Contents

1. A Theory of the Foundations of Thermodynamics (1903)

Annalen der Physik 11 (1903): 170–187. [*Collected Papers* Vol. 2, Doc. 4].

Unaware of Josiah Willard Gibbs's writings at the time, Einstein here gives an alternative view on the foundations of statistical mechanics, supplementing Gibbs's better-known and more influential work and effectively introducing the concept of a canonical ensemble himself.

In a recently published paper,[a] I showed that the laws of thermal equilibrium and the concept of entropy can be derived with the help of the kinetic theory of heat. The question that then arises naturally is whether the kinetic theory is really necessary for the derivation of the above foundations of the theory of heat, or whether, perhaps, assumptions of a more general nature may suffice. In this article, it shall be demonstrated that the latter is the case, and it shall be shown by what kind of reasoning one can reach the goal.

§1. On a general mathematical representation of the processes in isolated physical systems

Let the state of some physical system that we consider be uniquely determined by very many (n) scalar quantities $p_1, p_2 \dots p_n$, which we call state variables. The change of the system in a time element dt is then determined by the changes $dp_1, dp_2 \dots dp_n$ that the state variables undergo during that time element.

Let the system be isolated, i.e., the system considered should not interact with other systems. It is then clear that the state of the system at a given instant of time uniquely determines the change of the system in the next time element dt, i.e., the quantities $dp_1, dp_2 \dots dp_n$. This statement is equivalent to a system of equations of the form

$$\frac{dp_i}{dt} = \varphi_i(p_1 \dots p_n) \qquad (i = 1 \dots i = n), \tag{1}$$

where the φ's are unique functions of their arguments.

In general, for such a system of linear differential equations, there does not exist an integral of the form

$$\psi(p_1 \ldots p_n) = \text{const.},$$

which does not contain the time explicitly. However, for a system of equations that represents the changes of a physical system closed to the outside, we must assume that at least one such equation exists, namely the energy equation

$$E(p_1 \ldots p_n) = \text{const.}$$

At the same time, we assume that no further integral of this kind that is independent of the above equation is present.

§2. On the stationary distribution of state of infinitely many isolated physical systems of almost equal energies

Experience shows that, after a certain time, an isolated system assumes a state in which no perceptible quantity of the system undergoes any further changes with time; we call this state the stationary state. Hence, it will obviously be necessary for the functions φ_i to fulfill a certain condition so that equations (1) may represent such a physical system.

If we now assume that a perceptible quantity is always represented by a time average of a certain function of the state variables $p_1 \ldots p_n$ and that these state variables $p_1 \ldots p_n$ keep on assuming the same systems of values with always the same unchanging frequency, then it necessarily follows from this condition, which we shall elevate to a postulate, that the averages of all functions of the quantities $p_1 \ldots p_n$ must be constant; hence, in accordance with the above, all perceptible quantities must also be constant.

We will specify this postulate precisely. Starting at an arbitrary point of time and throughout time T, we consider a physical system that is represented by equations (1) and has the energy E. If we imagine having chosen some arbitrary region Γ of the state variables $p_1 \ldots p_n$, then, at a given instant of time T, the values of the variables $p_1 \ldots p_n$ will lie within the chosen region Γ or outside it; hence, during a fraction of the time T, which we shall call τ, they will lie in the chosen region Γ. Our condition then reads as follows: If $p_l \ldots p_n$ are state variables of a physical system, i.e., of a system that assumes a stationary state, then, for each region Γ, the quantity $\frac{\tau}{T}$ has a definite limiting value for $T = \infty$.

For any infinitesimally small region, this limiting value is infinitesimally small.

The following consideration can be based on this postulate. Let there be very many (N) independent physical systems, all of which are represented by the same system of equations (1). We select an arbitrary instant t and inquire after the distribution of the possible states among these N systems, assuming that the energy E of all systems lies between E^* and the infinitesimally close value $E^* + \delta E^*$. From the postulate introduced above, it follows immediately that the probability that the state variables of a system randomly selected from among N systems will lie within the region Γ at time t has the value

$$\lim_{T=\infty} \frac{\tau}{T} = \text{const.}$$

The number of systems whose state variables lie within the region Γ at time t is thus

$$N \cdot \lim_{T=\infty} \frac{\tau}{T},$$

i.e., a quantity independent of time. If g denotes a region of the coordinates $p_1 \ldots p_n$ that is infinitesimally small in all variables, then the number of systems whose state variables fill up the arbitrarily chosen infinitesimally small region g at an arbitrary time will be

$$dN = \varepsilon(p_1 \ldots p_n) \int_g dp_1 \ldots dp_n. \tag{2}$$

The function ε is obtained by expressing, in symbols, the condition that the distribution of states expressed by equation (2) is a stationary one. Specifically, the region g shall be chosen such that p_1 shall lie between the definite values p_1 and $p_1 + dp_1$, p_2 between p_2 and $p_2 + dp_2, \ldots p_n$ between p_n and $p_n + dp_n$; then we have, at the time t,

$$dN_t = \varepsilon(p_1 \ldots p_n) \cdot dp_1 \cdot dp_2 \ldots dp_n,$$

where the subscript of dN denotes the time. Taking into account equation (1), one obtains furthermore, at time $t + dt$ and the same region of the state variables,

$$dN_{t+dt} = dN_t - \sum_{\nu=1}^{\nu=n} \frac{\partial(\varepsilon\varphi_\nu)}{\partial p_\nu} \cdot dp_1 \ldots dp_n \cdot dt.$$

However, since $dN_t = dN_{t+dt}$, because the distribution is stationary, we have

$$\sum \frac{\partial(\varepsilon\varphi_\nu)}{\partial p_\nu} = 0.$$

This yields

$$-\sum \frac{\partial \varphi_\nu}{\partial p_\nu} = \sum \frac{\partial(\log \varepsilon)}{\partial p_\nu} \cdot \varphi_\nu = \sum \frac{\partial(\log \varepsilon)}{\partial p_\nu} \cdot \frac{dp_\nu}{dt} = \frac{d(\log \varepsilon)}{dt},$$

where $\frac{d(\log \varepsilon)}{dt}$ denotes the change of the function $\log \varepsilon$ with respect to time for an individual system, taking into account the changes with time of the quantities p_ν. One obtains further

$$\varepsilon = e^{-\int dt \sum\limits_{\nu=1}^{\nu=n} \frac{\partial \varphi_\nu}{\partial p_\nu} + \psi(E)} = e^{-m+\psi(E)}.$$

The unknown function ψ is the time-independent integration constant which may depend on the variables $p_1 \ldots p_n$, but can contain them, according to the assumptions made in §1, only in the combination in which they appear in the energy E.

However, since $\psi(E) = \psi(E^*) =$ const. for all N systems considered, the expression for ε reduces, in our case, to

$$\varepsilon = \text{const.}\, e^{-\int dt \sum\limits_{\nu=1}^{\nu=n} \frac{\partial \varphi_\nu}{\partial p_\nu}} = \text{const.}\, e^{-m}.$$

According to the above, we now have

$$dN = \text{const.}\, e^{-m} \int_g dp_1 \ldots dp_n.$$

For the sake of simplicity, we now introduce new state variables for the system considered; they shall be denoted by π_ν. We then have

$$dN = \text{const.}\, \frac{e^{-m}}{\frac{D(\pi_1 \ldots \pi_n)}{D(p_1 \ldots p_n)}} \int_g d\pi_1 \ldots d\pi_n,$$

where the symbol D denotes the functional determinant.—We now want to choose the new coordinates such that

$$e^{-m} = \frac{D(\pi_1 \ldots \pi_n)}{D(p_1 \ldots p_n)}.$$

This equation can be satisfied in infinitely many ways, e.g., by setting

$$\begin{array}{ll} \pi_2 = p_2 & \\ \pi_3 = p_3 & \\ \quad\ldots & \pi_1 = \int e^{-m} \cdot dp_1. \\ \pi_n = p_n & \end{array}$$

Using the new variables, we thus obtain

$$dN = \text{const.} \int d\pi_1 \ldots d\pi_n.$$

Henceforth, we will always suppose that such variables have been introduced.

§3. On the distribution of states of a system in contact with a system of relatively infinitely large energy

We now assume that each of the N isolated systems is composed of two partial systems, Σ and σ, in interaction. Let the state of the partial system Σ be determined by the values of the variables $\Pi_1 \ldots \Pi_\lambda$, and that of the system σ by the values of the variables $\pi_1 \ldots \pi_l$. Further, let the energy E—which, for each system, shall lie between the values E^* and δE^*, i.e., shall equal E^* up to the infinitesimally small—be composed of two terms, of which the first, H, shall be determined only by the values of the state variables of Σ, and the second η only by the state variables of σ so that, except for the relatively infinitesimally small, one has

$$E = H + \eta.$$

Two systems in interaction that satisfy this condition will be called two systems in contact. We also assume that η is infinitesimally small compared with H.

For the number dN_1 of the N-systems whose state variables $\Pi_1 \ldots \Pi_\lambda$ and $\pi_1 \ldots \pi_l$ lie between Π_1 and $\Pi_1 + d\Pi_1$, Π_2 and $\Pi_2 + d\Pi_2$, $\ldots$ Π_λ and $\Pi_\lambda + d\Pi_\lambda$, and π_1 and $\pi_1 + d\pi_1$, π_2 and $\pi_2 + d\pi_2$, $\ldots$ π_l and $\pi_l + d\pi_l$, we get the expression

$$dN_1 = C \cdot d\Pi_1 \ldots d\Pi_\lambda \cdot d\pi_1 \ldots d\pi_l,$$

where C can be a function of $E = H + \eta$.

However, since, according to the above assumption, the energy of each of the systems considered up to the infinitesimally small has the value E^*, we can replace C by const. $e^{-2hE^*} = \text{const. } e^{-2h(H+\eta)}$ without causing any changes in the result, where h is a constant still to be defined precisely. Hence, the expression for dN_1 becomes

$$dN_1 = \text{const. } e^{-2h(H+\eta)} \cdot d\Pi_1 \ldots d\Pi_\lambda \cdot d\pi_1 \ldots d\pi_l.$$

The number of systems whose state variables π lie between the indicated limits while the values of the variables Π are not subjected to any restrictive condition may thus be represented in the form

$$dN_2 = \text{const. } e^{-2h\eta} \cdot d\pi_1 \ldots d\pi_l \int e^{-2hH} d\Pi_1 \ldots d\Pi_\lambda$$

where the integral is to be extended over all values of Π to which correspond values of the energy H lying between $E^* - \eta$ and $E^* + \delta E^* - \eta$. Had the integration been carried out, we would have found the distribution of the state of the systems σ. This is, in fact, possible.

We put

$$\int e^{-2hH} \cdot d\Pi_1 \ldots d\Pi_\lambda = \chi(E),$$

where the integral on the left-hand side is to be extended over all values of the variables for which H lies between the definite values E and $E + \delta E^*$. The integral that appears in the expression dN_2 then assumes the form

$$\chi(E^* - \eta),$$

or, since η is infinitesimally small compared with E^*,

$$\chi(E^*) - \chi'(E^*) \cdot \eta.$$

Thus, if h can be chosen such that $\chi'(E^*) = 0$, the integral reduces to a quantity that is independent of the state of σ.

It is possible to put, up to the infinitesimally small,

$$\chi(E) = e^{-2hE} \int d\Pi_1 \ldots d\Pi_\lambda = e^{-2hE} \cdot \omega(E),$$

where the integration limits are the same as above and where ω denotes a new function of E.

The condition for h now assumes the form

$$\chi'(E^*) = e^{-2hE^*} \cdot \{\omega'(E^*) - 2h\omega(E^*)\} = 0;$$

consequently:

$$h = \frac{1}{2}\frac{\omega'(E^*)}{\omega(E^*)}.$$

If h is chosen in this way, the expression for dN_2 will assume the form

$$dN_2 = \text{const. } e^{-2h\eta} d\pi_1 \ldots d\pi_l. \tag{3}$$

With suitable choice of the constant, this expression represents the probability that the state variables of a system in contact with another system of relatively infinitely large energy will lie within the indicated limits. The quantity h depends only on the state of the above system Σ of relatively infinitely large energy.

§4. On absolute temperature and thermal equilibrium

Thus, the state of the system σ depends only on the quantity h and the latter only on the state of the system Σ. We call the quantity $\frac{1}{4}h\kappa = T$ the absolute temperature of the system Σ, where κ denotes a universal constant.

If we call the system σ "thermometer," then we can immediately advance the following propositions:

1. The state of the thermometer depends only on the absolute temperature of the system Σ and not on the kind of contact of the systems Σ and σ.
2. If, in case of contact, two systems Σ_1 and Σ_2 impart the same state to a thermometer σ, then they have the same absolute temperature and will also impart the same state to another thermometer σ' in case of contact.

Further, suppose two systems Σ_1 and Σ_2 are in contact and Σ_1 is also in contact with a thermometer σ. The distribution of states of σ depends then only on the energy of the system $(\Sigma_1 + \Sigma_2)$, i.e., on the quantity $h_{1,2}$. If the interaction between Σ_1 and Σ_2 is imagined to decrease infinitely slowly, this does not change the expression for the energy $H_{1,2}$ of the system $(\Sigma_1 + \Sigma_2)$, which can be readily seen from our definition of contact and the expression for the quantity h that we formulated in the last section. Finally, if the interaction had ceased completely, the distribution of states of σ, which does not change during the separation of Σ_1 and Σ_2, will now depend on Σ_1, i.e., on the quantity h_1, where the index denotes association with the system Σ_1 alone. Hence, we have

$$h_1 = h_{12}.$$

By an analogous line of argument, one could have obtained

$$h_2 = h_{12}\,;$$

hence

$$h_1 = h_2,$$

or, in words: If one separates two systems Σ_1 and Σ_2 in contact, which form an isolated system $(\Sigma_1 + \Sigma_2)$ of absolute temperature T, then the now-isolated systems Σ_1 and Σ_2 will have the same temperature after separation. We imagine a given system in contact with an ideal gas. This gas shall be completely describable in terms of the kinetic theory of gases. As the system σ, we consider a single monoatomic gas molecule of mass μ, whose state shall be completely determined by its orthogonal coordinates x, y, z and the velocities ξ, η, ζ. In accordance with §3, we obtain, for the probability that the state variables of this molecule lie between the limits x and $x + dx$, $\ldots \zeta$ and $\zeta + d\zeta$, the well-known Maxwellian expression

$$dW = \text{const. } e^{-h\mu(\xi^2+\eta^2+\zeta^2)} \cdot dx \ldots d\zeta.$$

By integration, one obtains from this, for the mean kinetic energy of this molecule,

$$\overline{\frac{\mu}{2}(\xi^2 + \eta^2 + \zeta^2)} = \frac{1}{4h}.$$

However, the kinetic theory of gases teaches that, at a constant volume of the gas, this quantity is proportional to the pressure exerted by the gas. The latter is, by definition, proportional to the quantity designated in physics as absolute temperature. Thus, the quantity we designated as absolute temperature is nothing else but the temperature of a system measured by the gas thermometer.

§5. On infinitely slow processes

Until now, we have only considered systems that are in a stationary state. Now, we are also going to investigate changes of stationary states, though only those that proceed so slowly that the distribution of states existing at an arbitrary instant differs only infinitesimally from the stationary distribution; or, more precisely, that, up to the infinitesimally small, the probability that the state variables lie in a certain region G can be represented, at any moment, by the formula found above. We call such a change an infinitesimally slow process.

If the functions φ_ν (equation (1)) and the energy E of a system are specified, then, according to the above, its stationary state distribution is also specified. An infinitely slow process will thus be specified either by a changing E, by the functions φ_ν containing the time explicitly, or by both circumstances simultaneously, but in such a way that the corresponding differential quotients, with respect to time, are very small.

We assumed that the state variables of an isolated system change according to equations (1). However, conversely, if there exists a system of equations (1), according to which the state variables of a system are changing, this system does not always have to be an isolated one. For it can happen that a system under consideration is influenced by other systems in such a way that this influence depends only on such functions of the variable coordinates of the influencing systems, which do not change when the distribution of states of the influencing system is constant. In this case, the change of the coordinates p_ν of the system considered can also be represented by a system having the form of equations (1). However, the functions φ_ν will then depend not only on the physical nature of the system in question, but also on certain constants that are defined through the influencing systems and their distributions of states. This kind of influence on the system under consideration we call adiabatic. It is easy to see that as long as the distributions of state of the adiabatically influencing systems do not change, there exists an energy equation for the equations (1) in this case as well. If the states of the adiabatically influencing systems do change, then the functions φ_ν of the systems considered change explicitly with time, with equations (1) maintaining their validity at all times. Such a change of the distribution of states of the system under consideration we call an adiabatic one.

We now consider a second kind of changes of the state of a system Σ. Consider a system that can be influenced adiabatically. We assume that, at time $t=0$, the system Σ enters into such an interaction with a system P of a different temperature that we called "in contact" above, and we remove the system P after the time necessary for the equalization of the temperatures of Σ and P. The energy of Σ has then changed. The equations (1) of Σ are invalid during the process but valid before and after it, while the functions φ_ν are the same before and after the process. Such a process we call "isopycnic," and the energy supplied to Σ, "heat supplied."

It is evident that, up to the infinitesimally small, it is possible to construct each infinitely slow process from a succession of infinitesimally small adiabatic and isopycnic processes, so that in order to get a general overview, we have to study the latter ones only.

§6. On the concept of entropy

Let there be a physical system whose instantaneous state shall be completely determined by the values of the state variables $p_1 \cdots p_n$. Let this system undergo a small, infinitely slow process, in which the systems that influence this system adiabatically experience an infinitesimally small change of state; energy is being supplied to the system considered by systems in contact. We take account of the adiabatically influencing systems by stipulating that in addition to the $p_1 \cdots p_n$, the energy E of the system considered shall also depend on some parameters $\lambda_1, \lambda_2 \cdots$, whose values shall be determined by the distributions of states of the systems that influence adiabatically the system considered. In purely adiabatic processes, there holds, at any instant, a system of equations (1) whose functions φ_ν depend not only on the coordinates p_ν, but also on the slowly changing quantities λ; for adiabatic processes, too, there will hold, at any instant, the energy equation, whose form is

$$\sum \frac{\partial E}{\partial p_\nu} \varphi_\nu = 0.$$

We now investigate the energy increase of the system during an arbitrary infinitesimally small, infinitely slow process.

For each element dt of the process, we have

$$dE = \sum \frac{\partial E}{\partial \lambda} d\lambda + \sum \frac{\partial E}{\partial p_\nu} dp_\nu. \tag{4}$$

For an infinitesimally small isopycnic process, all $d\lambda$ vanish in each time element, and thus the first term of the right-hand side vanishes too.

However, since, according to the previous section, in an isopycnic process, dE is to be considered as heat supplied for such a process, the heat supplied dQ is represented by the expression

$$dQ = \sum \frac{\partial E}{\partial p_\nu} dp_\nu .$$

However, for an adiabatic process, during which equations (1) are always valid, we have, according to the energy equation,

$$\sum \frac{\partial E}{\partial p_\nu} dp_\nu = \sum \frac{\partial E}{\partial p_\nu} \varphi_\nu \, dt = 0.$$

On the other hand, according to the previous section, $dQ = 0$ for an adiabatic process, so that one can put

$$dQ = \sum \frac{\partial E}{\partial p_\nu} dp_\nu$$

for an adiabatic process as well. Hence, this equation must be considered as valid for any arbitrary process during each time element. Thus, equation (4) becomes

$$dE = \sum \frac{\partial E}{\partial \lambda} d\lambda + dQ. \tag{4'}$$

This expression represents the energy change of the system occurring during the whole infinitesimally small process at changed values of $d\lambda$ and dQ as well.

At the beginning and the end of the process, the distribution of states of the system considered is stationary, and when the system is in contact with a system of relatively infinitely large energy before and after the process, this assumption having formal significance only, this distribution is defined by the equation having the form

$$\begin{aligned} dW &= \text{const. } e^{-2hE} \cdot dp_1 \ldots dp_n \\ &= e^{c-2hE} \cdot dp_1 \ldots dp_n, \end{aligned}$$

where dW denotes the probability that the values of the system's state variables lie within the limits indicated at any arbitrarily chosen moment. The constant c is defined by the equation

$$\int e^{c-2hE} \cdot dp_1 \ldots dp_n = 1, \tag{5}$$

where the integration has to be extended over all values of the variables.

Specifically, if equation (5) holds before the process under consideration, then afterwards we have

$$\int e^{(c+dc)-2(h+dh)\left(E+\sum \frac{\partial E}{\partial \lambda} d\lambda\right)} \cdot dp_1 \ldots dp_n = 1, \tag{5'}$$

and the two last equations yield

$$\int \left(dc - 2Edh - 2h \sum \frac{\partial E}{\partial \lambda} \cdot d\lambda\right) \cdot e^{c-2hE} \cdot dp_1 \ldots dp_n = 0,$$

or, since the expression in parentheses can be taken as a constant during integration because the system's energy E never differs markedly from a fixed average value before and after the process, and taking into account equation (5),

$$dc - 2Edh - 2h \sum \frac{\partial E}{\partial \lambda} d\lambda = 0. \tag{5''}$$

However, according to equation (4'), we have

$$-2hdE + 2h \sum \frac{\partial E}{\partial \lambda} d\lambda + 2hdQ = 0,$$

and, by adding these two equations, one obtains

$$2h \cdot dQ = d(2hE - c),$$

or, since $\frac{1}{4h} = \kappa \cdot T$,

$$\frac{dQ}{T} = d\left(\frac{E}{T} - 2\kappa c\right) = dS.$$

This equation states that dQ/T is a total differential of a quantity that we will call the entropy S of the system. Taking into account equation (5), one obtains

$$S = 2\kappa(2hE - c) = \frac{E}{T} + 2\kappa \log \int e^{-2hE} dp_1 \ldots dp_n,$$

where the integration has to be extended over all values of the variables.

§7. On the probability of distributions of states

In order to derive the second law in its most general form, we have to investigate the probability of distributions of states. We consider a very large number (N) of isolated systems, all of which can be represented by the same system of equations (1) and whose energies coincide up to the infinitesimally small. The distribution of states of these N systems can then be represented by an equation of the form

$$dN = \varepsilon(p_1 \dots p_n, t)dp_1 \dots dp_n, \tag{2'}$$

where in general ε depends explicitly on the state variables $p_1 \dots p_n$ and also on time. Here, the function ε completely characterizes the distribution of states.

It follows from §2 that when the distribution of states is constant—which, according to our assumptions, is always the case at very large values of t—we must have $\varepsilon =$ const., so that, for a stationary distribution of states, we will have

$$dN = \text{const. } dp_1 \dots dp_n.$$

From this, it follows immediately that the expression for the probability dW, for the values of the state variables of a system randomly chosen from among the N systems to lie in the infinitesimally small region g of the state variables located within the assumed energy limits, is given by

$$dW = \text{const.} \int_g dp_1 \dots dp_n.$$

This proposition can also be formulated as follows: If the whole pertinent region of state variables that is determined by the assumed energy limits is divided into l partial regions $g_1, g_2 \dots g_l$ such that

$$\int_{g_1} = \int_{g_2} = \dots = \int_{g_l},$$

and if one denotes by W_1, W_2, etc., the probabilities that the values of the state variables of the arbitrarily chosen system lie within $g_1, g_2 \dots$ at a certain instant, then

$$W_1 = W_2 = \dots = W_l = \frac{1}{l}.$$

The probability that, at a given moment, the system considered will belong to a specific region from among these $g_1 \dots g_l$ regions is thus just as great as the probability that it will belong to any other of these regions.

The probability that, at a randomly chosen time, ε_1 of N systems considered will belong to the region g_1, ε_2 to region g_2, and $\dots \varepsilon_l$ to region g_l is hence

$$W = \left(\frac{1}{l}\right)^N \frac{N!}{\varepsilon_1!\varepsilon_2!\ldots\varepsilon_n!},$$

or also, since $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n$ are to be thought of as very large numbers:

$$\log W = \text{const.} - \sum_{\varepsilon=1}^{\varepsilon=l} \varepsilon \log \varepsilon.$$

If l is sufficiently large, one can put, without noticeable error,

$$\log W = \text{const.} - \int \varepsilon \log \varepsilon \, dp_1 \ldots dp_n.$$

In this equation, W denotes the probability that a given distribution of states, which is expressed by the numbers $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_l$ or else, by a specific function ε of $p_1 \ldots p_n$ according to equation (2'), prevails at a given time.

If, in this equation, ε were constant, i.e., independent of the p_ν's within the energy limits considered, then the distribution of states considered would be stationary and, as can easily be proved, the expression for the probability W of the distribution of states would be a maximum. If ε depends on the values of the p_ν's, then it can be shown that the expression for log W for the distribution of states considered does not have an extremum, i.e., that there exist distributions of states differing infinitesimally from the considered one for which W is larger.

If we follow the N systems considered for an arbitrary time interval, the distribution of states, and thus also W, will continually change with time, and we will have to assume that always more probable distributions of states will follow upon improbable ones, i.e., that W increases until the distribution of states has become constant and W a maximum.

It will be shown in the following sections that the second law of thermodynamics can be deduced from this proposition.

First of all, we have

$$-\int \varepsilon' \log \varepsilon' \, dp_1 \ldots dp_n \geqq -\int \varepsilon \log \varepsilon \, dp_1 \ldots dp_n,$$

where the function ε determines the distribution of states of the N systems at a certain time t, the function ε' determines the distribution of states at a certain later time t', and the integration on both sides is to be extended over all values of the variables. Further, if the quantities $\log \varepsilon$ and $\log \varepsilon'$ of the individual systems from among the N systems do not differ markedly from each other, then, since

$$\int \varepsilon \, dp_1 \ldots dp_n = \int \varepsilon' \, dp_1 \ldots dp_n = N,$$

the last equation becomes

$$-\log \varepsilon' \geqq -\log \varepsilon. \tag{6}$$

§8. Application of the results obtained to a particular case

We consider a finite number of physical systems $\sigma_1, \sigma_2 \ldots$ that together form an isolated system, which we shall call a total system. The systems $\sigma_1, \sigma_2 \ldots$ shall not interact markedly with each other thermally, but they might affect each other adiabatically. The distribution of states of each of the systems $\sigma_1, \sigma_2 \ldots$, which we shall call partial systems, shall be stationary up to the infinitesimally small. The absolute temperatures of the partial systems may be arbitrary and different from each other.

The distribution of states of the system σ_1 will not be markedly different from the distribution of states that would hold if σ_1 were in contact with a physical system of the same temperature. We can therefore represent its distribution of states by the equation

$$dw_1 = e^{c_{(1)} - 2h_{(1)} E_{(1)}} \int_g dp_1^{(1)} \ldots dp_{(n)}^{(1)},$$

where the indices (1) indicate affiliation with the partial system σ_1.

Analogous equations hold for the other partial systems. Since the instantaneous values of the state variables of the individual partial systems are independent of those of the other systems, we obtain, for the distribution of states of the total system, an equation of the form

$$dw = dw_1 \cdot dw_2 \cdots = e^{\sum c_\nu - 2h_\nu E_\nu} \int_g dp_1 \ldots dp_n, \tag{7}$$

where the summation is to be extended over all systems, and the integration over the arbitrary region g, which is infinitesimally small in all the variables of the total system.

We now assume that, after some time, the partial systems $\sigma_1, \sigma_2 \ldots$ enter into some arbitrary interaction with each other, but that during that process, the total system always remains an isolated one. After the lapse of a certain time, there shall arise a state of the total system in which the partial systems $\sigma_1, \sigma_2 \ldots$ do not affect each other thermally and, up to the infinitesimally small, exist in a stationary state.

Then, an equation completely analogous to that holding before the process will hold for the distribution of states of the total system:

$$dw' = dw'_1 \cdot dw'_2 \cdots = e^{\sum c'_\nu - 2h'_\nu E'_\nu} \int_g dp \ldots dp_n. \tag{7'}$$

We now consider N such total systems. Up to the infinitesimally small, equation (7) shall hold for each of these systems at time t, and equation (7') at time t'. Then, the distribution of states of the N total systems considered at times t and t' will be given by the equations

$$dN_t = N \cdot e^{\sum(c_\nu - 2h_\nu E_\nu)} \cdot dp_1 \ldots dp_n$$
$$dN_{t'} = N \cdot e^{\sum(c'_\nu - 2h'_\nu E'_\nu)} \cdot dp_1 \ldots dp_n.$$

To these two distributions of states, we now apply the results of the previous section. Neither the

$$\varepsilon = N \cdot e^{\sum(c_\nu - 2h_\nu E_\nu)},$$

nor the

$$\varepsilon' = N \cdot e^{\sum(c'_\nu - 2h'_\nu E'_\nu)}$$

for the individual systems among the N systems are here markedly different so that we can apply equation (6), which yields

$$\sum(2h'E' - c') \geqq \sum(2hE - c),$$

or, noting that, according to §6, the quantities $2h_1E_1 - c_1, 2h_2E_2 - c_2, \ldots$ are identical with the entropies $S_1, S_2 \ldots$ of the partial systems up to a universal constant,

$$S'_1 + S'_2 + \cdots \geqq S_1 + S_2 + \cdots, \tag{8}$$

i.e., the sum of the entropies of the partial systems of an isolated system, after some arbitrary process, is equal to or larger than the sum of the entropies of the partial systems before the process.

§9. Derivation of the second law

Let there be an isolated total system whose partial systems shall be called W, M, and $\Sigma_1, \Sigma_2 \ldots$ Let the system W, which we shall call heat reservoir, have an energy that is infinitely large compared with the system M (engine). Similarly, the energy of the systems $\Sigma_1, \Sigma_2 \ldots$, which interact adiabatically with each other, shall be infinitely large compared with that of M. We assume that all the partial systems M, W, Σ_1, $\Sigma_2 \ldots$ are in a stationary state.

Suppose that the engine M passes through a cyclic process, during which it changes the distributions of states of the systems $\Sigma_1, \Sigma_2 \ldots$ infinitely slowly through adiabatic influence, i.e., performs work and receives the amount of heat Q from the system W. The reciprocal adiabatic influence of the systems $\Sigma_1, \Sigma_2 \ldots$ at the end of the process will then differ from that before the process. We say that the engine M has converted the amount of heat Q into work.

We now calculate the increase in entropy of the individual partial systems during the process considered. According to the results of §6, the entropy increase of the heat reservoir W equals $-\frac{Q}{T}$ if T denotes the absolute temperature. The entropy of M is the same before and after the process because the system M has undergone a cyclic process. The systems $\Sigma_1, \Sigma_2 \ldots$ do not change their entropies during the process at all because these systems only experience an adiabatic influence that is infinitely slow. Hence, the entropy increase $S' - S$ of the total system has the value

$$S' - S = \frac{Q}{T}.$$

Since, according to the results of the last section, this quantity $S' - S$ is always $\geqq 0$, it follows that

$$Q \leqq 0.$$

This equation expresses the impossibility of the existence of a perpetuum mobile of the second kind.

Bern, January 1903.

(Received on 26 January 1903.)

Editorial Note

[a] Einstein 1902.

Index