

Vol. 3, 10a. “On Boltzmann’s Principle and Some of Its Direct Consequences”

[p. 1]

[Zurich, 2 November 1910]^[1]

~~Thermodynamics is (known to be) based on two principles, the energy principle (also called the 1st law) and the principle of the irreversibility of natural events (also called the 2nd law). (This latter principle by states.) The substance of this latter principle can be expressed (in the Planckian sense) thus, according to Planck)~~

All natural science^[2] is founded on the presumption of an (unbroken) entirely causal relation for all occurrences. Let us assume that Galileo had found from his pendulum experiments that the period of one oscillation of this pendulum changed in a very irregular way. Let us assume furthermore that this change could not be connected with any change in other observable relations. Then it would have been impossible for Galileo to gather his observations under a law. If all phenomena accessible to us had as irregular a character as we have just pictured in this fictitious case, people would certainly never have resorted to scientific endeavors.

~~Which characteristic must phenomena have in order for science to be possible? (To this, one might first want to reply somewhat as follows: If we put a system into a particular state, then provided this system is separated from other systems—such as by a large spatial distance—then over time the course of the states of this system is completely determined; i.e., if we put (two) arbitrarily many equally composed isolated systems into exactly the same state and leave these systems alone, then the evolution of the phenomena in time is exactly the same for all these systems.)~~

Now, how about the (unbroken) entirely causal connection between events according to our knowledge today? This question has to be specified more precisely before it can be solved. Let us do this right away by means of an example. Take a cube of copper of a given size. Within this cube (we imagine) we establish by external influences a very specific temperature distribution and then, after having enveloped it in a thermally insulating shell, we leave it be. We know that in the course of time a temperature equilibrium will then set in by the process of thermal conduction. The temperature gradient at all points of the cube will thereby prove to be “uniquely determined” by the initial state; by the expression “uniquely determined” we mean that we are always going to perceive the same temperature gradients, no matter how often we may repeat the experiment, that is, no matter how often we may set the starting temperature distribution and then leave the cube be. Does this unique determinability of the process, this entirely causal connection of events really exist? In order to address a pertinent objection but of no interest to us

[p. 2]

here, we rather pose this question in the following form: Do we always establish an entirely causal connection between the events to closer approximation, the more exactly we realize the initial state and the more exactly we follow the process in time by measurement?

[p. 3] The point of view of physicists toward this problem changed considerably in the last century. If we leave aside for the time being Brownian motion, radioactive fluctuations, and a few other phenomena, which came into the focus of scientific interest in the past few years, we arrive decidedly at the verdict that according to experience^[3] an entirely causal connection in the last sense indicated does exist. Nevertheless, physicists, more specifically, heat theoreticians, managed to deny the entirely causal connection between events, more precisely speaking, between events insofar as they can be the objects of observation. Let us throw a fleeting glance at this development! From the simple idea that gases are composed of material points (molecules) that essentially only act upon one another mechanically by contact (collision), Clausius was able to derive a relation between the specific heats and the constants of the equations of state of monatomic gases as well as a ratio between thermal conduction, internal friction, and the diffusion of gases, which magnitudes or phenomena, resp., were entirely unrelated without Clausius's theory.^[4] This major success prompted physicists to attribute heat phenomena to irregular motions by molecules. This kinetic theory of heat required, however, that the laws of conduction, etc., be taken as only approximately valid laws; according to this theory a precisely valid law of conduction is absolutely impossible, being instead just a law of averages. That deviations from these averaging laws must normally be very small is unimportant in principle.

The kinetic theory so broadly supported by experience is not just incompatible with the presumption that observable occurrences be related \langle precisely \rangle completely causally, however. The analyses carried out by \langle Cla \rangle Maxwell, Boltzmann, and Gibbs^[5] also show that randomly large deviations from those averaging laws would have to occur within the range of observation, even though this happens so rarely in most \langle groups of phenomena accessible to observation \rangle systems according to the theory that we are not really able to detect such deviations.

[p. 4] The following well-known consideration shows most concisely that the laws of conduction, just as all other laws concerning irreversible processes, cannot be precise. According to the kinetic theory of heat the reversal in time of any process of molecular motion is likewise a possible process of motion; hence no such thermal process exists that could not also run in [the] opposite direction. Therefore, from the point of view of the molecular theory of heat, it has to be regarded as possible that by mere thermal conduction heat would flow from a colder body into a warmer

one. Why do we not observe this? Doesn't this consideration show that the kinetic theory of heat has to be abandoned?

This question was answered by Boltzmann, specifically, in [the] following way: Observe some closed physical system having a particular given amount of energy. We signify by $Z_1, Z_2 \dots Z_l$ all observable states that this system can assume at the given energy value. In the example of the copper cube, each Z_v would thus mean a particular temperature distribution, where in total l distinguishable temperature distributions are possible. But now it is assumed that these states Z be of entirely different probabilities, such that from among all the states differing very little from a given state Z_a , one (Z_b) be far more probable than all the others, at least provided that Z_a differ substantially from the so-c[alled] state of thermodynamic equilibrium. Then, if brought into the state Z_a and then left to itself, the system is far more likely to change into the state Z_b than into any other states neighboring state Z_a . The probability that this would occur can come as close to unity (i.e., certainty) as you like, although it is excluded in principle that this transition be entirely certain. T[his] m[ean]: If we bring the system very frequently into the state Z_a , then in the great majority of cases, but by no means always, state Z_b will follow state Z_a ; a transition into every other neighboring state to state Z_a will also occasionally occur, even if only extremely rarely. What has been said about the transition from state Z_a into neighboring state Z_b is again valid for the change a system experiences from state Z_b in the following little segment of time. Thus one arrives at a conception of (apparently) irreversible processes. [p. 5]

This sketch of the Boltzmann conception is incomplete. Still needing to be answered are the questions: "How should the probability of individual states $Z_1, Z_2 \dots$ be understood?" and "Why is a transition from one state Z_a to the most probable neighboring state Z_b more probable than a transition to other neighboring states?"

For the first of these questions we note the following: According to the (molecular th) kinetic theory of heat, there cannot be a thermal equilibrium in the strict sense. The state we call a thermal equilibrium is one that a system left to itself for enormously long most frequently has. However, it is a consequence of the kinetic theory that over long periods of time the system will take on all possible states; in particular, the further a state is away from thermodynamic equilibrium, the more rarely does the system assume it. The copper cube left infinitely long to itself incessantly changes its temperature distribution, whereby it extremely seldom assumes temperature distributions differing considerably from the temperature distribution of thermal equilibrium. If we imagine a system under observation for an immensely long time T , there will be for most states Z_v an abnormally small portion τ of this

total time during which the system is just assuming the state Z_v . We shall call the ratio $\frac{\tau}{T}$ the probability W of the state concerned.

[p. 6] If this definition of the probability of a state is taken as a basis, it is generally understandable that a system changes on average from one state Z_a such that from this state follows with the greatest probability the neighboring state Z_b . I only have to mention this without going into the proof. (This is the answer to the second of the questions posed above.)

It is essential that the definition of the probability of a state be definable independent of the kinetic picture; probability W is a magnitude in principle accessible to observation, even though in most cases direct observation of it is excluded, owing to the brevity of the time at our disposal.

If a system in a state substantially differing from thermodynamic equilibrium is left alone, it successively assumes states of ever greater W . A state's probability W shares this property with the entropy S of a system, and Boltzmann found out that the relation between W and S

$$S = k \lg W$$

holds, where k is a universal constant, i.e., independent of the system chosen. (This is the important equation that ... the mathematical expression of the Boltzmann conception)

This Boltzmann equation can be applied in two different ways. There can be a more or less complete picture of molecular theory on the basis of which one can calculate the probability W . The Boltzmann equation then yields the entropy S . This was how Boltzmann's equation has mostly been applied hitherto.

[p. 7] Example.^[6] In a volume V , let there be N molecules, i.e., one gram molecule of a particular type. The volume is large enough compared to the eigenvolume of the N molecules and of the other existing matter besides the N molecules—provided such matter is there—distributed evenly over V_0 so that the various points of V_0 are equivalent for each of the N molecules. This is an incomplete expression of how we visualize an ideal gas or a dilute solution. How large is the probability W that at a randomly chosen instant all N molecules are within the partial volume V of volume V_0 ?

A simple consideration yields

$$W = \left(\frac{V}{V_0}\right)^N.$$

From this, using the Boltzmann (constant) equation, we find

$$S = kN \lg\left(\frac{V}{V_0}\right) = kN \lg V + \text{const.},$$

where the constant “const.” can depend on the temperature but not on the volume. From this we immediately obtain the force that the N molecules are able to exert on a wall that forces them to remain within volume V . For, if the energy of the system is independent of V , and if G signifies the work received upon infinitesimally enlarging volume V along a reversible path, then

$$pdV = G = +TdS = +kNT\frac{dV}{V}, \text{ holds,}$$

hence
$$pV = kNT.$$

We thus have the equation of ideal gases and osmotic pressure. At the same time it is revealed that the universal constant kN of this equation is equal to constant R of the gas equation.

In my opinion, the main importance of the Boltzmann equation does not lie in that with its help one is able to calculate the entropy for a known molecular scenario. Rather, the most important application is, conversely, that by means of Boltzmann's equation one can get the statistical probabilit[ies] of the individual states from the empirically established entropy function S . Thus it is possible to assess how much the systems' behaviors differ from the behavior required by thermodynamics. [p. 8]

Example.^[7] A particle that is slightly heavier than the liquid in which it is suspended and which it displaces.

Such a particle should, according to thermodynamics, sink to the bottom of the vessel and stay there. According to Boltzmann's equation, however, a probability W is ascribed to each height z above the bottom; the particle incessantly changes its height in an irregular way. We want to determine S and from it W . If μ is the particle's mass and μ_0 is that of the liquid it displaces, then the work $A = (\mu - \mu_0)gz$ has to be expended to raise the particle to height z above the bottom. In order for the energy of the system to stay constant, the amount of heat $G = A$ has to be removed from the system, whereupon the entropy diminishes by $\frac{G}{T} = \frac{A}{T}$. Therefore,

$$S = \text{const} - \frac{1}{T}(\mu - \mu_0)gz.$$

From the Boltzmann equation it follows, if one substitutes the value $\frac{R}{N}$ for k :

$$W = \text{const} e^{-\frac{N}{RT}(\mu - \mu_0)gz}.$$

If there are many particles present in a single liquid, then the right-hand side of the equation indicates the distribution density of the particles as a function of depth. Perrin tested this relation and found it confirmed.^[8]

From this relation the law of Brownian motion can very easily be deduced. For, from it, it immediately follows that the mean height \bar{z} of a particle above the vessel bottom is equal to

$$\frac{\int z e^{-\frac{N}{RT}(\mu - \mu_0)gz} dz}{\int e^{-\frac{N}{RT}(\mu - \mu_0)gz} dz} = \frac{RT}{N} \cdot \frac{1}{g(\mu - \mu_0)} .$$

Now, however, because of its greater density the particle drops downward, according to Stokes's law, by $D = \frac{g(\mu - \mu_0)}{6\pi\eta P} \tau$

[p. 9] in the time τ if η signifies the liquid's viscosity coefficient and P , the radius of the (spherically shaped) particle. But in the same time τ , as a consequence of the irregularity of the molecular thermal process, [it] is also shifted a distance Δ upwards or downwards, where positive and negative values for Δ appear equally frequently; so $\bar{\Delta}$ is = 0.

A particle that, before time τ has elapsed, is located at height z , is, after τ has elapsed, at the height $z - D + \Delta = z'$. As the distribution law of all the particles should not depend on time, the mean value of z'^2 must be equal to \bar{z}^2 , therefore,

$$\overline{(z - D + \Delta)^2} = \bar{z}^2 ,$$

or for sufficiently small τ , D^2 is negligible and $\bar{z}\bar{\Delta} = \overline{D\Delta} = 0$

$$\bar{\Delta}^2 = 2\bar{z}D = \frac{RT}{N} \cdot \frac{1}{3\pi\eta P} \tau .$$

This is the familiar law of Brownian motion, which has likewise been confirmed by experience.^[9]—

The just-described example of a particle suspended in a liquid offers a fitting depiction of Boltzmann's conception of irreversible processes. For, if we imagine a particle suspended in such a tall vessel, and that it is so much heavier than the displaced liquid that the expression for probability W is very small, even at a height z just barely above the bottom of the vessel when compared to the value W_0 for $z = 0$, then very rarely will the particle rise much from the bottom, once it has

reached the bottom (thermodynamic equilibrium). If we raise the particle to a considerable height z , then, obviously, it will with the greatest probability sink back down to the bottom (irreversible process) in order then to dance up and down, as before, in its proximity. If this sinking back did not occur in the overwhelming majority of cases, a probability function of the assumed quality could not be valid.–

Before I go into other applications of Boltzmann's equation, I would like to draw a general conclusion about it regarding the mean size of the fluctuations that the parameters of a system perform around the values for ideal thermodynamic equilibrium.^[10] $\lambda_1 \dots \gamma_n$ are parameters defining the state of a system. The null values for the λ 's are chosen so that at thermal equilibrium $\lambda_1 = \lambda_2 \dots$ is $= 0$. The work that, according to thermodynamics, would have to be performed in order to bring the system from the state of thermodynamic equilibrium into the state very close to thermodynamic equilibrium characterized by the values $\lambda_1 \dots \gamma_n$, is [p. 10]

$$A = \sum A_v = \sum_1^n \frac{a_v}{2} \lambda_v^2.$$

In order that the system's energy be the same as before, after the state has been established, the amount of heat $G = A$ must be removed from it, which corresponds to a reduction in the system's entropy by $\frac{G}{T} = \frac{A}{T}$. Thus, if the system has assumed the considered state on its own, its entropy is

$$S = \text{const} - \frac{1}{T} \sum_1^n \frac{a_v}{2} \lambda_v^2.$$

If this is plugged into the Boltzmann equation, one obtains

$$W = \text{const} e^{-\frac{N}{RT} \sum_1^n \frac{a_v}{2} \lambda_v^2}$$

In this case, therefore, Gauss's law of error distribution applies to the deviations of the individual parameters from the values for thermodynamic equilibrium. For the mean work \overline{A}_v that according to thermodynamics would have to be expended to bring the parameter λ_v from equilibrium to the temporal mean $\sqrt{\overline{\lambda_v^2}}$ in a reversible process, one obtains the value

$$\overline{A}_v = \frac{RT}{2N}.$$

[p. 12] This result can be expressed thus: Provided A can be described in the above-indicated manner in the vicinity of thermodynamic equilibrium, deviations from the state of ideal thermodynamic equilibrium set in of their own accord; these deviations are, on average, for each parameter so large that the work required by thermodynamics for arbitrary generation of the deviation is equal to one third of the mean kinetic energy of the propagating motion of a gas molecule at the same temperature. Perceptible deviations from the state of ideal thermodynamic equilibrium occur everywhere where the performance of work, however small, can achieve a perceptible effect. Measurement of each such deviation provides us with a determination of the energy of the monatomic gas molecule, hence also a determination of the absolute size of the atom.

Smoluchowski has indicated a very interesting application of this general result. According to classical thermodynamics, the individual components of a phase in the case of thermodynamic equilibrium are distributed evenly over the volume of the phase. From what was said before, irregularities must, on the contrary, occur in the spatial distribution of the matter, which are greater, the weaker are the forces opposing a change in the even distribution of the matter, or, resp., the separate, independent components. So the phase is in reality inhomogeneous, which makes itself noticeable by an optical opacity (opalescence). This opalescing is particularly strong in the vicinity of critical states (for uniform substances and for solutions), because in these cases only slight forces oppose a change in density or concentration. A short while ago I demonstrated that on the basis of the outlined interpretation by Smoluchowski an exact calculation of the light diffracted by opalescence is possible.^[11]

[p. 13] Finally, I would not like to leave unmentioned that by means of the Boltzmann equation the statistical properties of thermal radiation are derivable in a simple way from the laws of thermal radiation, more precisely speaking, without having to avail oneself of electromagnetic and kinetic heat theory. The problem is the following: In a cavity surrounded by opaque bodies at temperature T , there is radiation whose quality is determined solely by the temperature. At time τ a defined radiant energy E passes through a surface σ imagined to lie somewhere within the cavity; its directional range is indicated by a definite elementary cone $d\Omega$ and its frequency range is dv . If this radiant energy is imagined to occur often and moreover is measured very precisely, then one would not



always find the same value E but a magnitude $E = E_0 + \varepsilon$, slightly deviating from a mean E_0 . One would seek the quadratic mean $\overline{\varepsilon^2}$ of this magnitude ε . This problem is of essential interest because its solution contains a statement about the structure of thermal radiation.

I would just like to suggest the way this problem can be solved. If some arbitrary body K is thermally related with another one of relatively infinitely large thermal capacity, then according to thermodynamics K will take on the temperature of this second body and continue to maintain it. According to Boltzmann's principle, however, the temperature of K will constantly change, albeit rarely substantially far removed from the temperature of thermal equilibrium; the Boltzmann equation yields the mean of those temperature fluctuations. The temperature fluctuations thus obtained are entirely independent of the way in which the thermal exchange between K and the relatively infinitely large body takes place; the temperature fluctuation is of the calculated magnitude even when this thermal exchange takes place exclusively by way of radiation. One then only has to investigate the question: Which would have to be the statistical properties for the radiation to really generate the calculated temperature fluctuations? If this analysis is carried out as suggested, the result one obtains is that the temporal fluctuations of the thermal radiation at a low radiation intensity and great frequency are far greater than would be expected according to our current theory.^[12] [p. 14]

If we now in closing again ask the question: "Are observable physical phenomena connected entirely causally with one another?" we definitely have to negate this question. To even the most conscientious observer, the positions a particle undergoing Brownian motion in two time values separated by one second must always seem completely independent of each other; and the greatest mathematician will never succeed in calculating in advance in a specific case, even approximately, the path traveled in one second by such a particle. According to the theory, in order to be able to do so, one would have to know precisely the positions and velocities of all the individual molecules, which in principle appears to be out of the question. Nevertheless, the ever-reliable (exact) averaging laws as well as the statistical laws valid for these fields of the subtlest effects lead us by means of these fluctuations to the conviction that we must retain in the theory the assumption of a completely causal connection between events, even if we may not hope ever to attain by refined observations direct confirmation by Nature of this conception.