Concerning the Statistics of Classical Systems
by Arnulf Schlüter, Max Planck Institute for Physics, Gottingen
(note 1: excerpted from the first part of the author’s dissertation)
Z. Naturforschg. 3a, 350-360 (1948)
Received on 8 December 1947
[paragraph numbers in brackets for ease of locating material in the original]

[1] We first define the phase volume of a system as the volume contained within the energy surface in the phase space and define its phase density as its derivative with respect to energy. Because of the ergodic hypothesis, the system phase density can be obtained as the convolution product of the phase densities of the individual parts; and these can again, when the energy of the parts can be separated, be depicted as convolution products of the energy contributions of the assigned phase densities.

[2] Certain, especially important parts of these results can be carried out explicitly. These will require some distribution problems to be solved. In particular, the exact validity of the equipartition theorem will be proved. Out of that and the isotropic invariance of the phase volume, we conclude that the thermodynamic differential equations are strictly adhered to if the entropy is set proportional to the phase volume and the temperature is set inversely proportional to its logarithmic derivative. These results apply for a system of arbitrarily few particles.

Phase Space

[3] Consider a system obeying the laws of the Classical Mechanics. We imagine that this system is composed of many individual particles, but we won’t yet make use of this assumption. The instantaneous state (microstate) of the system is defined by as many generalized coordinates, \( Q \), and their canonically conjugate momenta, \( P \), as the number of degrees of freedom (note 2: upper case characters indicate vector quantities representing variables describing the whole system). The momenta and coordinates are thereby interpreted out of Cartesian coordinates into a symbolic many dimensional “phase space.” (Ehrenfest’s \( \Gamma \)-phase space)

[4] Only systems that will not dissociate in time should be investigated. This is assured if we require that the following integral exists

\[
J(E_0) = \int \ldots \int dP dQ,
\]

where \( P, Q \) represent all coordinates in phase space, and \( E(P, Q) \) is the total energy as a function of location in phase space.

[5] \( J(E_0) \) measures the volume of the system phase space with any energy smaller than \( E_0 \) and is called the phase volume. \( J(E_0) \) is canonically invariant, meaning it is independent of the choice of generalized coordinates.

[6] A mechanical system such as this is called an ergodic system if, at constant total energy, the time average of an arbitrary, not explicitly time-dependent function of \( P, Q \) over adequately long time is equal to the phase space average of this function for this energy, in the sense, that

\[
\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \int F(P(t), Q(t)) dt = \int \ldots \int \frac{F(P, Q) dP dQ}{E_0 < E(P, Q) < E_0 + dE_0}.
\]

Eq. (1)
where the last quotient can also be written as:

$$\frac{dE_0}{dE_0} \frac{d}{dE_0} \int \ldots \int \frac{F(P, Q) dP dQ}{G(E_0)} = \frac{1}{G(E_0)} \frac{d}{dE_0} \int \ldots \int F dP dQ,$$

Eq. (1a)

where the derivative of the phase volume with respect to energy is abbreviated as $G(E)$, which we will call the phase density of the system being considered.

[7] The dependability of this average calculation is as good as the quasiergodic hypothesis in Ehrenfest’s formulation. It would be sufficient if the interchangeability requirements for the average were only required for “almost all” initial conditions. In other words, a different relationship could be allowed for a set of initial conditions of lesser dimensionality (For example, for simple periodical movements that originate through accidental degeneracy). The additional logical consequences would then also be valid only “almost always.”

[8] In such an ergodic system, the time average of any function of the microstate is a well-defined function of the energy of the system. This doesn’t work for other quantities (it is not true for angular momentum, for example,) and means that the ergodic hypothesis, even in our cautious formulation, is surely not strong enough. However, we won’t go into these difficulties further, but rather simply postulate the validity of Eq. (1).

[9] If a system consists of multiple parts which are almost always independent of each other in the sense that the total energy of the system can be portrayed as the sum of the energies of the individual parts at any moment in time and the times during which the parts interact can be ignored, then these parts of the system can be said to comprise the system.

[10] Correspondingly, the state of a subsystem is described by a point in its phase space. It possesses in that instant a definite energy and, thereby, a corresponding definite phase volume and a definite phase density.

[11] We consider now a separation of the system into two different subsystems ($I$ and $II$). We inquire, after a relatively long time, in which the subsystem $I$ has an energy between $E_{I0}$ and $E_{I0} + dE_0$(note 3. Upper case letters with Roman numeral indices represent large subsystems). According to the ergodic hypothesis, instead of this time average, the corresponding average in the phase space could be calculated by:

$$W_I(E_{I0}) dE = \int \ldots \int \frac{dP_I dQ_I}{G_I(E_{I0})G_{II}(E_{II0})} + \int \ldots \int \frac{dP_{II} dQ_{II}}{G_{II0} G(E)} ,$$

Whereby the numerator originates from the fact that, as long as the subsystem $I$ has the energy $E_I$, the complementary subsystem $II$ must have the energy $E_{II}$, with $E_I + E_{II} = E$ being constant.

[12] One can bring the phase densities back in again here and get (by leaving the Index 0 out):
The function $W_E$ is the “energy distribution function” (of subsystem $I$). With its help, the average of a function of the energy of the subsystem $I$ can be immediately specified as:

$$F_I(E) = \int_{-\infty}^{+\infty} F_I(E)W_I(E)dE_I = \frac{\int_{-\infty}^{+\infty} F_I(E)G_II(E-E_I)dE_I}{G(E)}$$

The average value is (through $W_I$) a function of the total energy, $E$.

**The Convolution Method**

Using the definition of the energy distribution function, its integral must, when extended over all possible energies, be normalized:

$$\int_{-\infty}^{+\infty} W_I(E)dE_I = 1,$$

that means

$$\int_{-\infty}^{+\infty} G_I(E_I)G_II(E-E_I)dE_I = G(E). \quad \text{Eq. (4)}$$

Equation (4) clearly states that the phase density of the whole system, $G(E)$, can be obtained through integration over the energy of the product of the interrelated phase densities of the subsystems. This can be seen as a direct consequence of the ergodic hypothesis, through which the equiprobability of similar size phase regions consistent with the energy requirement was postulated.

An integral of the form of Eq. (4) is designated as a *convolution integral* and the symbol for it is a star.

$$G_I(E) * G_II(E) = \int_{-\infty}^{+\infty} G_I(E_I)G_II(E-E_I)dE_I = \int_{-\infty}^{+\infty} G_I(E-E_{II})G_II(E_{II})dE_{II} = G_II(E) * G_I(E)$$

on which product (“convolution product”) a special notation is justified, as the operation is commutative, and also, as a linear operation, distributive with respect to addition. One can show (note 4: simply, the Laplace transform of a convolution is equal to the usual product of the Laplace transforms of the separate factors, see G. Doetsch, *Theory and Application of the Laplace Transform*, Berlin, 1937, 8, 155ff.) that multiple operations are also associative:

$$(G_I * G_II) * G_III = G_I * (G_II * G_III).$$

Equation (4) is also especially applicable when one considers a system which is composed of only two single particles with phase densities $g_1(u)$ and $g_2(u)$ (note 5: to distinguish them from whole system variables, those describing individual particles are written in lower case characters with Arabic number subscripts), which are almost always independent of each other (in the sense defined above for subsystems):

$$G(E) = \int_{-\infty}^{+\infty} g_1(u)g_2(E-u)du = g_1(E) * g_2(E).$$

If the system consists in similar manner of 3 particles, then 2 particles can be viewed together as one subsystem of the system, whose phase density is generated through the convolution of the phase density of both particles. The second convolution with the phase density of the third particle produces the phase density of the entire system. Because of the
associated laws of convolution, the outcome is independent of the alternative order of creation of the subsystems, as required physically.

\[ G(E) = g_1(E) * g_2(E) * g_3(E). \]

So we can naturally proceed with additional particles, so that the following is the phase density \( G(E) \) of an ergodic system composed of \( N \) particles which are almost always independent. [with phase densities \( g_i(u) \):]

\[ G(E) = g_1(E) * ... * g_i(E) * ... * g_N(E), \]

which can be abbreviated:

\[ G(E) = \prod_i g_i(E); \quad i = 1, ..., N. \quad \text{Eq. (5)} \]

(Note 6: the index \( i \) runs over all particles in the system)

[19] The requirements for the validity of Eq. (5) are the ergodic character of the system and the capability of the total energy to be separated into energies which are only dependent on the position and momentum coordinates of each particle. If energies exist, for example because of Coulomb or van der Waals forces between the particles, which do not depend solely on the positions of the individual pieces, this possibility no longer exists. In the following, only systems should be considered which (almost always) exist as separate, independent pieces so that the convolution method applies. (Note 7: concerning the number of particles, it is previously assumed that a system contains at least one particle. Likewise when a partition of a system into two parts is discussed, it must contain at least two particles to avoid an empty partition).

[20] Often we encounter the situation where the energy of individual particles can also be separated into particle energies depending on only one position and momentum coordinate. If the energy of a particle is completely separable, then it can also be written:

\[ u(p,q) = \sum_k v_k(q_k) + t_k(p_k); \quad k = 1, ..., f \]

(\( v_k = \) components of the potential energy; \( t_k = \) components of the kinetic energy; \( f = \) number of degrees of freedom, note 8: the index \( k \) runs over individual particle degrees of freedom).

[21] Every energy contribution can then have a phase volume and a phase density assigned to it again, and we can construct the phase density of the entire system again out of the individual “phase densities.”
The meaning of this possibility is that the phase volumes \( j_k(v_k) \) or \( j_k(t_k) \) can be obtained easily, as long as the energy functions \( v_k(q_k) \) or \( t_k(p_k) \) are known. The phase volume, \( j_k(v_{k0}) \), for instance, measures the length of the interval of \( q_k \), for which \( v_k(q_k) < v_{k0} \) (see Fig. 1). It is then simply

\[
j_k(v_{k0}) = q_k^{(2)}(v_{k0}) - q_k^{(1)}(v_{k0}) + q_k^{(4)}(v_{k0}) - q_k^{(3)}(v_{k0}) + \ldots \quad \text{Eq. (6)}
\]

if \( q_k(v_{k0}) \) represent the values of ordered roots of the equation:

\[
v_k(q_k) = v_{k0}
\]

In all interesting cases, this equation has no roots below a minimum energy, above which it has exactly two roots. (Illustration 1 thus depicts a general situation, which doesn't create any particular difficulties.)

Differentiation with respect to the energy gives the phase density and, through convolution, of all phase densities (including those which belong to the kinetic energy portions, which can be displayed accordingly), one gets the phase density of all the particles.

\[
g(u) = \prod_{k=1}^{f} \left[ \frac{(2) dq_k(u)}{du} - \frac{(1) dq_k(u)}{du} \right] \ast \left[ \frac{(2) dp_k(u)}{du} - \frac{(1) dp_k(u)}{du} \right] \quad \text{Eq. (7)}
\]

or written differently as:

\[
g(u) = \prod_{k=1}^{f} \left[ \frac{1}{\partial q_k(\partial u)} - \frac{1}{\partial q_k(\partial u)} \right] \ast \left[ \frac{1}{\partial p_k(\partial u)} - \frac{1}{\partial p_k(\partial u)} \right]
\]

(1) \( q_k = q_k^{(1)}(u), p_k = p_k^{(1)}(u) \)

(2) etc. in like manner

\[
\quad \text{Eq. (7a)}
\]
In these equations, it is especially important to note that the convolution product is written with
the result variable, which is a parameter of the integral, as argument instead of using the
integration variable. [compare equations (4a) and (5)].

“Classical” Particles

[24] In many models, the individual energy contributions from particles with separable energy
have similar dependence on the associated coordinates. For example, the kinetic energy is
always a homogeneous quadratic function of the momenta (note 9: this assumes a nonrelativistic
classical treatment), it can therefore be portrayed as a sum of parts which are proportional to the
square of the momentum components. Similarly, the potential energy of harmonic oscillators is
a quadratic function of the position coordinates, while the potential energy in a homogeneous
gravitational field referenced to a lower edge (the floor of the container), at which it can be said
to be zero, is proportional to a position coordinate. The energy function of a degree of freedom
can be brought about in this and also essentially general cases through suitable choice of the
constants $a_k$ and $c_k$ in one of the two forms:

$$v_k(q_k) = 2\frac{\Gamma(a_k + 1)}{c_k}q_k^{\frac{1}{a_k}} \quad \text{for} \quad -\infty < q_k < \infty \quad \text{and} \quad a_k^{-1} > 0 \quad \text{and even}, \quad \text{Eq. (8a)}$$
or

$$v_k(q_k) = \begin{cases} \frac{\Gamma(a_k + 1)}{c_k}q_k^{\frac{1}{a_k}} & \text{for} \quad q_k \geq 0 \\ \infty & \text{for} \quad q_k < 0 \end{cases} \quad \text{for any value of} \quad \frac{1}{a_k} > 0 , \quad \text{Eq. (8b)}$$

where all $a_k, c_k > 0$, and $\Gamma$ represents the gamma function. For the kinetic energy the
appropriately specialized form (8a) is sufficient:

$$t_k(p_k) = \frac{\pi}{c_k}p_k^2 ; \quad k = 1, \ldots, f , \quad c_k' > 0. \quad \text{Eq. (9)}$$

[25] The use of Eq. (6) immediately gives the phase volume in agreement with (8a) and (8b):

$$j_k(v_k) = \frac{c_k}{\Gamma(a_k + 1)}0v^{a_k}_k ; \quad a_k, c_k > 0 \quad \text{Eq. (10)}$$

and similarly for the kinetic energy portions (9):

$$j_k(t_k) = \frac{2}{\sqrt{\pi}}c_k'0t^{1/2}_k ; \quad c_k' > 0 ; \quad k = 1, \ldots, f ; \quad \Gamma\left(\frac{3}{2}\right) = \frac{1}{2\sqrt{\pi}} \quad \text{Eq. (11)}$$

[26] By this, as also in the following, the prefixed Null (0) means that the function distribution
should be replaced with zero for negative arguments:

$$0f(x) = \begin{cases} f(x) & \text{for} \quad x \geq 0 \\ 0 & \text{for} \quad x < 0 \end{cases}. \quad \text{Eq. (10a)}$$

[27] By differentiating the phase volumes, one gets the corresponding phase densities.

$$g_k(v_k) = \frac{c_k}{\Gamma(a_k)}0v^{a_k-1}_k ; \quad a_k, c_k > 0 \quad \text{Eq. (10a)}$$
Using these forms of the single-particle phase densities, the convolutions for the assembly of the phase density of the collection of particles can be carried out with the help of the convolution rules

\[
\frac{c_1}{\Gamma(a_1)} u^{a_1-1} * \frac{c_2}{\Gamma(a_2)} u^{a_2-1} = \frac{c_1 c_2}{\Gamma(a_1 + a_2)} u^{a_1 + a_2-1}
\]

Eq. (12)

(note 10: This formula follows by easily verified explicit integration using elementary integrals. It is especially elegant, however, to prove with help of Laplace transformation that

\[
L \left\{ \frac{c}{\Gamma(a)} u^a \right\} = \int_{-\infty}^{\infty} e^{-u\beta} \frac{c}{\Gamma(a)} u^a du = c \beta^{-a}
\]

and the transformation of formula (12) simply states \( c_1 \beta^{-a_1} c_2 \beta^{-a_2} = c_1 c_2 \beta^{-(a_1 + a_2)} \).)

The convolution product of “one-sided” powers is thus such that its analytical form, according to our definition of the constants, depends very simply on the parameters of the factors.

One obtains the overall density of an entire collection of such particles with:

\[
g(u) = \frac{c}{\Gamma(a)} u^{a-1}
\]

Eq. (13)

with

\[
c = \prod_{k=1}^{f} c_k c_k' \quad ; \quad a = \sum_{k=1}^{f} a_k + \frac{f}{2} \quad (f = \text{number of mechanical degrees of freedom})
\]

and through integration of the energy of the phase volumes:

\[
j(u) = \frac{c}{\Gamma(a+1)} u^{a}
\]

Eq. (13a)

The form (13) or (13a) also remains preserved if some or all position coordinates bring no contribution to the potential energy, but getting from them instead simply a boundary which depends on the area available to the particle. The size of this area comes in as a factor in the phase density and phase volumes, contributing thus to the product of the \( c_k \) for the corresponding degrees of freedom, while the respective \( a_k \) must be set to null.

Particles with a phase density and phase volume of the form of equation (13) or (13a) will be called “classical particles” here (note 11: without assuming separability of the energy). We will carry “classical” in quotes because the applicability of classical mechanics is not sufficient for the validity of (13) and (13a); we will show, however, that most (in the common sense) classical models actually correspond to this type and that the “classical” form of phase volume and density is necessary (note 12: see dissertation, section B 12 (compare note 1)) and sufficient for the equipartition principle for the energy.

The exponent \( a \) of the energy in the phase volume function should be named the characteristic number of the corresponding particle, as this characteristic number is crucial for the statistical behavior of the particle. According to Eq. (13), the characteristic number is equal to the sum of the potential energy contribution plus half of the number of mechanical degrees of
freedom of the particle.

**Examples of “Classical” particles**

[34] *Ideal gases.* In an ideal gas without internal degrees of freedom, the energy of every particle is only dependent on its momentum:

\[ u = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) \]

[35] This is included in the “classical” form (9), if one sets

\[ c_x' = c_y' = c_z' = \sqrt{\frac{2m}{\pi}} \]

[36] The position coordinates simply provide the volume, \( V \), which is occupied by the gas, as a factor, so that one altogether gets the following as phase density and phase volume of an ideal gas:

\[
\begin{align*}
g(u) &= 2\pi(2m)^{\frac{3}{2}} V_0 u^{\frac{3}{2}} \quad \text{Eq. (14)} \\
j(u) &= \frac{4}{3} \pi(2m)^{\frac{3}{2}} V_0 u^{\frac{3}{2}} \quad \text{Eq. (14a)}
\end{align*}
\]

[37] *Sedimentation/layering equilibrium.* If the ideal gas is in a homogeneous gravitational field, then the energy of the particles is given by:

\[ u = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + m\gamma z \]

\( (\gamma = \text{acceleration of free fall}, \ z = \text{height, with up being chosen positive}) \).

[38] If the container has height \( z = 0 \) as floor and has, from just above that to infinite height, a constant cross section \( Q \), then the \( z \) dependence of the potential energy has the form (8b) with

\[ a_z = 1; \ c_z = \frac{1}{\gamma} \]

while the horizontal coordinates only factor is the cross section \( Q \). This results in phase density and volume:

\[
\begin{align*}
g(u) &= \frac{8}{3} \pi(2m)^{\frac{3}{2}} \gamma^{-1} Q_0 u^{\frac{5}{2}} \quad \text{Eq. (15)} \\
j(u) &= \frac{16}{15} \pi(2m)^{\frac{3}{2}} \gamma^{-1} Q_0 u^{\frac{5}{2}} \quad \text{Eq. (15a)}
\end{align*}
\]

The potential energy thus causes an increase of the characteristic number \( a \) of the particle from \( 3/2 \) to \( 5/2 \).

[39] Also in the general case where the cross section of the container has the form

\[ Q(z) = \alpha^b; \ b \geq 0 \]

\( (b \ not \ necessarily \ an \ integer) \) depending on the height \( z \), giving the “classical” phase density of the particle; we get the following distinct value for \( a \)

\[ a = \frac{5}{2} + b \]

(note 13: the occasional use of \( 2a \) as the number of statistical degrees of freedom leads to contradictions. The relationship \( 2a = f + 2\Sigma a_k \) give the correct number of degrees of freedom including potential energy.)
[40] **Linear harmonic oscillator.** The energy of the one-dimensional harmonic oscillator is

\[ u = \frac{1}{2m} p^2 + \frac{1}{2} a^2 q^2, \]

If the elastic constant \( a^2 \) is used. The dependence of the energy on momentum and position is of the same kind and gives the phase density

\[ g(u) = 2\pi a^{-1} \sqrt{m_0} u^0 \quad \left( u^0 = \{1 \text{ for } u \geq 0 \text{ and } 0 \text{ for } u \leq 0\} \right) \quad \text{Eq. (16)} \]

and the corresponding phase volume

\[ j(u) = 2\pi a^{-1} \sqrt{m_0} u \quad \text{Eq. (16a)} \]

**Statistics of “Classical” Systems**

[41] If a system consists of simple “classical” particles, we will call it a “classical system.” Then its phase density can be calculated through repeated use of the convolution theorem (12), and the phase density of the system has the same “classical” form:

\[ G(E) = \frac{C}{\Gamma(A)} E^{A-1}; \quad A, C > 0 \quad \text{Eq. (13')} \]

\[ C = \prod_i c_i; \quad A = \sum_i a_i; \quad i = 1, \ldots, N; \quad N > 1. \]

Accordingly, for the phase volume we get the following:

\[ J(E) = \frac{C}{\Gamma(A + 1)} E^A. \quad \text{Eq. (13a')} \]

[42] These convenient analytical formulas for phase volumes and densities permit the explicit evaluation of various distribution problems. For instance, one wants to know the distribution function for the energy of a particle with the phase density

\[ g(u) = \frac{c}{\Gamma(a)} u^{a-1} \]

in such a “classical” system. Equation (2) gives for this, after appropriate changes, the expression:

\[ w(u) = g(u) \frac{G_R(E-u)}{G(E)}, \]

where \( G_R \) designates the phase density of the remaining part of the system without the particle. In this formula, \( g(u) \) is the a priori distribution while the quotient gives the influence of the remaining pieces on the energy distribution. Substitution of the “classical” expressions gives

\[ w(u) = \frac{c}{\Gamma(a)} u^{a-1} * \frac{\Gamma(A)(E-u)^{A-a-1}}{\Gamma(A-a)cE^{A-1}}; \quad 0 < u < E. \quad \text{Eq. (17)} \]

[43] If the system consists of very many (“infinitely” many) particles, and the average energy per particle is finite, in other words when \( A \) and \( E \) both become infinite at the same rate, this distribution function asymptotically approached the expression

\[ \lim_{A,E \to \infty} w(u) = \frac{c}{\Gamma(a)} u^{a-1} \frac{e^{-uA/E}}{c(E/A)^u}; \quad u > 0. \]

This is, however, the Maxwell-Boltzmann distribution function, if one replaces \( E/A \) with \( kT \), which we will justify below.
The difference between the exact distribution (17) and the Boltzmann distribution gives the enlightening conclusion that no particle can have an energy larger than the total energy of the system \[w(u) \rightarrow 0 \text{ for } u \rightarrow E\].

The average one-particle energy can be obtained through elementary integration (note 14: the simplest evaluation of these is a reverse convolution using Eq. 12):

\[
\bar{u} = \frac{\bar{u}(u)}{a} \frac{A}{E} \left( = \frac{J(E)}{G(E)} \right).
\]

Equation (18) proves that the average energy of every particle is independently proportional to the total energy, and, in fact, this ratio is the same as the relationship of the characteristic number of the particle (namely: \(a\)) to the system \((A = \Sigma a)\). This is the classical equipartition theorem for energy, which is also strictly true, then, in systems with arbitrary numbers of particles as long as these are “classical.”

Equation (18) presents the correct generalization for considering the potential energy (2\(a\) = “potential” + “kinetic” degrees of freedom, compare note 13), contrasting to the usual formulation.

Likewise, the analysis of the integral:

\[
\langle u^2 \rangle = \int_{-\infty}^{\infty} u^2 w(u) du = \frac{A(a+1)}{(A+1)a} \frac{\bar{u}^2}{a} \xrightarrow{A \rightarrow \infty} \frac{a+1}{a} \bar{u}^2
\]
gives the average squared energy, and out of that follows the variance:

\[
\langle (\Delta u)^2 \rangle = \langle u^2 \rangle - \bar{u}^2 = \frac{A-a}{(A+1)a} \frac{\bar{u}^2}{a} \xrightarrow{A \rightarrow \infty} \frac{1}{a} \bar{u}^2
\]

The most likely energy of a particle is the location, \(u = \hat{u}\), for which \(w(u)\) is at a maximum, which results in:

\[
\hat{u} = \frac{A(a-1)}{(A-2)a} \bar{u} \quad \text{for } a-1 \geq 0; \quad A-2 > 0;
\]

\[
\hat{u} = 0 \quad \text{for } a-1 \leq 0; \quad A-a-1 > 0.
\]

\[
\hat{u} \rightarrow \frac{a-1}{a} \bar{u} \quad \text{for } a-1 \geq 0; \quad A \rightarrow \infty.
\]

According to the derivation, the preceding formula is only dependent on the analytical form of the single-particle phase density and that of the system (both must have the “classical” form), but not dependent on the single-particle phase density belonging to the entire system. The formulas are also valid when they are based on individually separable energy contributions, such as the potential energy of a gas particle in a homogeneous force field. Equation (17) then gives the distribution of the potential energy, and that gives the height (and energy) dependent average density of the gas.

Likewise these formulas produce the distribution, the average and so forth, of the energy of a subsystem consisting of multiple particles if one sets their phase density (given by the convolution product of the phase density of its particles)—which is its characteristic number (which is the sum of the characteristic numbers of its particles)—in the place of the
corresponding single-particle density.

**Thermodynamics of “classical” systems**

[52] The energy of a system in a certain (micro) state described by a point in the phase space of the system, or in other words described by specifying all $P, Q$ (where again the pair of variables represent the set), still depends on external parameters, which we have previously presumed to be constants. Such parameters are like the strength of the gravitation field for sedimentation equilibrium or the volume occupied by the particles of a gas. If the function $E = E(P, Q)$ still depends on the parameter $x_r$, then the phase volume and density are dependent on these, so that at length what must be written is (note 15: we write simply $x$ to stand for all $x_r$):

$$E = E(P, Q, x),$$

$$J = J(E; x), \quad G = G(E; x) = \frac{\partial J(E; x)}{\partial E}.$$

[53] The force, $X_r$, which follows from the parameter, $x_r$, depends on the respective values of the $P, Q$, and all $x$ from:

$$X_r(P, Q; x) = -\frac{\partial}{\partial x_r} E(P, Q, x).$$

[54] We get the following as energy balance for an infinitesimal parameter displacement:

$$0 = dE_{adibat} + \sum_r X_r(P, Q, x) dx_r = dE_{adibat} - \sum_r \frac{\partial E(P, Q, x)}{\partial x_r} dx_r.$$

[55] This relation defines when a parameter change may be considered an “adiabatic process.” This is in contrast to a general process where the total energy of the system is changed, not only through work done by external parameters, but also separately, perhaps, by combining for a time the ergodic system being considered with another body, in other words, that is by bringing it into thermal equilibrium with another body. The energy increase of the system from this contact is called the “heat transfer,” $Q$ (not to be confused with the symbol for position coordinates), and this gives for a general infinitesimal process:

$$dQ = dE + \sum_r X_r dx_r.$$  \hspace{1cm} \text{(First Law of Thermodynamics)}

[56] The statement of the **First Law**, or the energy conservation theorem, is because $dQ$ for the system being considered and that of the body with which the regarded system is brought into contact together total to zero.

[57] Especially important are processes in which the parameter displacement is done so slowly that, in calculating the energy increase, $-\sum X_r dx_r$, the external force can be replaced by its time average. Such processes may be called *quasi-static*.

[58] Because of the ergodic hypothesis [Eq. (1a)], the time average of an external force can be written:

$$\bar{X}_r(P, Q, x) = \bar{X}_r(E_0; x) = -\frac{1}{G(E_0; x)} \frac{\partial}{\partial x_r} \int\int \frac{\partial E(P, Q, x)}{\partial x_r} dPdQ$$

$$= \frac{1}{G(E_0; x)} \frac{\partial}{\partial x_r} \int dPdQ = \frac{1}{G(E_0; x)} \frac{\partial J(E_0; x)}{\partial x_r}.$$
Then for the heat applied during a quasi-static process the following applies:

\[ dQ_{\text{quasist}} = dE_{\text{quasist}} + \sum_r X_r dx_r = dE_{\text{quasist}} + \frac{1}{G} \sum_r \frac{\partial J}{\partial x_r} dx_r. \]  

Eq. (19a)

A process which is simultaneously adiabatic and quasi-static, which, in other words, only exists in an endlessly slow parameter displacement, is called “isentropic”. For an infinitesimal isentropic process the energy balance is then applicable:

\[ 0 = dE_{\text{isentrop}} + \frac{1}{G} \sum_r \frac{\partial J}{\partial x_r} dx_r. \]  

(transl. note: referred to as Eq. 19b below)

This equation is very valuable because it permits one to calculate the energy increase of the system without having to know how the external force parameters depend on the microstate, or the function \( X_r(P, Q; x) \), because only the derivatives of the “macroscopic” function \( J(E; x) \) enter.

Because

\[ dJ(E; x) = \frac{\partial J}{\partial E} dE + \sum_r \frac{\partial J}{\partial x_r} dx_r = \left[ dE + \frac{1}{G} \sum_r \frac{\partial J}{\partial x_r} dx_r \right] \]  

Eq. (20)

is equivalent to Eq. (19b) with the claim of the isentropic invariance of the phase volume:

\[ dJ_{\text{isentrop}} = 0. \]

(note 16: In the literature, usually one speaks of the adiabatic invariance of phase volumes, which is here reminiscent of rapidly completed adiabatic processes of macroscopic thermodynamics, whereas here the infinitely slow process is completely essential. Smekal has therefore proposed simply referring to “parameter invariance”; however, this is obviously not substantially better. Therefore, we introduced here an isentropic process through the definition of entropy, analogous to the customary introduction of an adiabatic process or an adiabatic change through the definition of heat.)

We can now proceed to define the absolute temperature exactly as it is done in macroscopic thermodynamics, first by Helmholtz (note 17: H. v. Helmholtz, J. Math. 97, 111,317 (1884)) and also by Caratheodory (note 18: Caratheodory, Math. Ann. 61, 355 (1909)). The absolute temperature is defined as the integrated denominator of the quasi-static (note 19: Tatyana Ehrenfest-Afanasyeva has shown that it is the quasistatic rather than the reversible nature of the heat transfer upon which this depends (Z. Physik 33, 933 (1925), 34, 638 (1925)) applied heat, which possesses the attributes of an empirical temperature. Thus, an empirical temperature is one in which the energy is dependent monotonically on the state variable and is the same for all bodies in thermal equilibrium.

An integrating factor for the quasi-static added heat is, for example, \( G(E) \), which according to (19a) and (20) is:

\[ G(E) dQ_{\text{quasist}} = dJ \]

In general, one more arbitrary (integrable) function, \( f(J) \), can be applied to \( G(E) \), because

\[ f(J) G(E) dQ_{\text{quasist}} = f(J) dJ = d \int f(J) dJ \]

likewise produces a complete differential. (Transl. note: \( f(J) = J^1 \) below)

For “classical” systems, which we now return to—the preceding implementations of this section are valid independent of the special dependence of the phase volume on the energy—we can also immediately define an empirical temperature: The average energy per characteristic number is the same for all particles and also the same for all subsystems of a “classical” system,
and specifically they are the same as the quotients from the phase volumes and phase densities (18):

\[
\frac{u_1}{a_1} = \frac{u_2}{a_2} = \ldots = \frac{E_I}{A_I} = \frac{E_{II}}{A_{II}} = \ldots = \frac{E}{A} = \frac{J(E)}{G(E)}. 
\]

Eq. (18a)

Because \(J/G\) is also an integrable denominator of the quasi-static added heat,

\[
\frac{J(E)}{G(E)} = kT 
\]

Eq. (21)

is the absolute temperature, whereby the initially undefined factor \(k\) can be identified through an example, for example an ideal gas, with the Boltzmann constant.

[66] Differentiation of the energy with respect to the temperature gives the specific heat or respectively the heat capacity of the system:

\[
\frac{du}{dT} = C_v = kA, 
\]

which is, except for the factor \(k\), equal to the characteristic number of the system.

[67] Through the temperature, the entropy differential is determined at the same time:

\[
dS = \frac{dQ_{\text{quasistat}}}{T} = kdJ = kd\ln J, 
\]

Eq. (22)

\[
\frac{S - S_0}{k} = \ln J = \ln C + A\ln E - \ln \Gamma(A + 1); \quad E > 0.
\]

If one substitutes the temperature as the variable instead of the energy, we get:

\[
\frac{S - S_0}{k} = \ln C + A\ln kT + A\ln A - \ln \Gamma(A + 1). 
\]

Eq. (22a)

Using Stirling’s approximation, we find in the limit that \(A \to \infty\), if one simultaneously eliminates the index \(A\) through the heat capacity:

\[
S - S_0 \quad \underbrace{\quad \frac{1}{A \to \infty} k\ln C + C_v\ln kT + C_v + \frac{k}{2}\ln 2\pi A}_{\text{transl. note: } k \text{ before the bracketed last term in Eq. (22b) is not in the original}}.
\]

Eq. (22b)

The bracketed element is proportionally insignificant compared to what remains and can be left out. For an ideal gas, (22b) now agrees with the known entropy formula. The remaining thermodynamic functions can be constructed out of entropy and temperature.

[68] After our introduction of the entropy, one arbitrary additive constant \((S_0)\) remains, which, for example, could still depend on the particle number. This indeterminacy is totally appropriate, because in macroscopic thermodynamics, the entropy constant is first calculated through the examination of phase or chemical equilibria, which obviously can’t be developed through our model.

[69] As a result of the entropy formula (22a), a slight entropy correction emerges if two systems of same temperature are brought into thermal contact. In the approximation formula (22b), this expresses itself in the non-linearity of the bracketed element. The effect disappears thus with increased particle number. The entropy increase contains the impossibility of being able to combine systems quasi-statically because that would require the ergodic character of every system to each other be maintained. The combination is accordingly also irreversible, because of the separation an energy dispersion occurs, as is calculated above (p. 357).
Conclusion

[70] The goal of this study is not to refine the usual statistical formulas for systems consisting of few particles. (Deviations in the derived formulas, for example, of the energy distribution (17) or the entropy formulas (22, 22a), disappear for large numbers of particles and are independent of the strict validity of the ergodic hypothesis.) It has rather been shown that a sequence of required postulates for statistical justification of thermodynamics is superfluous: It is in no way fundamentally required, that systems be regarded as consisting of very many (“infinitely many”) independent particles, or that all or some of these must be assumed to be identical in order to determine thermodynamic quantities. Neither is it necessary to use complex-valued distributions of phase volumes, but rather, already out of pure classical, constant basic approaches, we get very simply both of the Laws of Thermodynamics. (the Nernst Theorem can naturally not be derived.) Rather, this shows that the conclusion $S - S_0 = k \ln J$ is in opposition to the Boltzmann-Planck postulate $S = k \ln W$, for one can construe the phase volume $J$ if need be (after normalization with the corresponding potential using $h$) as the total of the realized possibilities for all configurations with a smaller energy than $E$.

[71] It is noteworthy, that the method of the convolution of canonical invariant phase densities accomplished here leads to a straightforward definition of quantities which fulfill the thermodynamic differential equations exactly and which, for a composite system, have a definite value and don't “statistically” fluctuate as is usually supposed.

Translation by Joseph Dittle and Randall B. Shirts, 2006