Dirk Helbing, Kai Nagel

The physics of traffic and regional development

Journal article | **Accepted manuscript (Postprint)**
This version is available at https://doi.org/10.14279/depositonce-8218

This is an Accepted Manuscript of an article published by Taylor & Francis in Contemporary Physics on 20 Feb 2007, available online: http://www.tandfonline.com/10.1080/00107510410001715944.

Helbing, D.; Nagel, K. (2004). The physics of traffic and regional development. Contemporary Physics, 45 (5), 405–426. https://doi.org/10.1080/00107510410001715944

Terms of Use
Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

WISSEN IM ZENTRUM UNIVERSITÄTSBIBLIOTHEK

Technische Universität Berlin

The Physics of Traffic and Regional Development

Dirk Helbing*
Institute for Economics and Traffic, Dresden University of Technology, Andreas-Schubert-Str. 23, 01062 Dresden, Germany

Kai Nagel†
Transport Systems Analysis and Transport Telematics, Technical University of Berlin, Salzufer 17–19 SG 12, 10587 Berlin, Germany
(Dated: April 17, 2004)

This contribution summarizes and explains various principles from physics which are used for the simulation of traffic flows in large street networks, the modeling of destination, transport mode, and route choice, or the simulation of urban growth and regional development. The methods stem from many-particle physics, from kinetic gas theory, or fluiddynamics. They involve energy and entropy considerations, transfer the law of gravity, apply cellular automata and require methods from evolutionary game theory. In this way, one can determine interaction forces among driver-vehicle units, reproduce breakdowns of traffic including features of synchronized congested flow, or understand changing usage patterns of alternative roads. One can also describe daily activity patterns based on decision models, simulate migration streams, and model urban growth as a particular kind of aggregation process.

I. INTRODUCTION

From the point of view of a physicist, cars on a highway or pedestrians in a mall are many particle systems. Yet, the particles of those systems are more complicated than typical particles in physical systems. A question therefore is in how far the methods from physics can be applied to systems with these more complicated particles.

It turns out that in fact quite a lot can be explained with the methods from physics, which combine simple microscopic modeling assumptions with advanced mathematics and/or advanced computational science. For example, those systems display a surprising variety of states, such as jammed, laminar, or oscillating. Those states can be explained using several concepts, from microscopic to fluid-dynamical. Physics methods have contributed significantly to better understanding those states, and to alleviate them where possible.

Traffic usually unfolds on *networks*, rather than on flat two-dimensional space. And indeed, traffic has some similarity to the steady-state flow of electrons in a fuse network with non-linear resistors. Besides the fact that the traffic dynamics can be more complicated than this steady-state behavior, as pointed out in the last paragraph, the main difference between human travelers and electrons in a network is that human travelers typically have an individual destination. This introduces an important additional non-linearity, and also brings game theory into the picture, in particular the concept of a Nash Equilibrium (NE). Game theory states that a system *is* at a NE when no player can gain by unilaterally switching strategies. What is more of an interest to physicists is that certain types of learning systems can have a NE as fixed point attractor – they can, however, also have other types of attractors, such as periodic or chaotic ones. And indeed, route choice experiments with human subjects confirm that in many cases the players approach a NE, but also that this NE does not have to be stable and displays behavior reminiscent of intermittency and volatility clustering.

This and other evidence demonstrates that the route choice behavior of real world people can be approximated by a fastest path algorithm, which is relatively cheap in terms of computational complexity. From a practical point of view, the next problem then becomes to generate the destinations of the travelers. Two different approaches are discussed in the paper: (i) gravity models, which make the probability to select a destination dependent on the distance; and (ii) activity-based demand generation, which attempts to generate complete daily plans for all travelers in a system. And once more the situation is similar to physics: While it may be hopeless to realistically describe the behavior of individual people in the system, there is some hope that the macroscopic ("emergent") behavior of many travelers together can be derived from simple microscopic rules, akin to the derivation of the ideal gas equation. Also, behavioral invariances such as the energy consumption of the human – not the vehicle – can sometimes be found and used.

Intuitively, the next important question is then how the city itself develops – where humans find residences, where retail and other industries find locations, and where the city grows when the space inside its current boundaries is no longer sufficient. Once more, models from physics, this time growth models based on the cellular automata paradigm, can offer important insight. The use of such models is related to the fact that some quantities of real-world cities, such as the area, the perimeter, or the size

*Electronic address: ;

†Electronic address:

distribtion, display fractal properties. An alternative approach is to use models from Synergetics, with its associated master equation, as the starting point to describe migration dynamics.

Similar to the molecular dynamics approach in physics, it is now possible to build large scale "agent-based" simulation systems of real world transport systems which are based on the "first principles" introduced above. Since a typical metropolitan area has about 10^7 inhabitants, this is well within the range of computational feasibility, although some computational efficiency is lost because of the more complicated particles. The main challenge currently is to integrate the different levels, once more a challenge that is similar to multi-scale modeling in physics.

Results achieved so far indicate that simulation models based on those principles are, albeit just at the beginning of their development, at least as good as established methods. In consequence, one can start to move those models into the realm of real-world infrastructure planning. In addition, one would expect that additional scientific insights, better implementation techniques allowing faster turnaround, and the increasingly better availability of electronic data will make them considerably better in the future.

II. FLUID-DYNAMIC AND GAS-KINETIC TRAFFIC MODELS

Every driver has probably encountered the widespread phenomenon of so-called "phantom traffic jams", for which there is no visible reason such as an accident or a bottleneck [1]. So, why are vehicles sometimes stopped although everyone likes to drive fast? [2]

The first well-known approach to describe density waves in traffic flows goes back to Lighthill and Whitham [3] in 1955. They start with the typical *continuity equation,*

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial Q(x,t)}{\partial x} = 0\,, \tag{1}$$

where $\rho(x,t)$ denotes density and $Q(x,t)$ denotes flow or throughput. The equation reflects that vehicles are not generated or lost in the absence of ramps, intersections, or accidents. As usual, one also has the relation $Q(x,t) = \rho(x,t)\,V(x,t)$, where $V(x,t)$ is the velocity at place x and time t.

Lighthill and Whitham postulated that the traffic flow $Q(x,t)$ could be specified as a function of the density $\rho(x,t)$ only. This assumes an *instantaneous* adaptation of $Q(x,t)$ to some equilibrium flow-density relation $Q_{\rm e}(\rho(x,t))$. The corresponding curve $Q_{\rm e}(\rho) = \rho V_{\rm e}(\rho)$ is often called the *"fundamental diagram"* and obtained as a fit to empirical data.

Because of $\partial Q_e(\rho)/\partial x = (dQ_e/d\rho)\,\partial\rho/\partial x$, one obtains

$$\frac{\partial \rho}{\partial t} + C(\rho)\frac{\partial \rho}{\partial x} = 0\ , \tag{2}$$

which is a non-linear wave equation where the propagation velocity

$$C(\rho) = \frac{dQ_{\rm e}(\rho)}{d\rho} = V_{\rm e}(\rho) + \rho\frac{dV_{\rm e}(\rho)}{d\rho} \le V_{\rm e}(\rho) \tag{3}$$

of so-called *kinematic waves* depends on the vehicle density only. Kinematic waves have the property that they keep their amplitude, while their shape changes until *shock waves* (i.e. discontinuous changes in the density) have developed. The densities ρ_+ and ρ_- immediately upstream and downstream of a shock front determine its propagation speed

$$S(\rho_+,\rho_-) = \frac{Q_{\rm e}(\rho_+) - Q_{\rm e}(\rho_-)}{\rho_+ - \rho_-} = \frac{\Delta Q_{\rm e}}{\Delta\rho}\,. \tag{4}$$

Note that Eq. (4) is just the discrete version of Eq. (3).

Experimental observations of traffic patterns show additional features which cannot be reproduced by the above traffic model. While traffic flow appears to be stable with respect to perturbations at small and large densities, there is a linearly unstable range $[\rho_{\rm c2},\rho_{\rm c3}]$ at medium densities, where already small disturbances of uniform traffic flow give rise to traffic jams. Between the stable and linearly unstable density ranges, one finds meta- or *multi-stable* ranges $[\rho_{\rm c1},\rho_{\rm c2})$ and $(\rho_{\rm c3},\rho_{\rm c4}]$, since there exists a density-dependent, critical amplitude $\Delta\rho_{\rm c}$, so that the resulting traffic pattern is path- or *history-dependent* [4–7] (see Fig. 1). While subcritical perturbations fade away, supercritical perturbations cause a breakdown of traffic flow ("nucleation effect"). Consequently, traffic displays *critical points, non-equilibrium phase transitions, noise-induced transitions, and fluctuation-induced ordering phenomena.* One may view this as non-equilibrium analogue of the phase transitions between vapour (free flow), water ("synchronized", queued traffic flow above freeway capacity), and ice (jams with standing vehicles and zero flow). Note that noise-like effects enter the above deterministic models only via the boundary conditions, but some models take care of them by additional fluctuation terms or by distinguishing different driver-vehicle types, i.e. partial flows with non-linear interactions.

According to Krauß [8], traffic models show the observed *hysteretic phase transition* related with metastable traffic and high flows, if the typical maximal acceleration is not too large and the deceleration strength is moderate. In such kinds of models, $\rho_{\rm c1}$ is the density where the flow-density relation

$$J(\rho) = \frac{1}{T}\left(1 - \frac{\rho}{\rho_{\rm jam}}\right) \tag{5}$$

of traffic with fully developed traffic jams (the so-called "jam line") intersects with the free branch of the flow-density diagram $Q_{\rm e}(\rho)$. T denotes the net time gap (time clearance) in congested traffic and $\rho_{\rm jam}$ the density inside of traffic jams. The intuitive interpretation of Eq. (5)

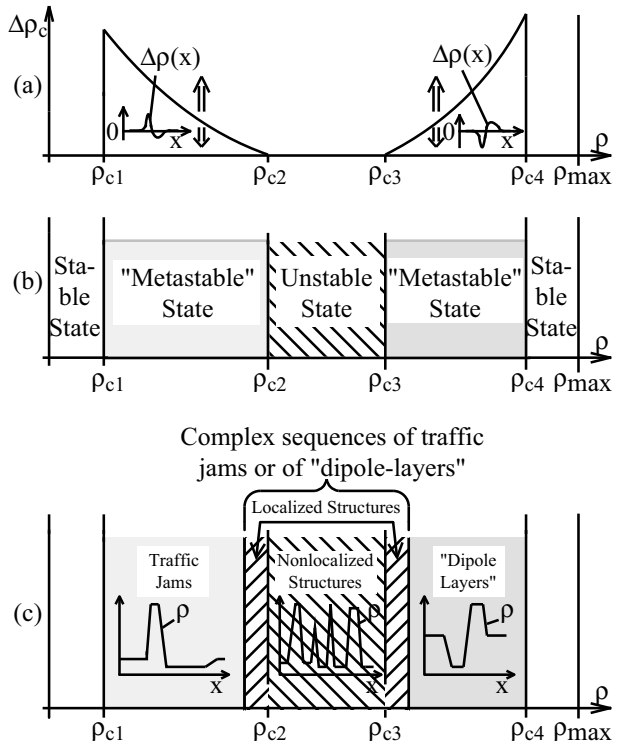

FIG. 1: Schematic illustration of (a) the chosen initial perturbation $\Delta\rho(x)$ and the density-dependent perturbation amplitudes $\Delta\rho_c$ required for jam formation in the meta-stable density regime, (b) the related instability diagram of homogeneous (and slightly inhomogeneous) traffic flow (predicting stable traffic at small and high densities, linearly unstable traffic at medium densities, and meta-stable traffic in between) and (c) the finally resulting traffic patterns $\rho(x)$ depending on the respective density regime. (Simplified diagram after [5, 6], see also [4].)

is that traffic in the congested regime is composed of moving areas and of areas where vehicles are stopped. The fraction of the moving areas is given by $1 - \rho/\rho_{\rm jam}$.

Moreover, the outflow $Q_{\rm out}$ from traffic jams is a self-organized constant [4, 7, 9], which lies between $Q_{\rm e}(\rho_{\rm c1})$ and $Q_{\rm e}(\rho_{\rm c2})$ [10]. The jam line corresponds to the flow-density relation for traffic patterns with a self-organized, stationary profile [4]. These propagate with the velocity $C = dJ/d\rho = -1/(T\rho_{\rm jam}) \approx -15$ km/h, which is another traffic constant. The explanation of this is the following: Once a traffic jam is fully developed, vehicles leave the downstream jam front at a constant rate, while new ones join it at the upstream front. This makes the jam move upstream with constant velocity.

In order to reproduce such emergent traffic jams and meta-stable density regimes, the Lighthill-Whitham model from Eqs. (1) to (4) needs to be generalized. For this, it is helpful that fluid-dynamic models can be derived from gas-kinetic models, which relate to models of interactive driver behavior [11, 12]. The *gas-kinetic approach* has been introduced by Prigogine *et al.* [13–15] and is inspired by *kinetic gas theory*. It describes the spatio-temporal change of the phase space density (= vehicle density × velocity distribution). The related equations are either of *Boltzmann-like* type (for point-like vehicles or low densities) or of *Enskog-like* type, if vehicular space requirements at moderate and high densities are taken into account (see [11] and references therein). They allow the systematic derivation of macroscopic equations for the vehicle density $\rho(x,t)$, the average velocity $V(x,t)$, the velocity variance $\Theta(x)$, etc. This hierarchy of equations is usually closed after the velocity or variance equation, although the separation of time scales assumed by the underlying approximations is weak. Nevertheless, the observed traffic dynamics is rather well reproduced by the resulting coupled partial differential equations. The density equation is just the continuity equation

$$\frac{\partial\rho}{\partial t} + \frac{\partial(\rho V)}{\partial x} = \nu_+ - \nu_- \,, \tag{6}$$

where ν_+ and ν_- denote on- and off-ramp flows, respectively. The velocity equation can be cast into the form

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = -\frac{1}{\rho}\frac{\partial P}{\partial x} + \frac{1}{\tau}(V^{\rm e} - V)\,. \tag{7}$$

In theoretically consistent macroscopic traffic models such as the gas-kinetic-based traffic model, the "traffic pressure" P and the velocity $V^{\rm e}$ are non-local functions of the density ρ, the average velocity V, and the variance Θ [16, 17]. Several well-known models are special cases of Eq. (7):

- The Lighthill-Whitham model of Eqs. (1) to (4) is obtained in the (unrealistic) limit $\tau \to 0$ of vanishing adaptation times τ. In that case $V(x,t) = V_e(\rho(x,t))$, and therefore $Q(x,t) = \rho(x,t)\,V_e(\rho(x,t)) =: Q_e(\rho(x,t))$, which depends on density only as was postulated for the LW model.

- Payne's macroscopic traffic model [18, 19] is obtained for $P(\rho) = [V_{\rm max} - V_{\rm e}(\rho)]/(2\tau)$ and $V^{\rm e} = V_{\rm e}(\rho)$, where $V_{\rm max}$ denotes the maximum speed (the average velocity at very low densities).

- Kerner's and Konhäuser's model [4, 20] is a variant of Kühne's model [21] and corresponds to the specifications $P = \rho\Theta_0 - \eta_0\partial V/\partial x$ and $V^{\rm e} = V_{\rm e}(\rho)$, where Θ_0 and η_0 are positive constants. The corresponding equation is a Navier-Stokes equation with a viscosity term $\eta_0\partial^2 V/\partial x^2$ and an additional relaxation term $[V_{\rm e}(\rho) - V]/\tau$ describing the delayed adaptation to the velocity-density relation $V_{\rm e}(\rho)$.

Both for the Payne model and for the Kerner-Konhäuser model, the linearly unstable regime $[\rho_{\rm c2}, \rho_{\rm c3}]$ is determined by the densities ρ fulfilling the condition

$$\rho\left|\frac{dV_{\rm e}(\rho)}{d\rho}\right| > \sqrt{\frac{dP(\rho)}{d\rho}}\,. \tag{8}$$

In consequence, both the Payne model and the Kerner-Konhäuser model have linearly unstable ranges if $dV_{\rm e}(\rho)/d\rho$ is large, while the Lighthill-Whitham model is marginally stable.

Beyond linear stability analysis, in some cases it is also possible to find analytical expressions for the large amplitude instability in the meta-stable ranges $[\rho_{\rm c1}, \rho_{\rm c2})$ and $(\rho_{\rm c3}, \rho_{\rm c4}]$ (Fig. 1). For these more complicated aspects, the interested reader is referred to [22].

In simple terms, the reason for emergent traffic jams is as follows: Due to the finite adaptation time (= reaction + acceleration time), a small disturbance in the traffic flow $Q_{\rm e}$ can cause an *overreaction* (overbraking) of a driver, if the safe vehicle speed $V_{\rm e}(\rho)$ (which satisfies $dV_{\rm e}(\rho)/d\rho \leq 0$) drops too rapidly with increasing vehicle density ρ. At high enough densities ρ, this will give rise to a *chain reaction* of the followers, as other vehicles will have approached before the original speed can be regained. This *feedback* can eventually cause the unexpected standstill of vehicles known as traffic jam.

III. CONGESTED TRAFFIC STATES IN THEORY AND EMPIRICAL DATA

With this knowledge one can understand the various congested traffic states observed on freeway sections with bottlenecks [10, 23–26]. Given that the flow Q and density ρ are measured *per lane,* a bottleneck due to ramp flows is determined by $\nu_+ = Q_{\rm rmp}/(IL)$, where L is the used length of the on-ramp and I the number of freeway lanes. The corresponding bottleneck strength is

$$\Delta Q = \frac{Q_{\rm rmp}}{I}\,. \tag{9}$$

A. Theoretical Phase Diagram of Traffic States

Let $Q_{\rm up}$ denote the traffic flow per lane upstream of the bottleneck and

$$Q_{\rm tot} = Q_{\rm up} + \Delta Q = Q_{\rm up} + \frac{Q_{\rm rmp}}{I} \tag{10}$$

the total capacity required downstream of the ramp. Then, we expect to always observe free traffic (FT) below the first instability threshold at density $\rho_{\rm c1}$, i.e. if

$$Q_{\rm tot} = Q_{\rm up} + \Delta Q < Q_{\rm e}(\rho_{\rm c1})\,. \tag{11}$$

Traffic flow will always be congested, if the maximum flow $Q_{\rm max} = \max_\rho Q_{\rm e}(\rho)$ (the capacity) is exceeded, i.e.

$$Q_{\rm tot} = Q_{\rm up} + \Delta Q > Q_{\rm max}\,. \tag{12}$$

At least if the density related with the maximum flow $Q_{\rm max}$ lies between $\rho_{\rm c1}$ and $\rho_{\rm c2}$, traffic states between the two diagonal lines $Q_{\rm up} + \Delta Q = Q_{\rm e}(\rho_{\rm c1})$ and $Q_{\rm up} + \Delta Q = Q_{\rm max}$ in the $Q_{\rm up}$-vs-ΔQ phase space can be either congested or free, depending on the initial and boundary conditions [26] (see Fig. 2). While homogeneous free flow may persist over long time periods, large perturbations tend to produce congested states. Extended congested traffic can emerge above the line

$$Q_{\rm up} = Q_{\rm out} - \Delta Q \tag{13}$$

in the phase diagram, i.e. if $Q_{\rm tot}$ is greater than the dynamic capacity $Q_{\rm out}$. This line does not have to be parallel to the previously mentioned phase boundaries, as $Q_{\rm out}$ may depend on the bottleneck strength ΔQ [1, 27]. For $Q_{\rm e}(\rho_{\rm c1}) \leq Q_{\rm tot} = Q_{\rm up} + \Delta Q < Q_{\rm out}$, congested traffic states are always localized, i.e. they never grow over long sections of the freeway, as the outflow can cope with the overall traffic volume.

The traffic flow $Q_{\rm cong}$ resulting in the congested area plus the inflow or bottleneck strength ΔQ are normally given by the outflow $Q_{\rm out}$, i.e.

$$Q_{\rm cong} = Q_{\rm out} - \Delta Q \tag{14}$$

(if vehicles cannot enter the freeway downstream of the congestion front). One can distinguish the following cases: *Homogeneous congested traffic* (HCT) such as typical traffic jams during holiday seasons can occur, if the density $\rho_{\rm cong}$ associated with the congested flow

$$Q_{\rm cong} = Q_{\rm e}(\rho_{\rm cong}) \tag{15}$$

lies in the stable or meta-stable range

$$Q_{\rm cong} < Q_{\rm e}(\rho_{\rm c3})\,, \quad \text{i.e.} \quad \Delta Q > Q_{\rm out} - Q_{\rm e}(\rho_{\rm c3})\,. \tag{16}$$

Oscillating forms of congested traffic can emerge, if

$$\Delta Q \leq Q_{\rm out} - Q_{\rm e}(\rho_{\rm c4}) \quad \text{and} \quad Q_{\rm up} > Q_{\rm out} - \Delta Q\,. \tag{17}$$

That is, lower bottleneck strengths tend to produce less serious congestion, namely oscillating rather than homogeneous congested flow. We either find *oscillating congested traffic* (OCT), *triggered stop-and-go traffic* (TSG), or *moving localized clusters* (MLC) [26]. In contrast to OCT, stop-and-go traffic is characterized by a sequence of moving jams, between which traffic flows freely. This state can either emerge from a spatial sequence of homogeneous and oscillating congested traffic [28], which is also called the *"pinch effect"* [29], or it can be caused by the inhomogeneity at the ramp. In the latter case, each traffic jam triggers another one by inducing a small perturbation in the inhomogeneous freeway section (see Fig. 2), which propagates downstream as long as it is small, but turns back when it has grown large enough *("boomerang effect").* This growth requires the downstream traffic flow to be linearly unstable. If it is metastable (when the traffic volume $Q_{\rm tot}$ is sufficiently small), small perturbations will fade away. Therefore, if

$$Q_{\rm e}(\rho_{\rm c1}) \leq Q_{\rm tot} = Q_{\rm up} + \Delta Q \leq Q_{\rm out} \leq Q_{\rm e}(\rho_{\rm c2})\,, \tag{18}$$

one expects to find localized traffic states, either a single *moving localized cluster* (MLC), or a *pinned localized*

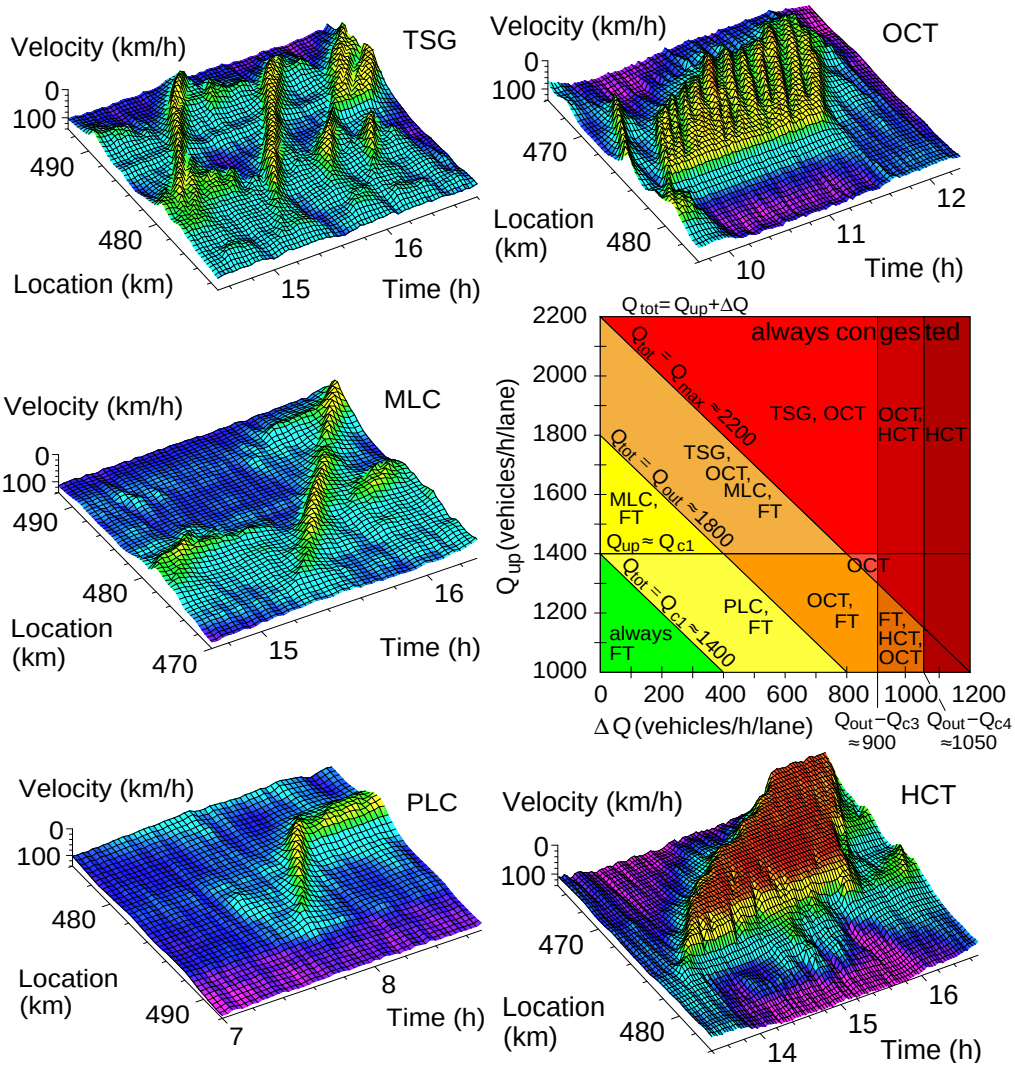

FIG. 2: Empirical representatives of different kinds of congested traffic observed on the German freeway A5 close to Frankfurt (after [2, 26]) and theoretical phase diagram of traffic states (after [26]) in the presence of one bottleneck as a function of the upstream flow $Q_{\rm up}$ and the bottleneck strength ΔQ (center right).

cluster (PLC) at the location of the ramp. The latter requires the traffic flow in the upstream section to be stable, i.e.

$$Q_{\rm up} < Q_{\rm e}(\rho_{\rm c1})\,, \tag{19}$$

so that no traffic jam can survive there. In contrast, moving localized clusters and triggered stop-and-go waves require [26]

$$Q_{\rm up} \geq Q_{\rm e}(\rho_{\rm c1})\,. \tag{20}$$

The simulation results displayed in Figure 2 summarize, for freeways with a single bottleneck, which states are typically found in different areas of the phase diagram. Apart from the detailed shape and exact location of the phase boundaries, this phase diagram is expected to be universal for all microscopic and macroscopic, stochastic and deterministic traffic models with the same *instability diagram* (with stable, meta-stable, and unstable density ranges) [1]. Results for more complex freeway geometries are available as well.

B. Wide Scattering of Traffic Data in Heterogeneous Traffic

In the congested traffic regime, flow-density data are characterized by a surprisingly wide scattering [30] (see Fig. 3a). This has led people to question the applicability of the jam line (5) [31–33]. However, it can be shown that, taking into account the variation of the average net time gap (time clearance) T, the variations in the data can be reproduced with a correlation coefficient of 0.92 (compared to 0.35 when only the density is varied).

A closer analysis reveals a large variation of time gaps between vehicles. The time gaps of trucks are particularly

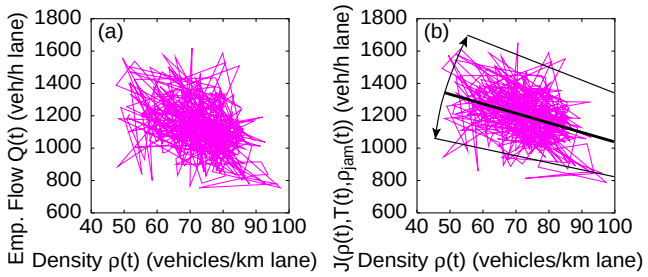

FIG. 3: The two-dimensional scattering of empirical flow-density data in synchronized traffic flow of high density $\rho \geq 45$ veh/km/lane (see (a)) is well reproduced by the jam relation (5), when not only the variation of the density ρ, but also the empirically measured variation of the average time gap T [and the maximum density $\rho_{\max}$] is taken into account (see (b)). The pure density-dependence $J(\rho)$ (thick black line) is linear and cannot explain a two-dimensional scattering. However, variations of the average time gap T change its slope $-1/(\rho_{\max}T)$ (see arrows), and about 95% of the data are located between the thin lines $J(\rho, T \pm 2\Delta T, 1/l) = (1-\rho l)/(T \pm 2\Delta T)$, where $l = 3.6$ m is the average vehicle length, $T = 2.25$ s the average time gap, and $\Delta T = 0.29$ s the standard deviation of T. The predicted form of this area is club-shaped, as demanded by Kerner [31–33]. (After [35].)

large. Simulations of heterogeneous traffic with different kinds of vehicles (i.e. different parameter values) suggest that at least part of the scattering of flow-density data may be explained by the mixture of cars and trucks [34] and of different driving behaviors (see Fig. 4). Although heterogeneous traffic can be treated by macroscopic models (see Fig. 4), a simulation of individual vehicles is much more easy, flexible, and efficient, in particular if one likes to simulate network traffic of drivers with different origins and destinations. Therefore, we will focus on some "microscopic" traffic models in the following.

IV. DRIVEN MANY-PARTICLE SYSTEMS AND MICROSCOPIC TRAFFIC SIMULATION

The observations in freeway traffic can not only be described by fluid-dynamic or "macroscopic" traffic models. They are also well reproduced by "microscopic" models [27]. Note that a theoretical connection between both approaches exists. It is called the *micro-macro link* [12].

A. Car-Following Models

Microscopic models are often *follow-the-leader models* specifying the acceleration dv_i/dt of each single vehicle i as a function F_i of their speed v_i, their distance headway $d_i = x_{i-1} - x_i$ with respect to the leading vehicle $i-1$, and/or their relative velocity $\Delta v_i = v_i - v_{i-1}$:

$$\frac{dv_i}{dt} = F_i(d_i, v_i, \Delta v_i) + \xi_i(t)\,. \tag{21}$$

$\xi_i(t)$ are random variations in the acceleration behavior, while F_i may be interpreted as average situation-dependent acceleration force. A typical example is the non-integer car-following model

$$\frac{dv_i(t+\Delta t)}{dt} = -\frac{\Delta v_i(t)}{\tau}\frac{[v_i(t+\Delta t)]^\ell}{[d_i(t)]^m} \tag{22}$$

with the reaction time $\Delta t \approx 1.3$s and the parameters $\tau \approx \Delta t/0.55$, $\ell \approx 0.8$, and $m \approx 2.8$ [36]. It has a linearly unstable range for $\tau/T > 1/2$. (The time delay Δt on both sides facilitates to determine analytical solutions.) A simpler model is the optimal velocity model

$$\frac{dv_i(t)}{dt} = \frac{1}{\tau}\Big[V\big(d_i(t)\big) - v_i(t)\Big]\,, \tag{23}$$

where $V(d_i)$ is the "optimal" velocity-distance relation and τ the adaptation time [37, 38]. This model has an unstable range for $dV(d_i)/dd_i > 1/(2\tau)$. The respective non-linearly coupled differential equations (or stochastic differential equations, if fluctuations are taken into account) are numerically solved as in molecular dynamics.

Let us now define the maximum, desired, or free velocity by $V_{\max} = \max V(d_i)$. The expression

$$f(s_i) = \frac{V(s_i + l_i) - V_{\max}}{\tau} \tag{24}$$

reflects the interaction force f as a function of the net distance (clearance) $s_i = d_i - l_i = (x_{i-1} - x_i - l_i)$ between two successive vehicles, where l_i denotes the length of vehicle i. With this, Eq. (23) can be reformulated in terms of the *"social force model"*

$$\frac{dv_i}{dt} = \frac{V_{\max} - v_i}{\tau} + f(s_i) + \xi_i(t)\,. \tag{25}$$

B. Interaction Forces Among Vehicles

The difficulty is now to specify the interaction forces among vehicles in an appropriate way. Their exact specification is essential for realistic traffic simulations, which are required for the design of efficient and reliable traffic optimization measures such as intelligent on-ramp controls, driver assistance systems, lane-changing assistants, etc. If the fluctuation term was zero (i.e. $\xi_i(t) = 0$), the interaction force could simply be determined from the velocity-density relation. Since the average of the headways $d_i = s_i + l_i$ determines the inverse density $1/\rho$, we would simply find the relation $f(s_i) = [V_{\rm e}(1/(s_i + l_i)) - V_{\max}]/\tau$.

However, things are considerably more difficult, when random fluctuations are not negligible. One may think of applying scattering theory to determine the interaction potential $U(s) = -df(s)/ds$ from statistical distributions, but learning about human interactions requires a somewhat different approach. Progress has recently

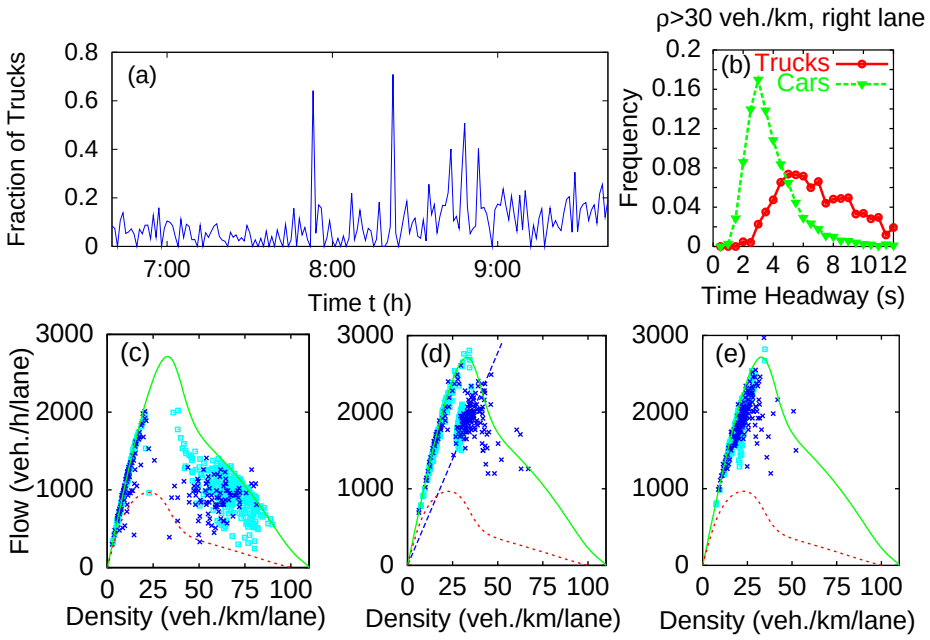

FIG. 4: (a) The empirical truck fraction varies considerably in the course of time. (b) The time headways of long vehicles ("trucks") are on average much higher than those of short vehicles ("cars"). (c)-(e) Assuming a fundamental diagram for cars (solid line), a separate one for trucks (dashed line), weighting them according to the measured truck fraction, and using empirical boundary conditions, allows one to reproduce the observations in a (semi-)quantitatively way [34]: (c) Free traffic (at low densities) is characterized by a (quasi-)one-dimensional curve. Data of congested traffic *upstream* of a bottleneck are widely scattered in a two-dimensional area. (d) *Immediately downstream* of the bottleneck, one observes homogeneous-in-speed states reflecting recovering traffic. (e) Further downstream, the data points approach the curve describing free traffic. Dark symbols correspond to empirical one-minute data, light ones to corresponding simulation results. (After [25].)

been made by means of *random matrix theory,* a powerful method from quantum and statistical physics. It allows one to determine the interaction potentials via net distance distributions of vehicles.

Random matrix theory [39] has been developed for many-particle systems exposed to a "thermal bath" of a given reciprocal temperature β, i.e. to random forces of a certain variance and statistics. The resulting velocity and net distance distributions allow one to draw conclusions about the interaction potential $U(s) = -df(s)/ds$, as this determines their shapes. Although a generalization to arbitrary potentials $U(s)$ is possible, here we will focus our attention on power-law potentials

$$U(s) \propto s^{-\alpha} \tag{26}$$

for simplicity. The exponent $\alpha > 0$ is a fit parameter and s is the net distance (clearance) between two successive cars. Such a potential describes the repulsive tendency of drivers to keep a safe distance to the respective car ahead. The related net distance distributions can be calculated to be

$$P(\rho s) = Ae^{-\beta(\rho s)^{-\alpha}} e^{-B\rho s}, \tag{27}$$

where ρ is the vehicle density and A, B represent normalization constants determined by requiring $\int_0^\infty P(\rho s)\, d(\rho s) = 1$ and $\langle \rho s \rangle \equiv \int_0^\infty \rho s P(\rho s)\, d(\rho s) = 1$.

Originally, the applied random matrix method is an equilibrium concept assuming the conservation of energy $\mathcal{H} = \mathcal{T} + \mathcal{V}$, i.e. a transformation of potential energy $\mathcal{V} = \sum_{i=1}^{N} U(|x_{i-1} - x_i - l_i|)$ into kinetic energy $\mathcal{T} = \sum_{i=1}^{N} m_i v_i^2/2$ with x_i, l_i, m_i, and v_i being the location, length, mass, and velocity of the ith vehicle. It can however be shown [40] that this is a reasonable approximation for an ensemble of N vehicles i in a stationary state at a given density ρ and inverse "temperature" (velocity variance) β, while the stop-and-go regime between 20 and 40 vehicles per kilometer and lane must be excluded from an investigation of this kind.

Milan Krbalek and one of us have separately analyzed eight density intervals of Dutch single vehicle data in the free low-density regime and eight density intervals in the congested traffic regime. In accordance with the theoretical predictions, the velocity distributions fit Gaussian distributions very well [41]. Moreover, the best fit of the net distance distributions is obtained for the integer parameter $\alpha = 4$ in the free traffic regime, while we find an excellent agreement with $\alpha = 1$ in the congested regime (see Figs. 5, 6). This is not only a strong support of studies questioning a uniform behavior of drivers in all traffic regimes [42]. It also offers an interpretation of the mysterious fractional distance-scaling exponent $\alpha + 1 \approx 2.8$ in classical follow-the-leader models [43], which average over the driver behavior in the free and congested regimes.

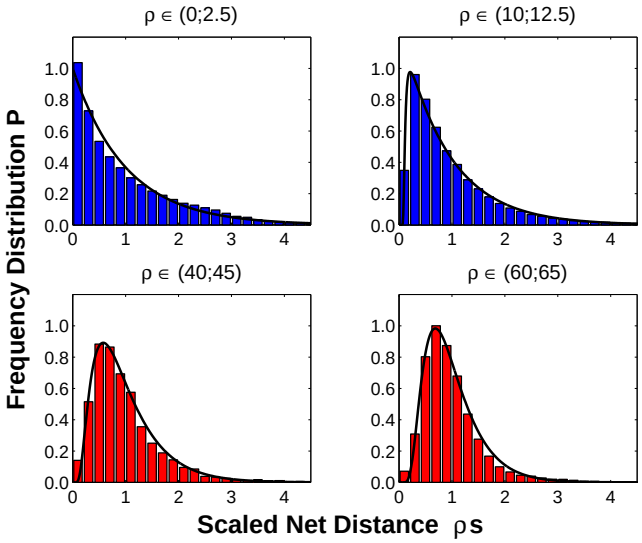

FIG. 5: Representative distributions of net distances (clearances) among successive cars for various densities ρ, determined from Dutch single-vehicle data (after [44]). The solid fit curves correspond to the theoretical distributions (27) for the empirical values of ρ and β. In the free traffic regime, the only fit parameter is $\alpha = 4$ (above, blue), while $\alpha = 1$ in the congested regime (below, red).

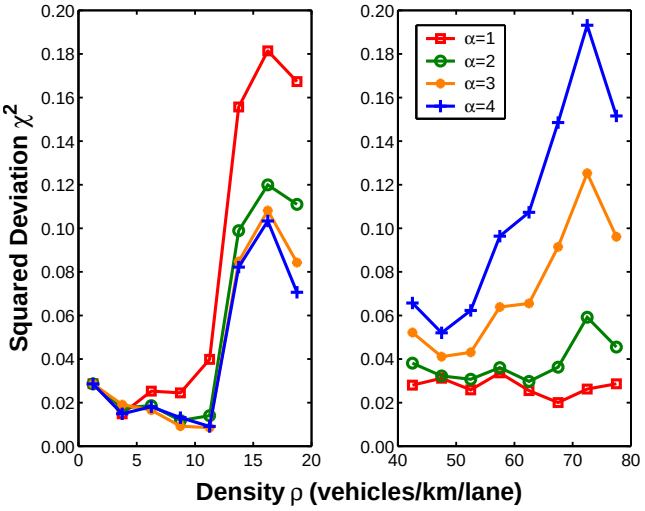

FIG. 6: Sum of squared deviations between the empirical and theoretical net distance distributions for various fit parameters α (after [44]). The best fit is $\alpha = 4$ in the free traffic regime and $\alpha = 1$ in the congested regime.

Note that the net distance distribution (27) agrees with the Pearson III type of distribution [45] which has been suggested to fit empirical time headway distributions already before our theoretical explanation had been found. Therefore, the determination of interaction potentials in freeway traffic contributes to the challenging problem of understanding time headway or distance distributions of cars [45, 46]. In addition, the identification of behavioral regimes points to adaptive driver behavior.

V. SIMULATION OF LARGE TRAFFIC NETWORKS WITH CELLULAR AUTOMATA

A. Cellular Automata Rules for Links

An alternative approach to car-following models are rule-based cellular automata (CA). For the CA technique, road links are divided into cells, say of length 7.5 m, which can contain at most one car each, and there is one array of cells for each lane, see Fig. 7(a). Movement is performed by jumping from one cell to another, where the new cell is determined by a set of "driving rules." A good update time step is 1 sec (justified by reaction time arguments). Taking this together means, for example, that a jump of 5 cells in a time step models a speed of 5×7.5 m/sec $= 135$ km/h.

Driving rules of traffic on a link consist of two parts: car following, and lane changing. Typical rules for car following are: deterministic speed calculation, randomization, and movement. The *deterministic speed calculation* rule first computes a new speed for each car based on its current speed and closeness to the car in front of it. An example for such a rule is

$$v_{\mathrm{safe},t} = \min[v_{t-1} + 1, v_{\max}, g_t] \ ,$$

where $v_{t-1}+1$ models acceleration, $v_{\max}$ is a speed limit, and g is the gap, i.e. the number of empty spaces to the vehicle ahead. See Fig. 7(a) for an illustration of gap. Next, the *randomization* rule introduces a probability p_{noise} for each car so that instead of following the above speed deterministically, it may instead drive one velocity level more slowly:

$$v'_t = \begin{cases} \max[0, v_{\mathrm{safe},t} - 1] & \text{with probability } p_{\mathrm{noise}} \\ v_{\mathrm{safe},t} & \text{else.} \end{cases} .$$

In the *movement* rule, each particle is moved forward:

$$x_{t+1} = x_t + v'_t \ .$$

To illustrate the above rules, Fig. 7(a) shows traffic moving to the right. Using a p_{noise} value of 0.2, the leftmost vehicle accelerates to velocity 2 with probability 0.8 and stays at velocity 1 with probability 0.2. The middle vehicle slows down to velocity 1 with probability 0.8 and to velocity 0 with probability 0.2. The rightmost vehicle accelerates to velocity 3 with probability 0.8 and stays at velocity 2 with probability 0.2. Velocities are in "cells per time step." All vehicles are then moved according to their velocities using the movement rule described above.

Fig. 8 shows, in a so-called space-time diagram for traffic, the emergence of a traffic jam. Lines show configurations of a segment of road in second-by-second time

steps, with time increasing in the downward direction; cars drive from left to right. Integer numbers denote the velocities of the cars. For example, a vehicle with speed "3" will move three sites (dots) forward. Via this mechanism, one can follow the movement of vehicles from left to right, as indicated by an example trajectory.

B. Lane Changing

Typical rules for lane changing (see Fig. 7(b)) include a reason to change lanes and a safety criterion for changing lanes. First, there needs to be a reason why a vehicle wants to change lanes, for example that the other lane is faster, or that it needs to get into a certain lane in order to make an intended turn at the end of the link. A possible rule to model the first is "check if the (forward) gap in the other lane is larger than the gap in the current lane." Second, the vehicle needs to check that there is really enough space in the destination lane. A possible rule is that the forward gap in the other lane needs to be larger than v_t, and the backward gap in the other lane needs to be larger than the velocity of the oncoming vehicle.

Figure 7(b) illustrates the lane changing rules. Only lane changes to the left lane are considered. In situation I, the leftmost vehicle on the bottom lane will change to the left because the forward gap on its own lane, 1, is smaller than its velocity, 3; the forward gap in the other lane, 10, is larger than the gap on its own lane, 1; the forward gap in the target lane is large enough; and the backward gap is large enough. In situation II, the second vehicle from the right on the right lane will not perform a lane change because the gap backwards on the target lane is not sufficient.

Due to the complexity of the dynamics, it is inconvenient to do both lane changing and car following in one parallel update. It is however possible, and convenient for parallel computing, to do the update in two completely parallel sub-timesteps: First, all lane changes are executed in parallel based on information from time t, resulting in intermediate information, at time $t+\frac{1}{2}$. Then, all car following rules are executed in parallel, based on information from time $t+\frac{1}{2}$, which results in information for time $t+1$.

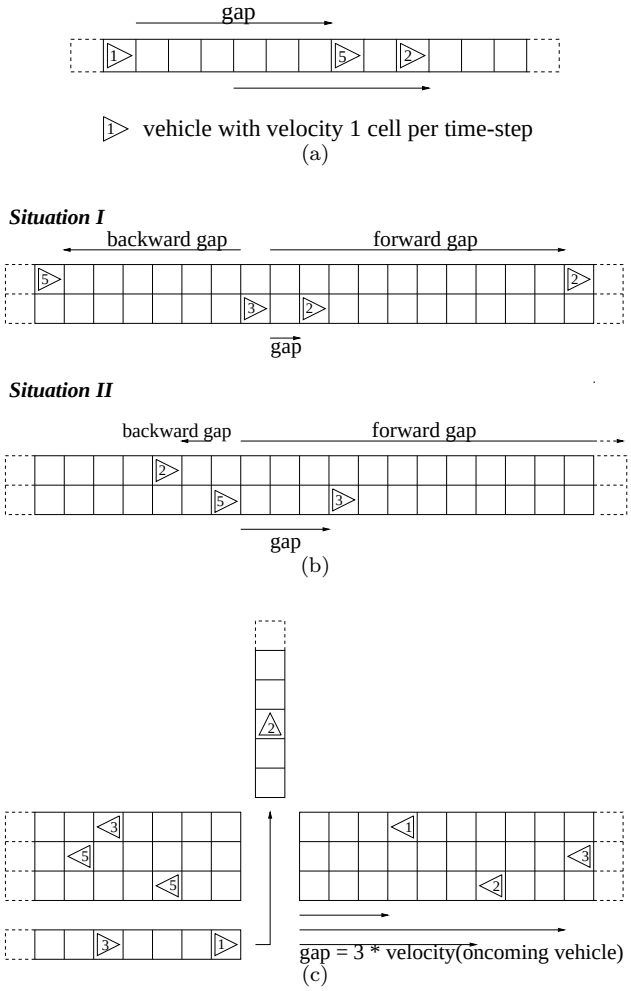

FIG. 7: Illustration of cellular automata driving rules (from [50]). (a) Definition of *gap* and one-lane update. (b) Lane-changing. (c) Left turn against oncoming traffic.

C. Intersections

The typical modeling substrate for transportation is the network, which is composed of links and nodes that model roads and intersections, respectively. Since the full simulation of an intersection is rather complicated, for large scale applications one attempts to reduce the complexity by resolving all conflicts at the entering of the intersection. Under such a simplification, it is still possible to model things like left turns against oncoming traffic, but the decision is made before the vehicle enters the intersection, not inside the intersection. In consequence, complications inside the intersection, such as the interesting dynamics of a large roundabout in Paris, are not modeled by this approach. Nevertheless, the simplified approach is quite useful for large scale applications.

Given this, there are only two remaining cases for traffic: Protected and unprotected turns. A protected turn means that a signal schedule takes care of possible conflicts, and thus the simulation does not have to worry about them. In this case, it is enough if the simulation makes vehicles stop when a signal is red, and lets them go when a signal is green.

In unprotected turns, conflicts are resolved by the legal rules, not by signals. For example, a vehicle making a left turn needs to give priority to all oncoming vehicles (Fig. 7(c)). This entails that one needs, for each turning direction, to encode which other oncoming lanes have the priority. Once this is done, a vehicle that intends to do a certain movement only needs to check if there is a vehicle approaching on any of these interfering lanes. If so,

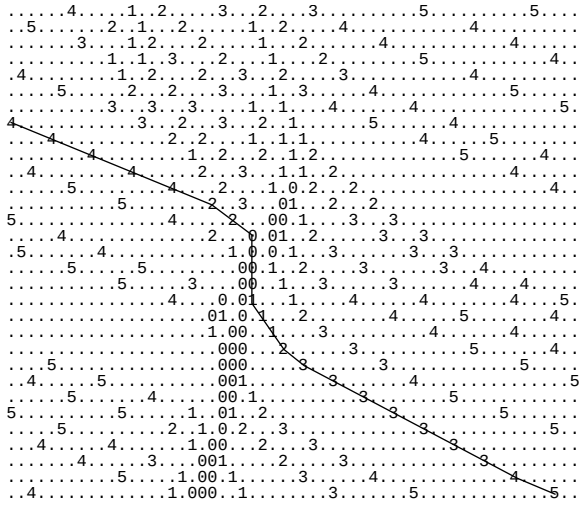

FIG. 8: Emergence of a jam "out of nowhere" ("phantom traffic jam"), simulated with a stochastic cellular automaton.

the vehicle under consideration has to stop, otherwise it can proceed. As said before, once the vehicle has entered the intersection, other vehicles are no longer regarded. Figure 7(c) shows an example of a left turn against oncoming traffic. The turn is accepted because on all three oncoming lanes, the gap is larger or equal to three times the first oncoming vehicle's velocity.

D. Modeling Queues

Sometimes, even the relatively simple CA rules are computationally too slow. Another factor of ten in the execution speed can often be gained by simplifying the dynamics even more. In the so-called queue model [47, 48], links have no internal structure any more, they can be considered as vehicle reservoirs. These reservoirs have a storage constraint, N_{max}, which varies from link to link, and is based on the physical attributes of the link, such as length and number of lanes. Once the storage constraint is exhausted, no more vehicles can enter the link until some other vehicles have left the link.

Vehicles can only leave the link if their destination link has space, if they have spent the time it takes to traverse the link, and if the flow capacity constraint is fulfilled. The flow capacity constraint is a maximum rate with which vehicles can leave the link.

Except for the storage constraint and its consequences, this is just a regular M/M/1 queuing model. The storage constraint does, however, necessitate one adaptation of the model: Since destination links can be full, there is now competition for space on the destination link. A good option is to give that space randomly to incoming links, with a bias toward incoming links with higher capacity [49].

With the queue model and using parallel computing, it is possible to simulate a full day of all car traffic of all of Switzerland (approx. 7 mio inhabitants) in less than 5 minutes [49]. Although data movement is still a challenge, this makes it possible to run completely microscopic (agent-based) studies of large metropolitan areas [117].

VI. ENTROPY LAWS AND DESTINATION CHOICE

The simulation of traffic in street networks requires data on the number V_{kl} of trips between origins k and destinations l as a function of the time of the day, the weekday etc. These origin-destination data are relatively scarce and expensive. However, the number $Q_k = \sum_l V_{kl}$ of trips starting in a city k or quarter of a city and the number $Z_l = \sum_k V_{kl}$ of trips ending in l is relatively well-known, as this requires only $2n$ instead of n^2 data, where n denotes the number of dinstinguished origins/destinations. Apart from this, Q_k and Z_l can be automatically measured by detectors, while the determination of the entries V_{kl} of the traffic flow matrix requires to obtain origin-destination pairs, i.e. to ask drivers.

Therefore, the relative frequencies $p_{kl} = V_{kl}/V$ from the origins k to the destinations l, where $V = \sum_{k,l} V_{kl}$, is often determined via the "most likely" distribution under the constraints

$$\sum_l p_{kl} = Q_k/V = p_k \quad \text{and} \quad \sum_k p_{kl} = Z_l/V = p_l \,. \tag{28}$$

This is obtained by minimizing the *"information gain"*

$$I = -\sum_{k,l} p_{kl} \ln \frac{p_{kl}}{b_{kl}} \tag{29}$$

compared to some "natural distribution" b_{kl}. This is analogous to maximizing entropy $H = -\sum_{kl} p_{kl} \ln p_{kl}$ in cases where b_{kl} is an equidistribution. The only difference is that (29) takes into account a "resistance function" b_{kl}, which reflects the "natural distribution" as a function of the distance and other behaviorally relevant variables in the absence of constraints like (28). The appropriate specification of b_{kl} will be discussed in Sec. VI A.

Instead of minimizing (29) under the constraints (28), it is easier to minimize

$$\begin{aligned} L \;=\; & -\sum_{k,l} p_{kl} \ln \frac{p_{kl}}{b_{kl}} \\ & + \sum_k \lambda_k \left(p_k - \sum_l p_{kl} \right) \\ & + \sum_l \mu_l \left(p_l - \sum_k p_{kl} \right) , \end{aligned} \tag{30}$$

where the additional terms vanish when the constraints (28) are fulfilled. Differentiation with respect to $p_{k'l'}$ gives

$$-\ln \frac{p_{k'l'}}{b_{k'l'}} - 1 - \lambda_{k'} - \mu_{l'} = 0 \tag{31}$$

or

$$p_{kl} = b_{kl} \underbrace{\mathrm{e}^{-\lambda_k - 1/2}}_{=g_k} \underbrace{\mathrm{e}^{-\mu_l - 1/2}}_{=h_l} . \qquad (32)$$

The $2n$ Lagrange parameters λ_k and μ_l are now determined in a way that the $2n$ constraints (28) are satisfied. This requires the use of numerical iteration procedures. For an overview of numerical solution algorithms see Refs. [51–53]. These references also suggest ways to treat the choice of destinations and transport modes simultaneously.

A. Energy Laws and Resistance Function

According to the last paragraphs, the distribution of traffic significantly depends on the "natural distribution" b_{kl}. Robert Kölbl and one of us have investigated the electronically available empirical data of the UK National Travel Survey during the years 1972–1998 [54]. When we distinguished between different daily modes of transport j, we found that the average modal travel time $\overline{t}_j$ per day and person remained almost constant over the 27 years of observation, despite variations by a factor 3.8 between different modes. More specifically, the average travel times $\overline{t}_j$ were 40 min during a day with walking ($j = \mathrm{w}$), without any usage of other means of transport, 42 min for cyclists ($j = \mathrm{c}$), 67 min for stage bus users ($j = \mathrm{b}$), 75 min for car drivers ($j = \mathrm{d}$), 59 min for car passengers ($j = \mathrm{p}$), and 153 min for train passengers ($j = \mathrm{t}$).

When we scaled the t-axis of the modal travel-time distributions $P_j(t)\,dt$ by the average travel times $\overline{t}_j$, we discovered that, within the statistical variation, the resulting distributions collapsed onto one single curve

$$P(t_j)\,dt_j \approx N \exp\left(-\frac{\alpha}{t_j} - \frac{t_j}{\gamma}\right) dt_j \qquad (33)$$

with $t_j = t/\overline{t}_j$, two fit parameters α and γ, and the normalisation constant $N = N(\alpha, \gamma)$ (see Fig. 9a). This implies a universal law of human travel behaviour and the resistance function or natural distribution

$$b_{kl} = N \exp\left(-\frac{\alpha}{w_{kl}} - \frac{w_{kl}}{\gamma}\right) , \qquad (34)$$

where $w_{kl} = t_{kl}/\overline{t}_j$ is the travel time t_{kl} from k to l divided by the average travel time $\overline{t}_j$ for transport mode j.

In the semi-logarithmic representation $\ln P(t_j) = \ln N - \alpha/t_j - t_j/\gamma$ of (33), the term α/t_k is relevant only for short time scales up to $t_j \approx 0.5$, while the linear relationship $\ln P(t_j) = \ln N - t_j/\gamma$ dominates clearly over a wide range of scaled travel times t_j (see Fig. 9b). Therefore, we suggest the following interpretation: The dominating term $P_j(t_j)\,dt_j \propto \exp(-t_j/\gamma)\,dt_j$ corresponds to the canonical energy distribution

$$P'(E_j)\,dE_j = N' \exp(-E_j/\bar{E})\,dE_j , \qquad (35)$$

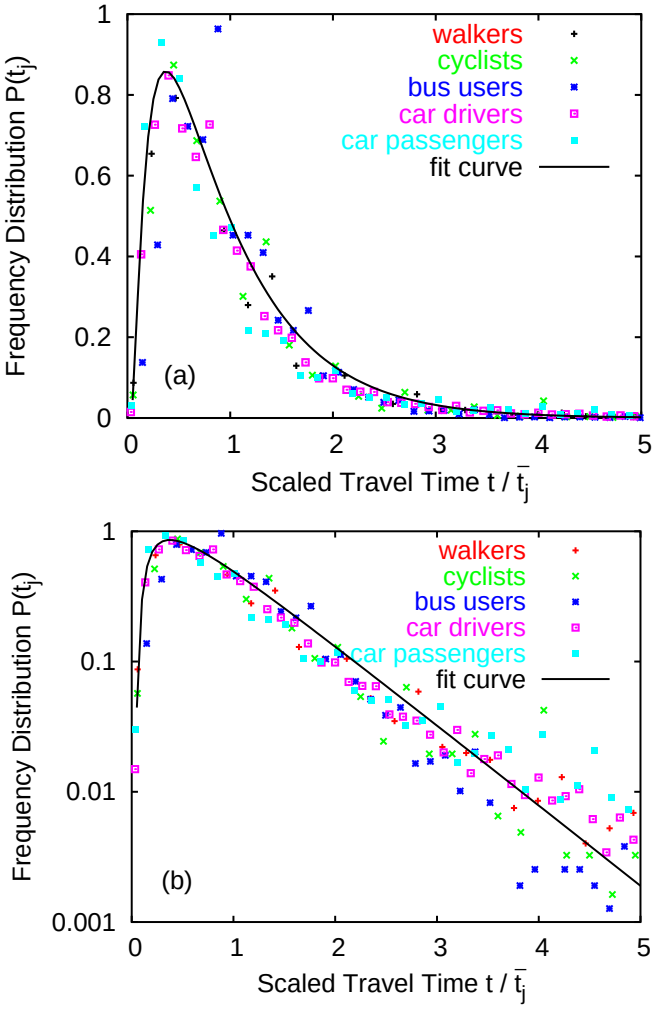

FIG. 9: Scaled time-averaged travel-time distributions for different modes of transport (a) in linear and (b) in semilogarithmic representation. Within the statistical variation (and rounding errors for small frequencies, which are magnified in the semilogarithmic plot), the mode-specific data could all be fitted by one universal curve (33) with $\alpha = 0.2$, $\gamma = 0.7$, and normalization constant $N(\alpha, \gamma) = 2.5$. The few railway data were not significant because of their large scattering. (From [59].)

where $E_j = p_j t$ with $p_j = \bar{E}/(\gamma \overline{t}_j)$ stands for the energy consumed by the physical activity of travelling with transport mode i for a time period t. The term $P_j(t_j) \propto \exp(-\alpha/t_j)$ possibly represents the additional amount of energy required for the mobilization of short trips, i.e. the Simonson effect [55]. With an average travel energy budget of $\bar{E} = 615$ kJ, this interpretation is consistent with values of the energy consumption p_j per unit time for certain activities obtained by ergonomic measurements of the related O_2-consumption [56–58].

Based on entropy maximization, the canonical distribution can be shown to be the most likely distribution, given that the average energy consumption $\bar{E}$ per day by

an ensemble of travellers is fixed for the area of investigation. This agrees well with the investigated data and generalizes the concept of a constant travel-time budget [60–67]. In addition, the above theory can be combined with trip distribution models [68–73], e.g. the multinomial logit model

$$P_j^* = \frac{\mathrm{e}^{\beta' V_j}}{\sum_{j'} \mathrm{e}^{\beta' V_{j'}}}\,, \tag{36}$$

where β' is a parameter (see Sec. VII B) and the mode-dependent energy consumption enters the utility function V_j as a (negative) cost term, which was shown to be a significant variable of human travel behaviour.

In summary, when daily travel time distributions by different modes of transport such as walking, cycling, bus or car travel are appropriately scaled, they turn out to have a universal functional relationship. A closer investigation reveals a law of constant average energy consumption by the physical activity of daily travelling. The advantage of the behavioral law (33) is its expected long-term validity over more than 25 years even under changing conditions. It will, therefore, prove to be important for urban, transport, and production planning: Compared to previous models, it facilitates improved quantitative conclusions about trip distributions, modal splits and induced traffic after the more reliable determination of fewer parameters, which are constant over typical planning horizons. It also helps to assess the increase in the acceptability of public transport, when the comfort of travel is improved, to forecast the change in travel behavior in an aging society, to predict the usage of new modes of transport, and to estimate potential market penetrations of new travel-related products.

VII. ACTIVITY PATTERNS

Sometimes, the models outlined in the previous sections are not sufficient to model demand generation for traffic. The first obvious shortfall is that they do not say anything about a time-of-day dynamics. This is sometimes introduced by assuming that all trips from residential to commercial locations represent morning traffic, and the afternoon/evening peak is then obtained by transposing the trip matrix.

Often, even that is not enough. This happens typically when there are significant correlations between travelers' attributes and their decisions. For example, the number of trips starting in a zone in general does not just depend on the number of residents, but also on their income or age; destination choice in general depends on vehicle availablity; the acceptance of a new mode of transport in general depends on the question if people can reorganize their entire day plans so that the new facility can be efficiently integrated.

A possible way out of this is to make the entire approach completely microscopic, or agent-based. This means that throughout the whole simulation, all travelers are maintained as individuals, and individual behavioral models are applied at all steps of the decision-making. For example, synthetic (= simulated) travelers may plan their activity patterns (e.g. home – work – shop – home – leisure – home), then the corresponding locations, then the modes of transportation, and finally the precise timing of their activities. The advantage of this approach over the ones discussed so far is that, at each step of the decision-making, all "internal" information of the agent is available.

Such an approach draws from many areas of science. Clearly, psychology is involved because it describes the behavior of human individuals. Yet, since the goal is not to describe each individual traveler correctly but to get correct aggregate results, economics and the social sciences are involved as well, and there is in fact a community of "travel behavior research" [74] which applies these principles to traffic. However, the approach to go to microscopic principles when an analysis on the macroscopic level fails is also deeply rooted in physics, and in fact, both the computing methods to large-scale problems and the general intuition that simple (and therefore individually wrong) microscopic models may lead to correct macroscopic results come to a large extent from physics.

Given its highly interdiscplinary nature plus the fact that only recently data and computing methods became available to do activity-based demand generation for large urban areas, it is understandable that the area is in rapid flux. There is one method, discrete choice theory [75], that is arguably better established than others, and it will be described in a bit more detail in the following. However, there are some drawbacks to that theory, and it is questionable whether it will be possible to correct them within the framework of that theory. For that reason, alternative, less well established approaches will also be discussed.

A. Discrete Choice Theory

Assume, for simplicity, a situation where an agent has a choice between several alternatives, j, for example between several locations for the same activity. Also assume that there are utilities U_j associated with those alternatives. Discrete choice theory starts from two assumptions:

- It is possible for the observer/modeler to measure/predict some part of the utilities. These measurable parts are called "systematic utilities" V_j.
- The differences between the V_j and the U_j are random and statistically independent. These differences are denoted as ε_j.

With these assumptions, one obtains essentially the following algorithm:

1. Compute the V_j based on your observations. For example, one might have observed that one loca-

tion is accessible only via some steep stairs, and therefore avoided by physically less fit persons.

2. Add the random components: $U_j = V_j + \varepsilon_j$.

3. Select the alternative that *now* has the highest utility. Note that, because of the random component, it can well be that the alternative with a smaller V_j in the end "wins".

There are at least two different interpretations where the random parts ε_j come from: Economic theory states that economic agents *always* select the option with the largest utility, and the difference between the observed/predicted utilities V_j comes from so-called unobserved attributes of the agent which are thus unknown to the observer/analysis. Psychology, in contrast, might say that there is just a random component to people's behavior.

B. Multinomial Logit Model (MNL)

The question now is how to obtain real-world numbers for the V_j and the ε_j. The typical way to progress is in two steps: (i) make some assumptions about the ε_j; (ii) assume that the V_j are additive in their contributions and do a maximum likelihood estimation of the parameters.

a. The random components. A conceptually easy way to proceed is to assume that the ε_j are normally distributed. The result of this is called a *(multinomial) probit model*. Its solution, however, cannot be written down in terms of simple functions. Another path is to assume that the ε_j follow so-called Gumbel-distributions with zero mean, i.e. their generating function is $F(\epsilon_j) = \exp[-\exp[-\mu_j\,\epsilon_j]]$, where μ_j is a width parameter. In this case, the result is that the probability for agent i to select option j follows the Boltzmann distribution

$$p_{i,j} = \frac{e^{\beta'\,V_{i,j}}}{\sum_k e^{\beta'\,V_{i,k}}}\ . \tag{37}$$

This is known as the *(multinomial) logit model*, cf. Eq. (36). β' results from the "widths" μ_j of the distributions. It models how "rational" the agents behave with respect to what is observed: a small β' makes the agents nearly random with respect to V_j; for $\beta' \to \infty$ one obtains an agent that always takes the option j with the best V_j.

b. The systematic components. In order to make further progress, one has to make an assumption about the functional form of the V_j, and its free parameters. The most typical assumption is that the V_j are linear in their components. For example, the utility for agent i to go swimming at location j might be

$$V_{i,j} = \beta_T\,T_{i,j} + \beta_C\,C_{i,j} + \beta_G\,G_i \tag{38}$$

where $T_{i,j}$ is the travel time to get there, $C_{i,j}$ is the monetary cost, and G_i is i's gender (e.g. 0 for male, and 1 for female). Normally, β_T and β_C will be negative. A positive β_G would express that, all other things being equal, females like to go swimming more than males. Note that attributes of the alternative as well as attributes of the person can occur, and that attributes can be continuous or discrete.

One now proceeds by asking a large number of people about their individual $T_{i,j}$, $C_{i,j}$, G_i, plus the decision that they actually made (revealed preference) or would make (stated preference) in the given situation. Maximum likelihood estimation then finds coefficients β_X such that the probability to obtain the decisions from the survey is maximized (X stands for the different possible indices). More precisely, it finds products $\beta'\,\beta_X$, meaning that the β' from Eq. (37) is already included. Note that it is *not* necessary for Eq. (38) to be linear; something like $\beta_{TT}\,T_{i,j}^2$ or $\beta_{CI}\,C_{i,j}/I_i$ is possible (where I_i is i's income). In the past, the utility function $V_{i,j}$ had to be linear in the β_X (because of the need for a partial derivative for the maximum likelihood estimation), but that requirement is now lifted [76].

C. Extensions

Multinomial logit models (MNLs) have a property called *independence from irrelevant alternatives* (IIA). It means that taking an alternative out of the choice set does not change the relative probabilities: $p_{i,j}/p_{i,k}$ does not depend on a third alternative m. This has sometimes odd consequences, typically explained in terms of the so-called "red bus blue bus" paradoxon. Assume that a traveler has the option to take a car (option 1) or to take the bus, and that she takes the bus with 50% probability. Now assume that the bus route is served by two buses, one of them red (option 2) and one of them blue (option 3), and both of them depart at exactly the same time, they have exactly the same service quality, and they are always nearly empty. Estimation of an MNL would result in a 25% probability for the red bus and a 25% probability for the blue bus.

So far, this is all fine. However, now assume that the blue bus gets taken out of service. The model for our traveler now predicts that she uses the car with probability 2/3, and the red bus with probability 1/3. This is due to the fact that the ratio of the probabilities of option 1 and option 2 has to remain unchanged.

The problem sits in the fact that for this example, the random components ε_2 and ε_3 are not independent. One way to deal with this problem is to make the model hierarchical ("nested"), where first the choice is between car and bus, and then about the type of bus. This reduces computational complexity, but demands that the analyst specifies the hierarchical structure. Another option is to use models in which correlations between the random variables are included as covariances. The results are more sophisticated than for the MNL model, but with increasingly better computers they enter the domain of

being useful for practical cases.

Another extension concerns the parameters β_X, where X again stands for the different possible indices. For example, in Eq. (38), it would make sense to assume that the ratio β_T/β_C, related to the "value of time", depended on income. The two options in this situation are (i) to assume an explicit income-dependent model such as $V_{ij} = \beta_T T_{i,j} + \beta_{CI}\, C_{i,j}/I_i ...$ (where once more β_T and β_{CI} would be expected to be negative) or (ii) to assume random taste variations across the population, which means that the coefficients β_X become random variables, with accordingly more free parameters to be estimated.

Software for the estimation of the models described in this section and pointers to relevant literature can be found at [76].

D. Discrete Choice Theory for Activities

Existing discrete choice models for the generation of daily activities use the nested MNL approach, which translates into a hierarchical decision tree: First a traveler decides on the daily pattern, then on activity locations, then on modes of transportation, and then on exact times [77, 78]. A problem is that decisions on the higher level depend on decisions on the lower levels; for example, the decision about the inclusion of an activity depends on how close its location is to the locations of other possible activities. The practical approach to that problem is to first compute the lower level options and then pass them on as "expected" utilities to the higher levels. For example, the utility of doing a trip both by car and by bus is computed, and (roughly) the maximum of both is passed on to the location choice module.

E. Alternative Methods for the Generation of Activity Patterns

In spite of the random term, discrete choice theory assumes extremely rational agents. In fact, a traveler modeled by a discrete choice model for activity planning will consider the utilities of *all* different options before making a decision. A consequence of this is that the computation is in fact *much more expensive than finding the best option*, since for a best option, suboptimal branches in the search tree can be pruned, while for discrete choice theory, all branches need to be followed to the leaves. In addition, discrete choice models for human activity planning need of the order of several hundred free parameters, and it is questionable how well so many parameters can be estimated from surveys.

For those reasons, it makes sense to look at alternative models. For example, one can make models where travelers do not look forward on the time axis [79]. Such a model is much easier to calibrate, but it is not sensitive to changes in the time structure: For example, the model will not make people get up later in the morning if opening times of shops are extended in the evening. An improvement of this is the use of Q-learning, where agents learn, by doing the same day over and over again, to back-propagate the temporal effects to the decision points [80]. Again, other models are entirely based on rules [81] or on genetic algorithms [82]. A newer development is the use of mental maps, where the agent remembers which parts of the system it knows [83–85].

VIII. LEARNING, FEEDBACK, AND EVOLUTIONARY GAME THEORY

Neither entropy-based destination choice (Sec. VI) nor complete activity patterns (Sec. VII) nor route choice (Sec. IX) can be sensibly modeled without including the effect of learning or adaptation. The simple reason for this is that there are nonlinear circular dependencies: Human plans depend on (expected) congestion, but congestion is the result of (the execution of) plans.

The obvious solution for this is any kind of relaxation method, i.e. some version of "make initial plans – compute congestion – adjust the plans – re-compute congestion" etc. Another interpretation of the same approach is that it models human learning from one day to the next.

It is interesting to note that there is actually some theory available to describe this process. For example, if all agents always play "best reply" (i.e. the best answer to the last iteration), and if the iterative dynamics converges to a fixed point attractor, than that fixed point is also a *Nash Equilibrium* (NE) [86]. For the traditional method of transportation planning, called static assignment (similar to flows in electric networks except that particles know their destinations), one can actually show that under some assumptions the NE is unique (in the link flows), and therefore the iterative process does not have to be explicitly modeled [87]. For the more complicated models discussed in this paper, no such mathematical statement is available, and therefore multiple fixed point attractors could be possible, or the iterative process could converge to a periodic or chaotic attractor. If the models are stochastic, then under mild assumptions they converge to a fixed steady-state state-space density [88]. In practice, only a small number of iterations is feasible, and effects such as broken ergodicity [89] need to be taken into account. A very illustrative example of these different types of possible learning dynamics will be provided in Sec. IX.

Real world scenarios are even more complicated. Decisions of different agents may happen on different time scales, which depending on their hierarchical order leads to different solutions [90, 91]. This is related to the issues of sequential games in game theory. Agents may have limited information, or they may not even use "best reply" at all, but rather some *constraint satisficing method,* in which they improve until they are satisfied. Technically, this means that some important real world results may lie in the transients rather than in the steady state be-

havior. A research agenda for the future needs to clarify these questions, and needs to identify the limits of predictability for such systems.

IX. ROUTE CHOICE

Once the activity pattern of an individual is selected (including the destination j and the mode of transport k [51–53]), one of the alternative routes m to the destination needs to be chosen. This is typically done by maximizing the utility U_k, which in most cases comes down to minimizing the overall travel time. As a consequence, the fastest connection is filled with vehicles, until the traffic density has increased so much that their travel times become comparable with the travel time on an alternative road with greater length or smaller speed limit. In this way, a distribution of traffic is produced, in which the travel times of all alternative routes are the same. This distribution is called the *Wardrop equilibrium* [92]. However, does this distribution really describe the route choice behavior of drivers? The following paragraphs will show that the Wardrop equilibrium systematically underestimates travel times, as drivers do not match the ideal traffic distribution due to coordination problems.

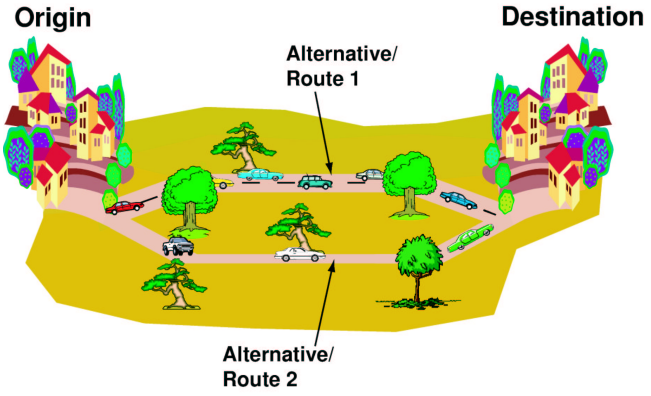

FIG. 10: Schematic illustration of the day-to-day route choice scenario (from [44, 93, 94]). Each day, the drivers have to decide between two alternative routes, 1 and 2. Note that, due to the different number of lanes, route 1 has a higher capacity than route 2. The latter is, therefore, used by less cars.

A. Decision Experiments

The efficient distribution of road capacities among users based on individual decisions is still a fundamental problem. As individuals normally have aggregate information (such as radio news) rather than complete information, one frequently observes far-from-optimal travel times. In order to learn more about this decision behavior, we have carried out a variety of route choice experiments [93, 94] (see Figs. 10 to 13).

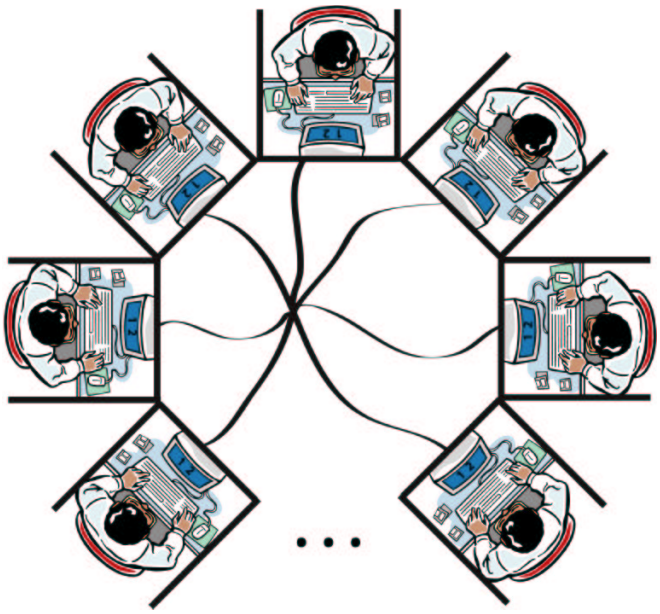

FIG. 11: Schematic illustration of the decision experiment (from [93, 94]). Several test persons take decisions based on the aggregate information their computer displays. The computers are connected and can, therefore, exchange information. However, a direct communication among players is suppressed.

In these experiments, N test persons had to repeatedly decide between two alternative routes $m \in \{1, 2\}$ and should try to maximize their resulting payoffs P_m, which were chosen proportionally to the inverse travel times. If the average vehicle speed V_m on route m is approximated by the linear relationship [95]

$$V_m(N_m) = V_m^0 \left(1 - \frac{N_m(t)}{N_m^{\max}}\right) , \qquad (39)$$

the inverse travel times obey the payoff relations $P_m(N_m) = P_m^0 - P_m^1 N_m$ with

$$P_m^0 = \frac{V_m^0}{L_m} \quad \text{and} \quad P_m^1 = \frac{V_m^0}{N_m^{\max} L_m} . \qquad (40)$$

Herein, V_m^0 is the maximum velocity (speed limit) and N_m the number of drivers on route m, L_m its length, and $N_m^{\max}$ its capacity, i.e. the maximum possible number of vehicles on route m. For an improved approach to determine the travel times in road networks see Ref. [96]. Note that alternative routes can reach comparable payoffs and travel times only when the total number $N = N_1 + N_2$ of vehicles is large enough to fulfil the relations $P_1(N) < P_2(0) = P_2^0$ and $P_2(N) < P_1(0) = P_1^0$. Our route choice experiment has focussed on this traffic regime. Furthermore, we have the capacity restriction

$N < N_1^{\max} + N_2^{\max}$, as $N = N_1^{\max} + N_2^{\max}$ is connected with a complete gridlock.

The Wardrop or user equilibrium corresponds to equal travel times and payoffs for both alternative decisions. It is found for a fraction

$$f_1^{\rm eq} = \frac{N_1}{N} = \frac{P_2^1}{P_1^1 + P_2^1} + \frac{1}{N}\frac{(P_1^0 - P_2^0)}{(P_1^1 + P_2^1)} \qquad (41)$$

of persons choosing alternative 1 and agrees with the system optimum only in the limit $N \gg 1$ of many participants. Small groups were chosen to investigate the fluctuations in the system. It was found that the test groups managed to adapt to the user equilibrium on average (see Fig. 12a). However, although it appears reasonable to stick to the same decision once the equilibrium is reached, the standard deviation stayed at a finite level. This was not only observed in *"treatment" 1*, where all players knew only their own (previously experienced) payoff, but also in *treatment 2*, where the payoffs $P_1'(N_1)$ and $P_2'(N_2)$ for both, 1- *and* 2-decisions, were transmitted to all players (analogous to radio news). Moreover, treatment 2 could increase the average payoffs, as the additional information allowed for a better adaptation without having to change decisions to explore the better performing alternative (see Ref. [97] and Fig. 12).

With *treatment 3*, we managed to reveal the mysterious persistence in the changing behavior and to achieve more than three times higher payoff increases. Every test person was informed about the own payoff $P_1'(N_1)$ [or $P_2'(N_2)$] and the *potential* payoff $P_2'(N - N_1 + \epsilon N) = P_2'(N_2) - \epsilon N P_2^1$ [or $P_1'(N - N_2 + \epsilon N) = P_1'(N_1) - \epsilon N P_1^1$] he or she would have obtained, if a fraction ϵ of persons had additionally chosen the other alternative (here: $\epsilon = 1/N = 1/9$). Consequently, in the user equilibrium with $P_1'(N_1) = P_2'(N_2)$ every player knew that he or she would *not* get *the same*, but a *reduced* payoff, if he or she would change the decision. That explains why treatment 3 could reach a better adaptation performance, reflected by a low standard deviation and close-to-optimal average payoffs. Moreover, even the smallest individual cumulative payoff exceeded the highest one in treatment 1. Therefore, treatment 3's way of information presentation is much superior to the ones used today.

Figure 13 clears up why players changed their decision in the user equilibrium at all. We discovered intermittent behavior, i.e. quiescent periods without changes, followed by turbulent periods with many changes. This is reminiscent of volatility clustering in stock market indices, where individuals also react to aggregate information summarizing all decisions (the trading transactions). Single individuals seem to change their decision speculating for above-average payoffs. In fact, although the cumulative individual payoff is anticorrelated with the average changing rate, some individuals receive higher payoffs with larger changing rates than others, i.e. they profit from the overreaction in the system: Once the system is out of equilibrium, all individuals respond in one way or another. Typically, there are too many decision

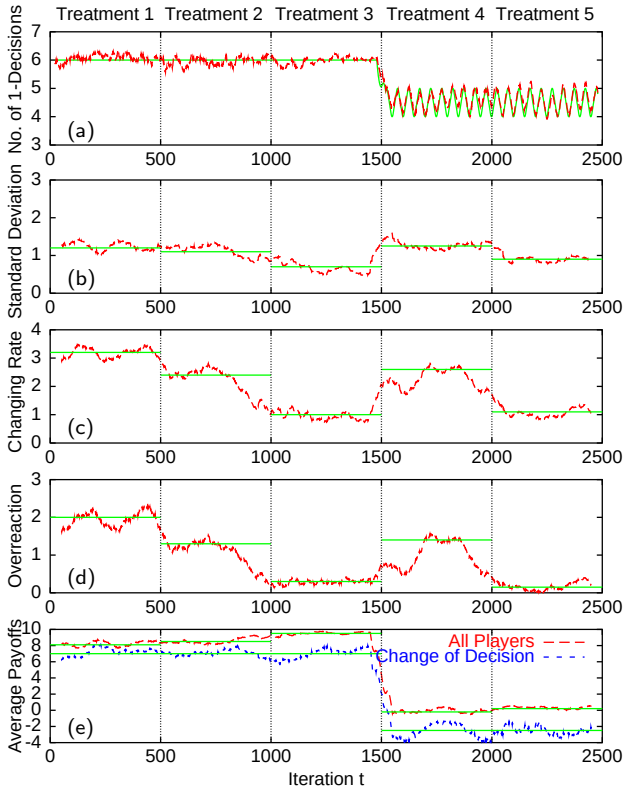

FIG. 12: Overview of treatments 1 to 5 (with $N = 9$ and payoff parameters $P_2^0 = 28$, $P_1^1 = 4$, $P_2^1 = 6$, and $P_1^0 = 34$ for $0 \le t \le 1500$, but a zick-zack-like variation between $P_1^0 = 44$ and $P_1^0 = -6$ with a period of 50 for $1501 \le t \le 2500$): (a) Average number of decisions for alternative 1 (fluctuating curve) compared to the user equilibrium (smooth curve), (b) standard deviation of the number of 1-decisions from the user equilibrium, (c) number of decision changes from one iteration to the next one, (d) overreaction, i.e., difference between the actual number of decision changes (changing rate) and the required one (standard deviation), (e) average payoff per iteration for players who have changed their decision and for all players. The latter increased with a reduction in the changing rate, but normally stayed below the payoff in the user equilibrium (which is 1 on average in treatments 4 and 5, otherwise 10). This payoff loss is caused by the overreaction in the system. The displayed moving time-averages [(a) over 40 iterations, (b-e) over 100 iterations] illustrate the systematic response to changes in the treatment every 500 iterations. Smooth lines in (b)–(e) are estimates of the stationary values after the transient period, while time periods around the dotted lines are not significant. Compared to treatment 1, treatment 3 managed to reduce the changing rate and to increase the average payoffs (even more than treatment 2 did). These changes were systematic for *all* players. In treatment 4, the changing rate and the standard deviation went up, since the user equilibrium changed in time. The user-specific recommendations in treatment 5 could almost fully compensate for this and managed to reach the minimum overreaction in the system. The above conclusions are also supported by additional experiments with single treatments. (After [93, 94].)

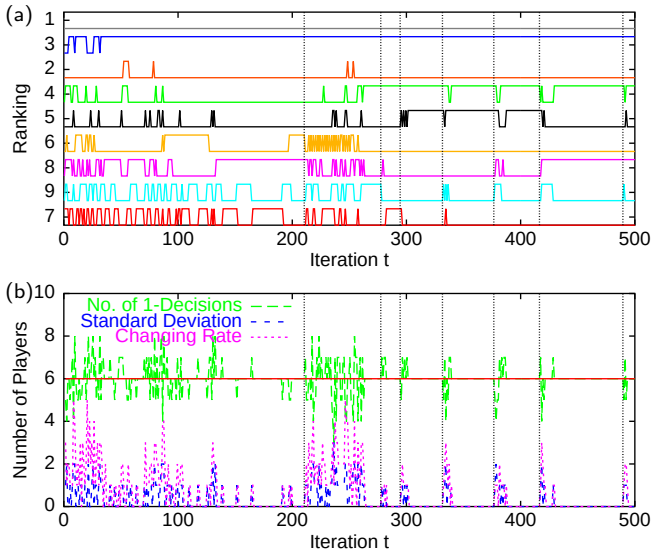

FIG. 13: Illustration of typical results for treatment 3 (from [93, 94]). (a) Decisions of all 9 players. Players are displayed in the order of increasing changing rate. Although the ranking of the cumulative payoff and the changing rate are anticorrelated, the relation is not monotonic. Note that turbulent or volatile periods characterized by many decision changes are usually triggered by individual changes after quiet periods (dotted vertical lines). (b) The changing rate is often larger than the (standard) deviation from the user equilibrium $N_1 = f_1^{\rm eq} N = 6$ (red horizontal line), indicating an overreaction in the system (see also Fig. 12d).

changes.

In *treatment 4*, we have tested the group performance under changing environmental conditions, when the participants got the same information as in treatment 3. That is, some of the payoff parameters were varied in time, implying a time-dependent user equilibrium. Even without recommendations, the group managed to adapt surprisingly well to the variable conditions, but the standard deviation and changing rate were approximately as high as in treatment 2 (see Fig. 12). This adaptability (the collective 'group intelligence') is based on complementary responses [94, 97]. That is, if some players do not react to the changing conditions, others will take the chance to earn additional payoff. This experimentally supports the behavior assumed in the theory of efficient markets, but here the efficiency is limited by overreaction.

To avoid overreaction, in *treatment 5* we have recommended a number $f_1^{\rm eq}(t+1)N - N_1(t)$ of players to change their decision and the other ones to keep it. These user-specific recommendations helped the players to reach the smallest overreaction of all treatments and a very low standard deviation, although the payoffs were changing in time as in treatment 4 (see Fig. 12).

X. MIGRATION AND THE LAW OF GRAVITY

Urban and regional evolution, in particular the change of street infrastructures, housing, industrial and business quarters, has a natural feedback on the spatial distribution of people. The temporal development of the population distribution $P(\boldsymbol{x},t)$ in space $\boldsymbol{x}$ can be described by migration models. Pioneers in this field have been Weidlich and Haag [98]. In the following, we will discuss a variant of their master equation model [99–101]

$$\frac{dP(\boldsymbol{x},t)}{dt} = \sum_{\boldsymbol{x}'} \Big[\underbrace{w(\boldsymbol{x}|\boldsymbol{x}',t)P(\boldsymbol{x}',t)}_{\text{Immigration into } \boldsymbol{x}} - \underbrace{w(\boldsymbol{x}'|\boldsymbol{x},t)P(\boldsymbol{x},t)}_{\text{Emigration out of } \boldsymbol{x}} \Big] . \tag{42}$$

According to this equation, the stream of immigrants into place $\boldsymbol{x}$ is given by the fraction $P(\boldsymbol{x}',t)$ of the population at place $\boldsymbol{x}'$ which could potentially move to place $\boldsymbol{x}$. The proportionality factor is the transition rate $w(\boldsymbol{x}|\boldsymbol{x}',t)$ from $\boldsymbol{x}'$ to $\boldsymbol{x}$, i.e. the migration probability to $\boldsymbol{x}$ per unit time, given one lives at place $\boldsymbol{x}'$ before. Analogously, the stream of emigrants leaving place $\boldsymbol{x}$ is proportional to the fraction $P(\boldsymbol{x},t)$ of the population living in $\boldsymbol{x}$, and the proportionality factor is the migration rate $w(\boldsymbol{x}'|\boldsymbol{x},t)$ from $\boldsymbol{x}$ to $\boldsymbol{x}'$. This migration rate is proportional to the fraction of people living at place $\boldsymbol{x}'$, as the number of acquaintances with friends, family members, or former collegues increases linearly with the number of inhabitants of a town $\boldsymbol{x}$. However, the migration rate decreases significantly with the distance $y = \|\boldsymbol{x} - \boldsymbol{x}'\|$ between two places $\boldsymbol{x}$ and $\boldsymbol{x}'$ (and/or with the transaction costs). Accordingly, we assume [102]

$$w(\boldsymbol{x}'|\boldsymbol{x},t)P(\boldsymbol{x},t) = \nu(t) \mathrm{e}^{U(\boldsymbol{x}',t)}\, \mathrm{e}^{-U(\boldsymbol{x},t)} \frac{P(\boldsymbol{x}',t)P(\boldsymbol{x},t)}{D(\|\boldsymbol{x}-\boldsymbol{x}'\|)} , \tag{43}$$

where $D(y)$ is a monotonously growing function in y. The parameter $\nu(t)$ reflects the mobility of the population and $U(\boldsymbol{x},t)$ the utility/attractiveness of living at place $\boldsymbol{x}$. Note that formula (43) reminds of the gravity law in physics [103], where $P(\boldsymbol{x}',t)$ and $P(\boldsymbol{x},t)$ have the meaning of the masses of two celestial bodies $\boldsymbol{x}$ and $\boldsymbol{x}'$ and $D(y) = y^2$. Moreover, Eq. (43) is analogous to Eq. (32) with $b_{\boldsymbol{x}\boldsymbol{x}'} = P(\boldsymbol{x}',t)P(\boldsymbol{x},t)/D(\|\boldsymbol{x}-\boldsymbol{x}'\|)$, $g_{\boldsymbol{x}}(t) = \mathrm{e}^{-U(\boldsymbol{x},t)}$, and $h_{\boldsymbol{x}'}(t) = \mathrm{e}^{U(\boldsymbol{x}',t)}$. That is, the gravity law can also be motivated by entropy considerations, but this time $P(\boldsymbol{x},t)$ is adapted rather than $g_{\boldsymbol{x}}(t)$ and $h_{\boldsymbol{x}'}(t)$. A great advantage of Eq. (42) is that it can describe disequilibrium situations such as resulting migration streams when borders between countries are opened.

XI. URBAN AND REGIONAL EVOLUTION

Weidlich and Haag have modeled urban and regional evolution with a master equation approach similar to Eq. (42), but with additional birth- and death-terms describing the generation or the removal of entities from

the system. Moreover, in their simulations, they distinguished different kinds of entities ("agents"): inhabitants, who may migrate, streets which may be expanded, industrial areas which may grow or decay, building blocks in which people may live, and leisure areas such as parks. Positive and negative feedbacks between these different entities were represented by interaction terms in the utility functions determining the corresponding transition rates w. In this way, they could describe the empirically observed differentiation in spatial usage patterns and interdependencies between them. Moreover, they have applied their approach to the evolution of cities in China [104].

An advantage of models based on the master equation is their huge flexibility. A disadvantage is that there is a large number of coefficients that need to be defined. Therefore, there is also a search for models that naturally include (usually geometric) constraints and by doing so can explain some aspects of urban dynamics with a much smaller number of free parameters. These models originate from the observation that some properties of cities are fractal. For example, the fractal dimension of the bulk of cities seems to be around 1.9, and the fractal dimension of the (outer) perimeter of cities seems to be around 1.3 [105, 106]. This implies that one might attempt to describe city growth with local growth models. There are both attempts to explain the fractal dimension by variations of models originally from physics [107, 108] and attempts to re-create (and thus predict) the growth of real-world cities [109, 110]. Yet, not all aspects of urban growth can be explained by entirely local rules. For example, construction of a large mall or a golf course certainly depends on the availability of a similar facility *somewhere* in the city. Unfortunately, this leads to models and simulations where everything depends on everything, which means complicated models and long simulation times. Yet, plausibly, this is also not correct: Although large and important facilities have a far-reaching impact, the effect still decays with distance. In consequence, hierarchical models, as they are known from particle dynamics with long range interactions, are now also transfered to urban dynamics [111].

As was the case at other places in this paper, the higher level dynamics of urban growth is not independent from lower level dynamics such as traffic congestion. In consequence, the ultimate challenge of this area may be to present a uniform view that bridges all scales from car driving via human activity planning to urban evolution. In such a view, the models presented so far would be limiting cases for the situation where a separation of (time) scales is possible. There are three approaches to achieve this bridging of the scales:

- *Parameterization:* If a separation of (time) scales is not possible, then lower level processes need to be taken into account at least via parameterizations, i.e. simplified aggregated models of the lower level processes. As is well known, parameterization needs to be preceeded by scientific understanding.
- *Model coupling:* Instead of using a parameterization, i.e. a macroscopic description of the lower level dynamics, a (computational) model of the lower level dynamics could be used directly. Such model coupling is another big challenge, not only for computational reasons, but also because the combined model systems are once more dynamical systems with a behavior that is usually not well understood.
- *Microscopic across all scales:* A further alternative is to construct a model that is microscopic or *agent-based* across all scales, that is, at all levels all decisions are made by synthetic versions of the respective actors, usually humans. This is similar to a molecular dynamics simulation in physics. It also faces the same challenge: It is difficult to simulate large scale systems in spite of all the computational progress we had in recent times [112]. Nevertheless, the first partial systems of this approach are emerging [113–115], and validations of partial aspects look rather promising [116, 117].

Once such models are available and operational, they could be applied to a rather wide range of questions, ranging from zoning via infrastructure planning to the deliberate attraction of commercial companies and the effects of migration. And although such models cannot take away the responsibility for decisions from politicians or from society, they might help that such decisions are based on a more informed understanding of the system than today.

XII. SUMMARY AND OUTLOOK

Physics is a science that has a tradition in deriving large scale emergent phenomena from microscopic rules. Sometimes, these derivations can be done analytically, and in such cases the thermodynamic limit is useful. In other situations, the microscopic rules are too complicated to make analytical progress, and the computer is used. However, even with today's computational power, only up to about 10^9 particles can be simulated directly.

The situation in traffic and regional development is similar in the sense that most effects are also caused by the combined behavior of many particles. In contrast to materials science, the number of particles in regional simulations, typically several millions, is well within the range that is computationally accessible. On the other hand, the rules that each individual particle follows are considerably more complicated.

In this situation, it makes sense to look at different scenarios one by one. The first question is to identify those areas where very simple models already lead to good levels of understanding and prediction. This refers, for example, to the area of traffic flow, where already models which just include limits on acceleration, braking, and excluded volume generate a wide range of dynamical phenomena that can all be identified in the real world.

Similar models are tried in the areas of route choice, activity generation, migration, or urban growth.

If these simple models fail, then more complicated models are in order. These models typically contain more behavioral parameters, and because the behavior of humans is difficult to predict, these models are more difficult to calibrate or validate. The challenge is, as so often, to keep the models as simple as possible, but no simpler. Also, one has to keep in mind that we are interested in macroscopic, "emergent" phenomena such as traffic jams, or general urban patterns, and those should be easier to predict than the behavior of individual humans. Nevertheless, one has to get used to validation error bars that are much larger than in physics.

With respect to real-world prediction, models that cover a large range of scales would be useful. Most notably, there is a rather old dream for integrated land use transport models, in which the location choices of inhabitants and commercial entities lead to traffic, which in turn triggers revised location choices, etc. [118] It has, however, been difficult to put that dream into practice [118], and one could argue that the main obstacle has been that the interactions between the scales were not well understood but rather included in an ad-hoc manner. Simultaneously, computing power in the past has not really been sufficient for sound solutions.

However, in our view the situation is now changing. Based on the combination of computational, analytical, and experimental work, there has been considerable progress with respect to understanding aspects at the different levels, leading to better and faster submodels in the future. At the same time, it is now possible to build computational models that bridge large ranges of scales while remaining microscopic, which, although computationally demanding, again solves a large variety of conceptual problems.

Finally, one should also mention that data availability in the field of geo-science has made a quantum jump forward. Satellite-derived high resolution spatial data is now available for nearly everywhere; decent census data is available in nearly every industrialized country; vector data for the infrastructure is increasingly becoming available; and a whole range of telematics devices is capable of collecting important data about spatial behavior, although aspects of privacy need to be valued against scientific interest. This also means that the traditional approach of physics, where theory, computation, and real-world measurement enhance each other, can now be applied to traffic and regional systems.

Overall, it is our impression that the new computing and data collection technologies have had a huge impact in the area of traffic and regional systems, and that both the computational and the theoretical work are struggling to keep up with the new possibilities. This, combined with the obvious societal relevance of the area, makes this a very exciting field.

XIII. AUTHORS' BIOGRAPHIES

Dirk Helbing (*19/Jan/1965) is the Managing Director of the Institute for Economics and Traffic at Dresden University of Technology. Originally, he studied Physics and Mathematics in Göttingen, Germany, but soon he got fascinated in interdisciplinary problems. Therefore, his master thesis dealt with physical models of pedestrian dynamics, while his Ph.D. thesis at the University of Stuttgart focussed on modeling interactive decisions and behaviors with methods from statistical physics and the theory of complex systems. After his habilitation on the physics of traffic flows, he received a Heisenberg scholarship and worked at Xerox PARC in Silicon Valley, the Weizmann Institute in Israel, and the Collegium Budapest—Institute for Advanced Study in Hungary. He gets excited when physics meets traffic or social science, economics, or biology, and if the results are potentially relevant for everyday life.

Kai Nagel (*17/Sep/1965) is Professor for Transport Systems Analysis and Transport Telematics at the Institute of Land and Sea Transport Systems at the Technical University Berlin in Germany. He studied physics and meteorology at the University of Cologne and the University of Paris, with one master's thesis in the area of cellular automata models for cloud formation and another one in the area of large scale climate simulations. His Ph.D., in computer science at the University of Cologne, was about cellular automata models for large scale traffic simulations. He then was postdoc, staff member, and research team leader at Los Alamos National Laboratory, working on the TRANSIMS (Transportation ANalysis and SIMulation System) project. In 1999–2004, he was assistant professor for computer science at ETH Zurich in Switzerland. His interests are in large scale simulation and in the simulation and modeling of socio-economic systems.

[1] Helbing, D., 2001, Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, **73**(4), 1067–1141.

[2] Helbing, D., 2004, Traffic. In *Encyklopedia of Nonlinear Science*, (London: Taylor & Francis).

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

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

[5] Kerner, B. S., Konhäuser, P., and Schilke, M., 1995, Deterministic spontaneous appearance of traffic jams in slightly inhomogeneous traffic flow. *Physical Review E*, **51**, R6243–R6246.

[6] Kerner, B. S., Konhäuser, P., and Schilke, M., 1996,

'Dipole-layer' effect in dense traffic flow. *Physics Letters A*, **215**, 45–56.

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

[8] Krauß, S., 1998, *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics* (Cologne: DLR—Deutsches Zentrum für Luft- und Raumfahrt e.V.).

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

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

[11] Helbing, D., Hennecke, A., Shvetsov, V., and Treiber, M., 2001, Macroscopic traffic simulation based on a gas-kinetic, non-local traffic model. *Transportation Research B*, **35**, 183–211.

[12] Helbing, D., Hennecke, A., Shvetsov, V., and Treiber, M., 2002, Micro- and macrosimulation of freeway traffic. *Mathematical and Computer Modelling*, **35**, 517–547.

[13] Prigogine, I., and Andrews, F. C., 1960, A Boltzmann-like approach for traffic flow. *Opns. Res.*, **8**, 789–797.

[14] Prigogine, I., 1961, A Boltzmann-like approach to the statistical theory of traffic flow. In *Theory of Traffic Flow*, edited by R. Herman (Amsterdam: Elsevier), pp. 158–164.

[15] Prigogine I., and Herman, R., 1971, *Kinetic Theory of Vehicular Traffic* (New York: Elsevier).

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

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

[18] Payne, H. J., 1971, Models of freeway traffic and control. In *Mathematical Models of Public Systems*, edited by G. A. Bekey (La Jolla, CA: Simulation Council), Vol. 1, pp. 51–61.

[19] Payne, H. J., 1979, A critical review of a macroscopic freeway model. In *Research Directions in Computer Control of Urban Traffic Systems*, edited by W. S. Levine, E. Lieberman, and J. J. Fearnsides (New York: American Society of Civil Engineers), pp. 251–265.

[20] Kerner, B. S., and Konhäuser, P., 1993, Cluster effect in initially homogeneous traffic flow. *Phys. Rev. E*, **48**, R2335–R2338.

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

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

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

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

[25] Helbing, D., *et al.*, 2003, Critical discussion of "synchronized flow", pedestrian evacuation, and optimal production. In *Traffic and Granular Flow '01*, edited by M. Fukui, *et al.*, (Berlin: Springer), pp. 511–530.

[26] Schönhof, M., and Helbing, D., 2004, Empirical features of congested traffic states and their implications for traffic modeling. In preparation for *Transportation Science*.

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

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

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

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

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

[32] Kerner, B. S., 2000, Phase transitions in traffic flow. In *Traffic and Granular Flow '99*, edited by D. Helbing, H. J. Herrmann, M. Schreckenberg, and D. E. Wolf (Berlin: Springer), pp. 253–284.

[33] Kerner, B. S., 2000, Theory of breakdown phenomenon at highway bottlenecks. *Transpn. Res. Rec.*, **1710**, 136–144.

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

[35] Nishinari, K., Treiber, M., and Helbing, D., 2003, Interpreting the wide scattering of synchronized traffic data by time gap statistics. *Physical Review E*, **68**, 067101.

[36] Gazis, D. C., Herman, R., and Rothery, R. W., 1961, Nonlinear follow the leader models of traffic flow. *Opns. Res.*, **9**, 545–567.

[37] Bando, M., Hasebe, K., Nakayama, A., Shibata, A., and Sugiyama, Y., 1994, Structure stability of congestion in traffic dynamics. *Jpn. J. Industr. Appl. Math.*, **11**, 203–223.

[38] Bando, M., Hasebe, K., Nakayama, A., Shibata, A., and Sugiyama, Y., 1995, Dynamical model of traffic congestion and numerical simulation. *Phys. Rev. E*, **51**, 1035–1042.

[39] Mehta, K., 1991, *Random Matrices* (New York: Academic Press).

[40] Helbing, D., and Treiber, M., 2003, Fokker-Planck equation approach to vehicle statistics. Preprint http://arXiv.org/abs/cond-mat/0307219.

[41] Krbalek, M. and Helbing, D., 2004, Determination of interaction potentials in freeway traffic from steady-state statistics. *Physica A*, **333**, 370–378.

[42] Neubert, L., Santen, L., Schadschneider, A., Schreckenberg, M., 1999, *Physical Review E*, **60**, 6480.

[43] May, Jr., A. D., and Keller, H. E. M., 1967, Non-integer car-following models. *Highway Res. Rec.*, **199**, 19–32.

[44] Helbing, D., and Nagel, K., 2003, Verkehrsdynamik und urbane Systeme. *Physik Journal*, **2**(5), 35–41.

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

[46] Hoogendoorn, S. P., and Bovy, P. H. L., 1998, *Motorway Traffic Flow Analysis*, edited by P. H. L. Bovy (Delft: University of Technology), pp. 71.

[47] Gazis, D. C., 1974, Modeling and optimal control of congested transportation systems. *Networks*, **4**, 113–124.

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

[49] Cetin, N., and Nagel, K., 2003, A large-scale agent-based traffic microsimulation based on queue model. *Proceedings of Swiss Transport Research Conference (STRC)* (Monte Verita, CH).

[50] Nagel, K., and Raney, B., 2004, Complex systems applications for transportation planning. In *The Real and Virtual Worlds of Spatial Planning*, edited by M. Koll-Schretzenmayr, M. Keiner, and G. Nussbaumer (Berlin: Springer-Verlag), chapter 16.

[51] Schnabel, W., and Lohse, D., 1997, *Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung*, Band 2 (Berlin: Verlag für Bauwesen).

[52] Lohse, D., Teichert, H., Dugge, B., and Bachner, G., 1997, Ermittlung von Verkehrsströmen mit n-linearen Gleichungssystemen – Verkehrsnachfragemodellierung. In *Schriftenreihe des Instituts für Verkehrsplanung und Straßenverkehr, TU Dresden*, Heft 5/1997.

[53] Lohse, D., and Schneider, R., 1997, Vergleichende Untersuchungen der aggregierten und disaggregierten Verkehrsplanungsmodelle. In *Schriftenreihe des Instituts für Verkehrsplanung und Straßenverkehr, TU Dresden*, Heft 3/1997.

[54] DETR/DTLR Data sn2852, sn2853, sn2854, sn2855, sn3288, sn4108, 1998, 2000, (Essex: The Data Archive).

[55] Hettinger, T., 1989, Physiologische Leistungsgrundlagen. In *Handbuch der Ergonomie*, 2nd ed., edited by H. Schmidke (Munich: Hanser), Vol. 2.

[56] Spitzer, H., Hettinger, T., and Kaminsky, G., 1982, *Tafeln für den Energieumsatz bei körperlicher Arbeit*, 6th ed. (Berlin: Beuth).

[57] Goodwin, P. B., 1976, Human effort and the value of travel time. *Journal of Transport Economics and Policy*, **10**, 3–15.

[58] Rowland, T. W., 1998, The biological basis of physical activity. *Medicine & Science in Sports & Exercise*, 392-399.

[59] Kölbl, R., and Helbing, D., 2003, Energy laws in human travel behaviour. *New Journal of Physics*, **5**, 48.1.–48.12.

[60] Zahavi, Y., and Talvitie, A., 1980, Regularities in travel time and money expenditure. *Transpn. Res. Rec.*, **750**, 13–19.

[61] Zahavi, Y., and Ryan, J., 1980, Stability of travel components over time. *Transpn. Res. Rec.*, **750**, 19–26.

[62] Roth, G. J., and Zahavi, Y., 1981, Travel time "budgets" in developing countries. *Transpn. Res. A*, **15**, 87–95.

[63] Supernak, J., and Zahavi, Y., 1982, Travel-time budget: A critique (discussion). *Transpn. Res. Rec.*, **28**(879), 15–28.

[64] Tanner, J. C., 1981, Expenditure of time and money on travel. *Transpn. Res. A*, **15**, 25–38.

[65] Schafer, A., 1998, The global demand for motorized mobility. *Transpn. Res. A*, **32**, 455–477.

[66] Schafer, A., 2000, Regularities in travel demand: An international perspective. *Journal of Transportation and Statistics*, **3**(3), 1–31.

[67] Goodwin, P. B., 1981, The usefulness of travel budgets. *Transpn. Res. A*, **15**, 97–106.

[68] Domencich, T. A., and McFadden, D., 1975, *Urban Travel Demand. A Behavioral Analysis*, pp. 61–69; Ortúzar, J. d. D., and Willumsen, L. G., 1994, *Modelling Transport*, 2nd ed. (New York: Wiley).

[69] Wilson, A. G., 1967, A statistical theory of spatial distribution modes. *Transpn. Res. A*, **1**, 253–269.

[70] Wilson, A. G., 1970, *Entropy in urban and regional modelling* (London: Pion).

[71] Wilson, A. G., 1998, Land-use/Transport Interaction Models. Past and Future. *Journal of Transport Economics and Policy*, **32**, 3–26.

[72] Ben-Akiva, M., and Lerman, S. R., 1997, *Discrete Choice Analysis: Theory and Application to Travel Demand* (Cambridge, MA: MIT Press).

[73] Ben-Akiva, M., McFadden, D. M. *et al.*, 1999, Extended framework for modeling choice behavior. *Marketing Lett.*, **10**, 187–203.

[74] Hensher, D., and King, J., 2001. In *The Leading Edge of Travel Behavior Research*, edited by D. Hensher, and J. King (Oxford: Pergamon).

[75] Ben-Akiva, M., and Lerman, S. R., 1985, *Discrete choice analysis* (Cambridge, MA: The MIT Press).

[76] BIOGEME www page. www.roso.epfl/biogeme, accessed 2004.

[77] Bowman, J. L., 1998, The day activity schedule approach to travel demand analysis, PhD thesis (Cambridge, MA: Massachusetts Institute of Technology).

[78] Bowman, J. L., Bradley, M., Shiftan, Y., Lawton, T. K., and Ben-Akiva, M., 1999, *Demonstration of an activity-based model for Portland*, Vol. 3 (Oxforf: Elsevier).

[79] Kitamura, R., Fujii, S., Kikuchi, A., and Yamamoto, T., 1997, An application of a micro-simulator of daily travel and dynamic network flow to evaluate the effectiveness of selected tdm measures for CO_2 emissions reduction.

[80] Charypar, D., Graf, P., and Nagel, K., 2004, Q-learning for flexible learning of daily activity plans. *Proceedings of Swiss Transport Research Conference (STRC)* (Monte Verita, CH), see: www.strc.ch.

[81] Arentze, T., Hofman, F., van Mourik, H., and Timmermans, H., 2000, ALBATROSS: A multi-agent rule-based model of activity pattern decisions, Paper 22, Transportation Research Board Annual Meeting, Washington, D.C..

[82] Charypar, D., and Nagel, K., 2003, Generating complete all-day activity plans with genetic algorithms. *Proceedings of the meeting of the International Association for Travel Behavior Research (IATBR)* (Lucerne, Switzerland), see http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html.

[83] Unger, H., 2002, *Modellierung des Verhaltens autonomer Verkehrsteilnehmer in einer variablen staedtischen Umgebung*, PhD thesis, TU Berlin, 2002.

[84] Arentze, T., and Timmermans, H. J. P., 2003, Representing mental maps and cognitive learning in micro-simulation models of activity-travel choice dynamics. *Proceedings of the meeting of the International Association for Travel Behavior Research (IATBR)* (Lucerne, Switzerland), see http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html.

[85] Kistler, D., 2004, Mental maps for mobility simulations of agents, Master's thesis, ETH Zurich.

[86] Hofbauer, J., and Sigmund, K., 1998, *Evolutionary games and replicator dynamics* (Cambridge: University Press).

[87] Sheffi, Y., 1985, *Urban transportation networks: Equilibrium analysis with mathematical programming methods* (Englewood Cliffs: Prentice-Hall).

[88] Cascetta, E., and Cantarella, C., 1991, A day-to-day and within day dynamic stochastic assignment model. *Transportation Research A*, **25**, 277–291.

[89] Palmer, R., 1989, Broken ergodicity. In *Lectures in the Sciences of Complexity*, edited by D. L. Stein (Redwood City, CA: Addison-Wesley), Vol. I of *Santa Fe Institute Studies in the Sciences of Complexity*, pp. 275–300.

[90] van Zuylen, H. J., and Taale, H., in press. In *Traffic and granular flow '03*, edited by P. Bovy et al.

[91] Nagel, K., Strauss, M., and Shubik, M., in press, *Physica A*.

[92] Wardrop, J. G., 1952, Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers II*, Vol. 1, pp. 325–378.

[93] Helbing, D., Schönhof, M., and Kern, D., 2002, Volatile decision dynamics: Experiments, stochastic description, intermittency control, and traffic optimization. *New Journal of Physics*, **4**, 33.1.–33.16.

[94] Helbing, D., 2004, Dynamic decision behavior and optimal guidance through information services: Models and experiments. In *Human Behaviour and Traffic Networks*, edited by M. Schreckenberg, and R. Selten (Berlin: Springer).

[95] Greenshields, B. D., 1935, A study of traffic capacity. *Proceedings of the Highway Research Board*, (Washington, D. C.: Highway Research Board), Vol. 14, pp. 448–477.

[96] Helbing, D., 2003, A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks. *J. Phys. A: Math. Gen.*, **36**, L593–598.

[97] Schreckenberg, M., Selten, R., Chamura, T., Pitz, T., and Wahle, J., 2001, Experiments on day-to-day route choice. www.trafficforum.org, e-print 01080701.

[98] Weidlich, W., and Haag, G., 1988, *Interregional Migration* (Berlin: Springer).

[99] Weidlich, W., and Haag, G., 1983, *Concepts and Models of a Quantitative Sociology. The Dynamics of Interacting Populations* (Berlin: Springer).

[100] Weidlich, W., 1991, Physics and social science—The approach of synergetics. *Physics Reports*, **204**, 1–163.

[101] Weidlich, W., 2000, *Sociodynamics. A Systematic Approach to Mathematical Modelling in the Social Sciences* (Amsterdam: Harwood Academic).

[102] Helbing, D., *Quantitative Sociodynamics. Stochastic Methods and Models of Social Interaction Processes* (Dordrecht: Kluwer Academic).

[103] Zipf, G. K., 1946, The $P1P2/D$ hypothesis on the intercity movement of persons. *Amer. Sociol. Rev.*, **11**, 677–686.

[104] Weidlich, W., and Haag, G., 1999, *An Integrated Model of Transport and Urban Evolution* (Berlin: Springer).

[105] Frankhauser, P., 1994, *La fractalité des structrures [sic] urbaines* (Paris: Anthropos).

[106] Batty, M., and Longley, P., 1994, *Fractal Cities: A Geometry of Form and Function* (Academic Press).

[107] Schweitzer, F., 1997, *Self-organization of complex structures: From individual to collective dynamics* (London: Gordon and Breach).

[108] Makse, H. A., Andrade, Jr., J. S., Batty, M., Havlin, S., and Stanley, H. E., 1998, Modeling urban growth patterns with correlated percolation. *Physical Review E*, **58**, 7054–7062.

[109] White, R., Engelen, G., and Uljee, I., 1997, The use of constrained cellular automata for high-resolution modeling of urban landscape dynamics. *Environment and Planning B*, **24**, 323–343.

[110] Rabino, G., and Laghi, A., 2002, Identification of cellular automata: theoretical remarks. *Proceedings of the Annual Conference of The European Regional Science Association (ERSA)* (Dortmund), see: www.ersa.org.

[111] Andersson, C., Lindgren, K., Rasmussen, S., and White, R., 2002, Urban growth simulation from 'first principles'. *Physical Review E*, **66**, 026204.

[112] Nagel, K., and Marchal, F., 2003, Computational methods for multi-agent simulations of travel behavior. *Proceedings of the meeting of the International Association for Travel Behavior Research (IATBR)* (Lucerne), see http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html.

[113] MATSIM www page. MultiAgent Transportation SIMulation. see www.matsim.org, accessed 2004.

[114] Salvini, P. A., and Miller, E. J., 2003, ILUTE: An operational prototype of a comprehensive microsimulation model of urban systems. *Proceedings of the meeting of the International Association for Travel Behavior Research (IATBR)* (Lucerne), see http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html.

[115] TRANSIMS www page. TRansportation ANalysis and SIMulation System. transims.tsasa.lanl.gov, accessed 2004, Los Alamos National Laboratory, Los Alamos, NM.

[116] Esser, J., and Nagel, K., 2001, Iterative demand generation for transportation simulations. In [74], pp. 689–709.

[117] Raney, B., Cetin, N., Völlmy, A., Vrtic, M., Axhausen, K., and Nagel, K., 2003, An agent-based microsimulation model of Swiss travel: First results. *Networks and Spatial Economics*, **3**, 23–41.

[118] Timmermans, H. J. P., 2003, The saga of integrated land use-transport modeling: How many more dreams before we wake up? *Proceedings of the meeting of the International Association for Travel Behavior Research (IATBR)* (Lucerne), see http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html.