PRAMANA — journal of physics Vol. 64, No. 5 May 2005 pp. 635–643

Boltzmann and Einstein: Statistics and dynamics – An unsolved problem

E G D COHEN
The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA
E-mail: egdc@rockefeller.edu

Abstract. The struggle of Boltzmann with the proper description of the behavior of classical macroscopic bodies in equilibrium in terms of the properties of the particles out of which they consist will be sketched. He used both a dynamical and a statistical method. However, Einstein strongly disagreed with Boltzmann's statistical method, arguing that a statistical description of a system should be based on the dynamics of the system. This opened the way, especially for complex systems, for other than Boltzmann statistics. The first non-Boltzmann statistics, not based on dynamics though, was proposed by Tsallis. A generalization of Tsallis' statistics as a special case of a new class of superstatistics, based on Einstein's criticism of Boltzmann, is discussed. It seems that perhaps a combination of dynamics and statistics is necessary to describe systems with complicated dynamics.

Keywords. Boltzmann; Einstein; dynamics; statistics.

PACS Nos 02.50.-r; 05.20.-y; 47.27.-c

1. Boltzmann's struggle

In his second paper in 1866, at the age of 22, Boltzmann wrote an article *On the mechanical meaning of the second law of the theory of heat* [1] in which he claimed to prove the second law of thermodynamics purely from dynamics, i.e., the second law was just a consequence of Newton's equations of motion.

At that time he was not even aware of the necessity of introducing probability concepts in order to prove or at least understand the second law. However, two years later in 1868 in a paper *Studies on the equilibrium of the kinetic energy between moving material points* [2] he follows Maxwell in introducing probability concepts into his mechanical considerations and discusses a generalization of Maxwell's equilibrium velocity distribution function for point particles in free space to the very general case that *a number of material points move under the influence of forces for which a potential function exists*, i.e., he derives the Maxwell–Boltzmann equilibrium distribution function. One would think that Boltzmann had now realized the relevance of probabilistic concepts to understand the second law for a system in equilibrium which follows from the Maxwell–Boltzmann distribution.

Yet in 1872 when he derived in his paper: *Further studies on thermal equilibrium between gas molecules* [3], what we now call the Boltzmann equation, he used, following Clausius and Maxwell, the so-called Stoszzahl Ansatz ('assumption of molecular chaos'), and he does not seem to have realized the statistical, i.e., probabilistic nature of this assumption, i.e., of the assumption of the independence of the velocities of two molecules which are going to collide. The same probabilistic assumption affected the H-theorem through which he proved the approach to equilibrium. He says, however: "One has therefore rigorously proved that, whatever the distribution of the kinetic energy at the initial time might have been, it will, after a very long time, always necessarily approach that found by Maxwell" [4]. In other words, like in 1866, he thought he had derived the approach to equilibrium and the second law purely from dynamics. Only later, forced in particular by Loschmidt's reversibility paradox around 1876 [5] and then still later by Zermelo's recurrence paradox around 1896 [6], based on Poincaré's recurrence theorem, did Boltzmann clearly state the probabilistic nature of the Stoszzahl Ansatz and the H-theorem in that they only hold for disordered states of the gas, which, to be sure, exceeded vastly the number of ordered ones. In other words, the Stoszzahl Ansatz might be a very good approximation for a dilute gas – which Boltzmann always considered – but nevertheless contained an essential probabilistic element, which prevented a purely mechanical derivation of the approach to equilibrium and the second law.

The polemics surrounding the H-theorem and its proper interpretation depressed Boltzmann enormously and in 1877 he proposed a completely new approach, in a paper: *On the relation between the second law of thermodynamics and probability theory with respect to the laws of thermal equilibrium* [7]. This approach contained no mechanical component at all and led to a relation between the thermodynamic entropy S and the probability W for the dynamical states of the system at a given total energy in phase space. This was later codified by Planck in 1906 [8] by the formula: $S = k \log W + \text{constant}$, where Boltzmann's constant k was first introduced. It seemed to connect the second law with probability alone, the opposite of what Boltzmann originally thought. While Boltzmann struggled with the proper roles played by dynamics and statistics, which we will encounter again later in this lecture, Gibbs in his book of 1902, *Elementary principles in statistical mechanics* [9], solved the problem 'by cutting the Gordian knot' using a 'hypothesis of equal *a priori* probability for equal regions in the phase space of the system' [10], thus eliminating all dynamics and postulating the canonical ensemble by a simple probabilistic assumption. This avoided all complications associated with the approach to equilibrium via the dynamics of the collisions, but restricted his work to equilibrium.

As an aside I might remark that this was anathema to Ehrenfest and no doubt also to Boltzmann, since for them the crucial question of statistical mechanics was not the equilibrium state, but the approach to equilibrium, which clearly had to involve the dynamics of the system. As I will explain, Einstein went a step further and argued that even in equilibrium, dynamics was essential.

2. Einstein

From 1904, the formula $S = k \log W$ was repeatedly criticized by Einstein [11]. Einstein argued that one cannot possibly obtain *a priori* the probabilities to find the dynamical system in a certain region in its phase space since these probabilities, W, must follow *a posteriori* from the dynamics of the system. In other words, the frequency that a phase point visits a region in phase space is determined by its dynamics and cannot be given *a priori*, without using the equations of motion [11,12].

Boltzmann's statistical approach $S = k \log W$ only applied to equilibrium and was in fact derived by him only for an ideal gas in μ-space, the phase space of a single molecule. Einstein never used Boltzmann's formula $S = k \log W$, but only its inverse, i.e., $W = \mathrm{e}^{S/k}$. This allowed him to determine W from the entropy S, which, like all physical quantities, is determined by the dynamics of the system [13]. Gibbs later generalized Boltzmann's formula in μ-space to $S = -k \int f(\Gamma) \log f(\Gamma) \mathrm{d}\Gamma$ [14], i.e. to Γ-space, the phase space of the entire system but, to the best of my knowledge, this formula is equivalent to Boltzmann's in equilibrium and its generalization outside equilibrium remains, as far as I know (like Boltzmann's), an unsolved problem.

Recently Gallavotti and I [15] argued that the entropy *content* S of a system can only be defined in an equilibrium state, but not in a non-equilibrium stationary state, similar but different from the impossibility to define the heat content of a system in equilibrium [16]. What can be defined, at least in non-equilibrium stationary states are, it seems, entropy changes, like the entropy production rate or the entropy transfer rate but not the entropy content.

Boltzmann's statistical approach was criticized repeatedly by Einstein [17], from 1904 to 1920, since no one listened to his critique anyway. Nevertheless, Einstein argued, in my opinion, completely correctly, that one cannot in principle fix the weights of the various regions in phase space without using the dynamics, i.e., the equations of motion of the system. I quote Einstein (1910) in an abbreviated translation [17]:

"Usually W is set equal to the number of ways (complexions) in which a state, which is incompletely defined in the sense of a molecular theory (i.e. coarse grained), can be realized. To compute W one needs a complete theory (something like a complete molecular-mechanical theory) of the system. For that reason it appears doubtful whether Boltzmann's principle alone, i.e. without a complete molecular-mechanical theory (elementary theory) has any real meaning. The equation $S = k \log W + \text{const.}$ appears (therefore), without an elementary theory – or however one wants to say it – without content from a phenomenological point of view."

It should be remarked that earlier (in 1868) [18] Boltzmann also mentioned Einstein's idea of basing the probabilities W of a system in equilibrium on the dynamics of the system, by postulating that the probability to find the system in a certain region in phase space would be proportional to the time spent by the system in that region, which clearly depends on the equations of motion. However, he never distinguished this from his statistical idea, based on $S = k \log W$ of 1877, as Einstein did. As we will see, Einstein's remark opened the way to other classical statistics than that of Boltzmann and Gibbs.

3. Tsallis

The first classical non-Boltzmann statistics of physical systems was proposed by Tsallis [19]. It contained an unknown real parameter q – a 'bias parameter' – which for $q = 1$ (no bias) gave Boltzmann's statistics. Tsallis derived his statistics based on a probabilistic entropy, similar but different from the Rényi entropy [19]. Then, Tsallis maximized this entropy under similar conditions as Boltzmann had done for his statistics (normalized probabilities for the various energy states of the system in phase space and a fixed total energy of the system), thus making it a physical rather than a mathematical entropy. He then obtained a new canonical-like distribution function, which for $q \neq 1$ established a power law decay of the system's energy states for large energies, while for $q = 1$, the distribution reduced to that of Maxwell–Boltzmann, with an exponential decay for large energies. Tsallis' distribution showed unusual properties, like non-additivity and non-extensivity of the entropy and would in general apply to systems with a more complicated dynamics than those that follow Boltzmann's statistics, which include systems consisting of particles with short-range forces. In particular, it would be applicable to non-equilibrium rather than equilibrium states, where non-extensivity is omnipresent. A great variety of systems appear to be well-described numerically by Tsallis' statistics [20,21]. However, a prediction or a physical understanding of this remarkable agreement is still lacking.

A different, more physical, derivation of Tsallis' statistics was given by Beck [22] and can be 'based' on what we later proposed together as a 'superstatistics'. This is a generalization of Tsallis' statistics, in the spirit of Einstein and has been applied in particular to systems in non-equilibrium stationary states, like, e.g., turbulent systems.

4. Superstatistics [23]

We consider a driven non-equilibrium system, here for simplicity a Brownian particle with a time-dependent friction in a fluid, which is thought to be partitioned in physically infinitesimal regions (cells or partitions), which exhibit spatio-temporal fluctuations of an intensive quantity. This could be the temperature or the chemical potential in a cell or, in the case of turbulence, the energy dissipation rate in a Kolmogorov cascade.

We will discuss mainly the case of fluctuations of the inverse temperature β. We assume for simplicity that the spatio-temporal fluctuations of β in the fluid can be described by a probability distribution function $f(\beta)$ and that local equilibrium in the cells is established on a much shorter time-scale τ (determined by a characteristic value of the friction in the Langevin equation) than the long time-scale T, characteristic of the fluctuations of β. Then, due to the motion of the Brownian particle through the cells each with a possible different β, a stationary probability distribution will arise for long times ($t \gg T$) as a superposition of local (cell) equilibrium Maxwell–Boltzmann distribution functions $\mathrm{e}^{-\beta E}$, weighted according to the distribution function $f(\beta)$. If E is the energy of a state associated with each cell, there will be an effective Boltzmann factor for the non-equilibrium system of the form:

$$B(E) = \int_0^\infty \mathrm{d}\beta f(\beta) \,.\, \mathrm{e}^{-\beta E}.$$

Here $B(E)$ can be considered as a superstatistics in the sense that it is 'a (global) statistics – induced by $f(\beta)$ – of a (local) statistics – given by the Boltzmann factor $e^{-\beta E}$' – i.e. 'a statistics of a statistics'.

It appears that the generalized superstatistical Boltzmann factors $B(E)$ are physically relevant for large classes of dynamically complex (such as turbulent) systems, where they can describe a variety of observed phenomena (e.g. correlation functions) in a rather convincing way.

Coming back to Einstein, although the dynamics of the system should determine the statistics of the system in phase space i.e. $f(\beta)$ here, in practice, this is an impossible task for many particle systems. Therefore, one has tried to describe the dynamics of the system by guessing a proper statistics, i.e. a normalized $f(\beta)$, which would then incorporate the relevant aspects of the system's full dynamics for the properties (e.g. the correlation functions) in which one is interested and which can be measured. Therefore, one hopes that one can convert the dynamics of the system into a statistics, which will describe, for instance, physical quantities of the system, which are of interest.

Or, vice versa, how do probability functions, of the type I will discuss later, represent certain physical features as a result of the dynamics of the system? Clearly this is not precisely what Einstein had in mind, but for systems consisting of many particles there seems to be no other way to proceed at present, if one follows Einstein's idea. Therefore, it would be extremely useful in this connection to have a physical theory of how dynamics can be converted into probability functions. For the case of the χ^2- (or Γ-) distribution [23] – a superposition of squared independent Gaussian random variables, all with mean zero – one could somehow physically imagine that such a superposition would represent in some crude sense some aspects of a turbulent fluid motion. For a number of correlation functions measured in cylindrical Couette flows, this has indeed been successfully carried out by Beck and others, by choosing appropriate distribution functions [24,25].

An interesting consequence of the superstatistical Boltzmann statistics $B(E)$ appears when one expands it to second order in powers of E. For, then all mathematical representations for $f(\beta)$ used to represent the dynamics of all the systems with complicated dynamics studied so far – such as the χ^2- [23] or Γ-distribution, the log-normal distribution [23] or the F distribution [23] – the small E behavior can be proved to be universal. That is, all can be written in the form $B(E) = e^{-\beta_0 E}[1 + \sigma^2 E^2/2 + O(\beta_0^3 E^3)]$. Here $\beta_0 = \langle\beta\rangle$ is the mean and $\sigma^2 = \langle\beta^2\rangle - \langle\beta^2\rangle$ is the variance of (β). In fact, Beck showed that the Tsallis and the χ^2 statistics for $f(\beta)$ are identical in general, i.e. to all orders of E.

Higher order terms in E than E^2, allow then, in principle, to distinguish between the various mathematical functions needed for a statistical representation of the relevant features of the dynamics of different systems. As a by-product of this universal behavior of the superstatistical distribution functions, one can obtain from them a simple universal expression for the analogue of Tsallis' bias parameter q from the universal coefficient of the term $\sim E^2$, of the form: $q = \langle\beta^2\rangle/\langle\beta\rangle^2$, i.e. the ratio of the average of β^2 to the average of β squared.

5. Experiment

1. To make these considerations concrete I will give one experimental example: the advective motion of a Brownian particle (a polystyrene sphere of 46 μm in diameter) swept around a highly turbulent circular Couette flow (with Reynolds number Re $\approx$ 63,000) between two oppositely moving concentric cylinders (figure 1). This experiment was carried out by Bodenschatz *et al* [24]. The χ^2 or perhaps even better the log-normal based superstatistics seem to be able to describe the fluctuations of a scaled acceleration component a of the Brownian particle as a function of a very well (figure 2) [21].
2. More details on superstatistical turbulence models can be found in [25]. Some other superstatistics noted in the literature are: Swinney *et al* [26, 27] for the longitudinal velocity difference fluctuations in a highly turbulent Couette flow; Castaing *et al* for the turbulent flow in a jet or in a wind tunnel [28]; Sattin and Salasnick for the electron density fluctuations in fusion plasma physics [29] and Boghosian for the drift turbulence in a pure magnetized electron plasma in a cylindrical tube [19, 30].

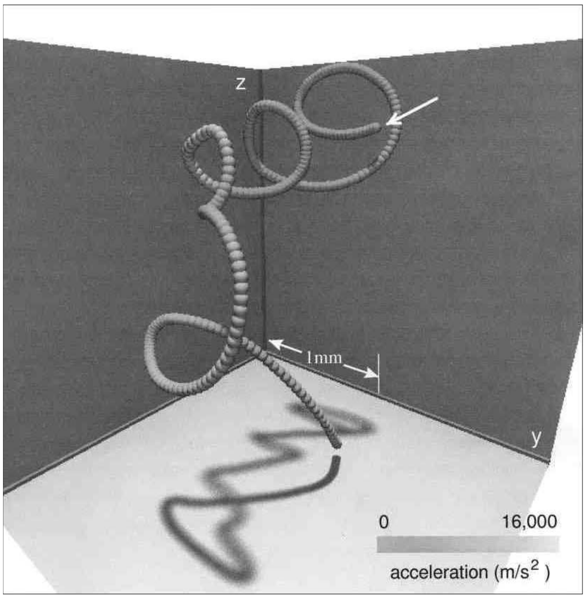

Figure 1. Trajectory of a 46 μm diameter polystyrene tracer particle at a Reynolds number of Re $\approx$ 63.000 in a turbulent fluid between two oppositely rotating concentric cylinders. The little spheres in the figure represent positions of the particle along its trajectory. The arrow indicates where the particle was injected in the fluid. The color of the trajectory indicates the magnitude of the acceleration (indicated by the scale), which can reach $\approx$ 1600 g in one of the bends [24] (Courtesy of E Bodenschatz).

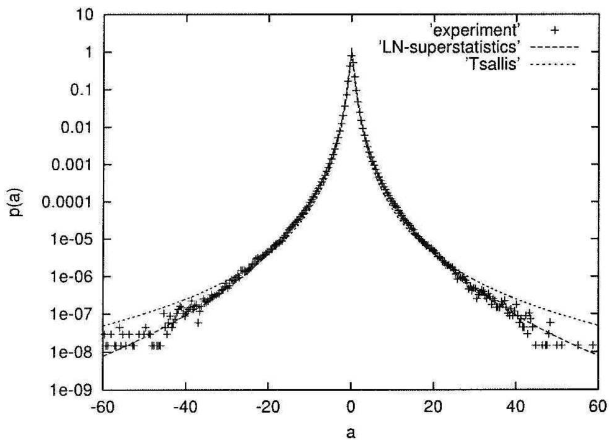

Figure 2. The probability density $p(a)$ of a scaled transverse component a of the acceleration of the tracer particle of figure 1, as a function of a. The large dashed lines (lower curve) give the theoretical results of using a log normal superstatistics for the dynamics, while the small dashed lines (upper curve) give the results of using a Tsallis distribution for the dynamics of the particle. The + signs represent the experimental data for three Reynolds numbers (Re = 63,000, 31,700, 26,700), all with variances equal to one [21] (Courtesy of C Beck).

3. Finally, I want to comment on a difference between this superstatistical approach to turbulence and the conventional one based on the Navier–Stokes equations. In the latter case one approaches the problem of turbulent motion 'from the outside in', by studying the Fourier components of a Fourier analysis of the Navier–Stokes equations in space–time of the entire fluid. In the former case one approaches the problem of turbulent motion 'from the inside out', by introducing probability functions, which apparently are able to represent certain aspects of the internal dynamics of the turbulent fluid, based on the dynamical idea of Einstein, but mixed with a good grain of statistics to make it work in practice. A connection between the two approaches is an open question.

6. Conclusion

Let me end this very brief survey of some ideas and results on the description of complex macroscopic systems, obtained dynamically or statistically or by both, in a historical perspective, by a quote from Boltzmann. This quote is taken from Boltzmann's lecture *On recent developments of the methods of theoretical physics*, given in 1899 in München, Germany [31]. Here, when he asks about the future of the physics of his days, he says:

"Will the old (classical) mechanics with the old forces ... in its essence remain, or live on, one day, only in history ... superseded by entirely different notions?

Will the essence of the present molecular theory, in spite of all amplifications and modifications, yet remain, or will one day an atomism totally different from the present prevail. . . . (Or, I could add, will entropy and statistics continue to play their present dominant role or will dynamics become more essential for the description of the properties of macroscopic systems?)" Boltzmann concludes his lecture by: "Indeed interesting questions! One almost regrets to have to die long before they are settled.

O! immodest mortal! Your fate is the joy of watching the ever-shifting battle!" (not to see its outcome).

Acknowledgements

The author would like to express my deep appreciation to the IUPAP Commission on Statistical Physics for awarding me the Boltzmann medal 2004. The author is also indebted for financial assistance to the Organizers of STATPHYS 22, T V Ramakrishnan, Sushanta Dattagupta and Rahul Pandit, as well as to the Office of Basic Energy Sciences of the US Department of Energy under contract DE-FG02-88-ER13847. Valuable suggestions of C Beck are also gratefully acknowledged.

References

[1] L Boltzmann, Über die mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie, *Wien. Ber.* **53**, 195–220 (1866)
F Hasenöhrl (ed.), *Wissenschaftliche Abhandlungen (WA)*, (Chelsea Publ. Co., NY, 1968), Band I, pp. 9–33
[2] L Boltzmann, Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten, *Wien. Ber.* **58**, 517–560 (1868); *WA*, Band I, pp. 49–96; ID., Löung eines mechanisches Problems, *Wien. Ber.* **58**, 1035–1044 (1868); *WA*, Band I, p. 97
[3] L Boltzmann, Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, *Wien. Ber.* **66**, 275–370 (1872); *WA*, Band I, pp. 316–402
[4] See ref. [3]; *WA*, Band I, p. 345
[5] J Loschmidt, *Wien. Ber.* **73**, 128, 366 (1876); ID., **75**, 287; ID., **76**, 209 (1877)
L Boltzmann, Bemerkungen über einige Probleme der mechanischen Wärmetheorie, *Wien. Ber.* **74**, 62–100 (1877) (section II)
[6] (a) L Boltzmann, Entgegnung auf die Wärmetheoretischen Betrachtungun des Hrn. Zermelo, *Wied. Ann.* **57**, 778–784 (1896); *WA*, Band III, pp. 567–578
(b) ID., Zu Hrn. Zermelos Abhandlung über die mechanische Erklärung irreversibler Vorgänge, *Wied. Ann.* **60**, 392–398 (1897); *WA*, Band III, pp. 579–586
(c) ID., Über einen mechanischen Satz Poincaré's, *Wien. Ber.* **106**, 12–20 (1897); *WA*, Band III, pp. 587–595
[7] L Boltzmann, Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über des Wärmegleichgewicht, *Wien. Ber.* **76**, 373–435 (1877); *WA*, Band II, pp. 164–223
[8] M Planck, *The theory of heat radiation* (Dover, 1951) p. 119
[9] J W Gibbs, *Elementary principles in statistical mechanics* (Yale University Press, 1902); also (Dover Publications, NY, 1960)

[10] R C Tolman, *The principles of statistical mechanics* (Oxford, 1938) p. 59
[11] A Pais, *Subtle is the Lord...* (Oxford, 1982) Chapter II
[12] E G D Cohen, *Physica* **A305**, 19 (2002)
[13] Ref. [11], p. 74
[14] (a) Cf. ref. [9], p. 45
(b) P Ehrenfest and T Ehrenfest, *Begriffliche Grundlagen der statistischen Auffassung in der Mechanik* in: *Enzykl. d. Mathem. Wiss.*, IV, 2, II, Heft 6, p. 59; or in P Ehrenfest, *Collected papers*, (North-Holland, Amsterdam, 1959) p. 269
(c) English translation by MJ Moravcsik: P Ehrenfest and T Ehrenfest, *The conceptual foundations of the statistical approach in mechanics* (Cornell University Press, Ithaca, NY, 1959) p. 51
[15] G Gallavotti and E G D Cohen, *Phys. Rev.* **E69**, 035104 (2004)
[16] L D Landau and E M Lifshitz, *Statistical physics* (Pergamon, London, 1958) p. 43
[17] Ref. [11], p. 72
A Einstein, *Ann. d. Phys.* **33**, 1275 (1910)
[18] L Boltzmann, Studien über des Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten, *Wien. Ber.* **58**, 517–560 (1868)
WA, F Hasenöhrl (ed.) (Chelsea Publ. Co. NY, 1968) Band I, p. 49
[19] C Tsallis, *Stat. Phys.* **52**, 479 (1988); ID., Nonextensive statistical mechanics and thermodynamics in: Lecture notes in physics edited by S Abe and Y Okamoto (Springer, Berlin, 2001) pp. 3–98
[20] I Bediaga, E M F Curado and de J M Miranda, *Physica* **A286**, 156 (2000)
C Tsallis, J C Anjos and E P Borges, *Phys. Lett.* **A310**, 372 (2003)
W M Alberico, A Lavagno and P Quarati, *Nucl. Phys.* **A680**, 94 (2000)
S Abe and N Suzuki, *J. Geophys. Res.* **108**, 2113 (2003)
[21] C Beck, *Cont. Mech. Thermodyn.* **16**, 293 (2003); see also ref. [24(a)]
[22] C Beck, *Phys. Rev. Lett.* **87**, 180601 (2001)
[23] C Beck and E G D Cohen, *Physica* **A322**, 267 (2003)
[24] (a) A La Porta, G A Voth, A M Crawford, J Alexander and E Bodenschatz, *Nature (London)* **409**, 1017 (2001)
(b) N Mordant, A M Crawford and E Bodenschatz, *J. Fluid Mech.* **469**, 121 (2002)
[25] (a) C Beck, *Europhys. Lett.* **64**, 151 (2003)
(b) A Reynolds, *Phys. Rev. Lett.* **91**, 804503 (2003)
[26] G S Lewis and H L Swinney, *Phys. Rev.* **E59**, 5457 (1999)
[27] C Beck, G S Lewis and H L Swinney, *Phys. Rev.* **E63**, 035303(R) (2001)
[28] B Castaing, Y Gagne and E J Hopfinger, *Physica* **D46**, 177 (1990)
[29] F Sattin and L Salisnick, *Phys. Rev.* **E65**, 035106(R) (2002)
[30] B M Boghosian, *Phys. Rev.* **E53**, 4754 (1996)
[31] L Boltzmann, Über die Entwicklung der Methoden der theoretischen Physik in neuerer Zeit in: *Populäre Schriften*, second ed. (Barth, Leipzig, 1919) pp. 226–227