Deterministic microscopic three-phase traffic flow models

2006 J. Phys. A: Math. Gen. 39 1775

(http://iopscience.iop.org/0305-4470/39/8/002)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 1775–1809 doi:10.1088/0305-4470/39/8/002

Deterministic microscopic three-phase traffic flow models

Boris S Kerner[1] **and Sergey L Klenov**[2]

[1] DaimlerChrysler AG, REI/VF, HPC: G021, 71059 Sindelfingen, Germany
[2] Department of Physics, Moscow Institute of Physics and Technology, 141700 Dolgoprudny, Moscow Region, Russia

Received 15 July 2005, in final form 28 November 2005
Published 8 February 2006
Online at stacks.iop.org/JPhysA/39/1775

Abstract

Two different deterministic microscopic traffic flow models, which are in the context of the Kerner's there-phase traffic theory, are introduced. In an acceleration time delay model (ATD model), different time delays in driver acceleration associated with driver behaviour in various local driving situations are explicitly incorporated into the model. Vehicle acceleration depends on local traffic situation, i.e., whether a driver is within the free flow or synchronized flow or else wide moving jam traffic phase. In a speed adaptation model (SA model), vehicle speed adaptation occurs in synchronized flow depending on driving conditions. It is found that the ATD and SA models show spatiotemporal congested traffic patterns that are adequate with empirical results. In the ATD and SA models, the onset of congestion in free flow at a freeway bottleneck is associated with a first-order phase transition from free flow to synchronized flow; moving jams emerge spontaneously in synchronized flow only. Differences between the ATD and SA models are studied. A comparison of the ATD and SA models with stochastic models in the context of three-phase traffic theory is made. A critical discussion of earlier traffic flow theories and models based on the fundamental diagram approach is presented.

PACS numbers: 89.40.+k, 47.54.+r, 64.60.Cn, 64.60.Lx

1. Introduction

Theoretical studies of freeway traffic flow dynamics are one of the rapid developing fields of statistical and nonlinear physics (see the reviews [1–6], the book [7] and the conference proceedings [8–15]). For a mathematical description of freeway traffic flow, a huge number of different microscopic and macroscopic traffic flow models have been introduced. In macroscopic models, individual dynamic vehicle behaviour is averaged, i.e., these models describe dynamics of average traffic flow characteristics such as average vehicle

speed and density (see, e.g., [16–20]).[3] Microscopic traffic flow models describe individual dynamic vehicle behaviour, which should simulate empirical spatiotemporal features of phase transitions and congested patterns in freeway traffic. In this paper, we restrict ourselves to a consideration of *microscopic* traffic flow models *only*.

There are two types of microscopic traffic flow models: deterministic models and stochastic models [1–7]. In deterministic models, some dynamic rules of vehicle motion in traffic flow are responsible for spatiotemporal features of traffic patterns that the models exhibit. Contrastingly, stochastic models, in addition to dynamic rules of vehicle motion, exhibit model fluctuations, which play a fundamental role for traffic pattern features.

There are at least two classes of deterministic traffic flow models [1, 3–6]. In the first class, driver time delays in vehicle acceleration (deceleration) a are explicitly taken into account. An example is the classic model of Herman, Montroll, Potts, and Rothery [23]: if the vehicle speed v or the speed difference between the vehicle speed and the speed of the preceding vehicle v_ℓ or else the net distance g (space gap) between vehicles changes, then the driver accelerates (decelerates) with a time delay τ [23]:

$$a(t+\tau) = f(v(t), v_\ell(t), g(t)). \tag{1}$$

Based on (1), Gazis, Herman and Rothery [24] have developed a microscopic traffic flow model, which is capable of describing traffic beyond of instabilities; steady-state solutions of this model lie on a one-dimensional curve in the flow–density plane (the fundamental diagram) (see the review by Nagel *et al* [6] for more details). Recall that steady-state solutions are hypothetical model solutions in which all vehicles move at the same time-independent speed and the same space gap between vehicles. One of the mathematical descriptions of this model class first proposed by Nagatani and Nakanishi [25] and further developed by Lubashevky *et al* [26] reads as follows:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{f(v(t), v_\ell(t), g(t)) - a(t)}{\tau}. \tag{2}$$

In both models [24–26], steady-state model solutions in the flow–density plane lie on the fundamental diagram.

There is also another class of deterministic microscopic models in which the vehicle speed satisfies the equation [1, 3–6]

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \phi(v(t), v_\ell(t), g(t)). \tag{3}$$

Examples are optimal velocity (OV) models of Newell [27], Whitham [28], Bando, Sugiyama *et al* [29], and the intelligent driver model (IDM) of Treiber and Helbing [30, 31]. Steady-state solutions of this model class that obviously satisfy the conditions $\phi(v, v_\ell, g) = 0$ and $v = v_\ell$ lie on the fundamental diagram in the flow–density plane.

If functions and model parameters in the models (2) and (3) are chosen in an appropriated way, then there is a range of vehicle density in which steady-state model solutions for free flow are unstable. This instability, which should explain the onset of congestion, leads to wide moving jam emergence in free flow (F $\rightarrow$ J transition) [2–6].

However, as explained in the book [7], the above models that are in the context of the fundamental diagram approach, as well as all other traffic flow models reviewed in [1–6] cannot explain the fundamental empirical feature of traffic breakdown, i.e., that the onset of congestion in free flow at a bottleneck is associated with a local first-order phase transition

[3] It should be noted that transferring the information delivered from one vehicle interacting with the neighbouring ones requires one to deal carefully with a complex averaging process by derivation of a macroscopic traffic flow model. The related mathematical theory is developed in [21, 22].

from free flow to synchronized flow (F → S transition) [32–34] rather than with an F → J transition. For this reason, Kerner introduced a three-phase traffic theory. In this theory, there are three traffic phases: (i) free flow, (ii) synchronized flow and (iii) wide moving jam.

The first microscopic models in the context of three-phase traffic theory introduced in 2002 are stochastic models [35, 36]. As in empirical observations [33, 34], in these models wide moving jams emerge spontaneously only in synchronized flow (S → J transition), i.e., the models exhibit the sequence of F → S → J transitions leading to wide moving jam emergence in free flow; in addition, the models show all types of congested patterns found in empirical observations [7, 35–39]. Recently, some new microscopic models based on three-phase traffic theory have been developed [40–42]. However, there are no deterministic models in the context of three-phase traffic theory, which can exhibit the F → S → J transitions found in empirical observations and the diagram of congested patterns of three-phase traffic theory [7, 34]. In stochastic models [7, 35–38], driver time delays in acceleration (deceleration) are simulated mainly through the use of model fluctuations. Therefore, a development of deterministic models based on three-phase traffic theory is important for a more realistic theory of car following behaviour.

In this paper, two deterministic microscopic three-phase traffic models are presented. In an **a**cceleration **t**ime **d**elay model (ATD model for short; section 2), an explicit description of driver time delays in vehicle acceleration (deceleration) is used. In a **s**peed **a**daptation model (SA model for short; section 3), vehicle speed adaptation occurs in synchronized flow depending on driving conditions. In sections 4 and 5, we show that these models exhibit the F → S → J transitions and congested patterns associated with results of empirical observations. In addition, a stochastic SA model is introduced and compared with the deterministic SA model of section 3. In section 6, the deterministic microscopic three-phase traffic models of sections 2 and 3 are compared with earlier deterministic models and a critical discussion of models in the context of the fundamental diagram approach is performed.

2. Acceleration time delay model

2.1. Driver behavioural assumptions and empirical basis of ATD model

A deterministic three-phase traffic flow model with driver time delays (ATD model) is based on the following empirical features of phase transitions and congested patterns as well as driver behavioural assumptions of three-phase traffic theory (sections 2.3, 2.4, and 8.6 of the book [7]):

(i) In synchronized flow, a driver accepts a range of different hypothetical steady states with various space gaps g at the same vehicle speed v, i.e., steady states of synchronized flow cover a two-dimensional region in the flow–density plane.

(ii) To avoid collisions, in the steady states a driver does not accept the vehicle speed that is higher than some safe speed (denoted by $v_s(g, v_\ell)$) that depends on the speed of the preceding vehicle v_ℓ. In contrast with earlier models in which a safe speed determines a multitude of steady states on the fundamental diagram [1, 3–6, 43], in the ATD model the safe speed determines the upper boundary of the two-dimensional region for the steady states in the flow–density plane [7].

(iii) If a driver cannot pass the preceding vehicle, then the driver tends to adjust the speed to the preceding vehicle within a synchronization gap $G(v, v_\ell)$, i.e., at

$$g \leqslant G(v, v_\ell) \tag{4}$$

a speed adaptation effect occurs. The synchronization gap determines the lower

boundary of the two-dimensional region for the steady states in the flow–density plane. In the ATD model, the speed adaptation effect is modelled through a driver acceleration $K(v, v_\ell)(v - v_\ell)$ adjusting the speed to the preceding vehicle under conditions (4); $K(v, v_\ell)$ is a sensitivity.

(iv) In traffic flow with greater space gaps, a driver searches for the opportunity to accelerate and to pass. This leads to driver over-acceleration, which is modelled through a driver acceleration $A(V^{(\mathrm{free})}(g) - v)$ adjusting the vehicle speed at

$$g > G(v, v_\ell) \tag{5}$$

to a gap-dependent optimal speed in free flow $V^{(\mathrm{free})}(g)$, where A is a sensitivity of this effect.

A competition between the speed adaptation effect and driver over-acceleration simulates a first-order F $\to$ S transition leading to the onset of congestion in real traffic flow (see explanations in section 2.4 in [7]).

(v) In empirical observations, due to an F $\to$ S transition there is a maximum point of free flow associated with the maximum density $\rho_{\max}^{(\mathrm{free})}$, maximum flow rate $q_{\max}^{(\mathrm{free})}$, and maximum speed $v_{\min}^{(\mathrm{free})}$ given by the formula $v_{\min}^{(\mathrm{free})} = q_{\max}^{(\mathrm{free})} \big/ \rho_{\max}^{(\mathrm{free})}$ (section 2.3 in [7]). This maximum point is modelled through F $\to$ S transition, which occurs already due to infinitesimal local perturbations in steady states of free flow associated with the optimal speed in free flow $V^{(\mathrm{free})}(g)$ at the density $\rho_{\max}^{(\mathrm{free})}$.

(vi) In high density flow, a driver decelerates stronger than it is required to avoid collisions if the preceding vehicle begins to decelerate unexpectedly (driver over-deceleration). In the ATD model, the over-deceleration effect, which explains and simulates moving jam emergence in synchronized flow, is modelled by a driver time delay τ in reduction of a current driver deceleration (denoted by $\tau = \tau_1^{(\mathrm{dec})}(v)$). The longer the $\tau_1^{(\mathrm{dec})}$, the stronger the over-deceleration effect. In empirical observations, the lower the synchronized flow speed, the greater the probability for moving jams emergence (section 2.4 in [7]). For this reason, $\tau_1^{(\mathrm{dec})}(v)$ is chosen to be longer at lower speeds than at higher ones.

(vii) At the downstream front of a wide moving jam or a synchronized flow region, a driver within the jam or the synchronized flow region does not accelerate before the preceding vehicle has begun to accelerate. In the ATD model, this effect is modelled through the use of a mean driver time delay in acceleration at the downstream front of the synchronized flow region, which depends on a time delay in driver acceleration (denoted by $\tau = \tau_0^{(\mathrm{acc})}$) and on the sensitivity $K(v, v_\ell)$ at $v < v_\ell$. At the downstream front of a wide moving jam, a mean time delay in acceleration from a standstill $v = 0$ within the jam should be longer than the mean driver time delay in synchronized flow [7]. To simulate this longer mean time delay in vehicle acceleration, in addition to two above-mentioned model effects, a vehicle within the jam does not accelerate before the condition

$$g \geqslant g_{\max}^{(\mathrm{jam})} \tag{6}$$

is satisfied, in which $g_{\max}^{(\mathrm{jam})}$ is the maximum space gap within the wide moving jam phase.

(viii) Moving in synchronized flow of lower speeds, a driver comes closer to the preceding vehicle than the synchronization gap G. In empirical observations, this self-compression of synchronized flow is called the pinch effect (section 12.2 in [7]). In the ATD model, the pinch effect is simulated through the use of two model assumptions. Firstly, a time delay in reduction of a current driver acceleration (denoted by $\tau = \tau_1^{(\mathrm{acc})}$) increases if the speed decreases. Secondly, the sensitivity $K(v, v_\ell)$, which describes the speed adaptation effect (item (iii)), is chosen at $v \geqslant v_\ell$ different from $K(v, v_\ell)$ at $v < v_\ell$ (item (vii)). Specifically, $K(v, v_\ell)$ at $v \geqslant v_\ell$ is chosen to be smaller at low speeds than at

higher ones. As a result, at lower speeds vehicles choose smaller space gaps than the synchronization gap G.

(ix) At the upstream front of a wide moving jam or a synchronized flow region, a driver begins to decelerate after a time delay denoted by $\tau = \tau_0^{(\mathrm{dec})}$. This delay time should describe realistic velocities of deceleration fronts in congested traffic patterns.

2.2. Main equations

An ATD model reads as follows:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v, \tag{7}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = a, \tag{8}$$

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \begin{cases} (a^{(\mathrm{free})} - a)/\tau & \text{at} \quad g > G \text{ and } g > g_{\max}^{(\mathrm{jam})}, \\ (a^{(\mathrm{syn})} - a)/\tau & \text{at} \quad g \leqslant G \text{ and } g > g_{\max}^{(\mathrm{jam})}, \\ (a^{(\mathrm{jam})} - a)/\tau & \text{at} \quad 0 \leqslant g \leqslant g_{\max}^{(\mathrm{jam})}, \end{cases} \tag{9}$$

where x is the vehicle space coordinate; $g = x_\ell - x - d$; the lower index ℓ marks variables related to the preceding vehicle; all vehicles have the same length d, which includes the minimum space gap between vehicles within a wide moving jam; $a^{(\mathrm{free})}$, $a^{(\mathrm{syn})}$ and $a^{(\mathrm{jam})}$ are vehicle accelerations (deceleration) in the free flow, synchronized flow and wide moving jam phases, respectively. If the condition (5) is satisfied, then a vehicle moves in accordance with the rules for free flow. Within synchronized flow associated with the condition $g_{\max}^{(\mathrm{jam})} < g \leqslant G(v, v_\ell)$, the vehicle tends to adapt the speed to the preceding vehicle. Within a wide moving jam, the space gap is small, specifically $g \leqslant g_{\max}^{(\mathrm{jam})}$, and the vehicle decelerates[4].

2.3. Driver acceleration

The accelerations (decelerations) $a^{(\mathrm{free})}$, $a^{(\mathrm{syn})}$ and $a^{(\mathrm{jam})}$ are found from the condition

$$a^{(\mathrm{phase})} = \min(\max(\tilde{a}^{(\mathrm{phase})}, a_{\min}), a_{\max}, a_{\mathrm{s}}), \tag{11}$$

the superscript 'phase' in (11) means either 'free', or 'syn', or else 'jam' for the related traffic phase; a_{s} is a deceleration related to safety requirements; $a_{\min}$ and $a_{\max}$ ($a_{\min} < 0, a_{\max} \geqslant 0$) are, respectively, the minimum and maximum accelerations for cases in which there are no safety restrictions. In (11), functions $\tilde{a}^{(\mathrm{free})}$, $\tilde{a}^{(\mathrm{syn})}$ and $\tilde{a}^{(\mathrm{jam})}$ associated with driver acceleration within the related traffic phase—free flow, or synchronized flow, or else wide moving jam—are determined as follows[5]:

$$\tilde{a}^{(\mathrm{free})}(g, v, v_\ell) = A(V^{(\mathrm{free})}(g) - v) + K(v, v_\ell)(v_\ell - v), \tag{12}$$

$$\tilde{a}^{(\mathrm{syn})}(g, v, v_\ell) = A \min\left(V_{\max}^{(\mathrm{syn})}(g) - v, 0\right) + K(v, v_\ell)(v_\ell - v), \tag{13}$$

$$\tilde{a}^{(\mathrm{jam})}(v) = -K^{(\mathrm{jam})} v. \tag{14}$$

[4] Since the vehicle speed v cannot be negative, the following condition is also used for equations (8), (9):

$$a(t) \geqslant 0 \quad \text{at} \quad v(t) = 0. \tag{10}$$

To satisfy this condition in numerical simulation, the acceleration $a(t)$ is replaced by the value $\max(a(t), 0)$ if $v(t) = 0$ at time t.

[5] In the paper, large enough flow rates on the main road are considered at which congested patterns can occur at a bottleneck. For this reason, in (12) K is chosen to be independent on g in the free flow phase. At considerably smaller flow rates in free flow, specifically, if g increases, K in (12) should tend towards zero when $g \gg G$.

Here, $V_{\max}^{(\mathrm{syn})}(g)$ is a gap-dependent maximum vehicle speed in steady states of synchronized flow, $K^{(\mathrm{jam})}$ is a sensitivity.

2.4. Safety conditions

Safety deceleration with a deceleration a_s can be applied, if the vehicle speed becomes higher than the safe speed $v_\mathrm{s}(g, v_\ell)$. We use safety deceleration found from the condition

$$a_\mathrm{s}(g, v, v_\ell) = A_\mathrm{s}(v_\mathrm{s}(g, v_\ell) - v), \tag{15}$$

where A_s is the sensitivity related to safety requirements.

The speed $v_\mathrm{s}(g, v_\ell)$ in (15) is found based on the safety condition of Gipps [43]:

$$v_\mathrm{s}T_\mathrm{s} + v_\mathrm{s}^2\big/(2b_\mathrm{s}) \leqslant g + v_\ell^2\big/(2b_\mathrm{s}), \tag{16}$$

where T_s is a safety time gap, b_s is a constant deceleration[6]. We use an approximated formula for $v_\mathrm{s}(g, v_\ell)$ derived from (16) in appendix A, which enables us to write $a_\mathrm{s}(g, v, v_\ell)$ (15) as follows:

$$a_\mathrm{s}(g, v, v_\ell) = A_\mathrm{s}^{(\mathrm{g})}(v_\ell)(g/T_\mathrm{s} - v) + K_\mathrm{s}(v_\ell)(v_\ell - v), \tag{18}$$

where

$$A_\mathrm{s}^{(\mathrm{g})}(v_\ell) = A_\mathrm{s}T_\mathrm{s}(T_\mathrm{s} + v_\ell/(2b_\mathrm{s}))^{-1}, \tag{19}$$

$$K_\mathrm{s}(v_\ell) = A_\mathrm{s}(T_0 + v_\ell/(2b_\mathrm{s}))(T_\mathrm{s} + v_\ell/(2b_\mathrm{s}))^{-1}, \tag{20}$$

T_0 is a constant.

2.5. Physics of driver time delays

In equations (9), the time delay τ is chosen as

$$\tau = \begin{cases} \tau_\mathrm{s} & \text{at} \quad a_\mathrm{s} < \min(0, \tilde{a}^{(\mathrm{phase})}, a), \\ \tilde{\tau} & \text{otherwise.} \end{cases} \tag{21}$$

Here, τ_s is a short driver time delay associated with a finite driver reaction time that must be taken into account in the cases when the driver should decelerate unexpectedly to avoid collisions, $\tilde{\tau}$ is a time delay in other traffic situations, which is chosen differently depending on whether the vehicle accelerates or decelerates:

$$\tilde{\tau} = \begin{cases} \tau^{(\mathrm{acc})} & \text{at} \quad a > 0, \\ \tau^{(\mathrm{dec})} & \text{at} \quad a \leqslant 0. \end{cases} \tag{22}$$

In turn, $\tau^{(\mathrm{acc})}$ and $\tau^{(\mathrm{dec})}$ in (22) depend on the acceleration a:

$$\tau^{(\mathrm{acc})} = \begin{cases} \tau_0^{(\mathrm{acc})} & \text{at} \quad a < \tilde{a}^{(\mathrm{phase})}, \\ \tau_1^{(\mathrm{acc})} & \text{otherwise,} \end{cases} \tag{23}$$

[6] Note that equations (9) of the ATD model can also be written without the term $a^{(\mathrm{jam})}$ as follows:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \begin{cases} (a^{(\mathrm{free})} - a)/\tau & \text{at} \quad g > G, \\ (a^{(\mathrm{syn})} - a)/\tau & \text{at} \quad g \leqslant G. \end{cases} \tag{17}$$

In (17), the speed $v_\mathrm{s}(g, v_\ell)$ in (15) is found based on the Gipps condition (16) when $g \geqslant g_{\max}^{(\mathrm{jam})}$ and the speed $v_\mathrm{s}(g, v_\ell) = 0$ when $g < g_{\max}^{(\mathrm{jam})}$. In the latter case, formula (15) with $v_\mathrm{s}(g, v_\ell) = 0$ plays the role of vehicle deceleration within the wide moving jam phase.

$$\tau^{(\mathrm{dec})} = \begin{cases} \tau_0^{(\mathrm{dec})} & \text{at} \quad a \geqslant \tilde{a}^{(\mathrm{phase})}, \\ \tau_1^{(\mathrm{dec})} & \text{otherwise.} \end{cases} \tag{24}$$

The driver time delays $\tau_0^{(\mathrm{acc})}$, $\tau_0^{(\mathrm{dec})}$, $\tau_1^{(\mathrm{dec})}$ and $\tau_1^{(\mathrm{acc})}$ in (23), (24) are associated with human expectation of local driving conditions, in particular, with spatial and temporal anticipation of a driver in accordance with local adaptation to those traffic situations in which the driver takes into account both the current and expected future behaviour of many vehicles ahead (see also section 2.1).

$\tau_0^{(\mathrm{acc})}$ is the mean time delay when a driver starts to accelerate or wants to increase the acceleration. This can often occur at the downstream front of a wide moving jam or a synchronized flow region, i.e., when the speed in traffic flow downstream of the vehicle is higher than the current vehicle speed. In these cases, after the preceding vehicle has begun to accelerate, the driver also begins to accelerate, however, after a time delay to have a desired time gap to the preceding vehicle.

$\tau_0^{(\mathrm{dec})}$ is the mean time delay when the driver starts to decelerate or wants to decelerate harder in cases in which the driver approaches a region of a lower speed downstream.

$\tau_1^{(\mathrm{acc})}$ corresponds to situations in which the driver accelerates currently but wants either to stop the acceleration or to reduce it. Thus, $\tau_1^{(\mathrm{acc})}$ is the mean driver time delay in interruption or reduction of driver acceleration in cases in which the driver recognizes that current acceleration is greater than a desired acceleration in the current driving situation.

$\tau_1^{(\mathrm{dec})}$ corresponds to situations in which the driver decelerates currently but wants either to stop the deceleration or to reduce it. Thus, $\tau_1^{(\mathrm{dec})}$ is the mean time delay in interruption or reduction of driver deceleration in cases in which the driver recognizes that current deceleration is more negative than a desired deceleration in the current driving situation.

2.6. Model of road with on-ramp bottleneck

'Open' boundary conditions are applied on the main road of the length L_0. At the beginning of the road, free flow conditions are generated for each vehicle one after another at equal time intervals $\tau_{\mathrm{in}} = 1/q_{\mathrm{in}}$ where q_{in} is the flow rate in the incoming boundary flow. To satisfy safety conditions, a new vehicle appears only if the distance from the beginning of the road ($x = x_{\mathrm{b}}$) to the position $x = x_\ell$ of the farthest upstream vehicle in the lane exceeds the distance $v_\ell T_{\mathrm{s}} + d$. The speed v and coordinate x of a new vehicle are $v = v_\ell$ and $x = x_{\mathrm{b}}$, respectively. After a vehicle has reached the end of the road, it is removed; before this, the farthest downstream vehicle maintains its speed. In the initial state ($t = 0$), all vehicles have the same initial speed $v = V^{(\mathrm{free})}(g)$ and space gap g, and $q_{\mathrm{in}} = v/(g + d)$.

An on-ramp bottleneck on the main road is considered. The on-ramp consists of two parts: (i) the merging region of the length L_{m} that begins at $x = x_{\mathrm{on}}$. Within this region, vehicles can merge onto the main road from the on-ramp. (ii) The part of the on-ramp lane of length L_{r} upstream of the merging region at which vehicles move according to the model equations for a homogeneous road with the maximum speed $v_{\mathrm{free,on}} = 90$ km h^{-1}. At the beginning of the on-ramp lane, the flow rate to the on-ramp q_{on} is given as the flow rate on the main road q_{in}.

The following rules are applied for vehicle merging within the merging region. A speed $\hat{v}$ is calculated corresponding to formula

$$\hat{v} = \min\left(v^+, v + \Delta v_{\mathrm{r}}^{(1)}\right) \tag{25}$$

and then it is used in the merging rules

$$g^+ > g_{\min}^{(\mathrm{on})} + \gamma \hat{v} T_{\mathrm{s}}, \qquad g^- > g_{\min}^{(\mathrm{on})} + \gamma v^- T_{\mathrm{s}}. \tag{26}$$

Table 1. ATD-model functions.

Synchronization gap
$G(v, v_\ell) = v \max(0, T^{(\mathrm{syn})}(v) + \kappa(v, v_\ell)(v - v_\ell))$
$\kappa(v, v_\ell) = \begin{cases} \kappa^{(\mathrm{acc})} & \text{at} \quad v < v_\ell \\ \kappa^{(\mathrm{dec})} & \text{at} \quad v \geqslant v_\ell \end{cases}$
$T^{(\mathrm{syn})}(v) = T_0^{(\mathrm{syn})}(1 - 0.85(v/V_0)^2)$
Sensitivities
$K(v, v_\ell) = \begin{cases} K^{(\mathrm{acc})} & \text{at} \quad v < v_\ell \\ K^{(\mathrm{dec})} & \text{at} \quad v \geqslant v_\ell \end{cases}$
$K^{(\mathrm{dec})}(v) = K_1^{(\mathrm{dec})}(1 - \lambda(v)) + K_2^{(\mathrm{dec})}\lambda(v)$
$\lambda(v) = (1 + \exp((v/v_\mathrm{c} - 1)/\epsilon))^{-1}$
Characteristic speed functions
$V^{(\mathrm{free})}(g) = V(g) \quad \text{at} \quad g \geqslant g_{\min}^{(\mathrm{free})}$
$V_{\max}^{(\mathrm{syn})}(g) = V(g) \quad \text{at} \quad g < g_{\min}^{(\mathrm{free})}$
$V(g) = V_0 \tanh((g + 2)/(V_0 T))$
$g_{\min}^{(\mathrm{free})}$ is the minimum space gap in free flow found as a solution of the set of the equations $v = V(g)$ and $g = G(v, v_\ell)$ at $v = v_\ell$
Speed dependensies of time delays
$\tau_1^{(\mathrm{dec})}(v) = \begin{cases} 0.5\mathrm{s} & \text{at} \quad v \geqslant v_\mathrm{c} \\ 0.7\mathrm{s} & \text{otherwise} \end{cases}$
$\tau_1^{(\mathrm{acc})}(v) = \begin{cases} 0.57\mathrm{s} & \text{at} \quad v \geqslant v_\mathrm{c} \\ 0.87\mathrm{s} & \text{otherwise} \end{cases}$

Here, superscripts g^+ and g^- are the space gaps to the preceding vehicle and the trailing vehicle on the main road, respectively; v^+ and v^- are the speeds of the preceding vehicle and the trailing vehicle, respectively; γ, $g_{\min}^{(\mathrm{on})}$ and $\Delta v_\mathrm{r}^{(1)}$ are the constants, where $g_{\min}^{(\mathrm{on})}$ is the minimum gap at which vehicle merging is possible, $\Delta v_\mathrm{r}^{(1)}$ describes the maximum possible increase in speed after vehicle merging. Note that the finite increase $\Delta v_\mathrm{r}^{(1)}$ in the vehicle speed (25) is used to simulate a complex driver behaviour during merging onto the main road, especially in synchronized flow: in some cases, before merging the driver has to accelerate abruptly to adjust the speed to the speed of the preceding vehicle.

If conditions (26) are satisfied, then the vehicle merges onto the main road. After merging the vehicle speed v is set to $\hat{v}$ (25) and the vehicle coordinate does not change. If the conditions (26) are not satisfied, the vehicle does not merge onto the main road. In this case, the vehicle moves in the on-ramp lane until it comes to a stop at the end of the merging region.

2.7. Model functions and parameters

Model functions and parameters are shown in tables 1 and 2, respectively. As explained in section 2.1, driver time delays $\tau_1^{(\mathrm{dec})}$ and $\tau_1^{(\mathrm{acc})}$ are chosen to be functions of the vehicle speed; additionally, the synchronization gap $G(v, v_\ell)$ and sensitivity $K(v, v_\ell)$ are chosen to be asymmetric speed functions depending on whether the vehicle speed v is higher or lower than the speed v_ℓ. Explanations of the function $K(v, v_\ell)$ have been made in items (vii) and (viii) of section 2.1.

Speed dependence and an asymmetric function for the synchronization gap $G(v, v_\ell)$ are explained by driver behaviour as follows. The synchronization gap is the space gap at which

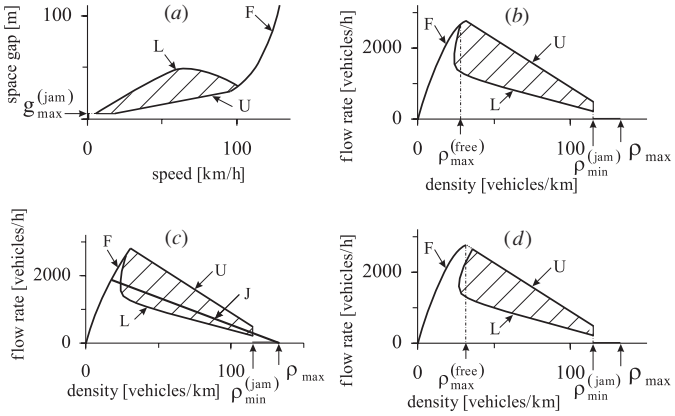

Figure 1. Model steady states for the ATD models: (*a*) in the space-gap–speed plane. (*b*) In the flow–density plane. (*c*) Steady states and the line J (for explanation of the line J see the book [7]). (*d*) ATD model with separated steady states in free flow and synchronized flow.

Table 2. ATD-model parameters.

$V_0 = 33.3$ m s^{-1} (120 km h^{-1}), $T = 0.9$ s
$A = 0.5$ s^{-1}, $K^{(\mathrm{acc})} = 0.8$ s^{-1}, $K^{(\mathrm{jam})} = 1$ s^{-1}
$K_1^{(\mathrm{dec})} = 0.95$ s^{-1}, $K_2^{(\mathrm{dec})} = 0.48$ s^{-1}, $v_c = 15$ m s^{-1}, $\epsilon = 0.15$
$T_0^{(\mathrm{syn})} = 2.5$ s, $\kappa^{(\mathrm{acc})} = 0.5$ s^{2}m^{-1}, $\kappa^{(\mathrm{dec})} = 0.55$ s^2 m^{-1}
$g_{\max}^{(\mathrm{jam})} = 0.95$ m, $\tau_0^{(\mathrm{dec})} = 1$ s, $\tau_0^{(\mathrm{acc})} = 0.75$ s, $\tau_s = 0.4$ s
$a_{\max} = 1$ms^{-2}, $a_{\min} = -1$ m s^{-2}
$A_s = 1.25$ s^{-1}, $b_s = 2$ m s^{-2}, $T_s = 1$ s, $T_0 = 0.42$ s
$L_0 = 25$ km, $x_b = -5$ km, $x_{on} = 16$ km
$L_m = 300$ m, $L_r = 500$ m, $\gamma = 0.22$, $\Delta v_r^{(1)} = 8$ m s^{-1}, $g_{\min}^{(\mathrm{on})} = 0$

a driver adapts its speed to the speed of the preceding vehicle. Firstly, the synchronization gap is an increasing function of speed: the lower the speed, the smaller the maximum gap at which the driver can comfortably move in synchronized flow. Secondly, if $v < v_\ell$, the driver accelerates and he/she can start speed adaptation at a smaller space gap than in the opposite case $v > v_\ell$. The function $T^{(\mathrm{syn})}(v)$ is used to have a difference in vehicle space gap in steady states of free flow and synchronized flow at a given flow rate. This space gap difference, which is used for simulation of a first-order F $\to$ S transition, tends towards zero when the density in free flow approaches the maximum point for free flow $\rho_{\max}^{(\mathrm{free})}$ (figures 1(*a*)–(*c*)); see also item (v) of section 2.1).

2.8. Steady states

In steady states, all vehicles have the same speed $v = v_\ell$ and the same space gap g, and all accelerations and their time derivatives are zero, and the density ρ and the flow rate q are related to the space gap g and to the speed v by the obvious conditions

$$\rho = 1/(x_\ell - x) = 1/(g + d), \qquad q = \rho v = v/(g + d). \tag{27}$$

According to (8)–(15) and formulae for $V^{(\mathrm{free})}(g)$, $V_{\max}^{(\mathrm{syn})}(g)$ (table 1) for steady states, we get

$$v = V(g) \quad \text{at} \quad g > G(v) \quad \text{and} \quad g > g_{\max}^{(\mathrm{jam})}, \tag{28}$$

$$v \leqslant V(g) \quad \text{at} \quad g \leqslant G(v) \quad \text{and} \quad g > g_{\max}^{(\mathrm{jam})}, \tag{29}$$

$$v = 0 \quad \text{at} \quad g \leqslant g_{\max}^{(\mathrm{jam})}, \tag{30}$$

$$v \leqslant v_{\mathrm{s}}(g, v). \tag{31}$$

According to (28)–(31), the model steady states consist of the curve $v = V(g)$ (28) at $g \geqslant g_{\min}^{(\mathrm{free})}$ (curve F in figure 1(a)) for free flow, a two-dimensional region in the space-gap–speed plane for synchronized flow determined by inequalities in (29), (31) and the line $v = 0$ at $g \leqslant g_{\max}^{(\mathrm{jam})}$ (30) for wide moving jams (figure 1(a)).

The two-dimensional region for steady states of synchronized flow is limited by the following boundaries: the boundary U, the curve L, the curve $v = V(g)$ at $g < g_{\min}^{(\mathrm{free})}$ and the horizontal line $g = g_{\max}^{(\mathrm{jam})}$. The boundary U is associated with the safe speed, i.e., this is determined by condition (31) when it is an equality. This leads to the condition for the boundary U

$$g = vT_{\mathrm{s}}. \tag{32}$$

The boundary L is found from the condition that the vehicle space gap is equal to the synchronization gap

$$g = G(v). \tag{33}$$

In the flow–density plane, free flow (curve F in figure 1(b)) is found from

$$q = \rho V_{\mathrm{F}}(\rho) \tag{34}$$

at $\rho \leqslant \rho_{\max}^{(\mathrm{free})}$ where $V_{\mathrm{F}}(\rho) = V(g)|_{g=\rho^{-1}-d}$, $\rho_{\max}^{(\mathrm{free})} = \left(g_{\min}^{(\mathrm{free})} + d\right)^{-1}$. A wide moving jam is associated with the horizontal line $q = 0$ at $\rho_{\min}^{(\mathrm{jam})} \leqslant \rho \leqslant \rho_{\max}$ (figure 1(b)), where $\rho_{\min}^{(\mathrm{jam})} = \left(g_{\max}^{(\mathrm{jam})} + d\right)^{-1}$, $\rho_{\max} = d^{-1}$. The boundaries of a two-dimensional region for steady states of synchronized flow are: the upper line U determined by the condition $q = (1-\rho d)/T_{\mathrm{s}}$, the lower curve L determined by the condition $\rho G(q/\rho) = 1-\rho d$, the curve (34) at $\rho > \rho_{\max}^{(\mathrm{free})}$, and the vertical line $\rho = \rho_{\min}^{(\mathrm{jam})}$.[7]

3. Speed adaptation model

3.1. Empirical $F \to S \to J$ transitions as physical basis of speed adaptation model

The fundamental hypothesis of three-phase traffic theory, which postulates that hypothetical steady states of synchronized flow cover a two-dimensional region in the flow–density plane, is also one of the basic hypotheses of the ATD model presented above (figures 1(a) and (b)). In contrast with the ATD model, in a speed adaptation model (**s**peed **a**daptation model, SA model for short) hypothetical steady states of synchronized flow are associated with a curve (curve S in figures 2(a) and (b)), i.e., they cover a one-dimensional region in the flow–density plane. The curve S is associated with an averaging of an infinite number of steady states of synchronized flow to one synchronized flow speed for each vehicle space gap. A gap

[7] We have also studied another version of the ATD model in which there is a separation of steady states in free flow and synchronized flow in the flow–density plane, i.e., the maximum point for free flow $\rho_{\max}^{(\mathrm{free})}$ is related to the intersection point of the line U and the curve F (figure 1(d)). Simulations of this version of the ATD model qualitatively show the same features of phase transitions and congested patterns as those discussed in section 4.

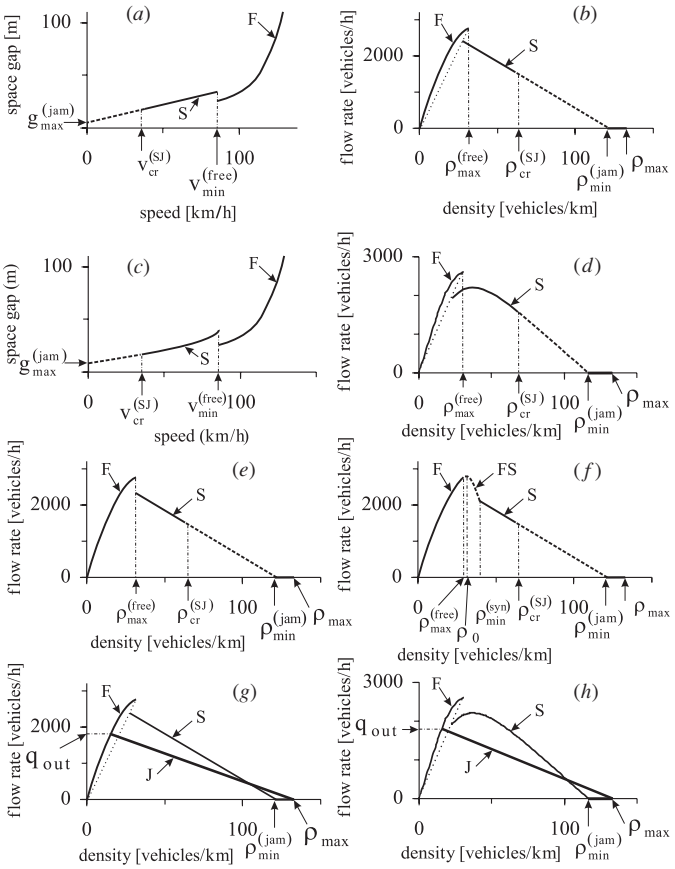

Figure 2. Model steady states for SA models: (*a*, *b*) in the space-gap–speed (*a*) and the flow–density plane (*b*) for the SA model (35)–(41). (*c*, *d*) The SA model (35)–(40), (42). (*e*, *f*) Other possible SA models (see appendix B). Dashed parts of curves for model steady states in (*a*–*f*) are associated with unstable model steady states. (*g*, *h*) Steady states and the line *J* for the SA model (*a*, *b*) and the SA model (*c*, *d*).

dependence of the average speed in synchronized flow steady states on the curve *S* is denoted by $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$. The basis hypothesis of the SA model is associated with the sequence of F $\rightarrow$ S $\rightarrow$ J transitions, which determine moving jam emergence in empirical observations [7, 33].

Note that as in the models and theories in the context of the fundamental diagram approach [1, 3–6], in the SA model steady-state model solutions cover a one-dimensional region(s) in the flow–density plane. However, in the models and theories reviewed [1, 3–6], which claim to show spontaneous moving jam emergence, the F $\rightarrow$ J transition governs the onset of congestion. This is inconsequent with empirical results [7]. In contrast, in the SA model the onset of congestion is associated with an F $\rightarrow$ S transition, whereas moving jams occur spontaneously only in synchronized flow, in accordance with empirical results.

The SA model is simpler than the ATD model. However, due to this simplification the SA model cannot show some features of congested patterns of the ATD model (section 5.2), which are observed in empirical observations. The purpose of the SA model is to simulate an F $\rightarrow$ S transition and features of the sequence of F $\rightarrow$ S $\rightarrow$ J transitions, as observed in empirical observations [7, 33], in a simple way. This confirms an assumption of three-phase

traffic theory that if rather than the fundamental hypothesis the hypothesis about the F → S → J transitions is the basis of a mathematical model, then the model can show and predict some important empirical features of the phase transitions (see footnote 4 of section 4.3.4 in [7]).

In the SA models, an F → S transition is modelled through two effects: (i) discontinuity of steady speed solutions (figures 2(*a*), (*c*) and (*e*)) or their instability (curve FS in figure 2(*f*)) in the vicinity of the maximum point of free flow $v_{\min}^{(\mathrm{free})}$, $\rho_{\max}^{(\mathrm{free})}$. (ii) The speed adaptation effect is modelled through the term $K(v, v_\ell)(v - v_\ell)$ that adjusts the speed to the preceding vehicle in synchronized flow.

Moving jam emergence is simulated through an instability of some of the synchronized flow model steady states associated with the curve $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$. This instability occurs in synchronized flow at lower speeds and greater densities (i.e., smaller space gaps). The associated critical density and speed of the synchronized flow steady states are denoted by $\rho_{\mathrm{cr}}^{(\mathrm{SJ})}$ and $v_{\mathrm{cr}}^{(\mathrm{SJ})}$, respectively (figure 2). To simulate this instability, as in the ATD model (item (viii) of section 2.1), in the SA models the sensitivity $K(v, v_\ell)$ at $v \geqslant v_\ell$ is a decreasing speed function. Similarly with the ATD model, to simulate the mean time delay in acceleration at the downstream jam front in the SA model, a vehicle within the jam does not accelerate before (6) is satisfied (item (vii) of section 2.1).

3.2. Basic equations

There can be different possibilities for a separation of steady states of free flow and synchronized flow in SA models, which all qualitatively exhibit the same features of the F → S → J transitions. To illustrate this, here we consider two variants of SA models; in appendix B other possible variants of SA models are discussed. All these variants of the SA models exhibit very similar features of phase transitions and spatiotemporal congested traffic patterns that are associated with the same physics of these SA models.

A formulation for the SA model reads as follows:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v, \tag{35}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \begin{cases} a^{(\mathrm{free})} & \text{at} \quad v \geqslant v_{\min}^{(\mathrm{free})} \text{ and } g > g_{\max}^{(\mathrm{jam})}, \\ a^{(\mathrm{syn})} & \text{at} \quad v < v_{\min}^{(\mathrm{free})} \text{ and } g > g_{\max}^{(\mathrm{jam})}, \\ a^{(\mathrm{jam})} & \text{at} \quad 0 \leqslant g \leqslant g_{\max}^{(\mathrm{jam})}. \end{cases} \tag{36}$$

The vehicle acceleration $a = \mathrm{d}v/\mathrm{d}t$ in (36) is supposed to be limited by the maximum acceleration $a_{\max}$, i.e., in (36)

$$a^{(\mathrm{phase})} = \min(\tilde{a}^{(\mathrm{phase})}, a_{\max}). \tag{37}$$

Here and below the associated designations of functions and parameters have the same meaning as those in the ATD model (section 2).

3.3. Vehicle acceleration

Functions $\tilde{a}^{(\mathrm{free})}(g, v, v_\ell)$, $\tilde{a}^{(\mathrm{syn})}(g, v, v_\ell)$ and $\tilde{a}^{(\mathrm{jam})}(v)$ in (37) are chosen as follows:

$$\tilde{a}^{(\mathrm{free})}(g, v, v_\ell) = A^{(\mathrm{free})}(V^{(\mathrm{free})}(g) - v) + K(v, v_\ell)(v_\ell - v), \tag{38}$$

$$\tilde{a}^{(\mathrm{syn})}(g, v, v_\ell) = A^{(\mathrm{syn})}\left(V_{\mathrm{av}}^{(\mathrm{syn})}(g) - v\right) + K(v, v_\ell)(v_\ell - v), \tag{39}$$

$$\tilde{a}^{(\mathrm{jam})}(v) = -K^{(\mathrm{jam})}v. \tag{40}$$

Two versions of functions $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$ in (39) that lead to two different versions of the SA models are considered:

$$V_{\mathrm{av}}^{(\mathrm{syn})}(g) = \tilde{g}(g)\big/T_{\mathrm{av}}^{(\mathrm{syn})} \tag{41}$$

and

$$V_{\mathrm{av}}^{(\mathrm{syn})}(g) = V_1\left[\tanh\left(\frac{\tilde{g}(g)}{T_{\mathrm{av}}^{(\mathrm{syn})}V_1}\right) + c\tilde{g}(g)\right], \tag{42}$$

where $\tilde{g}(g) = g - g_{\mathrm{max}}^{(\mathrm{jam})}$; $T_{\mathrm{av}}^{(\mathrm{syn})}$, V_1 and c are the constants.

3.4. Steady states and model parameters

In the SA models, in accordance with (36) there are three isolated curves for steady states of the SA models associated with the three traffic phases: free flow, synchronized flow and wide moving jam (figures 2(*a*) and (*b*)).

Steady states of free flow are related to a curve $v = V_{\mathrm{F}}(\rho)$ and formula (34) (the curve F in figures 2(*a*)–(*d*)) associated with the condition

$$v = V^{(\mathrm{free})}(g) \quad \text{at} \quad v \geqslant v_{\mathrm{min}}^{(\mathrm{free})}. \tag{43}$$

Steady states of synchronized flow are related to a curve S in the space-gap–speed plane (figures 2(*a*) and (*c*)) given by the condition

$$v = V_{\mathrm{av}}^{(\mathrm{syn})}(g) \qquad \text{at} \quad v < v_{\mathrm{min}}^{(\mathrm{free})} \quad \text{and} \quad g > g_{\mathrm{max}}^{(\mathrm{jam})}. \tag{44}$$

In terms of the flow rate q and density ρ, the formula (44) reads

$$q = \rho V_{\mathrm{S}}(\rho) \qquad \text{at} \quad \rho_{\mathrm{min}}^{(\mathrm{syn})} < \rho < \rho_{\mathrm{min}}^{(\mathrm{jam})}, \tag{45}$$

where $V_{\mathrm{S}}(\rho) = V_{\mathrm{av}}^{(\mathrm{syn})}(g)|_{g=\rho^{-1}-d}$, $\rho_{\mathrm{min}}^{(\mathrm{syn})} = \left(g_{\mathrm{max}}^{(\mathrm{syn})} + d\right)^{-1}$, $g_{\mathrm{max}}^{(\mathrm{syn})}$ is found from the equation $V_{\mathrm{av}}^{(\mathrm{syn})}\left(g_{\mathrm{max}}^{(\mathrm{syn})}\right) = v_{\mathrm{min}}^{(\mathrm{free})}$.

In the case of the function $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$ given by (41), formula (45) yields the equation for a curve S with a negative slope in the flow–density plane (figure 2(*b*))

$$q = \left(1 - \rho\big/\rho_{\mathrm{min}}^{(\mathrm{jam})}\right)\big/T_{\mathrm{av}}^{(\mathrm{syn})} \quad \text{at} \quad \rho_{\mathrm{min}}^{(\mathrm{syn})} < \rho < \rho_{\mathrm{min}}^{(\mathrm{jam})}. \tag{46}$$

When the function $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$ is given by formula (42), the curve S has a maximum in the flow–density plane (figure 2(*d*)).

Steady states for a wide moving jam are the same as those in the ATD model, i.e., they are given by a horizontal line

$$q = 0 \quad \text{at} \quad \rho_{\mathrm{min}}^{(\mathrm{jam})} \leqslant \rho \leqslant \rho_{\mathrm{max}} \tag{47}$$

in the flow–density plane (figures 2(*b*) and (*d*)).

Parameters of the SA models are shown in table 3.

4. Diagram of congested traffic patterns at on-ramp bottleneck in ATD model

Numerical simulations of the ATD model show that congested patterns (figure 3), which appear on the main road upstream of the bottleneck, are qualitatively the same as those for the stochastic models of [35–37] reviewed in the book [7]. However, dynamics of phase transitions leading to congested pattern formation and a diagram of these patterns in the flow–flow plane with coordinates being q_{on} and q_{in} (figure 3(*a*)) exhibit some important peculiarities in comparison with the stochastic models [35–37]. These peculiarities are associated with a deterministic character of the ATD model. To understand this, firstly features of an F $\to$ S transition at the bottleneck in the deterministic ATD model should be considered.

Table 3. SA-model parameters.

$V^{(\mathrm{free})}(g) = V(g)$
$V(g) = V_0 \tanh(g/(V_0 T))$, $V_0 = 33.3$ m s^{-1} (120 km h^{-1}), $T = 0.85$ s
$a_{\max} = 2$ m s^{-2}; $A^{(\mathrm{free})} = 0.4$ s^{-1}
$A^{(\mathrm{syn})} = 0.1$ s^{-1}, $K^{(\mathrm{jam})} = 2.2$ s^{-1}
$K(v, v_\ell)$ is given in table 1
$K^{(\mathrm{dec})}(v)$ is given in table 1 at $v_{\mathrm{c}} = 10$ m s^{-1}, $\epsilon = 0.07$
$g_{\min}^{(\mathrm{on})} = g_{\max}^{(\mathrm{jam})}$, $\gamma = 0.25$
In the SA model with function $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$ (41), we use
$v_{\min}^{(\mathrm{free})} = 22.22$ m s^{-1} (80 km h^{-1}), $K_1^{(\mathrm{dec})} = 0.95$ s^{-1} $K_2^{(\mathrm{dec})} = 0.64$ s^{-1}
$K^{(\mathrm{acc})} = 0.4$ s^{-1}; $T_{\mathrm{av}}^{(\mathrm{syn})} = 1.2$ s, $g_{\max}^{(\mathrm{jam})} = 0.7$ m, $\Delta v_{\mathrm{r}}^{(1)} = 3.5$ m s^{-1}
In the SA model with function $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$ (42), we use
$v_{\min}^{(\mathrm{free})} = 23.61$ m s^{-1} (85 km h^{-1}), $K_1^{(\mathrm{dec})} = 0.95$ s^{-1}, $K_2^{(\mathrm{dec})} = 0.75$ s^{-1}
$K^{(\mathrm{acc})}(v) = 0.3 + 0.4 \min(1, v/12)$ s^{-1}, $V_1 = 20$ m s^{-1}
$c = 0.007$ m^{-1}, $T_{\mathrm{av}}^{(\mathrm{syn})} = 1$ s, $g_{\max}^{(\mathrm{jam})} = 1$ m, $\Delta v_{\mathrm{r}}^{(1)} = 2.5$ m s^{-1}

4.1. Local perturbation and F → S transition in free flow at bottleneck

Vehicle merging results in an abrupt local space-gap reduction on the main road. This can lead to abrupt local vehicle deceleration. For this reason, a dynamic decrease in speed (figure 4(*a*)) and the associated increase in density in the on-ramp merging region appear. This local disturbance in the speed and density localized at the bottleneck can be considered a time-dependent dynamic perturbation in free flow. The dynamic nature of this perturbation (there are no random fluctuations in the deterministic ATD model) is explained by dynamic rules of vehicle motion and by a spatial non-homogeneity localized in the on-ramp merging region within which on-ramp inflow and flow on the main road merge. If the flow rate q_{in} is great enough, then due to dynamic merging rules of section 2.6 vehicles can merge onto the main road at different locations within the merging region. This complex dynamic vehicle merging behaviour causes the associated complex dynamic spatiotemporal dependence of the speed and, respectively, density within the dynamic perturbation (figure 4(*a*)).

If the speed and density within the perturbation are averaged over time with an averaging time interval that is considerably longer than time intervals between merging of vehicles, then spatial distributions of the speed and density within the associated average perturbation (figures 4(*b*) and (*c*)) can be considered a 'deterministic' perturbation localized at on-ramp bottleneck. At this time scale the deterministic perturbation is motionless, the total flow rate (across the main road and on-ramp lane) within the perturbation does not depend on spatial coordinate. This total flow rate in free flow is $q_{\mathrm{sum}} = q_{\mathrm{in}} + q_{\mathrm{on}}$. In contrast, the average speed and density spatially vary in free flow at the bottleneck. In particular, $q_{\mathrm{sum}} = v_{\mathrm{free}}^{(\mathrm{B})} \rho_{\mathrm{free}}^{(\mathrm{B})} = v^{(\mathrm{free})} \rho^{(\mathrm{free})}$, where $v_{\mathrm{free}}^{(\mathrm{B})}$ and $\rho_{\mathrm{free}}^{(\mathrm{B})}$ are the minimum speed and maximum density within the deterministic perturbation, respectively; $v^{(\mathrm{free})}$, $\rho^{(\mathrm{free})}$ are the speed and density downstream of the perturbation, respectively (figures 4(*b*) and (*c*)); $v_{\mathrm{free}}^{(\mathrm{B})} < v^{(\mathrm{free})}$, $\rho_{\mathrm{free}}^{(\mathrm{B})} > \rho^{(\mathrm{free})}$.

At a given q_{in}, the greater the q_{on}, the lower the speed $v_{\mathrm{free}}^{(\mathrm{B})}$ and the greater the density $\rho_{\mathrm{free}}^{(\mathrm{B})}$ within the perturbation, i.e., the greater the amplitude of the deterministic perturbation (figures 4(*b*) and (*c*)). This growth in the perturbation amplitude has a limit associated with an F → S transition that occurs spontaneously at the bottleneck when q_{on} gradually increases.

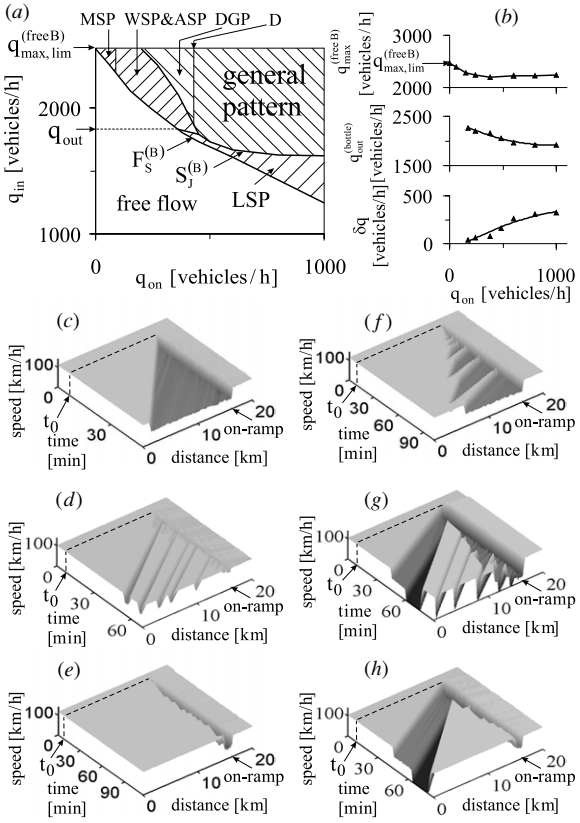

Figure 3. Congested patterns in the ATD model: (*a*) diagram of congested patterns. (*b*) Maximum capacity in free flow at the bottleneck $q_{\max}^{(\text{free B})}$, the discharge flow rate downstream of the congested bottleneck $q_{\text{out}}^{(\text{bottle})}$, and the capacity drop δq. (*c*–*h*) Congested patterns upstream of the bottleneck related to (*a*): (*c*–*f*) synchronized flow patterns (SPs) and (*g*, *h*) general patterns (GPs). (*c*) Widening SP (WSP). (*d*) Sequence of moving SPs (MSPs). (*e*) Localized SP (LSP). (*f*) Alternating synchronized flow pattern (ASP). (*g*) GP. (*h*) Dissolving GP (DGP). In (*c*–*h*), the flow rates (q_{on}, q_{in}) are: (*c*) (350, 2140), (*d*) (60, 2367), (*e*) (360, 1800), (*f*) (240, 2026), (*g*) (510, 2310), (*h*) (360, 2310) vehicles h^{-1}. $q_{\max,\text{lim}}^{(\text{free B})} \approx 2470$ vehicles h^{-1}. In (*b*), the discharge flow rate $q_{\text{out}}^{(\text{bottle})}$ is changed from 2270 to 1925 vehicles h^{-1}, $q_{\text{out}} \approx 1805$ vehicles h^{-1}. $T_{\text{ob}} = 30$ min for the boundary $F_{\text{S}}^{(\text{B})}$ and 60 min for the boundary $S_{\text{J}}^{(\text{B})}$. $t_0 = 7$ min.

The multitude of the flow rates q_{in} and q_{on}, at which the F $\to$ S transition occurs, determines the boundary $F_{\text{S}}^{(\text{B})}$ in the pattern diagram (figure 3(*a*)). At the boundary $F_{\text{S}}^{(\text{B})}$, a first-order F $\to$ S transition (see section 5.1) occurs spontaneously during a chosen time interval T_{ob} that is considerably longer than a time interval $\tau_{\text{determ}}^{(\text{grow B})}$ (about 60 s) required for the average speed to decrease from the speed within a dynamic perturbation in free flow at the bottleneck to a synchronized flow speed (see explanations in section 5.3.7 of [7]). The necessity of the time interval T_{ob} is associated with a time delay $T_{\text{FS}}^{(\text{B})}$ for an F $\to$ S transition found in the ATD model: after the time delay $T_{\text{FS}}^{(\text{B})}$, a time-dependent (dynamic) perturbation (figure 4(*a*)), which can cause a short-time decrease in the speed within the perturbation markedly lower than $v_{\text{free}}^{(\text{B})}$, can occur. This perturbation occurrence leads to the F $\to$ S transition. The boundary $F_{\text{S}}^{(\text{B})}$ is determined from the condition $T_{\text{ob}} \approx T_{\text{FS}}^{(\text{B})}$.

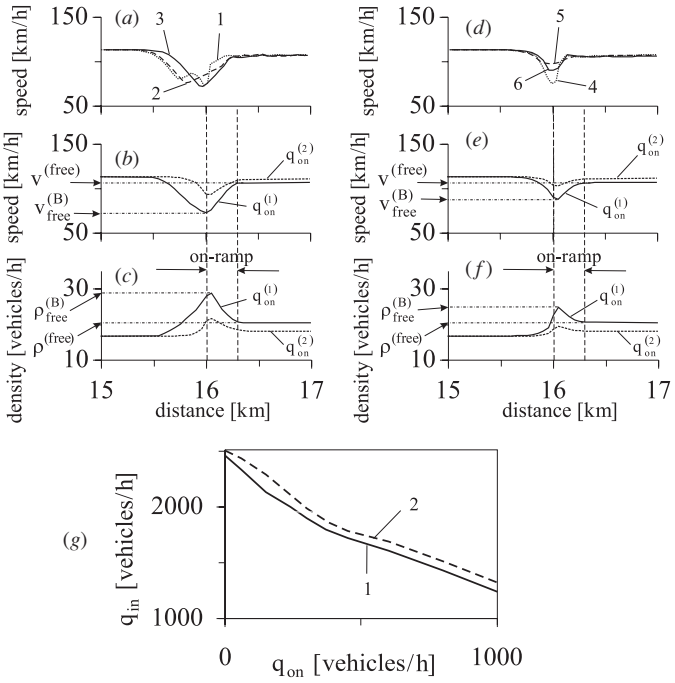

Figure 4. Peculiarities of F → S transitions in the ATD model: (*a*–*f*) space dependences of speed and density within local perturbations at on-ramp bottleneck at two different parameters of vehicle merging from on-ramp onto the main road $\Delta v_r^{(1)} = 8$ m s^{-1} (*a*–*c*) and $\Delta v_r^{(1)} = 12$ m s^{-1} (*d*–*f*). $q_{on} = q_{on}^{(1)} = 300$, $q_{on} = q_{on}^{(2)} = 120$, and $q_{in} = 1900$ vehicles h^{-1}. In (*a*, *d*) curves 1–6 are related to different time moments. (*g*) Boundaries $F_S^{(B)}$ of a first-order F → S transition at the bottleneck associated with (*a*) and (*d*), respectively. The boundary $F_S^{(B)}$ associated with curve 1 in (*g*) is taken from figure 3(*a*). In (*b*, *c*, *e*, *f*), 5 min averaging of traffic variables measured at virtual detectors at different locations are shown. Other model parameters are the same as those in figure 3.

In stochastic models [7], the boundary $F_S^{(B)}$ is also determined by the condition that an F → S transition occurs at given q_{in} and q_{on} after a time delay $T_{FS}^{(B)}$ during a chosen time interval T_{ob}. However, in the stochastic models $T_{FS}^{(B)}$ is a *random* value: in different realizations made at the same q_{in} and q_{on} various $T_{FS}^{(B)}$ are found. This stochastic model nature also enables us to calculate the probability for F → S transition occurrence [36, 7].

In contrast with the stochastic models [7], in the deterministic ATD model there are no random fluctuations. Time-dependent perturbations in free flow localized at the bottleneck (figure 4(*a*)) have dynamic nature explained above. For this reason, in the ATD model $T_{FS}^{(B)}$ is a *fixed* value at given q_{in} and q_{on}; consequently, the probability for F → S transition occurrence cannot be found.

In addition, numerical simulations of the ATD model show that a duration of a dynamic speed decrease within the perturbation below the speed $v_{free}^{(B)}$ is considerably shorter (1–3 s) than $\tau_{determ}^{(grow\,B)}$. As a result, it is found that at a given q_{in} the time delay $T_{FS}^{(B)}$ is a strong decreasing function of q_{on} in a neighbourhood of the boundary $F_S^{(B)}$: already a small increase in q_{on} behind the boundary $F_S^{(B)}$ leads to a decrease in $T_{FS}^{(B)}$ down to $\tau_{determ}^{(grow\,B)}$. Thus, we can suggest that in the ATD model the boundary $F_S^{(B)}$ is very close to the boundary for the deterministic F → S transition (see explanation of the deterministic F → S transition in section 5.3.7 of [7]).

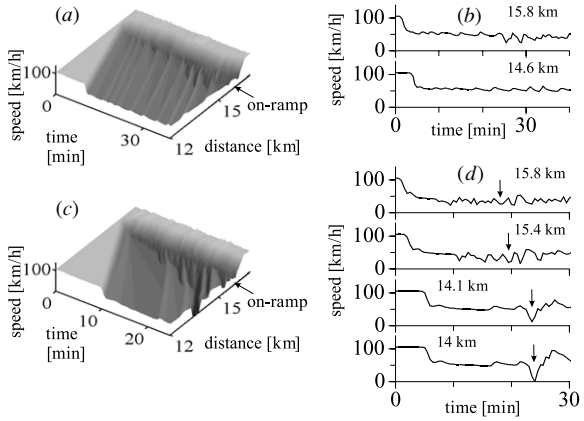

Figure 5. Peculiarities of S → J transitions in the ATD model: (*a*, *b*) spatiotemporal *decay* of dynamic speed waves that emerge in the merging on-ramp region during their upstream propagation through a widening SP (WSP) in space and time (*a*) and at different virtual detectors (*b*). (*c*, *d*) Spatiotemporal *growth* of dynamic speed waves that emerge in the merging on-ramp region during their upstream propagation through an initial WSP in space and time (*c*) and at different virtual detectors (*d*). $q_{on} = 300$ (*a*, *b*), 360 (*c*, *d*) and $q_{in} = 2310$ vehicles h^{-1}. Other model parameters are the same as those in figure 3.

The dynamic character of perturbations at the bottleneck, which is responsible for the above-mentioned physics of the boundary $F_S^{(B)}$ for an F → S transition in the ATD model, can clearly be seen, if smaller disturbances in speed and density occur due to vehicle merging. Smaller disturbances can be simulated by an increase in the parameter $\Delta v_r^{(1)}$ of vehicle merging (section 2.6). As a result, at the same q_{in} and q_{on} as those in figures 4(*a*)–(*c*) both time-dependent (figure 4(*d*)) and deterministic perturbations (figures 4(*e*) and (*f*)) become smaller. This leads to a shift of the boundary $F_S^{(B)}$ in the diagram of congested patterns to greater q_{on} (curve 2 in figure 4(*g*)).

4.2. Peculiarities of S → J transitions and congested patterns

In the ATD model, moving jam formation in synchronized flow (S → J transition), which occurs at the boundary $S_J^{(B)}$ in the congested pattern diagram (figure 3(*a*)), also exhibits some qualitative different features in comparison with the stochastic models [7].

As in the stochastic models [7], in the ATD model after a synchronized flow pattern (SP) occurs upstream of the bottleneck due to F → S transition at the bottleneck, a further increase in q_{on} leads to a subsequent decrease in the speed within the SP. This can cause an S → J transition with the following general pattern (GP) formation. In the stochastic models, a self-growth of random model fluctuations is mostly responsible for the S → J transition. In contrast, in the ATD model there are no random model fluctuations.

In the ATD model, dynamic merging of vehicles from the on-ramp lane onto the main road can cause dynamic speed and density waves that propagate upstream in synchronized flow of the SP (figure 5). It turns out that if the flow rate q_{on} is related to a point (q_{on}, q_{in}) between the boundaries $F_S^{(B)}$ and $S_J^{(B)}$ (figure 3(*a*)), then these dynamic waves decay during their upstream propagation within synchronized flow of the SP (figures 5(*a*) and (*b*)). In contrast, at the boundary $S_J^{(B)}$ the waves begin to self-growth in their amplitude leading wide

moving jam formation, i.e., one of GPs appears upstream of the bottleneck (figures 5(*c*) and (*d*)).

As in the KKW cellular automata (CA) model [7, 36], in the ATD model the maximum flow rate in free flow downstream of the bottleneck $q_{\max}^{(\text{free B})}$ is a decreasing function of q_{on} (figure 3(*b*)). Recall that the flow rate $q_{\max}^{(\text{free B})}(q_{\text{on}})$ is the flow rate in free flow downstream of the bottleneck associated with the boundary $F_{\text{S}}^{(\text{B})}$. After a congested pattern is formed at the bottleneck, the flow rate downstream of the congested bottleneck called discharge flow rate $q_{\text{out}}^{(\text{bottle})}$ (figure 3(*b*)) is usually smaller than the initial flow rate $q_{\max}^{(\text{free B})}$. The difference $\delta q(q_{\text{on}}) = q_{\max}^{(\text{free B})}(q_{\text{on}}) - q_{\text{out}}^{(\text{bottle})}(q_{\text{on}})$ called 'capacity drop' is an increasing function of q_{on} at the boundary $F_{\text{S}}^{(\text{B})}$ in the diagram of congested patterns.

In accordance with empirical results [7], in the ATD model moving jams do *not* emerge spontaneously in free flow. This is because in all states of free flow critical perturbations required for an F $\to$ S transition are considerably smaller than those for F $\to$ J transition. In the model, all synchronized flow states that are above the line J in the flow–density plane (figure 1(*c*)) are metastable ones against wide moving jam emergence.

5. Phase transitions and congested patterns in SA models

5.1. Nucleation and metastability effects of pattern formation

As the ATD model, the SA models exhibit a first-order F $\to$ S transition at the bottleneck, which is accompanied by nucleation and metastability effects, as well as by a hysteresis in SP emergence and dissolution. To illustrate these effects found for both the ATD and SA models, we restrict a consideration to the SA model (35)–(41) (figures 6 and 7). When an initial state at the bottleneck is free flow in which q_{in} is given and q_{on} increases gradually, then, as in the ATD model (figure 4(*a*)), a dynamic disturbance in free flow localized at the bottleneck appears spontaneously. A time averaging of spatial speed and density distributions within the perturbation leads to the associated deterministic perturbation (figures 6(*a*) and (*b*)). Deterministic perturbation features are the same as those for the ATD model (section 4.1).

The speed $v_{\text{free}}^{(\text{B})}$ within the deterministic perturbation decreases when q_{on} increases (from point 1 to 5 in figure 6(*c*)). Consequently, $\rho_{\text{free}}^{(\text{B})}$ increases. In the flow–density plane, the flow rate on the main road associated with this density increases too (from point 1 to 5 in figure 6(*d*)), whereas the flow rate upstream of the perturbation is equal to q_{in}, i.e., it does not change (from point 1 to 5 in figure 6(*e*)). This increase in the deterministic perturbation amplitude with q_{on} has a limit $q_{\text{on}} = q_{\text{on}}^{(\text{determ,FS})}$ associated with the deterministic F $\to$ S transition (dotted downarrow in figure 6(*c*)). However, non-homogeneous free flow dynamics (section 4.1), which is caused by vehicle merging, results in an F $\to$ S transition at a smaller q_{on} (point 5 in figure 6(*c*)) related to a point on the boundary $F_S^{(\text{B})}$ in the diagram of congested patterns (figure 8(*a*)). The speed decreases and density increases abruptly within the initial perturbation (arrows F $\to$ S from point 5 to 5′ in figures 6(*c*)–(*e*)) and a congested pattern emerges at the bottleneck. In the example, a widening SP (WSP) occurs upstream of the bottleneck due to the F $\to$ S transition (figure 8(*c*)).

If now q_{on} decreases, the speed within the WSP increases (from point 5′ to 3′ in figures 6(*c*)–(*e*)). This synchronized flow speed increase has a limit: the speed increases and density decreases abruptly within the synchronized flow (arrows S $\to$ F from the point 3′ to 3 in figures 6(*c*)–(*e*)) and free flow returns at the bottleneck.

Note that q_{in} in figure 6 is chosen to be greater than the threshold flow rate q_{th} for moving SP (MSP) existence. As a result, the initial motionless downstream front of synchronized flow at the bottleneck begins to move away upstream. Consequently, an MSP emerges (figure 8(*d*))

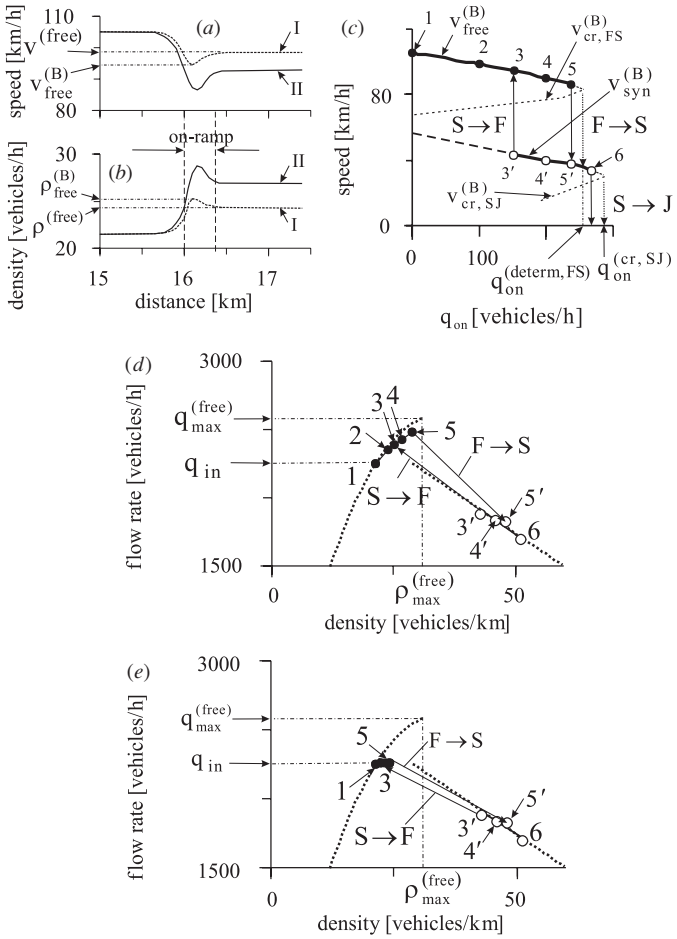

Figure 6. Nucleation and metastability of free flow and synchronized flow: (*a*, *b*) spatial dependence of the speed (*a*) and density (*b*) within the deterministic perturbation in free flow at the bottleneck for two flow rates $q_{\rm on} = 150$ (curve I) and 240 (curve II) vehicles $\rm h^{-1}$. (*c*) Double Z-characteristic in the flow-rate–speed plane for the sequence of F → S → J transitions. (*d*, *e*) Hysteresis effects in the flow–density plane due to F → S and reverse S → F transitions within the deterministic perturbation (*d*) and on the main road upstream of the bottleneck (*e*). Dotted curves in (*d*, *e*) are related to steady-state model solutions. $q_{\rm in} = 2252$ vehicles $\rm h^{-1}$.

(range of $q_{\rm on}$ within which MSPs occur is shown by a dashed part of the synchronized flow states $v_{\rm syn}^{\rm (B)}$ in figure 6(*c*)). At greater $q_{\rm on}$ on the dashed part of the synchronized flow states $v_{\rm syn}^{\rm (B)}$ in figure 6(*c*) this free flow at the bottleneck can persist for a short time only: a new F → S transition occurs spontaneously and a new MSP emerges at the bottleneck and so on. Due to this effect, a sequence of MSPs appears.

At $q_{\rm in} > q_{\rm th}$ an MSP can also be induced by applying a short-time local perturbation in free flow. The speed within this external perturbation should be lower than the critical speed $v_{\rm cr,FS}^{\rm (B)}$ associated with the critical branch on the Z-characteristic for the F → S and reverse S → F transitions at the bottleneck. As in the stochastic models [7], this Z-characteristic consists of the states for free flow associated with the deterministic perturbation at the bottleneck $v_{\rm free}^{\rm (B)}$, the critical branch $v_{\rm cr,FS}^{\rm (B)}$, and synchronized flow states $v_{\rm syn}^{\rm (B)}$ (figure 6(*c*)).

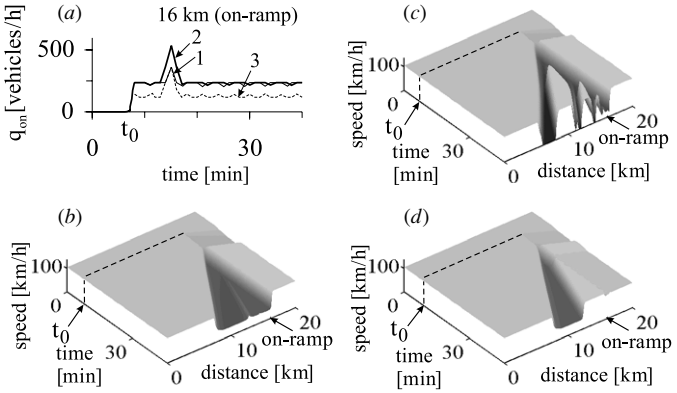

Figure 7. Congested pattern excitation in metastable free flow: (*a*) short-time (2 min) perturbations in the flow rate q_{on} used for pattern excitation. (*b–d*) WSP (*b*), GP (*c*) and MSP (*d*) induced by the related perturbations (curves 1, 2, and 3 in (*a*), respectively). $q_{in} = 2252$ vehicles h^{-1}. $q_{on} = 220$ vehicles h^{-1} in (*b*, *c*) and $q_{on} = 140$ vehicles h^{-1} in (*d*). Amplitudes of perturbations in (*a*) are: 270 for (*b*, *d*) and 500 vehicles h^{-1} for (*c*).

In accordance with this Z-characteristic, we get the associated hysteresis effects on the fundamental diagram (arrows F $\to$ S and S $\to$ F in figures 6(*d*) and (*e*)).

If in contrast q_{on} increases, the speed within the WSP decreases (figure 6(*c*)). This speed decrease has a limit associated with the flow rate $q_{on} = q_{on}^{(cr,SJ)}$ at which an S $\to$ J transition must occur (dotted downarrow S $\to$ J in figure 6(*c*)). However, because there are speed and density waves of a finite amplitude in synchronized flow, an S $\to$ J transition occurs already for $q_{on} < q_{on}^{(cr,SJ)}$ (point 6 and solid downarrow in figure 6(*c*)). As a result, a GP emerges. This is because in the SA models, synchronized flow steady states with the speed $v > v_{cr}^{(SJ)}$, which are above the line J in the flow–density plane, are metastable ones against wide moving jam emergence. This metastability can be seen from another Z-characteristic in the speed–flow plane associated with an S $\to$ J transition in synchronized flow. The Z-characteristic consists of the states for synchronized flow $v_{syn}^{(B)}$, the critical branch for critical perturbations in synchronized flow $v_{cr,SJ}^{(B)}$, and the line $v = 0$ for wide moving jams (figure 6(*c*)).

From the resulting double Z-characteristic (figure 6(*c*)), it can concluded that in a metastable free flow at the bottleneck (left of the boundary $F_S^{(B)}$ in the diagram in figure 8(*a*)) depending on amplitude of a time-limited perturbation caused, for example, by an increase in q_{on} (curves 1 and 2 in figure 7(*a*)), either a WSP (figure 7(*b*)) or a GP (figure 7(*c*)) can be induced. At smaller q_{on} (curve 3 in figure 7(*a*)), an MSP (figure 7(*d*)) can be excited in free flow. All results presented in figures 6 and 7 for the SA model remain qualitatively equal for the ATD model.

5.2. Comparison of congested patterns in ATD and SA models

The SA model (36)–(41) (figure 8) exhibits the following shortcoming in comparison with the ATD model (figure 3):

(i) If the flow rate q_{on} is within a flow rate range $q_{on}^{(FSJ)} < q_{on} < q_{on}^{(LSP)}$, then no SP can be formed at the boundary $F_S^{(B)}$ in the diagram (figure 8(*a*)): the sequence of F $\to$ S $\to$ J transitions occurs spontaneously at this boundary, leading to GP emergence. For this

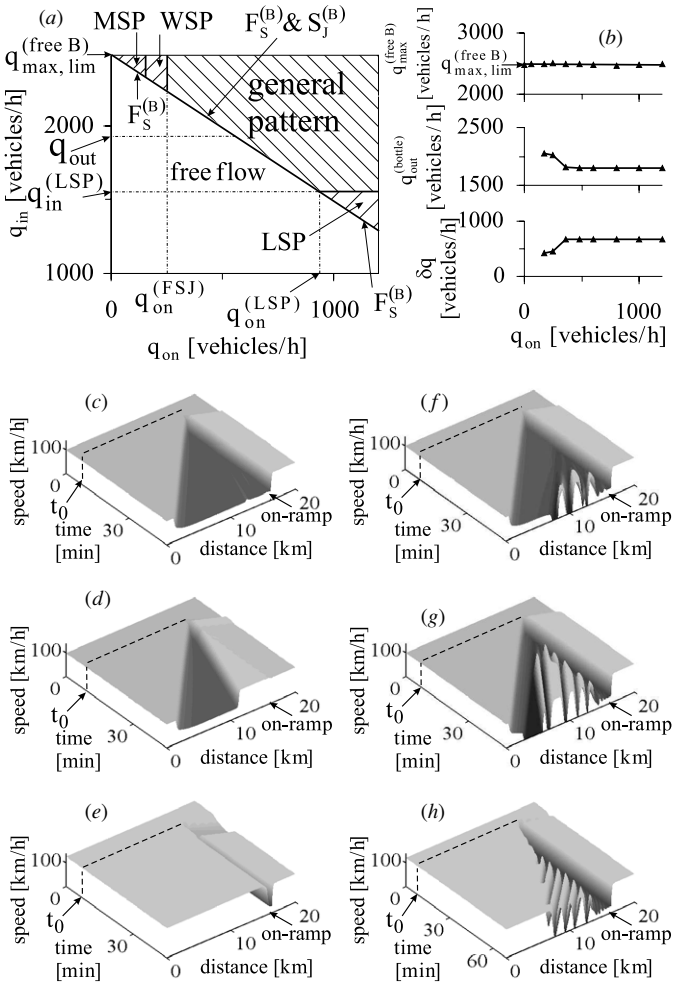

Figure 8. Diagram of congested patterns at the on-ramp bottleneck in SA model (35)–(41) (*a*), the maximum capacity in free flow $q_{max}^{(free\,B)}$, the discharge flow rate at the on-ramp $q_{out}^{(bottle)}$ and the capacity drop δq (*b*), and congested patterns (*c–h*) related to (*a*): (*c–e*) SPs and (*f–h*) GPs. (*c*) WSP. (*d*) MSP. (*e*) LSP. (*f*) GP arising from WSP at smaller q_{on}. (*g*) GP at $q_{in} > q_{out}$. (*h*) GP at $q_{in} < q_{out}$. In (*c–h*), the flow rates (q_{on}, q_{in}) are: (*c*) (200, 2353), (*d*) (140, 2378), (*e*) (940, 1520), (*f*) (250, 2353), (*g*) (400, 2353) and (*h*) (900, 1770) vehicles h^{-1}. $q_{max,lim}^{(free\,B)} \approx 2475$ vehicles h^{-1}.

reason, the related part of the boundary at which GPs emerge spontaneously in free flow at the bottleneck is labelled as $F_S^{(B)}$ and $S_J^{(B)}$.

(ii) If the flow rate q_{in} at this boundary decreases, another characteristic flow rate $q_{in} = q_{in}^{(LSP)}$ associated with the flow rate $q_{on} = q_{on}^{(LSP)}$ at this boundary is reached: at $q_{in} < q_{in}^{(LSP)}$ moving jams do not emerge in synchronized flow upstream of the bottleneck. As a result, at $q_{in} < q_{in}^{(LSP)}$ and right of the boundary $F_S^{(B)}$ only an LSP remains at the bottleneck. Within this LSP the speed is very low. This LSP has a qualitative different nature in comparison with an LSP of higher synchronized flow speed in the ATD model that occurs at considerably greater q_{in} (figure 3).

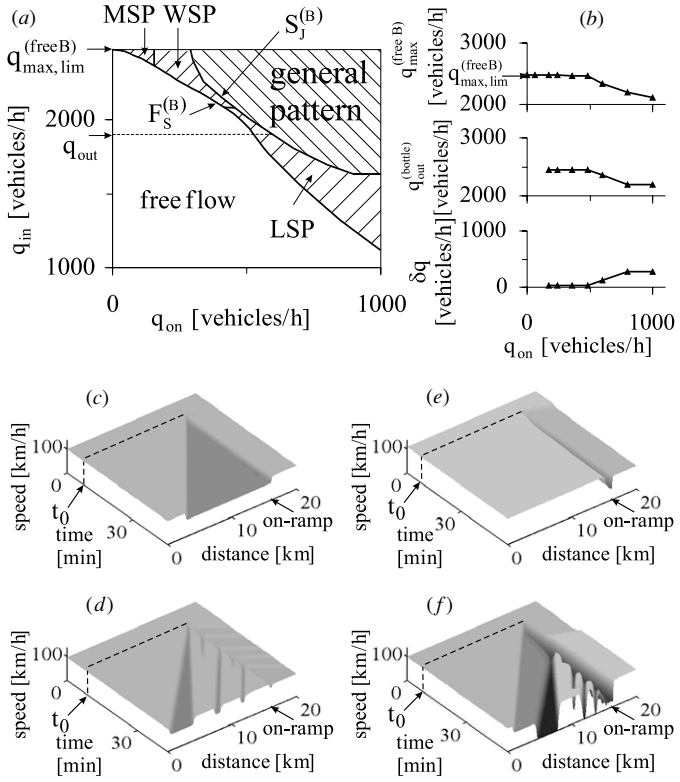

Figure 9. Diagram of congested patterns at the on-ramp bottleneck in the SA model (35)–(40), (42) (*a*), the maximum capacity in free flow $q_{max}^{(free\ B)}$, the discharge flow rate at the on-ramp $q_{out}^{(bottle)}$ and the capacity drop δq (*b*), and congested patterns (*c*–*f*) related to (*a*): (*c*–*e*) SP and (*f*) GP. (*c*) WSP. (*d*) Subsequence of MSPs. (*e*) LSP. (*f*) GP arising from WSP. In (*c*–*f*), the flow rates (q_{on}, q_{in}) are: (*c*) (200, 2400), (*d*) (90, 2450), (*e*) (360, 2115), (*f*) (300, 2400) vehicles h^{-1}. $q_{max,lim}^{(free\ B)} \approx 2475$ vehicles h^{-1}. In (*b*), $q_{out}^{(bottle)}$ is changed from 2450 to 2200 vehicles h^{-1}. $q_{out} \approx 1880$ vehicles h^{-1}.

In the SA model (35)–(40), (42), the branch for average synchronized flow states $V_{av}^{(syn)}$ has a part with a positive slope (figure 2(*d*)). Then LSPs of higher speeds appear in the diagram of congested patterns (figure 9). However, these LSPs are not related to LSPs observed in empirical observations. To explain this, note that these model LSPs are very narrow ones (figure 9(*e*)). They are localized within the merging region of the on-ramp and consist of two narrow fronts only (figure 10(*a*)): there is no region of synchronized flow between the fronts within these LSPs. This is regardless of the flow rates q_{in} and q_{on}. Conflictingly, in empirical observations rather than such narrow LSPs, an extended region of synchronized flow is usually observed within an empirical LSP. The LSP width (in the longitudinal direction) changes over time considerably. These empirical features of LSPs shown by the ATD model (figure 10(*b*)) are not found in the SA model.

(iii) In the ATD model (figures 10(*e*) and (*f*)) as in empirical observations, both free and synchronized flows can be formed between wide moving jams within a GP. In contrast, in the SA models only free flow can be formed between wide moving jams within the GP (figures 10(*c*) and (*d*)). The reason for this is as follows. The average branch for synchronized flow lies for speeds $v > v_{cr}^{(SJ)}$ above the line J (figure 2(*g*)). Flow states in

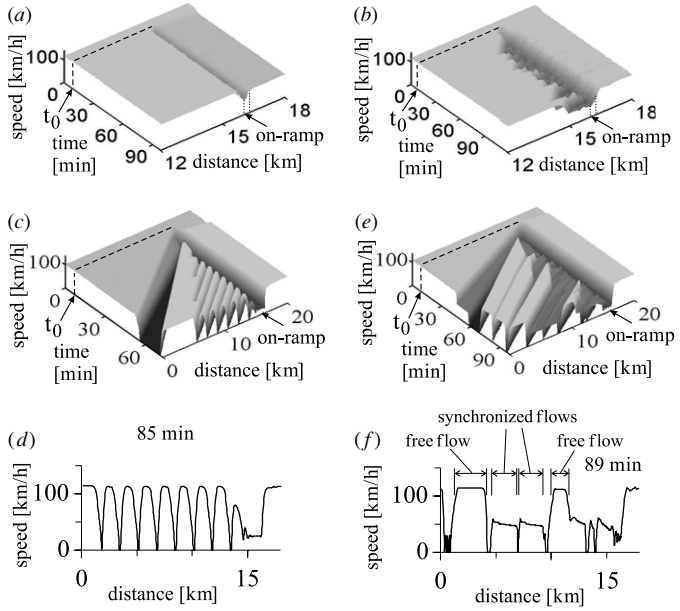

Figure 10. Comparison of LSPs (*a*, *b*) and GPs (*c*–*f*) in the SA model (35)–(40), (42) (*a*, *c*, *d*) and in the ATD model (*b*, *e*, *f*). LSPs and their parameters are the same as those in figures 9(*e*) and 3(*e*), respectively. In (*e*, *f*), $\Delta v_{\mathrm{r}}^{(1)} = 7.5$ m s^{-1}. (q_{on}, q_{in}) are: (*c*, *d*) (380, 2397), (*e*, *f*) (660, 2222) vehicles h^{-1}.

the jam outflow should be related to points on the line J. Thus, there are no synchronized flow states between the jams. This explains why only free flow can be formed between the jams in the SA models[8].

The mentioned shortcoming of the SA model results from the averaging of a 2D region of steady states for synchronized flow in the flow–density plane of the ATD model (figure 1(*a*)) to the branch for average synchronized flow states (curve S in figure 2).

At chosen SA model parameters the condition

$$q_{\mathrm{out}} > q^{(\mathrm{pinch})} \tag{48}$$

is satisfied, where $q^{(\mathrm{pinch})}$ is the flow rate within the pinch region of a GP in which narrow moving jams emerge. Under condition (48), no DGPs appear in the SA models (figures 8 and 9). At other parameters of the SA models, an opposite condition

$$q_{\mathrm{out}} < q^{(\mathrm{pinch})} \tag{49}$$

can be satisfied. Then DGPs appear in the SA models.

The maximum flow rate in free flow downstream of the bottleneck $q_{\mathrm{max}}^{(\mathrm{free\,B})}(q_{\mathrm{on}})$, the discharge flow rate $q_{\mathrm{out}}^{(\mathrm{bottle})}$ and the 'capacity drop' δq can sometimes exhibit different features as those in the ATD model (figure 3(*b*)) when the flow rate q_{on} changes (figures 8(*b*) and 9(*b*)).

[8] The only exclusion is the SA model with the average branch for synchronized flow (42), if parameters for the curve S and/or the line J, i.e., for wide moving jam propagation are chosen different from those as shown in figures 2(*c*), (*d*) and (*h*): these different parameters should lead to an intersection of the line J with the average branch for synchronized flow with a positive slope in the flow–density plane. However, in this specific case only one state of the synchronized flow, which is associated with the point of the latter intersection, is possible. This model effect is not agreed with empirical results, in which the flow rate and speed between wide moving jams within GPs can change over time considerably [7].

Particularly, in contrast with the ATD model, in the SA model (35)–(41) $q_{\max}^{(\text{free B})}$ does not depend on q_{on}, whereas in the SA model (35)–(40), (42) $q_{\max}^{(\text{free B})}$ depends on q_{on} but at considerably greater q_{on} than for the ATD model.

Simulations show that the SA models presented in appendix B qualitatively show the same features of the phase transitions and spatiotemporal congested patterns as those in the SA model (35)–(41).

5.3. Comparison with stochastic SA models

It is interesting to compare the deterministic SA models with possible *stochastic* SA models. Such models can be derived from the stochastic model of [35, 37], if 2D region of synchronized flow steady states is averaged to synchronized flow states related to a 1D region in the flow–density plane.

A stochastic SA model can easily be derived from the stochastic model of [37] based on the physics and ideas for the SA-model approach discussed in section 3. To reach this goal, in the part of the stochastic model of [37]

$$v_{n+1} = \max(0, \min(v_{\text{free}}, v_{\text{c,n}}, v_{\text{s,n}})), \tag{50}$$

$$x_{n+1} = x_n + v_{n+1}\tau \tag{51}$$

for a desired speed in synchronized flow $v_{\text{c,n}}$, rather than the formula (3) of [37] leading to a 2D region of synchronized flow steady states in the flow–density plane, the following equations associated with the physics of the SA models of section 3 are used:

$$v_{\text{c,n}} = v_n + \max(-b_n\tau, \min(a_n\tau, \Delta_n)), \tag{52}$$

$$\Delta_n = \begin{cases} A^{(\text{free})}(g_n)(v_{\text{free}} - v_n) + K(v_{\ell,n} - v_n) & \text{at} \quad g \geqslant g_{\min}^{(\text{free})}, \\ A^{(\text{syn})}\left(V_{\text{av}}^{(\text{syn})}(g_n) - v_n\right) + K(v_{\ell,n} - v_n) & \text{at} \quad g < g_{\min}^{(\text{free})}. \end{cases} \tag{53}$$

In (50)–(53), v_n and x_n are the speed and space coordinate of a vehicle; the index n corresponds to the discrete time $t = n\tau$, $n = 0, 1, 2, \ldots$; τ is the time step; v_{free} is the maximum speed in free flow, which is a constant; $v_{s,n}$ is the safe speed of [37]; $a_n \geqslant 0$ is acceleration, $b_n \geqslant 0$ is deceleration, which are taken as the same stochastic functions used in the model of [37]; the space gap $g_n = x_{\ell,n} - x_n - d$; the average speed in synchronized flow steady states $V_{\text{av}}^{(\text{syn})}$ is given by the formula (41) at $g_{\max}^{(\text{jam})} = 0$. Of course, other formulations for the average synchronized flow steady states $V_{\text{av}}^{(\text{syn})}$, for example, used in the deterministic SA models (figures 2(*b*), (*d*) and (*f*)) can also be applied.

In general, descriptions of random vehicle acceleration and deceleration are the same as those in the stochastic model of [35, 37]: at the first step, the preliminary speed $\tilde{v}_{n+1}$ is set to $\tilde{v}_{\text{n+1}} = v_{n+1}$ where the speed v_{n+1} is calculated from the equations (50)–(53). At the second step, a noise component ξ_n is added to the calculated speed $\tilde{v}_{n+1}$. Then the final speed is found from the condition [37]

$$v_{n+1} = \max(0, \min(v_{\text{free}}, \tilde{v}_{n+1} + \xi_n, v_n + a_{\max}\tau, v_{s,n})), \tag{54}$$

where $a_{\max}$ is the maximum acceleration.

However, in contrast with the stochastic model of [37], the noise component ξ_n in (54) is chosen to be different from zero *only* if the vehicle decelerates, specifically

$$\xi_n = \begin{cases} -b_{\max}\tau\theta(p_{\text{b}} - r) & \text{if } \tilde{v}_{n+1} < v_n - \delta \\ 0 & \text{otherwise,} \end{cases} \tag{55}$$

where $r = \text{rand}(0, 1)$, $\theta(z) = 0$ at $z < 0$ and $\theta(z) = 1$ at $z \geqslant 0$, $\delta \ll \tau a_{\max}$, $b_{\max}$, p_{b} are constants.

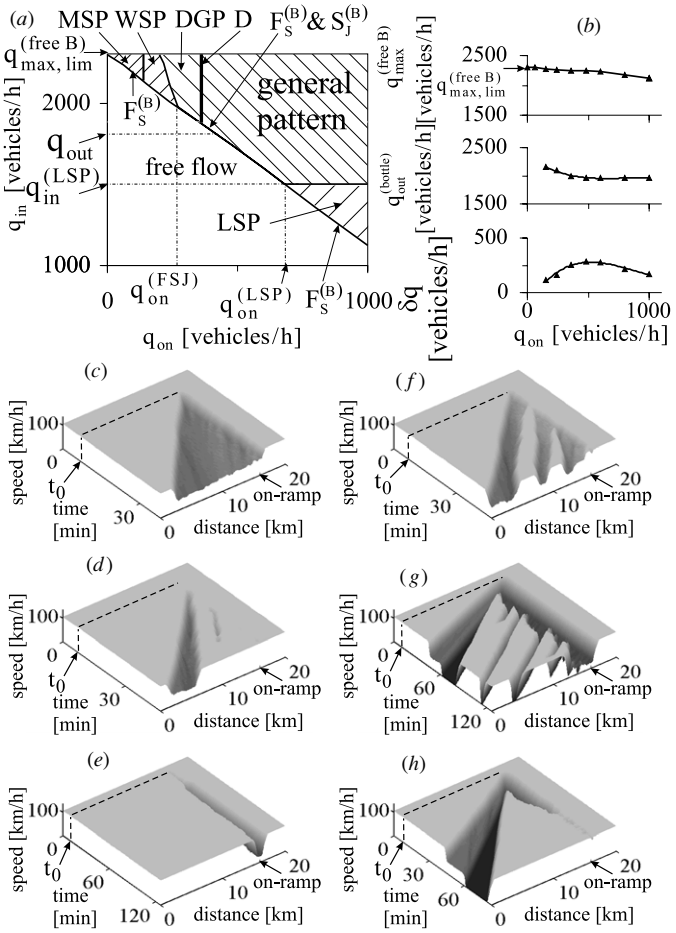

Figure 11. Diagram of congested patterns at the on-ramp bottleneck in the stochastic SA model (50)–(55) (*a*), the maximum capacity in free flow $q_{\max}^{(\text{free B})}$, the discharge flow rate at the on-ramp $q_{\text{out}}^{(\text{bottle})}$ and the capacity drop δq (*b*), and congested patterns (*c*–*h*) related to (*a*): (*c*–*f*) SPs and (*g*, *h*) GPs. (*c*) WSP. (*d*) MSP. (*e*) LSP. (*f*) ASP. (*g*) GP. (*h*) DGP. In (*c*–*h*), the flow rates $(q_{\text{on}}, q_{\text{in}})$ are: (*c*) (200, 2250), (*d*) (60, 2250), (*e*) (720, 1470), (*f*) (120, 2235), (*g*) (500, 2220) and (*h*) (352, 2235) vehicles h^{-1}. $q_{\max,\text{lim}}^{(\text{free B})} \approx 2300$ vehicles h^{-1}. Model parameters: $A^{(\text{free})}(g_n) = 0.5\min(1, (g_n - g_{\min}^{(\text{free})})/20)$, $g_{\min}^{(\text{free})} = 36$ m, $A^{(\text{syn})} = 0.1$, $T_{\text{av}}^{(\text{syn})} = 1.45$ s, $g_{\max}^{(\text{jam})} = 0$, $p_{\text{b}} = 0.02$. $p_1 = 0.55$, $p_2(v_n) = 0.5 + 0.48\theta(v_n - 15)$, $a_{\max} = b_{\max} = 0.5$ ms^{-2}, K = 1. The other parameters are the same as those in [37]. The on-ramp model from [37, 38] is used, in which the synchronization gaps $G_n^+ = G_n^- = g_{\min}^{(\text{free})}$.

Simulations show that the stochastic SA model (50)–(55) qualitatively exhibits similar spatiotemporal congested patterns at the on-ramp bottleneck (figure 11) as those in the associated deterministic SA models (figure 8). However, there are qualitative differences in the dynamics of first-order F → S and S → J transitions leading to pattern formation explained in section 5.1: in the stochastic SA model, random model fluctuations are important for phase transition nucleation, whereas in the deterministic SA models the F → S and S → J transitions are nucleated by dynamic perturbations emerging within the on-ramp merging region.

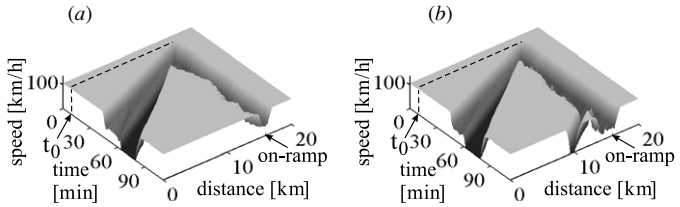

Figure 12. Transformation of DGP in figure 11(*h*) to GP with a very long time interval between wide moving jam emergence (*b*) through a DGP with a quasi-steady LSP (*a*) by a small increase in the flow rate $q_{\rm on}$ in a neighbourhood of the boundary D in the diagram of congested patterns in figure 11(*a*). The flow rates ($q_{\rm on}$, $q_{\rm in}$) are: (*a*) (357, 2235) and (*b*) (362, 2235) vehicles $\rm h^{-1}$.

Note that under the chosen model parameters in the stochastic SA model (50)–(55) condition (49) can be satisfied at smaller $q_{\rm on}$. As a result, there is a region in the diagram of congested patterns in which DGPs occur (region labelled DGP in figure 11(*a*)). After the wide moving jam of the DGP is upstream of the bottleneck as well as in the ATD model, an LSP remains at the bottleneck (figure 11(*h*)). In contrast with the ATD model, this LSP exists for a finite time interval only and free flow returns at the bottleneck.

However, in a small neighbourhood of the boundary labelled D in the diagram, which separates DGPs and GPs, there is a peculiarity in pattern formation under condition (49): if the flow rate $q_{\rm on}$ increases in the neighbourhood of the boundary D in the diagram, then the lifetime of an LSP, which occurs within a DGP increases and it tends towards infinity at the boundary D (figure 12(*a*)). This quasi-steady LSP is explained by a very long interval between wide moving jam emergence in the synchronized flow at the bottleneck (figure 12(*b*)). This interval tends towards the infinity at the boundary D.

6. Discussion

6.1. Comparison of ATD and SA models with OV models and other deterministic models

The first term in the formula for vehicle acceleration $\tilde{a}^{\rm (free)}$ (12), $A(V^{\rm (free)}(g) - v)$, describes the dynamics of the speed v in the vicinity of the optimal speed $V^{\rm (free)}(g)$ in free flow. At a time scale that is considerably greater than the time delay τ, this dynamic behaviour is the same as those in different OV models [3–5, 29], which can be written as follows:

$$\frac{{\rm d}v}{{\rm d}t} = A(g, v)(V(g) - v). \tag{56}$$

However, in (56) the vehicle acceleration $A(g, v)(V(g) - v)$ is valid for the whole possible space-gap range [3–5, 29]

$$g \geqslant 0. \tag{57}$$

In contrast with the OV models, in the ATD model this vehicle acceleration is applied for large space gaps (5) associated with free flow *only*.

The crucial difference of the ATD model with the OV models and all other deterministic microscopic traffic flow models (see references in the reviews [1, 3–6]) is that the vehicle acceleration behaviour qualitatively changes when the vehicle is within the synchronization gap, i.e., if the condition (4), which is opposite to the condition (5), is satisfied.

Condition (4) is associated with the synchronized flow phase in which there is *no* optimal speed in the ATD model. This conclusion follows from (13) and its analysis made in

section 2.8 in which it has been shown that for a given steady space gap in synchronized flow there are an infinity of steady vehicle speeds within a finite speed range (figure 1(*a*)).

The concept of safe speed $v_s(g, v_\ell)$ for vehicle collision prevention used in the ATD model is qualitatively different from the concept of optimal speed that is the basis of the deterministic approaches (2) and (3): the optimal speed is a desired one (this explains the term 'optimal' speed) for a driver to be reached (the driver moves comfortable with the optimal speed during a long time), whereas the safe speed is not an optimal one but a limiting speed that is still permitted (the driver should not move with this speed during a long time because this is strain for the driver and, therefore, non-comfortable). The qualitative difference of these two concepts is mathematically reflected in the dynamic model behaviour. In the ATD model, when the vehicle speed is higher than the safe speed and safe deceleration is applied, then a driver time delay is equal to a small driver reaction time: $\tau = \tau_s$ (21). In all other driving situations, which are not associated with safe speed, driver time delays are different from τ_s. This is because these driver time delays are associated mostly with qualitatively different expected events occurring within different traffic phases (section 2.5). As a result, in the ATD model driver deceleration to the safe speed occurs considerably quicker than in other driving situations. In contrast, in accordance with the concept of optimal speed, in OV models there is a driver time delay in deceleration that characterizes speed relaxation to the optimal speed [1, 4, 5, 27–29, 40].

The crucial differences between the SA models and all other traffic flow models in which steady states covering a one-dimensional region(s) in the flow–density plane (see references in the reviews [1, 3–6]) are as follows. In contrast with the models of [1, 3–6], in the SA models at each density of free flow states the critical amplitude of a local perturbation required for an F $\rightarrow$ S transition is considerably smaller than the critical amplitude of a local perturbation required for an F $\rightarrow$ J transition.

In the SA models, there are two ranges of model steady states separated from one another by a model discontinuity in vehicle space gap or in speed (figures 2(*a*)–(*e*)) or else due to instability of model steady states against infinitesimal non-homogeneous fluctuations (figure 2(*f*)) This simulates the hypothesis of three-phase traffic theory about a competition between over-acceleration and speed adaptation effect that is responsible for F $\rightarrow$ S and S $\rightarrow$ F transitions: the first range of steady states simulates free flow, whereas the second simulates synchronized flow. To simulate S $\rightarrow$ J transitions within synchronized flow, steady states associated with synchronized flow of higher speeds are metastable with respect to moving jam emergence, i.e., moving jams emerge in these synchronized flow states *only* if large enough amplitude local perturbations appear; synchronized flow states of lower speeds are unstable with respect to moving jam emergence. These requirements to the SA models lead to F $\rightarrow$ S $\rightarrow$ J transitions that are responsible for moving jam emergence found in empirical data [7].

6.2. Critical discussion of theories and models based on the fundamental diagram approach

In the OV model (56), as in other deterministic (and stochastic) traffic flow models in the context of the fundamental diagram approach reviewed in [1, 3–6], which claim to show spontaneous jam emergence, there is a range of the density on the fundamental diagram in which steady states on this diagram are unstable against infinitesimal perturbations[9]. This instability leads to wide moving jam emergence in these models both on homogeneous road and

[9] It should be noted that some of the models based on the fundamental diagram approach are not valid far from equilibrium. It is not simply a matter of a phase transition type that a model exhibits in steady conditions, but of the difficulty of closing equations, which should work in unsteady conditions by relations valid only in steady uniform conditions.

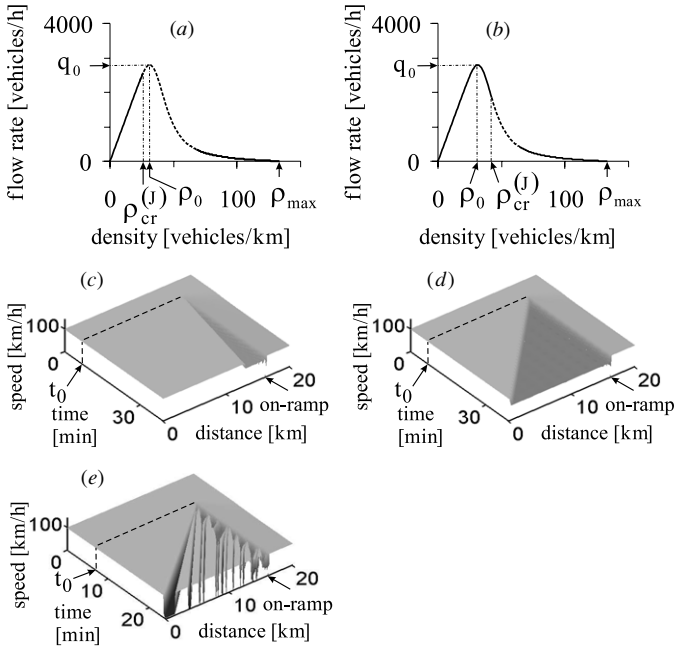

Figure 13. Traffic patterns in OV models at an on-ramp bottleneck: (*a*, *b*) fundamental diagrams of OV models when the condition (58) (*a*) and (59) (*b*) are satisfied, respectively. (*c*, *d*) Widening patterns of dense flow upstream of the bottleneck at two flow rates $q_{\rm on} = 160$ vehicles $\rm h^{-1}$ (*c*) and 400 vehicles $\rm h^{-1}$ (*d*). (*e*) Formation of wide moving jams within the dense flow upstream of the bottleneck at $q_{\rm on} = 700$ vehicles $\rm h^{-1}$. In (*c*–*e*), the flow rate $q_{\rm in} = 2676$ vehicles $\rm h^{-1}$. Parts (*c*)–(*e*) are related to the OV model (*b*) at $V(g) = V_0(\tanh((g - g_0)/g_1) + \tanh(g_0/g_1))$ at $V_0 = 14$ m $\rm s^{-1}$, $g_0 = 17$ m, $g_1 = 7$ m, the sensitivity $A(g, v) = A(v)$ in (56) is $A(v) = 5\rm s^{-1}$ at $v \geqslant 12$ m $\rm s^{-1}$ and $A(v) = 0.9\rm s^{-1}$ at $v < 12$ m $\rm s^{-1}$.

at a bottleneck. We denote the minimum density of this density range, in which infinitesimal fluctuations grow, by $\rho_{\rm cr}^{\rm (J)}$ (figures 13(*a*) and (*b*)).

There are two possibilities for the arrangement of the point of this instability $\left(\rho_{\rm cr}^{\rm (J)}, q_{\rm cr}^{\rm (J)}\right)$ on the fundamental diagram in the OV model and other models in the context of the fundamental diagram approach:

(i) The point $\left(\rho_{\rm cr}^{\rm (J)}, q_{\rm cr}^{\rm (J)}\right)$ lies left of the maximum point of the fundamental diagram (ρ_0, q_0), i.e., on the branch of the diagram with a positive slope (figure 13(*a*)):

$$\rho = \rho_{\rm cr}^{\rm (J)} < \rho_0. \tag{58}$$

(ii) The point $\left(\rho_{\rm cr}^{\rm (J)}, q_{\rm cr}^{\rm (J)}\right)$ lies right of the maximum point of the fundamental diagram (ρ_0, q_0), i.e., on the branch of the diagram with a negative slope (figure 13(*b*)):

$$\rho = \rho_{\rm cr}^{\rm (J)} > \rho_0. \tag{59}$$

Note that in both cases (i) and (ii) all states on the fundamental diagram, in which the density satisfies the condition

$$\rho_{\rm min} \leqslant \rho < \rho_{\rm cr}^{\rm (J)}, \tag{60}$$

where $\rho_{\rm min}$ is the density in the wide moving jam outflow associated with the flow rate $q_{\rm out}$, are metastable states with respect to moving jam emergence [44]. The case (i) has intensively been considered in the literature [3–6] and criticized in section 3.3.2 of the book [7].

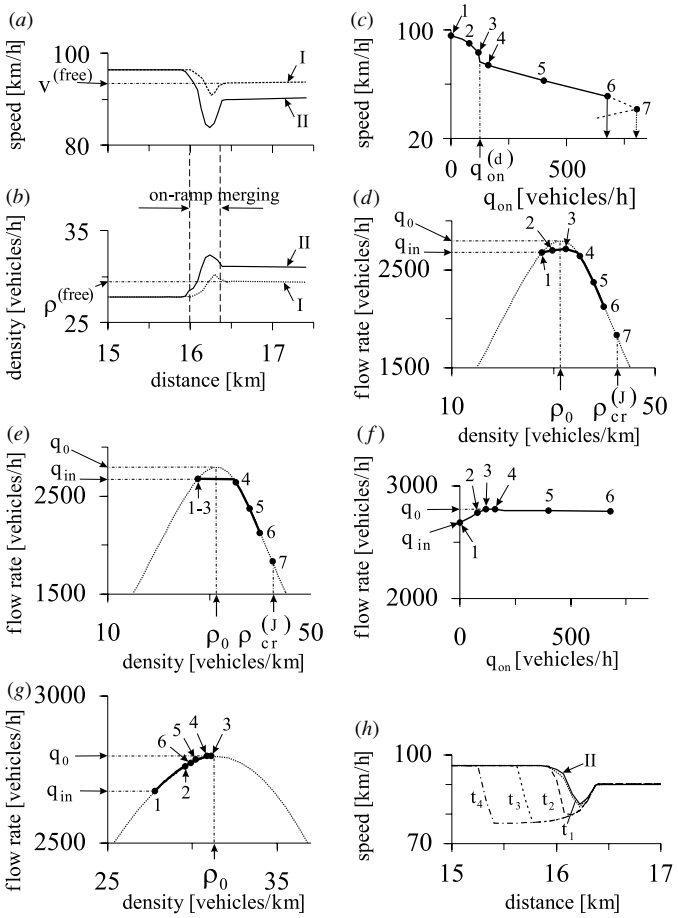

Figure 14. Pattern features in the OV model in figure 13(*b*) at an on-ramp bottleneck: (*a*, *b*) spatial dependences of the averaged speed (*a*) and density (*b*) on the main road within deterministic perturbations localized at the bottleneck at $q_{\rm on} = 80$ (curve I) and 117 (curve II) vehicles $\rm h^{-1}$. (*c*, *d*) On-ramp flow rate dependences of the average speed (*c*), flow rate and density (*d*) on the main road at locations of the minimum of the average speed. (*e*) On-ramp flow rate dependences of the flow rate and density on the main road at the location 200 m upstream of the begin of the on-ramp merging region. (*f*, *g*) On-ramp flow rate dependences of the flow rate (*f*) and density (in the flow–density plane) (*g*) on the main road in free flow downstream of the bottleneck. (*h*) Wave of dense flow that starts to propagate upstream with the velocity $v_{\rm d} \approx -0.7$ km $\rm h^{-1}$ under condition (62) at a very small value $\Delta q = 3$ vehicles $\rm h^{-1}$; $t_1 = 15, t_2 = 25, t_3 = 55, t_4 = 90$ min; $q_{\rm on} =$ 120 vehicles $\rm h^{-1}$; the deterministic perturbation (curve II) is the same as those in (*a*). Dotted curves on (*d*, *e*, *g*) show steady model states in the flow–density plane. $q_{\rm in} = 2676$ vehicles $\rm h^{-1}$, $q_0 = 2795$ vehicles $\rm h^{-1}$. 5 min averaged data.

In the case (ii) (figure 13(*b*)), the flow rate in free flow downstream of an on-ramp bottleneck $q_{\rm sum}$ cannot exceed the maximum flow rate on the fundamental diagram q_0. If $q_{\rm in}$ is a given large enough value and the flow rate $q_{\rm on}$ begins to increase, then a localized perturbation as that in the ATD and SA models (sections 4.1 and 5.1) appears at the bottleneck (figures 14(*a*) and (*b*)). The minimum speed within the time-averaged (deterministic) perturbation on the main road decreases when $q_{\rm on}$ increases (points 1–3 in figure 14(*c*)). In the OV model, when $q_{\rm on}$ increases, the location of the minimum speed within the deterministic perturbation on the

main road exhibits firstly a slight shift downstream and then upstream within the merging region of the on-ramp (figures 14(*a*) and (*b*)). For this reason, the flow rate on the main road at the location of the minimum speed within the deterministic perturbation on the main road firstly slightly increases and then decreases (points 1–3 in figure 14(*d*)), whereas the flow rate on the main road upstream of the on-ramp merging region is equal to $q_{\rm in}$ (points 1–3 in figure 14(*e*)). When $q_{\rm on}$ increases beginning from zero, the flow rate $q_{\rm sum}$ downstream of the bottleneck increases beginning from $q_{\rm sum} = q_{\rm in}$ (points 1–3 in figures 14(*f*) and (*g*)).

At a given large enough flow rate $q_{\rm in}$ this growth of the local perturbation in free flow at the bottleneck with $q_{\rm on}$ has a limit. This limit is reached when the flow rate $q_{\rm on}$ reaches some critical value $q_{\rm on} = q_{\rm on}^{\rm (d)}$ at which the flow rate $q_{\rm sum}$ is equal to the maximum flow rate on the fundamental diagram:

$$q_{\rm sum} = q_{\rm in} + q_{\rm on}^{\rm (d)} = q_0. \tag{61}$$

When the flow rate $q_{\rm on}$ increases further, i.e.,

$$\Delta q = q_{\rm in} + q_{\rm on} - q_0 > 0, \tag{62}$$

then the upstream front of the initial perturbation, which is motionless at the condition $q_{\rm sum} = q_{\rm in} + q_{\rm on} \leqslant q_0$ (curve II in figure 14(*h*)), begins to move upstream of the bottleneck, i.e., a wave of lower speed and greater density propagating upstream appears (spatial speed distributions related to the times t_1–t_4 in figure 14(*h*)). As a result, a dense flow associated with the branch of the diagram with a negative slope occurs upstream of the bottleneck (figures 13(*c*) and (*d*) and points 4–6 in figures 14(*c*)–(*e*)). At the critical point (61), the derivative of the minimum average speed on the main road on the flow rate $q_{\rm on}$ is discontinuous, whereas this speed is a continuous decreasing function of $q_{\rm on}$ (figure 14(*c*)). The greater the flow rate $q_{\rm on}$, specifically, the greater the Δq (62), the greater the absolute velocity of the wave of dense flow propagation $|v_{\rm d}|$ (figures 13(*c*) and (*d*)). In addition, the flow rate downstream of the bottleneck, which is equal to q_0 under condition (61), remains approximately to be equal to q_0, when $q_{\rm on}$ increases (points 4–6 in figures 14(*f*) and (*g*)).

It must be noted that the above-mentioned behaviour of the upstream front of the perturbation at the bottleneck in the OV model (56), (59) is qualitatively different from those for the upstream front of the perturbation at the bottleneck in the ATD and SA models. In the latter case, when the flow rate $q_{\rm on}$ reaches the critical value for an F $\rightarrow$ S transition, a wave of synchronized flow occurs abruptly and propagates upstream with a *finite* velocity. This is associated with a first-order F $\rightarrow$ S transition. In contrast, in the OV model there is *no* discontinuous change in the velocity $v_{\rm d}$ when due to an increase in $q_{\rm on}$ the condition (62) is satisfied: $|v_{\rm d}|$ increases continuously beginning from zero, when $q_{\rm on}$ first reaches and then exceeds the critical flow rate $q_{\rm on}^{\rm (d)}$ associated with condition (61). Specifically, we find that if $\Delta q \rightarrow 0$, then $|v_{\rm d}| \rightarrow 0$. Thus, in the OV model there is no first-order phase transition from free flow to dense flow.

The widening dense flow upstream of the bottleneck (figures 13(*c*) and (*d*) and 14(*h*)) can exist only, when the density $\rho_{\rm d}$ in the dense flow satisfies the condition

$$\rho_0 < \rho_{\rm d} < \rho_{\rm cr}^{\rm (J)}. \tag{63}$$

This is because at the density $\rho_{\rm d} = \rho_{\rm cr}^{\rm (J)}$ the dense flow loses its stability against wide moving jam emergence (point 7 and dotted downarrow in figure 14(*c*)). However, dynamic waves that emerge due to vehicle merging at the bottleneck propagate through the dense flow. For this reason, in numerical simulations this moving jam emergence occurs already at the density $\rho_{\rm d} < \rho_{\rm cr}^{\rm (J)}$ (point 6 and solid downarrow in figure 14(*c*)).

The congested patterns in figures 13(*d*) and (*e*) at the first glance resemble a widening SP and a GP, respectively. Indeed, in both cases a dense flow occurs upstream of the bottleneck

whose downstream front is fixed at the bottleneck. Thus, this dense flow should satisfy the macroscopic spatiotemporal objective criteria for the synchronized flow phase (section 1). This conclusion is, however, incorrect. To explain this, note that in *empirical* observations application of the objective criteria, which define the traffic phases in congested traffic, leads to clear distinction of the synchronized flow phase. This synchronized flow exhibits the following fundamental *empirical* feature: an F $\rightarrow$ S transition leading to synchronized flow emergence is a first-order phase transition. In contrast, in a traffic flow *model* an application of the objective criteria does not guarantee that dense flow occurrence in free flow is associated with a first-order phase transition, which is one of the requirements for the synchronized flow phase.

This conclusion concerns the OV model (56), (59) (figure 13(*b*)) as well as other models in the context of the fundamental diagram approach under condition (59). Whereas for the SA model there is a Z-shaped speed–flow characteristic associated with a first-order F $\rightarrow$ S transition in free flow at the bottleneck (figures 6(*c*)–(*e*)), for the OV model the on-ramp flow rate dependence of the speed at the bottleneck is a *monotonous* decreasing function (figure 14(*c*)): there is no first-order phase transition, when a dense flow related to the fundamental diagram with a negative slope is formed upstream of the bottleneck. Thus, the dense traffic flow in the case of the OV model and other models in the context of the fundamental diagram approach under condition (59) does not exhibit the important empirical feature of synchronized flow and, therefore, the dense flow is not associated with the synchronized flow phase.

There are also traffic flow models in the context of the fundamental diagram approach, in which there is no instability of steady model states on the fundamental diagram regardless of the vehicle density. Examples of this model class are as follows: (i) An OV model (56) in which the sensitivity $A(g, v)$ is great enough regardless of v and g. (ii) The Nagel–Schreckenberg cellular automata model in the deterministic model limit, i.e., when probability of model fluctuations in this model is equal zero ($p = 0$) [46]. (iii) The Lighthill–Whitham–Richards model [47] and the associated cell-transmission models [48]. In this model class, traffic patterns at a freeway bottleneck are qualitatively similar as those found in the OV model (56), (59) at the density considerably smaller than the critical density $\rho_{\mathrm{cr}}^{(\mathrm{J})}$ (the patterns associated with points 1–5 in figure 14(*c*)–(*g*)). These common model features are as follows: 1) the local perturbation at the bottlenecks at $\Delta q < 0$ (figures 14(*a*) and (*b*)); 2) widening dense flow upstream of the bottleneck at $\Delta q > 0$ (figures 13(*c*) and (*d*) and 14(*h*)); 3) there is no discontinuous change in speed (no speed breakdown) at the bottleneck when widening dense flow occurs; 4) with an increase in traffic demand at $\Delta q \geqslant 0$, the upstream front velocity of widening dense flow increases continuously beginning from zero. Thus, in this model class, there is no first-order F $\rightarrow$ S transition observed during the onset of congestion at the bottleneck, i.e., this dense flow has no relation to real freeway traffic.

6.3. Conclusions

(i) Two different deterministic microscopic traffic flow model classes in the context of three-phase traffic theory, the ATD and SA models, have been introduced in the paper.
(ii) The ATD and SA models reproduce important *empirical* spatiotemporal features of phase transitions in traffic flow and congested traffic patterns.
(iii) In contrast with all other known deterministic microscopic traffic flow models, in the ATD and SA models vehicles moving in free flow and vehicles moving in synchronized flow exhibit qualitatively different dynamic behaviour. This is a result of the introduction of two separated regions of steady-state model solutions for free flow and synchronized

flow in the ATD and SA models as well as different dynamic rules of vehicle motion in free flow and synchronized flow implemented in the models.

(iv) As in empirical observations, there is a first-order phase transition in the ATD and SA models from free flow to synchronized flow that explained the onset of congestion at bottlenecks in these models.

(v) The nature of the onset of congestion as a first-order F → S transition in free flow at the bottleneck, which the ATD and SA models show, is also associated with metastability of free flow at the bottleneck against external short-time disturbances in this flow in a neighbourhood of the bottleneck. As a result, there is multiple congested pattern emergence in an initial free flow at the bottleneck in the ATD and SA models: depending on an amplitude (or duration) of an external disturbance, one of the SPs or else a GP can be induced in free flow at the bottleneck at the same chosen model parameters.

(vi) In accordance with empirical results, in the ATD and SA models moving jams can emerge spontaneously in synchronized flow only, i.e., as a result of F → S → J transitions.

(vii) In addition to the above common behaviour of the ATD and SA models, these models also exhibit some qualitatively different features. This is because in the ATD model synchronized flow model steady states are related to a 2D region in the flow–density plane, whereas synchronized flow model steady states in the SA models belong to a 1D region (a curve) in the flow–density plane. In particular, the following differences of model features have been found:

  (1) The ATD model can show all types of spatiotemporal congested patterns at an on-ramp bottleneck observed in empirical observations.

  (2) In contrast, SA models cannot show LSPs associated with empirical results as well as some of empirical features of synchronized flow between wide moving jams within GPs.

(viii) Models in the context of the fundamental diagram approach reviewed in [1–6] cannot explain the onset of congestion in free flow, which in empirical observations is associated with a first-order F → S transition. Depending on the model type and model parameters, in these models either wide moving jam emergence is responsible for the onset of congestion at an on-ramp bottleneck rather than an empirically observed F → S transition or a widening dense traffic flow occurs upstream of the bottleneck when the density in free flow at the bottleneck exceeds the density associated with the maximum point on the fundamental diagram. In the latter case, in contrast with empirical observations there is *no* first-order phase transition from an initial free flow to this dense flow at the bottleneck: the dense flow results from non-homogeneity of a freeway in a neighbourhood of the bottleneck. Thus, these models cannot show a first-order F → S transition observed during the onset of congestion at the bottleneck in real freeway traffic, i.e., this dense flow has no relation to real freeway traffic. Indeed, the first-order F → S transition is a fundamental empirical feature of the onset of congestion in free flow with the subsequent synchronized flow phase emergence at the bottleneck.

Appendix A

To derive formula (18) [45], let us consider a solution of (16) when it is an equality:

$$v_{\rm s}(g, v_\ell) = \frac{2b_{\rm s}g + v_\ell^2}{b_{\rm s}T_{\rm s} + \sqrt{b_{\rm s}^2T_{\rm s}^2 + 2b_{\rm s}g + v_\ell^2}}. \tag{A.1}$$

From (16), (A.1), it can be seen that if $g = v_\ell T_s$, then the safe speed $v_s = v_\ell$; if in contrast $g < v_\ell T_s$, then the speed $v_s < v_\ell$. In particular, this ensures collisionless vehicle motion. To simplify formula (A.1), let us replace the space gap g in denominator of (A.1) by the value $v_\ell T_s$. This reduces the safe speed v_s at $g < v_\ell T_s$, therefore, the safety condition (16) remains valid. Then from formula (A.1), we get

$$v_s(g, v_\ell) = \frac{g + v_\ell^2/(2b_s)}{T_s + v_\ell/(2b_s)}. \tag{A.2}$$

To provide more comfortable vehicle deceleration, an anticipated gap $g^{(a)} = g + (v_\ell - v)T_0$ is used in formula (A.2) rather than the gap g. As a result, (A.2) takes the form

$$v_s(g, v_\ell) = \frac{g + (v_\ell - v)T_0 + v_\ell^2/(2b_s)}{T_s + v_\ell/(2b_s)}. \tag{A.3}$$

Substituting (A.3) into (15), we find formula (18) with coefficients (19), (20).

Note that we have also tested another formulation for the safe speed in the ATD model when the speed $v_s(g, v_\ell)$ in (15) is given by formula (A.1). Simulations of the ATD model show that both formulations (18) and (A.1) ensure collisionless vehicle motion at an appropriate choice of model parameters and lead to qualitatively the same features of phase transitions and congested patterns.

Appendix B

In this appendix, two further variants of the SA models are presented. In the first of these variants, the formula (36) reads as follows:

$$\frac{dv}{dt} = \begin{cases} a^{(\text{free})} & \text{at} \quad g \geqslant g_{\min}^{(\text{free})}, \\ a^{(\text{syn})} & \text{at} \quad g_{\max}^{(\text{jam})} < g < g_{\min}^{(\text{free})}, \\ a^{(\text{jam})} & \text{at} \quad 0 \leqslant g \leqslant g_{\max}^{(\text{jam})}, \end{cases} \tag{B.1}$$

where $a^{(\text{free})}$, $a^{(\text{syn})}$, $a^{(\text{jam})}$ are given by (37)–(41). In this SA model, steady states of free flow (the curve F in figure 2(e)) correspond to condition (43), averaged steady states of synchronized flow are related to a line S given by the condition

$$q = \left(1 - \rho/\rho_{\min}^{(\text{jam})}\right)/T_{\text{av}}^{(\text{syn})} \qquad \text{at} \quad \rho_{\max}^{(\text{free})} < \rho < \rho_{\min}^{(\text{jam})}, \tag{B.2}$$

steady states for a wide moving jam are associated with the condition (47) (figure 2(e)).

In another variant of SA model, formula (36) reads as follows:

$$\frac{dv}{dt} = \begin{cases} a^{(\text{free})} & \text{at} \quad g \geqslant g_{\min}^{(\text{free})}, \\ a^{(\text{FS})} & \text{at} \quad g_{\max}^{(\text{syn})} < g < g_{\min}^{(\text{free})}, \\ a^{(\text{syn})} & \text{at} \quad g_{\max}^{(\text{jam})} < g \leqslant g_{\max}^{(\text{syn})}, \\ a^{(\text{jam})} & \text{at} \quad 0 \leqslant g \leqslant g_{\max}^{(\text{jam})}. \end{cases} \tag{B.3}$$

In (B.3), $g_{\max}^{(\text{syn})}$ is the maximum space gap in synchronized flow; $a^{(\text{jam})}$, $a^{(\text{free})}$ and $a^{(\text{syn})}$ are given by (37) in which $\tilde{a}^{(\text{jam})}$ is taken from (40),

$$\tilde{a}^{(\text{free})}(g, v, v_\ell) = A^{(\text{free})}(V^{(\text{free})}(g) - v) + K^{(\text{free})}(v_\ell - v), \tag{B.4}$$

$$\tilde{a}^{(\text{syn})}(g, v, v_\ell) = A^{(\text{syn})}\left(V_{\text{av}}^{(\text{syn})}(g) - v\right) + K^{(\text{syn})}(v, v_\ell)(v_\ell - v), \tag{B.5}$$

where the sensitivity

$$K^{(\text{syn})}(v, v_\ell) = \begin{cases} K^{(\text{acc})} & \text{at} \quad v < v_\ell, \\ K^{(\text{dec})} & \text{at} \quad v \geqslant v_\ell, \end{cases} \tag{B.6}$$

$K^{(\mathrm{free})}$ is a sensitivity, $V_{\mathrm{av}}^{(\mathrm{syn})}(g)$ is given by (41). A function $a^{(\mathrm{FS})}(g, v, v_\ell)$ in (B.3) is taken as follows:

$$a^{(\mathrm{FS})}(g, v, v_\ell) = \min(a_{\max}, A^{(\mathrm{FS})}(V^{(\mathrm{FS})}(g) - v) + K^{(\mathrm{free})}(v_\ell - v)), \tag{B.7}$$

where the function $V^{(\mathrm{FS})}(g) = V(g)$ at $g_{\max}^{(\mathrm{syn})} < g < g_{\min}^{(\mathrm{free})}$, $A^{(\mathrm{FS})}$ is a sensitivity that in a general case can be different from the sensitivity $A^{(\mathrm{free})}$ in free flow.

In the SA model (B.3)–(B.7), steady states of free flow (the curve F in figure 2(f)) are associated with condition (43), averaged steady states of synchronized flow are related to a line S given by the condition

$$q = \left(1 - \rho \big/ \rho_{\min}^{(\mathrm{jam})}\right) \big/ T_{\mathrm{av}}^{(\mathrm{syn})} \qquad \text{at} \quad \rho_{\min}^{(\mathrm{syn})} \leqslant \rho < \rho_{\min}^{(\mathrm{jam})}, \tag{B.8}$$

where $\rho_{\min}^{(\mathrm{syn})} = 1 \big/ \left(g_{\max}^{(\mathrm{syn})} + d\right)$, and steady states for a wide moving jam are found from the condition (47) (figure 2(f)).

In contrast with the other SA models, the SA model (B.3)–(B.7) has a limited density range of steady states between steady states for free flow and synchronized flow, which are found from the condition

$$a^{(\mathrm{FS})} = 0 \qquad \text{at} \quad g_{\max}^{(\mathrm{syn})} < g < g_{\min}^{(\mathrm{free})}. \tag{B.9}$$

Equation (B.9) yields the following condition for these steady states in the flow–density plane (curve FS in figure 2(f))

$$q = \rho V_{\mathrm{F}}(\rho) \qquad \text{at} \quad \rho_{\max}^{(\mathrm{free})} < \rho < \rho_{\min}^{(\mathrm{syn})}, \tag{B.10}$$

where the density $\rho_{\max}^{(\mathrm{free})}$ at the maximum point for free flow is not greater than the density ρ_0 associated with the maximum point on the curve FS (figure 2(f)). To simulate a first-order F $\to$ S transition, the steady-state model solutions (B.10) should be *unstable* against infinitesimal non-homogeneous perturbations. This requirement to the SA model (B.3)–(B.7) is easy satisfied through an appropriated choice of the function $V^{(\mathrm{FS})}(g)$ and the sensitivities $A^{(\mathrm{FS})}$, $K^{(\mathrm{free})}$, $K^{(\mathrm{syn})}$. In this case, numerical simulations of the SA model (B.3)–(B.7) made show that this model exhibits F $\to$ S $\to$ J transitions in accordance with empirical results (we used the following parameters for the SA model (B.3)–(B.7): $V(g) = V_0(\tanh((g - g_0)/g_1) + \tanh(g_0/g_1))$ with $V_0 = 14$ m s^{-1}, $g_0 = 21$ m, $g_1 = 7$ m; $g_{\max}^{(\mathrm{syn})} = 24$ m; $A^{(\mathrm{free})} = A^{(\mathrm{FS})} = 0.1$ s^{-1}; $K^{(\mathrm{free})} = 0.6$ s^{-1}; $K^{(\mathrm{acc})} = 0.4$ s^{-1}; $K^{(\mathrm{dec})}$ is taken from table 1 with $K_1^{(\mathrm{dec})} = 1\mathrm{s}^{-1}$, $v_{\mathrm{c}} = 9$ m s^{-1}, $\epsilon = 0.05$; other parameters are the same as those in the SA model (35)–(41)).

References

[1] Gartner N H, Messer C J and Rathi A (ed) 1997 *Special Report 165: Revised Monograph on Traffic Flow Theory* (Washington, DC: Transportation Research Board)

[2] Wolf D E 1999 *Physica* A **263** 438

[3] Chowdhury D, Santen L and Schadschneider A 2000 *Phys. Rep.* **329** 199

[4] Helbing D 2001 *Rev. Mod. Phys.* **73** 1067–141

[5] Nagatani T 2002 *Rep. Prog. Phys.* **65** 1331–86

[6] Nagel K, Wagner P and Woesler R 2003 *Oper. Res.* **51** 681–716

[7] Kerner B S 2004 *The Physics of Traffic* (Berlin: Springer)

[8] Lesort J-B (ed) 1996 Transportation and traffic theory *Proc. 13th Int. Symp. on Transportation and Traffic Theory* (Oxford: Elsevier)

[9] Ceder A (ed) 1999 Transportation and traffic theory *Proc. 14th Int. Symp. on Transportation and Traffic Theory* (Oxford: Elsevier)

[10] Taylor M A P (ed) 2002 Transportation and traffic theory in the 21st century *Proc. 15th Int. Symp. Transportation and Traffic Theory* (Amsterdam: Elsevier)

[11] Wolf D E, Schreckenberg M and Bachem A (ed) 1995 Traffic and granular flow *Proc. Int. Workshop on Traffic and Granular Flow (October 1995)* (Singapore: World Scientific)
[12] Schreckenberg M and Wolf D E (ed) 1998 Traffic and granular flow '97 *Proc. Int. Workshop on Traffic and Granular Flow (October 1997)* (Singapore: Springer)
[13] Helbing D, Herrmann H J, Schreckenberg M and Wolf D E (ed) 2000 Traffic and granular flow '99 *Proc. Int. Workshop on Traffic and Granular Flow (October 1999)* (Heidelberg: Springer)
[14] Fukui M, Sugiyama Y, Schreckenberg M and Wolf D E (ed) 2003 Traffic and granular flow '01 *Proc. Int. Workshop on Traffic and Granular Flow (October 2001)* (Heidelberg: Springer)
[15] Hoogendoorn S P, Luding S, Bovy P H L, Schreckenberg M and Wolf D E (ed) 2005 Traffic and granular flow '03 *Proc. Int. Workshop on Traffic and Granular Flow (October 2003)* (Heidelberg: Springer)
[16] Payne H J 1971 *Mathematical Models of Public Systems* vol 1, ed G A Bekey (La Jolla, CA: Simulation Council)
Payne H J 1979 *Transp. Res. Record* **772** 68
[17] Kühne R 1991 *Highway Capacity and Level of Service* ed U Brannolte (Rotterdam: A A Balkema) p 211
[18] Kerner B S and Konhäuser P 1993 *Phys. Rev.* E **48** 2335–8
[19] Klar A, Kühne R and Wegener R 1996 *Surv. Math. Ind.* **6** 215–39
[20] Bellomo N, Delitala M and Coscia V 2002 *Math. Models Meth. Appl. Sci.* **12** 1801–44
[21] Darbha S and Rajagopal K R 2001 *Syst. Control Lett.* **43** 387–401
[22] Darbha S and Rajagopal K R 2001 *Math. Probl. Eng.* **7** 379–92
[23] Herman R, Montroll E W, Potts R B and Rothery R W 1959 *Oper. Res.* **7** 86–106
[24] Gazis D C, Herman R and Rothery R W 1961 *Oper. Res.* **9** 545–67
[25] Nagatani T and Nakanishi K 1998 *Phys. Rev.* E **57** 6415
[26] Lubashevsky I, Wagner P and Mahnke R 2003 *Eur. Phys. J.* B **32** 243–7
[27] Newell G F 1961 *Oper. Res.* **9** 209
[28] Whitham G B 1990 *Proc. R. Soc.* A **428** 49
[29] Bando M, Hasebe K, Nakayama A, Shibata A and Sugiyama Y 1994 *Japan. J. Appl. Math.* **11** 203
Bando M, Hasebe K, Nakayama A, Shibata A and Sugiyama Y 1995 *Phys. Rev.* E **51** 1035–1042
[30] Treiber M and Helbing D 1999 *Preprint* cond-mat/9901239
[31] Treiber M, Hennecke A and Helbing D 2000 *Phys. Rev.* E **62** 1805–24
[32] Kerner B S and Rehborn H 1997 *Phys. Rev. Lett.* **79** 4030
[33] Kerner B S 1998 *Phys. Rev. Lett.* **81** 3797
[34] Kerner B S 2002 *Phys. Rev.* E **65** 046138
[35] Kerner B S and Klenov S L 2002 *J. Phys. A: Math. Gen.* **35** L31
[36] Kerner B S, Klenov S L and Wolf D E 2002 *J. Phys. A: Math. Gen.* **35** 9971–10013
[37] Kerner B S and Klenov S L 2003 *Phys. Rev.* **68** 036130
[38] Kerner B S and Klenov S L 2004 *J. Phys. A: Math. Gen.* **37** 8753–88
[39] Kerner B S, Klenov S L and Hiller H 2005 *Preprint* physics/0507094 (http://arxiv.org/abs/physics/0507094)
[40] Davis L C 2004 *Phys. Rev.* E **69** 016108
[41] Lee H K, Barlović R, Schreckenberg M and Kim D 2004 *Phys. Rev. Let.* **92** 238702
[42] Jiang R and Wu Q S 2004 *J. Phys. A: Math. Gen.* **37** 8197–213
[43] Gipps P G 1981 *Trans. Res.* B **15** 105–11
Gipps P G 1986 *Trans. Res.* B **20** 403–14
[44] Kerner B S and Konhäuser P 1994 *Phys. Rev.* E **50** 54–83
[45] Kerner B S and Klenov S L 2005 *Preprint* physics/0507120 (http://arxiv.org/abs/physics/0507120)
[46] Nagel K and Schreckenberg M 1992 *J. Phys. (France) I* **2** 2221
[47] Lighthill M J and Whitham G B 1955 *Proc. R. Soc.* A **229** 317
Richards P I 1956 *Operations Res.* **4** 42
[48] Daganzo C 1999 Transportation and traffic theory *Proc. 14th Int. Symp. on Transportation and Traffic Theory* ed A Ceder (Oxford: Elsevier) pp 81–104