Angewandte Mathematik und Informatik
Universität zu Köln

Report No. 304

Metastable States in a Microscopic Model of Traffic Flow

by

S. Krauß, P. Wagner, C. Gawron

1997

Pysical Review E **55**, 5597

S. Krauß, P. Wagner
Deutsche Forschungsanstalt für Luft– und Raumfahrt,
Linder Höhe, 51170 Köln
C. Gawron
Zentrum für Paralleles Rechnen,
Universität zu Köln,
50931 Köln

1991 Mathematics Subject Classification: 90B20,60K30,82C70
Keywords: Highway traffic, Statistical processes: Applications, Transport processes

Metastable States in a Microscopic Model of Traffic Flow

S. Krauß, P. Wagner
Deutsche Forschungsanstalt für Luft– und Raumfahrt, Linder Höhe, 51170 Köln, Germany
Zentrum für Paralleles Rechnen, Universität zu Köln, Albertus–Magnus–Platz, 50923 Köln, Germany

C. Gawron
Zentrum für Paralleles Rechnen, Universität zu Köln, Albertus–Magnus–Platz, 50923 Köln, Germany

It is a well known fact that metastable states of very high throughput and hysteresis effects exist in traffic flow, which the simple cellular automaton (CA) model of traffic flow and its continuous generalization fail to reproduce. It is shown that the model can be generalized to give a one–parametric family of models, a part of which reproduces the metastable states and the hysteresis. The models having that property and those not having it are separated by a transition that can be clearly identified.

I. INTRODUCTION

When describing freeway traffic, one is hardly ever interested in the way individual cars move, but rather in the macroscopic properties of the system, meaning the properties that are expressed as probability distributions or averages taken over many cars. Therefore it is obviously inefficient to use very detailed models of individual driver behavior if only the flow–density relation or gap distributions, for example, are to be calculated.

This is the starting point for the cellular automaton model (CA) of traffic flow, first proposed by Nagel and Schreckenberg [1–3]. The model tries to reproduce the macroscopic properties of traffic with the simplest possible microscopic dynamics. Although the motion of individual cars has some unphysical peculiarities, these cancel out, when an average over a sufficient number of cars is performed, and typical phenomena like jamming can be reproduced qualitatively. However, it is not clear whether or not the model is in fact able to assume all the relevant macroscopic states that are found in real traffic flow.

In this work we want to refer to the highly ordered metastable state that can be found at densities shortly below the point where traffic breaks down and jams occur.

The existence of such metastable states manifests itself for example in a discontinuity and the existence of two branches in the fundamental diagram. Unfortunately, the simple CA model does not come up with either one of these properties, so there is definitely a discrepancy between the model and reality even on a macroscopic scale. We will see in this work, how this discrepany can be fixed by generalizing the model in a very natural way. When using the term CA in the sequel, we will no longer distinguish between the original model and its continuous generalization described in [4]. However it should be kept in mind that this work only investigates models discrete in time, but continuous in space whereas the original CA is discrete in both time and space.

II. FINITE SIZE EFFECTS

Before turning to the generalization of the model it should be mentioned that also the simple CA model reveals a discontinuity in the fundamental diagram, if the system is sufficiently small.

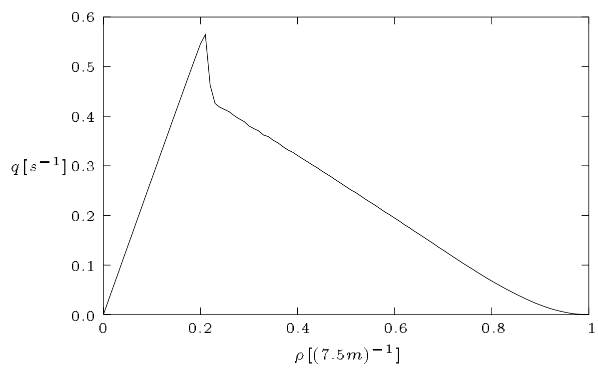

FIG. 1. Discontinuity due to finite–size–effects.

Fig. 1 shows the fundamental diagram for the CA model of a one–lane–ring of approximately $750m$ length. The fundamental diagram clearly reveals a capacity drop at the density of maximum throughput. A qualitative explanation for this can easily be given. As reported in [4] the appearance of traffic jams in partly constrained flow can be viewed as the coexistence of two phases of traffic flow in a dynamic equilibrium. The phases have different densities ρ_j and ρ_f in the jammed and the free phase respectively, which are uniquely determined by equilibrium conditions. The picture of coexisting phases, however, is only valid if both phases contain a "macroscopic" amount of cars.

This is exactly the reason for the finite size effect seen above. If the overall density in the system is only slightly above the equilibrium density of the free phase and the system is sufficiently small, any macroscopic traffic jam would absorb so many cars that the density in the free phase would drop below the equilibrium density ρ_f. So the coexistence of two phases is impossible in this case and the homogeneous state remains stable.

The capacity drop in the CA model is a finite size effect of limited interest. However, we will show that this effect can also appear in infinite systems if the model is only modified slightly. The stable homogeneous state of high throughput in the finite system then corresponds to a metastable state of the infinite system.

III. DRIVING WITH LIMITED DECELERATION

It is very plausible that the states of very high throughput that we try to model must be highly organized in some way. In the simple CA–model and its continuous generalization highly organized states of this kind cannot appear. The reason is that, loosely speaking, individual cars can destroy an ordered state easily by braking very hard (note that there is no limit to the maximum deceleration of the cars in the model).

Of course the unlimited braking capabilities are highly unrealistic, not only because they do not agree with the laws of physics, but rather, because everyday experience tells us that we obviously use a driving strategy that generally allows us to travel hundreds of miles without ever having to brake very hard. So allowing only small values for the deceleration in the model may appear as a sensible thing to do. Note however, that the CA–model of traffic flow is not designed to reproduce individual driver behavior correctly anyway, it rather represents a "minimal" model capable of explaining macroscopic properties of traffic flow. So refining the model can only be justified within this modeling philosophy by showing that the refinements do result in macrosopic effects and that the model can still be considered "minimal" in a proper sense.

Consider now a single lane road, on which vehicles move at speeds between 0 and $v_{\max}$. The velocity is not restricted to integer values. Assume that the system experiences a parallel update for a discrete succession of timesteps. One timestep can be identified with the reaction time of the individual drivers, because the velocity that a driver chooses at timestep t is determined by the velocity his predecessor has chosen at timestep $t-1$.

Now look at two cars following each other: If the second driver chooses the velocity v_{s} during the update, he thus determines the minimum distance d_{s} he has to travel before he can come to a complete stop. His choice will be based on the velocity v_{p} and the corresponding minimum braking distance d_{p} that his predecessor has before the update. If the gap between the cars is g, the condition for a safe velocity choice reads

$$d_{\mathrm{s}} \leq d_{\mathrm{p}} + g \ . \tag{1}$$

Any safe driving strategy has to satisfy this condition.

If no restrictions are put on the way cars can brake, the cars can stop immediately, so $d_{\mathrm{p}} = 0$ and $d_{\mathrm{s}} = v_{\mathrm{s}}$. Note that d_{s} is nonzero, because the successor travels the distance v_{s} before the next update. In this case the safety condition reduces to

$$v_{\mathrm{s}} \leq g \ , \tag{2}$$

which is the well known safety condition of the Nagel–Schreckenberg model. Note that the model is minimal in the sense that nothing but the absence of collisions has been assumed to establish the model.

Now consider the case that the maximum deceleration is limited to a value of b. Calculating the minimum braking distances for the two cars it can be derived in a straightforward way that the maximum safe velocity is given by

$$v_{\mathrm{s}} \leq v_{\mathrm{safe}} = b(\alpha_{\mathrm{safe}} + \beta_{\mathrm{safe}}) \ , \tag{3}$$

where α_{safe} and β_{safe} are given by

$$\alpha_{\mathrm{safe}} = \left\lfloor \sqrt{2\,\frac{d_{\mathrm{p}} + g}{b} + \frac{1}{4}} - \frac{1}{2} \right\rfloor \ , \tag{4}$$

$$\beta_{\mathrm{safe}} = \frac{d_{\mathrm{p}} + g}{(\alpha_{\mathrm{safe}} + 1)b} - \frac{\alpha_{\mathrm{safe}}}{2} \ . \tag{5}$$

Here $\lfloor x \rfloor$ denotes the integer part of the real number x. The braking distance d_{p} of the car ahead is given by

$$d_{\mathrm{p}} = b\left(\alpha_{\mathrm{p}}\beta_{\mathrm{p}} + \frac{\alpha_{\mathrm{p}}\,(\alpha_{\mathrm{p}} - 1)}{2}\right) \ , \tag{6}$$

where α_{p} and β_{p} are defined as the integer and the fractional part of v_{p}/b. The somewhat strange appearance of these formulas is a consequence of the fact that the system is updated in discrete timesteps. For a derivation the reader is referred to the Appendix.

Condition (3) does not try to model a certain driving strategy, it only states a condition that any strategy has to satisfy if collisions are to be avoided with limited braking capabilities.

Still we will, following a minimalistic modeling philosphy, adopt the above conditions as driving strategies and model deviations from optimal driving in this context as noise. The model corresponding to (3) is a simplified version of the well known Gipps model [5].

We thus acquire a family of models, each model characterized by the ratio $r = b/v_{\max}$. For $r = 1$ the braking rules of the Nagel–Schreckenberg model are resumed. The model is defined as follows:

To reduce the number of free parameters, the maximum acceleration and maximum deceleration were both set equal to b. For clarity the auxiliary variables v_0 and

v_1 are introduced.The variable v_1 denotes the optimal velocity for the next update, $v_1 - v_0$ the maximum deviation from v_1 due to imperfections in driving. $v_{\max}$ denotes the maximum velocity of the cars and v_{safe} the maximum safe velocity according to eq. (3).

The update rules are then given by

$$
\begin{aligned}
v_1 &= min[v(t) + b, v_{max}, v_{\text{safe}}] \ , \\
v_0 &= v_1 - \epsilon\,(v_1 - (v(t) - b)) \ , \\
v(t+1) &= v_{\text{ran},\text{v}_0,\text{v}_1} \ , \\
x(t+1) &= x(t) + v(t+1) \ .
\end{aligned}
\tag{7}
$$

The parameter ϵ was always chosen to be 0.4 for the calculations presented here, $v_{\text{ran},\text{v}_0,\text{v}_1}$ denotes a random number between v_0 and v_1. The unit length corresponds to $7.5m$, one timestep corresponds to $1s$.

Note that the amount by which the velocity of a car is perturbed randomly depends on how much the driver is forced to decelerate. On the one hand this is necessary to guarantee that the maximum deleration b is not exceeded due to random perturbations. On the other hand this feature crudely models the fact, that strong interactions between cars reduce the freedom of individual drivers to choose their velocity.

IV. THE HYSTERESIS

We now study the properties of the model for small values of r. As mentioned before, we expect the existence of highly ordered states, if r is sufficiently small.

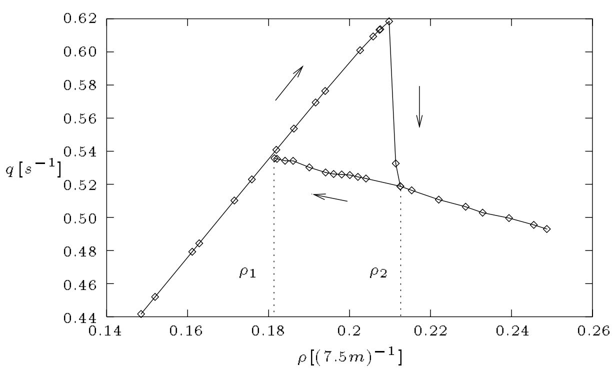

FIG. 2. The branched fundamental diagram

Fig. 2 depicts the fundamental diagram for $r = 1/30$ in the interesting range of densities. The value $r = 1/30$ means that it takes cars at least 30 timesteps to come to a complete stop from maximum velocity. The flow is measured in units of cars per timestep, where one timestep corresponds to roughly one second. The maximum flow therefore corresponds to about 2200 cars/h. The result shown corresponds to the case of an infinite system and was obtained analyzing the scaling of the fundamental diagram with system size, so finite size effects have been eliminated. The model reveals a discontinuous change in throughput at the density of maximum flow ρ_2, just like the CA model did for very small systems.

Also it can be seen that two branches of the fundamental diagram exist between the densities ρ_1 and ρ_2. The upper branch was calculated by adding cars to a homogeneous state, while the lower one was calculated by removing cars from a jammed state and allowing the system to relax after the intervention. In this way a hysteresis loop can be traced (arrows in Fig. 2). A hysteresis effect that appears similar at first glance has been found in the so called cruise control limit of the Nagel–Schreckenberg model [9]. Note however that in the cruise control limit random perturbations are simply switched off for cars traveling at maximum speed. In this way the system can be locked in a deterministic state of high throughput if it is started in an initially homogeneous configuration. Such states are essentially different from the metastable states found here, as will be shown in the chapter on the state of maximum flow.

Note that no hysteresis can exist if the discontinuity in the fundamental diagram is a mere finite size effect. So trying to perform the same procedure as described above for the CA of small system size does not reveal any new branch of the fundamental diagram.

V. OUTFLOW FROM JAMS

In the CA model the outflow from traffic jams is equal to the maximum flow (the capacity). This contrasts with empirical observations [6] as well as hydrodynamical models [7]. In the model proposed here the outflow q_{out} from a jam is lower than the maximum flow, namely

$$
q_{\text{out}} = q(\rho_1) \ . \tag{8}
$$

Fig. 3 shows a jam that is just developing. We see that the outflow region already influenced by the jam has a significantly lower throughput due to a reduced density. Clearly the partial reduction of the flow in the early stages of the evolution of the jam can be distinguished from the final flow reduction.

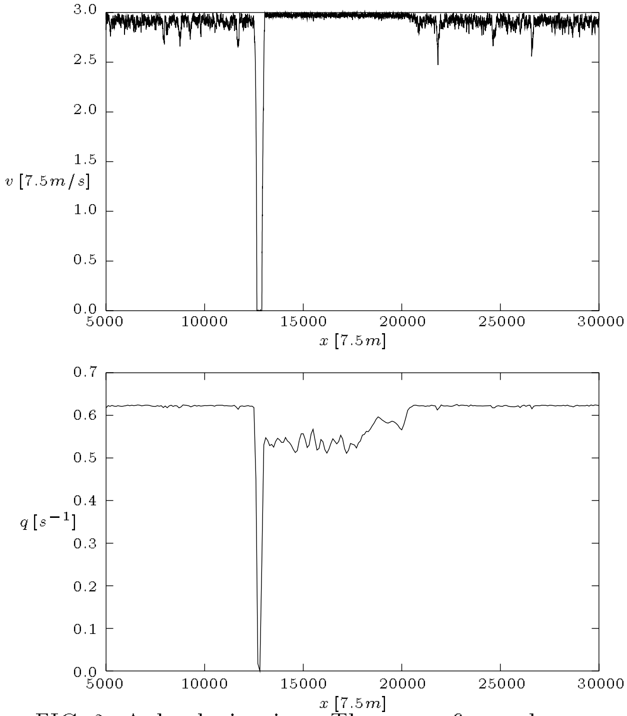

FIG. 3. A developing jam. The upper figure shows the velocity v, the lower one the flow q as a function of the space coordinate x.

This fact has important consequences for the dynamics of cluster formation. Assume that the system is in a state of oversaturation (i.e. $\rho > \rho_1$). The creation of microscopic jams can then be described in an analogous way to an unbiased random walk, just like in the CA [8,9]. This however is no longer true when a microscopic jam is large enough to reduce the outflow significantly. Once this is the case, the jam will grow inevitably, until inflow and outflow become equal again.

When the system is in an equilibrium state again, where inflow and outflow of all jams are equal, the random walk argument holds again, so small jams may eventually dissolve again. However, if this actually happens the cars of a dissolved jam will be swallowed by the next one, increasing its lifetime. So the stable equilibrium state will be a state with very few large jams with correspondingly large lifetimes.

VI. THE STATE OF MAXIMUM FLOW

We will now investigate the states close to the point of maximum flow.

It has been reported [10–12] that a rise in velocity variance is encountered within the free flow shortly before it collapses into a jammed state. Contrasting to this observation, in the CA a rise of the velocity variance is only found due to jamming phenomena.

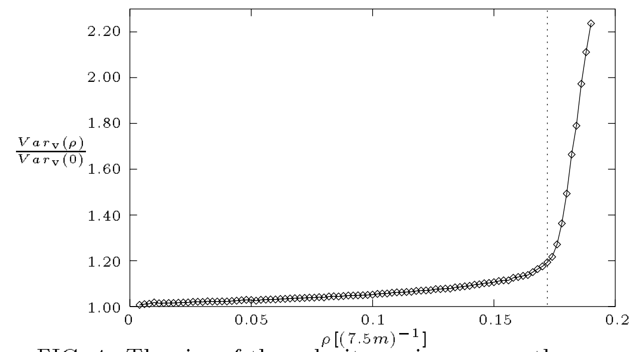

FIG. 4. The rise of the velocity variance near the jamming point (dashed line) in the CA.

In Fig. 4 the ratio of the actual velocity variance and the variance in the free flow due to the randomization step is displayed as a function of the density for the CA model with continuous space coordinates [4]. It can be seen that it rises only by a few percent in the homogeneous flow, before actual jamming occurs (dashed line). At that point the variance remains a continuous function of the density. We observe the same behavior in the model proposed here if r is set to 1. Note however that the case $r = 1$ is not completely equivalent to the CA model due to differences in the way cars accelerate.

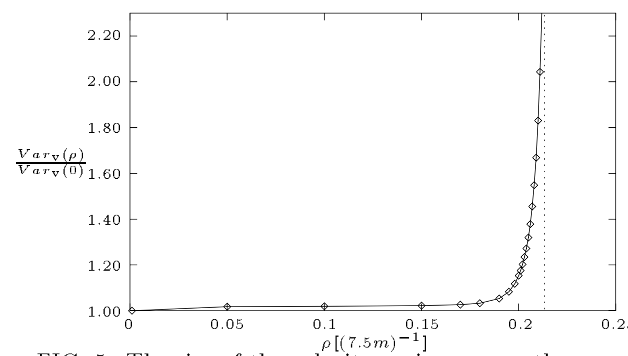

FIG. 5. The rise of the velocity variance near the jamming point (dashed line) for $r = 1/30$.

Now the same property is displayed for the model proposed here with $r = 1/30$ in Fig. 5. We see that the velocity variance rises considerably before any jamming is encountered. At the jamming point the velocity undergoes a discontinuous change and rises due to jamming. However, if the variance is calculated separately for the jammed and the free parts of the system, it drops approximately to the free flow value (see also Fig. 3). The fact that the point of maximum flow is also the point of maximum velocity variance clearly distinguishes our model from the state of maximum flow in the cruise control limit of the Nagel–Schreckenberg model.

The degree of self organization near ρ_2 can be estimated best by measuring the correlation length in the system. Therefore a correlation function G for the velocity fluctuations is defined, that depends on the distance of two cars with respect to their numbering. Note that this distance cannot be identified with any metric distance. So we define

$$G(j) = \frac{1}{N} \sum_{i=1}^{N} (v_i - \langle v \rangle)(v_{i+j} - \langle v \rangle) \quad j \in \mathbb{N}\,, \tag{9}$$

where $\langle v \rangle$ denotes the average velocity in the system and v_i is the velocity of the i-th car. The simulation results show that the correlation function can be represented in the form

$$G(j) \propto \frac{1}{j^{\alpha}} \, \exp\left(-\frac{j}{\xi}\right), \tag{10}$$

where the parameter of interest here is the correlation length ξ.

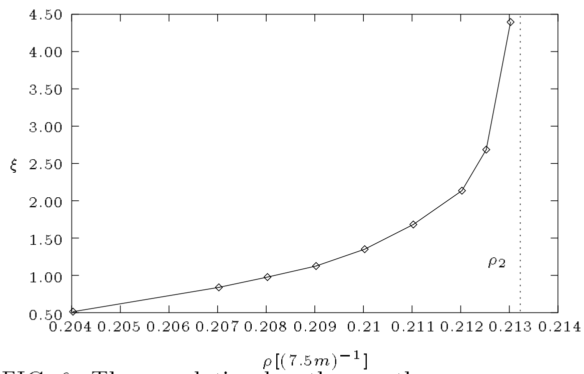

FIG. 6. The correlation length near the jamming point (dashed line)

Fig. 6 shows the correlation length as a function of the density ρ. We see that far from ρ_2 the correlation length practically vanishes whereas it increases significantly near the point where the cars start jamming.

VII. PROPERTIES OF THE MODEL FAMILY

So far we looked at a single model ($r = 1/30$) and investigated the properties of that model near ρ_2. The next natural step to be taken is the investigation of the way the properties at ρ_2 change when the parameter r, which characterizes the model, runs from zero to one.

As we already know from the properties of the CA–model the discontinuity in the fundamental diagram and along with it the ordered states of high flow have to disappear for some r_c between zero and one. A suitable parameter for characterizing the existence of the ordered states is the height of the jump Δq at ρ_2:

$$\Delta q(r) = \lim_{\rho \to \rho_2(r)-} q(\rho) - \lim_{\rho \to \rho_2(r)+} q(\rho)\,. \tag{11}$$

From the definition of r_c we have $\Delta q(r) = 0$ for $r \geq r_c$. When getting close to r_c, preparing homogeneous states becomes difficult and it is a laborious task to calculate lifetimes of such states with sufficient accuracy. Therefore we will limit ourselves to the following procedure:

For each r perfectly homogenous initial configurations are prepared at different densities. The highest density for which a homogenous state can evolve for more than 10^5 timesteps without developing any inhomogeneities, is then considered to be an approximation for $\rho_2(r)$. The flow of that homogeneous state is then compared to the flow of the jammed state that we get if we start the system with the same density and an initially slightly inhomogeneous state.

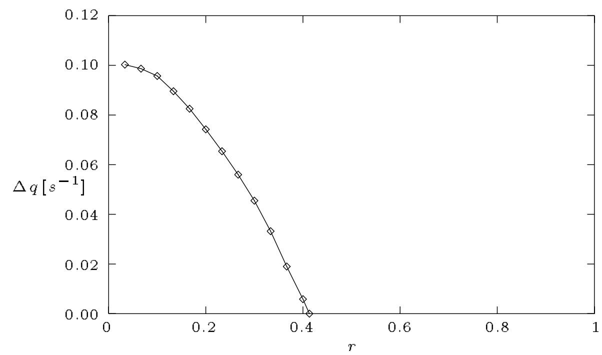

FIG. 7. The discontinuous change in throughput Δq as a function of $r = b/v_{\max}$.

Fig. 7 shows the function $\Delta q(r)$. It can be seen that the ordered states disappear for $r = 0.41$. Note however, that this value is not universal, but depends on details of the parameters used in the model, like the values of $v_{\max}$ and the noise parameter ϵ . Above this threshold the system assumes qualitatively the same properties as the CA–model. The important point about this observation is that the CA model and the models for small r presented here are separated by a transition that changes the macroscopic properties for states of high flow. The quantitative results concerning the way Δq goes to zero have to be considered with caution however, because the very details of the behavior depend on how accurately ρ_2 can be calculated. The algorithm used here certainly allows improvements in that respect.

VIII. CONCLUSIONS

A generalization of the Nagel–Schreckenberg model of traffic flow has been proposed that leads to a one–parametric family of models characterized by a parameter r that determines the braking capabilites of the cars.

It has been found that the macroscopic physical properties of the models within this family are not uniform. When scanning the properties of the models by varying the characteristic parameter r between 0 and 1 a transition is observed at $r = r_c$, where the model behavior for states of very high flow changes qualitatively. Below r_c the outflow from traffic jams is lower than the maximum flow and ordered states with very high flow and large correlation lengths can exist. The fundamental diagram displays a discontinuous change in throughput at the point of maximum flow. Above r_c the discontinuity in the fundamental diagram and the ordered states no longer exist. The outflow from the jams is approximately equal to the maximum flow again.

The changes concerning the outflow from jams have significant impact on the clustering dynamics. Above r_c the system displays small, comparatively short–lived, continuously branching jams, whereas below r_c the system assumes an equilibrium state with only few, but large and stable jams.

Note that we need to introduce considerable noise into the simple CA model using a randomization step to acquire phenomena of cluster formation. The noise, however, reduces the maximum flow considerably, so difficulties in attaining realistic capacities are encountered. In the model proposed here cluster formation also exists for very small artificial noise. In fact we found that the qualitative appearance of the state space pattern is relatively insensitive towards changes of the noise parameter. The value of this parameter mainly determines the relaxation time of the system (details on this topic will be published elsewhere). As a consequence we can easily simulate realistic road capacities of 2200 cars per hour, without losing jamming phenomena.

---

[1] K. Nagel and M. Schreckenberg. *J. Physique I*, 2:2221 (1992).
[2] A. Schadschneider and M. Schreckenberg. *Journal of Physics A*, 26, L679 (1993).
[3] M. Schreckenberg, A. Schadschneider, K. Nagel, and N. Ito, *Phys. Rev. E* **51**, 2939 (1995).
[4] S. Krauß, P. Wagner, and C. Gawron *Phys. Rev. E.*, **54** ,3707 (1996)
[5] P. G. Gipps, *Transpn. Res. B*, **15B**, 105–111 (1981)
[6] B. S. Kerner & H. Rehborn *Phys. Rev. E* ,**53**, 1297 R, (1996)
[7] B. S. Kerner and P. Konhäuser, *Phys. Rev. E* **50**, 54 (1994).
[8] K. Nagel *Int. J. Mod. Phys. C* **5(3)**, 567 (1994)
[9] K. Nagel and M. Paczuski. *Phys. Rev. E*, **51**, 2909 (1995).
[10] D. Helbing, *Phys. Rev. E* **51**, 3164 (1995).
[11] R. D. Kühne, in *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory*, edited by I. Volmuller and R. Hamerslag (VNU Science, Utrecht, 1984)
[12] R. D. Kühne, *Transp. Res. Rec.* **1320**, 251 (1991)

APPENDIX: THE SAFETY CONDITION

In a discrete time-step model, braking is modeled by subtracting in each time step one unit of the decelaration b from the velocity. With the above notation the braking distance of the first car is then given by

$$\begin{aligned} d_\mathrm{p} &= b\left((\alpha_\mathrm{p}+\beta_\mathrm{p}-1)+(\alpha_\mathrm{p}+\beta_\mathrm{p}-2)+\ldots+\beta_\mathrm{p}\right) \\ &= b\left(\alpha_\mathrm{p}\beta_\mathrm{p}+\frac{\alpha_\mathrm{p}\left(\alpha_\mathrm{p}-1\right)}{2}\right). \end{aligned} \tag{12}$$

Similarly, if the second driver chooses the velocity $v_\mathrm{safe}=b(\alpha_\mathrm{safe}+\beta_\mathrm{safe})$, his braking distance becomes

$$\begin{aligned} d_\mathrm{s} &= b\left((\alpha_\mathrm{safe}+\beta_\mathrm{safe})+(\alpha_\mathrm{safe}+\beta_\mathrm{safe}-1)+\ldots+\beta_\mathrm{safe}\right) \\ &= b\left(\left(\alpha_\mathrm{safe}+1\right)\beta_\mathrm{safe}+\frac{\alpha_\mathrm{safe}\left(\alpha_\mathrm{safe}+1\right)}{2}\right). \end{aligned} \tag{13}$$

If the expression is inserted into the safety condition, the resulting equation can be solved formally for α_safe to give

$$\alpha_\mathrm{safe}=f(\beta_\mathrm{safe})\ , \tag{14}$$

where the function $f(\beta)$ is given by

$$f(\beta)=\sqrt{2\,\frac{d_\mathrm{p}+gap}{b}+\left(\beta-\frac{1}{2}\right)^2}-\left(\beta+\frac{1}{2}\right)\ . \tag{15}$$

We know that α_safe is a nonnegative integer and β_safe a nonnegative real number smaller than 1. So the fact that $f(\beta)$ is a decreasing function of β and $f(0)-f(1)=1$ immediately yields

$$\alpha_\mathrm{safe}=\lfloor f(0)\rfloor\ . \tag{16}$$

β_safe can then be found from eq. (14).