Revista Brasileira de Ensino de Física, vol. 46, e20230322 (2024)
www.scielo.br/rbef
DOI: https://doi.org/10.1590/1806-9126-RBEF-2023-0322

Kinetic theory of gases: stochastic thermodynamics avant la lettre

Mário J. de Oliveira*[1]

[1]Universidade de São Paulo, Instituto de Física, São Paulo, SP, Brasil.

Received on October 25, 2023. Accepted on December 10, 2023.

We show that the kinetic theory of gases developed by Clausius, Maxwell, and Boltzmann can be understood as stochastic thermodynamics. This comprehension is based on the recognition that the dynamics used by Maxwell to demonstrate his distribution of velocities and by Boltzmann to derive his fundamental equation is a stochastic dynamics. An essential feature of the transition rates devised by Maxwell and used by Boltzmann is the preservation of energy and momentum. Since the dynamics is stochastic and not deterministic, we may say retrospectively that the objections to the kinetic theory raised by the paradox of irreversibility is resolved or is immaterial. Using these transition rates we obtain the Kolmogorov equation associated with the stochastic dynamics We show that it reduces to the Boltzmann equation if the velocities of distinct molecules can be considered to be statistically independent. We also show how the Boltzmann H-theorem is related to the production of entropy.

Keyword: kinetic theory of gases, Boltzmann equation, stochastic thermodynamics.

1. Introduction

The kinetic theory of gases [1–8] was advanced by Clausius in a paper of 1857 [7, 9] and developed by Maxwell and Boltzmann in the third quarter of the nineteenth century. They aimed to derive the macroscopic mechanical and thermal properties of gases from the microscopic structure of gases. This structure was assumed to be a collection of molecules that move inside a vessel colliding elastically with the walls of the vessel and with each other. Maxwell introduced the distribution of velocities that bears his name in a paper of 1860 [10, 11]. In this paper he gave an incipient proof of the distribution which was replaced in a subsequent paper of 1867 [12] by a notable demonstration. Boltzmann in a paper of 1872 [8, 13] advanced the fundamental equation of kinetic theory, the Boltzmann equation [14–20], and the concept of irreversibility through his theorem of irreversibility, known as the H-theorem.

The demonstrations and derivations carried out by Maxwell and Boltzmann employed reasonings that were not taken exclusively from mechanics as we might think at first sight. They employed the laws of mechanics intertwined with probabilistic reasonings sometimes in an implicit form. The concepts of probability were not used in the static sense only but also in the dynamic sense, in the sense that we call stochastic. That is, they implicitly considered that the system follows a stochastic dynamics [21, 22], a concept that is the basis of the theories of Markov and of Langevin, that appeared later, in the first decade of the twentieth century. The time evolution of the probability associated with the Markov theory is governed by the Kolmogorov equation or the master equation, and that associated with the Langevin equation is governed by the Fokker-Planck equation, a significant example of which is the equation used by Einstein to explain the Brownian motion [21, 22]

If the derivations carried out by Boltzmann came exclusively from the deterministic Newtonian equations of motion applied to a collection of interacting particles, one would conclude that irreversibility contained in the Boltzmann equation would arise from the deterministic motion. This cannot be true, and in fact, an objection to such a conclusion was raised a few years after the introduction of the Boltzmann equation, which became known as the paradox of irreversibility [18].

Strictly speaking, Boltzmann and Maxwell did not derive their results from the deterministic Newton equations of motion. They derived their results from a stochastic dynamics supplemented by the condition that the laws of conservation of energy and momentum are complied. An important point here is that the observation of these two laws by a system of interacting particles does not mean necessarily that the particles follow the Newton equations of motion. For instance, the equations of conservation of energy and of momentum of two particles are not in general sufficient to conclude that they obey the Newton equations of motion. An equivalent way of saying the same thing is to consider the collision of two hard spheres that follow the Newton equations of motion. It is well known that the laws of conservation of kinetic energy and momentum alone are not sufficient to determine the final velocities, except in one dimension.

---

*Correspondence email address: oliveira@if.usp.br

Understanding that the dynamics used by Boltzmann in deriving his equation is stochastic, and not deterministic, we may say, retrospectively, that the paradox of irreversibility is resolved, or rather, that it becomes immaterial. The stochastic aspect of kinetic theory that we advocate here is generally overlooked by authors analyzing the early development of kinetic theory. They tend to emphasize the deterministic Newtonian motion, supposedly adopted by Maxwell and Boltzmann, as well as the paradox of irreversibility to the detriment of the stochastic aspect.

The texts of Maxwell and Boltzmann sometimes appear to be exempt of probabilistic reasonings. This happens because in these cases the texts were written in a language in which the term probability is not explicit mentioned. For instance, a typical proposition of Maxwell [10, 11] is: “let the number of particles for which x lies between x and $x + dx$, be $Nf(x)dx$”. This should be understood as: “let $f(x)dx$ be the *probability* of the position of a particle liesbetween x and $x + dx$”. This peculiar language which is still used today does not mean that the authors did not wish to base kinetic theory on probability grounds. On the contrary. They employed this type of discourse simply because it was the probabilistic language of that time.

The key to recognize that the content of the propositions of kinetic theory is probabilistic lies in the distinction we must make between *probability* and *frequency* of an event [23]. The first term refers to an abstract concept which can be understood as a nonnegative number associated with a space of events. Its main property is that the sum of the probabilities of the elementary events that make up the whole space of events equals the unity. The space of events are usually identified as a set of real numbers in which case a variable that take any value on this set is called a random variable. When this variable depend on time, one uses the term stochastic variable. The second term, frequency, refers to the real counterpart of probability, or its interpretation, or yet how it is measured. Some probabilistic propositions, such as that of Maxwell mentioned above, were written using words akin to frequency of an event.

The probabilistic character of kinetic theory was pointed out by Jeans in his book of kinetic theory published in 1904 who considered it to be a consequence of *molecular chaos* [5]. The stochastic assumption implicit in the Boltzmann equation was called the collision number hypothesis (*Stosszahlansatz*) by Paul and Tatiana Ehrenfest [24, 25]. In a paper of 1907 [26], they expressed their views on the irreversibility of the Boltzmann equation, contained in the H-theorem, using an analogy with the urn model that they introduced and which is clearly an example of a stochastic dynamics. It is worth pointing out yet that the Gibbs statistical mechanics, which can be understood as kinetic theory of system in thermodynamic equilibrium, has a probabilistic character which is explicitly mentioned in the title of his book on the subject published in 1902 [27].

In the following we analyze the Maxwell derivation of the distribution of velocities that bears his name and the Boltzmann derivation of his equation as well as his theorem of irreversibility. After that, we show that the probabilistic and stochastic reasoning they used can be employed to derive a Kolmogorov equation associated with a stochastic dynamics [21, 22]. These results allow us to say that the kinetic theory can be understood as stochastic thermodynamics [28–32] avant la lettre.

2. Maxwell Distribution

The probabilistic approach to kinetic theory was explicitly stated by Maxwell in his comment concerning the collision of two molecules [33]. When dealing with an enormous number of collisions occurring in a small interval of time the problem becomes insurmountable and, says Maxwell [33]: “We are therefore obliged to abandon the strictly kinetic method and to adopt the statistical method”. The statistic method, which is founded on probability theory, was used in several sciences including physical sciences, one application of which was the description of errors. But it was Maxwell who used it as a description of the physical processes themselves [23, 34]. In this sense, Maxwell is the one who introduced the probabilistic approach in theoretical physics [23, 34].

Maxwell [10, 11] introduced his distribution of velocities by stating that the number of particle whose Cartesian coordinates lies between v_1 and $v_1 + dv_1$, v_2 and $v_2 + dv_2$, and v_3 and $v_3 + dv_3$ is $Nf(\mathbf{v})dv_1dv_2dv_3$ where

$$f(\mathbf{v}) = ce^{-(v_1^2+v_2^2+v_3^2)/\alpha^2}, \tag{1}$$

which is understood as the *probability* density distribution of the velocity $\mathbf{v} = (v_1, v_2, v_3)$, and $c = 1/\alpha^3\sqrt{\pi^3}$.

The probability distribution $h(v)$ of the absolute value of the velocity is related to $f(\mathbf{v})$ by $h(v)dv = f(\mathbf{v})dv_1dv_2dv_3$ from which follows

$$h(v) = 4\pi cv^2e^{-v^2/\alpha^2}. \tag{2}$$

To determine the probability distribution $g(\varepsilon)$ of the kinetic energy $\varepsilon = mv^2/2$ we use the relation $g(\varepsilon)d\varepsilon = h(v)dv$ from which we obtain

$$g(\varepsilon) = c'\sqrt{\varepsilon}\, e^{-\beta\varepsilon}, \tag{3}$$

$\beta = 2/m\alpha^2$ and $c' = 2\beta^{3/2}/\sqrt{\pi}$.

Maxwell made two attempts to derive his distribution of velocities. The first attempt, contained in his paper of 1860 [10, 11], is based on the assumption of isotropy which means that the probability distribution $f(\mathbf{v})$ cannot depend on the choice of directions of the Cartesian axes. Thus $f(\mathbf{v})$ should be a function of $v = |\mathbf{v}|$, which is invariant by any change of the Cartesian axes. From the isotropy, Maxwell concludes that the independent Cartesian components are also statistically independent leading him to the result that $f(\mathbf{v})$ should be a product of the probabilities of each one of the Cartesian components. Using the factorization of $f(\mathbf{v})$, he was able to derive the distribution (1).

Although the assumption of isotropy leads us to conclude that the velocity distribution is a function of the absolute value of the velocity, it does not necessarily implies that the independent Cartesian components of the velocity are statistically independent. Therefore one cannot necessarily conclude from isotropy that $f(\mathbf{v})$ factorizes into the distributions of the Cartesian components, and this demonstration of Maxwell cannot be considered to be a valid one [35]. If this reasoning were correct then any isotropic system of particles, whatever may be its dynamics, would obey the Maxwell distribution [35]. Despite its incorrectness this demonstration is still repeated in some text books of physics dealing with kinetic theory.

In the second attempt, contained in his paper of 1867 [12], Maxwell used probabilistic concepts that we recognize as the fundamental concepts of a stochastic process, and which we will examine in the following. One of the stochastic concepts used by Maxwell is the concept of transition rate probability which is implicit in the following proposition concerning the encounters of two molecules [12]:

> The number of such encounters in unit time will be $n_1 n_2 F de$ where n_1 and n_2 are the number of molecules of each kind under consideration, and F is a function of the relative velocity and of the angle θ, and de depends on the limits of variation within which we class encounters of the same kind.

According to this statement, F is a function of $|\mathbf{v}-\mathbf{u}| = |\mathbf{v}'-\mathbf{u}'|$ and of the angle θ which is half of the angle between the relative velocities $\mathbf{v}-\mathbf{u}$ and $\mathbf{v}'-\mathbf{u}'$. These two dependences is equivalent to saying that F is a function of $\mathbf{v}-\mathbf{u}$ and $\mathbf{v}'-\mathbf{u}'$. Maxwell then writes $n_1 = f_1(\mathbf{v},t)dV$ and $n_2 = f_2(\mathbf{u},t)dV$ from which follows that the number of encounters is proportional to

$$f_1(\mathbf{v},t) f_2(\mathbf{u},t) F(\mathbf{v}-\mathbf{u}, \mathbf{v}'-\mathbf{u}')\tau, \tag{4}$$

where τ is a small interval of time. This expression is interpreted as the probability density that at a given instant of time t the velocities of the molecules are $(\mathbf{v},\mathbf{u})$ *and* that at an earlier time $t-\tau$ they are $(\mathbf{v}',\mathbf{u}')$. The product $f_1(\mathbf{v},t)f_2(\mathbf{u},t)$ is the probability density that at time t the first particle has velocity $\mathbf{v}$ and the second has velocity $\mathbf{u}$. Finally, the quantity F is interpreted as the probability density that the particles have velocities $(\mathbf{v},\mathbf{u})$ at time t, given that that at an earlier time $t-\tau$ they had velocities $(\mathbf{v}',\mathbf{u}')$. In other words F is the transition probability rate of the process $(\mathbf{v},\mathbf{u}) \to (\mathbf{v}',\mathbf{u}')$,

In thermodynamic equilibrium, the state of a gas becomes unchanged. Maxwell tell us that this happens when the number of pairs of molecules that changes from $(\mathbf{v},\mathbf{u}) \to (\mathbf{v}',\mathbf{u}')$ is equal to the number of the reverse process $(\mathbf{v}',\mathbf{u}') \to (\mathbf{v},\mathbf{u})$. This is equivalent to the condition

$$f_1(\mathbf{v}) f_2(\mathbf{u}) F = f_1(\mathbf{v}') f_2(\mathbf{u}') F', \tag{5}$$

where here F stands for the transition probability rate $F(\mathbf{v}-\mathbf{u}, \mathbf{v}'-\mathbf{u}')$ of the direct process and F' stands for the transition probability rate $F(\mathbf{v}'-\mathbf{u}', \mathbf{v}-\mathbf{u})$ of the reverse transition. The equality between (5) is identified as the reversibility condition usually called detailed balance condition.

Next, Maxwell makes the *probabilistic assumption* that the transition probability rate of the direct process equal that of the reverse process, that is, $F = F'$, from which follows the relation

$$f_1(\mathbf{v}) f_2(\mathbf{u}) = f_1(\mathbf{v}') f_2(\mathbf{u}'). \tag{6}$$

Using this relation, Maxwell derives his distribution of velocities using the conservation of the kinetic energy before and after collision, given by (72). Comparing this expression with the logarithm of (6), one concludes that

$$\ln f_1(\mathbf{v}) = c_1 - \beta m_1 v^2/2, \quad \ln f_2(\mathbf{u}) = c_2 - \beta m_2 u^2/2, \tag{7}$$

which may be written in the form

$$f_1(\mathbf{v}) = c_1' e^{-\beta m_1 v^2/2}, \quad f_2(\mathbf{u}) = c_2' e^{-\beta m_2 u^2/2}, \tag{8}$$

which is the Maxwell distribution of velocities.

If we denote by v_1, v_2, v_3 the Cartesian components of $\mathbf{v}$, then the Maxwell probability density distribution is written as

$$f(v_1, v_2, v_3) = c e^{-\beta m (v_1^2 + v_2^2 + v_3^2)/2}, \tag{9}$$

where $c = (\beta m / 2\pi)^{3/2}$. From this distribution we may obtain the probability distribution of the absolute value $v = (v_1^2 + v_2^2 + v_3^2)^{1/2}$ of $\mathbf{v}$, which is

$$h(v) = 4\pi c \left(\frac{\beta m}{2\pi}\right)^{3/2} v^2 e^{-\beta m v^2/2}. \tag{10}$$

It is worth writing the Maxwell distribution in terms of the kinetic energy $x = mv^2/2$, which is

$$g(x) = c' \sqrt{x}\, e^{-\beta x}, \tag{11}$$

where $c' = 2\beta^{3/2}/\sqrt{\pi}$.

Let us write the reversibility condition (5) as

$$\frac{F'}{F} = \frac{f_1(\mathbf{v}) f_2(\mathbf{u})}{f_1(\mathbf{v}') f_2(\mathbf{u}')}. \tag{12}$$

We have seen above that Maxwell assumed $F = F'$ from which he obtained the relation (6) and from this relation he reached the distribution of velocities. We could of course assume the Maxwell distribution of velocities and then conclude that $F = F'$, provided that the kinetic energy is conserved.

Let us suppose that instead of the velocities we use the kinetic energy to describing the state of the molecules. In this case, the reversibility condition is written as

$$\frac{G'}{G} = \frac{g_1(x)g_2(y)}{g_1(x')g_2(y')}, \tag{13}$$

where x and y are the kinetic energies of the molecules before a collision and x' and y' after the collision, and the g_1 and g_2 are the distribution of energies given by (11), and G and G' are the transition rates of the direct and reverse processes. Using (13) and bearing in mind the conservation of energy, $x + y = x' + y'$, we find

$$\frac{G'}{G} = \frac{\sqrt{xy}}{\sqrt{x'y'}}, \tag{14}$$

which is the ratio between the transition rates when the states of the molecules is written in term of the kinetic energy instead of the velocity.

3. Boltzmann equation

In the very first paragraph of his paper of 1872 [8, 13], Boltzmann declares the probabilistic character of kinetic theory. He says that "the problems of the mechanical theory of heat is a problem of probability theory". He then continues stating that a result of the theory of probability, "like the result of any other theory, is a necessary consequence of definite premises," and that the result is confirmed on account of the enormous number of molecules involved. In other words, Boltzmann says that the theory of probability is an essential part of kinetic theory and not just a method for obtained incomplete or approximate laws.

In his paper of 1872 [8, 13], Boltzmann retakes the Maxwell problem of finding the final distribution of velocities of a gas. In this process he reaches the equation that bears his name. Instead of using the velocity $\mathbf{v}$ to define the dynamic state of a molecule, as did Maxwell, Boltzmann chose the kinetic energy $x = mv^2/2$. In a second part of the paper, however he used the velocity as the dynamic state of a molecule. In the following we derive his equation using the space of kinetic energies and after that, the space of velocities.

Let $g(x, t)dx$ be the number of molecules in unit volume such that at time t the molecules have kinetic energy between x and $x + dx$. The change in the number of molecules per unit volume during an interval of time τ is $g(x, t + \tau)dx - g(x, t)dx$. To find this variation we determine the number of molecules per unit volume that, in the interval τ, acquires the energy between x and $x + dx$, which we denote by dA, and the number of molecules per unit volume that, in the same interval τ, loses energies between x and $x + dx$, which we denote by dB. We may thus write

$$g(x, t + \tau)dx - g(x, t)dx = dA - dB. \tag{15}$$

The change in the energy of the molecules is due to the elastic collision between two molecules. These collisions change the kinetic energies of the molecules while keeping their sum constant. The determination of dA is thus reduced to the determination of the number of collisions occurring in the interval τ such that after the collision one of the molecules have energy between x and $x + dx$. Similarly, dB is determined by considering that before the collision the energy of one of the molecules have energy between x and $x + dx$.

Let x and y be the kinetic energies of two colliding molecule and x' the kinetic energy of one of the molecule after the collision. The kinetic energy y' of the other molecule becomes determined by the conservation of the kinetic energy $x' + y' = x + y$. Let us denote by dn the number of collisions that takes place during an interval τ such that the kinetic energies of one of the molecules *before* the collision lies between x and $x + dx$ and the other is between y and $y + dy$. Boltzmann argues that this number is proportional to $g(x, t)dx$, to $g(y, t)dy$, and to a factor $\psi(x, y; x', y')dx'$ which depends on the nature of the collision. Therefore,

$$dn = \tau g(x, t)dx g(y, t)dy \psi(x, y; x', y')dx', \tag{16}$$

where $y' = x + y - x'$. If we integrate in all possible values of y and x', determined by the restriction $0 \leq x' \leq x + y$, we find the number of collisions in unit volume occurring in the interval τ such that the kinetic energy is between x and $x + dx$, which is the quantity dB, that is

$$dB = \tau dx \int_0^\infty \int_0^{x+y} g(x, t)g(y, t)\psi(x, y; x', y')dy dx'. \tag{17}$$

To write an expression for dA, we have to consider the the number of collisions that takes place during an interval τ such that the kinetic energy of one of the molecules *after* the collision lies between x and $x + dx$, and the energies of the molecules before the collision lies between ξ and $\xi + d\xi$, and between ξ' and $\xi' + d\xi'$, which we denote by $d\nu$. Boltzmann argues that this quantity is given by

$$d\nu = \tau g(x', t)dx' g(y', t)dy' \psi(x', y'; x, y)dx, \tag{18}$$

where $y = x' + y' - x$, and that

$$dA = \tau dx \int_0^\infty \int_0^{x+y} g(x', t)g(y', t)\psi(x', y'; x, y)dx' dy, \tag{19}$$

where $y' = x + y - x'$.

Replacing the expressions for dA and dB in equation (15), dividing by τ and taking the limit $\tau \to 0$, we find

$$\frac{\partial}{\partial t} g(x, t) = \int \{g(x', t)g(y', t)\psi(x', y'; x, y) - g(x, t)g(y, t)\psi(x, y; x', y')\}dx' dy, \tag{20}$$

where y' should be replaced by $x + x' - y$ and the integral carried out in the region $y \geq 0$ and $x' \leq x + y$, which is the Boltzmann equation.

Let $N(t)$ be the integral of $g(x,t)$ in the variable x. If we take the time derivative of N, and replace $\partial f/\partial t$ by the right-hand side of (20) we find $dN/dt = 0$. It means that N is a constant in time, and we may choose its value to be the unity. We may write then

$$\int_0^\infty g(x,t)dx = 1, \tag{21}$$

and that $g(x,t)$ can indeed be understand as a probability density distribution.

The crucial quantity in the derivation of the Boltzmann equation is $\psi(x,y;x',y')$, understood as the transition rate probability from from (x,y) to (x,y'), that is, proportional to the conditional probability density that the molecules have energies x' and y' at a certain time t, given that they had energies x and y at an earlier time $t-\tau$. This quantity cannot be derived from pure mechanics reasoning alone, but should be consistent with the conservation of mechanical energy, which, in the present case of hard spheres is the kinetic energy, $x+y = x'+y'$.

It is worth mentioning that dn equals $\tau dxdydx'$ multiplied by

$$g(x,t)g(y,t)\psi(x,y;x',y'), \tag{22}$$

which is the first term of the integrand in (20), and that $d\nu$ equals $\tau dxdydx'$ multiplied by

$$g(x',t)g(y',t)\psi(x',y';x,y), \tag{23}$$

which is the second term of the integrand in (20). The first term above is proportional to the probability density that the molecules have energies x' and y' at a certain time t *and* that they have energies x and y at an earlier time $t-\tau$. A similar interpretation is obtained for the second term above by exchanging (x',y') and (x,y). The equalities of these two quantities is the condition of detailed balance used by Maxwell to reach his distribution of velocities, and also used by Boltzmann as we shall see below.

Boltzmann adopts a form for $\psi(x,y;x',y')$ which is equivalent to

$$\psi(x,y;x',y') = \frac{\phi(x,y;x',y')}{\sqrt{xy}}, \tag{24}$$

where $\phi(x,y;x',y') = \phi(x',y';x,y)$, from which we get

$$\frac{\psi(x,y;x',y')}{\psi(x',y';x,y)} = \frac{\sqrt{x'y'}}{\sqrt{xy}}, \tag{25}$$

and we see that ψ coincides with G given by equation (14). Replacing the form (24) in (20) gives

$$\frac{\partial}{\partial t}g(x,t) = \int_0^\infty \int_0^{x+y} \left\{ \frac{g(x',t)}{\sqrt{x'}}\frac{g(y',t)}{\sqrt{y'}} - \frac{g(x,t)}{\sqrt{x}}\frac{g(y,t)}{\sqrt{y}} \right\} \phi(x,y;x',y')dydx', \tag{26}$$

which Boltzmann calls the fundamental equation for the time variation of $g(x,t)$. We recall that in this equation one should make the replacement $y' = x + x' - y$.

The detailed balance condition is represented by the equality of the expressions (22) and (23) or equivalently by the vanishing of the integrand of (26), which is

$$\frac{g(x')}{\sqrt{x'}}\frac{g(y')}{\sqrt{y'}} = \frac{g(x)}{\sqrt{x}}\frac{g(y)}{\sqrt{y}}. \tag{27}$$

To find f from this equation, we should keep in mind that this equation should be consistent with the conservation of kinetic energy $x'+y' = x+y$. Comparing this equation with the logarithm of the equation above we find

$$\ln g(x) - \frac{1}{2}\ln x = c - \alpha x. \tag{28}$$

or

$$g(x) = C\sqrt{x}e^{-\alpha x}. \tag{29}$$

Recalling that $x = mv^2/2$ is the kinetic energy of a molecule, we see that this is no other than the Maxwell distribution of velocities written in terms of the kinetic energy.

It is worth writing the Boltzmann equation (26) as

$$\frac{\partial}{\partial t}g(x,t) = \int \left\{ \frac{g(x',t)}{\sqrt{x'}}\frac{g(y',t)}{\sqrt{y'}} - \frac{g(x,t)}{\sqrt{x}}\frac{g(y,t)}{\sqrt{y}} \right\} \phi(x,y;x',y')dydx'dy', \tag{30}$$

where now we are including y' as the third variable of integration, with the understanding that the integral is carried out over the surface $y+y'-x' = x$ of the space spanned by the variables x', y, and y', such that $y \geq 0$, $x' \geq 0$, and $y' \geq 0$.

We now consider the space of velocities. Let $f(\mathbf{v})d\mathbf{v}$ be the probability of finding a molecule with velocity $\mathbf{v}$ withe Cartesian coordinates v_1, v_2, and v_3 inside the element $d\mathbf{v} = dv_1dv_2dv_3$. If $(\mathbf{v},\mathbf{u})$ are the velocities of two molecules before a collision and $(\mathbf{v}',\mathbf{u}')$ the velocities after the collision, we assume the probability transition rate is that given by Maxwell which we denoted by $F(\mathbf{v}-\mathbf{u},\mathbf{v}'-\mathbf{u}')$ and we recall that $|\mathbf{v}-\mathbf{u}| = |\mathbf{v}'-\mathbf{u}'|$. In other words, if this condition is not fulfilled, F vanishes. We remember that this condition is a consequence of the conservation of the kinetic energy, which in the case of molecules with the same mass is

$$v^2 + u^2 = v'^2 + u'^2. \tag{31}$$

Using the same reasonings employed before in the case of the space of kinetic energy, we arrive at the following equation for the time evolution of $f(\mathbf{v},\mathbf{t}$,

$$\frac{\partial}{\partial t}f(\mathbf{v},t) = \int \{f(\mathbf{v}',t)f(\mathbf{u}',t) - f(\mathbf{v},t)f(\mathbf{u},t)\} \times F(\mathbf{v}-\mathbf{u},\mathbf{v}'-\mathbf{u}')d\mathbf{u}d\mathbf{v}'d\mathbf{u}', \tag{32}$$

where the integral is performed in the surface defined by (31).

The detailed balance condition is now represented by the equation

$$f(\mathbf{v}')f(\mathbf{u}') = f(\mathbf{v})f(\mathbf{u}), \tag{33}$$

which is the same as the Maxwell condition given by equation (6). Therefore, we obtain from this condition the Maxwell distribution of velocities as we did above from equation (6).

4. Theorem of Irreversibility

As we have seen, Boltzmann derived the Maxwell distribution (29) by considering the stationary solution of equation (26) which is given by the condition of irreversibility or detailed balance (27). He then asked whether (29) would be the asymptotic solution of (26) for any initial condition. To answer this question he introduced the quantity

$$E = \int_0^\infty g(x,t)\left\{\ln\frac{g(x,t)}{\sqrt{x}} - 1\right\}dx, \tag{34}$$

and demonstrated that this quantity can never decrease, $dE/dt \geq 0$. We call this proposition the irreversibility theorem but it is usually called Boltzmann H-theorem. Boltzmann used the letter E in his original paper of 1872, as we did above, but some years later someone denoted it by the letter H, and Boltzmann himself ended up using the letter H in his book on kinetic theory. Boltzmann remarks that one could use the expression

$$E_1 = \int_0^\infty g(x,t)\ln\frac{g(x,t)}{\sqrt{x}}dx, \tag{35}$$

which differs from E by a constant as the integral of f equals the unity. Deriving E with respect to time, we find

$$\frac{dE}{dt} = \int_0^\infty \frac{\partial g(x,t)}{\partial t}\ln\frac{g(x,t)}{\sqrt{x}}dx. \tag{36}$$

To reach this result we should bear in mind that the integral of $g(x,t)$ in x equals the unity.

Replacing (26) in (36), we may write

$$\frac{dE}{dt} = \int \ln\frac{g(x,t)}{\sqrt{x}} \times \{g_2(x',y') - g_2(x,y)\}\phi(x,y;x',y')dxdx'dydy', \tag{37}$$

where $g_2(x,y,t) = g(x,t)g(y,t)/\sqrt{xy}$ and we have dropped the explicit time dependence of g_2. If we perform the changes of variables $x \leftrightarrow y$ and $x' \leftrightarrow y'$, we find an expression which added which added with (37) gives

$$\frac{dE}{dt} = \frac{1}{2}\int \ln g_2(x,y) \times \{g_2(x',y') - g_2(x,y)\}\phi(x,y;x',y')dxdx'dydy'. \tag{38}$$

If we now we now perform the changes $x \leftrightarrow x'$ and $y \leftrightarrow y'$, we find an expression which added which added with the previous expression gives

$$\frac{dE}{dt} = \frac{1}{4}\int \ln\frac{g_2(x,y)}{g_2(x',y')} \times \{g_2(x',y') - g_2(x,y)\}\phi(x,y;x',y')dxdx'dydy'. \tag{39}$$

It should be remarked that all transformations above leaves $\phi(x,y;x',y')$ invariant.

Let us write the integrand as

$$\mathcal{I} = \ln\frac{g_2}{g_2'}(g_2' - g_2)\phi. \tag{40}$$

If $g_2' \geq g_2$ than $\ln(g_2/g_2') \leq 0$ and $\mathcal{I} \leq 0$ because $\phi \geq 0$. If $g_2' \leq g_2$ than $\ln(g_2'/g_2) \geq 0$ and and again $\mathcal{I} \leq 0$. Therefore we conclude that the $dE/dt \leq 0$, that is, E can never decrease, and the the Boltzmann irreversibility theorem becomes demonstrated.

We now consider the theorem of irreversibility written in terms of the probability defined in the space of velocities. In this case the quantity E is defined by

$$E = \int f(\mathbf{v})\ln f(\mathbf{v})d\mathbf{v}. \tag{41}$$

Deriving this expression with respect to time,

$$\frac{dE}{dt} = \int \frac{\partial f(\mathbf{v})}{\partial t}\ln f(\mathbf{v})d\mathbf{v}. \tag{42}$$

Replacing $\partial f/\partial t$ given by the Boltzmann equation (32), we get

$$\frac{dE}{dt} = \int \ln f(\mathbf{v})\{f(\mathbf{v}',t)f(\mathbf{u}',t) - f(\mathbf{v},t)f(\mathbf{u},t)\} \times F(\mathbf{v}-\mathbf{u},\mathbf{v}'-\mathbf{u}')d\mathbf{u}d\mathbf{v}'d\mathbf{u}'d\mathbf{v}. \tag{43}$$

Performing the exchange of variables $\mathbf{v} \leftrightarrow \mathbf{u}$ and $\mathbf{v}' \leftrightarrow \mathbf{u}'$, and defining $f_2(\mathbf{v},\mathbf{u},t) = f(\mathbf{v},t)f(\mathbf{u},t)$, we obtain the expression

$$\frac{dE}{dt} = \frac{1}{2}\int \ln f_2(\mathbf{v},\mathbf{u}) \times \{f_2(\mathbf{v}',\mathbf{u}') - f_2(\mathbf{v},\mathbf{u})\}Fd\mathbf{u}d\mathbf{v}'d\mathbf{u}'d\mathbf{v}. \tag{44}$$

If now we perform the exchange of variables $\mathbf{v} \leftrightarrow \mathbf{v}'$ and $\mathbf{u} \leftrightarrow \mathbf{u}'$, we reach the expression

$$\frac{dE}{dt} = \frac{1}{2}\int \ln\frac{f_2(\mathbf{v},\mathbf{u})}{f_2(\mathbf{v}',\mathbf{u}')} \times \{f_2(\mathbf{v}',\mathbf{u}') - f_2(\mathbf{v},\mathbf{u})\}Fd\mathbf{u}d\mathbf{v}'d\mathbf{u}'.d\mathbf{v} \tag{45}$$

We remark that the exchanges of variables just employed leaves $F(\mathbf{v}-\mathbf{u},\mathbf{v}'-\mathbf{u}')$. Again, using the reasoning used above, we conclude that the integrand is negative from which follows that $dE/dt \leq 0$ because $F \geq 0$.

5. Kolmogorov equation

The Boltzmann equation can be understood as an approximation to the Komolgorov equation, the fundamental equation of stochastic thermodynamics. This equation gives the time evolution of the probability distribution defined on the space of all the coordinates and velocities of the particles of a collection of molecules.

Let us consider a gas consisting of N molecules of mass m each, that occupy a vessel of volume V. We denote by $\mathbf{r}_i$ and $\mathbf{v}_i$ the position and velocity of the i-th molecule, respectively. The collections of variables $\{\mathbf{r}_i\}$ and $\{\mathbf{v}_i\}$ are denoted by $\mathbf{r}$ and $\mathbf{v}$, respectively, and the state of the gas is denoted by $(\mathbf{r}, \mathbf{v})$. We use an abbreviation in which $(\mathbf{r}_i, \mathbf{v}_i)$ is denoted by ξ_i and the full state $(\mathbf{r}, \mathbf{v})$ by ξ. The states ξ comprise a vector space called phase space.

As time goes by the state of the gas follows a trajectory

$$\xi' \to \xi'' \to \xi''' \to \xi^{\text{iv}} \to \dots \tag{46}$$

in phase space and we are considering that time is discretized in intervals τ, which we consider to be small. Given that the state of the system is ξ' at time t we ask for the conditional probability density that it is found in state ξ at time $t+\tau$. The stochastic dynamics that is the basis of stochastic thermodynamics is founded on the *assumption* that the aforementioned conditional probability depends only on ξ' in addition to depend on ξ, and which we denote by $\mathcal{P}(\xi|\xi')$. Within the theory of stochastic processes, this assumption is known as Markovian. This conditional probability may also depend on t, but here we consider that it is time independent. It is worth pointing out that

$$\int \mathcal{P}(\xi|\xi') d\xi = 1, \tag{47}$$

independently of ξ'.

Let us ask for the probability density $\mathcal{P}(\xi, \xi')$ that the system is in state ξ at time $t+\tau$ *and* in state ξ' at time t. It is given by

$$\mathcal{P}(\xi, \xi') = \mathcal{P}(\xi|\xi')\mathcal{P}(\xi', t), \tag{48}$$

where $\mathcal{P}(\xi', t)$ is the probability density that the system is in state ξ' at time t. To determine the probability $\mathcal{P}(\xi, t+\tau)$ that the system is in state ξ at time $t+\tau$, it suffices to determine the marginal distribution associated with ξ from the joint probability density (48). This amounts to integrate (48) in ξ'. The result is

$$\mathcal{P}(\xi, t+\tau) = \int \mathcal{P}(\xi|\xi')\mathcal{P}(\xi', t) d\xi'. \tag{49}$$

From the property (47) and from (49) it follows that

$$\int \mathcal{P}(\xi, t) d\xi = 1, \tag{50}$$

at any time t.

Equation (49) gives the evolution of the probability distribution in discrete time. To derive the continuous time version of this equation, we define, for $\xi \neq \xi'$, the transition probability rate $\mathcal{W}(\xi|\xi')$ as the limit when $\tau \to 0$ of the ratio $\mathcal{P}(\xi|\xi')/\tau$. If we subtract (49) from the identity

$$\mathcal{P}(\xi, t) = \int \mathcal{P}(\xi'|x)\mathcal{P}(\xi, t) d\xi', \tag{51}$$

and divide the result by τ we find, after taking the limit $\tau \to 0$, the equation

$$\frac{d}{dt}\mathcal{P}(\xi) = \int \{\mathcal{W}(\xi|\xi')\mathcal{P}(\xi') - \mathcal{W}(\xi'|\xi)\mathcal{P}(\xi)\} d\xi', \tag{52}$$

where we have dropped the explicit dependence in t. This equation gives the time evolution of the probability distribution $\mathcal{P}(\xi)$ and is known as the Kolmogorov equation or a master equation. The equation is completely defined once we are given the transition probability rate $W(\xi|\xi')$.

The transition rate used by Maxwell and Boltzmann is such that only the states of two molecules will change, whereas the state of the others remain the same. If we let i and j be these two molecules this type of transition rate is denoted by

$$w(\xi_i', \xi_j'|\xi_i, \xi_j), \tag{53}$$

and the total transition rate is the sum

$$\mathcal{W}(\xi|\xi') = \frac{1}{N}\sum_{i<j} w(\xi_i', \xi_j'|\xi_i, \xi_j). \tag{54}$$

Replacing this expression in (52), we obtain the equation

$$\frac{d}{dt}\mathcal{P}(\xi) = \frac{1}{N}\sum_{i<j}\int \{w(\xi_i, \xi_j|\xi_i', \xi_j')\mathcal{P}(\xi') - w(\xi_i', \xi_j'|\xi_i, \xi_j)\mathcal{P}(\xi)\} d\xi'. \tag{55}$$

Let us integrate over all variables ξ_i except ξ_1. We find

$$\frac{d}{dt}\mathcal{P}(\xi_1) = \int \{w(\xi_1, \xi_2|\xi_1', \xi_2')\mathcal{P}(\xi_1', \xi_2') - w(\xi_1', \xi_2'|\xi_1, \xi_2)\mathcal{P}(\xi_1, \xi_2)\} d\xi_1' d\xi_2' d\xi_2. \tag{56}$$

This is an equation that gives the time evolution for the one-particle probability density $\mathcal{P}(\xi_1)$. It is an exact equation as much as equation (55) is, but it cannot be solved alone because it is not closed in $\mathcal{P}(\xi_1)$.

If we approximate the two-particle probability density $\mathcal{P}(\xi_1, \xi_2)$ by the product of the one-particle probability densities $\mathcal{P}(\xi_1)$ and $\mathcal{P}(\xi_2)$, we reach the equation

$$\frac{d}{dt}\mathcal{P}(\xi_1) = \int \{w(\xi_1, \xi_2|\xi_1', \xi_2')\mathcal{P}(\xi_1')\mathcal{P}(\xi_2') - w(\xi_1', \xi_2'|\xi_1, \xi_2)\mathcal{P}(\xi_1)\mathcal{P}(\xi_2)\} d\xi_1' d\xi_2' d\xi_2, \tag{57}$$

which is a closed equation for the one-particle probability density and can be solved. In this approximation, the master equation is recognized as the Boltzmann equation. It should be remarked that the original Boltzmann equation, given by (32), is written in terms of velocities only. This happens because the transition probability rate considered by Boltzmann, as well as by Maxwell, depends only on velocities of the molecules and not on their positions.

Considering that the transition rate w depends on the velocities, we may integrate (57) in all positions to find

$$\frac{d}{dt}\mathcal{P}(\mathbf{v}_1) = \int \{w(\mathbf{v}_1, \mathbf{v}_2|\mathbf{v}_1', \mathbf{v}_2')\mathcal{P}(\mathbf{v}_1')\mathcal{P}(\mathbf{v}_2') - w(\mathbf{v}_1', \mathbf{v}_2'|\mathbf{v}_1, \mathbf{v}_2)\mathcal{P}(\mathbf{v}_1)\mathcal{P}(\mathbf{v}_2)\}d\mathbf{v}_1'd\mathbf{v}_2'd\mathbf{v}_2. \tag{58}$$

Using the property

$$w(\mathbf{v}_1', \mathbf{v}_2'|\mathbf{v}_1, \mathbf{v}_2) = w(\mathbf{v}_1, \mathbf{v}_2|\mathbf{v}_1', \mathbf{v}_2'), \tag{59}$$

it can be written as

$$\frac{d}{dt}\mathcal{P}(\mathbf{v}_1) = \int \{\mathcal{P}(\mathbf{v}_1')\mathcal{P}(\mathbf{v}_2') - \mathcal{P}(\mathbf{v}_1)\mathcal{P}(\mathbf{v}_2)\} \times w(\mathbf{v}_1, \mathbf{v}_2|\mathbf{v}_1', \mathbf{v}_2')d\mathbf{v}_1'd\mathbf{v}_2'd\mathbf{v}_2, \tag{60}$$

which is identical to the Boltzmann equation (32).

The above result allow us to conclude that the Boltzmann equation can be understood a master equation in the approximation of one-particle, which is an approximation equivalent to say that the dynamic variables of distinct molecules are statistically independent. This independence assumption was called molecular disorder (*molekularen Unordnung*) by the Ehrenfests [24, 25].

6. Entropy Production

One of the fundamental concepts of stochastic thermodynamics is that of entropy production, to be defined in the following. The entropy of the system described by the probability density $\mathcal{P}(\xi, t)$ is defined by

$$S = -\int \mathcal{P}(\xi)\ln \mathcal{P}(\xi)d\xi, \tag{61}$$

where we are dropping the explicit dependence on t, Its time derivative is

$$\frac{dS}{dt} = -\int \frac{\partial \mathcal{P}(\xi)}{\partial t}\ln \mathcal{P}(\xi)d\xi. \tag{62}$$

Replacing $\partial\mathcal{P}/\partial t$ given by (52), we find

$$\frac{dS}{dt} = \int \mathcal{W}(\xi'|\xi)\mathcal{P}(\xi)\ln\frac{\mathcal{P}(\xi)}{\mathcal{P}(\xi')}d\xi d\xi'. \tag{63}$$

The right-hand side of the equation above is split in two parts

$$\frac{dS}{dt} = \Pi + \Psi, \tag{64}$$

where Ψ is the entropy flux from the outside to the system, given by

$$\Psi = -\int \mathcal{W}(\xi'|\xi)\mathcal{P}(\xi)\ln\frac{\mathcal{W}(\xi'|\xi)}{\mathcal{W}(\xi|\xi')}d\xi d\xi', \tag{65}$$

and Π is the production of entropy, given by

$$\Pi = \int \mathcal{W}(\xi'|\xi)\mathcal{P}(\xi)\ln\frac{\mathcal{W}(\xi'|\xi)\mathcal{P}(\xi)}{\mathcal{W}(\xi|\xi')\mathcal{P}(\xi')}d\xi d\xi'. \tag{66}$$

We remark that the entropy is not a conserved quantity like energy. The time variation of the entropy is not in general equal to the flux of entropy Ψ. Entropy is also created by virtue of the irreversible processes occurring inside the system. This contribution is the entropy production rate Π and is always nonnegative.

To demonstrate that Π is a nonnegative quantity we write it as

$$\Pi = \frac{1}{2}\int \{\mathcal{W}(\xi'|\xi)\mathcal{P}(\xi) - \mathcal{W}(\xi|\xi')\mathcal{P}(\xi')\} \times \ln\frac{\mathcal{W}(\xi'|\xi)\mathcal{P}(\xi)}{\mathcal{W}(\xi|\xi')\mathcal{P}(\xi')}d\xi d\xi', \tag{67}$$

which is obtained performing the exchange of variables $\xi \leftrightarrow \xi'$ in the integral of equation (66). The integrand in (67) has the form $(a - b)\ln(a/b)$ which is a nonnegative expression, a result that allows us to conclude that $\Pi \geq 0$.

Let us consider the transition probability rate employed by Maxwell and Boltzmann. The transition rate they employed is invariant under the exchange of the velocities of the molecules before and after the collisions. This is equivalent to say that

$$\mathcal{W}(\xi'|\xi) = \mathcal{W}(\xi|\xi'). \tag{68}$$

Replacing this property in the expression (65) shows that the flux of entropy vanishes identically. Thus the Kolmogorov equation associated with the Boltzmann kinetic theory describes a closed system. There is no exchange of entropy nor of energy. As to the entropy production rate, wee see that it coincides with the variation of entropy, that is $\Pi = dS/dt$. Since $\Pi \geq 0$, it follows that

$$\frac{dS}{dt} \geq 0, \tag{69}$$

which can be understood as the Boltzmann theorem of irreversibility. To reach this conclusion, it suffices to see that the entropy S is the Boltzmann quantity E with the reversed sign.

To verify that equation (45) derived by Boltzmann is indeed identical to (63), it suffices to use (68), and write the probability distribution $\mathcal{P}(\xi)$ as a product of one-particle probability distributions.

7. Collision of two particles

Let $\mathbf{v}$ and $\mathbf{u}$ be the velocities of two molecules before the collision and $\mathbf{v}'$ and $\mathbf{u}'$ after the collision. If the kinetic energy and momentum are conserved before and after collision then the absolute value of the relative velocities are the same before and after the collision,

$$|\mathbf{v}-\mathbf{u}| = |\mathbf{v}'-\mathbf{u}'|. \tag{70}$$

Usually this result is demonstrated assuming that the particles obeys the Newton equations of motion. However, this is not necessary. It is a consequence of any type of motion, be it Newtonian or not, as long as it leads to the conservation of energy and momentum.

Our premises concerns the total momentum of the two molecules and the total kinetic energy. The total momentum is the same before and after the collision,

$$m_1\mathbf{v}' + m_2\mathbf{u}' = m_1\mathbf{v} + m_2\mathbf{u}, \tag{71}$$

as well as the kinetic energy, which we write as

$$m_1 v'^2 + m_2 u'^2 = m_1 v^2 + m_2 u^2, \tag{72}$$

where m_1 and m_2 are the masses of the molecules, and we are using a boldface letter to represent a vector and an ordinary letter to represent its absolute value.

The two equations above can be obtained from a parametric formulation as follows

$$\mathbf{v}' = \mathbf{c} + \frac{\mu}{m_1}|\mathbf{v}-\mathbf{u}|\mathbf{e}, \tag{73}$$

$$\mathbf{u}' = \mathbf{c} - \frac{\mu}{m_2}|\mathbf{v}-\mathbf{u}|\mathbf{e}, \tag{74}$$

where $\mu = m_1 m_2/(m_1+m_2)$, $\mathbf{e}$ is a unit vector, $|\mathbf{e}| = 1$, and $\mathbf{c}$ is the velocity of the center of mass, defined by

$$\mathbf{c} = \frac{m_1\mathbf{v} + m_2\mathbf{u}}{m_1+m_2}. \tag{75}$$

That is, given $\mathbf{v}$ and $\mathbf{u}$, then $\mathbf{v}'$ and $\mathbf{u}'$ given parametrically by the equations (73) and (74) obey the two conservation laws above for an arbitrary unitary vector $\mathbf{e}$.

Multiplying (73) by m_1 and summing with (74) multiplied by m_2, we find the result

$$m_1\mathbf{v}' + m_2\mathbf{u}' = (m_1+m_2)\mathbf{c}, \tag{76}$$

which is equivalent to the conservation of momentum (71) if we consider the definition of $\mathbf{c}$.

Taking the square of (73) and of (74), summing the resulting equations after multiplying by m_1 and m_2, respectively, we find

$$m_1 v'^2 + m_2 u'^2 = (m_1+m_2)c^2 + \mu|\mathbf{v}-\mathbf{u}|^2. \tag{77}$$

Replacing the definition of c in this equation we are led to the conservation of energy (72).

Subtracting the equations (73) and (74), we get

$$\mathbf{v}' - \mathbf{u}' = |\mathbf{v}-\mathbf{u}|\mathbf{e}, \tag{78}$$

from which follows the result (70) because $|\mathbf{e}| = 1$.

In the one-dimensional case,

$$v' - u' = |v-u|e \tag{79}$$

where e takes the values $+1$ if $v < u$ and -1 if $v > u$. In any case $v' - u' = u - v$ and v' and u' are uniquely given by

$$v' = c + \frac{\mu}{m_1}(u-v) = \frac{(m_1-m_2)v + 2m_2 u}{m_1+m_2}, \tag{80}$$

$$u' = c - \frac{\mu}{m_2}(u-v) = \frac{(m_2-m_1)u + 2m_1 v}{m_1+m_2}. \tag{81}$$

The transition rates that we consider here are associated with processes in which the velocities of two molecules changes while their momentum and their kinetic energy remains unchanged. The parametric formulation that we defined above becomes appropriate to describe such processes. The transition rate can then be calculate by considering that the unit vector $\mathbf{e}$ is an independent random variable with a given probability density distribution which we denote by $P(\mathbf{e})$. The conditional probability density of $(\mathbf{v}', \mathbf{u}')$ given $(\mathbf{v}, \mathbf{u})$ is obtained by

$$P(\mathbf{v}', \mathbf{u}'|\mathbf{v}, \mathbf{u})d\mathbf{v}'d\mathbf{u}'. = P(\mathbf{e})d\mathbf{e} \tag{82}$$

In one dimension the velocities v' and u' are uniquely determined by v and u. Therefore in one dimension the motion determined by the conservation of kinetic energy and momentum cannot be stochastic and remains deterministic. In fact, in this case the conservation of kinetic energy and momentum leads to the conclusion that the motion follows the Newton equations of motion.

Another result of the one dimensional case is as follows. If all particles have the same mass than from the equations (80) and (81), the particles exchange their velocities in a collision. Therefore the probability density distribution does not change and remains that of the initial condition. There is no relaxation to the Maxwell distribution, unless the initial condition is already a Maxwell distribution. In other words, both sides of the Boltzmamm equation vanish identically.

8. Conclusion

We have shown that the kinetic theory of gases developed by Clausius, Maxwell, and Boltzmann can be understood as stochastic thermodynamics. This result was accomplished by showing that the dynamics used by Boltzmann to reach his equation is a stochastic dynamics and not properly a deterministic Newton equation of motion. Using the probability transition rate devised by Maxwell and used by Boltzmann we constructed the

Kolmogorov equation, that gives the time evolution of the probability density distribution associated with the whole state of the system.

The Kolmogorov equation is the basic equation of the stochastic thermodynamics from which we determine the time evolution of the entropy as well as the flux of entropy and the entropy production rate. The Boltzmann equation is reached using an approximation which corresponds to consider the velocities of distinct molecules as independent statistically. Within this approximation we have also shown that the entropy production rate is associate with the Boltzmann H-theorem.

The Markov assumption is the fundamental hypothesis of the stochastic dynamics developed here. Another assumption is the one related to statistical independence of the molecules, which we have employed as an approximation to reach the Boltzmann equation from the Kolmogorov equation. These two distinct assumption are implicit in the paper of Boltzmann of 1872 but they are not clearly distinguished in the historical texts on the kinetic theory. Boltzmann himself did not distinguish them and considered them consequences of the gas being molecularly disordered (*molekular Ungeordnet*) [3]. The two assumptions were distinguished by Jeans but he considered them consequences of *molecular chaos* [5]. The Ehrenfests on the other hand not only distinguish them but gave them different names. The fundamental hypothesis concerning the stochastic dynamics they called the collision number hypothesis (*Stosszahlansatz*) [24, 25] and the one related to the statistical independence of velocities they called molecular disorder (*molekularen Unordnung*) [24, 25].

References

[1] H.W. Watson, *A Treatise on the Kinetic Theory of Gases* (Clarendon Press, Oxford, 1876).

[2] O.E. Meyer, *Die kinetische Theorie der Gase* (Maruschke und Berendt, Breslau, 1877).

[3] L. Boltzmann, *Vorlesungen über Gastheorie* (Barth, Leipzig, 1896–1898), v. 1–2.

[4] S.H. Burbury, *A Treatise on the Kinetic Theory of Gases* (Cambridge University Press, Cambridge, 1899).

[5] J.H. Jeans, *The Dynamical Theory of Gases* (Cambridge Universiyt Press, Cambridge, 1904).

[6] L. Boltzmann, *Lectures on Gas Theory* (University of California Press, Berkeley, 1964).

[7] S.G. Brush, *Kinetic Theory* (Pergamon Press, Oxford, 1965), v. 1.

[8] S.G. Brush, *Kinetic Theory* (Pergamon Press, Oxford, 1966), v. 2.

[9] R. Clausius, Annalen der Physik und Chemie **100**, 353 (1857).

[10] J.C. Maxwell, Philosophical Magazine **19**, 19 (1860).

[11] J.C. Maxwell, Philosophical Magazine **20**, 21 (1860).

[12] J.C. Maxwell, Philosophical Transactions of the Royal Society of London **157**, 49 (1867).

[13] L. Boltzmann, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften **66**, 275 (1872).

[14] K. Huang, *Statistical Mechanics* (Wiley, New York, 1963).

[15] Y.B. Rumer and M.S. Ryvkin, *Thermodynamics, Statistical Physics and Kinetics* (Mir, Moscow, 1980).

[16] L.E. Reichl, *A Modern Course in Statistical Mechanics* (University of Texas Press, Austin, 1980).

[17] C. Cercignani, *The Boltzmann Equation and its Applications* (Springer, New York, 1988).

[18] S.R.A. Salinas, *Introduction to Statistical Physics* (Springer, New York, 2001).

[19] G.M. Kremer, *Uma Introdução à Equação de Boltzmann* (EDUSP, São Paulo, 2005).

[20] G.M. Kremer, *An Introduction to the Boltzmann Equation and Transport Processes in Gases* (Springer, Berlin, 2010).

[21] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

[22] T. Tomé and M.J. de Oliveira, *Stochastic Dynamics and Irreversibility* (Springer, Cham, 2015).

[23] M.J. de Oliveira, Rev. Bras. Ens. Fis. **44**, e20220087 (2022).

[24] P. Ehrenfest and T. Ehrenfest *in* Encyklopädie der mathematischen Wissenschaften (Teubner, Leipzig, 1907–1914); volume 4, issue 4, article 32.

[25] P. Ehrenfest and T. Ehrenfest, *The Conceptual Foundations of the Statistical Approach in Mechanics* (Dover, New York, 1990).

[26] P. Ehrenfest and T. Ehrenfest, Physikalische Zeitschrift **8**, 311 (1907).

[27] J.W. Gibbs, *Elementary Principles in Statistical Mechanis* (Scribner, New York, 1902).

[28] T. Tomé and M.J. de Oliveira, Phys. Rev. E **82**, 021120 (2010).

[29] T. Tomé and M.J. de Oliveira, Phys. Rev. E **91**, 042140 (2015).

[30] T. Tomé and M.J. de Oliveira, J. Chem. Phys. **148**, 224104 (2018).

[31] M.J. de Oliveira, Phys. Rev. E **99**, 052138 (2019).

[32] M.J. de Oliveira, Rev. Bras. Ens. Fis. **42**, e20200210 (2020).

[33] E. Garber, S.G. Brush and C.W.F. Everitt, *Maxwell on Molecules and Gases* (MIT Press, Cambridge, 1986).

[34] M.J. de Oliveira, Rev. Bras. Ens. Fis. **41**, e20180307 (2019).

[35] A. Walstad, Am. J. Phys. **81**, 555 (2013).