Headways in traffic flow: Remarks from a physical perspective

Milan Krbalek*
Faculty of Nuclear Sciences and Physical Engineering, Trojanova 13, Prague, Czech Republic

Petr Šeba†
Institute of Physics, Czech Academy of Science, Cukrovarnicka 10, Prague, Czech Republic

Peter Wagner‡
German Aerospace Center (DLR), Transport Research Linder Höhe, 51170 Köln, Germany
(Received 25 December 2000)

Traffic flow can be understood as a realization of a broad class of one dimensional physical systems, where a hard core repulsive interaction competes with a longer ranged attraction between the particles. It can be shown rigorously that the statistical properties of such systems in thermal equilibrium are well described by a family of distributions that stems from the random matrix theory. Analyzing the traffic data from different sources, we show that traffic on real roads belongs to that class of random matrix distributions. Also, various traffic simulation models show a similar behavior. It is demonstrated in such a way that the headway distribution of a highway traffic, that serves usually as a paradigm of systems driven far from equilibrium, is reasonably well described by a distribution originating from equilibrium statistical physics.

I. RANDOM MATRIX THEORY AND TRAFFIC

Random matrix theory (RMT) appears to be a very universal instrument. Originally invented to model the energy levels of atomic nuclei it turns out to be useful in a wide range of different systems and occasions. Of special importance is the deep connection between classical chaotic systems and their quantum mechanical counterparts.

This paper reports on a work done on another connection, that between RMT and one dimensional many particle systems, with a special focus on traffic. For illustration, this relation will be discussed first with the help of the Dyson gas where it is known to be exact. The Dyson gas describes the equilibrium properties of a one dimensional system of \( N \) particles confined either to a ring or by a harmonic potential, interacting repulsively through a Coulomb potential,

\[
V = - \sum_{i=j+1,j+2,\ldots,j+h} \ln(|x_i - x_j|),
\]

(1)

where \( x_i \) is the coordinate of the \( i \)th particle and \( h \) denotes the number of the interacting neighbors. In the original model of Dyson the interaction extends over all particle pairs and hence \( h = N \) [1]. Recent modifications use, however, also \( h < N \) [2].

It is well known that the statistical properties of the Dyson gas in thermal equilibrium are exactly described by RMT. In particular, for the heat bath inverse temperature \( \beta = 1 \) or \( \beta = 2 \) the Dyson gas conforms to the orthogonal/unitary ensemble of random matrices, respectively. In this respect two prominent statistical distributions are commonly discussed: the spacing distribution and the number variance. The spacing distribution \( P(s) \) describes the probability density that two neighboring particles are found with mutual distance equal to \( s \). The distribution \( P(s) \) is scaled so that the mean distance equals one: \( \langle s \rangle = \int s P(s) ds = 1 \). It takes into account only the two particle correlations and is therefore very robust and not sensitive to the detailed properties of the system. A more sensitive measure for discussing the correlations between the particles is the number variance \( \Sigma^2(L) = \langle (N(L) - L)^2 \rangle \). Here \( N(L) \) denotes the number of particles contained in an interval of length \( L \). Note that \( \langle N(L) \rangle = L \), due to the fact that the mean particle spacing equals one.

For high temperatures of the gas (\( \beta \to 0 \)) the spacing distribution is Poissonian, \( P(s) = e^{-s} \) and \( \Sigma^2(L) = L \). Upon increasing the heat bath inverse temperature \( \beta \), the interaction between the particles causes what is known as level repulsion in the original context of the theory, that means the probability to find small distances between particles is suppressed and \( P(s) \approx s^\beta \) for small \( s \). For a fixed \( \beta \) the spacing distribution \( P(s) \) is well described by the heuristic formula [3]

\[
P_\beta(s) = A\beta \left( \frac{\pi s}{2} \right)^\beta \exp \left[ -\beta \frac{\pi^2}{16} s^2 - \left( B_\beta - \frac{\beta}{4} \pi \right) s \right],
\]

(2)

where the constant \( A_\beta \) and \( B_\beta \) are determined by requiring \( \int P(s) ds = 1 \) and \( \int s P(s) ds = 1 \). The number variance behaves as

\[
\Sigma^2(L) = L\beta^{-1}.
\]
\[ \Sigma^2(L) \approx \frac{2}{\beta \pi^2} \ln L + C_{\beta} \]  

for large \( L \). Restricting the interaction to neighboring pairs only \( (h = 1) \), the so-called semi-Poisson regime of the Dyson gas (called also a short ranged Dyson gas – SRDG) is encountered. This leads to [2]

\[ P_\beta(s) = \frac{(\beta + 1)^{\beta+1}}{\Gamma(\beta + 1)} s^\beta \exp[-(\beta + 1) s] \]  

and

\[ \Sigma^2(L) \approx \frac{1}{\beta + 1} L + \frac{\beta(\beta + 2)}{6(\beta + 1)^2} \]  

for large \( L \).

The aim is to use the statistical properties of the Dyson gas for the description of what is called headway distribution in the traffic science. Before doing this a few comments are necessary. The particles of the standard Dyson gas are regarded as dimensionless. This is not true in a realistic traffic situation, where the cars occupy (depending on their density) up to 30\% of the available space. Moreover, the potential Eq. (1) is of long-range type, a property that is not justified for realistic transport systems. The influence of this modifications on the equilibrium distribution can be checked using the Metropolis algorithm to reach the thermal equilibrium.

First, for reasonable lengths of the cars the spacing distribution, Eq. (4), holds true provided the spacing \( s \) is measured from bumper to bumper, i.e., excluding the part of the space occupied by the car. On the contrary, the slope of the number variance, Eq. (5), decreases with the increasing length of cars. This is understandable, since the excluded volume causes stronger correlations between the cars. Second, the system is not sensitive to the long range character of the potential Eq. (1). Replacing the potential \( V(x) = \ln(x) \) with \( \ln(x_w) \) for \( x > x_w \) does not change the statistical properties of the gas provided \( x_w \gg 1 \). Additionally, direct numerical integrations of the Dyson gas have been performed showing the same behavior. They are reported in a separate section of this paper.

For the above results the special character of the potential (Coulomb) is not important. It has been recognized that a gas with a different shape of the potential energy (for example, the so-called Yukawa gas [4], or Pechukas gas [5]) leads to the same statistics. It has to be stressed that the Dyson gas models is a one dimensional system. When comparing real data from a multilane traffic to the RMT predictions, it is necessary to exclude all the overtaking cars because they are not correlated with the others. This holds true mainly for free traffic. For the synchronized regime the number of the overtaking cars is considerably smaller (less than 0.5\% ). Also the fact that traffic is, on contrary to the Dyson gas, a system with open boundary and with active elements that do not conserve energy, is not significant for further discussion.

II. ANALYZING TRAFFIC DATA

In [6] it has been found that the time headway distribution between Mexican buses follows the prediction of RMT. In the following it is shown that the same holds true for other transport data sets as well. However, it would be very interesting to learn about further examples of this type of distributions in order to have a stronger argumentation for the connection between RMT and traffic.

In traffic science, distance distributions of the type described by the equation Eq. (4) are well known, but were not recognized as the result of RMT or a general result describing one dimensional dynamical systems. For instance, the most general function that traffic engineers use is the so called Pearson-type-III distribution that reads (see [7] and references therein)

\[ P(g) = \left( \frac{g-a}{b} \right)^{k-1} \frac{1}{b\Gamma(k)} \exp\left( -\frac{g-a}{b} \right), \]

where \( a \) is a location parameter, \( b \) is a scale parameter (basically the mean value if \( a = 0 \)), and \( k \) is a shape parameter. Note that for \( a = 0 \) and \( k = 1/b = \beta + 1 \) the Pearson distribution is identical with the distribution Eq. (4) discussed above. For traffic in the low density state, this reduces to the Poisson distribution. To traffic with higher flow, where the interaction between the cars could not be neglected, the full distribution has been applied successfully. Note that in traffic it is better to use the time headway distribution, since most observations of the traffic flow are measured from induction loop detectors, i.e., are local measurements. Additionally, as is well known, traffic can condense into a traffic jam, where the headway distributions for the normal and jammed traffic flow may stem from different stochastic processes. Therefore, it has to be made sure that the used data belong exclusively to one state, for instance by sampling the data from a certain velocity range only.

An important fact is that it is hardly possible to distinguish the different traffic models using the headway distribution only. For this purpose it is suitable to use the number variance statistics as well since it is much more sensitive to deviations.

Shown is the analysis of a data set recorded on the German freeway A1 near Cologne. The two detectors are some 9000 m apart, between them there is a number of on and off ramps. This data set has been used already in [8] for a different purpose. The Fig. 1 shows a comparison of the measured headway distributions [9] for the so-called synchronized traffic [10,11] and for the free traffic. It could be seen from these figures that the real traffic seems to follow the RMT nicely, except for some values near the maximum of the curve. The number of variance behavior is visible in Fig. 2. For the free flow, the distribution is Poissonian, for synchronized traffic and small interval lengths the number variance follows the prediction of the nearest neighbor Dyson gas with excluded volume.
III. SIMULATION MODELS

This section analyzes a number of well known microscopic traffic flow models from two different classes, fully discrete and time discrete. The case of time continuous models has been explored also, however, not with the rigour used in the case of the discrete models. Preliminary results suggest that the model in \([12,13]\) behaves differently, since its headway distribution is Gaussian for any density, while the model in \([14]\) shows a behavior similar to what is found for the discrete models below.

As an example of a fully discrete model the Nagel Schreckenberg cellular automaton model \([15,16]\) is used. For this model, named CA model in the following, a number of exact results is known at least for the case \(v_{\text{max}} = 1\) and for the time headway distribution.

The next model to be discussed is the so called SK model \([17,18]\). The basic idea is that cars drive as fast as possible, but avoid crashes. Therefore at time \(t + 1\), they have to choose their velocity \(v(t+1) \approx v_{\text{safe}}\) that takes into account the braking distances of the following and the preceding car, respectively. This means that the velocity has to fulfill the inequality \(d(v) + v \tau \leq d(\tilde{v}) + g\), where \(d(\cdot)\) is the braking distance of the cars at velocity \(v\) of the following or velocity \(\tilde{v}\) of the preceding car, respectively. The space headway between the cars is just \(g = x - x - \lambda\), where \(\lambda\) is the length of car. The braking capabilities of the cars are the same for all cars and are parameterized by the maximum deceleration \(b\).

FIG. 1. Scaled \((\langle g \rangle = 1)\) time headway distribution of highway traffic. Full lines represent the Poissonian curve and Eq. (4) for \(\beta = 3.24\). Plus signs and bars display the spacing distribution for the free flow regime and for the synchronized regime, respectively. Integrated time headway \(\int_0^T P_\beta(s')ds'\) is shown on the insert where the plus signs are the free flow states and the squares represent the synchronized states. The curves display integrated probability density for Poissonian and semi-Poissonian distributions, Eq. (4).

FIG. 2. Number variance of highway traffic. The full line represents a Poissonian number variance, the curve represents Eq. (5) for \(\beta = 3.24\). Plus signs, squares, and circles display number variance for free flow states, synchronized states, and stop-and-go states, respectively. Note that the synchronized states are more correlated than the stop-and-go states.
The parameter $a$ below is derived, which is essentially identical to the CA model. Before comparing Eq. (4) to similar results, it is necessary to shift the probability Eq. (11) by one timestep to the left in order to compare with the RMT result, Eq. (4). The results for $c = 0.5, c = 0.25$, respectively, are shown in Fig. 3. The interaction between the cars of the CA model leads to a short ranged repulsion, but additionally to a medium ranged attraction (in the vicinity of $x = 1$) that probably causes the discrepancy in the time headway statistics close to the maximum of the curve in Fig. 3.

Additionally, it is also possible to characterize the CA model by plotting the inverse temperature $\beta$ of SRDG as a function density $c$, which gives the increase from $\beta = 1$ to $\beta = 1.9$ for the densities $c \in [1/10, 1/2]$ (see Fig. 4). The difference between Eq. (4) and Eq. (11), described by $\chi^2$ test, is plotted on the inset of the same Fig. 4. The best agreement is reached for the density equal to 1/2.

For maximum velocities greater than one, no analytic solution of the model is known. Numerical simulations lead, however, to similar results.

With the results above in mind, the case of the SK model can be discussed. The most prominent feature of this model is that it has a transition from a free flow state into a jammed state at some density $\rho_c$. This transition has some similarity with a first order phase transition, therefore the transition point depends on the initial condition (hysteresis). The simulations shown below start always in the homogeneous state $g_i = 1/\rho - 1, v_i = \min[g_i, v_{\text{max}}]$. The critical density $\rho_c$ is just the maximum density where the model remains in the homogeneous state infinitely long. By using this initial condition it is made sure that the interesting high flow state that exists only near $\rho_c$ can be explored.

More specifically, the model displays, as function of density, the following set of states:

- $\rho \ll \rho_c$, a noninteracting low density regime, which upon increasing the density transforms into $\rho = \rho_c$
- a strongly interacting high flow state, that has some similarity to an almost deterministic ordered state, and which eventually is metastable.

This safety conditions can be transformed into a set of update rules as follows:

$$v_{\text{safe}} = \bar{v}(t) + 2 b \frac{g(t) - \bar{v}(t)}{2 b + v(t) + \bar{v}(t)},$$

$$v_{\text{des}} = \min\{v(t) + a, v_{\text{safe}}, v_{\text{max}}\},$$

$$v(t+1) = \max\{v_{\text{des}} - a \xi, 0\},$$

$$x(t+1) = x(t) + v(t+1).$$

The parameter $a$ is the maximum acceleration, the parameter $b$ measures the degree of randomness, $\xi$ is a random number, $\xi \in [0,1]$, while $v_{\text{max}}$ is just the maximum velocity. The choice of $v_{\text{safe}}$ is not unique, other formulas can be used as well, but give little differences.

The work for the CA model has been done already, see [16] and references therein. For this model with $v_{\text{max}} = 1$ a complicated time headway distribution formula has been derived, which is essentially identical (not in all respects, see below) with Eq. (4)

$$P(s) = \frac{q y}{c - y} \left(1 - \frac{q y}{c}\right)^{s-1} - q^2(s-1)p^{s-2} + \frac{q y}{1-c-y} - \frac{q y}{1-c}$$

$$= \frac{q y}{c - y} \left(\frac{q y}{c - y} + \frac{q y}{1-c-y}\right) p^{s-1},$$

where the short hand notations $q = 1 - p$ and

$$y = \frac{1}{2 q} \left(1 + \sqrt{1 - 4 q c (1-c)}\right)$$

have been used. Here, $c$ is the global density and $p$ is the only parameter of the model. Before comparing Eq. (11) with Eq. (4) it is important to take into account that the CA model
a free flow regime that coexists with the jammed flow. Within a jam, the cars move very slowly.

The model should belong to the case of the Dyson gas with short ranged interaction and excluded volume, therefore, it is to be expected that it displays a headway distribution and a number variance of the semi-Poissonian type. However, the complexities encountered in the case of the CA model happen to occur in this case, also.

The Fig. 5 shows one representative time headway distribution from each of the regimes above, except for the jammed regime. It can be seen that the results nicely follow the prediction of random matrix theory for \( b' = 4 \) given by Eq. (2). For small car densities, the time headway distribution follows a semi-Poisson with \( b < 1 \) (Poison regime), where \( \beta \) depends on the density \( \rho \) with \( \beta \rightarrow 0 \) if \( \rho \rightarrow 0 \).

For larger densities, the model reaches via a RMT regime with \( \beta = 1 \) and \( \beta = 2 \) (around density 0.1) the above-mentioned high flow state. This state did not directly fit into the framework of RMT, however. E.g., the time discrete headway distribution collapses into a very narrow distribution, which can hardly fit with the continuous formulas above. Considering a continuous distribution constructed from the dynamical variable \( \tau_i = (s_i + 1)/v_i \) would need a large value of \( \beta \) if fitted with the semi-Poisson formula or with the Wigner formula, Eq. (2).

For densities even larger, the free flow state that coexists with the jammed state displays a behavior very similar to...
what has been found in the simple CA, with a general con-
formance to RMT, but with the same deviations for inter-
mediate headways. The deviations at intermediate headways are
stronger than in the case of the simple CA. In order to learn
more about these deviations, the next section reports simula-
tions of the Dyson gas with a modified interaction potential
that includes such an intermediate attraction.

IV. INTEGRATING THE DYSON GAS

This section reports the numerical integration of the
Dyson gas. The numerical scheme has to be time invariant
and energy conserving, and retain the symplectic structure of
the phase space. The approach used here is described in [21],
a so called step size controlled Verlet algorithm. The energy
conservation $\Delta E/E$ achieved is about $10^{-4}, \ldots, 10^{-5}$,
which is enough for the purpose here. This scheme uses a
reparameterization of time where the function $R(x)$ that gov-
ers the evolution of the time parameterization has been
chosen simply as $R(x) = \max[\alpha(x)]$ with a maximum step size,
where $\alpha(x)$ is the acceleration (vector), $x$ is the vector of
space coordinates and $v$ the vector of velocities. The equa-
tions used read

$$
\frac{d}{ds}x = \frac{v}{R(x)},
$$

$$
\frac{d}{ds}v = - \frac{\nabla V(x)}{R(x)},
$$

$$
\frac{d}{ds}f = \frac{1}{R(x)},
$$

which are translated into the following numerical scheme:

$$
x_{n+1} = x_n + \frac{\Delta s}{\rho_{n+1/2}} v_{n+1/2},
$$

$$
v_{n+1/2} = v_{n-1/2} - \frac{\Delta s}{2} \nabla V(x_n) \left( \frac{1}{\rho_{n-1/2}} + \frac{1}{\rho_{n+1/2}} \right),
$$

$$
\rho_{n+1/2} = 2R(x_n) - \rho_{n-1/2},
$$

$$
t_{n+1} = t_n + \frac{\Delta s}{\rho_{n+1/2}}.
$$

Here $s$ is the new time variable, while $\rho$ is a short hand
notation for the step size control function $R(x_n)$.

Running the simulation then is standard in molecular dy-
namics. In order to reach thermal equilibrium from a non-
equilibrium initial condition, so called thermalization steps
have to be applied, that rescale the velocity until equilibrium
has been reached. Such a rescaling simply multiplies the velo-
cities with $1/\sqrt{2 \beta E_{\text{kin}}}$, where $E_{\text{kin}}$ is the current kinetic
energy of the sample. All measurements have been done in
equilibrium, of course, and give really impressive agreement
between simulation and RMT, see Fig. 6 for an example
[22].

Upon adding a small attractive interaction, which acts on
an intermediate scale only, e.g., using as the interaction po-
tential

$$
V = - \ln(\Delta x) - c_1 \frac{1}{1 - c_2 (\Delta x - c_0)^2}
$$

with $c_0 = 2$, $c_1 = 1$, $c_2 = -0.5$, the system displays a behavior
that has the same kind of deviations from the RMT predic-
tion as are present in the CA and SK model, see Fig. 6.

V. CONCLUSIONS

So far, it has been shown that traffic systems behaves
similar to a large class of one dimensional dynamical sys-
tems, whose representative is provided by the Dyson gas.
There is a deep connection between quantum systems that
displays level repulsion, and follows RMT, and their classical counterpart displaying chaos. The Dyson gas is known to display Hamiltonian chaos. This is why it can be speculated that traffic is a chaotic system, too. It has to be stressed that properties of the Dyson gas are described with the help of equilibrium statistical physics that serves as a link between this dynamical system and the theory of random matrices. On the other hand, the traffic is used to be investigated as a dynamical system and the theory of random matrices. Nevertheless, this work shows, that the time headway distributions observed in traffic flow can be understood as the consequence of a very general underlying theory (Dyson gas and RMT), which is not very sensitive to the details of the interaction between the particles. For the Dyson gas, it can be shown that a number of different interaction types leads to the same headway distribution. Of course, not any interaction leads to the semi-Poisson distribution. Since the statistical properties of traffic seem to be in line with RMT, the models aimed at describing traffic have to checked for the deviations between them and reality in the light of the results presented here.

ACKNOWLEDGMENTS

This work has been supported by the Foundation for Theoretical Physics in Slemeno, Czech Republic. It would not have been possible without the provision of data by the Landschaftverband Rheinland of the German Federal State Northrhine Westfalia, which we gratefully acknowledge. Discussions with Claudia Hertfelder and Georg Hertkorn from the German Aerospace Center have helped to clarify the issues presented here.

[9] Note, that there are differences between Fig. 1 and the corresponding figure in [8]. In this paper the raw data were used, i.e., instead of computing the time headway from the recorded distance to the car ahead, the time difference to the car ahead has been used. Furthermore, only those data have been used for the analysis, where the velocity difference to the car ahead has been smaller than 30 km/h and whose velocity stems from a window [60,80] km/h to indicate synchronized traffic flow.
[19] Indeed, the simulations with the SK-model reported in the following use the exact formula (if $b=\text{const}$). $v_{\text{add}} = -b + \sqrt{b^2 + \bar{v}^2(t) + 2bg(t)}$ simply because it is logically more appealing: it does not need the velocity of the car to be updated.
[20] If analyzing data from the jammed region, i.e., data where $v \leq 2a$ and $\bar{v} \leq 2a$ holds simultaneously, a bimodal distribution is encountered. One peak can be identified with the RMT, the second one is basically Gaussian. It shows that the criterion above was not able to separate completely jammed cars from nonjammed ones, jams have big holes inside, where cars can gain some speed. The Gaussian distribution can be understood as a consequence of how cars are added to the jam, they do not stop directly behind the last jammed car, but at a certain stochastic distance.
[22] Of course, the same result can be achieved by using just the right initial conditions, i.e., the correct initial kinetic energy. The system converges then rapidly toward the gap distribution given by RMT and a Maxwellian velocity distribution.