Philosophical Magazine, 2013
Vol. 93, No. 20, 2576–2612, http://dx.doi.org/10.1080/14786435.2013.784402

COMMENTARY
The discovery of thermodynamics

Peter Weinberger*

Center for Computational Nanoscience, Seilerstätte 10/21, A1010 Vienna, Austria

(*Received 21 December 2012; final version received 6 March 2013*)

Based on the idea that a scientific journal is also an "agora" (Greek: market place) for the exchange of ideas and scientific concepts, the history of thermodynamics between 1800 and 1910 as documented in the Philosophical Magazine Archives is uncovered. Famous scientists such as Joule, Thomson (Lord Kelvin), Clausius, Maxwell or Boltzmann shared this forum. Not always in the most friendly manner. It is interesting to find out, how difficult it was to describe in a scientific (mathematical) language a phenomenon like "heat", to see, how long it took to arrive at one of the fundamental principles in physics: entropy. Scientific progress started from the simple rule of Boyle and Mariotte dating from the late eighteenth century and arrived in the twentieth century with the concept of probabilities. Thermodynamics was the driving intellectual force behind the industrial revolution, behind the enormous social changes caused by this revolution. The history of thermodynamics is a fascinating story, which also gives insights into the mechanism that seem to govern science.

Keywords: thermodynamics; thermotransport; thermomechanical

1. Introduction

Most likely, thermodynamics was and is the (technologically) most successful part of physics in the last 200 years. Of course, nowadays it is very easy to summarize formally the key equations of thermodynamics in just three lines,

$$F = U - TS, \tag{1}$$

$$G = H - TS, \tag{2}$$

$$H = U + pv, \tag{3}$$

where F is the Helmholtz free energy, G the Gibbs free energy, H the enthalpy, U the internal energy, T the temperature, S the entropy, p the pressure and v the volume, but back in the first half of the nineteenth century none of the quantities denoted by a capital letter had a name nor did they have any clear meaning. It took altogether almost a century till the last "mysterious" term, namely the entropy, was fully accepted by the physics community. To make absolutely clear what kind of gigantic progress was made, just imagine to sit – around 1820 – in front of a fire place with a nice coal fire burning and to wonder of what physical origin the experienced "heat" possibly might be. Even today it would not be easy at all to answer this question in simple terms.

*Email: peter@pwein.at

© 2013 Taylor & Francis

Quite clearly, at the beginning of the nineteenth century, which usually is recognized as being the beginning of the industrial revolution, the interest in gases, in particular steam, was mainly driven by the invention and subsequent use of steam engines (James Watt, around 1770, and others).

At least some relations for gases were already well known, namely Boyle's (1662) and Mariotte's (1676) law,

$$pv = k, \qquad \text{(Boyle-Mariotte)}$$

and Gay–Lussac's relation (1802),

$$V = k'(a + t), \qquad \text{(Gay-Lussac)}$$

where p is the pressure, v the volume, t the temperature in centigrades and a, k and k' are constants. Carnot's (Nicolas Léonard Sadi, 1796–1832) splendid explanation of steam engines in his *Reflections on the Motive Power of Fire* [1], although many times cited, remained slightly mysterious for a number of years.

In modern sociology of science, the term *agora* (Greek, market place) is frequently used for the social structure of scientific communities. A scientific journal can be viewed as an *agora*, as a place suitable for the exchange of ideas, as a market place of cultural importance. The Philosophical Magazine Archives give a wonderful account of the conceptual progress in thermodynamics made in the nineteenth century, since many famous papers by Joule, Thomson (Kelvin), Clausius, Maxwell, Boltzmann and many others got published there or were reproduced as translations. For this very reason, almost exclusively, only quotations and citations from the Philosophical Magazine are used in the following. Consequently, the conceptual history compiled below reflects only the progress of ideas in a single scientific auditorium (*agora*), it cannot claim to provide a complete account of the history of thermodynamics. There is a wide variety of books dedicated to this subject.

It should be noted that all equations specified by the name of a particular author, refer to the original formulation. Since some of the concepts are written in a formal language that might appear difficult to understand, two appendices, one dealing with the kinetic theory of gases, the other one with a more modern view of Boltzmann statistics, are provided.

2. Joule and the concept of heat

Perhaps one of the very first people (in Britain) to investigate various forms of heat experimentally in a scientific rigorous and systematic manner was Joule (James Prescott, 1818–1889), a brewer by profession. Very typically for his time, he first started out [2,3] to deal with heat generated in wet-chemical processes and electrochemical phenomena. From his carefully described experiments he concluded[1] *that the heat evolved by other chemical actions, besides which is called combustion, is caused by resistance to electric conduction …* [2]. While the heat of combustion remained a bit mysterious, the heat produced in magneto-electric experiments seemed to lead to first quantitative relations: (1) *the heat evolved by a bar of iron revolving between the poles of a magnet is proportional to the square of the inductive force* and (2) *the heat evolved by the coil of the magneto-electrical machine is proportional to the square of the current* [4].

Joule quickly realized that in obtaining a concept for heat at least a quantity that would be specific for individual (chemical) systems was needed, in particular, since an adequate unit for heat obviously was missing. The specific heat as formulated by Petit and Dulong [5],

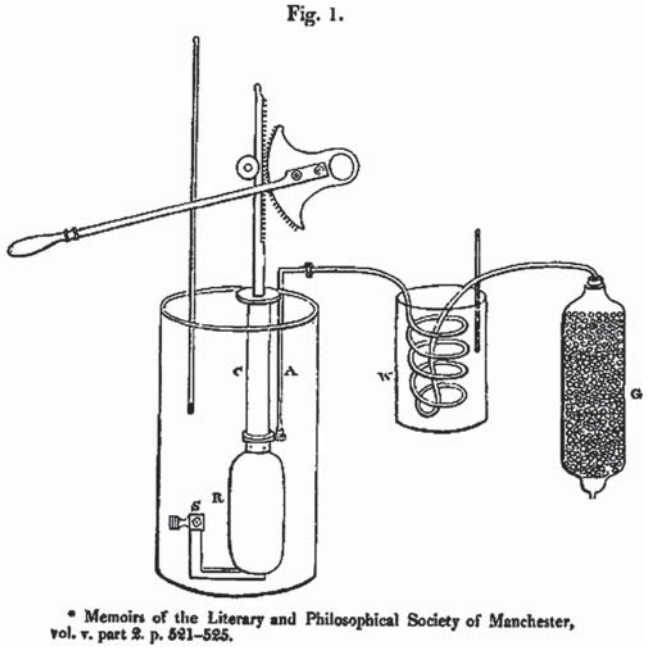


Figure 1. First part of Joule's original 1845 set-up to investigate *changes of temperature produced by the rare-faction and condensation of air* [7]: a pump C, a pipe A, a receiver R, a vessel W containing water, and a vessel G filled with small pieces of $CaCl_2$. Air dried by means of G is led into the pump. The temperature is measured in W and in the container of water into which C and R are immersed.

namely *that the specific heat of simple bodies is inversely proportional to their atomic weights* [6] seemed to provide a first key, but did not convince Joule completely.

Besides electricity and magnetism, the other very popular scientific topic in the first half of the nineteenth century was the handling of gases, in particular, since in galvanic processes gases were produced. In 1845, Joule constructed a remarkable device [7] consisting of a pump and one (Figure 1) or two receivers (Figure 2). In a first set of experiments he found that the work to operate the pump could be related to the heat caused by compressing air in a receiver. This work he termed *mechanical equivalent of heat*. Furthermore, in connecting that receiver with another receiver containing air under reduced pressure he managed to show that *no change of temperature occurs when air is allowed to expand in such a manner as to develop mechanical power* [7]. He even made sure that it did not matter whether the two receivers were immersed in one water tank or in separate ones, see Figure 2.

The experiments with the two receivers proved that obviously a kind of conservation law of heat had to be anticipated, since the heating up in one receiver and the cooling in the other showed a vanishing net gain in temperature of the water tank(s). The effect of cooling Joule encountered is now called Joule–Thomson effect and is the underlying principle of refrigerators, air conditioners, heat pumps, etc. The idea of having discovered an *equivalent relation between heat and the ordinary forms of mechanical power* led Joule to also investigate the heat of friction [8,9]. To show his *equivalent relation* he had a paddle wheel running in various kinds of liquids: it needed mechanical power to rotate the paddle wheel, while at the same time the heating up of the liquid could be observed. Like his

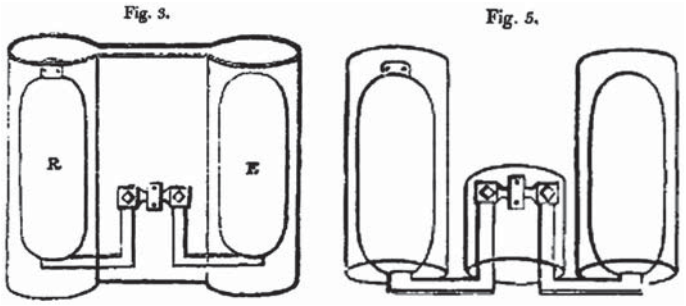


Figure 2. Second part of Joule's set-up [7]. In his *Figure 3*, *R* is a receiver filled with about 20 atmospheres of dry air and *E* a receiver exhausted by means of an air pump. In his *Figure 3* both receivers were put in the same water bath, in his *Figure 5* in separate baths. The two receivers were connected by means of stop-cocks.

compatriot Michael Faraday, Joule was mainly interested in the factual outcome of his experiments. He was perhaps less intrigued to attempt a formalization of the relations he discovered.

3. Clausius and the first fundamental law of thermodynamics

In 1851 Clausius (Rudolf von, 1822–1888), a German mathematician and physicist, published a series of papers in Poggendorff's Annalen, all of which were translated and published in the Philosophical Magazine [10–12]. In particular his paper entitled by *On the moving force of heat, and the laws regarding the nature of heat itself which are deducible there from* [11] made him internationally famous: it introduced a mathematically oriented language that in essence is still the backbone of modern (phenomenological) thermodynamics.

In this paper [11], he sets out by rephrasing Joule's *equivalent relation. In all cases where work is produced by heat, a quantity of heat proportional to the work done is expended; and inversely, by the expenditure of a like quantity of work, the same amount of heat may be produced* [11], namely as a hypothesis (*maxim*), from which a mathematical treatment can be deduced.

Let a certain quantity of permanent gas … be given, he starts his argumentation in [11], then in *a union of the laws of* [Boyle–] *Mariotte and Gay-Lussac*,

$$pv = R(a + t), \tag{Clausius 1}$$

where p is the pressure, v the volume, t the temperature in centigrades, a [is] *a constant equal for all gases and also R* [is] *a constant … . This last constant is … different for different gases … .* In principle, as he explains in terms of an example, it is clear that by investigating a gas (in ideal conditions) in two different states, say p_0, v_0, t_0 and p_1, v_1, t_1, the constant a assumes the (universal) value of $a = 273$.

Being a mathematician, Clausius knew very well how to express empirical laws into differential equations. He immediately concluded from the Boyle–Mariotte-Gay-Lussac law that [11] *the increase of pressure* [at constant volume v] … *due to the differential* $t - \tau = dt$ … *can be expressed in terms of* v *and* t … *as follows:*

$$dp = \frac{Rdt}{v}. \tag{Clausius 2}$$

If the increase of volume … be denoted by dv, we obtain …

$$\textit{The work produced} = \frac{Rdvdt}{v}. \qquad \text{(Clausius 3)}$$

Thus, obviously only one more quantity was needed to complete the description: let the *amount of heat which must be imparted to change the gas by a definite process from any given state to another, in which its volume is v and its temperature is t, be called Q* [11], then by neglecting all differentials of higher order, one obtains

$$\textit{The heat expended} = \left[\frac{d}{dt}\left(\frac{dQ}{dv}\right) - \frac{d}{dv}\left(\frac{dQ}{dt}\right)\right] dvdt. \qquad \text{(Clausius 4)}$$

Finally, by defining the following quotient

$$A = \frac{\textit{The heat expended}}{\textit{The work produced}}, \qquad \text{(Clausius 5)}$$

one gets

$$\frac{d}{dt}\left(\frac{dQ}{dv}\right) - \frac{d}{dv}\left(\frac{dQ}{dt}\right) = \frac{AR}{v}, \qquad \text{(Clausius 6)}$$

which in turn *can be brought under the form of a complete differential, thus:*

$$dQ = dU + AR\frac{a+t}{v}dv, \qquad \text{(Clausius 7)}$$

where U denotes an arbitrary function of v and t [11].

In considering the case of a vapor in equilibrium with its fluid, Clausius was even able to show that *the function U … cannot contain v, and hence the equation changes to*

$$dQ = cdt + AR\frac{a+t}{v}dv, \qquad \text{(Clausius 8)}$$

where c can only be a function of t … [11].

From this equation the consequences for the specific heat dQ/dt,

$$v = const., \quad dv = 0 : \frac{dQ}{dt} = c, \qquad \text{(Clausius 9)}$$

$$p = const., \quad dv = \frac{Rdt}{p} : \frac{dQ}{dt} = c' = c + AR, \qquad \text{(Clausius 10)}$$

can immediately be read off, namely *that the difference of both specific heats for every gas is a constant quantity AR* [11].

4. Thomson and the dynamical theory of heat

Clausius' approach was very quickly adopted by the scientific community, since only one year later, in 1852, Thomson, the later Lord Kelvin (William, 1824–1907), stated in an article entitled by *On the dynamical theory of heat, with numerical results deduced from Mr. Joule' equivalent of a thermal unit, and M. Regnault's observations on steam* that *the whole theory of the motive power of heat is founded on … two … propositions, due respectively to Joule, and to Carnot and Clausius* [13].

It is worthwhile to follow Thomson's argumentation, since, as compared to Clausius, his style of formulation is much more physics oriented.

Let *the quantity of work which it* [a mass of any substance] *will produce … be*

$$vdv\,; \tag{Thomson 1}$$

and the quantity of heat which must be added to it to make its temperature rise during the expansion to $t + dt$ may be denoted by

$$Mdv + Ndt. \tag{Thomson 2}$$

The mechanical equivalent of this is

$$J(Mdv + Ndt), \tag{Thomson 3}$$

if J denotes the mechanical equivalent of a unit of heat. Hence the mechanical measure of the external effect produced … is [13]

$$(p - JM)dv - JNdt. \tag{Thomson 4}$$

In here v, p and t are again the volume, the pressure and the temperature. *J is an absolute constant*, M and N seem to be function of t or v only, since

$$\frac{dp}{dt} = J\left(\frac{dM}{dt} - \frac{dN}{dv}\right). \tag{Thomson 5}$$

This equation, as Thomson insists, *expresses in a perfectly comprehensive manner the application of the first fundamental proposition to the thermal and mechanical circumstances of any substance whatever, … , when subjected to any possible variations of temperature, volume and pressure* [13].

As far as the second proposition is concerned, namely that due *to Carnot and Clausius* [13], it *is completely expressed by the equation*

$$\mu = \frac{1}{M}\frac{dp}{dt}, \tag{Thomson 6}$$

where μ denotes what is called "Carnot's function", a quantity which has an absolute value, the same for all substances for any given temperature … [13].

In the third part of his paper [13], he added that *N must be what is commonly called the "specific heat at constant volume"* [13] and if K denotes the specific heat under constant pressure, i.e. $(K - N)dt = Mdv$, then *M merely shows the meaning in terms of the two specific heats* …[13], which is what Clausius obtained in the case of an ideal gas, see Equations (Clausius 9) and (Clausius 10).

5. A flash forward

Before going ahead with further methodological achievements and historical details, it is perhaps quite useful to point out, how far Clausius and Thomson already got. Rewriting Clausius's main equation, see Equation (Clausius 7), in a more familiar notation ($c = c_v$, Equation (Clausius 9), U internal energy), assuming one Mol of an ideal gas ($A = 1$) and dividing by T, one gets

$$\frac{dQ}{T} = \underbrace{\frac{c_v}{T}\,dT + \frac{R}{v}\,dv}_{\text{Clausius, 1851}} \equiv \frac{dU + pdv}{T}. \tag{4}$$

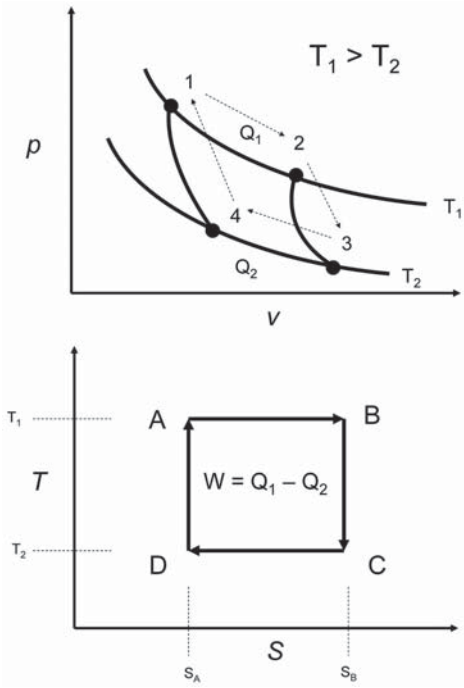


Figure 3. pv- and TS diagrams corresponding to a Carnot cycle.

Integrated between states A and B "simply" yields the change in *entropy*

$$\Delta S = S_B - S_A = \int_A^B \frac{dQ}{T}\text{, or, } Q = \int_A^B T dS. \tag{5}$$

"Simply" of course in terms of present days' knowledge, since the term **entropy** still had to be coined and its meaning to be discovered. It needed another almost 20 years till it was recognized that dQ/T is not a kind of renormalized heat.

From Figure 3 one can guess, why in terms of a pressure (p) vs. volume (v) description of the Carnot process seemed to be so complicated, but also that by using entropies an explanation can be almost read off. If T_1 denotes the temperature of a hot reservoir and T_2 that of a cold one, then the heat (thermal energy) transferred between these two reservoirs is given in terms of Equation (5) by

$$\Delta Q = Q_1 - Q_2 = (T_1 - T_2)(S_B - S_A), \tag{6}$$

and the efficiency of a Carnot machine – the then main topic of interest – by

$$\eta = \frac{\Delta Q}{Q_1} = 1 - \frac{T_2}{T_1}. \tag{7}$$

6. Clausius and the absolute temperature

In 1856 Clausius formulated a *theorem of the equivalence of transformations* [14] in which he introduced the quantity Q/T, which he called *equivalence-value*. In the presence of more than one reservoir the corresponding equivalence values had to be additive, i.e.

$$N = \sum_i \frac{Q_i}{T_i} = \int \frac{dQ}{T}, \qquad \text{(Clausius 11)}$$

and *if the process is reversible, then, however complicated it may be, we can prove … that the transformations which occur must exactly cancel each other, so that their algebraic sum is zero.* [14]

Furthermore, *the equation*

$$\int \frac{dQ}{T} = 0, \qquad \text{(Clausius 12)}$$

is the analytical expression of the second fundamental theorem in the mechanical theory of heat [14].

Although his analytical expression, Equation (Clausius 12), already looks like the one on the lhs of Equation (4), in Clausius' formulation [14] $f(T) = 1/T$ was still assumed to be an unknown function with values of $T_1, T_2, \ldots$ corresponding to the temperatures $t_1, t_2, \ldots$. Only in his second publication [15] of that year entitled by *On the Application of the Mechanical Theory of Heat to the Steam Engine* he explicitly states that

$$T = 273 + t, \qquad \text{(Clausius 13)}$$

and calls T the **absolute temperature**, a term, which became customary since then. For matters of curiosity, in Figure 4 Clausius' concept of a steam engine is displayed.

It is interesting to note that Joule and Thomson, who in the mean time got interested in the *œconomical production of mechanical effect from chemical forces* [16] *and the heat produced by an electric discharge* [17], respectively, rushed to comment on Clausius' paper [15], trying to point out their own achievements: *The results we* [Joule and Thomson] *have arrived at in our joint investigation have tended certainly to develope our views, and to give a more definite knowledge of the constitution of elastic fluids, but they do not contradict our original statements, published before the appearance of Prof. Clausius's papers.*[18]

7. Maxwell and the dynamical theory

Around 1850 or even much before first ideas came up to relate the pressure in gases and consequently also the heat to the motions of particles, considered as "hard spheres" or "elastic bodies". In the Philosophical Magazine, the stage was first set by Clausius, who – very unusual for his approach to physics – first tried to describe [19] the new set of problems in words and only then attempted to propose [20] an equation for the mean length of the path of a "molecule". It is hard to believe, but in 1860 Maxwell (James Clerk, 1831–1879) solved most problems related to the motion of elastic bodies in just one longish paper [21].

So many of the properties of matter, especially when in the gaseous form, can be deduced from the hypothesis that their minute parts are in rapid motion, the velocity increasing with the temperature, that the precise nature of this motion becomes a subject of rational curiosity. … certain phenomena, such as the internal friction of gases, the conduction of heat through a gas, and the diffusion of one gas through another, seem to indicate the possibility of determining accurately the mean length of path which a particle describes

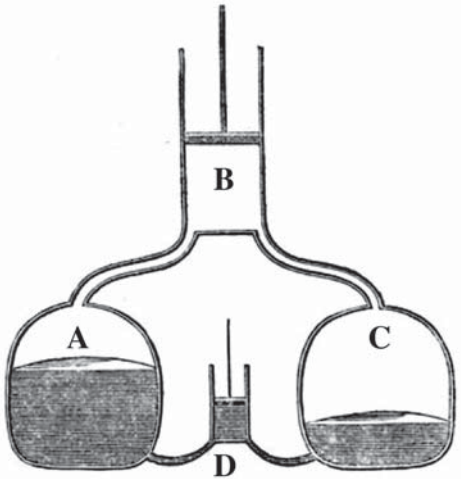


Figure 4. Clausius's sketch of a steam energine: *In the … figure … A represents the boiler whose contents are maintained by the source of heat at a constant temperature* T_1. *Apart of the C steam passes from the boiler to the cylinder B and raises the piston a certain height. The cylinder and boiler are next disconnected, and the vapour contained in the former raises the piston still higher by its own expansion. After this the cylinder is put in communication with the space C, which shall represent the condenser. We shall suppose the latter to be kept cold by external cooling …The constant temperature of the condenser shall be* T_o. *During the connexion of the cylinder with the condenser the piston retraces the whole of its former path, and thus all the vapour which did not immediately pass by itself into the condenser is driven into it, and there becomes condensed. In order to complete the cycle of operations, it is now necessary to convey the liquid produced by condensation back again into the boiler. This is done by means of the small pump D …* [15].

between two successive collisions. In order to lay the foundation of such investigations on strict mechanical principles, I shall demonstrate the laws of motion of an indefinite number of small, hard, and perfectly elastic spheres acting on one another only during impact. [21]

Retrospectively seen, Maxwell's paper [21] was probably of considerable importance for all future developments yet to come (including all of Maxwell's own improvements). It consists of three parts, organized as a sequence of "Propositions", see also Appendix A, each of them followed by a (at the time) complete mathematical treatment. The paper reads almost like a publication in pure mathematics: the "Propositions" being called there theorems, their formal handling corresponds to respective proofs; each "Proposition" follows logically from the previous one(s). As it is not possible to reproduce Maxwell's 31 pages of solid mathematics in a few lines, the reader is referred to Appendix B, which contains a kind of summary of the so-called kinetic theory of gases and of the Boltzmann–Maxwell distribution of velocities. Appendix A is meant to give at least an impression of the precise language Maxwell used.

8. Another flash forward

From Maxwell's paper [21] follows, e.g. that the averaged kinetic energy $\overline{\epsilon}_t$ of a gas molecule (particle) of mass m is given by

$$\overline{\epsilon}_t = \frac{m}{2}\overline{u}^2, \tag{8}$$

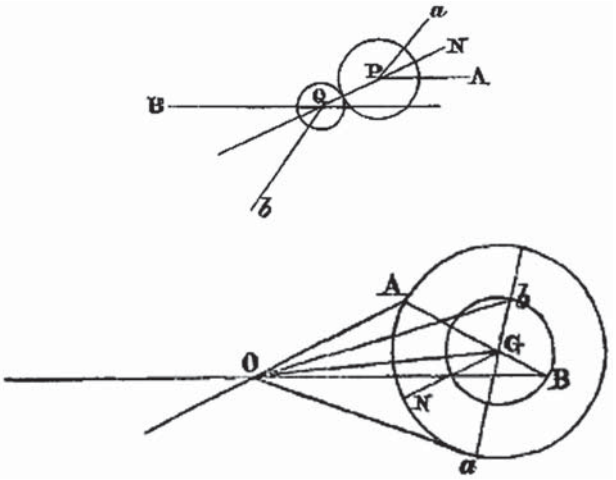


Figure 5. Some of Maxwell's original illustrations [21]. **Proposition I** (top): *Let P and Q be the position of the centres at impact; AP, BQ the directions and magnitudes of the velocities before impact; Pa, Qb the same after impact; then, resolving the velocities parallel and perpendicular to PQ the line of centres, we find that the velocities parallel to the line of centres are exactly reversed, while those perpendicular to that line are unchanged. Compounding these velocities again, we find that the velocity of each ball is the same before and after impact, and that the directions before and after impact lie in the same plane with the line of centres, and make equal angles with it.* **Proposition III** (bottom): *Let 0A, 0B represent the velocities before impact, so that if there had been no action between the bodies they would have been at A and B at the end of a second. Join A B, and let G be their centre of gravity, the position of which is not affected by their mutual action. Draw G N parallel to the line of centres at impact (not necessarily in the plane AOB). Draw a Gb in the plane AGN, making NGa = NGA, and Ga − GA and Gb = GB; then by Prop. I. Ga and Gb will be the velocities relative to G; and compounding these with OG, we have Oa and Ob for the true velocities after impact.*

where $\overline{u}^2$ denotes the averaged square velocity. The pressure p for an isotropic velocity distribution can then be written as

$$p = \frac{2}{3}\frac{N}{v}\overline{\epsilon}_t, \tag{9}$$

where N is the total number of particles and v the volume. Alternatively formulated,

$$pv = \frac{2}{3}N\overline{\epsilon}_t, \tag{10}$$

and compared with the equation for an ideal gas,

$$pv = nRT, \tag{11}$$

one immediately can deduce that for $n = 1$ and N referring to the Loschmidt (Avogadro) number L

$$\overline{\epsilon}_t = \frac{3}{2}\left(\frac{R}{L}\right)T = \frac{3}{2}kT, \tag{12}$$

or,

$$T = \frac{2}{3}k^{-1}\overline{\epsilon}_t, \tag{13}$$

where k is the Boltzmann constant.

The only one who immediately understood the importance of Maxwell's paper was Clausius [22]. In 1862, he published a long paper [23], in which he explicitly cited some of Maxwell's propositions. In dealing with problems similar to those of Maxwell, in some cases he found different results. In a much later publication, Maxwell in fact admitted [33] that *there were several errors in my theory of the conduction of heat in gases, which M. Clausius has pointed out in a eleborate memoir* [quoted as Ref. [23]] *on that subject.*

9. Scientific clashes: the first law of thermodynamics

Between 1863 and 1864, a formidable verbal battle between Tyndall (John, 1820–1893) on the one side and Thomson and Tait (Peter Guthrie, † 1901) on the other side was waged in the Philosophical Magazine *concerning the credit of being the first to establish in all its generality the principle of the Conservation of Energy*. By referring to an article in Liebig's *Annals of Chemistry and Pharmacy* that appeared in 1842, Tyndall claimed it was Mayer (Julius Robert von, 1814–1878), while the others insisted on Joule being the originator. The style of their contributions to the Philosophical Magazine probably can at best be called only "less than friendly", which, as should be known, is not unique to the nineteenth century.

Tyndall: ... *I was not aware of the existence of an article in the October Number of a journal called "Good Words", from the combined pens of … William Thomson and Tait, in which, though not mentioned by name, I am referred to … . When, however, it is known that the other articles in the number to which I refer, bear such titles as "The Childhood of Jesus", … &e., I think I may be excused if an article on Energy, in the scientific sense of the term, imbedded in such matter as those titles indicate, escaped my attention.*

Many will agree with me in thinking that it would hardly conduce to the interests or the dignity of science, if the habit were to become general of taking difficult and disputed points, which apparently involve imputations on individual character, into such a court as that chosen by Professor Thomson and Professor Tait. It is very laudable and very desirable that men in their high positions should instruct the readers of "Good Words" but these respectable persons are placed in a false position when they are virtually called upon to decide between the rival claims of Joule and Mayer, and to form an opinion as to the scientific morality of myself. [24]

This dispute made **Mayer** to produce a friendly, philosophically minded account of his achievements [25] by pointing out that his 1842 article even got translated into English and had appeared in 1862 [26].

Joule: *…but it is certain that the result of it as stated by M. Mayer, was not received or remembered by the scientific world, so that M. Regnault and myself were considered by others, as we considered ourselves, to be original inquirers on this subject. … This, however, so far from detracting from the merit of M. Mayer, must, on the contrary, increase our admiration of his wonderful prescience … [27].*

Tait: *I wish to make a few additional remarks on this subject: especially as it appears to me, after a careful perusal of all that Dr. Tyndall has written upon it, that he does not yet quite understand the points which Prof. Thomson and I wished to establish when, in consequence of his lecture on "Force," we wrote our article in "Good Words" nearly two years ago. Hence, so far as regards the question of heat alone, Mayer has no title to the position Dr. Tyndall claims for him. He did no more than repeat what Davy and Rumford had done better; and he has never, so far as even Dr. Tyndall's partisanship can show, attempted anything of the nature of either the theoretical or experimental developments which have*

advanced thermo-dynamic science during the present century. What, in point of fact, did Mayer do in thermo-dynamics? [28]

It seems that this dispute was a bit biased by national bride. One has to admit that to use the word "prescience" can indeed be regarded as an insult. Speaking strictly historically, Mayer surely was "the first one" to observe a conservation law for energy only he did not get the same scientific response (and audience) as Joule.

10. At last: the term entropy is introduced

Clausius knew that something in his description of "thermodynamic" processes was still missing. In 1862, he returned [29] to his *analytical expression of the second fundamental theorem* [14], see Equation (Clausius 12), and *stated that the relation*

$$\int \frac{dQ}{T} \geq 0 \qquad \text{(Clausius 14)}$$

must hold good for every circular process which is in any way possible [29]. Although this relation is an important statement *per se*, implying severe consequences for future developments, it was, however, only in1866 that he introduced [30] a notation and a terminology for dQ/T, which presently is still in use.

In [30] he starts out again with the first fundamental equation of thermodynamics, namely

$$\begin{aligned} dQ &= dU + dw, \\ dw &= AdW, \end{aligned} \qquad \text{(Clausius 15)}$$

where *A denotes the thermal equivalent of work …, d Q the quantity of heat* [infinitely] *absorbed and dW the external work simultaneously performed*. Since *according to W. Thomson, who subsequently introduced this same magnitude into his equations, has proposed for it* [*U*] *the convenient name of Energy, I have adopted this expression.* [30]

Immediately from there he addresses the second fundamental theorem in the form of his Equation (Clausius 12) by coining a new term S,

$$\frac{dQ}{T} = dS, \qquad \text{(Clausius 16)}$$

which he calls entropy. *I have formed, namely, from the Greek word* $\tau\rho o\pi\eta$*, change, the word* ***Entropy****, which expresses the meaning of the magnitude S, in the same way as the word Energy denotes that of the magnitude U.* [30] [en + tropein, Greek, meaning "transformation content" ("Verwandlungsinhalt")]

Combining Equation (Clausius 15) with Equation (Clausius 16) yields

$$-dw = dU - TdS, \qquad \text{(Clausius 17)}$$

which formally already looks like the very first equation in the introduction.

Suppose now that U, S and w are functions of the independent variables x and y, then, for example,

$$dw = \left(\frac{dw}{dx}\right) dx + \left(\frac{dw}{dy}\right) dy,$$

and thus the differentials dw, dU and dS can be formulated in terms of respective partial derivatives. Equation (Clausius 17) can indeed be regarded as one of the key equations of phenomenological thermodynamics, since at a given temperature T all occurring changes are caused only by changes in the energy U and the entropy S.

It should be mentioned that the term entropy seems to have been quickly accepted by the scientific community, as can be judged, e.g. from a paper [31] by Rankine (William John Macquorn, 1820–1872), one of the major actors in the amazing history of thermodynamics.

11. The second law of thermodynamics

The second law of thermodynamics remained very much under discussion. In particular, 1868 must have been an inspiring year, since not only Clausius published [32] a by then incredible theorem, see Equation (Clausius 14), but also Maxwell surprised the scientific community with a modification of his *dynamical theory of gases* considering *the molecules of a gas, not as elastic spheres of definite radius, but as small bodies or groups of smaller molecules repelling one another with a force whose direction always passes very nearly through the centres* of gravity of the molecules*, and whose magnitude is represented very nearly by some function of the distance of the centres of gravity* [33].

Twelve years after his first ideas [14] about the additive character of the entropy, see Equation (Clausius 11), and a subsequent proof of its positive definite character, Equation (Clausius 14), Clausius could finally give a very general interpretation of this quantity, which turned out to become one of the corner stones of physics:

Hence we must conclude that in all the phenomena of nature the total entropy must be ever on the increase and can never decrease; and we thus get as a short expression for the process of transformation which is everywhere unceasingly going on the following theorem:

The entropy of the universe tends towards a maximum

The more the universe approaches this limiting condition in which the entropy is a maximum, the more do the occasions of further changes diminish; and supposing this condition to be at last completely attained, no further change could evermore take place, and the universe would be in a state of unchanging death [32].

12. Hamilton and thermodynamics

By around 1870, the (phenomenological) equation reflecting both, the first and the second fundamental theorem of thermodynamics, seemed to be clear and generally accepted. However, the meaning of Joule's mechanical equivalent of heat remained unresolved. Within 10–15 years, a series of papers [34–45], appeared in the Philosophical Magazine, in which such a connection was proposed. This "mechanical equivalent" seems to have been almost an obsession and turned out to be a battle in classical mechanics. The main actors on the stage were Boltzmann (Ludwig, 1844–1906), Clausius and Szily (Coloman), being joined a couple of years later by Burbury (Samuel Hawksley, 1831–1911).

Since quite a few new concepts were introduced in these publications, it is worthwhile to summarize briefly some of them. The first one to embark in hitherto unknown fields was (again) Clausius.

Let there be any system whatever of material points in stationary motion. By stationary motion I mean one in which the points do not continually remove further and further from

their original position, and the velocities do not alter continuously in the same direction … [34]. Then:

$$\underbrace{\frac{1}{2}\sum_i m_i \overline{v}^2}_{\text{mean “vis viva”}} = \underbrace{-\frac{1}{2}\sum_i \overline{(X_i x_i + Y_i y_i + Z_i z_i)}}_{\text{“virial”}} \qquad \text{(Clausius 18)}$$

$$= \underbrace{\frac{1}{2}\sum_i r\phi_i(r)}_{\text{“ergal”}}. \qquad \text{(Clausius 19)}$$

In here $\overline{v}^2$ is the mean square velocity at a particular time t, (x_i, y_i, z_i) are the coordinates of materials points of mass m_i and X_i, Y_i and Z_i are the components of the forces acting on these material points. By the way, the term “*vis viva*” (latin, meaning living force)

$$\frac{1}{2}\sum_i m_i v_i^2 = \frac{1}{2}\sum_i \left(\left(\frac{dx_i}{dt}\right)^2 + \left(\frac{dy_i}{dt}\right)^2 + \left(\frac{dz_i}{dt}\right)^2\right)$$

was coined by Leibnitz (Gottfried Wilhelm, 1676–1689). For the term on the rhs of Equation (Clausius 18), *we will …give … the name of Virial of the system, from the Latin word vis (force)* [34].

For the rhs of Equation (Clausius 19) Clausius invented the name *ergal*, *from the Greek word* $\epsilon\rho\gamma o\upsilon$ *(work)* [35]. It is worthwhile to mention that the terms *vis viva*, *virial* and *ergal* were frequently used in this period of time, see for example [44].

Two important theorems are connected with the *ergal* and the *virial* [34]:

(1) the sum of the vis viva and the ergal is constant, and
(2) the mean vis via is equal to the virial,

which sound slightly mysterious unless “translated” into modern language. Theorem 1 states that the total energy is the sum of the kinetic and the potential energy, while, as already mentioned, Theorem 2 is nothing but the virial theorem[2].

For heat (h) – *the vis viva of the internal motions* [34] – one therefore obtains the expression,

$$h = \frac{1}{2}\sum_i r\phi_i(r) + \frac{3}{2}pv, \qquad \text{(Clausius 20)}$$

where as before p and v denote the pressure and the volume, respectively.

At last heat, the quantity that was so much seeked for since the beginning of the nineteenth century, got now a meaning, a proper theoretical description, namely as the sum of the (internal) potential energy (ergal) and the external work.

The quoted series of papers were aimed at proving, disproving, agreeing or disagreeing with the above claims. The main argument seems to have been whether or not Hamilton’s principles were correctly or if a Jacobi-type description of (classical) mechanics should have been applied. Szily disagreed [40,43,45], Boltzmann agreed mostly [44], while Clausius tried to generalize and prove his claims in detail [36–39,41,42]. This dispute, however, was not really surprising, since the mathematical treatments of mechanics by Hamilton (William Rowan, 1805–1865) and Jacobi (Jacques Simon, 1804–1851) were also rather very new: Clausius, e.g. was a contemporary of Hamilton and Jacobi.

For matters of curiosity, it should be mentioned that in 1875, in a paper [42] describing the motion of n material points along closed paths, Clausius changed his notation in response to the at that time ongoing discussions. Let $q_1, q_2, \ldots q_n$ be arbitrary variables. Then – according to him – the *ergal* (potential energy) U can be considered as a function of these variables

$$U = U(q_1, q_2, \ldots, q_n), \tag{Clausius 21}$$

and the *vis viva* (kinetic energy) is given by

$$2T = \sum_{i=1}^{n} p_i dq_i, \quad p_i = \frac{dT}{dq_i}. \tag{Clausius 22}$$

Note that the use of the capital letters U and T for the potential and kinetic energy, and p for the momentum became almost standard since then.

13. Boltzmann and his distribution of energies

The English speaking scientific community became aware of Boltzmann's contributions in the field of thermodynamics rather late. Early references such as [44,48,49] can at best be regarded only as abstracts or very short summaries. The first of his papers [52] that got translated and published in the Philosophical Magazine is from 1882. For this reason, use has to be made of the seemingly first paper [47] by a British scientist, whose treatise (*memoir*) relied completely on Boltzmann's 1871 reports to the Viennese Academy of Science [46], namely Burbury (Samuel Hawksley, 1831–1911).

Burbury sets very carefully the stage for his derivations [47]:

In a memoir by Boltzmann [46] *is contained an analytical proof of the second law of thermodynamics, founded on results obtained in the earlier part of the same memoir. The writer's treatment of this part of the subject is not characterized by his usual lucidity, and has given much trouble to some mathematicians. It is worth while to show that the second law can be deduced by a simple process from Boltzmann's results.*

Boltzmann has shown that the mean kinetic energy is the same for every particle. Let T be this mean kinetic energy, and let $T = 3/2h$. The mean kinetic energy of the system is then $3\lambda/2h$.

Let the positions of the particles at any instant be defined by reference to three rectangular axes passing through the centre of the vessel, the coordinates of the several particles being $x_1, y_1, z_1, \ldots x_\lambda, y_\lambda, z_\lambda$. Let v be the volume of the vessel.

The pressure …, if the motion be stationary, … must have on the average of any long time a mean value P per unit area ... .

Let χ be the ergal for any given positions of the particles of all the forces acting on the particles that is, of all the forces which act in any manner on the system … except the external force P.

If in the case of a stationary motion M denotes a particular configuration of positions such that

$$\begin{aligned}\chi(x_1, y_1, z_1, \ldots x_\lambda, y_\lambda, z_\lambda) \\ = \chi(x_1 + dx_1, y_1 + dy_1, z_1 + dz_1, \ldots x_\lambda + dx_\lambda, y_\lambda + dy_\lambda, z_\lambda + +dz_\lambda)\end{aligned}$$

then [47]*, as as Boltzmann has shown, the time during which on the average of any very long time the system will be within configuration M is proportional to*

$$e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda. \tag{Burbury 1}$$

Its actual value is therefore

$$\frac{e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda}{\int \cdots \int e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda}. \tag{Burbury 2}$$

The mean value of any function, f, of the position of the particles (which we denote as usual $\bar{f}$) is [47]

$$\bar{f} = \frac{\int \cdots \int f\, e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda}{\int \cdots \int e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda}, \tag{Burbury 3}$$

and therefore

$$\bar{\chi} = \frac{\int \cdots \int \chi\, e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda}{\int \cdots \int e^{-h_\chi} dx_1 dy_1 dz_1 \ldots dx_\lambda dy_\lambda dz_\lambda}. \tag{Burbury 4}$$

One has to admit, Boltzmann's ideas, as demonstrated by Burbury in the Philosophical Magazine, were a huge step forward as now a concept of probabilities was introduced. Calling the expression in Equation (Burbury 2) simply W, then quite obviously $0 \leq W \leq 1$.

In using this concept, Burbury managed to show that the total change in u is given by

$$\delta u = -\bar{\chi}\delta h + \lambda \frac{\delta v}{v} - h\overline{\frac{d\chi}{dv}}\delta v, \tag{Burbury 5}$$

and that

$$\frac{\delta Q}{T} = \delta\left\{-\lambda \log h + \frac{2}{3}h_{\bar{\chi}} + \frac{2}{3}u\right\} \tag{Burbury 6}$$

is an exact differential [47]. Note that the last two equations are still only of formal significance, since, e.g. λ, the total number of particles was – as usual for this period of time – assumed to be arbitrary.

In an attempt to compare his approach with that of Maxwell, Boltzmann added another interesting step [52] in 1882, namely when the number of particles becomes very large:

$$\frac{Z_1}{Z_2} = \frac{e^{-k/2K} dk}{k\sqrt{2\pi}} \quad , \quad K = \frac{E - V}{n}, \tag{Boltzmann 1}$$

where K is the mean value of the kinetic energy per particle, E being the total and V the potential energy. In the above expression, Z_1 is the number of configurations of n particles with coordinates $\{x_i + dx_i, y_i + dy_i, z_i + dz_i, i = 1, n\}$ and a kinetic energy corresponding to a momenta $k \leq p_q \leq k + dk, q \leq n$, while Z_2 is the number of configurations with the same specification for the coordinates, but arbitrary kinetic energy. Note that the conceptual step of considering a very large number of particles became important once it

was generally acknowledged that n was indeed a natural constant (Avogadro (Loschmidt) number).

14. Scientific clashes

14.1. *The kinetic theory of gases*

Interestingly enough, neither of Maxwell's two publications in 1860 and 1868 [21,33] raised any objections by British physicists at the time they appeared, it was only after Boltzmann came forward with his ideas about the kinetic theory of gases and his concept of probabilities that almost an opposition front formed. This front saw Tait, Thomson (Lord Kelvin) and to some extent Raleigh (John William Strutt, 3. Baron Rayleigh, 1842–1919) on the one side, and Burbury, Boltzmann and Clausius on the other side. The battle ground was mostly the Philosophical Magazine. This battle extended well into the twentieth century and most likely stopped only because internationally the scientific focus points had moved to Rutherford's radiation phenomena, Einstein's relativity theory and to quantum phenomena.

Of course in this dispute not the mathematical details brought forward are of interest (as is well known, they were settled in favour of Maxwell and Boltzmann, anyhow), but the wording, which was again less than friendly. In 1886, i.e. almost 20 years after Maxwell's second article [33], Tait's first attack appeared [53], followed immediately by answer by Burbury [54].

Tait: *In this Magazine for 1860 Clerk-Maxwell published his papers on the "Collision of Elastic Spheres," ... .In the ardour of his research, Maxwell here and there contented himself with very incomplete proofs (we can scarcely call them more than illustrations) of some of the most important of his results. This is specially the case with the investigation of the law of ultimate partition of energy in a mixture of smooth spherical particles of two different kinds.*

If it be shown ... that certain preliminary conditions are absolutely necessary to the proof even of Maxwell's Theorem, and if these cannot be granted in the more general case treated by Boltzmann, it is clear that Boltzmann's Theorem must be abandoned [53].

Burbury: *I think Professor Tait has postulated rather more than is necessary. ...*

I think it follows [from the proofs he gives in this paper] *that collisions tend to bring about Maxwell's distribution* [54].

Tait: *Mr. Burbury now ... raises an objection, which I do not understand, to an assumption made (after Clerk-Maxwell and Clausius) in my second paper. This objection is based upon a Theorem which Mr. Burbury supposes to have been established by Prof. Boltzmann.*

If Mr. Burbury can show that solution [Tait's view [55] on the kinetic theory of gases] *to be erroneous, he may possibly make out a presumptive case against the assumption on which it was founded :–but not otherwise* [58].

14.2. *Avogadro's law*

Avogadro (Amadeo, 1776–1858) was an Italian physicist, whose ideas [50] about gases were only very slowly accepted by the scientific community. His law,

$$v = nk, \qquad \text{(Avogadro)}$$

namely that the volume v of a gas is proportional (k) to the "amount of substance" n in the gas, only became prominent once Loschmidt (Johann Joseph, 1821–1895) managed in 1866

to actually evaluate the number of "gas molecules" per volume. This number, $N_L = N_A =$ 2,686 7805 [10^{25} m^{-3}], is usually nowadays termed Avogadro number, sometimes also Loschmidt constant.

Since Tait in his 1886 paper stated that *in seeking for a proof of Maxwell's Theorem, I found it absolutely essential to the application of the statistical method to premise : (A) That the particles of the two systems are thoroughly mixed. (B) That, in any region containing a very large number of particles, the particles of each kind acquire and maintain the error-law distribution of velocities* [53], he had of course also touched the formal problem behind Avogadro's law.

This was exactly the point Boltzmann picked up in a 29 (!) page-long article [56] entitled by *On the Assumptions necessary for the Theoretical Proof of Avogadro's Law,* in which he addresses Tait directly: *Professor Tait, by making a number of special assumptions, has given a very exact proof of Avogadro's Law – or, rather, of the proposition that in the case of the heat-equilibrium of two mixed gases the mean potential energy of the molecules must be equal. Prof. Tait does not appear to me, however, to have given the least proof that the special assumptions which he makes as the basis of this proposition are necessary, or even that the more general propositions stated by myself and Maxwell are incorrect.* [56]

In this paper, Boltzmann not only treated in every detail the case of two impinging elastic spheres of mass m and M and velocity v and V, see Figure 6, but also tried to generalize the problem by considering the time evolution of the velocity distributions and – **most importantly** – of their products[3]

$$\frac{\partial f(v,t)}{\partial t}, \quad \frac{\partial F(V,t)}{\partial t} \text{ and } \frac{\partial f(v',t)}{\partial t}, \quad \frac{\partial F(V',t)}{\partial t}, \qquad \text{(Boltzmann 2)}$$

before and after the collision.

Perhaps besides proofing that Burbury's answer [54] to Tait [53] was correct, the main aspect of his paper is that he uses the entropy, denoted E by him,

$$E = \int_0^\infty f(v,t)(\ln f(v,t) - 1)v^2 dv + \int_0^\infty F(V,t)(\ln F(V,t) - 1)V^2 dV, \quad \text{(Boltzmann 3)}$$

its time evolution

$$\frac{dE}{dt} = \int_0^\infty \ln f(v,t)\frac{\partial f(v,t)}{\partial t} v^2 dv + \int_0^\infty \ln F(V,t)\frac{\partial F(V,t)}{\partial t} V^2 dV, \qquad \text{(Boltzmann 4)}$$

and the condition that (in equilibrium) dE/dt must vanish.

Tait's reply [57] to Boltzmann's paper on the Avogadro Law was almost rude:

Prof. Boltzmann says that I do not expressly state that my work applies only to hard spheres. This is an absolutely unwarrantable charge, as I have taken most especial care throughout to make this very point clear.

Prof. Boltzmann, while objecting to my remark about "playing with symbols," has unwittingly furnished a very striking illustration of its aptness. His paper bristles throughout with formulae, not one of which has the slightest direct bearing on the special question he has raised! …

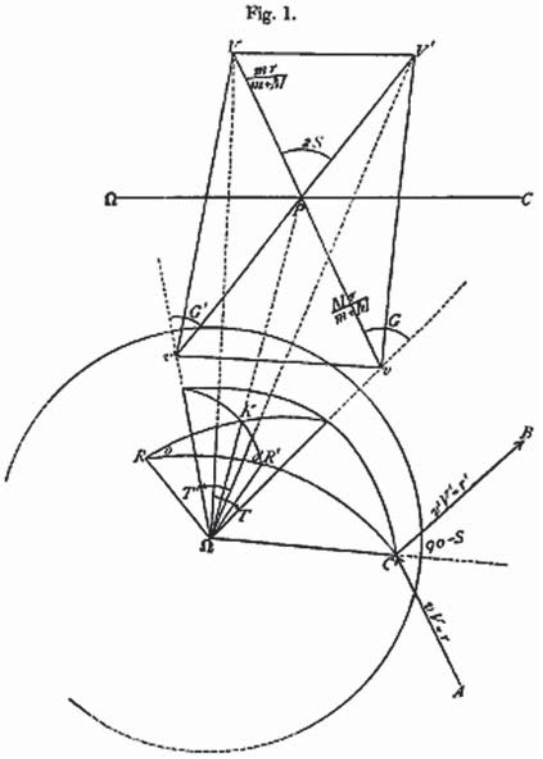


Figure 6. Boltzmann's Figure 1: *Let two elastic spheres (molecules) of masses m and M impinge upon each other. Let* $v = \Omega v$ *and* $V = \Omega V$ *be their velocities before impact (see Figure 1), and* $v\prime = \Omega v'$ *and* $V\prime = \Omega V'$ *their velocities after impact,* δ *the sum of radii of the spheres,* $r = vV$ *their relative velocity before impact, and suppose their relative velocity after impact to have the direction* $\prime = v\prime V'$. *Let* ΩC *be the line of centre of the spheres at the moment of impact. Let* $\measuredangle vV = T$, $\measuredangle vr = G$, $\measuredangle rr' = 2S$, $\measuredangle r, \Omega C = 90° + S$, $\measuredangle v\prime V\prime = T\prime$, $\measuredangle v\prime r\prime = G\prime$. *Let O be the angle of the planes r, v and r,* ΩC; *and* O' *the angle of the planes* r', v',*and* r', ΩC. *... For the sake of clearness I include in fig. 1 those lines of fig. 2 of my Theory of Gas Diffusion,..., which are here necessary.* [56] ($\sphericalangle ab$ means angle between a and b).

I need only say that Prof. Boltzmann, while causelessly attributing to me a silly mathematical mistake, has evidently overlooked the special importance which I attach to the assumed steadiness of the "average behaviour of the various groups of a community." [57]

Needless to say that Burbury replied [60] almost immediately to Tait, who in turn answered [61] quickly. In 1890, Burbury summarized [62] the whole dispute following essentially Boltzmann's arguments, but using substantial more transparent mathematics than Boltzmann.

Perhaps the final and most decisive support for Maxwell and Boltzmann came from Raleigh in an article [63] cleverly entitled by *Remarks on Maxwells's Investigations, respecting Botzmann's theorem.*

Having found Maxwell's demonstration of the fundamental theorem

$$dq'_1 \dots dq'_n dp'_1 \dots dp'_n = dq_1 \dots dq_n dp_1 \dots dp_n \qquad \text{(Raleigh 1)}$$

difficult to follow, I have sought to simplify it by an arrangement such that the initial and final times $t\prime$ and t may be considered as absolutely fixed throughout the discussion. [63]

In here the $\{q_s\}$ and $\{p_s\}$ are the coordinates and momenta of the particles at a given time t, the corresponding dashed quantities at t'. Making now use [63] of *Hamilton's "principal function" S*,

$$\delta S = \sum_{i=1}^{n} p_i dq_i - \sum_{i=1}^{n} p'_i dq'_i, \qquad \text{(Raleigh 2)}$$

where all the motions contemplated are unconstrained, and occupy the fixed time $t - t'$. The total energy … is variable from one motion to another, and S is to be regarded as a function of the $q's$ and $(q')\prime s$.
The initial and final momenta are thus expressed by means of S in the form

$$p'_r = -\frac{dS}{dq'_r}, \quad p_r = -\frac{dS}{dq_r}, \qquad \text{(Raleigh 3)}$$

so that

$$\frac{dp'_r}{dq_s} = -\frac{d^2 S}{dq'_r dq_s} = -\frac{dp_s}{dq'_r}, \qquad \text{(Raleigh 4)}$$

Thus, using S with $t - t\prime$ constant … we get [63]

$$dq'_1 \ldots dq'_n dp'_1 \ldots dp'_n \qquad \text{(Raleigh 5)}$$
$$= dq'_1 \ldots dq'_n dq_1 \ldots dq_n \det \begin{pmatrix} \frac{dp'_1}{dq_1} & \cdots & \frac{dp'_n}{dq_1} \\ \vdots & & \vdots \\ \frac{dp'_1}{dq_n} & & \frac{dp'_n}{dq_n} \end{pmatrix},$$

$$dq_1 \ldots dq_n dp_1 \ldots dp_n \qquad \text{(Raleigh 6)}$$
$$= dq'_1 \ldots dq'_n dq_1 \ldots dq_n \det \begin{pmatrix} \frac{dp_1}{dq'_1} & \cdots & \frac{dp_n}{dq'_1} \\ \vdots & & \vdots \\ \frac{dp_1}{dq'_n} & & \frac{dp_n}{dq'_n} \end{pmatrix}.$$

Since both determinants are equal, Raleigh gave a very elegant proof[4] of Equation (Raleigh 1), i.e. of Maxwell's important result.

One would have thought that after Raleigh's perfectly clear proof no more doubts would be raised with respect to Maxwell's approach. In vain, Kelvin was not convinced at all. In a short communication [64] entitled by *On a Decisive Test-case disproving the Maxwell–Boltzmann Doctrine regarding Distribution of Kinetic energy*, he proposed a situation (a test case) in which the average kinetic energy would not be uniform.

This conjecture provoked Boltzmann to reformulate [65] Maxwell's theory (*I can hardly doubt that Lord Kelvin will be satisfied with this result of his test-case* [65]) and his own formulation for the entropy in general coordinates. At least in the view of Boltzmann, within the mathematical language he had applied, there was nothing more to be added. Ref. [65], by the way, turned out to be his last paper in the Philosophical Magazine. From there on, it was essentially Burbury that carried the "probability torch" [66–76], seconded to some extent by Natanson with a *kinetic interpretation of the dissipation function* [77]. The last of

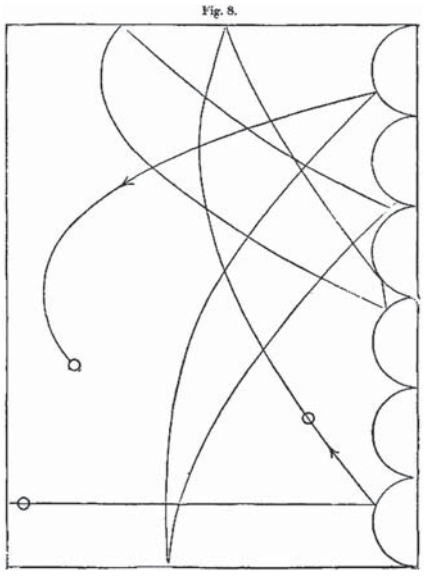


Figure 7. Lord Kelvin's Figure 8 in [78], showing the trajectory of a single particle in a box with one rippled wall.

Burbury's papers, *Boltzmann's Law of Probability* e^{-h_x} [76], perhaps gave rise to use the term "The H theorem" or "Boltzmann's H theorem".

15. Nineteenth-century clouds over the dynamical theory of heat

The beauty and clearness of the dynamical theory, which asserts heat and light to be modes of motion, is at present obscured by two clouds. (I) The first came into existence with the undulatory theory of light, and was dealt with by Fresnel and Dr. Thomas Young … . (II) The second is the Maxwell–Boltzmann doctrine regarding the partition of energy. …

Maxwell's 1860 proof has always seemed to me quite inconclusive, and many times I urged my colleague, Professor Tait, to enter on the subject. This he did, and in 1886 he communicated … a paper on the foundations of the kinetic theory of gases, which contained a critical examination of Maxwell's 1860 paper, highly appreciative of the great originality and splendid value, for the kinetic theory of gases, of the ideas and principles set forth in it; but showing that the demonstration of the theorem of the partition of energy in a mixed assemblage of particles of different masses was inconclusive, and successfully substituting for it a conclusive demonstration. …

With magnificent boldness Boltzmann and Maxwell declared that the ratio [between the averaged kinetic energy of rotation and that of translation] *is equality; Boltzmann having found what seemed to him a demonstration of this remarkable proposition, and Maxwell having accepted the supposed demonstration as valid. …*

The simplest way of arriving at this desired result [some escape from the destructive simplicity of the general conclusion] *is to deny the conclusion; and so, in the beginning of the twentieth century, to lose sight of a cloud which has obscured the brilliance of the molecular theory of heat and light during the last quarter of the nineteenth century.* [78]

Table 1. Units or constants related to thermodynamics.

Property or constant	Unit	Name
Heat, work	[J]	Joule
Energy transfer	[W] = [J][s]$^{-1}$	Watt
Temperature	[K]	Kelvin
k, k_B	$1.3806488(13)\times10^{-23}$ [J][K]$^{-1}$	Boltzmann constant
N_A, N_L	2,686 7805 [10^{25} m^{-3}]	Avogadro number Loschmidt constant
Entropy	1 [Cl] = 4.1868 [J][K]$^{-1}$	Clausius, no longer in use

In 1901, Kelvin summarized in a very long paper (42 pages!) [78] his disgust about the direction theoretical molecular dynamics had taken. By repeating or setting up new cases, in which Maxwell's and/or Boltzmann's approach would fail, and by compiling various critical citations from other authors, he seemingly looked forward to the moment the dark clouds over "thermodynamics" would lift.

Figure 7, in which one of his "special" cases, namely the trajectory of a single particle in "box" with one rippled wall, is reproduced, illustrates best his fear that in future simplicity will be lost in a theoretical concept based only on probabilities. If Maxwell and Boltzmann were right, then all the exotic cases – the one displayed in Figure 7 was just one of them – would become meaningless.

16. Summary

Kelvin's "farewell" paper [78] to molecular dynamics based on classical mechanics appeared almost 100 years after Gay-Lussac had published his law. As is well known, some physical units or universal constants, see Table 1, are named in honour of famous scientists in the field of thermodynamics.

Like many others, Kelvin [78] witnessed at the arrival of the twentieth century the end of an era in physics, which was dominated by the beauty of classical mechanics, by the fascination of being able to describe the movements of (unspecified) particles of some mass *m* entirely in terms of differential equations in a monocausal manner [79]. A new type of physics was about to take over: radiation phenomena, quantum and relativity effects, and probability concepts started to dominate all areas of interest. The famous Boltzmann statistics, for example, that was so much disputed, turned out to be a limiting case of the Bose–Einstein statistics (1924), see also Appendix C. There is a lot more to be said as far as the history of thermodynamics is concerned in the twentieth century: elastic hard spheres became "real" atoms, molecules were represented by oscillators, etc. What remained from the nineteenth-century approaches is the mathematical formulation of phenomenological thermodynamics, see Equations (1)–(3), which is still of great importance in technology.

However, there is one additional point that has to be made. In viewing the conceptual history of thermodynamics from 1801 till 1910 as published in the Philosophical Magazine, there is not the slightest trace of the political and social changes that took place during the nineteenth century, neither in England [80,81] nor in continental Europe. Perhaps, Kelvin regretted also something else, namely the end of a Victorian type of society. The slow social

changes in the last decades of the nineteenth century caused by the revolution in 1848, the liberation of up-to-then politically marginal (and rightless) strata of the population, women, jews, etc., had changed society fundamentally. Kelvin most likely felt that the age of the well-to-do, of Victorian gentlemen being exclusively in charge of science, had ended. The *agoras* of the twentieth century reflected a completely different social structure.

Notes

1. all quotations are shown in italics.
2. Virial Theorem:
$$2\langle T\rangle = -\sum_{i=1}^{N}\langle \mathbf{F}_i \cdot \mathbf{r}_i\rangle$$
where $\langle T\rangle$ is the time average over the total kinetic energy of a stable system consisting of N particels located at $\mathbf{r}_i$ and the $\mathbf{F}_i$ are the forces acting on each of the particles.
3. The product AB of two incidences A and B is an incidence AB where A as well as B occur. If the incidences A and B are "statistically" independent, i.e. the probability of A is independent of the probability of B, then
$$p(AB) = p(A)p(B) \quad .$$
Similar definitions apply for probability densities.
4. Taking for matters of simplicityin Equation (Raleigh 4) $r = s = i$ and using Clausius expression for the kinetic energy, Equation (Clausius 22), is probably the best way to illustrate Raleigh's proof:
$$\begin{aligned} T(t) &= \sum T_i(t), \\ T_i(t) &= T_i = (p_i + dp_i)(q_i + dq_i), \\ T_i(t') &= T_i' = \left(p_i' + dp_i'\right)\left(q_i' + d'q_i\right), \end{aligned}$$
$$\begin{aligned} T_i' - T_i = \Big[& p_i'\left(q_i' + dq_i'\right) - p_i\left(q_i + dq_i\right) \\ & + \underbrace{\left(p_i' dq_i' - p_i dq_i\right)}_{\delta S_i\text{, see Eq.(Raleigh 2)}} + \underbrace{\left(dp_i' dq_i' - dp_i dq_i\right)}_{=0\text{ (Maxwell)}}\Big]. \end{aligned}$$

References

[1] S. Carnot, (1824). Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance, Bachelier, Paris, 1824. Available at http://books.google.com/books?id=YcY9AAAAMAAJ
[2] J.P. Joule, Phil. Mag. Series 3 20 (1842) p.98.
[3] J.P. Joule, Phil. Mag. Series 3 20 (1843) p.204.
[4] J. P. Joule, Phil. Mag. Series 3 23 (1843) p. 263; Phil. Mag. Series 3 23 (1843) p.347; Phil. Mag. Series 3 23 (1843) p.435.
[5] A.-T. Petit and P.-L. Dulong, Annales de Chimie et de Physique 10 (1819) p.395.
[6] J.P. Joule, Phil. Mag. Series 3 25 (1844) p.334.
[7] J.P. Joule, Phil. Mag. Series 3 26 (1845) p.369.
[8] J.P. Joule, Phil. Mag. Series 3 27 (1845) p.205.
[9] J.P. Joule, Phil. Mag. Series 3 31 (1847) p.173.
[10] R. Clausius, Phil. Mag. Series 4 1 (1851) p.398.
[11] R. Clausius, Phil. Mag. Series 4 2 (1851) p. 1; Phil. Mag. Series 4 2 (1851) p.102.
[12] R. Clausius, Phil. Mag. Series 4 2 (1851) p.483.

[13] W. Thomson, Phil. Mag. Series 4 4 (1852) p. 8; Phil. Mag. Series 4 4 (1852) p.105; Phil. Mag. Series 4 4 (1852) p.168; Phil. Mag. Series 4 4 (1852) p.424.
[14] R. Clausius, Phil. Mag. Series 4 12 (1856) p.81.
[15] R. Clausius, Phil. Mag. Series 4 12 (1856) p. 241; Phil. Mag. Series 4 12 (1856) p.338.
[16] J.P. Joule, Phil. Mag. Series 4 5 (1853) p.1.
[17] W. Thomson, Phil. Mag. Series 4 7 (1854) p.347.
[18] J.P. Joule, Phil. Mag. Series 4 12 (1856) p.385.
[19] R. Clausius, Phil. Mag. Series 4 14 (1857) p.108.
[20] R. Clausius, Phil. Mag. Series 4 17 (1859) p.81.
[21] J.C. Maxwell, Phil. Mag. Series 4 19 (1860) p. 19; Phil. Mag. Series 4 20 (1860) p.21. See also his 1868 paper [33].
[22] R. Clausius, Phil. Mag. Series 4 19 (1860) p.434.
[23] R. Clausius, Phil. Mag. Series 4 23 (1862) p. 417; Phil. Mag. Series 4 23 (1862) p.512.
[24] J. Tyndall, Phil. Mag. Series 4 25 (1863) p. 220; Phil. Mag. Series 4 25 (1863) p.368.
[25] J.R. Mayer, Phil. Mag. Series 4 25 (1863) p.493.
[26] J.R. Mayer, Phil. Mag. Series 4 24 (1862) p.371.
[27] J.P. Joule, Phil. Mag. Series 4 26 (1863) p.145.
[28] P.G. Tait, Phil. Mag. Series 4 28 (1864) p. 288.
[29] R. Clausius, Phil. Mag. Series 4 24 (1862) p. 81; Phil. Mag. Series 4 24 (1862) p.201.
[30] R. Clausius, Phil. Mag. Series 4 32 (1866) p.1.
[31] W.J.M. Rankine, Phil. Mag. Series 4 30 (1865) p.241.
[32] R. Clausius, Phil. Mag. Series 4 35 (1868) p.405.
[33] J.C. Maxwell, Phil. Mag. Series 4 35 (1868) p. 129; Phil. Mag. Series 4 35 (1868) p.185.
[34] R. Clausius, Phil. Mag. Series 4 40 (1870) p.122.
[35] The unit "erg" is derived from the same Greek word, 1[erg] = 10^{-7}[J].
[36] R. Clausius, Phil. Mag. Series 4 42 (1871) p.161.
[37] R. Clausius, Phil. Mag. Series 4 44 (1872) p.365.
[38] R. Clausius, Phil. Mag. Series4 46 (1873) p.1.
[39] R. Clausius, Phil. Mag. Series 4 46 (1873) p. 236; Phil. Mag. Series 4 46 (1873) p.266.
[40] C. Szily, Phil. Mag. Series 4 46 (1873) p.426.
[41] R. Clausius, Phil. Mag. Series 4 48 (1874) p.1.
[42] R. Clausius, Phil. Mag. Series 4 50 (1875) p. 26; Phil. Mag. Series 4 50 (1875) p.101; Phil. Mag. Series 4 50 (1875) p.191.
[43] C. Szily, Phil. Mag. Series 5 1 (1876) p.22.
[44] L. Boltzmann, Phil. Mag. Series 4 50 (1875) p.495.
[45] C. Szily, Phil. Mag. Series 5 2 (1876) p.254.
[46] L. Boltzmann, Sitzungsberichte der kaiserlichen Akademie der Wissenschaften Wien 63 (1871) p. 670; Sitzungsberichte der Wiener Akademie der Wissenschaften 63 (1871) p.712.
[47] S.H. Burbury, Phil. Mag. Series 5 1 (1876) p.61.
[48] L. Boltzmann, Phil. Mag. Series 5 3 (1877) p.320.
[49] L. Boltzmann, Phil. Mag. Series 5 6 (1878) p.236.
[50] A. Avogadro, Journal de Physique, de Chimie, d'Histoire naturelle et des arts 73 (1811) p. 58.
[51] J.J. Loschmidt, Sitzungsberichte der kaiserlichen Akademie der Wissenschaften Wien 52 (1866) p.395.
[52] L. Boltzmann, Phil. Mag. Series 5 14 (1882) p.299.
[53] P.G. Tait, Phil. Mag. Series 5 21 (1886) p.343.
[54] S.H. Burbury, Phil. Mag. Series 5 21 (1886) p.481.
[55] P.G. Tait, Phil. Mag. Series 5 23 (1887) p.141.
[56] L. Boltzmann, Phil. Mag. Series 5 23 (1887) p.305.
[57] P.G. Tait, Phil. Mag. Series 5 23 (1887) p.433.
[58] P.G. Tait, Phil. Mag. Series 5 25 (1887) p.38.

[59] L. Boltzmann, Phil. Mag. Series 5 25 (1887) p.81.
[60] S.H. Burbury, Phil. Mag. Series 5 25 (1887) p.129.
[61] P.G. Tait, Phil. Mag. Series 5 25 (1887) p.172.
[62] H. Burbury, Phil. Mag. Series 5 30 (1890) p.298.
[63] Lord Rayleigh, Phil. Mag. Series 5 33 (1892) p.356.
[64] Lord Kelvin, Phil. Mag. Series 5 33 (1892) p.466.
[65] L. Boltzmann, Phil. Mag. Series 5 35 (1893) p.153.
[66] S.H. Burbury, Phil. Mag. Series 5 37 (1894) p.143.
[67] S.H. Burbury, Phil. Mag. Series 5 37 (1894) p.574.
[68] S.H. Burbury, Phil. Mag. Series 5 50 (1900) p.584.
[69] S.H. Burbury, Phil. Mag. Series 6 2 (1901) p.403.
[70] S.H. Burbury, Phil. Mag. Series 6 3 (1902) p.225.
[71] S.H. Burbury, Phil. Mag. Series 6 5 (1903) p.134.
[72] S.H. Burbury, Phil. Mag. Series 6 6 (1903) p.251.
[73] S.H. Burbury, Phil. Mag. Series 6 8 (1904) p.43.
[74] S.H. Burbury, Phil. Mag. Series 6 11 (1906) p.455.
[75] S.H. Burbury, Phil. Mag. Series 6 17 (1909) p.1.
[76] S.H. Burbury, Phil. Mag. Series 6 19 (1910) p.712.
[77] L. Natanson, Phil. Mag. Series 5 39 (1895) p. 455; Phil. Mag. Series 5 39 (1895) p.501.
[78] Lord Kelvin, Phil. Mag. Series 6 2 (1901) p.1.
[79] P. Weinberger, Phil. Mag. 92 12 (2012) p.382.
[80] D. Thomson, England in the Nineteenth Century, 1815–1914, Penguin BooksHarmondsworth, 1967.
[81] H. Pelling, A History of British Trade Unionism, Penguin BooksHarmondsworth, 1973.

Appendix A. Maxwell's propositions

In order to show the perfect logical order in Maxwell's first paper [21], and also to give an impression of his minimalist use of words, in the following a list of his "propositions" is compiled.

Part I. On the Motions and Collisions of Perfectly Elastic Spheres

Prop. I. Two spheres moving in opposite directions with velocities inversely as their masses strike one another; to determine their motions after impact.

Prop. II. To find the probability of the direction of the velocity after impact lying between given limits.

Prop. III. Given the direction and magnitude of the velocities of two spheres before impact, and the line of centres at impact; to find the velocities after impact.

Prop. IV. To find the average number of particles whose velocities lie between given limits, after a great number of collisions among a great number of equal particles.

Prop. V. Two systems of particles move each according to the law stated in Prop. IV.; to find the number of pairs of particles, one of each system, whose relative velocity lies between given limits.

Prop. VI. Two systems of particles move in the same vessel; to prove that the mean vis viva of each particle will become the same in the two systems.

Prop. VII. A particle moves with velocity r relatively to a number of particles of which there are N in unit of volume; to find the number of these which it approaches within a distance s in unit of time.

Prop. VIII. A particle moves with velocity v in a system moving according to the law of Prop. IV.; to find the number of particles which have a velocity relative to the moving particle between r and $r + dr$.

Prop. IX. Two sets of particles move as in Prop. V.; to find the number of pairs which approach within a distance s in unit of time.

Prop. X. To find the probability of a particle reaching a given distance before striking any other.

Prop. XI. In a mixture of particles of two different kinds, to find the mean path of each particle.

Prop. XII. To find the pressure on unit of area of the side of the vessel due to the impact of the particles upon it.

Prop. XIII. To find the internal friction in a system of moving particles.

Part II. On the Process of Diffusion of two or more kinds of moving particles among one another

Prop. XIV. In a system of particles whose density, velocity, &c. are functions of x, to find the quantity of matter transferred across the plane of yz, due to the motion of agitation alone.

Prop. XV. The quantity transferred, in consequence of a mean motion of translation V, would obviously be $Q = V\rho$.

Prop. XVI. To find the resultant dynamical effect of all the collisions which take place in a given stratum.

Prop. XVII. To find the resultant effect of the collisions upon each of several different systems of particles mixed together.

Prop. XVIII. To find the mechanical effect of a difference in the mean velocity of translation of two systems of moving particles.

Prop. XIX. To find the law of diffusion in the case of two gases diffusing into each other through a plug made of a porous material, as in the case of the experiments of Graham.

Prop. XX. To find the rate of diffusion between two vessels connected by a tube.

Prop. XXI. To find the amount of energy which crosses unit of area in unit of time when the velocity of agitation is greater on one side of the area than on the other.

Part III. On the Collision of Perfectly Elastic Bodies of any Form

Prop. XXII. Two perfectly elastic bodies of any form strike each other: given their motions before impact, and the line of impact, to find their motions after impact.

Prop. XXIII. To find the relations between the average velocities of translation and rotation after many collisions among many bodies.

Appendix B. The kinetic theory of gases and the Maxwell–Boltzmann (velocity) distribution

In this Appendix a short, simplified summary of the Maxwell–Boltzmann approach to the kinetic theory of gases and their famous velocity distribution is compiled.

B.1. *The "molecular" distribution function*

Let

$$f(\mathbf{r}, \mathbf{u}, t)\underbrace{dxdydz}_{d\mathbf{r}}\underbrace{du_x du_y du_z}_{d\mathbf{u}} \tag{B.1}$$

be the number of molecules (atoms, hard spheres), which are at the time t at $\mathbf{r}$ in the volume element $d\mathbf{r}$ and whose component velocities are in the intervals $[u_x, u_x+du_x]$, $[u_y, u_y+du_y]$ and $[u_z, u_z+du_z]$. The six-dimensional space of the coordinates x, y, z, u_x, u_y, u_z is called **phase space**. Integrating over all velocities,

$$\iiint\limits_{\{u_x,u_y,u_z\}} f(\mathbf{r}, \mathbf{u}, t)d\mathbf{r}d\mathbf{u} = n(\mathbf{r}, t)d\mathbf{r}\ , \tag{B.2}$$

yields the so-called **particle density**, namely the number of (velocity independent) particles which at the time t are located within the volume element $d\mathbf{r}$. Integration over the volume element $d\mathbf{r}$,

$$\iiint n(\mathbf{r}, t)d\mathbf{r} = N(t), \tag{B.3}$$

determines the total number of particles at a given time t. Quite clearly the molecular distribution function $f(\mathbf{r}, \mathbf{u}, t)$ is not normalized to unity, but to the number of particles. In the following, only equilibrium situations shall be considered. Obviously, the individual gas molecules can move around quite a bit, the averaged number of particles per volume, however, shall be assumed to be not time dependent ("steady state approximation", "stationary state"), i.e. to be a constant.

B.2. *Atomistic picture of the pressure of an ideal gas*

Suppose one wants to calculate the pressure p exerted by a monoatomic gas on the walls of the container. Phenomenologically, the averaged force $\overline{F}_\perp$ perpendicular to a surface element ΔA is given by

$$\overline{F}_\perp = p\Delta A. \tag{B.4}$$

This force can also be viewed as a time average in the interval $[t, t+\Delta t]$ of the forces exerted by the individual molecules hitting the wall:

$$p\Delta A = (\overline{F}_\perp)_{gas\ on\ wall} = \frac{1}{\Delta t}\int\limits_{\Delta t} (F_\perp)_{gas\ on\ wall}\, dt. \tag{B.5}$$

If $p_\perp$ denotes the momentum of a gas particle perpendicular to ΔA, then the force the wall exerts on to this particle is given by

$$(F_\perp)_{wall\ on\ gas} = \frac{dp_\perp}{dt}, \tag{B.6}$$

and oppositely the force exerted by the gas on the wall by

$$(F_\perp)_{gas\ on\ wall} = -\frac{dp_\perp}{dt}, \tag{B.7}$$

The equilibrium obviously is given by

$$(\overline{F}_\perp)_{wall\ on\ gas} = -\,(\overline{F}_\perp)_{gas\ on\ wall}. \tag{B.8}$$

Combining, therefore, (B.5) and (B.7) one gets

$$p\Delta A\Delta t = \int\limits_{\Delta t}\left(-\frac{dp_\perp}{dt}\right)dt = (p_\perp)_1 - (p_\perp)_2 = -\,\Delta p_\perp, \tag{B.9}$$

where $(p_\perp)_1$ is the perpendicular component of the momentum of the particle before the collision with the wall and $(p_\perp)_2$ after the collision. From *Figure 8*, one can see that for one particular velocity $\mathbf{u}$, whose direction forms an angle θ, $0 \le \theta \le \pi/2$, with the surface normal, all those particles with velocity $\mathbf{u}$ that are within the length $|\mathbf{u}|\,\Delta t$ from the wall will hit the wall within the time interval Δt, i.e. one can see that all those molecules that are within the volume element $dV = \Delta A\,|\mathbf{u}|\,\Delta t\cos\theta$ will arrive at the wall within Δt. If one assumes now a distribution function of velocities, $f(\mathbf{u})$, the number of particles actually hitting the wall is given by

$$f(\mathbf{u})\underbrace{\Delta A\,|\mathbf{u}|\,\Delta t\cos\theta}_{dV}d\mathbf{u}. \tag{B.10}$$

Let m denote the mass of the particles then the momentum perpendicular to the wall for one particular particle is given by $m\,|\mathbf{u}|\cos\theta$ and the total momentum therefore by this momentum times the total number of particles, namely

$$m\,|\mathbf{u}|\cos\theta\ \left(\ f(\mathbf{u})\Delta A\,|\mathbf{u}|\,\Delta t\cos\theta d\mathbf{u}\ \right) = m\Delta t\Delta A\left(|\mathbf{u}|^2\cos^2\theta\right)f(\mathbf{u})d\mathbf{u}.$$

The total momentum before the collision ($(p_\perp)_1$, (B.9)) is therefore given by

$$(p_\perp)_1 = m\Delta t\Delta A\iiint |\mathbf{u}|^2\cos^2\theta f(\mathbf{u})d\mathbf{u}\quad ;\quad 0 \le \theta \le \pi/2\,, \tag{B.11}$$

whereas after the collision with the wall the individual momentum is given by $-\,m\,|\mathbf{u}|\cos\theta$ and the total momentum by

$$(p_\perp)_2 = -\,m\Delta t\Delta A\iiint |\mathbf{u}|^2\cos^2\theta f(\mathbf{u})d\mathbf{u}\quad ;\quad \pi/2 \le \theta \le \pi\ . \tag{B.12}$$

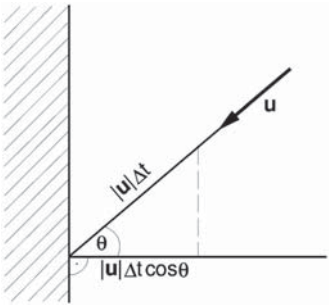


Figure 8. Atomistic picture of the pressure of an ideal gas.

Going now back to (B.9), one can see that by combining the last two equations,

$$p\Delta A\Delta t = (p_\perp)_1 - (p_\perp)_2 \tag{B.13}$$

$$= m\Delta A\Delta t \iiint |\mathbf{u}|^2 \cos^2\theta f(\mathbf{u}) d\mathbf{u} \quad ; \quad 0 \le \theta \le \pi \ , \tag{B.14}$$

or, that the pressure p can be written as

$$p = mn \iiint |\mathbf{u}|^2 \cos^2\theta \underbrace{f(\mathbf{u})n^{-1}} d\mathbf{u} \quad ; \quad 0 \le \theta \le \pi \ , \tag{B.15}$$

where n is the particle density as given by the total number of particles N divided by the volume V and N is according to the "steady state approximation" a constant. By including the particle density n to the integral, the quantity $f(\mathbf{u})n^{-1}$ is a probability density, since it is normalized to the " total" particle density, i.e. since

$$0 \le f(\mathbf{u})n^{-1} \le 1.$$

Denoting now by $\overline{u}_\perp^2$, the average of the square of the velocity in z-direction (direction of the surface normal), then the pressure is given by

$$p = mn\overline{u}_\perp^2. \tag{B.16}$$

If one finally assumes that the velocity distribution is **isotropical**, i.e. all directions of the velocity, u_x, u_y and u_z, are equally probable,

$$\overline{u}_\perp^2 = \overline{u}_x^2 = \overline{u}_y^2 = \overline{u}_z^2 = \frac{1}{3}(\overline{u_x^2 + u_y^2 + u_z^2}) = \frac{1}{3}\,\overline{u}^2, \tag{B.17}$$

one gets

$$p = \frac{1}{3} nm\,\overline{u}^2. \tag{B.18}$$

Denoting the averaged kinetic energy of a gas molecule (particle) by $\overline{\epsilon}_t$,

$$\overline{\epsilon}_t = \frac{m}{2}\,\overline{u}^2, \tag{B.19}$$

for an isotropic velocity distribution the pressure is then defined by

$$p = \frac{2}{3}\,n\overline{\epsilon}_t = \frac{2}{3}\frac{N}{V}\,\overline{\epsilon}_t. \tag{B.20}$$

Since the total number of particles N can be written as a multiple of the **Avogadro** (**Loschmidt**) number L , $N = \nu L$, the pressure can be written also in the following form

$$p = \frac{2}{3}\frac{\nu L}{V}\,\overline{\epsilon}_t, \tag{B.21}$$

or, alternatively one gets

$$pV = \frac{2}{3}\nu L\overline{\epsilon}_t. \tag{B.22}$$

If one compares this equation with the famous equation for an ideal gas,

$$pV = \nu RT, \tag{B.23}$$

where T is the temperature, one immediately can deduce that

$$\overline{\epsilon}_t = \frac{3}{2}\left(\frac{R}{L}\right)T = \frac{3}{2}kT, \tag{B.24}$$

or alternatively find an "atomistic-like" definition for the temperature

$$T = \frac{2}{3}k^{-1}\overline{\epsilon}_t. \tag{B.25}$$

In the last two equations k is the famous **Boltzmann constant**. Using (B.19) and (B.24),

$$\overline{u}^2 = \frac{3kT}{m} = \frac{3RT}{mL} = \frac{3RT}{\mu}, \tag{B.26}$$

where $\mu = mL$ is the molecular weight [g].

B.3. *Maxwell–Boltzmann distribution of velocities*

Suppose the velocity distribution $f(\mathbf{r}, \mathbf{u}, t)$ is rewritten as a product of the particle density $n(\mathbf{r}, t)$ and a distribution $F(\mathbf{r}, \mathbf{u}, t)$ normalized to unity,

$$f(\mathbf{r}, \mathbf{u}, t) = n(\mathbf{r}, t)\underbrace{n(\mathbf{r}, t)^{-1} f(\mathbf{r}, \mathbf{u}, t)} = n(\mathbf{r}, t)F(\mathbf{r}, \mathbf{u}, t). \tag{B.27}$$

In equilibrium, this new distribution function has to be time and space independent, i.e. $F(\mathbf{r}, \mathbf{u}, t) \equiv F(\mathbf{u})$. By imposing the following two conditions,

$$F(\mathbf{u}) = F(|\mathbf{u}|) \equiv F(u), \tag{B.28}$$

$$F(u) = F_1(u_x)F_2(u_y)F_3(u_z), \tag{B.29}$$

namely that $F(\mathbf{u})$ is isotropical and that the components of the velocity are mutually independent, $F(\mathbf{u})$ can be determined.

From (B.29) follows immediately that

$$\ln F(u) = \ln F_1(u_x) + \ln F_2(u_y) + \ln F_3(u_z), \tag{B.30}$$

which partially differentiated with respect to u_x gives

$$\frac{\partial \ln F(u)}{\partial u_x} = \frac{d \ln F(u)}{du}\frac{\partial u}{\partial u_x} = \frac{u_x}{u}\frac{d \ln F(u)}{du} = \frac{d \ln F_1(u_x)}{du_x}, \tag{B.31}$$

since $u = \sqrt{u_x^2 + u_y^2 + u_z^2}$. From the similar derivatives with respect to u_y and u_z follows

$$\frac{1}{u}\frac{d \ln F(u)}{du} = \frac{1}{u_x}\frac{d \ln F_1(u_x)}{du_x} = \frac{1}{u_y}\frac{d \ln F_2(u_y)}{du_y} = \frac{1}{u_z}\frac{d \ln F_3(u_z)}{du_z}. \tag{B.32}$$

These expressions can only be equal to each other, if they are equal the same constant, say -2γ. For example,

$$\frac{1}{u_x}\frac{d\ln F_1(u_x)}{du_x} = -2\gamma. \tag{B.33}$$

The solution of this differential equation is simply given by

$$F_1(u_x) = a\exp(-\gamma u_x^2), \tag{B.34}$$

where a and γ are constants yet to be determined. The distribution function $F(u)$ is therefore given by

$$F(u) = a^3\exp(-\gamma(u_x^2+u_y^2+u_z^2)) = a^3\exp(-\gamma u^2). \tag{B.35}$$

Since $F(u)$ is normalized to unity, i.e.

$$\int F(u)d\mathbf{u} = 1 \quad , \quad (d\mathbf{u} = u^2du\sin\theta d\theta d\phi), \tag{B.36}$$

a^3 has to be $(\gamma/\pi)^{3/2}$.

Finally, the constant γ can be evaluated by determining the averaged square of the velocity $\overline{u}^2$,

$$\begin{aligned}\overline{u}^2 &= \int u^2F(u)d\mathbf{u}\\ &= (\gamma/\pi)^{3/2}\int_0^\infty\int_0^\pi\int_0^{2\pi} u^4\exp(-\gamma u^2)du\sin\theta d\theta d\phi\\ &= 4\pi(\gamma/\pi)^{3/2}\int_0^\infty u^4\exp(-\gamma u^2)du = \frac{3}{2\gamma}.\end{aligned} \tag{B.37}$$

If one associates now this result with that for $\overline{u}^2$ in (B.26), one gets

$$\gamma = \frac{m}{2kT}, \tag{B.38}$$

i.e. $F(u)$ is given by

$$F(u) = \left(\frac{m}{2\pi kT}\right)^{3/2}\exp(-\frac{mu^2}{2kT}) = \left(\frac{m}{2\pi kT}\right)^{3/2}\exp\left(-\frac{\epsilon_t}{kT}\right), \tag{B.39}$$

where $\epsilon_t = mu^2/2$ is the translational energy. Equation (B.39) is nothing but the famous **Maxwell–Boltzmann (velocity) distribution**.

Appendix C. Statistical mechanics

Since nowadays Boltzmann's statistics is usually derived as a special case of the Bose–Einstein statistics, this appendix serves as a short summary of statistical mechanics in a more modern sense and hopefully will make Boltzmann's statistics more transparent. Experienced readers can of course skip this appendix.

C.1. *Phase space*

For one particle the phase space, that is the collection of space (**r**)- and momentum (**p**) coordinates of this particle, is six dimensional. For N particles, the dimension of the phase space is $6N$. If one considers the phase space to be partitioned into cells of the volume $d\phi$,

$$d\phi = dxdydzdp_xdp_ydp_z, \tag{C.1}$$

then each particle (atoms, molecules) has to belong to one cell. If furthermore i numbers these cells, then in each cell the number of particles N_i can be assumed to be much larger than one, $N_i \gg 1$, since $d\phi$ only has to be small as compared to the macroscopical dimensions of the system. The fundamental problem of statistical mechanics is to find out how these numbers N_i can be related to the phase space coordinates.

C.2. *Microstates and macrostates*

As is well known from quantum mechanics, because of the uncertainty relation, the space and momentum coordinates of a microscopical particle cannot be determined simultaneously exact:

$$\Delta x \Delta p_x = \Delta y \Delta p_y = \Delta z \Delta p_z \simeq h. \tag{C.2}$$

This implies that the space- and momentum coordinates of such a particle can only be pinned down to the volume h^3 in phase space. Such a volume shall be called a **subcell**. It should be noted that while the volume of a cell, $d\phi$, is arbitrary, the volume of a subcell is fixed to h^3. In general, the volume of a cell can be assumed to be large enough so as to contain a large number of subcells. Quite clearly, the number of subcells in a particular cell is given by

$$g = \frac{d\phi}{h^3}. \tag{C.3}$$

The complete specification of all six coordinates of the subcell in which a particular particle is located is called a microstate. The specification of the number of particles in each cell of the phase space is called a macrostate.

Up to now implicitly only monoatomic gases were considered for which it is sufficient to specify the subcell into which a particular atom is located. For molecules, for example, one has to deal also with internal degrees of freedom in addition to the translational energies.

C.3. *Thermodynamical probabilities*

The number of microstates that belong to each macrostate is called thermodynamical probability. However, before this number can be determined, one has to distinguish between bosons (arbitrary number of phase points per cell) and fermions (only two phase points per cell). As for the Boltzmann statistics, the Fermi-Dirac statistics is of no relevance, it is not dealt with in here.

C.4. *Bose–Einstein statistics*

In general for an arbitrary number M of cells i with individual probabilities W_i the total probability W is given by

$$W = \prod_{i=1}^{M} W_i. \tag{C.4}$$

Suppose that the subcells in one particular cell i can be labelled by $1, 2, \ldots, g_i$, whereas the phase points (particles) in this cell are labelled by $a, b, c, \ldots, N_i$. The number of permutations is then given by $(N_i + g_i - 1)!$ The number of sequences that start with a number is therefore $g_i(N_i + g_i - 1)!$ All permutations, however, that result from a pairwise interchange of phase points have to be excluded. There are altogether $N_i!$ such pairwise permutations. Furthermore, the numbering of cells has to be in natural order, otherwise $g_i!$ double countings occur. The total number of microstates in cell i is therefore given by the following individual probability W_i,

$$W_i = \frac{g_i(N_i + g_i - 1)!}{g_i! N_i!} = \frac{(N_i + g_i - 1)!}{(g_i - 1)! N_i!}, \tag{C.5}$$

and consequently the thermodynamical probability is defined by

$$W = \prod_{i=1}^{M} \frac{(N_i + g_i - 1)!}{(g_i - 1)! N_i!}. \tag{C.6}$$

By taking the logarithm of W,

$$\ln W = \sum_{i=1}^{M} \left\{ \ln(N_i + \underbrace{g_i - 1}_{\simeq g_i})! - \ln(\underbrace{g_i - 1}_{\simeq g_i})! - \ln(N_i!) \right\}, \tag{C.7}$$

and using the Stirling formula,

$$\ln(n!) = n \ln n - n, \quad n \gg 1, \tag{C.8}$$

one obtains the following expression,

$$\ln W \simeq \sum_{i=1}^{M} \{(N_i + g_i) \ln(N_i + g_i) - g_i \ln g_i - N_i \ln N_i\}, \tag{C.9}$$

where $(g_i - 1)$ was replaced by g_i, since $g_i \gg 1$.

Since particles in motion change their positions in phase space with respect to time, the number of phase points N_i in the various cells also changes with time. Quite clearly, if W has a maximum W^0, then also $\ln W$ is at its maximal value. Considering, therefore, small finite changes δN_i (finite, because the N_i are integer numbers!) and remembering that the number of subcells $g_i = d\phi_i / h^3$ is fixed by the uncertainty relation, such changes results in the following change $\delta \ln W$ for $\ln W$:

$$\delta \ln W = \sum_{i=1}^{M} \{\ln(N_i + g_i)\delta N_i + \underbrace{(N_i + g_i)\delta \ln(N_i + g_i)} - \ln N_i \delta N_i - \underbrace{N_i \delta \ln N_i}\}. \tag{C.10}$$

From the relations

$$\delta \ln N_i = \frac{\delta N_i}{N_i} \quad , \quad \delta \ln(N_i + g_i) = \frac{\delta N_i}{N_i + g_i}, \tag{C.11}$$

one easily can see that

$$\sum_{i=1}^{M} N_i \delta \ln N_i = \sum_{i=1}^{M} (N_i + g_i)\delta \ln(N_i + g_i) = \sum_{i=1}^{M} \delta N_i, \tag{C.12}$$

which reduces Equation (C.10) to

$$\delta \ln W = \sum_{i=1}^{M} \{\ln(N_i + g_i)\delta N_i - \ln N_i \delta N_i\} = \sum_{i=1}^{M} \ln\left(\frac{N_i + g_i}{N_i}\right) \delta N_i. \tag{C.13}$$

The set of N_i for which $\delta \ln W^0 = 0$, shall be denoted by $\{N_i^0\}$,

$$\delta \ln W^0 = \sum_{i=1}^{M} \ln\left(\frac{N_i^0 + g_i}{N_i^0}\right) \delta N_i = 0. \tag{C.14}$$

and obviously corresponds to the thermodynamical equilibrium. The variation of the N_i, however, is restricted by the following two conditions,

$$\sum_{i=1}^{M} N_i = N = const \quad \rightarrow \quad \sum_{i=1}^{M} \delta N_i = 0, \tag{C.15}$$

$$\sum_{i=1}^{M} N_i \epsilon_i = U = const \quad \rightarrow \quad \sum_{i=1}^{M} \delta N_i \epsilon_i = 0, \tag{C.16}$$

namely that total number particles and the total energy of the system has to be constant, ϵ_i being the energy of an individual particle. This implies that Lagrange parameters have to be used of the form

$$\lambda_1 \sum_{i=1}^{M} \delta N_i = 0 \quad , \quad \lambda_2 \sum_{i=1}^{M} \delta N_i \epsilon_i = 0, \tag{C.17}$$

such that

$$\sum_{i=1}^{M} \left\{ \ln \left(\frac{N_i^0 + g_i}{N_i^0} \right) + \lambda_1 + \lambda_2 \epsilon_i \right\} \delta N_i = 0. \tag{C.18}$$

Since in Equation (C.18) each factor of δN_i has to vanish, by choosing for matters of convenience

$$\lambda_1 = -\ln B \quad , \quad \lambda_2 = -\beta, \tag{C.19}$$

one easily can see that

$$\ln \left(\frac{N_i^0 + g_i}{N_i^0} \right) - \ln B_1 - \beta \epsilon_i = 0, \tag{C.20}$$

from which immediately follows that

$$\ln \left(\frac{N_i^0 + g_i}{B N_i^0} \right) = \beta \epsilon_i \quad \rightarrow \quad \frac{g_i}{N_i^0} = B \exp(\beta \epsilon_i) - 1. \tag{C.21}$$

The ratio of the number of phase points N_i per cell and the number of subcells (g_i) is therefore given by

$$\frac{N_i^0}{g_i} = \frac{1}{B \exp(\beta \epsilon_i) - 1}. \tag{C.22}$$

This ratio is called the Bose–Einstein distribution function.

C.5. *Maxwell–Boltzmann statistics*

If the number of particles per cell is much smaller than the number of subcells, i.e.

$$N_i^0 \ll g_i \;\rightarrow\; \frac{N_i^0}{g_i} \ll 1 \rightarrow\; B \exp(\beta \epsilon_i) - 1 \gg 1 \rightarrow\; B \exp(\beta \epsilon_i) - 1 \simeq B \exp(\beta \epsilon_i), \tag{C.23}$$

then

$$\frac{N_i^0}{g_i} = \frac{1}{B} \exp(-\beta \epsilon_i) \quad , \; B \gg 1. \tag{C.24}$$

This last equation is exactly of the form of the Maxwell–Boltzmann distribution, see also Appendix B. The inequalities in Equation (C.23) imply that within a gedanken experiment where the subcells become smaller and smaller, i.e. position and momentum of a particle becomes sharper and sharper, the Bose–Einstein distribution reduces to the Maxwell–Boltzmann distribution.

C.6. *Partition function*

The constant B appearing in Equations (C.22) and (C.24) can be obtained from the condition that the number of particles has to be constant. Assuming now always maximal probability, the superscript specifying the "optimal" set of numbers N_i can be dropped. From Equation (C.15) follows immediately that

$$\sum_{i=1}^{M} N_i = N = \frac{1}{B}\sum_{i=1}^{M} g_i \exp(-\beta\epsilon_i). \tag{C.25}$$

The sum appearing on the rhs of this equation is called partition function and traditionally is abbreviated by Z or Q

$$Z = \sum_{i=1}^{M} g_i \exp(-\beta\epsilon_i). \tag{C.26}$$

With this notation, the constant B is given by

$$B = \frac{Z}{N}, \tag{C.27}$$

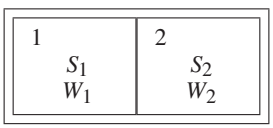
and the Maxwell–Boltzmann distribution can be formulated as

$$N_i = \frac{N}{Z} g_i \exp(-\beta\epsilon_i). \tag{C.28}$$

The evaluation of the constant β will be given below.

C.7. *Entropy and thermodynamical probability*

Consider an isolated thermodynamical system that is partitioned into two isolated subsystems 1 and 2 with thermodynamical probabilities W_1 and W_2.



Let S_1 and S_2 be two functions such that in thermodynamical equilibrium

$$S = S_1 + S_2. \tag{C.29}$$

S is called entropy. If there is a connection between this phenomenological quantity S and the thermodynamical probability $W = W_1 W_2$ (in thermodynamical equilibrium) then S has to be a function of W such that

$$S = S_1 + S_2 = f(W) = f(W_1 W_2) = f(W_1) + f(W_2). \tag{C.30}$$

Quite clearly, S can only be proportional to the logarithm of the thermodynamical probability W, if

$$S = k \ln W. \tag{C.31}$$

Going now back to Equation (C.9)

$$\ln W^0 \equiv \ln W = \sum_{i=1}^{M} \underbrace{(N_i + g_i)\ln(N_i + g_i)} - g_i \ln g_i - N_i \ln N_i, \tag{C.32}$$

the first term on the rhs of this equation can be rewritten as follows

$$(N_i + g_i)\ln(N_i + g_i) = g_i\left(1 + \frac{N_i}{g_i}\right)\ln\left(g_i\left(1 + \frac{N_i}{g_i}\right)\right)$$
$$= g_i\left(1 + \frac{N_i}{g_i}\right)\left(\ln g_i + \ln\underbrace{\left(1 + \frac{N_i}{g_i}\right)}\right). \tag{C.33}$$

Using the following expansion of $\ln(1 + x) = x - x^2/2 + x^3/3 \pm \cdots$; $(-1 \le x \le 1)$, it is obvious that for $g_i \gg N_i$ (Maxwell–Boltzmann distribution)

$$\ln\left(1 + \frac{N_i}{g_i}\right) \simeq \frac{N_i}{g_i},$$

and therefore

$$(N_i + g_i)\ln(N_i + g_i) = g_i \ln g_i + N_i \ln g_i + N_i + \frac{N_i^2}{g_i} \simeq g_i \ln g_i + N_i \ln g_i + N_i. \tag{C.34}$$

In the case of the Maxwell–Boltzmann distribution, Equation (C.32) reduces therefore to

$$\ln W = \sum_{i=1}^{M} N_i \ln g_i + N_i - N_i \ln N_i = \sum_{i=1}^{M} N_i\left(\ln\frac{g_i}{N_i} + 1\right). \tag{C.35}$$

By using now in this equation the definition for the partition function Z, Equation (C.28), which was derived from the condition that the number of particles has to stay constant,

$$\frac{N_i}{g_i} = \frac{N}{Z}\exp(-\beta\epsilon_i) \rightarrow \ln\frac{g_i}{N_i} = \ln\frac{Z}{N} + \beta\epsilon_i \quad , \quad Z = \sum_{i=1}^{M} g_i \exp(-\beta\epsilon_i), \tag{C.36}$$

one gets for the entropy S:

$$\begin{aligned} S &= k\ln W \\ &= k\sum_{i=1}^{M} N_i\left(\ln\frac{Z}{N} + \beta\epsilon_i + 1\right) \\ &= k\ln\frac{Z}{N}\sum_{i=1}^{M} N_i + k\beta\underbrace{\sum_{i=1}^{M}\epsilon_i N_i}_{=U} + k\sum_{i=1}^{M} N_i \\ &= kN\ln\frac{Z}{N} + k\beta U + kN. \end{aligned} \tag{C.37}$$

From classical (phenomenological) thermodynamics, however, it is known that in the case of only one type of particles the partial derivative of the entropy with respect to the energy U at constant volume v is given by the inverse of the temperature T,

$$\left(\frac{\partial S}{\partial U}\right)_v = \frac{1}{T}. \tag{C.38}$$

In evaluating this derivative,

$$\left(\frac{\partial S}{\partial U}\right)_v = \frac{kN}{Z}\frac{dZ}{d\beta}\left(\frac{\partial\beta}{\partial U}\right)_v + k\beta + kU\left(\frac{\partial\beta}{\partial U}\right)_v, \tag{C.39}$$

one can make use of the fact that

$$\frac{dZ}{d\beta} = -\sum_{i=1}^{M} \epsilon_i g_i \exp(-\beta\epsilon_i) = -\frac{UZ}{N}, \tag{C.40}$$

since from the definition of the partition function

$$N_i = \frac{N}{Z} g_i \exp(-\beta\epsilon_i)$$

it follows that

$$\sum_{i=1}^{M} \epsilon_i N_i = U = \frac{N}{Z} \sum_{i=1}^{M} g_i \epsilon_i \exp(-\beta\epsilon_i). \tag{C.41}$$

Using now the expression for $dZ/d\beta$, Equation (C.40) in Equation (C.39) one gets

$$\left(\frac{\partial S}{\partial U}\right)_v = -\frac{kN}{Z}\frac{UZ}{N}\left(\frac{\partial \beta}{\partial U}\right)_v + k\beta + kU\left(\frac{\partial \beta}{\partial U}\right)_v = k\beta. \tag{C.42}$$

Comparing finally this result with the classical expression in Equation (C.38) one can see that obviously the constant β has to be of the form

$$\beta = \frac{1}{kT}. \tag{C.43}$$

Now all Lagrange parameters are evaluated and the Boltzmann distribution in Equation (C.28) – and for that matter also the Bose–Einstein distribution – can be completed:

$$N_i = \frac{Ng_i}{Z} \exp\left(-\frac{\epsilon_i}{kT}\right). \tag{C.44}$$

The partition function Z is then given by

$$Z = \sum_{i=1}^{M} g_i \exp\left(-\frac{\epsilon_i}{kT}\right), \tag{C.45}$$

and the total energy U by

$$U = \frac{N}{Z} \sum_{i=1}^{M} g_i \epsilon_i \exp\left(-\frac{\epsilon_i}{kT}\right) = \frac{NkT^2}{Z}\frac{dZ}{dT} = NkT^2 \frac{d\ln Z}{dT}. \tag{C.46}$$

C.8. *Partition functions*

In the Maxwell–Boltzmann statistics, which shall be applied in the following, the partition function Z, Equation (C.26), is given by a sum over the exponential functions only

$$Z = \sum_{i=1}^{M} \exp(-\epsilon_i/kT). \tag{C.47}$$

C.9. *Summary of thermodynamical functions for weakly interacting particles*

In summarizing all the above expressions for the various thermodynamical quantities or functions, one can see at one glance the overall importance of the partition function in statistical thermodynamics:

$$U = -nR\left(\frac{\partial \ln Z}{d(1/T)}\right)_v,$$

$$c_v = \frac{R}{T^2}\left(\frac{\partial^2 \ln Z}{\partial(1/T)^2}\right)_v,$$

$$S = nR\left(\ln\left(\frac{Z}{N}\right) - \frac{1}{T}\left(\frac{\partial \ln Z}{\partial(1/T)}\right)_v + 1\right),$$

$$F = -nRT\left(\ln\left(\frac{Z}{N}\right) + 1\right),$$

$$p = nRT\left(\frac{\partial \ln Z}{\partial v}\right)_T,$$

$$G = -nRT\left(\ln\left(\frac{Z}{N}\right) + 1 - v\left(\frac{\partial \ln Z}{\partial v}\right)_T\right),$$

$$H = nRT\left(\left(\frac{\partial \ln Z}{\partial \ln T}\right)_v + \left(\frac{\partial \ln Z}{\partial \ln v}\right)_T\right).$$