2 Empirical facts and modeling approaches

Like in every other field of physics, theoretical models for traffic flow have to be compared with empirical findings. The aspired accordance strongly depends on the goal of the particular application. E.g., in order to use a certain model for traffic forecasts in realistic networks, the agreement has to be as precise as possible. On the other hand it is also useful to use oversimplified model approaches in order to concentrate on particular aspects of traffic flow phenomena [103]. But in any case it is important to know which aspects of real traffic are described by a certain model. A fundamental difficulty for a systematic comparison is the fact that one cannot perform laboratory experiments. In contrast empirical data has to be established from real and often quite complex traffic networks, that complicates the interpretation of traffic data. Nevertheless, in recent years several investigations were carried out, which led to a much more profound understanding of the empirical facts.

2.1 Measurement techniques and observables

The analysis of traffic flow characteristics is based on empirical data which provides information about the driving state of a car. The complete information of all cars allows therefore the representation of the complete traffic state. However, since the empirical data recording relies on observations of real life traffic, it is not possible to have access to all the information needed. Nevertheless, there are several ways in order to collect traffic data that try to overcome these problems all having their advantages and disadvantages. In general, data can be collected by moving or by stationary devices.
In car-following experiments [84, 109, 110], a number of measurement vehicles (floating cars) is equipped with detectors that allow the recording of the single vehicle behavior. If the test-drivers imitate a mean driving behavior of the surrounding traffic, the calculation of some macroscopic quantities is also possible whereas the density cannot be determined. However, it is difficult to get representative data since the drivers may drive biased. Moreover, for financial and practical reasons, normally only a few measurement vehicles exist.
In contrast, stationary detectors are able to give information about all vehicles passing the measurement location, but no information of the vehicle's driving state off the measurement location is provided. Therefore, stationary detectors with an increased spatial extension exist, e.g., video cameras, which improve the information of the single vehicle data, but, nevertheless, these measurement devices also cover only short sections of the highway.
The most widely used observation technique (that was applied in almost any empirical investigation) is by means of inductive loops. Inductive loops usually count the number of passing vehicles and provide information about the time the vehicle arrives at the detector and the duration it takes to pass the measurement section. Additionally, two successive inductive loops are able to calculate the velocity and the length of the vehicles. The data provided by inductive loops are for the sake of storage reasons usually given as temporally

aggregated data. The typical averaging intervals vary from 30 seconds up to several minutes. In case of speed and distance measurements, the aggregation of traffic data is unfavorable, because it hides important details of their distribution. Contrary, averages in time are the only way to quantify the (local) density. Nevertheless, some inductive loops offer the possibility of single-vehicle data. For the later discussion it should be stressed that the measurements provided by counting loops are local, that has to be taken into account when comparing them with model results.
The flow $J(t)$ gives the number $\mathcal{N}(t)$ of vehicles that pass the detector in a given time interval $[t, t+\mathcal{T}]$ at time t:

$$J(t) = \frac{1}{\mathcal{T}}\mathcal{N}(t). \tag{2.1}$$

The velocity of a vehicle is calculated by the time the vehicle needs to move from the first to the second detector and the known distance between the two detectors. However, this definition of the velocity is only valid under the assumption of a constant velocity between the detectors. In situations with strong fluctuations of the velocity on short distances, e.g., in congested traffic, this may not be true. The arithmetic average $v(t)$ of the velocities v_n of all vehicles passing the detector at time t_n in the time $\mathcal{T}$ where n denotes the n-th vehicle is given by

$$v(t) = \frac{1}{\mathcal{N}(t)} \sum_{t \le t_n < t+\mathcal{T}} v_n(t_n). \tag{2.2}$$

In contrast to the velocity averaged over a highway section, this locally averaged velocity is systematically overestimated since fast vehicles cross a highway section more frequently than slow vehicles. This can be compensated by the harmonic average of the velocity which is defined by

$$v(t) = \frac{\mathcal{N}(t)}{\sum_{t \le t_n < t+\mathcal{T}} \frac{1}{v_n(t_n)}}. \tag{2.3}$$

Using this definition, the velocity becomes smaller but more sensible to errors in the measurement of small velocities. Thus, most of the detectors calculate the arithmetic mean of the velocity.
If the detector provides both, the velocity of a vehicle and the time Δt_n the vehicle n occupies the measurement device, it is possible to determine the length l_n of a car. This allows (under certain restrictions) the classification of vehicles into cars, trucks, etc. Note that the length only gives the electrical length that is normally smaller than the physical length of the vehicle. Since the inductive loop records the time the vehicle passes the measurement location, the temporal headway t_n^h of a car n is given by the time lag to its predecessor $n-1$ under the assumption of a constant v_{n-1}:

$$t_n^h = t_n - t_{n-1} - \frac{l_{n-1}}{v_{n-1}} \tag{2.4}$$

t_n denotes the time the n-th vehicle passes the detector, l_n and v_n its length and velocity. By means of the temporal headway it is possible to calculate the netto distance headway d_n, that is the distance from bumper-to-bumper:

$$d_n = v_n t_n^h - l_{n-1}. \tag{2.5}$$

The usage of the velocity v_n is somehow arbitrary since v_{n-1} can also be taken, which leads to the calculation of the headway in upstream direction rather than downstream. However, the difference between both results is negligible because on small distances the

surroundings upstream and downstream of the measurement location should be homogeneous. Again, the calculation of the distance headway is based on the assumption that the velocity of both vehicles does not change significantly, which is the fact for decreasing temporal headways.
Up to now, all quantities measured are local. In contrast, the density gives the number of vehicles occupying a given section of the highway which, in principle, does not allow the determination of the density by an inductive loop. Nevertheless, two approaches try to approximate the spatial density of the highway by locally measurable quantities.
The hydrodynamical relation between the flow J, the velocity v and the density ρ allows the calculation of the density via

$$\rho(t) = \frac{J(t)}{v(t)}. \tag{2.6}$$

However, since the detection is often an event-driven process, only moving vehicles will be measured. Thus, for large densities in congested traffic, the mean velocity is systematically overestimated, while the density therefore is systematically underestimated. E.g., a compact jam leads to a density of zero instead of one, since no vehicle moves.
Occupancy (*occ*) rather than density is used as an estimation of the macroscopic density because of the ease of measuring compared with the density. It is given by the relative time an inductive loop is covered by vehicles:

$$occ(t) = \frac{1}{\mathcal{T}} \sum_{t \le t_n < t+\mathcal{T}} \Delta t_n = \frac{1}{\mathcal{T}} \sum_{t \le t_n < t+\mathcal{T}} \frac{l_n}{v_n}. \tag{2.7}$$

Often, occupancy is given in percent. The occupancy can simply be related to the density, under the assumption that the occupancy is proportional to it. An occupancy of one means that vehicles are driving bumper-to-bumper, which leads to a density $\rho_{\max}$ of

$$\rho_{\max} = \frac{\mathcal{N}}{\sum_{t \le t_n < t+\mathcal{T}} l_n} = \frac{1}{\mathcal{L}} \tag{2.8}$$

with the mean length $\mathcal{L}$ of the vehicles and the number $\mathcal{N}$ of vehicles passing the detector in the time $\mathcal{T}$. Thus, for the density estimated by occupancy it holds [1]:

$$\rho(t) = occ(t) \cdot \rho_{\max} = occ(t) \cdot \frac{\mathcal{N}}{\sum_{t \le t_n < t+\mathcal{T}} l_n} = \frac{occ(t)}{\mathcal{L}}. \tag{2.9}$$

In contrast to the calculation of the density via the hydrodynamical relation, occupancy leads to a wide range of densities with the same flow which becomes visible at a triangular shape of the congested branch of the fundamental diagram. On one hand, very large values of occupancies are measured due to a jammed vehicle blocking the detector for a long time, which may overestimate the density since large gaps between standing vehicles are ignored. On the other hand, gaps in a compact jam can lead to small occupancies, and therefore small densities are obtained although a jam can be found in the vicinity of the detector. Thus, the definition of density by local rather than by global quantities results in large fluctuations and is not an explicit function of the number of vehicles. Nevertheless, most of the local measurement devices use occupancy since it can easily and directly be determined.
Global quantities that are provided by measurement sections covering the whole highway can be calculated in simulations of traffic flow. As an advantage, these observables give

[1] The calculation of the spatial density by means of local measurements is only valid under the assumption of a constant velocity of the vehicles during the averaging process.

the complete and correct information about the extended traffic states, while local measurements are only approximations of these states. Now, the density ρ_{global} can directly be obtained by counting the number N of vehicles on a highway section of length L via

$$\rho_{\text{global}} = \frac{N}{L}. \tag{2.10}$$

The average velocity v_{global} is then defined as

$$v_{\text{global}} = \frac{1}{N}\sum_{n=1}^{N} v_n. \tag{2.11}$$

Again, the hydrodynamical relation allows the calculation of the flow

$$J_{\text{global}} = \rho_{\text{global}} v_{\text{global}} = \frac{N}{L}\cdot\frac{1}{N}\sum_{n=1}^{N} v_n = \frac{1}{L}\sum_{n=1}^{N} v_n. \tag{2.12}$$

In addition, microsimulations provide information about the driving state of all vehicles like the velocity, the headways, and the length.

2.2 Spatio-temporal structures of traffic flow

Since the first aerial photography of a traffic jam by Treiterer [138], much progress has been made in the analysis of traffic flow characteristics. From a theoretical point of view, traffic data allows the identification of traffic phases [67] and the characterization of the transitions between these phases. These results help to calibrate and validate traffic flow models [80, 81] and to improve existing model approaches. From a practical point of view, the understanding of traffic flow lays the foundations for the improvement of road capacities and traffic forecasts [65, 66, 77].
The development of the measurement techniques and the improved equipment of roads with detectors in the last few years provides not only an increasing amount of data and a more detailed analysis of traffic flow, but also increases the statistical relevance of these analyzes. In particular, recent empirical studies [73] were based on a series of successive inductive loops in order to facilitate the analysis of spatial and temporal structures. In addition, inductive loops providing single-vehicle data [110] give insight into the microscopic driving behavior of single vehicles and help to improve microscopic traffic flow models and to understand the underlying jamming mechanisms.

2.2.1 Traffic phases

The empirical observation of highway traffic revealed the existence of very complex spatio-temporal structures [38, 67, 69, 70]. So it is now widely believed that three traffic states exist, i.e., (i) free flow, (ii) synchronized traffic and (iii) wide moving jams [73]. Often, the synchronized state and wide moving jams are summarized by the term "congested traffic". Free flow traffic is characterized by a high average velocity of the vehicles, e.g., close to the speed limit, while congested traffic is simply defined to be the opposite of free flow. Wide moving traffic jams are upstream moving structures consisting of two fronts separated by a region of negligible velocity and flow. They can be characterized by the upstream propagation velocity, their density and the outflow from the jam if free flow is formed in the outflow region. These parameters of the jam are quite robust and are only determined

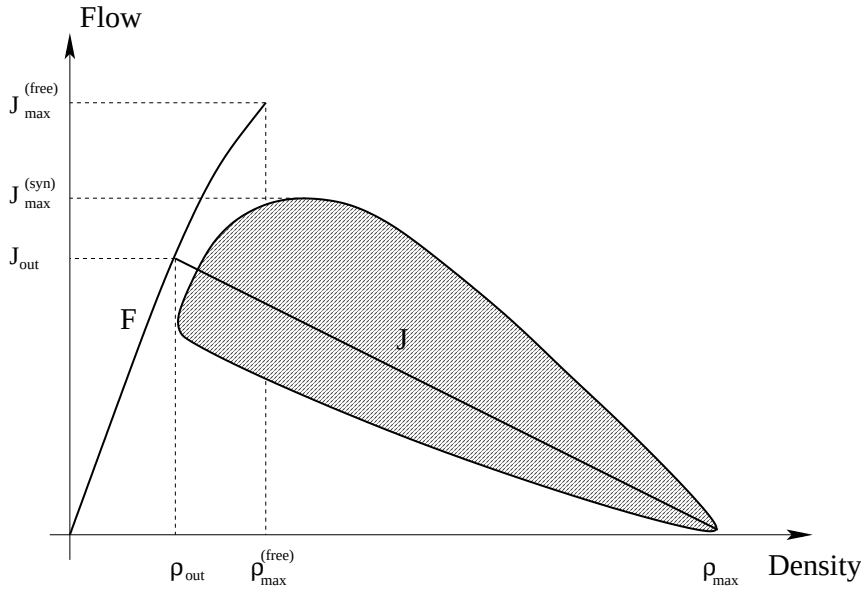

Figure 2.1: Schematic plot of the fundamental diagram for a multi-lane road. Free flow states are given by the line F while synchronized states are illustrated by the hatched region. The line J is determined by the characteristics of wide moving jams. Below the line J no wide jams can exist for a long time while above the states are meta-stable in respect to jam formation.

by the road conditions: Empirically it has been observed [63] that two jams can move in parallel through the system, which implies that their characteristic parameters agree. In contrast, in the vicinity of hindrances, or if synchronized traffic is formed in the outflow, the outflow from a jam and the density and the velocity in the outflow region are no longer characteristic properties [68] and perform complex slides along the line J (Fig. 2.1).
For a long time it was believed that these states are the only stable traffic states. This commonly accepted picture was enhanced by establishing a second stable congested state, i.e., synchronized traffic. Synchronized traffic [38, 67, 69, 70, 73], which is typically observed at on- and offramps, is characterized by a large variance in flow and density measurements and a velocity that is significantly lower than in free flow traffic. Compared to free flow traffic, the variance of the velocity in the synchronized state is considerably lower, which can be explained by a bunching of vehicles [78], and waves (e.g., narrow jams) of the density and the flow can propagate in both directions, upstream and downstream.
Three types of synchronized flow can be distinguished. In type 1, i.e., the stationary and homogeneous state, the flow as well as the velocity are nearly stationary during a relatively long time interval (2–5 min). In type 2, called homogeneous-in-speed states, the flow and the density strongly fluctuate whereas the velocity is nearly stationary. Waves of the flow and the density have a positive velocity. In multi-lane traffic, different waves may propagate in different lanes. Non-stationary and non-homogeneous states are given in type 3, that is the most commonly observed state. Waves of the flow and the density with positive and negative velocity can be found. Note that often a spatio-temporal sequence of these types of states is observed [73].
The origin of the notation "synchronized traffic" is based on the fact that the time-series of measurements on different lanes are highly correlated [73, 109, 110]. Besides, empirical observations have shown that this is the case even for wide moving jams [73]. But more

important is the apparent absence of a functional flow-density form, i.e., the measurements of the flow, density and velocity of the traffic are distributed over a wide area [64] (Fig. 2.1). This observation has been confirmed quantitatively [110] by means of vanishing cross-correlations between these quantities. In particular, while strong correlations of the flow and the density measurements can be found in free flow traffic and wide jams, synchronized traffic can be characterized by negligible values of the cross-correlation. Therefore, the applicability of the fundamental diagram, that illustrates the functional relation of the flow, the density and the velocity, is questioned (cf. [68]).
While in free flow traffic the flow increases proportional to the density since the average velocity is nearly constant, synchronized traffic covers a wide area of possible states (Fig. 2.1). These states are divided by the line J that intersects the free flow curve at a point determined by the outflow from a wide jam giving the characteristic density ρ_{out} and flow J_{out}. The slope of the line J is obtained by the velocity of the downstream front of a wide jam. Below the line J free flow is stable, that is no wide jam can exist for a long time, while at larger values free flow is in a meta-stable state with respect to the jam formation. Typical values for the outflow from a wide jam are $J_{\text{max}}^{\text{free}}/J_{\text{out}} \approx 1.5$, and the velocity of the downstream front has a value of about 15 km/h [63, 68]. In contrast, the outflow from synchronized traffic can strongly fluctuate and exceed considerably the outflow from a jam. These fluctuations are related to changes in the upstream flows rather than to fluctuations of the discharge flow itself [72].
Synchronized states below the line J are stable in the sense that no moving jam can be excited or exist, while states on and above the line J are meta-stable with respect to the formation of wide jams [70]. Moreover, states of free flow and synchronized traffic overlap in a certain density region (Fig. 2.1). Since the velocity in free flow is considerably higher than the maximum velocity in synchronized traffic, there is a gap between the maximal flow rate of both states. This gap decreases with decreasing maximum velocity in free flow, i.e., from the left to the right lane [73], and vanishes completely on single-lane roads.
Synchronized traffic and wide moving jams differ also in their behavior at bottlenecks. If synchronized flow is generated at a bottleneck, it's downstream front is pinned at the bottleneck. In contrast, the downstream front of wide moving jams propagates with constant velocity in upstream direction. This velocity (about 15 km/h) and the corresponding flow rate are only determined by the density inside a wide jam and the delay-time between two vehicles leaving the jam. Therefore the velocity does not depend on the traffic state wide jams cross and is even unchanged if they pass a bottleneck. This property is an objective criterion for the identification of wide jams and is responsible for the coexistence of wide moving jams and synchronized traffic [73].

2.2.2 Microscopic properties

The empirical results of the last section are based on time-averaged quantities. However, the analysis of single-vehicle data allows a more detailed characterization of the three traffic states, free flow, synchronized traffic, and wide moving jams and provides information about the microscopic driving behavior of the vehicles. This is necessary for the validation and calibration of micro-simulations in order to reproduce the traffic states. In particular, the time-headway distribution and the speed-distance relation can directly be compared with simulation results.
For the analysis of single-vehicle data it is necessary to distinguish not only between the traffic states, but also between the density regimes. In the free flow regime, the time-headway distribution shows a maximum at about 0.9 s. However, considerably smaller

headways can be measured, which is a result of small platoons of vehicles moving with large speed but small distance-headway. In contrast, in congested traffic the maximum of the time-headway distribution is shifted to larger values, while the weight of small time-headways is decreased due to safety reasons [110, 133]. However, the shape of the distribution within a traffic state does not change significantly. Thus, the driving state of a vehicle seems to depend not only on the density, but also on the traffic state.
This is confirmed by the dependence of the velocity of individual vehicles on the distance headway. From a theoretical point of view, this function is also of great importance in the so-called optimal velocity models [5]. In the free flow regime it is obvious that the asymptotic velocity is reached already for small distances. In contrast, in the congested regime the asymptotic velocity is reduced considerably. Therefore, vehicles do not drive as fast as the distance-headway allows [110, 133]. Moreover, in both traffic states the distance to the vehicle in front is minimized if both cars move with the same velocity [110].
In summary, the driving behavior of a vehicle is determined not only by the distance and the velocity difference to the predecessor, but also by the car's velocity and, more important, by the traffic state: more dense traffic requires a more careful driving strategy. Nevertheless, in free flow small platoons of vehicles driving bumper-to-bumper with the same speed exist, while the length of these platoons is considerably larger in synchronized traffic, which is confirmed by a slow decrease of the velocity correlation of the vehicles [110]. However, in contrast to free flow, in synchronized traffic the distance between the platoon vehicles strongly fluctuates.
A more detailed analysis of single-vehicle data is presented in chapter 3.

2.3 Phase transitions in traffic flow

The transitions between the three traffic states are local first-order phase transitions that are accompanied by a sharp decrease of the traffic observables [64, 67, 71, 78]. The notation *first order transition* is borrowed from the terminology of thermodynamic phase transitions. The analogy is simply in the discontinuous change of the velocity signal which is similar to the first derivative of a thermodynamic potential in a first order transition.
There are three possible types of transitions: (i) from free flow to synchronized traffic, (ii) from synchronized traffic to wide jams, and (iii) from free flow to wide jams. The inverse transitions show hysteresis, since free flow states and synchronized states overlap in the fundamental diagram (Fig. 2.1). This means that, e.g., a transition from free flow to synchronized traffic occurs at considerably larger densities than the transition back from synchronized traffic to free flow.
Empirical observations suggest [69, 70, 73] that jams rarely form in free flow, and the transition from free flow to jams occurs by, first, the transition from free flow to synchronized traffic and then from synchronized traffic to jams. In general the sequence of this transition happens at spatially different locations with a time delay.
The transition from free flow to synchronized traffic is mainly observed at bottlenecks, but it can also be found far away from ramps [72, 92]. It is accompanied by a sharp decrease of the average velocity, whereas the flow remains nearly as high as in free flow. In contrast, at the transition from synchronized traffic to wide jams, a sharp decrease of the velocity *and* the flow occurs. A perturbation of the flow can cause a random transition (e.g., a random fluctuation of the flow in the vicinity of the bottleneck) or an induced transition (a wide jam or an upstream moving synchronized region) to synchronized traffic. As a result, the downstream front of the synchronized region is fixed at the bottleneck. At the upstream boundary either wide jams are formed in the so-called pinch region [73],

or synchronized flow transforms to free flow. In the pinch region a self-compression of synchronized flow occurs, and small narrow jams can form. In contrast to wide jams, narrow jams (waves in synchronized flow) consist of only two fronts where the vehicle speed and the density sharply changes, whereas the width of the jam fronts of wide jams is negligible compared to the bulk region between the fronts. In addition to that, there are no characteristic parameters: the outflow depends on the inflow, and the upstream velocity is larger compared to wide jams. Therefore they cannot propagate undisturbed through all traffic states and without impact on the states they pass. However, inside the pinch region the narrow jams can merge and thus form wide jams which suppress the formation of new narrow jams in between [42, 84]. Thus, the formation of wide jams in initially free flow traffic is a result of the pinch effect in synchronized flow. Since narrow jams emerge in the pinch region, the flow rate out of this region is considerably low. As a result, synchronized traffic can exist a very long time and can even self-maintain if the flow into the pinch region strongly fluctuates. Moreover, local perturbations in synchronized flow can cause continuous spatial-temporal transitions between the different types of synchronized traffic [73].

2.4 Multi-lane characteristics

The analysis of empirical highway data provides much information about the traffic phases and their transitions. Moreover, single-vehicle data gives insight into the driving strategies of the cars. It turned out that the synchronization of the vehicles' speed on different lanes in synchronized traffic and wide jams is a consequence of lane changes. If the density increases on the right lane, more and more vehicles change to the neighboring lane. As a result, the densities and thus the velocities adapt to each other, leading to a synchronization of the movement on both lanes. Thus, lane changes seem to play an important role in the context of traffic phases and their transitions. Nevertheless, only a few empirical studies [12, 17, 39, 130] were carried out to observe especially the multi-lane characteristics of highways and the lane changing behavior of vehicles.

The lane usage curve represents the density or the flow distribution on the lanes. This distribution has a strong dependence on legal restrictions. Normally, the flow is distributed nearly equally on all lanes: in three lane traffic, the right-lane shows a smaller flow than the middle and the left lane [12, 17, 39]. This is a result of the driver's intention to give way to the vehicles squeezing on the highway from an onramp. In contrast, studies on German highways [130] have revealed the so-called lane-usage inversion. At small densities nearly all vehicles are on the right lane. With increasing density, this density difference decreases, until the majority of the vehicles are on the left lane, and thus the flow on the left lane will be larger than on the right lane. This can be explained by the right lane preference rule and the right lane overtake ban that are applied in Germany.

More important than the flow and the density distribution on the lanes is the lane change behavior of the vehicles. Unfortunately, lane changes are very difficult to observe with the common measurement techniques so that only a few empirical results exist [12, 17, 130]. Obviously, the number of lane changes increases with increasing density since the undisturbed movement of a vehicle becomes more and more difficult. At large densities, the number of lane changes is very small due to insufficient gaps, but lane changes are still possible. At intermediate densities, the lane change curve exhibits a maximum that can be found in the vicinity of the maximum flow (Fig. 2.2).

As a result of the large number of lane changes, the velocity on different lanes in the congested traffic state becomes synchronized.

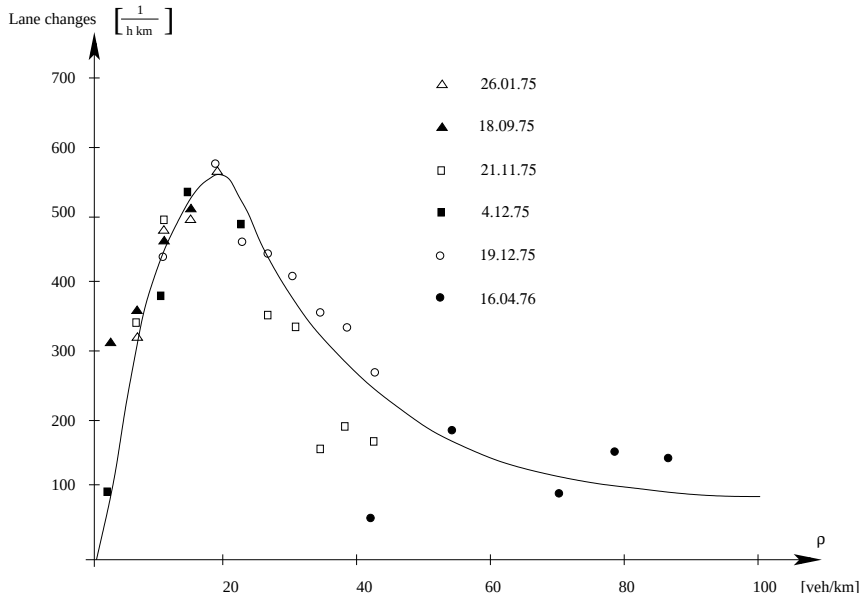

Figure 2.2: Number of lane changes versus density. The symbols represent the different observation days (from [130]).

2.5 Macroscopic theories of vehicular traffic

In the last decades several approaches have been proposed for modeling vehicular traffic (for an overview see, e.g., [18, 41]). These models can be classified by the way the vehicle movement is considered. Microscopic models try to reproduce the dynamics of each single vehicle. Once the trajectories of all vehicles are known, it is easy to obtain the aggregated quantities by simply averaging the single-vehicle data. In contrast, for the comparison with minute averaged data provided by inductive loops, the movement of each vehicle is not important so that averaged quantities are sufficient. Thus, macroscopic models are restricted to the collective vehicle dynamics and are formulated in analogy to the hydrodynamical description of compressible viscous fluids, considering a coarse-grained picture where the microscopic degrees of freedom are ignored, and different vehicles are not distinguished.

These models characterize traffic by means of macroscopic quantities like the flow J and the density ρ. The observables are connected by the continuity-equation that gives simply the conservation of vehicles. In the absence of sources and sinks, the following local condition has to be fulfilled:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = 0. \tag{2.13}$$

Since it is not possible to determine the flow and the density explicitly by the solution of the continuity-equation, one needs a second independent equation for $v(x,t)$ (note, that $v(x,t)$ and $J(x,t)$ are coupled via the hydrodynamical relation $J(x,t) = \rho(x,t)v(x,t)$) or the number of independent variables has to be reduced by one. However, the solution of the macroscopic models requires explicit assumptions for the stationary fundamental diagram. This can be done by using empirical flow-density relations or by calculating the fundamental diagram by means of microscopic simulations.

Lighthill and Whitham [95, 96] have assumed that the flow is primarily determined by the density, i.e., velocity adapts instantaneously to the surrounding density and thus the

continuity-equation can be treated as a function of the density only. The simplest functional relationship

$$J(x,t) = J(\rho(x,t)) \tag{2.14}$$

allows to solve the continuity-equation. Note that information about the fundamental diagram is still needed. For example, first measurements [37] suggested the following form that exhibits a maximum of flow:

$$J = \rho v(\rho) = \rho \cdot v_0 \left(1 - \frac{\rho}{\rho_{\max}}\right) \tag{2.15}$$

with the velocity v_0 at small densities and the maximal density $\rho_{\max}$. The solutions are so-called kinematic waves (since their properties are determined by the continuity-equation only) and allow the reproduction of shock fronts at bottlenecks, which corresponds to the formation of jams waves in real traffic. Nevertheless, the kinematic waves are stable, and traffic flow is always in equilibrium so that no spontaneous jam formation occurs.
The consideration of an additional velocity equation in order to include inertia allows a more realistic description of vehicular traffic. This equation assumes that the adaption of the velocity may be delayed with some relaxation time τ. Often, the velocity equation can be written in the following Navier-Stokes like form:

$$\frac{\partial v(x,t)}{\partial t} + v\frac{\partial v(x,t)}{\partial x} = -\frac{1}{\rho}\frac{\partial P(x,t)}{\partial x} + \nu(\rho)\frac{\partial^2 v(x,t)}{\partial x^2} + \frac{v_{\mathrm{e}}(\rho) - v(x,t)}{\tau(\rho)}. \tag{2.16}$$

The term $-\frac{1}{\rho}\frac{\partial P(x,t)}{\partial x}$ is usually denoted as traffic pressure, while $P(x,t)$ is often given by the velocity variance. With the viscosity ν, the term $\nu(\rho)\frac{\partial^2 v(x,t)}{\partial x^2}$ leads to the reduction of spatial inhomogeneities of the velocity and increases the numerical robustness. In particular, the viscosity term reduces the occurrence of discontinuities at shock fronts, that would lead to numerical problems. The relaxation of the velocity to the equilibrium velocity $v_{\mathrm{e}}(\rho)$ is described by $\frac{v_{\mathrm{e}}(\rho)-v(x,t)}{\tau(\rho)}$, with the delay time τ.
Several model approaches [52, 74–76, 87, 88, 143] have been suggested that use a velocity equation of this general form. In contrast to the Lighthill and Whitham theory, these models show the existence of unstable and meta-stable traffic states. In unstable states, density clusters that can be related to traffic jams in reality emerge and move with constant velocity upstream. In meta-stable traffic states, traffic is stable with respect to infinitesimal perturbations. As a result of these perturbations, small jams form which dissolve leading to a density pattern similar to stop-and-go traffic.
The consideration of on- and offramps in the hydrodynamical models adds a source and a sink term to the right hand side of the continuity-equation (2.13). By doing so, periodic oscillations of the flow and the density that are localized in the vicinity of an onramp have been found, which can be related to characteristics of synchronized traffic [90, 91]. In [11], however, the interpretation of this traffic state to be synchronized traffic is questioned.
In order to extend the macroscopic models to multi-lane traffic, additional lane-changing terms in the density and velocity equations have to be introduced. Several approaches have been proposed that give quite realistic results [43, 54, 99].
Nevertheless, in the context of the hydrodynamical description of vehicular traffic, the origin of the phenomena observed in traffic flow cannot be related to the human driving behavior. In particular, the enormous number of free parameters are not deducible from first principles. Moreover, there are no empirical analogies for, e.g., the viscosity term. Another important fact is that macroscopic models that use a diffusion term in equation (2.13) or a pressure term in equation (2.16) for smoothing shocks lead under certain

boundary or initial conditions (e.g., at the upstream end of a jam) to vehicles moving backwards [23]. But one of the main problems of a macroscopic view of vehicular traffic is the lack of information on single-vehicle data. Especially this type of data is needed in modern applications of traffic forecasts and dynamic route guidance systems.
The derivation of macroscopic models from "microscopic" equations is the goal of the gas-kinetic approach. In the kinetic theory, traffic is treated as a gas of interacting vehicles, that is described by the distribution $P(x, v, t)$. $P(x, v, t)\ dx\ dv$ is the probability density and gives the number of vehicles at a time t with a position between x and dx and a velocity between v and dv. The time-evolution of the distribution $P(x, v, t)$ can be obtained by the Boltzmann-like equation [117]

$$\frac{\partial P}{\partial t} + v\frac{\partial P}{\partial x} = \left(\frac{\partial P}{\partial t}\right)_{\text{rel}} + \left(\frac{\partial P}{\partial t}\right)_{\text{col}}. \tag{2.17}$$

Here, $\left(\frac{\partial P}{\partial t}\right)_{\text{rel}}$ denotes the relaxation within time τ of P towards a desired velocity distribution $P_{\text{des}}(v)$ that drivers try to achieve in the absence of vehicle interactions. In [117]

$$\left(\frac{\partial P}{\partial t}\right)_{\text{rel}} = -\frac{P(x, v, t) - \rho(x, t)P_{\text{des}}(v)}{\tau} \tag{2.18}$$

with the density ρ and the relaxation time τ was suggested. The distribution P_{des} has to be calculated from empirical observations of the velocity distribution of vehicles with a large distance-headway. The interactions of the vehicles are modeled by

$$\begin{aligned}\left(\frac{\partial P}{\partial t}\right)_{\text{col}} &= \int_v^\infty dv'(1 - f_{\text{pass}})(v' - v)P(x, v', t)P(x, v, t) \\ &\quad - \int_0^v dv'(1 - f_{\text{pass}})(v - v')P(x, v, t)P(x, v', t).\end{aligned} \tag{2.19}$$

Fast vehicles with velocity v' meet slower ones with velocity v at a rate that is proportional to $P(x, v', t)P(x, v, t)$ and cannot overtake these vehicles with a probability $1 - f_{\text{pass}}$ thus increasing P. Analogously, P is decreased if vehicles with velocity v cannot overtake slower ones with v'.
Analyzing the gas-kinetic approach, a transition from free flow to congested traffic was found [118] that could be identified by the corresponding velocity distributions. Moreover, it is possible to derive macroscopic equations like the continuity equation and the velocity equation from the kinetic models under the assumptions of several approximations. However, the derived macroscopic equations are also based on first principles since even the gas-kinetic approach is of phenomenological nature.

2.6 Microscopic theories of vehicular traffic

In contrast to the macroscopic models which use a coarse-grained picture for the description of vehicular traffic, microscopic models focus on the reproduction of the movement of individual vehicles. One of the advantages of microscopic models is that they provide information on the driving state of each vehicle. Empirical observables like the flow, the density and the mean velocity are therefore obtained by averaging the single-vehicle data.

2.6.1 Car-following models

Car-following models [33, 49, 114, 120, 121] give the trajectories of the vehicles by solving an equation of motion. In analogy to Newtonian mechanics, the acceleration $\ddot{x}_n$ of a vehicle

n is considered as a response to a stimulus f. In general, the stimulus f is determined by the velocity v_n of the vehicle, the distance Δx_n and the velocity difference Δv_n to its predecessor:

$$\ddot{x}_n = S\ f(v_n, \Delta x_n, \Delta v_n). \tag{2.20}$$

S determines the sensitivity a vehicle reacts to the stimulus.
The car-following models differ in the choice of the stimulus-function. In the earliest car-following models, it was assumed [114, 120, 121] that drivers tend to have a safe distance to the preceding vehicle that increases with the velocity. However, the response to the stimulus should be retarded [16], so that the response of a driver at time t depends on the stimulus of the preceding vehicle $n+1$ at time $t-T$ where T is a typical reaction time that leads to the following equation of motion:

$$\ddot{x}(t+T) = S(\dot{x}_{n+1}(t) - \dot{x}_n(t)). \tag{2.21}$$

Within this approach, the vehicles tend to equilibrate their velocity so that in the stationary state all vehicles have the same velocity. A more realistic approach considers a generalized sensitivity $S = S_0 \cdot S_n(\Delta x_n, v_n)$. Thus, the general form of the car-following models reads [33]:

$$\ddot{x}_{n+1}(t+T) = S_0\, \frac{(v_{n+1}(t+T))^m}{(x_n(t) - x_{n+1}(t))^l}\ (\dot{x}_n(t) - \dot{x}_{n+1}(t))\,. \tag{2.22}$$

The parameters m and l are used in order to calibrate the model with empirical data. A good agreement with empirical data of American highways is given by $m = 0.8$ and $l = 2.8$ [53, 98].
The car-following models allow the calculation of the fundamental diagram by means of first principles of driving. Nevertheless, the dependencies of the stimulus and the sensitivity on m and l cannot be related to the human driving behavior. Therefore, on a microscopic level, car-following models give quite unsatisfactory descriptions of vehicular traffic. E.g., clustering cannot be reproduced since vehicles are driving with the same velocity in the stationary state.

2.6.2 Optimal-velocity models

In the car-following models, the driving state of a car is clearly determined by its predecessor. Therefore, the movement of a single vehicle cannot be described within this approach since a leading vehicle is missing.
In the so-called optimal velocity models [2–5], a vehicle does not adopt the velocity of its predecessor but of an optimal velocity V_{opt}, that in general depends on the distance. The equation of motion then reads:

$$\ddot{x}_n(t) = S_0\ [V_{\text{opt}}(\Delta x_n(t)) - v_n(t)]\,. \tag{2.23}$$

The optimal velocity function has to fulfill the conditions $V_{\text{opt}}(\Delta x) \to 0$ for $\Delta x \to 0$ and $V_{\text{opt}}(\Delta x) \to V_0$ for $\Delta x \to \infty$. One choice [4] for the optimal velocity function is

$$V_{\text{opt}}(\Delta x) = \tanh(\Delta x - \Delta x_s) + \tanh(\Delta x_s) \tag{2.24}$$

with the constant safety distance Δx_s. A homogeneous steady-state solution of equation (2.23) is

$$x_n = \frac{L}{N} \cdot n + V_{\text{opt}}\left(\frac{L}{N}\right) \cdot t \tag{2.25}$$

with the length L of the road and the number N of vehicles. Here, vehicles are driving with the constant headway $\frac{L}{N}$ and the velocity $V_{\text{opt}}\left(\frac{L}{N}\right)$. This solution becomes unstable if the condition

$$\left.\frac{\partial V_{\text{opt}}}{\partial \Delta x}\right|_{\Delta x=\frac{L}{N}} > \frac{S_0}{2} \tag{2.26}$$

is fulfilled [5, 131]. In detail, one can distinguish five different density regimes with respect to the stability of the homogeneous solution. At small and at large densities the homogeneous states are stable. At intermediate densities jammed states exist that are meta-stable in two density regimes, that is, the homogeneous state is stable with respect to infinitesimal perturbations, whereas in one density region only the jammed states are stable. The stability of traffic flow can be enhanced by introducing vehicles with an adaptive cruise control which reduces the distance-headway to the predecessor at a given velocity. As a result, vehicles moving in groups increase the flow, and jams are avoided [89].
A generalization of the optimal velocity models takes the movement of more than one predecessor into account [93]. These multi-vehicle interactions can be expressed in the form

$$\ddot{x}_n = \sum_{i=1}^{m} S_i \left(V_{\text{opt}} \left(\frac{x_{n+i} - x_n}{i} \right) - v_n \right) \tag{2.27}$$

with the sensitivity coefficients S_i. The increased anticipation of the predecessor's velocity stabilizes the traffic flow and leads to the formation of vehicle platoons.
Optimal velocity models allow the reproduction of the fundamental diagram and give quite realistic results compared to empirical data. However, due to the possibility to chose an arbitrary optimal velocity function, the dimension of the parameter space becomes infinite. Furthermore, recent empirical observations [110] suggest that the optimal velocity function strongly depends on the traffic state (cf. section 2.2.2). This complicates the analysis of phase transitions in these models, since a change of the optimal velocity function changes the traffic phase. In addition, since the acceleration is determined by the optimal velocity function, the model is not free of collisions. If a car approaches a vehicle at rest [48], it may have not enough time to slow down, which results in a crash because of the delayed adaption of the velocity corresponding to the distance. Even special initial conditions will produce accidents. This problem becomes more serious in the discretized version of the optimal velocity model [47] that will be discussed in section 4.3.2.

2.6.3 Intelligent driver model

In order to avoid collisions in the optimal velocity model, the velocity of a vehicle has to be dependent not only on the distance, but also on the velocity difference to the predecessor. Simulations with the so-called intelligent driver model [137] have shown that the consideration of velocity differences plays an important role in stabilizing traffic. The equation of motion reads:

$$\frac{dv_n}{dt} = \alpha \left[1 - \left(\frac{v_n}{v_0} \right)^{\delta} - \left(\frac{\Delta x_n^*(v_n, \Delta v_n)}{\Delta x_n} \right)^2 \right] \tag{2.28}$$

with the velocity v_n, the desired velocity v_0 and the distance Δx_n of vehicle n, and the velocity difference Δv_n to its predecessor. For the desired distance Δx_n^* it holds

$$\Delta x_n^* = s_0 + \max\left(v_n t^s + \frac{v_n \Delta v_n}{2\sqrt{a \cdot b}} \right) \tag{2.29}$$

with the safe time headway t^s , the maximum acceleration a, the desired deceleration b and the acceleration exponent δ. The desired distance is calculated by the jam distance

s_0 to a standing vehicle and an additional safety distance $v_n t^s$. The movement of a vehicle is determined by the interplay of the tendency to accelerate with $a[1-(v_n/v_0)^\delta]$ and the tendency to brake with the deceleration $-a(\Delta x^*/\Delta x_n)^2$. Thus, two mechanisms are responsible for the braking of a vehicle. On one hand, at high approach rates (that is, at large velocity differences to the predecessor) the deceleration necessary to avoid a collision is compared with the desired deceleration b. This comparison can be referred to as intelligent driving behavior. In emergency situations, i.e., b is less than the necessary deceleration, drivers overreact while otherwise the deceleration is smaller than b. On the other hand, the comparison of the distance headway Δx_n with the desired distance Δx_n^* allows the reduction of the velocity even if the velocity difference is negligible.
The approach of the intelligent driver model leads to quite realistic results compared with empirical data and shows in the presence of inhomogeneities [136] various traffic states similar to synchronized traffic. However, a recent study [11] of the optimal-velocity model with an explicit onramp questions these interpretations of the traffic states found and doubts the existence of these states in reality.

2.6.4 Nagel-Schreckenberg model

Cellular automata models have received much interest in the last few years. In contrast to continuously formulated models, in cellular automata models space and time and the state variable are discrete. Thus, these models are by design well suited for large scale computer simulations.
One of the first cellular automaton model for the simulation of traffic flow was introduced by Cremer [22]. However, the most prominent cellular automaton model describing vehicular traffic is the model of Nagel and Schreckenberg (hereafter cited as NaSch-model) [106], which is an extension of the Asymmetric Simple Exclusion Process [86]. The NaSch-model is a probabilistic cellular automaton. The road is divided into cells of length 7.5 m that can be either empty or occupied by just one vehicle n with a velocity $v_n = 0, 1, ..., v_{\max}$. Cars move from the left to the right on a lane with periodic boundary conditions. The configuration of the system at time $t+1$ can be obtained by applying the following rules to all vehicles simultaneously, that is by applying a parallel update:

1. Acceleration: $v_n\left(t+\frac{1}{3}\right) = \min\left(v_n(t)+1, v_{\max}(t)\right)$
2. Deceleration: $v_n\left(t+\frac{2}{3}\right) = \min\left(v_n\left(t+\frac{1}{3}\right), d_n\right)$
3. Noise: $v_n(t+1) = \max\left(v_n\left(t+\frac{2}{3}\right)-1, 0\right)$ with probability p_{dec}
4. Motion: $x_n(t+1) = x_n(t) + v_n(t+1)$.

v_n denotes the velocity and x_n the position of the n-th vehicle while d_n specifies the number of empty cells in front of car n. The maximum velocity $v_{\max}$ is the same for all vehicles. Typical parameters are $v_{\max} = 5$ and $p_{\text{dec}} = 0.5$. Thus, a time-step corresponds to 1 s in real time. In the special case $v_{\max} = 1$ the NaSch-model is reduced to the Asymmetric Simple Exclusion Process with parallel update and periodic boundary conditions [31].
The driving strategy of the NaSch-model comprises 3 aspects. In the first two steps of the update scheme, a vehicle tries to move as fast as possible. First, a car increases its velocity by one unit (step 1) if it is not already driving with v_{max}. If the gap to the predecessor is smaller than this velocity, the car has to brake in order to avoid a collision (step 2). Step 3 takes different behavioral patterns of the drivers into account and introduces a stochastic element into the vehicle dynamics: the velocity is reduced by one unit with the probability

p_{dec}. As a consequence, the velocity of free moving vehicles fluctuates, braking vehicles may overreact and the acceleration of the vehicles may be retarded.
The NaSch-model is able to describe basic properties of traffic flow like the spontaneous formation of jams. Also a realistic fundamental diagram in agreement with empirical data can be obtained. However, for the description of more complex traffic patterns, e.g., metastable states [6, 8] and synchronized traffic [80, 81], more sophisticated update rules are necessary [2]. In particular, the NaSch model is a minimal model in the sense that every further simplification leads to a loss of realism. Nevertheless, due to its simplicity the simulation of large scale traffic networks in real time is possible [30, 61, 62, 142].

[2] For a detailed analysis of the degree of realism of the most prominent cellular automaton models and a discussion of their properties see chapter 4.