatiotemporal dynamics of extended traffic jams.

The paper is structured as follows: Section 2 presents a scheme to extract

quantitative aspects of the spatiotemporal evolution of congested traffic patterns from cross-sectional data. In order to provide input for the theory to be developed, the scheme is applied to a database containing several hundred instances of jams. Section 3 gives analytical criteria for limits of absolute and convective string instability that are applicable to a wide range of microscopic and macroscopic traffic models. Moreover, the relative locations of the stability limits are connected to the observed patterns by a newly formulated “class diagram” in parameter space. As an example, the full parameter space of the Intelligent Driver Model (Treiber et al., 2000) is explored. Section 4 gives approximate analytic solutions to the equations of the considered models for realistic congested situations, i.e., open systems with local sustained noise that may represent lane changes at bottlenecks. The implications of these results are discussed in Sec. 5.

2 Extracting quantitative properties of wave propagation from traffic data

In this section, we propose quantitative estimates for the wavelength, propagation velocity and growth rate of oscillations in congested traffic. Since these quantities depend on the severity of the bottleneck, we give data-driven estimates for the “bottleneck strength” as well. Here, we will only give a summary; details are published elsewhere (Treiber and Kesting, 2010a).

2.1 The proposed scheme

We assume that aggregated data for the flow Q and speed V (arithmetic averages) are available from several consecutive detector cross sections i at locations x_i , counted in the direction of traffic flow. Specifically, Q_{it} denotes the lane-averaged flow for cross section i during the t -th time interval, and V_{it} denotes the corresponding speed.

In order to get valid results, it is crucial that the spatiotemporal area in which the data are evaluated is restricted to the region of an extended jam. Furthermore, since most of the estimates are based on cross correlations, only time series over compatible time windows can be used. Together with the Qualitative Facts 1 and 3 (Sec. 1), this leads to the two parallelogram-shaped regions depicted in Fig. 1(a). In this figure, deviations from the average speed (horizontal baselines) are drawn as oscillating curves. The detector spacings are proportional to the spacings of the baselines. Figure 1(b) visualizes the spatiotemporal dynamics using a dedicated reconstruction method (Treiber et al., 2010b). This shows more explicitly that the two parallelogram-shaped

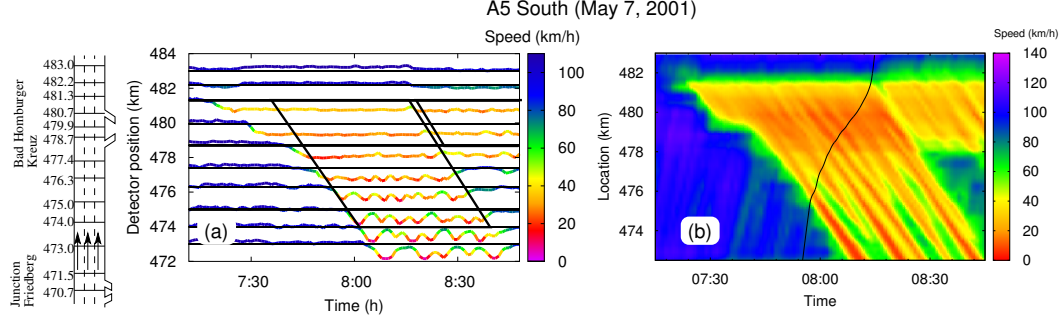


Fig. 1. (a) Speed time series of an extended jam. (b) Visualization of the spatiotemporal dynamics.

regions belong to congestions with different characteristics which justifies to investigate them separately.

The parallelogram-shaped regions can be determined semi-automatically as follows (for details we refer to Ref. (Treiber and Kesting, 2010a)): The downstream and upstream edges of the parallelogram are given, respectively, by the detector location nearest to the bottleneck, and by the location where the waves begin to saturate. The other two edges are parallel to the direction defined by the propagation velocity $\Delta x / \Delta t = c_{\text{cong}}$ of perturbations in congested traffic. Finally, the vertices are defined such that the parallelogram has a maximum area subject to the condition that there is no free traffic anywhere inside. As can be seen in Fig. 1, this may lead to several parallelogram-shaped regions adjacent to each other.

For practical purposes, each parallelogram must be of sufficient size such that the parallelogram includes at least three cross sections, and each cross section records at least three oscillations. Furthermore, for the sake of simplicity, we restrict ourselves to congestions that do not spread over multiple bottlenecks. When propagating over additional bottlenecks, new phenomena may occur such as oscillations with decreasing rather than increasing amplitude. These phenomena are outside the scope of the present paper and possible may be explained otherwise (Treiber et al., 2010a).

One may wonder if these stringent filtering criteria leave enough samples to represent the full spectrum of the oscillation patterns. This is essentially ensured by the qualitative facts for most instances of extensive congestions, i.e., congested traffic covers a spatial range which is significantly larger than the bottleneck region and a time period which is significantly larger than the period of the oscillations. Such patterns include oscillatory congested traffic (OCT) with clearly visible oscillations as in Fig. 1, and essentially stationary patterns (homogeneous congested traffic, HCT) where possible oscillations cannot be distinguished from noise (Treiber et al., 2010a). They also include a part of congested patterns emitting isolated stop-and-go waves (triggered

stop-and go waves, TSG). While, due to their inherently nonlinear nature, the TSG waves itself are not included, the filtering criteria capture the region of smaller oscillations near the bottleneck. As already observed in Ref. (Kerner and Rehborn, 1996), such a “pinch region” is part of most TSG patterns.

The filtering criteria do exclude localized congestions which can be observed either in the form of a single standing wave at a bottleneck (pinned localized cluster, PLC), or as isolated stop-and-go waves (moving localized clusters, MLC), similarly to the stop-and-go waves of the TSG state. In summary, the filter criteria retain nearly all instances where wavelengths, propagation velocities, and growth rates are defined in a meaningful way. Once suitable instances are selected from the database, we determine, for each sample, the characteristic quantities as follows.

The propagation velocity c is determined by maximizing the cross correlations,

$$c = \arg \max_{c'} \sum_i \sum_{j>i} \text{Corr} \left[V_i(t), V_j \left(t + \frac{x_i - x_j}{c'} \right) \right]. \quad (1)$$

In contrast to other investigations using the cross-correlation method (Mika et al., 1969; Coifman and Wang, 2005; Zielke et al., 2008), we take into account all pairs of time series in the relevant region in order to make the measure more robust and to tackle discretization errors (for details, see Ref. (Treiber and Kesting, 2010a)). Pairings of time series with distinct oscillations are automatically weighted more than pairings where one or both of the time series essentially shows noise. The continuous-in-time function $V_i(t)$ is defined by a piecewise linear interpolation of the detector time series V_{it} . The qualitative facts imply that c lies in a very small range around c_{cong} which is confirmed by Fig. 3(b). This makes it possible to define the spatiotemporal region using the a-priori estimate c_{cong} without danger of circular reasoning.

The average wavelength L is determined via the average period τ of the oscillations by the relation $L = |c|\tau$. The period, in turn, is given by the first nontrivial maximum of the autocorrelation function (AKF),

$$\tau = \arg \max_{\tau'>0} \text{Corr} [V_1(t), V_1(t + \tau')]. \quad (2)$$

Since waves may merge during their propagation, the period is not unique and therefore has to be estimated at a defined development stage for which we chose the limit of saturation, i.e., the time series at the upstream edge of the parallelogram-shaped relevant region. Figure 2(a) shows the AKF for several detectors for the situation corresponding to Fig. 1(a). All AKF show a first maximum between 6 and 7 min. However, for the most upstream detector at 473 km, the second maximum is higher than the first maximum indicating the beginning of merges (cf. Fig. 1). Consequently, this detector is outside the parallelogram and the detector at location 474 km is used to evaluate Eq. (2).

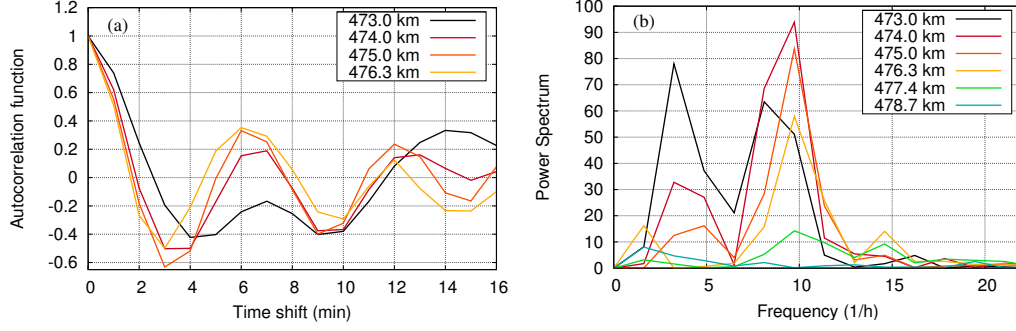


Fig. 2. (a) Autocorrelation function, and (b) spectrum of the speed detector data shown in Fig. 1. The input is restricted to the parallelogram-shaped region shown in this figure.

The same information can also be obtained by a spectral analysis: Figure 2(b) shows the predominant linear mode at a frequency of 9.5 h^{-1} (corresponding to the first nontrivial maximum of the AKF). Furthermore, a nonlinear mode with a significantly lower frequency can be seen for some of the detectors, corresponding to the second nontrivial maximum of the AKF.

A third property, the growth rate, has (to our knowledge) not yet measured systematically. In a first step, the temporal growth rate σ is expressed by the spatial growth rate $\tilde{\sigma}$ via $\sigma = c\tilde{\sigma}$. The spatial growth rate, in turn, is given by the slope of the linear regression of the logarithm of the amplitude,

$$\tilde{\sigma} = \frac{\sum_i x_i \ln |A_i| - n\bar{x} \overline{\ln |A_i|}}{\sum_i x_i^2 - n\bar{x}^2}. \quad (3)$$

The amplitude A_i of the oscillations at the n detector locations x_i is approximated by the standard deviation of the speed time series V_{it} inside the parallelogram. To minimize the risk of biased results, the linear trend is removed from the time series before application of Eq. (3). Nevertheless, $\tilde{\sigma}$ is only a rough approximation of the true spatial growth rate. Particularly, $\tilde{\sigma}$ is influenced by non-collective fluctuations caused by random noise and driver-vehicle heterogeneity as well as by saturation effects. Moreover, the standard deviation gives a strongly biased result if the time series contains periods of free traffic (which should be excluded by the filtering criteria). All these effects systematically decrease the estimated growth rate but do not invalidate the overall picture if they are taken into account when evaluating the simulations as well. Notice that, in general, $\tilde{\sigma}$ is negative, and σ is positive.

According to the qualitative facts mentioned in the Introduction (Sec. 1), the characteristic quantities L , τ , σ , and $\tilde{\sigma}$ depend on the severity of the bottleneck. So it is necessary to quantify this property as well. Theoretically, the bottleneck strength is defined by a local drop of the capacity that is available for the mainroad traffic (Schönhof and Helbing, 2007). However, this quantity is problematic to measure, particularly, if the bottleneck includes merg-

ing/diverging traffic (e.g., junctions, intersections, or lane closings). Instead, we use as correlated measurable quantity the arithmetic speed average over the downstream edge of the parallelogram (cross section n),

$$\bar{V} = \bar{V}_n. \quad (4)$$

According to Qualitative Fact 4, the oscillations are minimal at this location making \bar{V} a well-defined quantity.

2.2 Application to freeway data

In order to provide input for the theoretical analysis of the next sections, the characteristic quantities of oscillations have been calculated for all applicable instances of traffic congestions observed on a section of the German freeway A5-South near Frankfurt for April and May 2001. Images of the corresponding spatiotemporal speed fields can be viewed at the website www.traffic-states.com. An example is shown in the left parallelogram of Fig. 1. This specific instance of a jam has the characteristic quantities $c_{\text{cong}} \approx -16 \text{ km/h}$, $\tau \approx 6 \text{ min}$, $L_{\text{wave}} \approx 1.6 \text{ km}$, $\tilde{\sigma} \approx -0.4 \text{ km}^{-1}$, and $\sigma \approx 6.4 \text{ h}^{-1}$.

Figure 3 displays the result in form of scatter plots of the dynamic properties as a function of the bottleneck strength which is characterized by the speed near the bottleneck. Each data point corresponds to one parallelogram-shaped region of congested traffic behind one of the junction or intersection bottlenecks shown in the schematics of the site. For comparison, some data points of the propagation velocity and growth rate of perturbations in free traffic (sampled from spatiotemporal areas outside the parallelogram-shaped regions of congested traffic) have been added. Figure 3(a) confirms the qualitative observation that the wavelength decreases with the severity of the bottleneck, i.e., with decreasing speed at the bottleneck. However, the wavelength does not drop below 1 km. Plot 3(b) confirms that the propagation velocity of perturbations in congested traffic is essentially independent of the average vehicle speed (Li et al., 2010; Orosz et al., 2010; Windover and Cassidy, 2001), in contrast to the situation in noncongested traffic. Most instructive is Fig. 3(c): In agreement with the qualitative observations, the growth rate is generally positive, apart for very low average velocities \bar{V} corresponding to extended congestions behind severe bottlenecks. Such congestions typically are a consequence of accident-caused lane closures. Remarkably, it is not confirmed that congested traffic flow behind minor bottlenecks ($\bar{V} > 50 \text{ km/h}$) may be stable as well. However, since this result is based on only four data points, further investigations are necessary.

To conclude this section, we emphasize that the investigated instances of con-

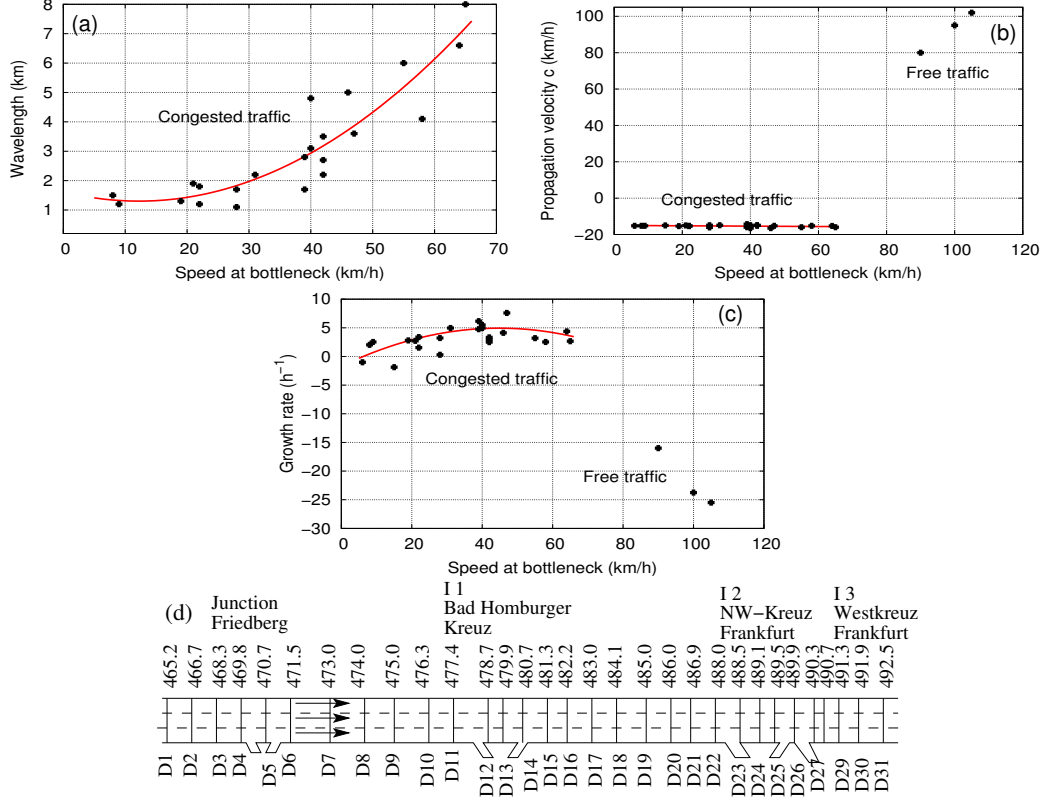


Fig. 3. Characteristic properties of the propagation of perturbations as a function of the speed near the bottleneck.

gested traffic do not spread over multiple bottlenecks. However, they may be caused by different bottlenecks, specifically, by the Junction Friedberg and the intersections I_1 and I_2 , cf. Fig. 3(d). Nevertheless, the data points collapse onto each other within statistical uncertainty. This is strong evidence for our approach to characterize the global dynamical effect of a bottleneck by a single property, the bottleneck strength measured in terms of the speed immediately upstream of it.

3 Stability analysis

In this section, we start with the main theoretical investigation by giving a systematic stability analysis, including an analytic criterion for convective instability that can be evaluated for a wide range of microscopic and macroscopic traffic models.

3.1 Definitions of relevant stability types

For extended open systems with external noise, the analysis is conceptionally much more demanding than for systems with a few degrees of freedom, even on a linear level (Wilson, 2008; Huerre and Monkewitz, 1990). In particular, one has to distinguish between local and string (or flow) instability. Local instability implies that sustained oscillations already arise when following a single vehicle. It can be shown that, for all sensible microscopic models, this type of instability is absent provided that the reaction time is sufficiently low (Orosz and Stépán, 2006; Treiber and Kesting, 2010b). Much more relevant and widespread is the string instability of a platoon of vehicles following each other (Fig. 4), which is also equivalent to the flow instability of macroscopic models. This collective phenomenon may be present even if each driver is perfectly able to follow a single leader with damped or even without oscillations (Kesting and Treiber, 2008).

In extended *open* systems, the string instability can be subdivided into convective and absolute instabilities. In the presence of a convective instability depicted in Fig. 4(a), small oscillations grow but the range of growing amplitudes propagates in only one direction. Consequently, in the absence of sustained noise, the oscillations will eventually leave the road section under consideration. In the traffic flow context, the propagation direction is generally opposite to the driving direction (type “up”, Fig. 5(a)), but a linear propagation purely in driving direction is possible as well, at least theoretically (type “down”, Fig. 5(b)). However, after some time, nonlinearities will destroy the convective nature of the type “down” (Fig. 5(b) after about $t \approx 20$ min), but not of the type “up”. Therefore, convective instability of the “down” type is not relevant for the observed traffic patterns. (In fact, it is never observed in reality.) When the string instability is absolute (Fig. 4(b)), the range of growing perturbations propagates in both directions. Therefore, a single localized perturbation eventually leads to sustained oscillations in the whole region. Notice that, for obvious reasons, the distinction between convective and absolute instability does not exist in closed systems such as ring roads (Sugiyama et al., 2008). Furthermore, all the instabilities discussed above may be linear (can be triggered by an infinitesimal perturbation), or nonlinear (a finite perturbation is necessary). In the latter case, one also speaks of metastability.

Mathematically, the different stability concepts can be defined in terms of the response $U(x, t)$ to a localized initial perturbation of homogeneous traffic flow of density ρ_e in an infinite system (Fig. 5),

$$U(x, 0) = \begin{cases} \epsilon & \text{if } |x| < \frac{1}{2\rho_e}, \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

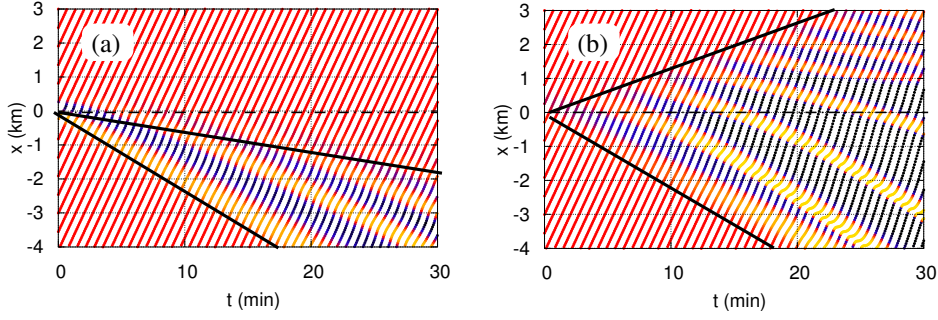


Fig. 4. Visualization of (a) convective upstream instability, and (b) absolute instability by vehicle trajectories in a space-time plot. Shown are IDM simulations with parameters as in Fig. 6, but the acceleration parameter a changed to 1.1 m/s^2 and 0.9 m/s^2 for the plots (a) and (b), respectively. Shown are the trajectories of every 20th vehicle.

where $\epsilon > 0$ and $U(x, 0)$ denotes the initial speed perturbation. The right-hand side of Eq. (5) is defined such that, microscopically, the perturbation corresponds to an instantaneous speed change of a single vehicle (the vehicle nearest to the location $x = 0$) by the amount of ϵ . This follows from the fact that (i) the perturbation integrated over all vehicles gives $\int U(x, t) \rho(x, t) dx = \epsilon$, and (ii) the perturbation is nonzero only in a road section of length $\Delta x = 1/\rho_e$ corresponding to the space attributed to one vehicle. Traffic flow is linearly string unstable if

$$\lim_{t \rightarrow \infty} \max_x U(x, t) > 0 \quad \text{for any } \epsilon > 0, \quad (6)$$

and nonlinearly string unstable (metastable) if a finite $\epsilon_{\text{nl}} > 0$ exists such that

$$\lim_{t \rightarrow \infty} \max_x U(x, t) = \begin{cases} U_0 > 0 & \text{if } \epsilon > \epsilon_{\text{nl}}, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Furthermore, traffic flow is convectively unstable if it is linearly unstable and if, additionally,

$$\lim_{t \rightarrow \infty} U(0, t) = 0. \quad (8)$$

This is illustrated by the simulated trajectories of Fig. 4 showing the response to a localized and instantaneous perturbation at $t = 0$ near $x = 0$. In the left graphics showing the case of (upstream-propagating) convective instability, the cone-shaped region of oscillations does not include the dash-dotted location of the initial perturbation ($U(0, t) = 0$ after some time). In contrast, in Fig. 4(b) displaying absolute string instability, the perturbations eventually propagate everywhere such that $U(0, t)$ does not revert to zero after some time.

In analogy to definition (7), the latter condition can also be generalized on a

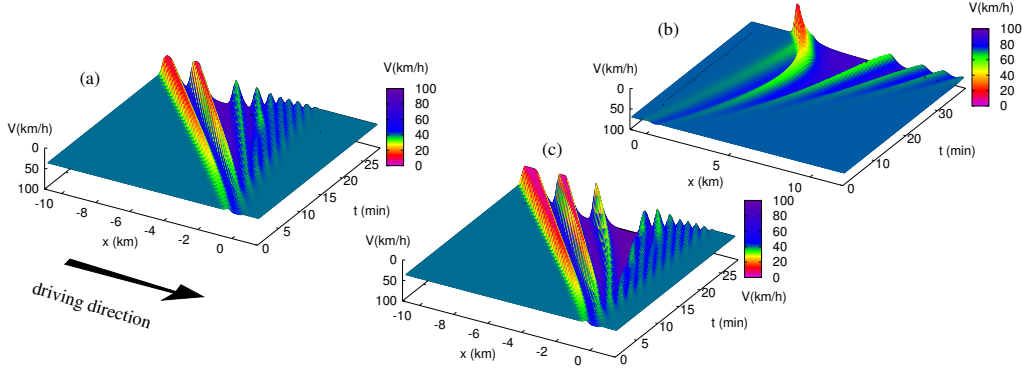


Fig. 5. Speed response functions for the initial localized perturbation (5). Shown are IDM simulations of (a) convective instability propagating upstream, (b) linear convective instability propagating downstream which is destroyed by nonlinearities, and of (c) the limit between convective and absolute instability.

nonlinear level (“convective metastability”). Notice that the above definitions are formulated in terms of directly observable quantities and explicitly include the effects of nonlinearities. For example, the response shown in Fig. 5(b) corresponds to a (downstream-propagating) convective instability in the linear regime, but does not satisfy the general condition (8). In contrast, both the existence and the threshold of upstream-propagating convective instabilities (Fig. 5((a) and (c), respectively), are not altered by nonlinearities and therefore can be determined by a linear analysis.

The above definitions take into account nonlinearities at two stages: Nonlinearities at the initial perturbation (giving rise to metastability) are defined by $\epsilon_{nl} > 0$ while nonlinearities at later stages (possibly destroying the initial convective nature of the perturbation) are described by the response field $U(x, t)$.

3.2 String instability in traffic flow models

We now give analytical criteria for a wide class of car-following and macroscopic models. The considered car-following models are formulated in continuous time by dynamical equations for the speed $v_\alpha(t)$ of each vehicle α (counted against the direction of the traffic flow), and for its (bumper-to-bumper) gap $s_\alpha(t)$ to the leading vehicle $\alpha - 1 = l$. We consider models of the mathematical form

$$\frac{ds_\alpha}{dt} = v_{\alpha-1} - v_\alpha \equiv v_l - v_\alpha, \quad \frac{dv_\alpha}{dt} = a_{\text{mic}}(s_\alpha(t), v_\alpha(t), v_l(t)). \quad (9)$$

The microscopic acceleration function $a_{\text{mic}}(s, v, v_l)$ characterizes the specific microscopic model and must satisfy the plausibility conditions (Wilson and Ward, 2011; Treiber and Kesting, 2010b)

$$\frac{\partial a_{\text{mic}}}{\partial s} \geq 0, \quad \frac{\partial a_{\text{mic}}}{\partial v} < 0, \quad \frac{\partial a_{\text{mic}}}{\partial v_l} \geq 0, \quad (10)$$

as well as

$$a_{\text{mic}}(\infty, v_0, v_l) = a_{\text{mic}}(0, 0, v_l) = 0. \quad (11)$$

From this it follows that, for identical driver-vehicle units, the considered model class possesses a one-parameter family of homogeneous-stationary equilibrium states characterized by the “microscopic fundamental diagram” $v_e(s)$ in which the equilibrium speed function is defined by $a_{\text{mic}}(s, v_e(s), v_e(s)) = 0$ and satisfies $v_e(0) = 0$, $v'_e(s) \geq 0$, and $v_e(\infty) = v_0$.

In order to investigate string instability, we consider traffic flow in an infinite homogeneous system which initially is in a steady state defined by $s_\alpha(t) = s_e$ and $v_\alpha(t) = v_e$. This is only a solution to (9) if s_e and v_e are related by the already mentioned microscopic fundamental diagram. Now we introduce perturbations in the form of linear modes with a common real-valued wavenumber k and amplitudes \tilde{s} and \tilde{v} :

$$\begin{aligned} s_\alpha(t) &= s_e + \tilde{s}e^{\lambda t + i\alpha k}, \\ v_\alpha(t) &= v_e + \tilde{v}e^{\lambda t + i\alpha k}, \end{aligned} \quad (12)$$

where $i = \sqrt{-1}$. Inserting the perturbations into (9) shows that this ansatz is only consistent with the general car-following model if the complex growth rate $\lambda(k)$ depends on the wavenumber k according to the quadratic equation

$$\lambda^2 - \left(\frac{\partial a_{\text{mic}}}{\partial v} \Big|_e + e^{-ik} \frac{\partial a_{\text{mic}}}{\partial v_l} \Big|_e \right) \lambda + (1 - e^{-ik}) \frac{\partial a_{\text{mic}}}{\partial s} \Big|_e = 0. \quad (13)$$

Here, the subscript e stands for “taken at equilibrium”, i.e., $s = s_e$, and $v = v_l = v_e$. The stationary state is string stable if and only if the real part of λ is nonpositive for both solutions of the quadratic equation and for all possible wavenumbers k . As elaborated in detail elsewhere (Wilson and Ward, 2011; Treiber and Kesting, 2010b), this leads to following criterion for string instability:

$$v'_e(s_e) > \frac{1}{2} \left(\frac{\partial a_{\text{mic}}}{\partial v_l} \Big|_e - \frac{\partial a_{\text{mic}}}{\partial v} \Big|_e \right). \quad (14)$$

A typical example of such a model is the Intelligent Driver Model (IDM) (Treiber et al., 2000). It is characterized by the acceleration function

$$a_{\text{IDM}}(s, v, v_l) = a \left[1 - \left(\frac{v}{v_0} \right)^\delta - \left(\frac{s^*}{s} \right)^2 \right], \quad s^* = s_0 + vT + \frac{v(v - v_l)}{2\sqrt{ab}}, \quad (15)$$

where v_0 is the desired speed, a and b are comfortable accelerations and decelerations, δ is the acceleration exponent, s_0 denotes the minimum gap, and T gives the desired time headway in car-following situations. For this model, the instability criterion (14) becomes (Treiber and Kesting, 2010b)

$$(v'_e)^2 > \frac{a(s_0 + v_e T)}{s_e^2} \left[\frac{s_0 + v_e T}{s_e} + \frac{v_e v'_e}{\sqrt{ab}} \right]. \quad (16)$$

Obviously, string instability requires sufficiently low values of the acceleration parameter a and time headway T . For the limiting case of very congested traffic, $s_e \rightarrow s_0$, condition (16) reduces to $a < s_0/T^2$.

The class of considered macroscopic models consists of second-order models that are formulated in continuous space and time by the equations

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho V)}{\partial x} = 0, \quad (17)$$

$$\frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} + \frac{1}{\rho} \frac{\partial P(\rho)}{\partial x} = A(\rho, V, \rho_a, V_a, \rho_x, V_x, \rho_{xx}, V_{xx}). \quad (18)$$

Here, the “traffic pressure” $P(\rho)$ comprises the macroscopic kinematic effects of a finite microscopic speed variance (Treiber et al., 1999). The macroscopic acceleration function $A(\cdot)$ may depend on the density $\rho(x, t)$, the speed field $V(x, t)$, and on the gradients, $\rho_x = \frac{\partial \rho(x, t)}{\partial x}$, $\rho_{xx} = \frac{\partial^2 \rho(x, t)}{\partial x^2}$, v_x , and v_{xx} (defined in analogy). Furthermore, the acceleration function may depend on the fields calculated at an “anticipated” position,

$$\rho_a = \rho(x_a, t), \quad V_a = V(x_a, t), \quad \text{where } x_a > x. \quad (19)$$

A typical representative of this class of nonlocal models is the gas-kinetic based traffic model (GKT) (Treiber et al., 1999).

For local macroscopic models such as the Payne model (Payne, 1979) or the Aw-Rascle model (Aw and Rascle, 2000), the linear long-wavelength analysis leads to a generalization of the string instability criterion proposed in Ref. (Helbing, 2001),

$$(\rho_e V'_e)^2 \leq P'_e - \rho_e \left(V'_e \frac{\partial A}{\partial V_x} + \frac{\partial A}{\partial \rho_x} \right)_e, \quad (20)$$

where $P'_e = P'(\rho_e)$ and $V'_e = V'(\rho_e)$. The other quantities are evaluated at equilibrium (subscript e) as well. Remarkably, the second derivatives ρ_{xx} and V_{xx} do not enter this criterion. For nonlocal models, there is no need for gradient terms in the acceleration function. A long but straightforward calculation

results in the string instability criterion

$$(\rho_e V_e')^2 \leq P_e' - \rho_e s_a \left(V_e' \frac{\partial A}{\partial V_a} + \frac{\partial A}{\partial \rho_a} \right)_e, \quad (21)$$

where $s_a = x_a - x$ denotes the anticipation distance in the stationary situation. Simulations with the GKT model confirmed the correctness of this result (Fig. 7).

3.3 Convective instability

In deriving approximate analytic criteria for the linear case for both microscopic and macroscopic models, we directly evaluate the definition (8) by calculating a linear response function (Green's function) $G(x, t)$ to the ideally localized instantaneous unit perturbation

$$U(x, 0) = G(x, 0) = A_0 \delta(x), \quad A_0 = 1 \text{ m}^2/\text{s}, \quad (22)$$

and determining the growth rate of its amplitude for the ray $x(t) = 0$. The δ -distribution (satisfying $\int_a^b \delta(x) dx = 1$ if $0 \in [a, b]$ and zero, otherwise) is an abstraction of the initial condition (5). We first expand the solution into Fourier modes. For the macroscopic case, we set

$$\begin{pmatrix} \rho(x, t) - \rho_e \\ V(x, t) - V_e \end{pmatrix} \equiv \begin{pmatrix} \rho_1(x, t) \\ G(x, t) \end{pmatrix} = \text{Re} \left[\int_{-\infty}^{\infty} dk e^{ikx} \begin{pmatrix} \hat{\rho}(k) \\ \hat{G}(k) \end{pmatrix} e^{\lambda^*(k)t} \right], \quad (23)$$

where λ^* is the complex conjugate of λ . Here, we have only considered the branch of “slow modes” containing the instability (there is another set of fast relaxing modes). A similar ansatz applies for the microscopic models. Inserting Eq. (23) into the model equations (17) and (18) leads to a complex growth rate of the functional form

$$\lambda(k) = \sigma(k) + i\omega(k) = -\frac{p(k)}{2} \pm \sqrt{\frac{p^2(k)}{4} - q(k)}, \quad (24)$$

where the sign is chosen to obtain the slow mode. The complex coefficients depend on the model class. For the nonlocal macroscopic models, the coefficients are given by

$$p(k) = - \left(2ikV_e + \frac{\partial A}{\partial V} + \frac{\partial A}{\partial v_a} e^{-iks_a} \right)_e, \quad (25)$$

$$q(k) = ikV_e \left(\frac{\partial A}{\partial V} + \frac{\partial A}{\partial V_a} e^{-iks_a} \right)_e - ik\rho_e \left(\frac{\partial A}{\partial \rho} + \frac{\partial A}{\partial \rho_a} e^{-iks_a} \right)_e + k^2(P_e' - V_e^2) \quad (26)$$

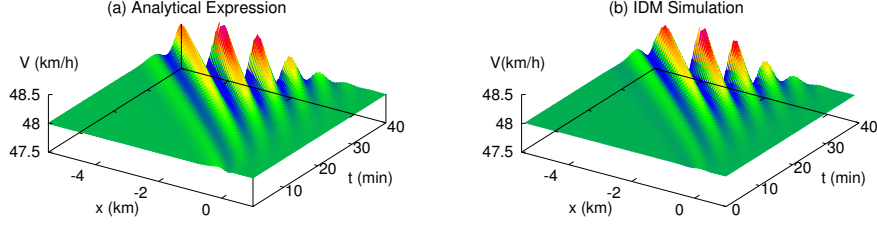


Fig. 6. Traffic flow at the limit between convective and absolute string instability for the Intelligent Driver Model (IDM). Shown is the linear response to the perturbation (22) of the stationary state $V_e = 48$ km/h (a) according to the analytical expression (34), (b) for the IDM simulation. The IDM parameters are $v_0 = 120$ km/h, $T = 1.5$ s, $s_0 = 2$ m, $a = 1.04$ m/s², and $b = 1.5$ m/s². The vehicle length is set to 5 m.

Similar expressions can be found for the local macroscopic models. For microscopic models of the form (9), the coefficients

$$p_{\text{mic}}(k) = - \left(\frac{\partial a_{\text{mic}}}{\partial v} + \frac{\partial a_{\text{mic}}}{\partial v_l} e^{-ik} \right)_e, \quad (27)$$

$$q_{\text{mic}}(k) = (1 - e^{-ik}) \left(\frac{\partial a_{\text{mic}}}{\partial s} \right)_e \quad (28)$$

relate to the “microscopic” growth rate $\lambda_{\text{mic}} = -p_{\text{mic}}/2 \pm \sqrt{p_{\text{mic}}^2/4 - q_{\text{mic}}}$ which is connected to the growth rate of the macroscopic solutions via

$$\lambda(k) = \lambda_{\text{mic}}(k/\rho_e) + i v_e k. \quad (29)$$

The Fourier amplitudes $\hat{\rho}(k)$ and $\hat{G}(k)$ are determined by the initial conditions. For the localized initial condition (22), one obtains $\hat{G}(k) = A_0(2\pi)^{-1}$ while $\hat{\rho}(k)$ depends on the ratio of the density and speed amplitudes of the eigenvector for the corresponding slow mode. With (23), this leads to

$$G(x, t) = \text{Re}[\tilde{G}(x, t)], \quad \tilde{G}(x, t) = \frac{A_0}{2\pi} \int_{-\infty}^{\infty} dk \exp[ikx + \lambda^*(k)t]. \quad (30)$$

In order to approximatively solve this complex integral, we assume that only modes near the fastest growing mode (growth rate $\sigma_0 > 0$, wavenumber k_0) contribute significantly. This allows for an expansion of $\lambda(k)$ around this mode,

$$\lambda(k) = \sigma_0 + i[\omega_0 + v_g(k - k_0)] + \frac{1}{2}(\sigma_{kk} + i\omega_{kk})(k - k_0)^2 + \mathcal{O}[(k - k_0)^3]. \quad (31)$$

Here,

$$k_0 = \arg \max_k (\text{Re } \lambda(k)), \quad (32)$$

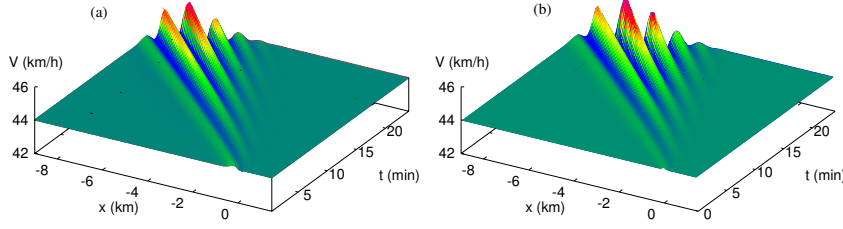


Fig. 7. Convectively unstable traffic flow for the GKT model for a homogeneous density $\rho_e = 38$ veh./km. Shown is the linear response to the localized perturbation (22) (a) according to the analytical expression (34), (b) for the simulation.

and the expansion coefficients are given by

$$\begin{aligned}\sigma_0 &= \text{Re}(\lambda(k_0)), \quad \omega_0 = \text{Im}(\lambda(k_0)), \quad v_g = \text{Im}(\lambda'(k_0)), \\ \sigma_{kk} &= \text{Re}(\lambda''(k_0)), \quad \omega_{kk} = \text{Im}(\lambda''(k_0)).\end{aligned}\tag{33}$$

Inserting this in Eq. (30) and solving the resulting complex Gaussian integral gives the final expression for the complex response function,

$$\tilde{G}(x, t) = \frac{A_0}{\sqrt{2\pi\lambda_2 t}} \exp[i(k_0 x - \omega_0 t)] \exp\left[\left(\sigma_0 - \frac{(v_g - \frac{x}{t})^2}{2\lambda_2}\right)t\right],\tag{34}$$

where $\lambda_2 = i\omega_{kk} - \sigma_{kk}$.

In order to test the accuracy, we have compared the analytical result (34) with simulated response functions for the microscopic IDM (Fig. 6) and the macroscopic GKT model (Fig. 7). For the GKT model equations see Ref. (Treiber et al., 1999). The GKT model parameters are set to $V_0 = 128$ km/h, $T = 1.6$ s, $\tau = 31$ s, $\rho_{\max} = 140$ veh./km, and the anticipation factor $\gamma = 1.2$. The parameters of the velocity variance function (again, see Ref. (Treiber et al., 1999)) are $\alpha_0 = 0.008$, $\Delta\alpha = 0.012$, $\rho_c = 0.27 \rho_{\max}$, and $\Delta\rho = 0.1 \rho_{\max}$. In spite of all the approximations, we have found a nearly perfect agreement including fine details for both the IDM and the GKT model.

An analytic criterion for convective instability is obtained from the response function by inserting it into the defining condition (8). Traffic flow is convectively unstable, if it is linearly unstable ($\sigma_0 > 0$), but the real growth rate σ_{conv} of the amplitude of $\tilde{G}(x, t)$ along the ray $x(t) = 0$ is negative. From (34), one obtains

$$\sigma_{\text{conv}} = \sigma_0 - \frac{v_g^2}{2D_2}, \quad D_2 = -\sigma_{kk} \left(1 + \frac{\omega_{kk}^2}{\sigma_{kk}^2}\right).\tag{35}$$

Consequently, the criterion for convective instability can be expressed by the

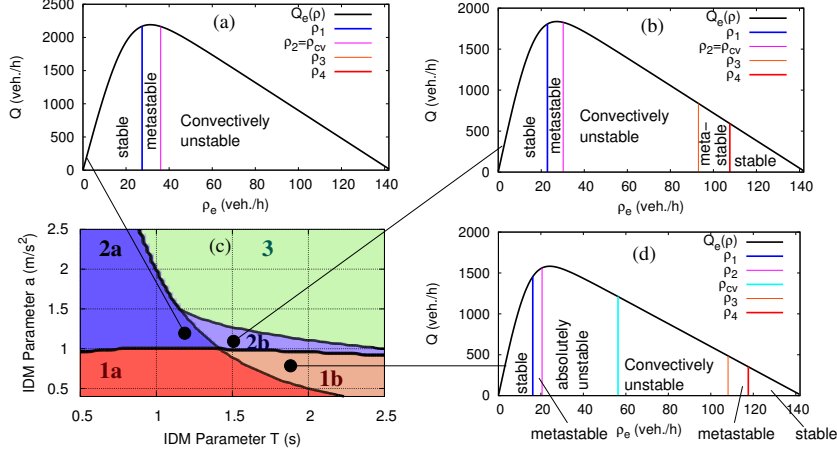


Fig. 8. Class diagram of the IDM as a function of the time gap T and acceleration a , and stability diagrams for three points of the class diagram corresponding to the class 1b, 2a, and 2b. The other IDM parameters and the vehicle length are the same as in Fig. 6.

conditions

$$0 < \sigma_0 \leq \frac{v_g^2}{2D_2}. \quad (36)$$

This concludes the derivation of an analytic criterion for convective instability. It should be noted that criterion (36) implies that the linear string instability always starts as a convective one. Often, the instability remains convective for the whole range of equilibrium densities (Fig. 8(a) and (b)). For typical congested situations, the group velocity v_g is of the order of (but not equal to) the propagation velocity c_{cong} , and the dispersion coefficient D_2 is of the order of $500 \text{ m}^2/\text{s}$.

Simulations with the IDM and the GKT model show that criterion (36) defines the actual density range of convective instability, e.g. that of Fig. 8(d), with errors below 1%. For illustration, Fig. 6 plots the response to the perturbation (22) for traffic flow modeled by the IDM at the limit of convective instability. In order to fulfill the threshold condition $\sigma_0 = v_g^2/(2D_2)$ implied by (36), the IDM acceleration parameter has been set to $a = 1.04 \text{ m/s}^2$. The simulation result displayed in Fig. 6(b) shows that (i) the analytical expression (30) is essentially identical to the simulation result (Fig. 6(a)), (ii) both the simulation and the analytic formula lead to a stationary downstream boundary of the congestion (although no bottleneck is present) which confirms that traffic flow is, in fact, at the threshold between convective and absolute instability.

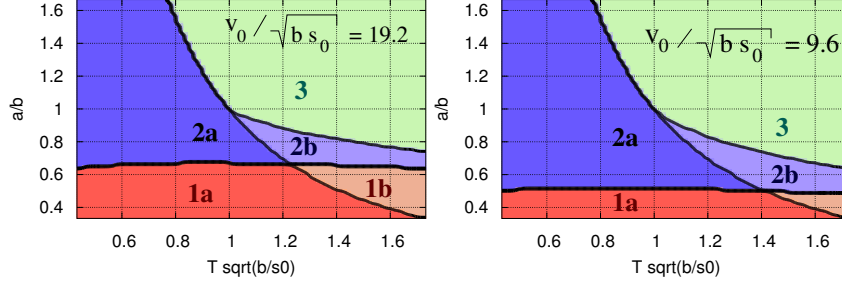


Fig. 9. Class diagram for the complete IDM parameter space. The stability class depends only on the parameter combinations $v'_0 = v_0/\sqrt{bs_0}$, $T' = T\sqrt{b/s_0}$, and $a' = a/b$. The left diagram ($V' = 19.2$) relates to freeway situations, and the right diagram ($V' = 9.6$) to city traffic.

3.4 Stability classes and class diagram

The overall stability properties of a traffic flow model are given by the stability diagram denoting the stability type for all possible equilibrium states parameterized by the associated density $\rho_e = 1/(s_e + l)$. Generally, one obtains several density regions with the boundaries $0 < \rho_1 \leq \rho_2 \leq \rho_{cv} \leq \rho_3 \leq \rho_4 \leq \rho_{\max}$, see Fig. 8. Homogeneous traffic flow is absolutely stable for $\rho < \rho_1$ or $\rho \geq \rho_4$, metastable for $\rho \in [\rho_1, \rho_2]$ or $[\rho_3, \rho_4]$, convectively unstable for $\rho \in [\rho_2, \rho_{cv}]$, and absolutely unstable for $\rho \in [\rho_{cv}, \rho_3]$. Furthermore, the number and types of traffic patterns do not only depend on the set of observed stability types, but also on the relative position of the stability boundaries with respect to the density at (static) capacity $\rho_K = \arg \max_{\rho} Q_e(\rho)$. However, not all combinations of the relative positions of the stability limits ρ_i are relevant. In fact, only five “stability classes” lead to significantly different sets of traffic patterns (cf. Fig. 8):

- Class 1a: Traffic at capacity (density ρ_K) is (linearly) unstable, $\rho_2 < \rho_K$, and remains so for all higher densities, $\rho_3 = \rho_4 = \rho_{\max}$. Furthermore, $\rho_K < \rho_{cv} < \rho_{\max}$, i.e., the instability is absolute for comparatively light congested traffic and generally becomes convective for higher densities.
- Class 1b: As class 1a, but traffic restabilizes for heavily congested traffic, $\rho_3 < \rho_{\max}$.
- Class 2a: Traffic at capacity is metastable or stable, and unstable for sufficiently high densities, $\rho_K < \rho_2 < \rho_3 = \rho_{\max}$. In most cases, the instability is purely convective, but a small region of absolute instability is possible as well.
- Class 2b: As class 2a, but with restabilization, $\rho_3 < \rho_{\max}$.
- Class 3: Unconditionally stable, $\rho_1 = \rho_{\max}$.

The stability class depends on the model and on the parameters. Some models can even be “tuned” to any stability class. Figure 8 displays such a “class diagram” for the IDM.

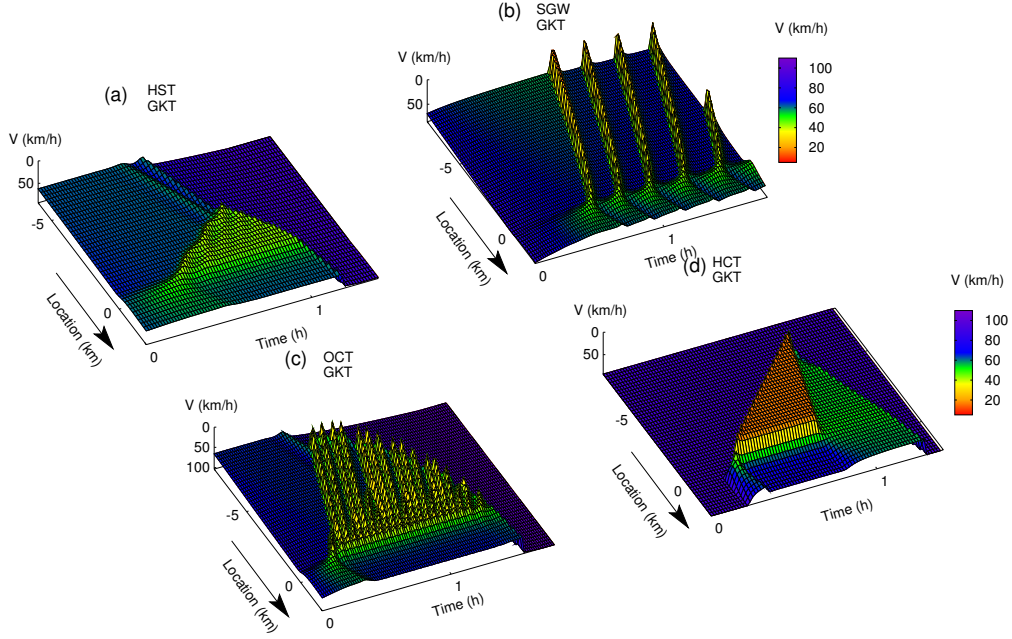


Fig. 10. Simulations of the relevant elementary traffic patterns in an open system with an onramp bottleneck using the GKT model with the same parameters as in Fig. 7 relating to stability class 2b. (a) Low onramp flow such that $\rho_{\text{cong}} < \rho_2$; (b) and (c) intermediate onramp flows such that $\rho_2 < \rho_{\text{cong}} < \rho_3$; (d) severe bottleneck (high ramp flows) such that $\rho_{\text{cong}} > \rho_3$.

By scaling space in units of the minimum distance, and time in units of $\sqrt{s_0/b}$, the number of IDM parameters is reduced to three, and class diagrams can be given for the complete parameter space (Fig. 9). Since, in general, the desired acceleration a and comfortable deceleration b are similar in magnitude ($a' = a/b \approx 1$), the IDM exhibits stability classes 2a or 2b for realistic parameterizations (or class 3 for city traffic where b is higher and s_0 smaller).

The stability and class diagrams are crucial for a possible explanation of the observed extended traffic patterns. The reasoning is as follows: After a traffic breakdown at a bottleneck, the average traffic flow of the congested state upstream is determined by the dynamic capacity $K_{\text{dyn}} = K(1 - \epsilon)$ of the bottleneck (which is smaller than the static capacity by a capacity drop of the order of $\epsilon = 10\%$). Consequently, the congestion pattern depends on the stability properties of congested traffic at flow K_{dyn} .

For stability class 1, minor bottlenecks (high values of K_{dyn}) correspond to absolutely unstable traffic, i.e., congested traffic is non-stationary everywhere and consists of a sequence of stop-and-go waves triggered by the bottleneck (TSG) (Helbing et al., 2009). The less severe the bottleneck, the lower the difference $Q_{\text{out}} - K_{\text{dyn}}$ between the outflow of the waves and the bottleneck capacity, i.e., the longer it takes until there is a sufficient number of vehicles to

trigger the next stop-and-go wave. Conversely, the wave frequency increases with the severity of the bottleneck until the stop-and-go waves are no longer isolated but part of oscillatory congested traffic (OCT). For class 1b, traffic becomes stable for very severe bottlenecks resulting in homogeneous congested traffic (HCT). For class 1b, OCT (with very high frequencies) persists until the complete standstill.

For stability class 2 (see Fig. 10), congested traffic is linearly stable for minor bottlenecks, so a new high-flow homogeneous state (homogeneous synchronized traffic, HST) is possible (Helbing et al., 2009). However, since traffic is generally metastable in this situation, stronger perturbations can lead to stop-and-go waves as well. When increasing the bottleneck strength, one obtains OCT (class 2a) or the sequence OCT-HCT (class 2b), similarly to stability class 1. Unlike class 1, however, the instability is nearly always a convective one, so perturbations grow, but only in upstream direction. As will be shown analytically in Sec. 4, this leads to essentially stationary traffic flow near the bottleneck even if sustained finite perturbations are present. Since this sequence of patterns (stationary traffic at the bottleneck, oscillations further upstream) is observed persistently, we conclude that the dynamics of traffic flow is described by the stability classes 2a or 2b.

For stability class 3, congested traffic is always essentially homogeneous, and the dynamics similar to that of first-order models (Lighthill and Whitham, 1955; Richards, 1956).

4 Sustained perturbations

In the previous section, we have defined convective instability in the traffic context and provided an analytical criterion for its presence in a wide range of models. As defining criterion, oscillations (triggered by a temporary and local perturbation) grow but eventually propagate out of homogeneous open systems. In this section, we consider realistic extended open systems with bottlenecks and noise and show analytically that, instead of propagating out of the system, the convective instability manifests itself in congested pattern with stationary traffic near the bottleneck, and growing oscillations further upstream.

To this end, we investigate the statistical properties of solutions to models of the class (9) with an additional localized noise term in the acceleration equation, caused, e.g., by mergings at onramps or offramps,

$$\frac{dv_\alpha}{dt} = a_{\text{mic}} [s_\alpha(t), v_\alpha(t), v_{\alpha-1}(t)] + \xi_\alpha(t) v_\alpha \delta[x_\alpha(t)]. \quad (37)$$

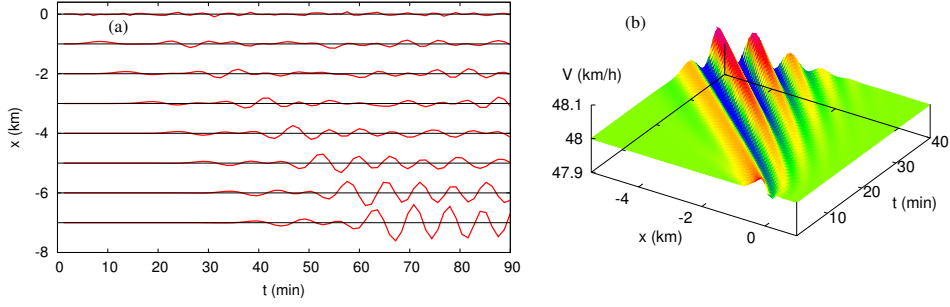


Fig. 11. (a) A realization of the solution (38) to the stochastic equation (37) for the IDM with $a = 1.10 \text{ m/s}^2$ and the other parameters as in Fig. 6. (b) The corresponding Green's function.

Here, $\xi_\alpha(t)$ are realizations of independent standard-normally distributed stochastic variables. The distribution $\delta[x_\alpha(t)]$ models an instantaneous change of the speed by ξ_α when this vehicle crosses the location $x = 0$.

Since we are interested in the collective properties of oscillations, it is sufficient to investigate the statistical properties of an approximate macroscopic speed field $U(x, t)$ representing the local deviations from the equilibrium speed v_e . In the linear regime, the realizations of this velocity field are given by

$$U(x, t) \propto \int_0^\infty d\tau G(x, \tau) \xi(t - \tau), \quad \langle \xi(t) \rangle = 0, \quad \langle \xi(t) \xi(t') \rangle = \delta(t - t'), \quad (38)$$

where $\langle \cdot \rangle$ denotes averaging over the realizations, and the Green's function $G(x, t)$ is given by Eq. (30) with (34). The integral extends over time, but not over space, representing perturbations that are sustained in time but localized at $x = 0$. Typically, such perturbations occur near bottlenecks. As an example, they may represent disturbances caused by lane changes of entering or leaving vehicles near onramp or offramp bottlenecks.¹

The stochastic properties of the solution (38) (a realization is shown in Fig. 11 (a)) can be described by the structure function

$$S(x, x', t, t') = \langle U(x, t) U(x', t') \rangle. \quad (39)$$

Inserting Eqs. (38) and (34) and dropping some algebraic prefactors leads, after a lengthy calculation, to an approximate analytic expression for the speed

¹ Simulations indicate that extended spatiotemporal noise will not lead to significant changes.

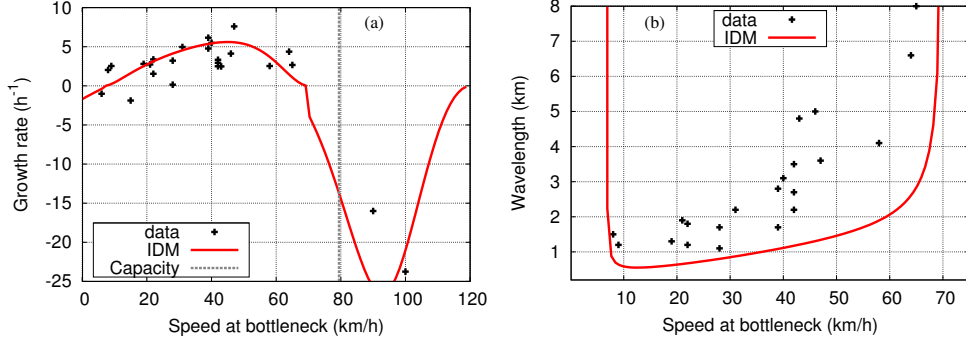


Fig. 12. (a) Effective linear growth rate and (b) wavelength of speed oscillations as a function of the speed near the bottleneck characterizing the severity of the bottleneck. The analytical IDM expression $-c\tilde{\sigma}$ for sustained perturbations, Eq. (41), is compared with the data of Fig. 3. The IDM parameters a and δ have been changed from their standard values to $a = 1.23 \text{ m/s}^2$ and $\delta = 7$, respectively.

variance:

$$\sigma_V^2(x, t) \equiv S(x, x, t, t) \sim \frac{1}{\beta} \exp \left[\frac{2xv_g}{D_2} (1 - \beta) \right] \quad \text{with} \quad \beta = \sqrt{1 - \frac{2D_2\sigma_0}{v_g^2}}. \quad (40)$$

This expression is valid for sufficiently large t under the condition $x \ll 2D_2/(v_g\beta)$, i.e., sufficiently far away from the threshold to absolute instability and from the source of the perturbations. (The right-hand side of the inequality evaluates to -850 m for the parameters of Fig. 11.)

Figure 11 displaying a realization of the solution (38) to the stochastic equation (37) for the IDM illustrates this result. In spite of linear instability and sustained noise, the oscillations remain finite if the instability is convective ($0 < \beta < 1$). Furthermore, a stationary situation is reached in the stochastic sense after some time (approximately 70 min in Fig. 11). Consequently, the variance (40) depends only on x . Finally, the oscillations grow in the upstream direction (notice that $v_g < 0$), in agreement with most of the observations. From Eq. (40), one can extract a quantity that is observable from the realizations $U(x, t)$, Fig. 11(a), or from traffic data, Fig. 3(a): The distance scale \tilde{L} in which the amplitude grows by a factor of e is given by

$$\tilde{L} = -\frac{1}{\tilde{\sigma}} = -\frac{D_2}{v_g(1 - \beta)}. \quad (41)$$

Using the observable phase velocity $c = \omega_0/k_0$, this can be translated into an effective growth rate,

$$\sigma_{\text{eff}} = -\frac{c}{\tilde{L}} \quad (42)$$

For the IDM parameters and the steady-state speed of Fig. 11, one obtains a physical wavelength $L = 2\pi k_0/\rho_e = 1.5 \text{ km}$, a distance scale $\tilde{L} = 3.5 \text{ km}$, a maximum growth rate $\sigma_0 = 3.0 \text{ h}^{-1}$, and the effective growth rate $\sigma_{\text{eff}} =$

4.4 h^{-1} . Notice that the effective growth rate inferred from the growth scale is higher than the maximum linear growth rate σ_0 . First investigations of the statistic properties of the realizations (Fig. 11(a)) indicate that this surprising result (and, more generally, Eq. (40)) are quantitatively correct, in spite of the numerous approximations made during its derivation.

In the last paragraphs, we have shown that the analytic calculations and the simulations agree quantitatively. However, a model validation implies testing the model predictions against the data, for a wide range of bottleneck strengths. Specifically, for a given microscopic or macroscopic model and a given parameter set, the theory gives the predictions

- $c = v_e + \omega_0/(k_0\rho_e)$ and $c_{\text{mac}} = \omega_0/k_0$ for the phase velocity of microscopic and macroscopic models, respectively,
- $L = 2\pi/(k_0\rho_e)$ and $L_{\text{mac}} = 2\pi/k_0$ for the wavelength of microscopic and macroscopic models, and
- Equation (42) for the effective growth rate

as a function of the speed v_e of the underlying stationary state. Since this is our indicator for the bottleneck strength, each model and each parameter combination gives three quantities (the propagation velocity, the wavelength, and the growthrate) as a function of the bottleneck strength which then can be compared with the data.

Figure 12 shows such a comparison for the IDM. By varying only two parameters with respect to the previously used settings, we obtain agreeing values (within statistical uncertainty) for the growth rate (42) and the propagation velocity (not shown) over the complete range of bottleneck strengths. The increase of the wavelength with decreasing bottleneck strength (i.e., increasing speed at the bottleneck) is predicted on a qualitative level, but the predicted wavelength is consistently too small by a factor of about two. First simulations indicate that this deficiency is resolved by extending the models to include more specific human properties such as reaction time or looking several vehicles ahead (Treiber et al., 2006).

5 Conclusion

In order to gain qualitative and quantitative insight into the nature of instabilities of traffic flow, we have analyzed a large database of congested traffic on a German freeway and give quantitative detail of the key factors of congested traffic instabilities. Notably, we found positive perturbation growth rates for most instances of congestions and no empirical evidence against the existence of linear instabilities. Investigations of traffic data around the world (Zielke

et al., 2008) come to similar conclusions.

To understand the observations, one must relate them to mathematical models and theoretical concepts, preferably analytically and quantitatively. To this end, we define all relevant theoretical stability concepts and give analytic criteria for convective and absolute string instability for the wide class of car-following and macroscopic models that are formulated in continuous time and possess a fundamental diagram. Although some approximations have been used in deriving the analytical threshold between convective and absolute instability, it has been verified by simulations within observation uncertainty (about 1 % of the width of the convectively unstable region).

We found new insight into the nature of traffic flow instabilities in open systems on two levels. On the qualitative level, we could relate the set of simulated spatiotemporal patterns (“dynamic phases”) to three “stability classes” (and two subclasses) which are derived from the stability diagram, i.e., from the set of density regions where a certain stability type applies. The original “phase diagram” of Ref. (Helbing et al., 1999) relates to stability class 1b, but other stability classes lead to different phase diagrams. Moreover, we have found that the observed patterns of congested traffic are consistent with the classes 2a and 2b. The stability class depends not only on the model but also on the parameters. In particular, the Intelligent Driver Model (IDM) can assume any stability class which we have represented by a “class diagram” in parameter space. Scanning the complete parameter space reveals that, for realistic values, the IDM assumes class 2a or 2b, in agreement with the observations.

In order to get further insight into the nature of convectively unstable flow, we have analytically related the properties of extended congested patterns with convectively unstable flow in an open system with sustained local noise (caused, e.g., by mandatory lane changes at lane closings, onramps, or off-ramps). The resulting structure function corresponds to stationary traffic flow near the bottleneck and, simultaneously, to growing oscillations further upstream if and only if traffic flow is convectively unstable. This means, the most common congested pattern is directly related to convective instability.

The predictions of the analytical calculation are quantitatively verified by simulations with the IDM and the GKT model in the presence of sustained localized noise. Furthermore, the results agree semi-quantitatively with the data: While the associated growth rates and propagation velocities agree with the data within statistical uncertainty for all values of the bottleneck strength, the wavelengths generally remain too short. This, however, is not a deficiency of the analytic theory but of the specific model to which this theory is applied. In turn, the theory provides a powerful tool for validating the dynamic aspects of models for traffic flow.

In conclusion, there is much evidence that the largely neglected concept of convective instability provides the final missing link in explaining the dynamics of congested traffic flow.

Acknowledgements

The authors would like to thank the *Hessisches Landesamt für Straßen- und Verkehrswesen* for providing the freeway data.

References

- Ahn, S., Cassidy, M. J., 2007. Freeway traffic oscillations and vehicle lane-change maneuvers. In: Allsop, R. E., Bell, M. G. H., Heydecker, B. G. (Eds.), *Transportation and Traffic Theory 2007*. Elsevier, Ch. 29, pp. 691–710.
- Aw, A., Rascle, M., 2000. Resurrection of “second order” models of traffic flow. *SIAM Journal on Applied Mathematics* 60 (3), 916–938.
- Bertini, R. L., Leal, M. T., 2005. Empirical study of traffic features at a freeway lane drop. *Journal of Transportation Engineering* 131 (6), 397–407.
- Coifman, B. A., Wang, Y., 2005. Average velocity of waves propagating through congested freeway traffic. In: Mahmassani, H. S. (Ed.), *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction*. Elsevier, pp. 165–179.
- Daganzo, C. F., 2002a. A behavioral theory of multi-lane traffic flow. part I: Long homogeneous freeway sections. *Transportation Research Part B* 36 (2), 131–158.
- Daganzo, C. F., 2002b. A behavioral theory of multi-lane traffic flow. part II: Merges and the onset of congestion. *Transportation Research B* 36 (2), 159–169.
- Helbing, D., 2001. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics* 73, 1067–1141.
- Helbing, D., Hennecke, A., Treiber, M., 1999. Phase diagram of traffic states in the presence of inhomogeneities. *Physical Review Letters* 82, 4360–4363.
- Helbing, D., Treiber, M., Kesting, A., Schönhof, M., 2009. Theoretical vs. empirical classification and prediction of congested traffic states. *The European Physical Journal B* 69, 583–598.
- Hoogendoorn, S., Bovy, P., 2001. State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 215 (4), 283–303.
- Huerre, P., Monkewitz, P., 1990. Local and global instabilities in spatially developing flows. *Annual Review of Fluid Mechanics* 22 (1), 473–537.
- Kerner, B., 2002. Empirical macroscopic features of spatio-temporal traffic patterns at highway bottlenecks. *Physical Review E* 65, 046138.

- Kerner, B., Rehborn, H., 1996. Experimental properties of complexity in traffic flow. *Physical Review E* 53, R4275–R4278.
- Kesting, A., Treiber, M., 2008. How reaction time, update time and adaptation time influence the stability of traffic flow. *Computer-Aided Civil and Infrastructure Engineering* 23, 125–137.
- Li, X., Peng, F., Ouyang, Y., 2010. Measurement and estimation of traffic oscillation properties. *Transportation Research Part B* 44 (1), 1–14.
- Lighthill, M., Whitham, G., 1955. On kinematic waves: II. A theory of traffic on long crowded roads. *Proc. Roy. Soc. of London A* 229, 317–345.
- Mauch, M., Cassidy, M. J., 2002. Freeway traffic oscillations: observations and predictions. In: Taylor, M. A. (Ed.), *Transportation and traffic theory in the 21st century: Proceedings of the Symposium of Traffic and Transportation Theory*. Elsevier, pp. 653–672.
- Mika, H. S., Kreer, J. B., Yuan, L. S., 1969. Dual mode behavior of freeway traffic. *Highway Research Record* 279, 1–12.
- Orosz, G., Stépán, G., 2006. Subcritical Hopf bifurcations in a car-following model with reaction-time delay. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 462 (2073), 2643–2670.
- Orosz, G., Wilson, R. E., Stépán, G., 2010. Traffic jams: dynamics and control. *Philosophical Transactions of the Royal Society A* 368, 4455–4479.
- Payne, H., 1979. A macroscopic simulation model of freeway traffic. *Transportation Research Record* 772, 68–77.
- Pipes, L., 1953. An operational analysis of traffic dynamics. *Journal of Applied Physics* 24, 274.
- Reuschel, A., 1950. Fahrzeugbewegungen in der Kolonne. *Österreichisches Ingenieur-Archiv* 4, No. 3/4, 193–215.
- Richards, P., 1956. Shock waves on the highway. *Operations Research* 4, 42–51.
- Schönhof, M., Helbing, D., 2007. Empirical features of congested traffic states and their implications for traffic modeling. *Transportation Science* 41, 1–32.
- Schönhof, M., Helbing, D., 2009. Criticism of three-phase traffic theory. *Transportation Research Part B* 43 (7), 784–797.
- Sugiyama, Y., Fukui, M., Kikuchi, M., Hasebe, K., Nakayama, A., Nishinari, K., Tadaki, S., Yukawa, S., 2008. Traffic jams without bottlenecks – experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics* 10, 033001.
- Treiber, M., Hennecke, A., Helbing, D., 1999. Derivation, properties, and simulation of a gas-kinetic-based, non-local traffic model. *Physical Review E* 59, 239–253.
- Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E* 62, 1805–1824.
- Treiber, M., Kesting, A., 2010a. Validation of traffic flow models with respect to the spatiotemporal evolution of congested traffic patterns submitted (Preprint arXiv:1008.1639).
- Treiber, M., Kesting, A., 2010b. Verkehrsdynamik und -simulation – Daten,

- Modelle und Anwendungen der Verkehrsflussdynamik. Springer, (In German) ISBN: 978-3-642-05227-9. URL www.verkehrsdynamik.de
- Treiber, M., Kesting, A., Helbing, D., 2006. Delays, inaccuracies and anticipation in microscopic traffic models. *Physica A* 360, 71–88.
- Treiber, M., Kesting, A., Helbing, D., 2010a. Three-phase traffic theory and two-phase models with a fundamental diagram in the light of empirical stylized facts. *Transportation Research Part B* 44 (8-9), 983–1000.
- Treiber, M., Kesting, A., Wilson, R. E., 2010b. Reconstructing the traffic state by fusion of heterogenous data. *Computer-Aided Civil and Infrastructure Engineering* preprint arxiv.org/abs/0909.4467, in press.
- Treiterer, J., Myers, J., 1974. The hysteresis phenomenon in traffic flow. In: Buckley, D. (Ed.), *Proc. 6th Int. Symp. on Transportation and Traffic Theory*. Elsevier, New York, p. 13.
- Wagner, P., Nagel, K., 2008. Comparing traffic flow models with different number of ‘phases’. *The European Physical Journal B* 63 (3), 315–320.
- Ward, J. A., Wilson, R. E., 2011. Criteria for convective versus absolute string instability in car-following models. *Proceedings of the Royal Society A*, in press.
- Wilson, R., 2008. Mechanisms for spatio-temporal pattern formation in highway traffic models. *Philosophical Transactions of the Royal Society A* 366 (1872), 2017–2032.
- Wilson, R. E., Ward, J. A., 2011. Car-following models: fifty years of linear stability analysis - a mathematical perspective. *Transportation Planning and Technology* 34, 3–18.
- Windover, J., Cassidy, M., 2001. Some observed details of freeway traffic evolution. *Transportation Research Part A* 35 (10), 881–894.
- Zielke, B., Bertini, R., Treiber, M., 2008. Empirical Measurement of Freeway Oscillation Characteristics: An International Comparison. *Transportation Research Record* 2088, 57–67.