Phase Diagram of Traffic States in the Presence of Inhomogeneities

Dirk Helbing[1,2], Ansgar Hennecke[1], and Martin Treiber[1]

[1]*II. Institute of Theoretical Physics, University of Stuttgart, Pfaffenwaldring 57/III, 70550 Stuttgart, Germany*

[2]*Department of Biological Physics, Eötvös University, Budapest, Puskin u 5–7, H–1088 Hungary*

Abstract

We present a phase diagram of the different kinds of congested traffic that are triggered by disturbances when passing ramps or other spatial inhomogeneities of a freeway. The simulation results obtained by the nonlocal, gas-kinetic-based traffic model are in good agreement with empirical findings. They allow to understand the observed transitions between free and various kinds of congested traffic, among them localized clusters, stop-and-go waves, and different types of "synchronized" traffic. We also give analytical conditions for the existence of these states which suggest that the phase diagram is universal for a class of different microscopic and macroscopic traffic models. 05.70.Fh,47.55.-t,51.30.+i,89.40.+k

For some years now, the various observed nonlinear states of freeway traffic flow, including traffic jams and stop-and-go waves, have attracted the interest of a rapidly growing community of physicists [1,2]. The same applies to simulation approaches that allow to reproduce the transitions between the different collective states of motion [2–5]. A particular attention received the recent discovery of a first-order phase transition to "synchronized" congested traffic [6,7], which stimulated an intense investigation of the complex phenomena associated with ramps, yielding an explanation of the hysteretic phase transition [8,5]. It is also known from other effectively one-dimensional many-body systems with short-range interactions that localized inhomogeneities can cause a variety of phenomena, including phase transitions [9] and spontaneous symmetry breaking [10].

To characterize the parameter-dependence of the possible states of a system resulting in the long run, phase diagrams are a very powerful method. They are of great importance in thermodynamics with various applications in metallurgy, chemistry, etc. Moreover, they allow to compare very different kinds of systems like equilibrium and nonequilibrium ones, or microscopic and macroscopic ones, whose equivalence cannot simply be shown by transformation to normal forms [11,12]. Defining universality classes by mathematically equivalent phase diagrams, one can even classify so different systems as physical, chemical, biological and social ones, as done in systems theory.

In the following, we will present and explain the phase diagram that we obtained for the nonlocal, gas-kinetic-based traffic model [5], when we studied a freeway with a ramp in the presence of a single perturbation [Fig. 1]. By systematical variation of the inflow at the upstream freeway boundary *and* the ramp, we were able to reproduce the observed transitions between different traffic states, some of which have not been explained before. We will also present analytical conditions for the existence of the various traffic states. Since we are interested in the generic properties of the model, we will formulate it in dimensionless quantities [13] by measuring times in units of the relaxation time $\tau \approx 40\,\mathrm{s}$, and distances in units of the inverse $1/\rho_{\mathrm{max}}$ of the maximum vehicle density $\rho_{\mathrm{max}} \approx 140\,\mathrm{vehicles/km}$. The equation for the vehicle density $\rho(x,t)$ at position x and time t is a continuity equation with

a sink/source term and reads

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho V)}{\partial x} = \frac{Q_{\mathrm{rmp}}(x,t)}{L_{\mathrm{rmp}}} \, .$$ (1)

V denotes the average vehicle velocity, L_{rmp} is the length of the ramp, Q_{rmp} is the flow of vehicles entering the freeway ($Q_{\mathrm{rmp}} > 0$) or leaving it ($Q_{\mathrm{rmp}} < 0$) along the ramp of length L_{rmp}, divided by the number n of freeway lanes. The velocity equation contains a convection term (due to the movement of the velocity profile with velocity V), a pressure term (reflecting dispersion effects due to a finite velocity variance θ), a relaxation term (describing an adaptation to a desired velocity V_0), and an interaction term (corresponding to braking maneuvers) [13]:

$$\frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} = -\frac{1}{\rho} \frac{\partial(\rho\theta)}{\partial x} + (V_0 - V)$$
$$- PA(\rho_{\mathrm{a}}) \frac{(\rho_{\mathrm{a}} V_{\mathrm{a}})^2}{(1 - \rho_{\mathrm{a}})^2} B(\delta_V) \, .$$ (2)

V_0 has a meaning similar to the Reynolds number (since it allows for traffic instabilities when it exceeds the value $V_{\mathrm{c}} = 61$). In addition, P is a scaled cross section.

$$A(\rho) = 0.171 + 0.417\{\tanh[10(\rho - 0.27)] + 1\}$$ (3)

represents a structure factor, which is of order unity and determines the variance via the constitutive relation $\theta = A(\rho)V^2$, but also the form of the flow-density relation in equilibrium [13] (see Fig. 2). Finally,

$$B(\delta_V) = 2\left[\delta_V \frac{\mathrm{e}^{-\delta_V^2/2}}{\sqrt{2\pi}} + (1 + \delta_V^2) \int_{-\infty}^{\delta_V} dy \, \frac{\mathrm{e}^{-y^2/2}}{\sqrt{2\pi}}\right]$$ (4)

with the dimensionless velocity difference $\delta_V = (V - V_{\mathrm{a}})/\sqrt{\theta + \theta_{\mathrm{a}}}$ is a Boltzmann factor arising from the vehicle interactions. An index "a" indicates that the respective quantity is evaluated at the advanced "interaction point" $x_{\mathrm{a}} = x + \gamma(1 + TV)$ rather than at the actual position x. The related nonlocality reflects the anticipative driver behavior and is essential for the realistic properties of the model and its robust and efficient numerical solution. γ is an anticipation factor and T about the safe time headway. In our simulations, we used the

parameter values $V_0 = 171$, $P = 0.31$, $\gamma = 1.2$, and $T = 0.043$, which are typical for Dutch freeways [13], but our results are not very sensitive to their particular choice.

The nonlocal, gas-kinetic-based traffic model reproduces the characteristic properties of traffic flows [13] formulated by Kerner and Rehborn [14]. Furthermore, for a homogeneous freeway *without* ramps, it shows the phase ("instability") diagram [13] that was postulated by Kerner and Konhäuser on the basis of macroscopic simulations [4]: For a given density ρ, there exists a state of homogeneous traffic with equilibrium velocity $V_e(\rho)$ and equilibrium flow $Q_e(\rho) = \rho V_e(\rho)$ (see Fig. 2), and there are four critical densities ρ_{ci}. For densities $\rho < \rho_{c1}$ and $\rho > \rho_{c4}$, homogeneous traffic is stable with respect to localized perturbations, and for a range $\rho_{c2} < \rho < \rho_{c3}$ of intermediate densities, it is linearly unstable, giving rise to cascades of traffic jams ("stop-and-go traffic"). For the two density regimes $\rho_{c1} \leq \rho \leq \rho_{c2}$ and $\rho_{c3} \leq \rho \leq \rho_{c4}$ between the stable and the linearly unstable regions, it is metastable, i.e., it behaves nonlinearly unstable with respect to perturbations exceeding a certain critical amplitude, but otherwise stable. For the discussion of the phase diagram, an additional range $\rho_{cv} \leq \rho \leq \rho_{c3}$ of convective stability will be important, where homogeneous traffic is linearly unstable, but the growing perturbations are convected away from any fixed location [11].

We will denote by $Q_{ci} = Q_e(\rho_{ci})$ the equilibrium flows corresponding to the critical densities introduced above. Moreover, there exists a characteristic outflow Q_{out} from traffic jams, stop-and-go waves, etc. [13,14], that is nearly independent of the surrounding density and the type of congested traffic, at least if there is no ramp. If traffic relaxes at the location of an on-ramp, the observed outflow is $\tilde{Q}_{out}(Q_{rmp}, L_{rmp}) \leq Q_{out}$, but we find $\tilde{Q}_{out} \to Q_{out}$ for $L_{rmp} \to \infty$ or $Q_{rmp} \to 0$ [5]. For our model parameters, we obtain $\rho_{c1} = 0.11$, $\rho_{c2} = 0.12$, $\rho_{cv} = 0.34$, $\rho_{c3} = 0.36$, and $\rho_{c4} = 0.40$, related to $Q_{c1} = 16.2$, $Q_{c2} = 17.5$, $Q_{cv} = 14.6$, $Q_{c3} = 14.1$, and $Q_{c4} = 13.2$. If $L \geq 50$, we obtain $\tilde{Q}_{out} \approx Q_{c2}$, otherwise $\tilde{Q}_{out}$ decreases with growing Q_{rmp} and decreasing L_{rmp}. For these values (which, with the exception of $\tilde{Q}_{out}$, are determined from situations *without* ramps!), the equations for the phase boundaries suggested below agree quantitatively with the ones in the phase diagram in Fig. 1.

In our simulations, we investigated the congested traffic states that formed near an on-ramp when passed by a perturbation on the freeway (Fig. 3). For the perturbation we chose a fully developed density cluster that did not change its amplitude or shape anymore and travelled upstream with constant speed. When the density cluster reached the ramp, it induced different kinds of congested states (Figs. 1 through 3), depending on the inflows Q_{in} and Q_{rmp} at the upstream boundary of the freeway and the ramp, respectively.

Now, we will explain the different states in the phase diagram starting with relatively high ramp flows. When traffic breaks down (a sufficient criterium for this is given by $(Q_{\text{out}} + Q_{\text{rmp}}) > Q_{\text{max}}$, where Q_{max} denotes the maximum of the equilibrium flow), we find the formation of a growing, extended region of congested traffic with average flow $Q_{\text{cong}} = (\tilde{Q}_{\text{out}} - Q_{\text{rmp}})$ in front of the downstream end of the on-ramp [5]. Because of the fixed downstream end, any perturbation will be convected out of the congested region, if the homogeneous solution of density ρ_{cong} and flow $Q_{\text{cong}} = Q_{\text{e}}(\rho_{\text{cong}})$ is convectively stable, i.e. if $Q_{\text{cong}} \leq Q_{\text{cv}}$. In this case, we have homogeneous congested traffic (HCT) [Fig. 3(a)], which corresponds to an equilibrium solution on the high-density branch of the flow-density diagram [Fig. 2(a)]. Otherwise, oscillatory congested traffic (OCT) forms [Figs. 3(b) and 2(b)]. Thus, the boundary between HCT and OCT is given by

$$\text{HCT-OCT:} \quad Q_{\text{rmp}} = \tilde{Q}_{\text{out}} - Q_{\text{cv}} \,. \tag{5}$$

The oscillation amplitudes grow with decreasing ramp flow, until they reach the low-density part of the flow-density diagram associated with free traffic [Fig. 2(c)]. This means that we have an alternation between free and congested traffic which defines the state of triggered stop-and-go waves (TSG) [Fig. 3(c)]. It turns out, that TSG is characterized by $Q_{\text{TSG}} \geq Q_{\text{TSG}}^{\text{min}}$, where Q_{TSG} denotes the average flow of the TSG state. Hence, the boundary between TSG and OCT is given by

$$\text{OCT-TSG:} \quad Q_{\text{rmp}} = \tilde{Q}_{\text{out}} - Q_{\text{TSG}}^{\text{min}} \,. \tag{6}$$

However, the precise value of $Q_{\text{TSG}}^{\text{min}}$ is hard to determine.

The minimal inflows to sustain the TSG states can be derived from the triggering mechanism illustrated in Fig. 3(c). When passing the ramp, the initial localized perturbation causes a secondary perturbation travelling downstream [15]. Since the amplitude of this secondary perturbation is always small, regardless of the amplitude of the primary perturbation, it only grows on the instability condition $Q_{\text{down}} > Q_{\text{c2}}$, where $Q_{\text{down}} = (Q_{\text{in}} + Q_{\text{rmp}})$ denotes the flow downstream of the ramp. With growing amplitude, the triggered perturbation changes its propagation speed, reverses its direction, and finally induces another small perturbation when passing the ramp, etc. In this way, the interplay of the perturbations with the ramp defines an intrinsic timescale and wavelength of the triggered stop-and-go waves. If the above instability condition is not fulfilled, instead of TSG we find a single moving localized cluster (MLC) [Figs. 3(d) and 2(d)], similar to the transition between stop-and-go waves and localized clusters in homogeneous traffic [4]. The respective boundary is given by

$$\text{TSG-MLC:} \quad Q_{\text{rmp}} = Q_{\text{c2}} - Q_{\text{in}} \,. \tag{7}$$

The MLC continues the course of the initial perturbation which, however, changes its propagation speed at the ramp, probably because of $\tilde{Q}_{\text{out}} < Q_{\text{out}}$. If $Q_{\text{in}} < Q_{\text{c1}}$, the upstream flow cannot sustain localized clusters, since it is stable to perturbations, and the MLC becomes a pinned (standing) localized cluster (PLC [or SLC]) [15] at the on-ramp [Fig. 3(e)]:

$$\text{MLC-PLC:} \quad Q_{\text{in}} = Q_{\text{c1}} \,. \tag{8}$$

A pinned localized cluster will grow and, thereby, give rise to a an extended form of congested traffic (CT), if the sum $(Q_{\text{in}} + Q_{\text{rmp}})$ of the inflows exceeds the self-organized outflow $\tilde{Q}_{\text{out}}$ from CT behind ramps. The corresponding boundaries are given by

$$\text{PLC-CT:} \quad Q_{\text{rmp}} = \tilde{Q}_{\text{out}} - Q_{\text{in}} \,. \tag{9}$$

They are the same for HCT and OCT, so that there could be also stationary and oscillating variants of PLC (SPLC and OPLC).

Finally, no congested state can be formed, if $Q_{\text{down}} < Q_{\text{c1}}$, since traffic flow is stable to any perturbation, then. This defines the boundary between the pinned localized cluster state and free traffic (FT):

$$\text{PLC-FT:} \quad Q_{\text{rmp}} = Q_{\text{c1}} - Q_{\text{in}} \,. \tag{10}$$

Due to the metastability between ρ_{c1} and ρ_{c2}, the transition to free traffic depends on the amplitude of the perturbation. For small perturbation amplitudes, stable traffic flow associated with free traffic expands up to $Q_{\text{down}} < Q_{\text{c2}}$. Hence, the MLC and PLC regimes may disappear in cases of small disturbances, leading to the boundaries FT-TSG, FT-OCT, and FT-HCT. In conclusion, the transition between free and congested traffic is of first order (i.e. hysteretic) [5], while the other transitions seem to be continuous.

Notice that the triple points A(TSG-MLC-PLC), B(OCT-TSG-PLC), and C(HCT-OCT-PLC) must each satisfy the conditions of *two* boundaries. In particular, points A and C are determined uniquely by dynamic properties on a homogeneous road *without* ramps or other inhomogeneities: $Q_{\text{in}}^A = Q_{\text{c1}}$, $Q_{\text{rmp}}^A = Q_{\text{c2}} - Q_{\text{c1}}$, $Q_{\text{in}}^C = Q_{\text{cv}}$, and $Q_{\text{rmp}}^C = \tilde{Q}_{\text{out}} - Q_{\text{cv}} \approx Q_{\text{c2}} - Q_{\text{cv}}$.

Summarizing our results, we have presented a phase diagram of traffic states developing close to an on-ramp when passed by a perturbation on the freeway. Similar results are found for disturbances on the ramp (not displayed). Moreover, the same phase diagram is expected for other kinds of inhomogeneities as caused by gradients, changes in the number of lanes, etc. In such cases, Q_{rmp} corresponds to the capacity drop along the inhomogeneity. Despite of the complex behavior of the model, it was possible to determine the phase boundaries analytically and in good agreement with numerical investigations. As the analytical relations for the phase boundaries do not contain any details of the nonlocal, gas-kinetic traffic model used for the simulations, they are expected to be also valid for other microscopic and macroscopic traffic models which have (a) the same instability diagram or, more exactly, critical densities ρ_{c1}, ρ_{c2}, and ρ_{cv}, and (b) a characteristic outflow $\tilde{Q}_{\text{out}}(Q_{\text{rmp}}, L_{\text{rmp}})$ that is independent of the type of congestion at the ramp. A sensitive dependence of $\tilde{Q}_{\text{out}}$ on the

surrounding traffic situation could lead to different results.

There is some evidence [5,16] that the homogeneous congested state can be identified with the observed synchronized traffic of type (i) according to the classification by Kerner and Rehborn [6], whereas synchronized traffic of types (ii) and (iii) probably corresponds to the oscillating congested state. Stop-and-go waves and localized clusters have also their empirical counterparts [6,14]. Our findings suggest that inhomogeneities may be the generic reason for the formation of the different congested states [17]. The dependence of the forming states on the inflows Q_{in} and Q_{rmp} allows to understand the observed transitions between localized clusters, stop-and-go waves, free and synchronized flow [6,7,14,18]. It is also of importance for designing on-ramp controls. The dependence of the traffic states on the ramp length L_{rmp} (because of the dependence of $\tilde{Q}_{\mathrm{out}}$ on L_{rmp}) is relevant for an optimal dimensioning of ramps and explains why the observable traffic states are dependent on the observation site.

REFERENCES

[1] D. E. Wolf *et al.*, *Traffic and Granular Flow* (World Scientific, Singapore, 1996); M. Schreckenberg and D. E. Wolf (eds.) *Traffic and Granular Flow '97* (Springer, Singapore, 1998).

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

[3] K. Nagel and M. Schreckenberg, J. Phys. I France **2**, 2221 (1992).

[4] B. S. Kerner and P. Konhäuser, Phys. Rev. E **50**, 54 (1994).

[5] D. Helbing and M. Treiber, Phys. Rev. Lett. **81**, 3042 (1998).

[6] B. S. Kerner and H. Rehborn, Phys. Rev. E **53**, R4275 (1996).

[7] B. S. Kerner and H. Rehborn, Phys. Rev. Lett. **79**, 4030 (1997).

[8] H. Y. Lee *et al.*, Phys. Rev. Lett. **81**, 1130 (1998).

[9] J. Krug, Phys. Rev. Lett. **67**, 1882 (1991).

[10] M. R. Evans *et al.*, Phys. Rev. Lett. **74**, 208 (1995).

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

[12] J.-P. Eckmann *et al.*, Annales de l'Institut Henri Poincare Physique Theorique **58**, 287 (1993); Y. Kuramoto, Progress in Theor. Phys. Suppl. **99**, 244 (1989).

[13] M. Treiber *et al.*, Phys. Rev. E **59**, 239 (1999).

[14] B. S. Kerner and H. Rehborn, Phys. Rev. E **53**, R1297 (1996).

[15] B. S. Kerner *et al.*, Phys. Rev. E **51**, 6243 (1995).

[16] M. Treiber and D. Helbing, J. Phys. A: Math. Gen. **32**, L17 (1999).

[17] C. F. Daganzo *et al.*, Transpn. Res. B, in print (1999).
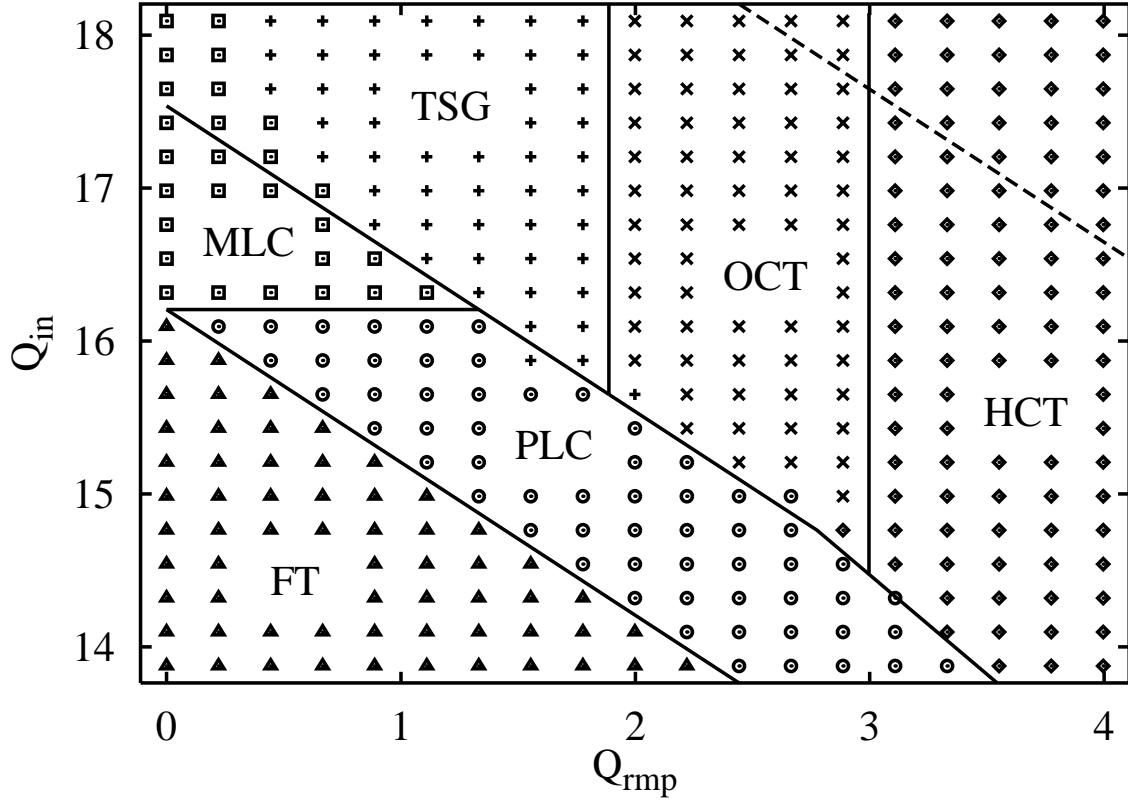
[18] D. Helbing, Phys. Rev. E **55**, R25 (1997).

FIGURES

FIG. 1.   Phase diagram of the traffic states in the vicinity of an on-ramp as a function of the inflows Q_{in} and Q_{rmp} on the main road and the on-ramp for fixed ramp length $L = 56$. The different states are classified at $t = 135$, i.e., after a sufficiently long transient period. Displayed are homogeneous congested traffic ($\diamond$), oscillatory congested traffic ($\times$), triggered stop-and-go traffic ($+$), moving localized clusters ($\square$), pinned localized clusters ($\odot$), and free traffic ($\triangle$). The states are triggered by a fully developed localized cluster travelling upstream and passing the ramp (cf. Fig. 3). Solid lines indicate the theoretical phase boundaries. The dashed line represents the condition $(Q_{\text{in}} + Q_{\text{rmp}}) = Q_{\text{max}}$ that characterizes the maximum downstream flow for which a (possibly unstable) equilibrium solution exists. Above this line, extended congestion develops even without any perturbation.
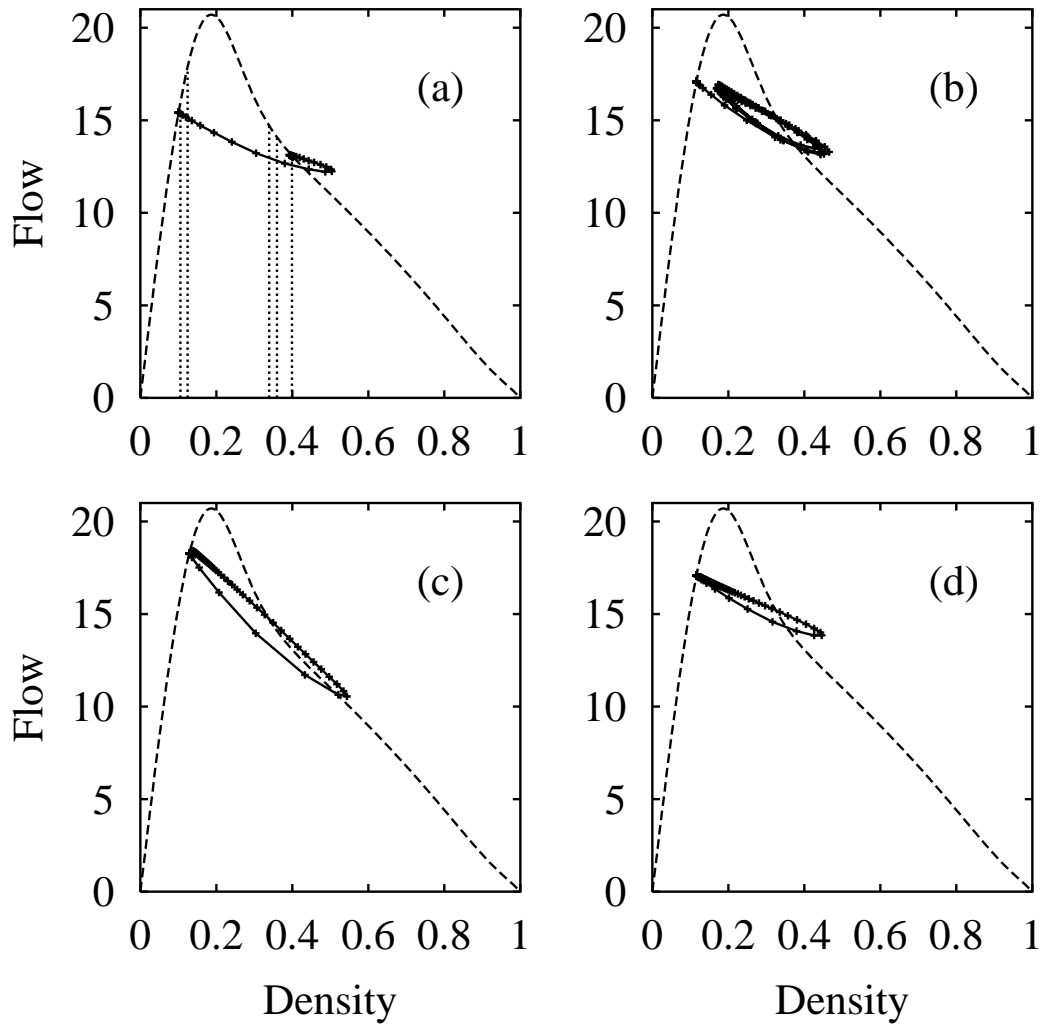
FIG. 2.   Dynamics in the flow-density space at the fixed position $x = -560$ upstream of the ramp. Dashed lines represent the flow-density diagram (i.e. the equilibrium flow-density relation), dotted lines the position of the critical densities ρ_{ci}. (a) Shown is the transition from free traffic (symbols at the left side of the flow-density diagram) to homogeneous congested traffic (symbols at the right side). The other illustrations display (b) oscillatory congested traffic, (c) triggered stop-and-go traffic, and (d) a moving localized cluster.
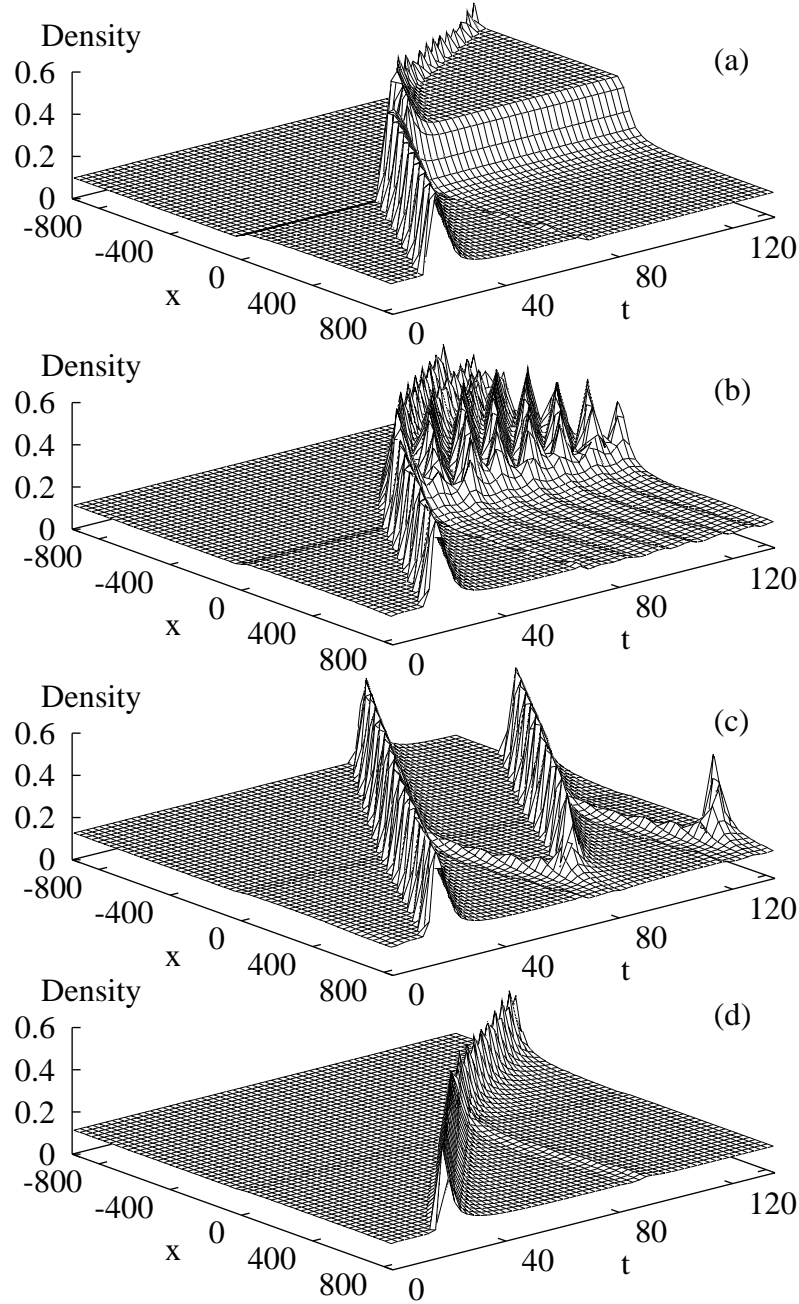
FIG. 3. Spatiotemporal dynamics of typical representatives of the states depicted in Fig. 1. The respective states are (a) homogeneous congested traffic (HCT), (b) oscillatory congested traffic (OCT), (c) triggered stop-and-go traffic (TSG), and (d) a moving localized cluster (MLC). $x = 0$ corresponds to the middle of the ramp.