Interpreting the wide scattering of synchronized traffic data by time gap statistics

Katsuhiro Nishinari,[1,2,3] Martin Treiber,[1] and Dirk Helbing[1]

[1]*Institute for Economics and Traffic, Dresden University of Technology, 01062 Dresden, Germany*
[2]*Department of Applied Mathematics and Informatics, Ryukoku University, Shiga 520-2194, Japan*
[3]*Institute for Theoretical Physics, University of Cologne, 50923 Köln, Germany*

Based on the statistical evaluation of experimental single-vehicle data, we propose a quantitative interpretation of the erratic scattering of flow-density data in synchronized traffic flows. A correlation analysis suggests that the dynamical flow-density data are well compatible with the so-called jam line characterizing fully developed traffic jams, if one takes into account the variation of their propagation speed due to the large variation of the netto time gaps (the inhomogeneity of traffic flow). The form of the time gap distribution depends not only on the density, but also on the measurement cross section: The most probable netto time gap in congested traffic flow upstream of a bottleneck is significantly increased compared to uncongested freeway sections. Moreover, we identify different power-law scaling laws for the relative variance of netto time gaps as a function of the sampling size. While the exponent is -1 in free traffic corresponding to statistically independent time gaps, the exponent is about $-2/3$ in congested traffic flow because of correlations between queued vehicles.

Recently, the statistical physics and nonlinear dynamics of driven many-particle systems have been key disciplines in the discovery, interpretation, and simulation of phenomena in traffic flows [1–3]. The observed instability mechanisms, jamming, segregation, breakdown, and clustering phenomena are now viewed as a paradigm for similar phenomena in granular and colloidal physics [3–5], in biology (social insects) [6], logistics (instability of supply chains), and economics (business cycles) [7,4]. Empirical investigations [8–11] have particularly stimulated the development of traffic theory [11–13]. For example, it was possible to identify constants of traffic dynamics such as the propagation speed $C \approx -15$ km/h of wide jams [14]. According to analytical calculations by Kerner *et al.* for a macroscopic traffic model assuming uniformly behaving driver-vehicle units with identical parameters, fully developed wide traffic jams should be characterized by a self-organized flow-density relation

$$J(\rho, T, \rho_{max}) = \frac{1}{T}\left(1 - \frac{\rho}{\rho_{max}}\right), \qquad (1)$$

where ρ denotes the vehicle density, while the average netto time gap T (the time clearance) and the maximum density ρ_{max} are assumed to be fixed parameters [15]. The dependence of the flow J on the density ρ in Eq. (1) is sometimes referred to as "jam line" $J(\rho)$ and does not necessarily agree with the high-density part of the steady-state flow-density relation $Q_f(\rho)$ for stationary and homogeneous traffic flow, the so-called fundamental diagram [15]. We note, however, that a linear flow-density relation for wide moving jams is not yet fully confirmed by empirical measurements and that aggregation methods used to determine macroscopic traffic data from single-vehicle data have a nontrivial impact on the measurement results [16,17]. Nevertheless, the slope $C = dJ(\rho)/d\rho = -1/(\rho_{max}T)$ is usually identified with the average propagation velocity of wide moving jams [15]. In contrast to wide moving jams, synchronized traffic flow is a

form of congested flow which is mostly found upstream of inhomogeneities (e.g., freeway bottlenecks) and claimed to show a completely different behavior [9]. It is characterized by an erratic motion of time-dependent flow-density data in a two-dimensional area (and a synchronization of the time-dependent average vehicle velocities among neighboring lanes) [8]. More specifically, in synchronized flow, the slopes

$$C(t_{k+1}) = [Q(t_{k+1}) - Q(t_k)]/[\rho(t_{k+1}) - \rho(t_k)] \qquad (2)$$

of the lines connecting flow-density data measured at a given freeway cross section at successive times t_k and t_{k+1} erratically take on positive and negative values, characterizing a complex spatiotemporal dynamics [8]. This is, in fact, one of the most controversial subjects in traffic theory. It has not only been the reason for Kerner's fundamental criticism of all traffic models assuming a fundamental (steady-state) relation $Q_f(\rho)$ between the flow Q and the density ρ (i.e., of the vast majority of models suggested up to now) [18,1]. It has also triggered a flood of publications in physics journals with various suggestions how to describe this wide scattering. The proposed interpretations reach from shock waves propagating forward or backward with speed $C(t)$ [8,18], changes in the driving behavior in response to the traffic situation (including "frustration effects") [19,17], anticipation effects [20,21], nonunique solutions [22], a trapping of vehicles [11], or multiple metastable oscillating states [13]. Nonunique solutions are also expected for car-following models, in which the vehicle acceleration dv/dt is not a unique function of the distance or velocity [23]. Other explanations based on a mixture of different vehicle types such as cars and trucks [24] or a heterogeneity of time headways [25] are easier accessible to empirical verification. In the following, we will therefore focus on these. The particular approach of this study is its restriction to empirical analyses to avoid theory-driven interpretations. Although it has been seriously questioned that relation (1) would also be applicable

to flow-density data of synchronized congested traffic flow, by correlation analysis we will surprisingly find a strong empirical support for the validity of relation (1), if the average netto time gap is treated as a dynamical variable.

This requires single-vehicle data, which were recorded by double induction-loop detectors for the Dutch freeway A9. The measured data include, for each lane, the passage time t_i^0 of vehicle i, its velocity v_i, and its length l_i. From this, we determine the individual netto time gaps as $T_i = t_i^0 - t_{i-1}^0 - l_i/v_i$. As discussed in Ref. [27], we obtain macroscopic quantities by averaging over $N=50$ successive vehicles, which gives a better statistics than averaging over a fixed time interval. Thus, we obtain the empirical traffic flow at time $t_k = (1/N)\Sigma_{i=(k-1)N+1}^{kN} t_i^0$ by $Q(t_k) = 1/\langle t_i^0 - t_{i-1}^0 \rangle = N/(t_{kN}^0 - t_{(k-1)N}^0)$ and the average netto time gap by

$$T(t_k) = \langle T_i \rangle = \frac{1}{N} \sum_{i=(k-1)N+1}^{kN} T_i = T_{(k-1)N}^{(N)}. \qquad (3)$$

Likewise, we define the vehicle density ρ and the maximum density $\rho_{\max}$ by

$$\frac{1}{\rho(t_k)} = \frac{1}{N} \sum_{i=(k-1)N+1}^{kN} v_i(t_i^0 - t_{i-1}^0), \qquad (4)$$

$$\frac{1}{\rho_{\max}(t_k)} = \langle l_i \rangle = \frac{1}{N} \sum_{i=(k-1)N+1}^{kN} l_i. \qquad (5)$$

For a given loop detector on the left lane, we consider all time intervals with congested traffic [defined by $\rho(t_k) \geq 45$ vehicles per kilometer and lane] and compare the temporal changes $\Delta Q^{(k+1)} = [Q(t_{k+1}) - Q(t_k)]$ of the empirically measured flow with the changes predicted by Eq. (1). We will compare three hypotheses by correlation analysis: (i) ρ is treated as an independent variable defined by (4), while T and $\rho_{\max}$ are treated as parameters, (ii) ρ and $1/T$ are independent variables defined directly from the single-vehicle data via Eqs. (4) and (3), while $\rho_{\max}$ is a parameter, (iii) ρ, $1/T$, and $1/\rho_{\max}$ are independent variables defined by Eqs. (3)–(5). The first-order Taylor approximations for the temporal change $\Delta J_{(\alpha)}^{(k+1)} = [J_{(\alpha)}(t_{k+1}) - J_{(\alpha)}(t_k)]$ of the corresponding theoretical vehicle flows are obtained by differential of relation (1) with respect to the independent variables, but not the parameters:

$$\Delta J_{(i)}^{(k+1)} = -\frac{1}{\rho_{\max}T}[\rho(t_{k+1}) - \rho(t_k)], \qquad (6)$$

$$\Delta J_{(ii)}^{(k+1)} = \Delta J_{(i)}^{(k+1)} + \left[1 - \frac{\rho(t_k)}{\rho_{\max}}\right]\left[\frac{1}{T(t_{k+1})} - \frac{1}{T(t_k)}\right],$$

$$\Delta J_{(iii)}^{(k+1)} = \Delta J_{(ii)}^{(k+1)} - \frac{\rho(t_k)}{T(t_k)}\left[\frac{1}{\rho_{\max}(t_{k+1})} - \frac{1}{\rho_{\max}(t_k)}\right].$$

Our statistical analysis gives the following correlations Corr $(\Delta Q, \Delta J_{(\alpha)}) = (\Sigma_k \Delta Q^{(k)} \Delta J_{(\alpha)}^{(k)})/[(\Sigma_k \Delta Q^{(k)2})(\Sigma_k \Delta J_{(\alpha)}^{(k)2})]^{1/2}$ for the hypotheses $\alpha = $ (i)–(iii): Corr$(\Delta Q, \Delta J_{(i)}) = 0.347$,
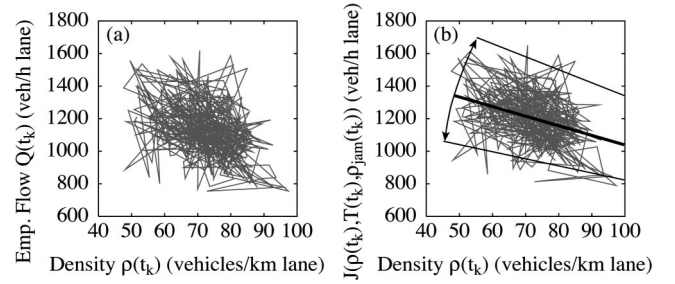

FIG. 1. (Color online) The two-dimensional scattering of empirical flow-density data in synchronized traffic flow of high density $\rho \geq 45$ veh/km/lane [see (a)] is well reproduced by the jam relation (1), when not only the variation of the density ρ, but also the empirically measured variation of the average time gap T and the maximum density $\rho_{\max}$ are taken into account [see (b)]. The similarity between the experimental data and relation (1) is partly because the density $\rho(t_k)$ plotted in (a) and (b) (the x value) is determined with the same formula (4), but the agreement of the empirical flow $Q = 1/\langle T_i + l_i/v_i \rangle$ and of the theoretical relation $J(\rho, T, \rho_{\max}) = [1 - \langle l_i \rangle/\langle v_i(T_i + l_i/v_i)\rangle]/\langle T_i \rangle$ (the y values) is not trivial, as even low-order approximations of these formulas differ. The pure density dependence $J(\rho)$ (thick black line) is linear and cannot explain a two-dimensional scattering. However, variations of the average time gap T change its slope $-1/(\rho_{\max}T)$ (see arrows), and about 95% of the data are located between the thin lines $J(\rho, \bar{T} \pm 2\Delta T, 1/l) = (1 - \rho l)/(\bar{T} \pm 2\Delta T)$, where $l = 3.6$ m is the average vehicle length, $\bar{T} = 2.25$ s the average time gap, and $\Delta T = 0.29$ s the standard deviation of T. The predicted form of this area is club shaped, as demanded by Kerner [18].

Corr$(\Delta Q, \Delta J_{(ii)}) = 0.918$, Corr$(\Delta Q, \Delta J_{(iii)}) = 0.938$. As a consequence, hypothesis (i) assuming a linear variation of the flow with the density is in fact a poor description of "synchronized" congested traffic data due to their two-dimensional scattering. However, hypothesis (ii), taking into account the dynamical variation of the average netto time gap T, yields a strong correlation, which indicates that we have identified the main reason for the wide scattering (see Fig. 1). Taking into account a variation of $\rho_{\max}$ with the truck fraction improves the correlation coefficient only a little. Hence, the maximum density, which influences the density offset of the jam line, is an unimportant explanatory variable. In contrast, the netto time gap T determining the slope of the jam line [see Eq. (1)] is highly significant. For this reason, we will investigate the characteristic features of the netto time gap distribution more closely.

The variation of the average time gaps with time can only be relevant, when it is large. The time gap distribution is, in fact, surprisingly wide, and it has so-called heavy tails (see Fig. 2). To investigate this, let us calculate the variance of the sample-averaged time gaps as a function of the sampling size N,

$$\text{var}(T) = \frac{1}{I-N+1} \sum_{j=0}^{I-N} (T_j^{(N)} - \bar{T})^2, \qquad (7)$$

where $(j+1)$ runs over the first-vehicle indices of all possible samples of N consecutive vehicles in the dataset of size
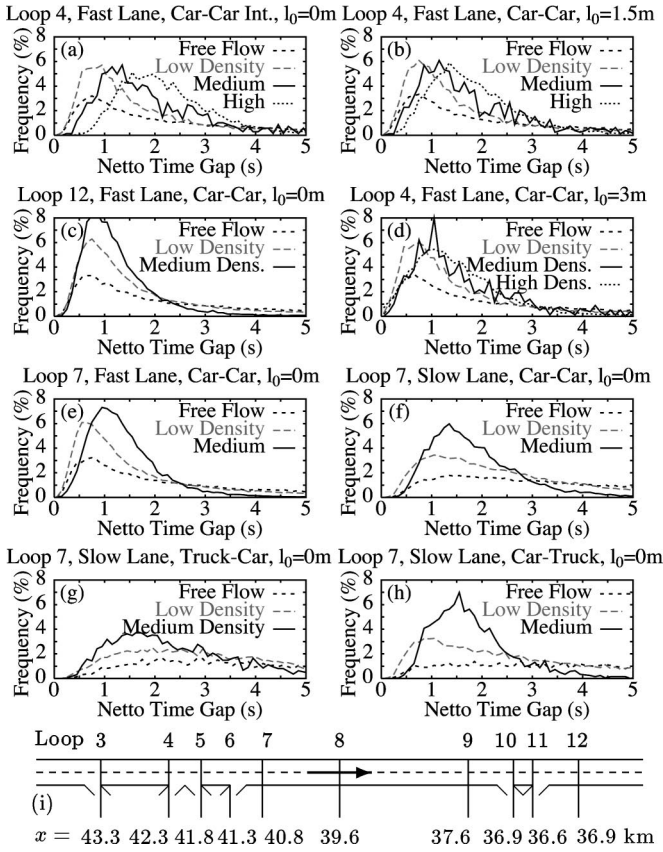
FIG. 2. (Color online) (Netto) time gap distributions at different cross sections of the Dutch freeway A9 [see the sketch in (i)], separately measured for the fast and the slow lane, for car-car, car-truck, and truck-car interactions, and for different density regimes, namely free flow ($\rho<10$ veh/km/lane), low densities (free traffic with $10<\rho<25$ veh/km/lane), medium densities (mostly congested traffic with $25<\rho<50$ veh/km/lane), and high densities (congested traffic with $\rho>50$ veh/km/lane, only observed upstream of serious bottlenecks). Densities were determined as 50-vehicle averages, and trucks were defined by a vehicle length $l_i \geq 6$ m. The time gaps of trucks with respect to cars are, on average, considerably higher than the ones of cars with respect to other vehicles, which causes a large variation of individual time gaps. However, even the variation of time gaps among cars is considerable. A remarkable point is the increase in the most probable time gap from 0.75 s in free traffic (on the fast lane) to considerably higher values in congested traffic upstream of a bottleneck [1.2 s at medium densitites and 1.9 s at high densities at loop 4 compared to 0.75 s at loop 12 in all density regimes and 1 s at loop 7 at medium densities, see (a), (c), (e); the values for the slow lane are somewhat higher]. This "frustration effect" is still slightly active immediately downstream of a bottleneck (loop 7), in contrast to freeway sections, which are never seriously congested (loop 12). It is also observed for finite minimum bumper-to-bumper distances of, for example, $l_0=1.5$ or 3 m, which is considered by replacing l_i by (l_i+l_0) everywhere.

I, $T_j^{(N)}=(1/N)\Sigma_{i=j+1}^{j+N}T_j$ is defined as moving average of the time gaps of the N next vehicles (Veh), and $\bar{T}=\Sigma_{i=1}^{I}T_i/I$ $=T_0^{(I)}$ is the global average. To account for artifacts caused by the daily variation of traffic flow, we applied a high-pass filter (with cutoff period $N_c=500$) to the single-vehicle data
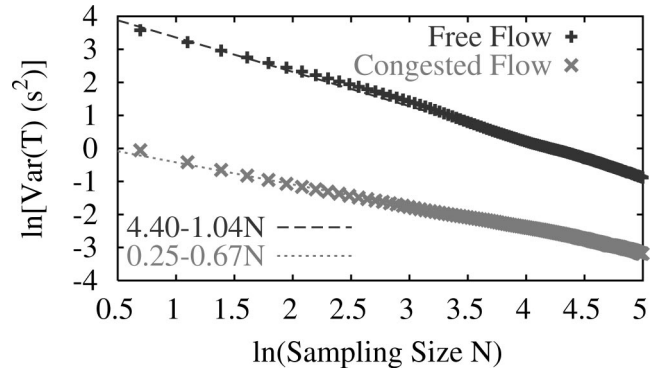

FIG. 3. (Color online) Variance of the average time gaps T as a function of the sampling size N, separately measured for free and congested flows at loop 4 on the fast lane. While in free flow, the relative variance is consistent with the expected power-law exponent of $\gamma=-1$, it decreases much slower with N in the congested traffic regime, namely with an exponent of $\gamma\approx-0.67$, causing the significant variation of T.

before calculating Eq. (7), which limits the scaling range of relation (8). For a given N, variance (7) decreases with increasing density (see Figs. 2 and 3), as less and less space is available for time gaps larger than the safe time gap, but the variance is also a function of the sampling size N. For free traffic flow (with $\rho<15$ veh/km/lane), we observe the power-law behavior $\mathrm{var}(T)\propto1/N$, as expected for statistically independent time gaps in free traffic. At high vehicle densities (with $\rho>35$ veh/km/lane), however, we find the power law

$$\mathrm{var}(T)\propto N^\gamma \tag{8}$$

with an exponent $\gamma\approx-0.67$ (see Fig. 3). That is, the relative variance decreases much slower with the sampling size than expected, implying that *the time gaps do not average well and fluctuations of average time gaps remain significant for reasonable sampling sizes N.* This is related to surprisingly wide time gap distributions and results from correlations between queued vehicles, which are probably due to platoon formation [11], but may also be caused by dynamical vehicle interactions. The reproduction of this scaling law will be a serious test for microscopic traffic simulation models, which are needed to reveal the detailed mechanism behind it.

According to Fig. 2, the time gap distribution is not only dependent on the density, but also on the measurement cross section (more specifically, on the relative location with respect to bottlenecks of the freeway). It has been suggested [26] that the increase of the most probable *brutto* time gap (time headway) $\Delta t_i^0=t_i^0-t_{i-1}^0=T_i+l_i/v_i$ in congested traffic compared to free traffic [10,27] is due to the drop of the vehicle velocity v_i, increasing l_i/v_i. However, an increase is also observed for the most probable *netto* time gaps T_i (see Fig. 2). This increase is most pronounced at cross sections upstream of bottlenecks (e.g., on ramps), where vehicles are queuing. We interpret this increase of the netto time gaps as *"frustration effect"* of drivers after a considerable queuing time, which is supported by other observations as well

[3,28]. The frustration effect is most significant at freeway sections, where the accumulated time of driving under congested conditions is high, and it decays with increasing recovery time. At freeway sections, where serious congestion never occurs, there is no relevant increase in the most probable netto time gap with growing density (see Fig. 2). Finally, we note that the time gaps in front of long vehicles ("trucks") are much higher, on average, than in front of short vehicles ("cars"), as expected (see Fig. 2). This supports the theory suggested in Ref. [24].

In conclusion, we have found an interpretation of the wide scattering of flow-density data in synchronized congested traffic based on the jam relation $J(\rho)$, but taking into account the time-dependent variation of empirical netto time gaps among successive cars. This variation is related to time-dependent changes of the slope $C = -1/(\rho_{max}T)$ of the jam line and causes the two-dimensional scattering together with the variation of the vehicle density. The surprisingly strong variation of the average time gaps was due to the fact that the power law relating the variance of the time gaps with the sampling size had a considerably smaller exponent ($\gamma \approx -2/3$) than expected ($\gamma = -1$). Hence, distinguishing different forms of congested traffic based on the scattering is questionable. We also found that the increase in the most probable time gap in congested traffic is site dependent and not attributed to a drop in the vehicle velocity, indicating a "frustration effect." Traffic models considering the dynamical changes of the netto time gaps can reproduce empirical observations more realistically [24,17].

[1] D. Chowdhury, L. Santen, and A. Schadschneider, Phys. Rep. **329**, 199 (2000); T. Nagatani, Rep. Prog. Phys. **65**, 1331 (2002).

[2] D. Helbing, Rev. Mod. Phys. **73**, 1067 (2001).

[3] *Traffic and Granular Flow '01*, edited by M. Fukui, Y. Sugiyama, M. Schreckenberg, and D.E. Wolf (Springer, Berlin, 2003).

[4] *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics*, edited by D. Helbing, H.J. Herrmann, M. Schreckenberg, and D.E. Wolf (Springer, Berlin, 2000).

[5] J. Dzubiella, G.P. Hoffmann, and H. Löwen, Phys. Rev. E **65**, 021402 (2002).

[6] I.D. Couzin and N.R. Franks, Proc. R. Soc. London, Ser. B **270**, 139 (2003).

[7] C. Daganzo, *A Theory of Supply Chains* (Springer, New York, 2003); D. Helbing, New J. Phys. **5**, 90.1 (2003); Physica A **287**, (3/4) (2000); special issue on economic dynamics from the physics point of view, edited by F. Schweitzer and D. Helbing.

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

[9] B.S. Kerner and H. Rehborn, Phys. Rev. Lett. **79**, 4030 (1997); **81**, 3797 (1998).

[10] L. Neubert, L. Santen, A. Schadschneider, and M. Schreckenberg, Phys. Rev. E **60**, 6480 (1999).

[11] I. Lubashevsky, R. Mahnke, P. Wagner, and S. Kalenkov, Phys. Rev. E **66**, 016117 (2002).

[12] H.Y. Lee, H.-W. Lee, and D. Kim, Phys. Rev. Lett. **81**, 1130 (1998); D. Helbing and M. Treiber, *ibid.* **81**, 3042 (1998); D. Helbing, A. Hennecke, and M. Treiber, *ibid.* **82**, 4360 (1999); N. Mitarai and H. Nakanishi, Phys. Rev. Lett. **85**, 1766 (2000).

[13] E. Tomer, L. Safonov, and S. Havlin, Phys. Lett. A **84**, 382 (2000).

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

[15] B.S. Kerner and P. Konhäuser, Phys. Rev. E **50**, 54 (1994); B.S. Kerner, S.L. Klenov, and P. Konhäuser, *ibid.* **56**, 4200 (1997).

[16] M. Treiber and D. Helbing, Cooper@tive Tr@nsport@tion Dyn@mics **1**, 2.1 (2002). (The journal contents are freely available at www.trafficforum.org/journal/)

[17] M. Treiber and D. Helbing, Phys. Rev. E (to be published).

[18] B.S. Kerner, in *Proceedings of the 3rd International Symposium on Highway Capacity*, edited by R. Rysgaard (Road Directorate, Denmark, 1998), Vol. 2, pp. 621; B.S. Kerner, in *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics* in (Ref. [4]) p. 253; B.S. Kerner, Transp. Res. Rec. **1710**, 136 (2000).

[19] S. Krauß, P. Wagner, and C. Gawron, Phys. Rev. E **55**, 5597 (1997).

[20] C. Wagner, Physica A **260**, 218 (1998).

[21] W. Knospe, L. Santen, A. Schadschneider, and M. Schreckenberg, J. Phys. A **33**, L477 (2000).

[22] P. Nelson, Phys. Rev. E **61**, R6052 (2000); P. Nelson and A. Sopasakis, Transp. Res., Part B Methodol. **32**, 589 (1998).

[23] R. Wiedemann, *Simulation des Straßenverkehrsflusses* (Institut für Verkehrswesen, Universität Karlsruhe, Karlsruhe, Germany, 1974); B.S. Kerner and S.L. Klenov, J. Phys. A **35**, L31 (2002); B.S. Kerner, S.L. Klenov, and D.E. Wolf, *ibid.* **35**, 9971 (2002).

[24] M. Treiber and D. Helbing, J. Phys. A **32**, L17 (1999).

[25] J.H. Banks, Transp. Res. Rec. **1678**, 128 (1999).

[26] J.H. Banks, Transp. Res. Rec. **1320**, 234 (1991).

[27] B. Tilch and D. Helbing, in *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics* (Ref. [4]), p. 333.

[28] M. Krbalek and D. Helbing, e-print cond-mat/0301484.