Probabilistic description of traffic breakdowns

Reinhart Kühne*
German Aerospace Center, Institute of Transport Research, Rutherfordstraße 2, 12489 Berlin, Germany

Reinhard Mahnke†
Fachbereich Physik, Universität Rostock, D-18051 Rostock, Germany

Ihor Lubashevsky‡
Theory Department, General Physics Institute, Russian Academy of Sciences, Vavilov Street 38, Moscow 119991, Russia

Jevgenijs Kaupužs
Institute of Mathematics and Computer Science, University of Latvia, 29 Rainja Boulevard, LV-1459 Riga, Latvia
(Received 10 November 2001; published 26 June 2002)

We analyze the characteristic features of traffic breakdown. To describe this phenomenon we apply the probabilistic model regarding the jam emergence as the formation of a large car cluster on a highway. In these terms, the breakdown occurs through the formation of a certain critical nucleus in the metastable vehicle flow, which enables us to confine ourselves to one cluster model. We assume that, first, the growth of the car cluster is governed by attachment of cars to the cluster whose rate is mainly determined by the mean headway distance between the car in the vehicle flow and, maybe, also by the headway distance in the cluster. Second, the cluster dissolution is determined by the car escape from the cluster whose rate depends on the cluster size directly. The latter is justified using the available experimental data for the correlation properties of the synchronized mode. We write the appropriate master equation converted then into the Fokker-Planck equation for the cluster distribution function and analyze the formation of the critical car cluster due to the climb over a certain potential barrier. The further cluster growth irreversibly causes jam formation. Numerical estimates of the obtained characteristics and the experimental data of the traffic breakdown are compared. In particular, we draw a conclusion that the characteristic intrinsic time scale of the breakdown phenomenon should be about 1 min and explain the case why the traffic volume interval inside which traffic breakdown is observed is sufficiently wide.

DOI: 10.1103/PhysRevE.65.066125 PACS number(s): 02.50.Ey, 45.70.Vn, 05.70.Fh, 89.40.+k

I. INTRODUCTION: TRAFFIC BREAKDOWN AS A NUCLEATION PHENOMENON

The spontaneous formation of traffic jams on highways has attracted attention over the last years because of two reasons. The first is the importance of this problem for traffic engineering, especially concerning the feasibility of attaining the limit capacity of traffic networks and quantifying it. The second is the fact that the jam formation nicely exemplifies the existence of various phase states and transitions between them in statistical systems comprising elements with motivated behavior, which is a novel branch of physics. According to the recent notion proposed by Kerner (see, e.g., Refs. [1–3]) based on the experimental data [4–7] the jam formation is of sufficiently complex nature. In particular, it proceeds mainly through the sequence of two phase transitions: free flow (F)$\rightarrow$ synchronized mode (S)$\rightarrow$ stop-and-go pattern (J) [7]. Both of these transitions are of the first order, i.e., they exhibit breakdown, hysteresis, and nucleation effects [6]. The $F\rightarrow J$ transition can occur directly if only the synchronized mode is suppressed by a road heterogeneity [3]. The recent analysis of single-vehicle data by Neubert *et al.* [8], in particular, confirmed these features and also discovered fundamental microscopic properties distinguishing synchronized mode from other traffic states.

Theoretical description of the jam formation is far from being developed well because the synchronized mode possesses extremely complex structure [5]. For example, it comprises a certain continuum of quasistable states, so, matches a whole two-dimensional region on the phase plane ‘‘vehicle density–traffic volume’’ in contrast to the free flow state. However, tackling the question of how to regulate traffic flow on highways, for example, by controlling the speed limitation in order to prevent the jam formation we may rough out the problem. Indeed, for this purpose it is sufficient to analyze the conditions giving rise to jams rather than the jam evolution itself. Such a standpoint is justified, in part, by the aforementioned phase transitions being of the first order. The free flow state, presumably, should have the feasibility to exist at the given car density or inside its certain neighborhood. This might be the necessary requirement for jam formation at a fixed vehicle density, or for appearance of both the jam phase and the synchronized mode at a fixed traffic volume. Second, the jam formation proceeds via the nucleation mechanism but not in a regular manner. Therefore, the key point in the emergence of a jam is the random occurrence of its critical nucleus inside the synchronized mode or free flow.

*Electronic address: reinhart.kuehne@dlr.de
†Electronic address: reinhard.mahnke@physik.uni-rostock.de
‡Electronic address: ialub@fpl.gpi.ru

The jam formation manifests itself in the traffic breakdown, i.e., in a sharp drop of the traffic volume to a substantially lower value. Detecting these events one can get the rate of the critical nucleus generation depending on the road conditions and the traffic state. In this way the main attention is shifted to the experimental and theoretical analysis of the probabilistic features of jam formation regarding the characteristic mean values of the traffic volume as phenomenological parameters [9–12]. Such a probabilistic description of the traffic breakdown is the main purpose of the present paper.

At the first glance the problem seems hopeless until the model of the synchronized mode is developed. Nevertheless, there are circumstances enabling us to make a step towards this problem right now (see also Ref. [13]). The matter is that the $F\rightarrow S$ transition is of another nature than the $S\rightarrow J$ transition. The former is due to a sharp decrease in the overtaking frequency, giving rise to the synchronized mode, whereas the latter is caused by the pinch effect (see, e.g., Refs. [1–3]). Thus fluctuations in the vehicle density and velocity are not the main causes of the $F\rightarrow S$ transition, but other characteristic parameters of traffic flow are relevant (cf. also Ref. [14]). By contrast, just these fluctuations give rise to the jam in the synchronized traffic flow. As a result, the threshold of the $F\rightarrow S$ transition turns out to be remarkably less than that of the jam formation and can be attained at lower values of the vehicle density. So, the generation rate of critical nuclei for the former transition has to be great in comparison with the latter one. Thus, on time scales characterizing the occurrence of critical jam nuclei the traffic state with respect to the transitions between the free flow and synchronized mode is quasistationary. Therefore the formation of a critical jam nucleus is the leading nonequilibrium process limiting the traffic breakdown. The latter feature allows us to confine our consideration solely to the jam nucleus generation and to regard the synchronized mode and the free flow phase (if they coexist in the case under consideration) as one traffic state. Moreover, since a jam forms actually inside the synchronized mode where the vehicle motion in different lanes is strongly correlated we may apply to a single-lane road an approximation that treats all the cars moving in different lanes on a multilane highway and being neighbors across the highway as a single effective macrovehicle consisting of many cars. The macrovehicle concept is partly justified by the empirically observed fact that fluctuations in the downstream flow leaving a freeway bottleneck can proceed without the traffic state change, even their amplitude attains 30% of the mean traffic volume [9–13]. In any case, fluctuations in the traffic flow volume near its breakdown are of macroscopic nature and the critical nucleus of traffic jam has to include many vehicles. This feature is also pointed to by the observed breakdown near an on-ramp occurring each time after a large vehicle cluster entered the freeway stream [9].

It should be pointed out that the real structure of congested traffic near a highway bottleneck is sufficiently complex, it contains the region of synchronized mode located in the close proximity of the bottleneck, the preceding upstream region of moving narrow jams transformed into wide jams [13]. However, it is quite reasonable to consider this structure as being induced by the traffic breakdown processes arising inside the ''head'' of this complex jam, in the region of synchronized mode adjacent the bottleneck. Therefore the main characteristics of the breakdown phenomenon may be related to intrinsic processes taking place inside the synchronized mode on not too large spatial scales. The latter justifies our attempt to describe traffic breakdown ignoring the complex spatial structure of the metastable traffic state inside which critical jam nuclei originate.

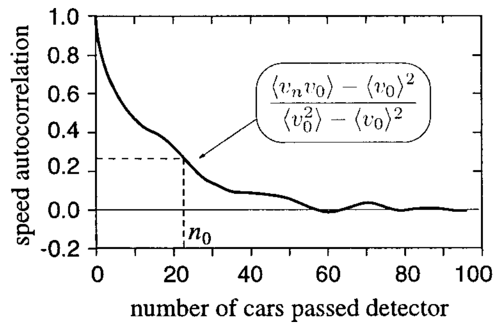

FIG. 1. Illustration of the speed autocorrelation vs the number of cars that have passed a fixed detector. Based on the observations by Neubert *et al.* [8].

Processes similar to the traffic breakdown are widely met in physical systems. For example, water condensation in supersaturated vapor proceeds via formation of small atom clusters of a critical size. Keeping in mind this analogy between the traffic breakdown and the phase transitions in physical systems, Mahnke *et al.* [15,16] proposed a kinetic approach based on the stochastic master equation describing the jam formation in terms of the attachment of individual cars to their cluster. However, the particular form of the developed master equation does not allow for the jam formation being the first order phase transition and, thus, the traffic breakdown. In the present paper we generalize this kinetic approach to describe the latter phenomenon.

However, before passing directly to the model we make two comments clarifying its original part and the feasibility of comparing the results to be obtained within the model and the available experimental data.

First, we recall the experimental data enabling us to estimate the characteristic size n_0 of vehicle clusters that are small enough so the behavior of drivers inside them seems to be special. From our point of view the multilane correlations in the vehicle motion are due to the drivers taking into account the behavior of all the cars, including also the cars in the neighboring lanes, which are inside the region accessible to observation. Therefore the synchronized mode has to exhibit strong correlations in this region. Figure 1 shows the speed autocorrelation function vs the number of cars that have passed a fixed detector that was experimentally found in the synchronized mode [8] (see also Ref. [17]). We see that the car velocities are strongly correlated over scales spanning some ten vehicles, i.e., a car cluster of this size, $n_0 \gtrsim 20$, makes up actually a certain fundamental unit of the synchronized mode. In the free flow no such long-distance correlations have been observed.

Second, the model to be developed considers the breakdown phenomenon for traffic flow on a homogeneous road, whereas the available experimental data were obtained for traffic flow near highway bottlenecks. Therefore it could be thought at first glance that their comparison is not justified. It would be unjustified indeed if the traffic breakdown near highway bottlenecks and far from them, i.e., on homogeneous segments of highways, would proceed by *different* mechanisms. However, applying to similar phenomena in physical media we see that a new phase inside a metastable one arises through generation of the critical nuclei in homogeneous as well as heterogeneous systems. The critical nucleus size is determined by the competition between the increase in the surface free energy and the decrease in the bulk one when a nucleus of new phase appears. The observed quantitative difference in the rate of phase transition in the same homogeneous and heterogeneous medium is caused by the critical nucleus form affected by the system boundary. So, the general properties of these phase transitions are similar in both cases and can be analyzed using the same physical concepts. Roughly speaking, the heterogeneous nucleation is singled out mainly by an individual form factor only.

The physics of traffic breakdown is currently far from being well understood. Therefore if it does proceed by the nucleation mechanism and the high probability of jam emergence near a highway bottleneck is due to the effect the bottleneck on the particular properties of car clusters only, then a model for homogeneous (or quasihomogeneous) road may be applied to its description. There is a natural way to verify this assumption. It is to compare the characteristic properties predicted by such a model and the observed data. If the model does predict something new or previously unclear then its application to the traffic breakdown also near highway bottleneck is justified, at least, to understand the basic features of the phenomenon. This question will be discussed in detail in Sec. III.

II. PROBABILISTIC MODEL FOR THE CAR AGGREGATION

A. Discrete description: Master equation

We consider traffic flow on a single-lane road and study the spontaneous formation of a jam regarded as a large car cluster arising on the road. Instead of dealing with a certain road part of length L and imposing some boundary conditions at its entries and exits we examine a circular road of length L with N cars moving on it. All the cars are assumed to be identical vehicles of effective length l_{car} and can make up two phases. One of them is the set of ''freely'' moving cars and the other is a single cluster. The cluster is specified by its size n, the number of aggregated cars. Its internal parameters, namely, the headway distance h_{clust} (i.e., the distance between the front bumper of a chosen car and the back bumper of the following one) and, consequently, the velocity v_{clust} of cars in the cluster are treated as fixed values independent of the cluster size n. We note that in the model under consideration there can be only one cluster on the road. The ''free'' flow phase is specified also by the corresponding headway distance $h_{\text{free}}(n)$ that, however, depends on the car cluster size n. The larger the cluster is, the less is the number

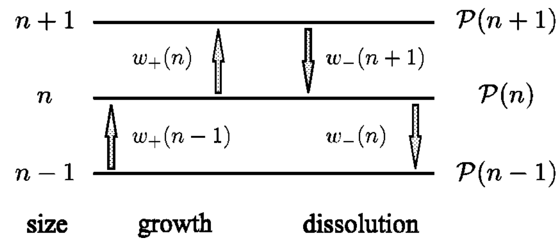

FIG. 2. Schematic illustration of the cluster transformations.

$(N-n)$ of the ''freely'' moving cars and therefore the longer is the headway distance $h_{\text{free}}(n)$.

When a cluster arises on the road its further growth is due to the attachment of the ''free'' cars to its upstream boundary, whereas the cars located near its downstream boundary accelerate to leave it, which decreases the cluster size. These processes are treated as random changes of the cluster size n by ± 1 (Fig. 2) and the cluster evolution is described in terms of time variations of the probability function $\mathcal{P}(n,t)$ for the cluster to be of size n at time t. Then following Mahnke *et al.* [15,16] we write the master equation governing the cluster growth,

$$\begin{aligned}\partial_t\mathcal{P}(n,t)=&\,w_+(n-1)\mathcal{P}(n-1,t)\\&+w_-(n+1)\mathcal{P}(n+1,t)-[w_-(n)\\&+w_+(n)]\mathcal{P}(n,t),\end{aligned}\tag{1}$$

where the cluster size n meets the inequality $1\leqslant n\leqslant(N-1)$. The transition rates $w_+(n)$ and $w_-(n)$ are illustrated in Fig. 2, depending generally on the cluster size n. The formation and dissolution of the maximum possible cluster containing all the cars is described by the equation

$$\begin{aligned}\partial_t\mathcal{P}(N,t)=&\,w_+(N-1)\mathcal{P}(N-1,t)\\&-w_-(N)\mathcal{P}(N,t),\end{aligned}\tag{2}$$

whereas the emergence of the jam seed, the cluster consisting of one car called below precluster, obeys the equation

$$\partial_t\mathcal{P}(0,t)=w_-(1)\mathcal{P}(1,t)-w_+(0)\mathcal{P}(0,t).\tag{3}$$

Here the function $\mathcal{P}(0,t)$ is the probability of no cluster on the road. At the initial time $t=0$ no cluster is assumed to be on the road,

$$\mathcal{P}(n,0)=\delta_{n0},\tag{4}$$

where $\delta_{nn'}$ is the Kronecker's delta. The system of equations (1)–(3) subject to the initial condition (4) provides the probabilistic description of the cluster formation.

Special attention should be paid to the question as to what the precluster is. The model proposes the following. When there is no cluster on the road, i.e., all the cars move ''freely,'' the velocity of one of them can randomly drop down to its value v_{clust} in the cluster. Such a car is regarded as the precluster, a size-one cluster. When a precluster has arisen its further evolution follows the scheme shown in Fig. 2. The precluster concept may be justified by recalling the

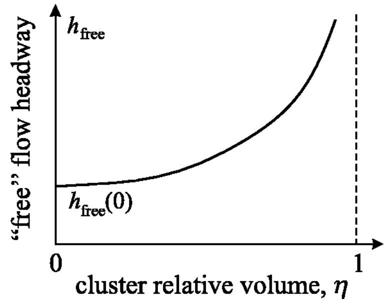

FIG. 3. The headway distance h_{free} in the "free" flow phase vs the cluster relative volume $\eta = n/N$. A qualitative sketch.

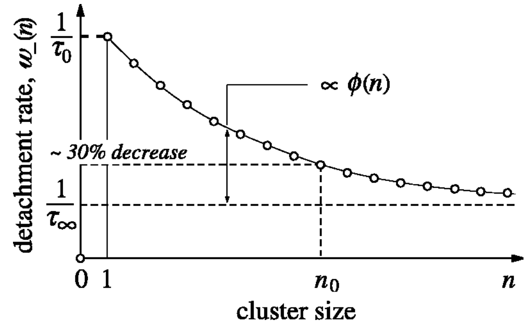

FIG. 4. The detachment rate $w_-(n)$ vs the cluster size n. A qualitative sketch.

problem we deal with initially, i.e., the breakdown processes in multilane traffic flow. The cars under consideration actually match small vehicle clusters of the synchronized mode, macrovehicle, arising in traffic flow on a multilane highway and comprised of real vehicles moving synchronously at different lanes. Therefore the precluster is actually as a certain sufficiently small cluster of the synchronized mode. Keeping in mind the relatively low threshold of the $F \rightarrow S$ transition, we will assume the precluster generation as well as the precluster dissipation to be intensive processes so that the "free" flow phase, $n=0$, and the precluster state, $n=1$, come into quasiequilibrium on time scales needed for the critical cluster nucleus to arise. In particular, in no case the precluster emergence limits the cluster evolution, so, the particular details of the precluster formation has no substantial effect on the traffic breakdown.

At the next step we should specify the transition rates $w_+(n)$ and $w_-(n)$. Let us apply to the optimal velocity model assuming the velocity v of the "freely" moving cars as well as the clustered cars to be determined directly by the corresponding headway distance h according to the formula

$$v = \vartheta(h) := v_{\max} \frac{h^p}{h^p + D_{\text{opt}}^p}. \tag{5}$$

Here D_{opt} is the characteristic value of the headway distance at which drivers begin to feel themselves "free" and their velocity approaches the maximum $v_{\max}$. The parameter $p > 1$ allows for different shapes of the function $\vartheta(h)$. The larger the value of p is, the sharper is the dependence $\vartheta(h)$. Case $p=2$ is often used [15,16]. A car attaches itself to the cluster as fast as the distance to the last car in the cluster decreases down to the cluster headway h_{clust}, enabling us to write the following ansatz for the attachment rate to the cluster of size $n \geqslant 1$:

$$w_+(n) = w_+^{\text{ov}}(n) := \frac{\vartheta[h_{\text{free}}(n)] - \vartheta[h_{\text{clust}}]}{h_{\text{free}}(n) - h_{\text{clust}}}. \tag{6}$$

Applying to a simple geometrical consideration and assuming $N \gg 1$ as well as $N - n \gg 1$ we get the relationship (illustrated also in Fig. 3)

$$h_{\text{free}}(n) = h_{\text{clust}} + (l_{\text{car}} + h_{\text{clust}}) \frac{(\rho_{\text{lim}} - \rho)}{\rho(1-\eta)}, \tag{7}$$

which together with Eq. (6) gives the attachment rate as a function of the cluster size n. Here we have introduced the following traffic flow parameters: $\rho = N/L$ being the mean value of the car density on the road, its maximum possible value $\rho_{\text{lim}} = 1/(l_{\text{car}} + h_{\text{clust}})$ for the given road, and the relative volume $\eta = n/N$ of the cluster with respect to the initial volume of the "free" flow state.

In order to compare the cluster growth due to the car attachment with the precluster generation we specify its rate in terms of

$$w_+(0) = \epsilon w_+^{\text{ov}}(0), \tag{8}$$

where ϵ is a phenomenological factor and we formally set $n=0$ in expression (6). Keeping in mind the aforesaid about the precluster emergence, we assume the factor ϵ to be about unit, $\epsilon \sim 1$, or at least not to be small enough to limit the cluster formation, so its particular numerical value is of no importance.

The rate of the cars escaping from the cluster at its downstream front is written as (see also Fig. 4)

$$w_-(n) = \frac{1}{\tau(n)} := \frac{1-\phi(n)}{\tau_\infty} + \frac{\phi(n)}{\tau_0}, \tag{9}$$

where the value $\tau(n)$ can be interpreted as the characteristic time needed for the first car in the cluster to leave it and to go out from its downstream boundary at a distance about the headway distance $h_{\text{free}}(n)$ in the current "free" flow state. The function $\phi(n)$ allows for the dependence of the detachment time $\tau(n)$ on the cluster size n. We note that expression (9) is the main original part of the model under consideration.

When the cluster is sufficiently large, $n \gg 1$, it is reasonable to regard the characteristic time $\tau(n) \approx \tau_\infty$ as a constant [i.e., $\phi(n) \rightarrow 0$ for $n \gg 1$] as was done in papers [15,16] for all the values of n.

For small clusters the $\tau(n)$ dependence, however, requires special attention. The matter is that the car attachment rate $w_+(n)$ is considered to be directly determined by the local characteristics of the "free" flow phase and the car cluster. Thus the dependence of the attachment rate $w_+(n)$ on the cluster size n arises via the headway distance $h_{\text{free}}(n)$ in the "free" flow being a function of n, i.e., $w_+(n) = w_+[h_{\text{free}}(n), h_{\text{clust}}]$. Therefore the attachment rate is actually an explicit function $w_+(\rho,\eta)$ of the mean car density ρ and the cluster relative volume η only and, so, exhibits minor variations on scales $\delta n \ll N$. As will be seen below, exactly this feature is essential rather than the particular form of $w_+(\rho,\eta)$ given here. To describe traffic breakdown at least one of the kinetic coefficients $w_+(n)$ and $w_-(n)$ has to be a direct function of the cluster size n for its relatively small values corresponding to the formation of the critical nucleus. We associate this dependence with the escaping rate $w_-(n)$ that, in contrast to the attachment rate $w_+(n)$, exhibits its substantial variations in the region $n \lesssim n_0 \sim 20$.

The parameter n_0 actually divides the car clusters into the large cluster group, $n \gg n_0$, for which the escaping rate is constant, $\phi(n)\to 0$, and the group of small clusters, $n \lesssim n_0$, whose dissolution is affected substantially by the size n. This assumption is based on the fact that there should be a variety of possible manoeuvres for a driver to escape from a sufficiently small cluster on a multilane highway when the lanes are not too crowded.

Expression (9) takes into account this effect via the function $\phi(n)$ running from 1 to 0 as the cluster size n increases, so, $\phi(1)\simeq 1$ and $\phi(n)\to 0$ as $n\to\infty$. In particular, for a small neighborhood of the precluster size, $n\sim 1$, the value τ_0 of $\tau(n)$ gives us actually the lifetime of the preclusters and is assumed to be less than the escaping time from a large cluster, i.e., $\tau_0 < \tau_\infty$. Naturally, for the case of no cluster on the road we have to set $w_-(0)=0$. The main results will be obtained below actually applying to the general properties of the dependence $w_-(n)$, however, for the sake of simplicity, we will adopt the following ansatz for $n \geq 1$:

$$\phi(n) := \phi[x]\big|_{x=n/n_0} := \frac{1}{(1+x)^q}, \tag{10}$$

where the exponent $q>1$ is regarded as a given constant. We point out once more that the dependence of the characteristic time $\tau(n)$ on the cluster size is crucial because it is responsible for the existence of the metastable "free" flow phase.

The system of equations (1)–(3) subject to the initial condition (4) with the relationships (6), (8), and (9) forms the proposed probabilistic model for the car aggregation. Within this model we will analyze the characteristic features of the large cluster emergence and the shape of the fundamental diagram, i.e., the "flow volume–car density" relation in the vicinity of traffic breakdown. In particular, in the adopted terms the flow volume $j(n)$ for the given traffic flow state, i.e., when a cluster of size n arises on the road, is written as [15,16]:

$$j(n) = (1-\eta)\rho\vartheta[h_{\text{free}}(n)] + \eta\rho\vartheta[h_{\text{clust}}]. \tag{11}$$

We will get the fundamental diagram $j=j(\rho)$ by averaging expression (11) with respect to the distribution $\mathcal{P}(n,t)$.

B. Equilibrium distribution

To clarify the characteristic features of the cluster formation let us analyze, first, the stationary cluster size distribution $\mathcal{P}_{\text{eq}}(n)$. The system of equations (1)–(3) subject to the initial condition (4) admits the stationary solution $\mathcal{P}_{\text{eq}}(n)$ meeting the zero "probability" flux in the cluster size space,

$$w_+(n-1)\mathcal{P}_{\text{eq}}(n-1) - w_-(n)\mathcal{P}_{\text{eq}}(n) = 0.$$

It is evident that

$$\frac{\mathcal{P}_{\text{eq}}(n)}{\mathcal{P}_{\text{eq}}(n-1)} = \frac{w_+(n-1)}{w_-(n)}$$

holds, which enables us to write the expression

$$\mathcal{P}_{\text{eq}}(n) \propto \exp\{-\Omega(n)\},$$

where the function $\Omega(n)$ (called below the car growth potential) is specified for $n\geq 2$ by the formula

$$\Omega(n) = -\sum_{n'=1}^{n-1} \ln[\tau_\infty w_+^{\text{ov}}(n')] + \sum_{n'=2}^{n} \ln\left[1 + \frac{(\tau_\infty - \tau_0)}{\tau_0}\phi(n')\right]. \tag{12}$$

Both terms in Eq. (12) vary weakly as the argument n changes by one, enabling us to convert sum (12) into an integral with respect to the cluster size n treated as a continuous variable,

$$\Omega(n) = \Omega_\infty(n) + \Omega_0(n), \tag{13}$$

where

$$\Omega_\infty(n) \simeq -\int_0^n dn' \ln\{\tau_\infty w_+^{\text{ov}}[h_{\text{free}}(n')]\}, \tag{14}$$

$$\Omega_0(n) \simeq \int_0^n dn' \ln\left[1 + \frac{(\tau_\infty - \tau_0)}{\tau_0}\phi(n')\right]. \tag{15}$$

The former term in Eq. (12) or Eq. (13), i.e., the component $\Omega_\infty(n)$ called below the growth potential mainly characterizes whether a stable car cluster can arise on the road under the given conditions and specifies its size because it exhibits substantial variations on large scales exceeding substantially the size n_0. By contrast the latter one, the component $\Omega_0(n)$, describes the formation of the critical cluster nucleus and, so, the breakdown phenomenon. Indeed, as follows from Eqs. (10) and (15) the potential $\Omega_0(n)$ is constant for $n \gg n_0$ and, thus, cannot affect the growth of a large cluster already formed on the road. Besides, within the continuum approximation we have ignored the details of the cluster distribution in the region including both the points $n=0$ and $n=1$ and expand the cluster space $n\geq 1$ to the whole axis $n \geq 0$.

Let us, first, analyze the condition of the cluster emergence dealing with the potential $\Omega_\infty(n)$ only. Applying to Fig. 5 we can see that a large cluster can arise on the road, in principle, if there exists a value of the headway distance h_c meeting the equality

$$\tau_\infty w_+^{\text{ov}}[h_c] = 1, \tag{16}$$

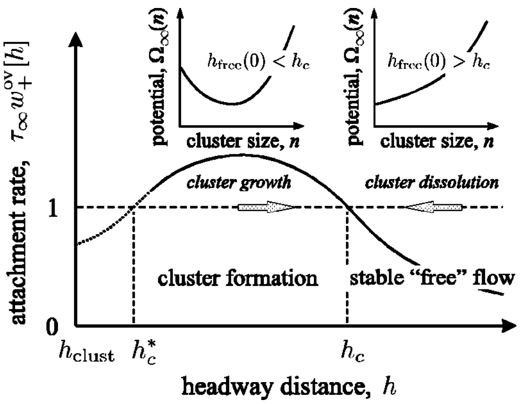

FIG. 5. The attachment rate $w_+^{ov}[h]$ vs the headway distance h and the stability regions of the "free" flow phase.

which will be assumed to hold beforehand. In particular, within approximation (6) together with Eq. (5) for $D_{opt} \gg h_{clust}$ and $p=2$ this assumption holds if $\tau_\infty v_{max} > 2D_{opt}$ and the critical headway reads

$$h_c^{(2)} = \frac{1}{2}\left(\tau_\infty v_{max} + \sqrt{(\tau_\infty v_{max})^2 - 4D_{opt}^2}\right),$$

whereas for $p \to \infty$ and $\tau_\infty v_{max} > D_{opt}$ we have

$$h_c^{(\infty)} = \tau_\infty v_{max}.$$

The "free" flow phase will be stable if the initial headway distance $h_{free}(0) > h_c$ and unstable otherwise.

Let us justify these statements. The growth potential $\Omega(n)$ is actually the sum of $\ln[w_-(n)/w_+(n)]$ over n [see formula (12)]. So, in the region where the integrand of Eq. (14) meets the inequality $\tau_\infty w_+^{ov}[h] < 1$ and the potential $\Omega_\infty(n)$ is an increasing function of n, the cluster dissolution is more intensive than the car attachment. Under these conditions the cluster size on the average decreases in time. The same concerns the time dependence of the headway distance $h_{free}(n)$ in the "free" flow phase because the value of $h_{free}(n)$ decreases as the cluster becomes smaller (Fig. 3), which is also illustrated by arrows in Fig. 5. Since $\tau_\infty w_+^{ov}[h] < 1$ for $h > h_c$ any randomly arising cluster tends to disappear and, consequently, the "free" flow phase is stable when $h_{free}(0) > h_c$. In this case the potential $\Omega_\infty(n)$ possesses one minimum located at the boundary point $n=0$ (or $n=1$ what is the same in the continuum description).

Otherwise, $h_{free}(0) < h_c$, there is a region $h_{free}(0) < h < h_c$ where $\tau_\infty w_+^{ov}[h] > 1$ and the car attachment rate exceeds that of the cluster dissolution and a cluster occurring in the corresponding "free" flow state tends to grow, inducing the further increase in the headway distance $h_{free}(n)$. In this case the "free" flow phase is unstable and the cluster will continue to grow until the value of $h_{free}(n)$ reaches the critical point h_c, where the car attachment and the cluster dissolution balance each other. Whence it follows, in particular, that the developed cluster is of the size n_{clust} obeying the equation

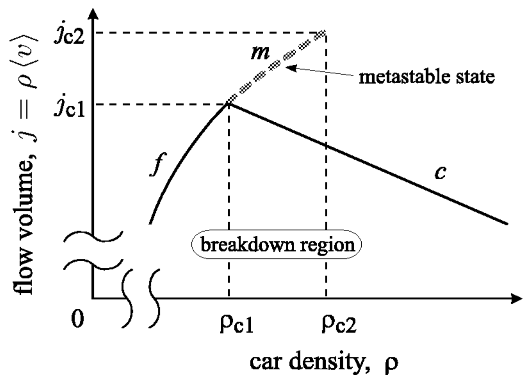

FIG. 6. The fundamental diagram in the vicinity of the breakdown region.

$$h_{free}(n_{clust}) = h_c \tag{17}$$

and the $\Omega_\infty(n)$ has a minimum at the internal point $n = n_{clust}$. In the present paper we will ignore the existence of another region where the equality $\tau_\infty w_+^{ov}[h] < 1$ also holds for very dense traffic flow, which has been considered in papers [15,16].

Relationship (7) enables us to rewrite the instability conditions in terms of the mean car density ρ. The critical value ρ_c of the car density is the solution of the equation $h_{free}(0) = h_c$, whence we immediately get

$$\rho_{c1} = \rho_{lim}\frac{l_{car} + h_{clust}}{l_{car} + h_c}. \tag{18}$$

Then the stable state of the "free" flow corresponds to the inequality $\rho < \rho_{c1}$ and it loses the stability when $\rho > \rho_{c1}$. In the latter case a large cluster of size $n_{clust}(\rho) = \eta_{clust}(\rho)N$ arises on the road and its relative volume is

$$\eta_{clust}(\rho) = \frac{h_c + l_{car}}{h_c - h_{clust}}\,\frac{\rho - \rho_{c1}}{\rho}. \tag{19}$$

In the given analysis we have ignored the dependence of the cluster dissolution rate $w_-(n)$ on the size n and, thereby, the considered picture describes actually the "free" flow cluster transition of the second order. It does not allow for the metastable state of the "free" flow phase and corresponds to the continuous transition from the traffic state without cluster on the road to the formation of a certain cluster whose relative volume changes continuously from zero as the car density penetrates deeper in the instability region [see formula (19)]. Consequently, this approximation cannot explain the traffic breakdown and on the phase diagram matches solely the stable branches "f" and "c" of the "free" flow and the traffic with a developed cluster, respectively (Fig. 6). Nevertheless, exactly the given approximation describes the stable branches of the fundamental diagram and, moreover, the metastable branch is a continuation of the branch "f" into the instability region. Keeping in mind the latter, we present also the expression specifying these branches,

$$j_{fc}(\rho)=\begin{cases}\rho\vartheta[h_c+(h_c+l_{\rm car})(\rho_{c1}-\rho)/\rho] & \text{if} \quad \rho<\rho_{c1}\\ j_{c1}-G(\rho-\rho_{c1})/\rho_{c1} & \text{if} \quad \rho>\rho_{c1},\end{cases}$$

where the constants are

$$j_{c1}=\rho_{c1}\vartheta[h_c],$$

$$G=\frac{(h_c+l_{\rm car})}{(h_c-h_{\rm clust})}\frac{\rho_{c1}}{\rho_{\rm lim}}(\rho_{c1}\vartheta[h_c]-\rho_{\rm lim}\vartheta[h_{\rm clust}]).$$

It should be noted that this expression has been obtained by substituting the maximum probability value $n_{\rm clust}$ of the cluster size into expression (11) instead of averaging it over the distribution $\mathcal{P}_{\rm eq}(n)$. The latter is justified because the effect of the cluster size fluctuations is ignorable due to $N\gg1$.

Now we analyze the possible metastable states of the ''free'' flow phase. In order to do this we should take into account both the components of the growth potential $\Omega(n)$. Since the function $\Omega_0(n)$ exhibits remarkable variations in the region $n\lesssim n_0$ only and, thus, the size n_c of the critical nucleus also belongs to this region, we may confine our consideration to clusters whose size n is much less than the final cluster size $n_{\rm clust}$ attained after the instability development. In addition, for simplicity we will regard the value $(\tau_\infty-\tau_0)/\tau_\infty$ as a small parameter, which enables us to examine solely a small neighborhood of the instability boundary, $0<\rho-\rho_{c1}\ll\rho_{c1}$.

In this case the value of $\tau_\infty w_+^{\rm ov}[h_{\rm free}(n)]$ is practically constant and can be approximated by the expression

$$\ln\{\tau_\infty w_+^{\rm ov}[h_{\rm free}(n)]\}\simeq g\frac{\rho-\rho_{c1}}{\rho_{c1}}, \tag{20}$$

where the coefficient

$$g=\frac{(l_{\rm car}+h_c)}{h_c}\left|\frac{d\ln w_+^{\rm ov}[h]}{d\ln h}\right|_{h=h_c}$$

is about unity, $g\sim1$, in the general case. In particular, we have the rigorous equality $g=1$ for stepwise dependence of $\vartheta(h)$ [if we set $p=\infty$ in expression (5)] and $D_{\rm opt}\gg l_{\rm car}$, $h_{\rm clust}$. Expression (20) together with formula (10) allows us to represent the dependence of the growth potential $\Omega(n)$ on the cluster size n as

$$\frac{d\Omega(n)}{dn}\approx\frac{(\tau_\infty-\tau_0)}{\tau_0}\phi(n)-g\frac{(\rho-\rho_{c1})}{\rho_{c1}}$$
$$=\frac{(\tau_\infty-\tau_0)}{\tau_0}\left(\frac{n_0}{n_0+n}\right)^q-g\frac{(\rho-\rho_{c1})}{\rho_{c1}}. \tag{21}$$

The first term on the right-hand side of Eq. (21) is due to the increase in the cluster dissolution rate for $n\lesssim n_0$, whereas the latter one is proportional to the cluster growth rate in the region of large values of n. The resulting value of the derivative $d\Omega(n)/dn$ characterizes the direction of the cluster evolution. If it is positive, $d\Omega(n)/dn>0$, i.e., the potential $\Omega(n)$ is an increasing function of n, then the cluster dissolution is the dominant process and the cluster of size n tends to shrink. Otherwise, when $d\Omega(n)/dn<0$, it will grow.

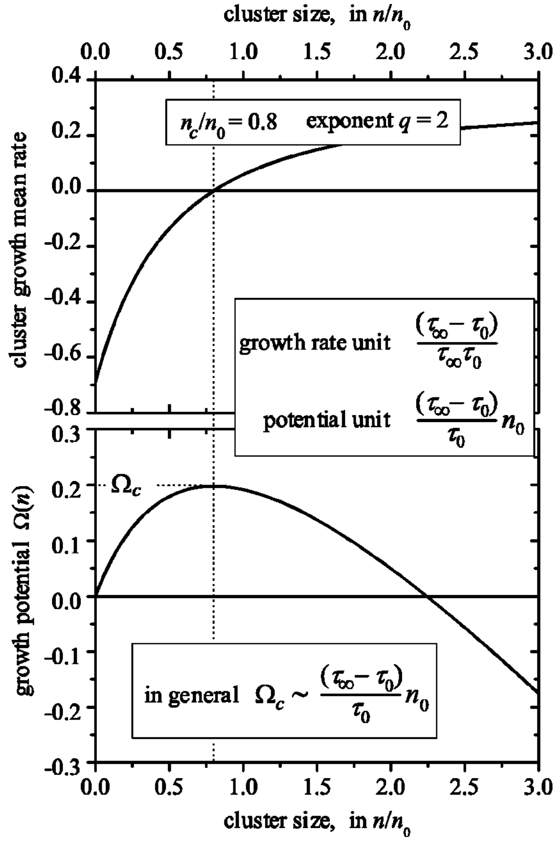

FIG. 7. The form of the cluster growth potential $\Omega(n)$ in the breakdown region (*upper window*), the value of $-d\Omega(n)/dn$ proportional to the mean rate of cluster growth vs the cluster size n and (*lower window*) the growth potential $\Omega(n)$ itself. The present figure has been obtained using ansatz (10) with the exponent $q=2$ and the chosen value of the vehicle density ρ gives the ratio of the critical nucleus size n_c to the characteristic value n_0 equal to $n_c/n_0=0.8$.

The former term attains its maximum at $n=0$, so, according to Eq. (21) the derivative $d\Omega(n)/dn$ is negative for all the possible values of the cluster size under consideration $0\leqslant n\ll n_{\rm clust}$ when

$$\rho>\rho_{c2}:=\rho_{c1}\left[\frac{(\tau_\infty-\tau_0)}{g\tau_0}+1\right]. \tag{22}$$

In this case the ''free'' flow phase becomes absolutely unstable. Under the opposite condition, $\rho_{c1}<\rho<\rho_{c2}$ there is a certain value n_c at which the derivative $d\Omega(n)/dn$ changes the sign (Fig. 7). Setting the left-hand side of Eq. (21) equal to zero we get the relationship

$$\phi(n_c)=\frac{\rho-\rho_{c1}}{\rho_{c2}-\rho_{c1}}, \tag{23}$$

which together with ansatz (10) gives the estimate

$$n_c=n_0\left[\left(\frac{\rho_{c2}-\rho_{c1}}{\rho-\rho_{c1}}\right)^{1/q}-1\right], \tag{24}$$

well justified except for small neighborhoods of the boundary points ρ_{c1} and ρ_{c2}. If $n<n_c$ the derivative is positive and the cluster should decrease in size, i.e., the "free" flow phase is stable with respect to the emergence of such small clusters. However, if a certain cluster of size $n>n_c$ has already formed, for example, due to random fluctuations, then it will grow and a large cluster of size n_{clust} arises on the road because $d\Omega(n)/dn<0$ in the region $n>n_c$.

In other words, we have shown that the dependence of the dissolution rate $w_-(n)$ on the cluster size $n\lesssim n_0$ makes the "free" flow phase metastable when the car density belongs to the interval $\rho\in(\rho_{c1},\rho_{c2})$ (branch "m" in Fig. 6). The formation of a large cluster, $n\gg n_0$, proceeds via generation of the critical nucleus whose size n_c is estimated by expression (24). In order to find the generation rate of the critical nuclei and, thus, the breakdown frequency we should consider the transient processes in the cluster growth, which is the subject of the following section.

C. Continuum approximation: The breakdown probability

In order to apply well developed techniques of the escaping theory (see, e.g., Ref. [19]) to the analysis of the traffic breakdown probability we approximate the discrete master equation (1) by the corresponding Fokker-Planck equation. It is feasible because in the case under consideration the kinetic coefficients $w_+(n)$, $w_-(n)$, first, vary smoothly on scales about unity and, second, are approximately equal to each other, $|w_+(n)-w_-(n)|\ll w_-(n)$. The latter conditions enable us to treat the argument n as a continuous variable and to expand the functions $w_+(n\pm1)$, $w_-(n\pm1)$, and $\mathcal{P}(n\pm1,t)$ into the Taylors series. In this way and, in addition, taking into account expression (9) we reduce Eq. (1) to the following Fokker-Planck equation:

$$\tau_\infty\partial_t\mathcal{P}(n,t)=\partial_n[\partial_n\mathcal{P}(n,t)+\mathcal{P}(n,t)\partial_n\Omega(n)], \tag{25}$$

where the potential $\Omega(n)$ is given by formula (13) in the general form. However, in the case under consideration the ratios n/N, $(\rho-\rho_{c1})/\rho_{c1}$, and $(\tau_\infty-\tau_0)/\tau_0$ are regarded to be sufficiently small and it is possible to expand the potential $\Omega(n)$ in the three parameters and to remain the leading terms only. In this way we get

$$\Omega(n)=\frac{(\tau_\infty-\tau_0)}{\tau_0}n_0\left\{\int_0^{n/n_0}dx\,\phi[x]-\phi[x_c]\frac{n}{n_0}\times\left[1-\frac{n}{2n_{\text{clust}}(\rho)}\right]\right\}, \tag{26}$$

where $x_c=n_c/n_0$ and we have set $\Omega(0)=0$. Equation (2) transforms into the boundary condition at infinitely distant points, which is imposed on the probability flux

$$J(n):=-\partial_n\mathcal{P}(n,t)-\mathcal{P}(n,t)\partial_n\Omega(n)$$

and requires it to be equal to zero, $J(\infty)=0$. Equation (3) describing the precluster generation is reduced, in turn, to the zero boundary condition imposed on the probability flux $J(n)$ formally at $n=0$, i.e., $J(0)=0$. The latter is justified by the assumed quasi-equilibrium between the "free" flow phase and the preclusters. And, finally, the initial condition (4) can be rewritten as

$$\int_0^\infty dn\,\mathcal{P}(n,t)=1.$$

When the car density belongs to the interval $\rho\in(\rho_{c1},\rho_{c2})$ and the "free" flow phase is metastable the form of the growth potential $\Omega(n)$ in the region $n\lesssim n_0$ is shown in Fig. 7. The "free" flow phase being in quasiequilibrium with preclusters matches the local minimum at $n=0$ separated from the region of the stable cluster growth $n>n_c$ by the potential barrier Ω_c. The value of this potential barrier is estimated as

$$\Omega_c\simeq\frac{(\tau_\infty-\tau_0)}{\tau_0}n_0\omega\left[\frac{n_c}{n_0}\right], \tag{27}$$

where the function ω is defined as

$$\omega[x_c]:=\int_0^{n_c/n_0}dx\,x\left(-\frac{d\phi[x]}{dx}\right). \tag{28}$$

In particular, for ansatz (10) with exponent $q=2$ expression (27) becomes

$$\Omega_c\simeq\frac{(\tau_\infty-\tau_0)}{\tau_0}n_0\frac{x_c^2}{(1+x_c)^2}. \tag{29}$$

Moreover, in the limit $x_c\ll1$ we have

$$\omega[x_c]\simeq\frac{1}{2}rx_c^2,\ \text{ with the constant }\ r=-\left.\frac{d\phi[x]}{dx}\right|_{x=0},$$

as follows from expression (28), and the general formula for the potential Ω_c can be written as

$$\Omega_c\simeq\frac{(\tau_\infty-\tau_0)}{2\tau_0}rn_0x_c^2. \tag{30}$$

In the same limit expression (23) gives us

$$x_c\simeq\frac{(\rho_{c2}-\rho)}{r(\rho_{c2}-\rho_{c1})}. \tag{31}$$

The main much deeper minimum of the potential $\Omega(n)$ is located at $n=n_{\text{clust}}\gg n_0$.

We have demonstrated that a precluster must climb over the potential barrier Ω_c at the point n_c to convert into a large stable cluster. It is implemented through random fluctuations carrying the cluster size up to the critical value n_c. In these terms the traffic breakdown is the classical problem of escaping from a potential well described by the Fokker-Planck equation (25). The latter analogy enables us to write down the estimate for the frequency ν_{bd} of the traffic breakdown processes depending on the given vehicle density in the "free" flow state. Namely, as shown in Appendix

$$\nu_{\mathrm{bd}} \simeq \frac{1}{\sqrt{2\pi n_0}\tau_\infty}\left(\frac{\tau_\infty-\tau_0}{\tau_0}\right)^{3/2}(1-\phi[x_c])|\phi'[x_c]|^{1/2} \times \exp\left\{-\frac{(\tau_\infty-\tau_0)}{\tau_0}n_0\omega[x_c]\right\}, \tag{32}$$

which is well justified for the car density ρ belonging to the interval $\rho_{c1}<\rho<\rho_{c2}$ except for certain sufficiently small neighborhoods on the critical points ρ_{c1}, ρ_{c2}. Ansatz (10) with the exponent $q=2$ together with formula (23) enables us to rewrite expression (32) as

$$\nu_{\mathrm{bd}} \simeq \frac{1}{\sqrt{\pi n_0}\tau_\infty}\left(\frac{\tau_\infty-\tau_0}{\tau_0}\right)^{3/2}(1-\Delta)\Delta^{3/4} \times \exp\left\{-\frac{(\tau_\infty-\tau_0)}{\tau_0}n_0\frac{(1-\Delta)^2}{(1+\Delta^{1/2})^2}\right\}. \tag{33}$$

Here we have introduced the quantity

$$\Delta := \frac{\rho-\rho_{c1}}{\rho_{c2}-\rho_{c1}} \tag{34}$$

treated as a dimensionless overcriticality measure showing how deep the system penetrates into the metastability region. $\Delta=0$ corresponds to the value ρ_{c1} of the vehicle density where a jam can emerge in principle and $\Delta=1$ matches the vehicle density ρ_{c2} after exceeding which no traffic state except for jams can exist at all (Fig. 6).

D. Frequency of traffic breakdown during a fixed time interval

Experimentally, traffic breakdown is typically analyzed detecting a significant drop in the vehicle speed during a certain fixed time interval T_{obs} about several minutes and then drawing the relative frequency of these events vs the traffic volume [9–12]. In order to compare this representation with the obtained results let us consider them in more detail.

As follows from expression (22) the density interval (ρ_{c1},ρ_{c2}) inside which the traffic jam emerges by the nucleation mechanism is of the thickness

$$(\rho_{c2}-\rho_{c1}) = \rho_{c1}\frac{(\tau_\infty-\tau_0)}{g\tau_0}.$$

According to the experimental data [9–13] the thickness of the traffic volume interval inside which the traffic breakdown demonstrates the probabilistic behavior is about its low boundary in magnitude. So we have to regard the ratio $(\tau_\infty-\tau_0)/\tau_0$ also as a value about unity,

$$\frac{(\tau_\infty-\tau_0)}{g\tau_0} \sim 1. \tag{35}$$

Thereby, setting $n_0=20$ we conclude that in the general case where $n_c \sim n_0$ the potential barrier is $\Omega_c \sim 5$ corresponding to the exponential factor $\exp\{-\Omega_c\} \sim 0.7\times10^{-2}$. Then setting $\tau_\infty \sim 2$ s and estimating the preceding cofactor as $1/(\sqrt{2\pi n_0}\tau_\infty)$ we find the characteristic rate of the traffic breakdown to about 1/50 min^{-1} in the general case. So the real traffic breakdown events seem to be observed in cases where the vehicle density comes to the upper boundary ρ_{c2}. The latter allows us to confine our analysis formally to the limit case

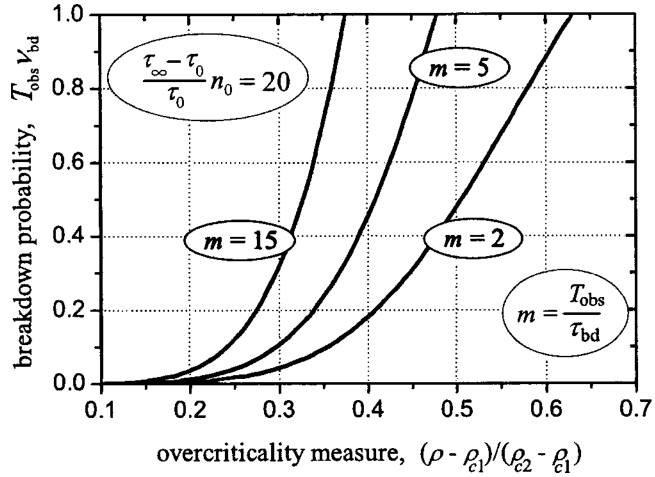

FIG. 8. The traffic breakdown probability vs the depth $(\rho-\rho_{c1})/(\rho_{c2}-\rho_{c1})$ of penetration into the metastability region. In calculating ansatz (10) with the exponent $q=2$ has been used.

$$x_c \ll 1 \Leftrightarrow (\rho_{c2}-\rho) \ll (\rho_{c2}-\rho_{c1}). \tag{36}$$

Then estimating the probability $\mathcal{F}_{\mathrm{bd}}$ of detecting a traffic breakdown during the observation time interval T_{obs} as $\mathcal{F}_{\mathrm{bd}} = T_{\mathrm{obs}}\nu_{\mathrm{bd}}$ we obtain from Eq. (32) the expression

$$\mathcal{F}_{\mathrm{bd}}(\Delta) = \frac{T_{\mathrm{obs}}}{\tau_{\mathrm{bd}}}2\Omega_c^{1/2}\exp\{-\Omega_c\}, \tag{37}$$

where

$$\Omega_c = \frac{(\tau_\infty-\tau_0)}{2r\tau_0}n_0(1-\Delta)^2 \tag{38}$$

and we have introduced the time scale

$$\tau_{\mathrm{bd}} = \frac{2\sqrt{\pi}\tau_0}{r(\tau_\infty-\tau_0)}n_0\tau_\infty, \tag{39}$$

giving us the characteristic time of the breakdown emergence. In deriving Eq. (37) we have also taken into account formulas (10), (30), and (31) and number 2 remained directly as cofactor because the maximum of the function $z^{1/2}\exp(-z)$ is about 0.43. Naturally, we have to confine ourselves to such values of the vehicle density for which $\mathcal{F}_{\mathrm{bd}} \leqslant 1$ holds because traffic flow with higher values of the vehicle density cannot exist on these time scales. For the following values $r=2$, the ratio $(\tau_\infty-\tau_0)/\tau_0 \sim 1$, $\tau_\infty \sim 2$ s, and $n_0 \sim 20$ expression (39) gives us the estimate $\tau_{\mathrm{bd}} \sim 1$ min of the characteristic breakdown time. It should be pointed out that the latter estimate does not contradict the evaluation of the breakdown rate given at the beginning of the present section because it holds only in the region $\Omega_c \gg 1$.

Figure 8 illustrates the obtained results depicting the breakdown probability for different values of the observation time $T_{\rm obs}$ measured in units of $\tau_{\rm bd}$ vs the depth of penetrating into the metastability region. It should be pointed out that in drawing Fig. 8 we have applied formula (33) rather than Eq. (37) in order to have a possibility to go out of the frameworks of the formal limit (36). The latter is considered here to clarify the obtained results only. To make the form of the $\mathcal{F}_{\rm bd}(\Delta)$ dependence more evident we apply again the formal limit case (36) assuming the ratio $m := T_{\rm obs}/\tau_{\rm bd}$ to be a large parameter. Then analyzing a small neighborhood of the point ρ_m^* specified by the equality

$$1-\Delta_m^* \approx \left[\frac{2r\ln(2m)\tau_0}{(\tau_\infty-\tau_0)n_0}\right]^{1/2} \tag{40}$$

we get

$$\mathcal{F}_{\rm bd}(\rho) = \exp\left(\frac{\rho-\rho_m^*}{\overline{\rho}_m}\right), \tag{41}$$

where the vehicle density scale is

$$\overline{\rho}_m = (\rho_{c2}-\rho_{c1})\left[\frac{r\tau_0}{2\ln(2m)(\tau_\infty-\tau_0)n_0}\right]^{1/2}. \tag{42}$$

Therefore, in a rough approximation the $\mathcal{F}_{\rm bd}(\rho)$ dependence is a simple exponential function whose scale $\overline{\rho}_m$ is approximately a constant value [because the function ln(m) shows weak variations for $m \gg 1$]. Changing the observation duration $T_{\rm obs}$ practically shifts the cutoff point ρ_m^* only.

III. COMPARISON WITH EXPERIMENTAL RESULTS

We recall once more that we compare the model developed for traffic flow on a homogeneous road with the experimental data obtained near highway bottlenecks (see also Introduction). So this comparison is justified if the model can explain the characteristic features of the observed phenomenon based on the general properties of traffic flow, which is the main goal of the present section.

Experimental investigations of the traffic breakdown regarded as a probabilistic phenomenon have been carried out by several authors (see, e.g., Refs. [9–12,18]). Elefteriadou *et al.* [9] actually pointed out the fact that traffic breakdown at ramp merge junctions occurs randomly without precise relation to a certain fixed value of traffic volume. A more detailed analysis of the breakdown probability has been made in papers [10–12,18]. These observations show that traffic breakdown can occur inside a wide interval of traffic volume from about 1500 veh/h/l (vehicles per hour per lane) up to 3000 veh/h/l. The real dynamics of traffic breakdown near bottlenecks and the developed structure of the congested traffic flow are sufficiently complex as was exhibited by Kerner [13]. In particular, Kerner demonstrated that the synchronized mode of traffic flow in the vicinity of highway bottleneck is locally metastable under the discharged downstream traffic flow of volume j varying in the same interval. The latter enables us to estimate the detachment time τ_∞ playing significant role in the presented model. In fact, ignoring the velocity $\vartheta(h_{\rm clust})$ of cars in the cluster as well as the headway distance $h_{\rm clust}$ in ansatz (6) we find that the lower boundary ρ_{c1} of the metastability region meets the traffic volume

$$j_{c1} = \rho_{c1}\vartheta(h_c) \approx w_+^{\rm ov}[h_c] = \frac{1}{\tau_\infty} \sim 1800 \ \ \text{veh/h/l}.$$

We find immediately the estimate $\tau_\infty \sim 2$ s, which is in agreement with the value adopted previously in papers [15,16]. Like in the preceding section, here we use the estimates of the quantities $n_0 \sim 20$ according to the experimental data depicted in Fig. 1, taking $(\tau_\infty-\tau_0)/\tau_0 \sim 1$ from the general consideration.

In order to compare the obtained results and the available experimental data we have applied the latest materials presented in detail by Lorenz and Elefteriadou [12]. The breakdown phenomenon was investigated in traffic flow near two bottlenecks of Highway 401, one of Toronto's primary traffic arteries. The detectors were located right after the on ramps within several hundred meters downstream. So the dynamics of traffic breakdown observed at these places seems to be mainly due to local internal properties of traffic flow discussed in the present paper. The complex spatial structure of the induced congested phase including moving wide and narrow jams reported by Kerner [13] should emerge above the detectors upstream. We consider in detail the data obtained for one of these bottlenecks (site "A" in paper [12]). The paired detectors were located in each of the three lanes and were instrumented to provide vehicle count and speed estimates continuously at 20 s intervals. A breakdown event was fixed via the velocity drop below 90 km/h, the middle point of a certain gap in the velocity field visually separating the congested and free traffic flow states. Besides, only those disturbances that caused the average speed over all the lanes to drop below 90 km/h for a period of 5 min or more were considered a true breakdown. The latter enabled the authors to filter out large amplitude fluctuations in the mean vehicle velocity not leading to the traffic breakdown. Figure 9 exhibits the obtained probability (relative frequency) of the traffic breakdown events during 5 min and 15 min intervals vs the traffic volume partitioned within 100 veh/h/l steps. We note that Fig. 9 does not show the available 1 min interval data because the corresponding breakdown probability is not significant for all the observed values of traffic volume except for the upper boundary 2800 veh/h/l, certainly due to its rare occurrence.

The continuous curves in Fig. 9 present our attempts to fit the obtained theoretical dependence in the given experimental data. Namely, these curves describe the breakdown probability estimated as $\mathcal{F}_{\rm bd} = T_{\rm obs}\nu_{\rm bd}$, where the latter cofactor is given by the expression (33) within the replacement $\Delta \Leftarrow (j - j_{c1})/(j_{c2}-j_{c1})$ [see expression (34)] and we have used the following values of the critical traffic volumes $j_{c1} = 1200$ veh/h/l and $j_{c2} = 3400$ veh/h/l, the characteristic breakdown time $\tau_{\rm bd} = 2.5$ min [see expression (39)], and set the product $n_0(\tau_\infty-\tau_0)/\tau_0 = 25$. Keeping in mind the aforesaid, all the adopted values are quite reasonable, including the estimate $j_{c2} = 3400$ veh/h/l that seems extremely high at

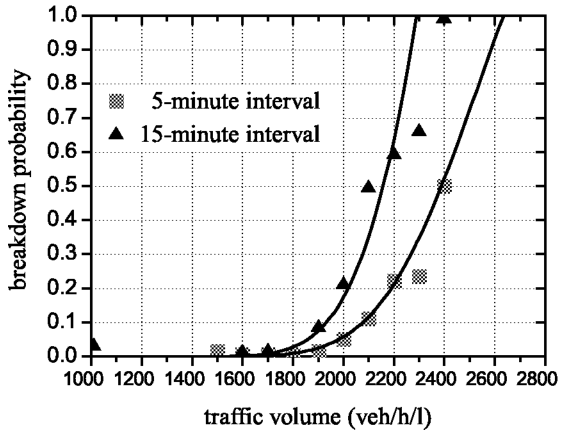

FIG. 9. The traffic breakdown probability during 5 min and 15 min intervals vs the traffic volume (after Lorenz and Elefteriadou [12]). The continuous curves present the dependence (33) fitting these experimental data in the frameworks of the replacement $\Delta \Leftarrow (j-j_{c1})/(j_{c2}-j_{c1})$ [see expression (34)] under the following values of the critical traffic volumes $j_{c1}=1200$ veh/h/l and $j_{c2}=3400$ veh/h/l, the characteristic breakdown time $\tau_{\rm bd}=2.5$ min, and the product $n_0(\tau_\infty-\tau_0)/\tau_0=25$.

first glance. Indeed, this value is no more than a result of approximating the $j(\rho)$ dependence by a linear function formally into the region of high vehicle densities.

IV. SUMMARY AND CONCLUSION

We have considered the traffic breakdown phenomenon regarded as a random process developing via the nucleation mechanism. The origin of critical jam nuclei proceeds in a metastable phase of traffic flow and seems to be located inside a not too large region on a highway, for example, in the close proximity of a highway bottleneck [3,6]. The induced complex structure of the congested traffic phase is located upstream of the bottleneck [13]. Keeping these properties in mind, we have applied the probabilistic model regarding the jam emergence as the development of a large car cluster on highway. In these terms the traffic breakdown proceeds through the formation of a certain car cluster of critical size in the metastable vehicle flow, which enabled us to confine ourselves to the single cluster model.

We assumed that, first, the growth of the car cluster is governed by attachment of cars to the cluster whose rate is mainly determined by the mean headway distance between the cars in the vehicle flow and, may be, also by the headway distance in the cluster. Second, the cluster dissolution is determined by the car escape from the cluster whose rate depends on the cluster size directly. To justify the latter assumption we apply the modern notion of the traffic flow structure (see Refs. [1–3]). Namely, the jam emergence goes mainly through the sequence of two phase transitions: free flow → synchronized mode → stop-and-go pattern [7]. Both of these transitions are of the first order, i.e., they exhibit breakdown, hysteresis, and nucleation effects [6]. Therefore considering the final stage of the jam emergence we have to regard the synchronized mode as the metastable phase exactly inside which a critical jam nucleus appears due to random fluctuations. The synchronized mode is characterized by strong multilane correlations in the car motion and, as a result, all the vehicles in a certain effective cluster spanning over all the highway lanes move as a whole. Thus the proposed probabilistic description deals with actually macrovehicles comprising many individual cars. The available single-vehicle experimental data [8] present the correlation characteristics of the synchronized mode, which have enabled us to estimate the characteristic dimension $n_0 \sim 20$–30 of the car cluster entering the dependence of the car detachment rate on the cluster size. Namely, for small car clusters, $n \lesssim n_0$, the characteristic detachment time τ_0 should be substantially less than this time τ_∞ for large clusters, $n \gg n_0$.

We have written the appropriate master equation for the cluster distribution function and have analyzed the formation of the critical car cluster due to the climb over a certain potential barrier. The inequality $n_0 \gg 1$ has opened for us the way to convert the discrete master equation to the appropriate Fokker-Planck equation and find all the required characteristics of the traffic breakdown.

The obtained results are compared with the available experimental data and, in detail, with the probability of traffic breakdown in the vicinity of bottlenecks vs the traffic volume presented by Lorenz and Elefteriadou [12]. It turned out that the theoretical curves can be fitted closely to the given experimental data using values of the main parameters chosen based on the general properties of the traffic flow not related directly to the breakdown dynamics. In particular, first, we have demonstrated that the characteristic internal time scale $\tau_{\rm bd}$ of the breakdown development is about $\tau_{\rm bd} \sim n_0 \tau_\infty$ (we recall that $\tau_\infty \sim 2$ s is the characteristic time during which a car can individually leave a cluster). Hence we get the estimate of the breakdown time scale about 1 min. The latter justifies the widely used probabilistic technique of the breakdown investigation based on fixing this event during a time interval of several minutes. Second, the proposed model explains why the traffic breakdown as a probabilistic phenomenon is observed inside a sufficiently wide interval of the traffic volume, namely, the thickness Δj of this layer can attain its low boundary j_{c1} in magnitude. The matter is that $\Delta j / j_{c1} \sim (\tau_\infty - \tau_0)/\tau_0 \sim 1$.

Both of these results have been obtained practically without any fitting parameters applying only to the general assumptions on the traffic flow theory. Namely, we assumed the traffic breakdown to develop inside the synchronized mode of traffic flow and used its characteristic correlation properties. Thereby a highway bottleneck seems to affect mainly the particular properties of the critical jam nuclei, which increases the probability of the jam emergence near highway bottlenecks. The physics of the traffic breakdown must be the same for homogeneous and heterogeneous roads.

Concluding the aforesaid, we state that traffic breakdown is a mesoscopic process, as it must be for the synchronized mode, whose characteristic spatial and temporal scales correspond to car clusters made of a large number of vehicles.

ACKNOWLEDGMENTS

The authors would like to thank Peter Wagner (Berlin) for useful comments and criticism. This work was also sup-

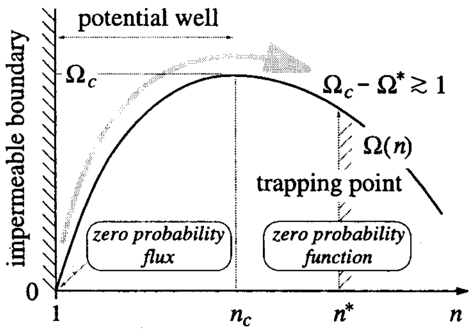

FIG. 10. Escaping problem simulating the critical nucleus formation

ported in part by the Russian Foundation of Basic Research, Grant No. 01-01-00439, and the Russian High School Scientific Program, Grant No. 05.03.01.30.

APPENDIX: ESCAPING RATE FROM A BOUNDARY WELL

In Sec. II C we have obtained the Fokker-Planck equation (25) governing the evolution of car clusters treated as random wandering in the space of their size n. It has turned out that near the threshold the precluster domain is separated from the large cluster region by a potential barrier Ω_c, therefore the formation of the clustered phase should proceed through the nucleation mechanism (Fig. 7). In other words, for a large cluster to emerge on the road its critical nucleus n_c has to arise via random fluctuations of the cluster size in the precluster region. Thereby in order to describe the cluster formation we need an expression specifying the rate of the critical nucleus generation, being the subject of the present appendix.

Mathematically the description of the critical nucleus generation is equivalent to the problem of a particle escaping form the corresponding potential barrier (Fig. 10). Thereby the rate of the critical nucleus generation, i.e., the frequency of the traffic breakdown ν_{bd} is represented in terms of the probability density $\mathcal{F}(t)$ for this particle to escape from the potential well at a given time t provided initially, $t=0$, it has been placed near the local minimum (here $n=0$). Namely,

$$\nu_{\mathrm{bd}}=\mathcal{F}(+0), \tag{A1}$$

where the value $+0$ of the argument t means that we consider time scales exceeding substantially the duration of all the transient processes during which the distribution of the particle inside the potential well attains locally quasiequilibrium.

Since the potential relief under consideration is rather special we prefer to recall briefly the way of deriving the probability $\mathcal{F}(t)$ referring a reader to the specific literature (see, e.g., Ref. [19]) for details.

The concept of potential well implies that the barrier is sufficiently high, $\Omega_c \gg 1$, therefore the particle can climb over it due to rare fluctuations lifting the particle to points at the potential barrier where $\Omega(n) \gg 1$. If such an event does not lead to escape, the particle will drift back to the neighborhood of the local minimum $n=0$ whose thickness is specified by the inequality $\Omega(n) \lesssim 1$. Thereby the subsequent attempts of escaping may be considered as being mutually independent. After the particle has climbed over the barrier the force $-d\Omega(n)/dn$ carries it away to distant points, making the return impossible. So from this point of view we may refer to the particle being inside the potential well or having escaped from it as two of its possible states without specifying the particular position. Therefore the probability $\mathcal{P}(t-t')$ that the particle remains inside the well at time t, if it has been placed there at time t', obeys the equation

$$\mathcal{P}(t)=\mathcal{P}(t-t')\mathcal{P}(t') \quad \text{for} \quad 0<t'<t.$$

We get the general expression for the function $\mathcal{P}(t)$,

$$\mathcal{P}(t)=\exp\left(-\frac{t}{\tau_{\mathrm{life}}}\right),$$

where τ_{life} is a certain constant specified by the particular properties of a potential well. The latter formula gives us immediately the general form of the escape probability

$$\mathcal{F}(t)=-\frac{d\mathcal{P}(t)}{dt}=\frac{1}{\tau_{\mathrm{life}}}\exp\left(-\frac{t}{\tau_{\mathrm{life}}}\right). \tag{A2}$$

In order to find the lifetime τ_{life} we will deal with the Laplace transform $\mathcal{F}_L(s)$ of the escape probability $\mathcal{F}(t)$,

$$\mathcal{F}_L(s):=\int_0^\infty dt \exp(-st)\mathcal{F}(t)=\frac{1}{1+s\,\tau_{\mathrm{life}}}, \tag{A3}$$

whence it follows that in the expansion of $\mathcal{F}_L(s)$ with respect to s around the point $s=0$,

$$\mathcal{F}_L(s)=1-s\,\tau_{\mathrm{life}}+\cdots, \tag{A4}$$

the first order term directly contains the desired lifetime as the coefficient.

Following the standard approach [19] we reduce the escaping problem to finding the first passage time probability. In other words, we assume the particle never comes back to the potential well if, after climbing the barrier, it reaches points where $\Omega_c-\Omega(n)\gtrsim 1$ (Fig. 10). The particle may be withdrawn from the consideration or, what is the same, it will be trapped when it reaches for the first time any fixed point n^* in this region. The time it takes for the particle to reach the point n^* after overcoming the barrier at the critical point n_c is ignorable in comparison with the characteristic waiting time for critical fluctuations. Thereby the function $\mathcal{F}(t)$ specifies actually the probability of passing (reaching) the point n^* for the first time at the time moment t. This construction enables us to introduce a more detailed relative function $\mathcal{F}(n,t)$ giving the probability for the particle initially placed at the point $0<n<n^*$ to reach first the right boundary n^* of the region under consideration at the time moment t. The left boundary $n=0$ is impermeable for the particle. Then using the standard technique based on the backward Fokker-Planck equation conjugated with Eq. (25) we obtain the governing equation for the function $\mathcal{F}_L(n,s)$,

$$\tau_\infty s \mathcal{F}_L = \partial_n^2 \mathcal{F}_L - [\partial_n \Omega(n)][\partial_n \mathcal{F}_L] \tag{A5}$$

subject to the boundary conditions

$$\mathcal{F}_L(0,s) = \mathcal{F}_L(n^*,s) = 1. \tag{A6}$$

It directly follows that the first order term $\varphi(n)$ in the expansion of the Laplace transform $\mathcal{F}_L(n,s)$ with respect to s,

$$\mathcal{F}(n,s) = 1 - s\varphi(n),$$

obeys in turn the equation

$$\partial_n^2 \varphi(n) - [\partial_n \Omega(n)][\partial_n \varphi(n)] = -\tau_\infty \tag{A7}$$

subject to the boundary conditions

$$\partial_n \varphi(0) = 0 \quad \text{and} \quad \varphi(n^*) = 0. \tag{A8}$$

The solution of the system (A7) and (A8) has the form

$$\varphi(n) = \tau_\infty \int_n^{n^*} dn' e^{\Omega(n')} \int_0^{n'} dn'' e^{-\Omega(n'')} \tag{A9}$$

and the value $\varphi(0)$ gives us the desired lifetime,

$$\tau_{\text{life}} = \varphi(0). \tag{A10}$$

Inside the potential well the function $\varphi(n)$ takes practically a constant value mainly contributed by the points n'' belonging to the well bottom, i.e., to the region $\Omega(n'') \lesssim 1$ and by the points n' located near the top of the potential barrier where $\Omega(n_c) - \Omega(n') \lesssim 1$. This feature leads us immediately to the approximation

$$\tau_{\text{life}} \approx \sqrt{2\pi}\, \tau_\infty [|\partial_n^2 \Omega(n_c)|]^{-1/2} [\partial_n \Omega(0)]^{-1} e^{\Omega(n_c)}, \tag{A11}$$

which is the main result of the present appendix.

In particular, for the potential $\Omega(n)$ specified by expression (21) or (26), formula (A11) gives

$$\tau_{\text{life}} \approx \sqrt{2\pi n_0}\, \tau_\infty \left(\frac{\tau_0}{\tau_\infty - \tau_0} \right)^{3/2} \times (1 - \phi[x_c])^{-1} (|\phi'[x_c]|)^{-1/2} e^{\Omega(n_c)}. \tag{A12}$$

Formulas (A1), (A2), (A12), and (27) give us expression (32).

[1] B.S. Kerner, Phys. World **12**, 25 (1999).
[2] B.S. Kerner, Transp. Res. Rec. **1678**, 160 (1999).
[3] B.S. Kerner, J. Phys. A **33**, L221 (2000).
[4] B.S. Kerner and H. Rehborn, Phys. Rev. E **53**, R1297 (1996).
[5] B.S. Kerner and H. Rehborn, Phys. Rev. E **53**, R4275 (1996).
[6] B.S. Kerner and H. Rehborn, Phys. Rev. Lett. **79**, 4030 (1997).
[7] B.S. Kerner, Phys. Rev. Lett. **81**, 3797 (1998).
[8] L. Neubert, L. Santen, A. Schadschneider, and M. Schreckenberg, Phys. Rev. E **60**, 6480 (1999).
[9] L. Elefteriadou, R.P. Roess, and W.R. McShane, Transp. Res. Rec. **1484**, 80 (1995).
[10] R. Kühne, and N. Ansttet, in *Proceedings of the 14th International Symposium on Transportation and Traffic Theory,* edited by A. Ceder (Pergamon, New York, 1999), p. S177.
[11] B. Persaud, S. Yagar, and R. Brownlee, Transp. Res. Rec. **1634**, 64 (1998).
[12] M. Lorenz and L. Elefteriadou, Transp. Res. Circ. **E-C018**, 84 (2000).
[13] B.S. Kerner, Transp. Res. Rec. **1710**, 136 (2000).
[14] I.A. Lubashevsky and R. Mahnke, Phys. Rev. E **62**, 6082 (2000).
[15] R. Mahnke and N. Pieret, Phys. Rev. E **56**, 2666 (1997).
[16] R. Mahnke and J. Kaupužs, Phys. Rev. E **59**, 117 (1999).
[17] D. Helbing, A. Hennecke, V. Shvetsov, and M. Treiber, Transp. Res. Rec. **35**, 183B (2001).
[18] P. Schick and R. Kühne, *Arbeiten aus dem Institut für Straßen- und Verkehrswesen 2000/2001* (*Research at the Institute of Road and Transportation Studies*) (Universität Stuttgart, Stuttgart, 2001), Vol. 28, p. 1.
[19] C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer, Berlin, 1994).