Derivation of Boltzmann Principle

Michele Campisi[1, *] and Donald H. Kobe[2]

[1]*Institute of Physics, University of Augsburg, Universitätsstrasse 1, D-86153 Augsburg, Germany*
[2]*Department of Physics, University of North Texas,*
P.O. Box 311427, Denton, TX 76203-1427, USA
(Dated: November 11, 2009)

We present a derivation of Boltzmann principle $S_B = k_B \ln \mathcal{W}$ based on classical mechanical models of thermodynamics. The argument is based on the heat theorem and can be traced back to the second half of the nineteenth century with the works of Helmholtz and Boltzmann. Despite its simplicity, this argument has remained almost unknown. We present it in a modern, self-contained and accessible form. The approach constitutes an important link between classical mechanics and statistical mechanics.

I. INTRODUCTION

One of the most intriguing equations of modern physics is Boltzmann's celebrated principle

$$S_B = k_B \ln \mathcal{W}, \tag{1}$$

where k_B is Boltzmann's constant. Despite its unquestionable success in providing a means to compute the thermodynamic entropy of isolated systems based on counting the number of available microscopic states $\mathcal{W}$, its theoretical justification remains obscure and vague in most statistical mechanics textbooks. In this respect Khinchin commented:[1] "All existing attempts to give a general proof of this postulate must be considered as an aggregate of logical and mathematical errors superimposed on a general confusion in the definition of the basic quantities." This lack of a crystal clear proof of Boltzmann's principle puts physics students, teachers, and in indeed all physicists, in the uncomfortable position of being forced to accept the formula as a *postulate* that is necessary to link thermodynamic entropy to microscopic dynamics.

Recent studies in the field of history and foundations of statistical mechanics [2] have drawn the attention to the fact that a similar formula,

$$S = k_B \ln \Phi, \tag{2}$$

emerges naturally from classical mechanics if (a) the ergodic hypothesis is made, (b) the properties that entropy should satisfy are appropriately set and the basic quantities are consistently defined. The quantity Φ is the volume in phase space enclosed by a hyper-surface of constant energy E.

Equation (2) is valid for both small and large systems and coincides with the Boltzmann formula for large systems. Hence the derivation of Eq. (2) provides the missing proof of Eq. (1). The basic argument underlying the derivation of Eq. (2) can be traced back to as early as the second half of the nineteenth century in the works of Helmholtz and Boltzmann.[3, 4]

The purpose of this article is to provide a widely accessible modern presentation of the original argument of Helmholtz and Boltzmann [3, 4] and of its recent developments [2], that could be used in the classroom. We derive Boltzmann's principle from classical mechanics with one simple guiding principle, *viz.*, the heat theorem (see Statement 1, below), and one central assumption, namely, the ergodic hypothesis.

In Sec. II we briefly review basics of thermodynamics. We give concise formulations of the first and second laws of thermodynamics and introduce the heat theorem. We then construct a one-dimensional mechanical model of thermodynamics in Sec. III according to the work of Helmholtz.[3] The concepts of ergodicity and microcanonical probability distribution emerging naturally in this model are introduced in Sec. IV. We then generalize the one-dimensional model to more realistic Hamiltonian systems of N-particles in three-dimensions in Sec. V. At this stage the ergodic hypothesis is made and Eq. (2) is derived. In Sec. VI, we point out that the mechanical entropy of Eq. (2) does not change during quasi-static processes in isolated systems, in agreement with the second law of thermodynamics. Non quasi-static processes that can lead to an increase of entropy have been treated elsewhere.[5, 6] In Sec. VII the Boltzmann principle is derived. A summary and some remarks concerning the validity of the ergodic hypothesis are given in Sec. VIII.

In the text we present the line of reasoning and the main results, while proofs and problems are provided in the appendices.

II. CLAUSIUS ENTROPY

The purpose of this paper is to construct a *classical mechanical* analog of *thermodynamic entropy*. To this end it is necessary to give a clear account of the definition of entropy in thermodynamics. We now review the first and second laws of thermodynamics in the formulation

*Electronic address: Michele.Campisi@physik.uni-augsburg.de

given by Clausius (see Ref. [7]). The latter gives the definition of thermodynamic entropy.

First Law of Thermodynamics

In its differential form the first law of thermodynamics reads: [8]

$$dE = \delta Q + \delta W, \tag{3}$$

where dE is the change in internal energy, δQ is the heat added to the system and δW is the work done on the system during an infinitesimal transformation. The first law is the energy conservation law applied to a system in which there is an exchange of energy by both work and heat.

Of crucial importance for the understanding of the first law is that δQ and δW are *inexact* differentials, whereas dE is *exact*. The internal energy E is a *state variable*, namely a quantity that characterizes the thermodynamic equilibrium *state* of the system. On the other hand, W and Q are quantities that characterize thermodynamic energy *transfers* only and are not properties of the state of the system.[9]

Exactness of differentials is best understood in terms of their integral along a path in the system's state space. The differential is exact if and only if the integral depends only on the end points of the path. Accordingly, the integral of an inexact differential along two different paths with same end points may take on different values. The interested readers find a summary of the formal definition of differential forms and their major properties in Appendix A.

Second Law of Thermodynamics

The second law of thermodynamics can be conveniently summarized as three statements.

Statement 1. *The differential $\delta Q/T$ is exact.*

Statement 1, is one of the most important statements of thermodynamics: Although δQ is not an exact differential, when it is divided by the absolute temperature T, an exact differential is obtained.

This is equivalent to stating that that there exist a state function S, such that

$$\frac{\delta Q}{T} = dS. \tag{4}$$

The function S is called the *thermodynamic entropy* of the system.

Statement 1 can be expressed in an equivalent way also in *integral form*, by stating that the integral of $\delta Q/T$ along a path connecting a state A to a state B in the state variables' space, does not depend on the path but only on its endpoints A and B.[11] This in turn says that

there exists a state function S (i.e. the thermodynamic entropy), such that

$$\int_A^B \frac{\delta Q}{T} = S(B) - S(A). \tag{5}$$

From Eq. (3) it is $\delta Q = dE - \delta W$. In general, the work is performed by changing a certain number of external parameters λ_i, e.g. volume, magnetic field, electric field. Then the work δW is given by $-\sum_i F_i d\lambda_i$, where F_i denote the corresponding conjugate forces, i.e., pressure, magnetization, electric polarization, respectively. Therefore it is:

$$\delta Q = dE + \sum_i F_i d\lambda_i \tag{6}$$

Without loss of generality, in the following we will restrict ourselves to the case of only one external parameter V with conjugate force P:[12]

$$\delta Q = dE + PdV \tag{7}$$

In this case Statement 1 can be re-expressed as:

$$(dE + PdV)/T = exact\ differential = dS, \tag{8}$$

This is referred to in the literature as the *heat theorem*. [10] The heat theorem can be re-expressed in equivalent terms as: *there exists a function $S(E,V)$ such that:*[13]

$$\frac{\partial S}{\partial E} = \frac{1}{T}, \qquad \frac{\partial S}{\partial V} = \frac{P}{T}. \tag{9}$$

It is worth emphasizing that any inexact differential, like for example $\delta Q = dE + PdV$, does not enjoy the same property: it is impossible, in general, to find a function of state $Q(E,V)$ such that $\partial Q/\partial E = 1$ and $\partial Q/\partial V = P$.

The following two statements, regarding the function S complete Clausius's form of the second law:

Statement 2. *For a quasi-static process occurring in a thermally isolated system, the entropy change between two equilibrium states is zero,*

$$\Delta S = 0. \tag{10}$$

Statement 3. *For a non quasi-static process occurring in a thermally isolated system, the entropy change between two equilibrium states is nonnegative,*

$$\Delta S \geq 0. \tag{11}$$

A crucial point that must not be overlooked is that Statements 2 and 3 pertain to *thermally isolated* systems. This means that the system is not in contact with a thermal bath, by means of which one could in principle control either its temperature or its energy. Thus the processes mentioned therein are processes in which only the external parameter V is varied in a controlled way and there is no control over the variable E. In Statement 2 the change of the parameter V is so slow that at any instant of time the system is almost at equilibrium (quasi-static process). In Statement 3, this requirement is relaxed.

III. ONE-DIMENSIONAL MECHANICAL MODELS OF THERMODYNAMICS

In this section we construct a one-dimensional classical mechanical analogue of Clausius thermodynamic entropy. This construction dates back to Helmholtz [2, 3, 10] and is based on the heat theorem (8).

Consider a point particle of mass m and coordinate x moving in a U-shaped potential $\varphi(x)$, as illustrated in Fig. 1.

In order to allow for the possibility of doing work on the particle by means of an external intervention, we assume the potential φ to depend on some externally controllable parameter V: $\varphi = \varphi(x; V)$. As an example one could think of a pendulum whose length can be changed at will by an experimenter while the pendulum oscillates. In this case V would denote the length of the pendulum. The Hamiltonian of the system is:

$$H(x, p; V) = K(p) + \varphi(x; V), \qquad (12)$$

where $K(p) = p^2/2m$ is the kinetic energy and p is the momentum.

Now that our mechanical system is defined, we have to specify its "internal energy" E, "temperature" T, and the "force" P conjugate to the external parameter V.

For the internal energy we simply take the energy E given by the Hamiltonian. For a fixed V the particle's energy E is a constant of motion. For the sake of simplicity we chose the gauge of the potential in such a way that the minimum of the potential is 0 regardless of the value of V.

Once V and E are specified, the orbit in phase space of the particle is fully determined: we say that E and V are the system's state variables. [14] The particle moves back and forth between the two turning points $x_\pm(E, V)$, and draws closed orbits in phase space with a certain period $\tau(E, V)$. See Fig. 1.

Now that the state variables are fixed we have to define the corresponding temperature and conjugate force. In agreement with the common understanding of temperature as a measure of the speed of the particles, we take the temperature to be proportional to the kinetic energy averaged over one period:

$$T(E, V) := \frac{2}{k_B \tau(E, V)} \int_0^{\tau(E, V)} dt\, K(p(t; E, V)) \quad (13)$$

Thanks to the the factor $1/k_B$, T has the correct dimensions of a temperature. For the conjugate force we take the time average of $-\frac{\partial \varphi}{\partial V}$ [15], i.e.,

$$P(E, V) := -\frac{1}{\tau(E, V)} \int_0^{\tau(E, V)} dt \frac{\partial \varphi(x(t; E, V); V)}{\partial V} \qquad (14)$$

In Eqs. (13,14), $x(t; E, V)$ and $p(t; E, V)$ are the solution of Hamilton's equations of motion with a fixed V, and an arbitrary initial condition x_0, p_0 such that $H(x_0, p_0; V) = E$.
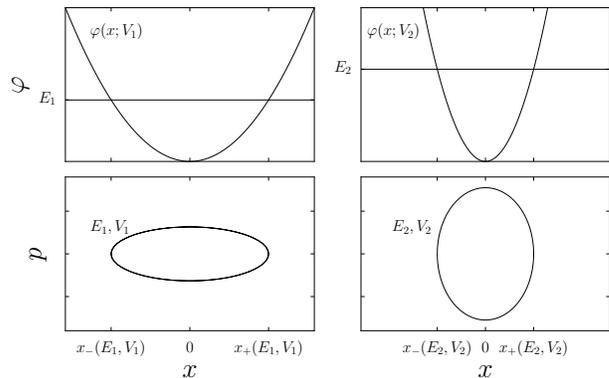

FIG. 1: Point particle in the U-shaped potential $\varphi(x; V) = mV^2 x^2/2$. Top left panel: shape of the potential for a certain $V = V_1$. Top right panel: shape of the potential for a certain $V = V_2$. Bottom left panel, phase space orbit corresponding to the potential $\varphi(x; V_1)$ at energy E_1. Bottom right panel, phase space orbit corresponding to the potential $\varphi(x; V_2)$ at energy E_2. The two quantities E, V, uniquely determine one "state", i.e., one closed orbit in the phase space.

Having identified the mechanical analogues of internal energy, external parameter, temperature and conjugate force with the quantities E, V, T, P, respectively, we can now ask whether a mechanical analogue of entropy, S, exists. In order to answer this question we must ask, in agreement with Statement 1 as expressed in Eq. (9), whether there exists a function $S(E, V)$ such that:

$$\frac{\partial S}{\partial E}(E, V) = \frac{1}{T(E, V)}, \quad \frac{\partial S}{\partial V}(E, V) = \frac{P(E, V)}{T(E, V)}. \quad (15)$$

The answer to this question is given by the following theorem.

Theorem 1 (Helmholtz). *A function $S(E, V)$ satisfying Eq. (15) exists and is given by*

$$S(E, V) = k_B \log 2 \int_{x_-(E, V)}^{x_+(E, V)} dx \sqrt{2m(E - \varphi(x; V))} \qquad (16)$$

The proof of the theorem is reported in Appendix B, see also Ref. [10] (pp 45–46).

The quantity $\sqrt{2m(E - \varphi(x; V))}$ represents the momentum of the particle when it is located at x. Solving the equation $E = p^2/2m + \varphi(x; V)$, and taking the positive root, we obtain:

$$p(x; E, V) = \sqrt{2m(E - \varphi(x; V))} \qquad (17)$$

Thus the entropy can be rewritten compactly as:

$$S(E, V) = k_B \log \oint p\, dx \qquad (18)$$

The integral $\oint p\, dx$ is called the *reduced action*.[15] It represents the area Φ enclosed by the orbit of energy E, and

parameter V in phase space:

$$S(E,V) = k_B \log \Phi(E,V) \quad (19)$$

where

$$\Phi(E,V) = \int_{H(x,p;V) \leq E} dpdx. \quad (20)$$

Theorem 1 says that there exists a mechanical counterpart of entropy, and that this is the logarithm of the phase space volume *enclosed* by the curve of constant energy $H(x,p;V) = E$.

The fact that there exists a function S satisfying Eq. (15) is a highly non trivial result which tells us that we have a consistent one-dimensional mechanical model of thermodynamics. Although this model nicely suggests the deep connection between classical mechanics and thermodynamic entropy, it is definitely too stylized to model a real thermodynamic system composed of as many as 10^{23} particles. It is necessary to generalize Helmholtz Theorem to multidimensional systems.

IV. ERGODICITY AND MICROCANONICAL ENSEMBLE

The main ingredients needed for the generalization of our model to more degrees of freedom are *ergodicity* and the *microcanonical* ensemble.

We use the one-dimensional example of the previous section to introduce these important concepts. Imagine we want to calculate the time average of a phase function $f(x,p)$ over the orbit specified by E and V:

$$\langle f \rangle_t := \frac{1}{\tau} \int_0^\tau dt f(x(t),p(t)). \quad (21)$$

For simplicity of notation we have dropped the explicit dependence of τ, $x(t)$, $p(t)$ and $\langle f \rangle_t$ on E, V.

Since $p = mv = mdx/dt$ the differential dt is:

$$dt = m\frac{dx}{p(x)}. \quad (22)$$

With this we obtain:

$$\langle f \rangle_t = \frac{2m}{\tau} \int_{x_-}^{x_+} \frac{dx}{p(x)} f(x,p(x)) \quad (23)$$

where the factor 2 stems from the fact that the particle goes from x_- to x_+, in a half period, i.e., $\tau/2$. Now consider the following integral

$$\int dp\delta \left(p^2/2m + \varphi(x;V) - E\right) \quad (24)$$

where δ denotes Dirac's delta function. Using the formula $\delta(f(p)) = \sum_i \delta(p - p_i)/|f'(p_i)|$, where the p_i's are the zeroes of $f(p)$, and $\int dp\delta(p - p_i)=1$, we get:

$$\int dp\delta \left(p^2/2m + \varphi(x;V) - E\right) = 2m/p(x) \quad (25)$$

Then Eq. (23) becomes:

$$\langle f \rangle_t = \frac{1}{\tau} \int dx \int dp\delta \left(p^2/2m + \varphi(x;V) - E\right) f(p,x) \quad (26)$$

where the integration extremes $x_\pm$ need not be specified, being implied by the Dirac δ. The period τ is given by:

$$\begin{aligned} \tau &= \int_0^\tau dt = 2\int_{x_-}^{x_+} \frac{dx}{p(x)} \\ &= \int dx \int dp\delta \left(p^2/2m + \varphi(x;V) - E\right) \quad (27) \end{aligned}$$

Hence we arrive at:

$$\langle f \rangle_t = \int dx \int dp\rho(x,p;E,V)f(p,x) \quad (28)$$

where we have introduced the phase space probability density function

$$\rho(x,p;E,V) = \frac{1}{\tau(E,V)}\delta \left(p^2/2m + \varphi(x;V) - E\right) \quad (29)$$

From Eq. (29) it is clear that τ is the normalization. This $\rho(x,p;E,V)$ is called the microcanonical distribution. Eq. (28) says that the time average of a phase space quantity $f(x,p)$ over one period, is equal to its microcanonical average. This property is called ergodicity. All one-dimensional systems with a U-shaped potential are ergodic.

V. MULTI PARTICLE MECHANICAL MODELS OF THERMODYNAMICS

We now can extend the previous treatment to systems of N-particles in three dimensions. The Hamiltonian for an interacting system of N-particles of mass m is

$$H_N(\mathbf{q},\mathbf{p};V) = K_N(\mathbf{p}) + \varphi_N(\mathbf{q};V), \quad (30)$$

where $K_N(\mathbf{p}) = \sum_{i=1}^{3N} p_i^2/2m$ is the kinetic energy, φ_N is the potential energy, and the coordinates $\mathbf{q} = \{q_i\}_{i=1}^{3N}$ and their conjugate canonical momenta $\mathbf{p} = \{p_i\}_{i=1}^{3N}$ are $3N$-dimensional vectors.

In analogy with one-dimensional systems with a U-shaped potential, we define the microcanonical probability distribution as

$$\rho_N(\mathbf{q},\mathbf{p};E,V) = \frac{1}{\Omega_N(E,V)}\delta \left(E - H_N(\mathbf{q},\mathbf{p};V)\right), \quad (31)$$

where $\Omega_N(E,V)$ is the normalization:

$$\Omega_N(E,V) = \int \cdots \int d\mathbf{q}d\mathbf{p}\delta \left(E - H_N(\mathbf{q},\mathbf{p};V)\right). \quad (32)$$

Continuing the analogy with one-dimensional systems, we make the following crucial assumption:

Ergodic Hypothesis. *For given E and V, the time average $\langle f \rangle_t$ of any function $f(\mathbf{q}, \mathbf{p})$ is uniquely determined and is equal to its microcanonical average $\langle f \rangle_\mu$, i.e.:*

$$\langle f \rangle_t = \int \cdots \int d\mathbf{q} d\mathbf{p} \; \rho_N(\mathbf{q}, \mathbf{p}; E, V) f(\mathbf{q}, \mathbf{p}) \doteq \langle f \rangle_\mu. \tag{33}$$

In analogy with Eqs.(13,14) we define the temperature as:

$$T_N(E, V) := \frac{2}{3Nk_B} \langle K_N \rangle_t \tag{34}$$

and the conjugate force as

$$P_N(E, V) := -\left\langle \frac{\partial \varphi_N}{\partial V} \right\rangle_t. \tag{35}$$

We ask, in agreement with Statement 1 as expressed in Eq. (9), whether there exists a function $S_N(E, V)$ such that

$$\frac{\partial S_N}{\partial E}(E, V) = \frac{1}{T_N(E, V)}, \quad \frac{\partial S_N}{\partial V}(E, V) = \frac{P_N(E, V)}{T_N(E, V)}. \tag{36}$$

The answer is given by the following theorem:

Theorem 2 (Helmholtz, Generalized). *A function $S_N(E, V)$ satisfying Eq. (36) exists and is given by:*

$$S_N(E, V) = k_B \log \Phi_N(E, V) \tag{37}$$

where

$$\Phi_N(E, V) := \int \cdots \int_{H_N(\mathbf{q}, \mathbf{p}) \leq E} d\mathbf{q} d\mathbf{p} \tag{38}$$

The proof, which is based on the equipartition theorem, is given in Appendix C. See also Ref. [2].

We draw the attention to the fact that, unlike temperature and conjugate force, S_N is not in the form of the time average of some phase function $f(\mathbf{q}, \mathbf{p})$

Theorem 2 says that ergodic systems constitute ideal mechanical models of thermodynamics. One can define their state variables by E and V as in thermodynamics. Moreover, one can define their temperature and conjugate force straightforwardly as functions of the state variables. Surprisingly, the heat differential $(dE + P_N dV)/T_N$ is exact, allowing for a consistent and logical definition of entropy S_N.

VI. ADIABATIC INVARIANCE

According to Theorem 2, S_N complies with the first law of thermodynamics and statement 1 of the second law of thermodynamics: Is the construction consistent with statements 2 and 3 of the second law of thermodynamics as well?

Let us focus on statement 2. In order for S_N to be consistent with this statement it is necessary that if the parameter V is varied very slowly in time (much slower than any time scale of the system dynamics) from a certain $V_0 = V(t_0)$ to a certain $V_f = V(t_f)$, the corresponding change of the entropy S_N is null. Note that, by allowing for a time dependence of V, the system's Hamiltonian now becomes time dependent, and energy is not conserved. Let the system be at $t = t_0$, in $\mathbf{q}_0, \mathbf{p}_0$. Under the time dependent Hamiltonian

$$H_N(\mathbf{q}, \mathbf{p}; V(t)) = K_N(\mathbf{p}) + \varphi_N(\mathbf{q}; V(t)), \tag{39}$$

it evolves to a new phase space point $\mathbf{q}_f(\mathbf{q}_0, \mathbf{p}_0), \mathbf{p}_f(\mathbf{q}_0, \mathbf{p}_0)$, where we made explicit the dependence on the initial condition of the evolved phase space point. Then the energy at time t_f is $E_f = H_N(\mathbf{q}_f(\mathbf{q}_0, \mathbf{p}_0), \mathbf{p}_f(\mathbf{q}_0, \mathbf{p}_0); V(t_f))$. It is known [16] that, for *ergodic systems*, the energy reached at the end of a very slow protocol depends only on the initial energy $E_0 = H_N(\mathbf{q}_0, \mathbf{p}_0; V(t_0))$, and is determined by solving the following equation for E_f:

$$\Phi_N(E_0, V_0) = \Phi_N(E_f, V_f) \tag{40}$$

That is, the quantity $\Phi_N(E, V)$ does not change in the course of time, when V is varied infinitely slowly. This property is called in classical mechanics *adiabatic invariance*. Since it is $S_N(E, V) = k_B \log \Phi_N(E, V)$, and $\Phi_N(E, V)$ is an adiabatic invariant, it is evident that S_N is an adiabatic invariant too. Namely, it does not change if V is changed very slowly in time. Thus S_N complies with Statement 2. For completeness in Appendix D we provide a proof of adiabatic invariance of Φ_N. See also Ref. [16] (pp. 27–30).

It is also possible to prove that, in an averaged sense, S_N complies with Statement 3, as well.[5, 6] In this case one has to consider the average change of entropy, because, for fast transformations, the final energy is not uniquely determined by the initial energy, and depending on the initial conditions, one ends up with different final energies, i.e., different final entropies.

VII. BOLTZMANN PRINCIPLE

For a system composed of a very large number N of particles which interact via short range forces, the phase space volume $\Phi_N(E)$ approaches an exponential behavior $\Phi_N(E) \propto e^E$. Since $\Omega_N = \partial \Phi_N / \partial E$ (see Eq. (32)), it is $\Phi_N \propto \Omega_N$ (see Ref. [17], p.148).

The quantity Ω_N, defined in Eq. (32), represents the measure of the shell of constant energy $H_N(\mathbf{q}, \mathbf{p}; V) = E$. As such it is proportional the number $\mathcal{W}$ of micro states compatible with the given energy E. (According to semiclassical theory each micro state occupies a volume h^{3N} of phase space, where h is Planck's constant.[18] By introducing an arbitrary energy scale ΔE, the number $\mathcal{W}$ is given by $\mathcal{W} = \Omega_N \Delta E / h^{3N}$). Thus, for very large N,

$$\Phi_N \propto \Omega_N \propto \mathcal{W}, \quad N \gg 1 \tag{41}$$

By taking the logarithm, we have

$$S_N \simeq k_B \ln \mathcal{W} = S_B, \quad N \gg 1 \qquad (42)$$

except for an irrelevant constant. Eq. (42) says that for large ergodic systems composed of particles interacting via short range forces, the differential of Boltzmann entropy is equal to the differential $\delta Q/T$. Hence it can be identified with Clausius entropy. This is a proof of Boltzmann principle.

VIII. CONCLUSIONS

Given an ergodic system, it is possible to specify its thermodynamic state by means of the total energy E and the external parameter V. Given the state E, V, we can unambiguously define the quantities $T_N(E, V)$ and $P_N(E, V)$. Once these are identified with the system temperature and conjugate force, one can ask whether, as prescribed by the heat theorem, the combination:

$$\frac{dE + P_N dV}{T_N}$$

is an exact differential. Surprisingly the answer is positive, meaning that there exist a function $S_N(E, V)$ which can be identified with the thermodynamic entropy of the system. The generalized Helmholtz theorem says that this is given by the logarithm of the volume $\Phi_N(E, V)$ of phase space *enclosed* by the hyper-surface of energy $H(\mathbf{q}, \mathbf{p}; V) = E$. For macroscopic systems this entropy coincides with Boltzmann entropy, thus revealing the rationale of Boltzmann principle.

The entropy in Eq. (37) is sometimes referred to in the literature as Hertz entropy.[19, 20] Hertz [21, 22] derived it from the requirement of adiabatic invariance (see also Refs. [16, 23, 24]), whereas we have derived it here from the heat theorem. Its fundamental character is also recognized in Ref. [25] where its property of being a *canonical invariant* is emphasized and in Ref. [26] which highlights its compliance with the equipartition theorem, and the fact that it is a positive and increasing function of the energy[27, 28]. The entropy in Eq. (37) also appears in Gibbs seminal book.[29] However its connection with the heat theorem has not previously recognized.

The most crucial point of the derivation of Boltzmann principle is the introduction of the ergodic hypothesis. Although this hypothesis is generally believed to be true for real macroscopic systems, its mathematical proof is a formidable challenge which has been achieved only in few special cases.[30] A proof that a gas of elastically colliding hard spheres is ergodic was announced in 1963 by Sinai. [31] However the full proof was not published and the problem is still open (see Ref. [32] for a more detailed discussion). Nonetheless ergodicity of hard spheres systems seems plausible as indicated also by recent numerical simulations (see Sec. IV of Ref. [33]). The hypothesis cannot be true in the case of real crystals. Here the nuclei remain close to their lattice site, preventing them from sampling the whole energy hypersurface homogeneously (this being a necessary condition for ergodicity).

In regard to these difficulties, it is worth pointing out that the present derivation of Boltzmann principle does not make use of the fact that the average of *any* arbitrary phase function be equal to its microcanonical average, as required by the ergodic hypothesis. It *only* requires that the time average of K and $-\partial \varphi / \partial V$, be equal to their microcanonical averages (see the proof of Theorem 2 in appendix C). Thus the ergodic hypothesis can be greatly relaxed by requiring the much less stringent condition that T_N and P_N, Eqs. (34,35), can be calculated as microcanonical averages.[34] In this case, the Clausius entropy can still be calculated via the Hertz formula (37).

APPENDIX A: DIFFERENTIAL FORMS: BRIEF REVIEW OF DEFINITIONS AND MAIN RESULTS

A differential form ω in a connected subset $\mathcal{A}$ of $\mathbb{R}^2$ is formally written as:

$$\omega = \mathcal{M}(x_1, x_2)dx_1 + \mathcal{N}(x_1, x_2)dx_2 \qquad (A1)$$

where $\mathcal{M}(x_1, x_2), \mathcal{N}(x_1, x_2)$ are two functions on $\mathcal{A}$ and (x_1, x_2) are the coordinates in $\mathbb{R}^2$. [35]

Given a curve $\psi : [s_0, s_1] \to \mathcal{A}$,

$$\psi(s) = (\psi_1(s), \psi_2(s)) \qquad (A2)$$

the integral of ω along the curve ψ, is defined as:

$$\int_\psi \omega = \int_{s_0}^{s_1} [\mathcal{M}(\psi(t))\psi_1'(t) + \mathcal{N}(\psi(t))\psi_2'(t)]dt \qquad (A3)$$

where $\psi_{1,2}'$ are the derivatives of $\psi_{1,2}$.

A differential form ω is said to be *exact*, if there exist a function $G : \mathcal{A} \to \mathbb{R}$ such that:

$$\omega = dG \qquad (A4)$$

that is:

$$\frac{\partial G}{\partial x_1}(x_1, x_2) = \mathcal{M}(x_1, x_2), \quad \frac{\partial G}{\partial x_2}(x_1, x_2) = \mathcal{N}(x_1, x_2) \qquad (A5)$$

G is called a *primitive* for the differential form.

Let $\Sigma(A, B)$ be the set of all curves connecting the point $A \equiv (a_1, a_2)$ to the point $B \equiv (b_1, b_2)$ in $\mathcal{A}$. A differential form is exact if and only if for any couple of points A and B in $\mathcal{A}$ and curves ψ and ϕ in $\Sigma(A, B)$, it is

$$\int_\psi \omega = \int_\phi \omega \qquad (A6)$$

The integral of an exact differential form along any curve ψ connecting A to B does not depend on the curve ψ, but only on the ending points, and is given by:

$$\int_\psi \omega = \int_\psi dG = G(B) - G(A). \qquad (A7)$$

The following statement also holds: A differential form is exact if and only if its integral along any closed curve is null.

If the functions $\mathcal{M}$ and $\mathcal{N}$ are of class C^1 (i.e. they are differentiable), then a necessary condition for the form ω to be exact is that:

$$\frac{\partial \mathcal{M}}{\partial x_2} = \frac{\partial \mathcal{N}}{\partial x_1} \qquad (A8)$$

In this case the differential form ω is said to be *closed*.

APPENDIX B: PROOF OF THEOREM 1

It is known [15] that in one-dimensional systems confined in a U-shaped potential, the period τ of the orbit is equal to the derivative of the area Φ in phase space enclosed by the orbit with respect to energy.

$$\tau = \frac{\partial \Phi}{\partial E} \qquad (B1)$$

One simple way to prove this relation is by expressing the area as $\Phi(E, V) = \int dp dx \theta(E - H(x, p; V))$, where $\theta(x)$ is Heaviside step function [$\theta(x) = 1$ if $x \geq 0$, $\theta(x) = 0$ if $x < 0$]. Taking the derivative with respect to E, and using the relation $\delta(x) = d\theta(x)/dx$, gives τ (see Eq. (27)). Using Eq. (B1) and Eq. (19) we obtain:

$$\frac{\partial S}{\partial E} = k_B \frac{\tau}{\Phi}. \qquad (B2)$$

From Eq. (23), we obtain the relation:

$$2 \langle K \rangle_t = \frac{\Phi}{\tau} \qquad (B3)$$

from which, using Eq. (B2), we get:

$$\frac{\partial S}{\partial E} = \frac{k_B}{2 \langle K \rangle_t} \qquad (B4)$$

Similarly we also get:

$$\frac{\partial S}{\partial V} = -\frac{k_B}{2 \langle K \rangle_t} \left\langle \frac{\partial \varphi}{\partial V} \right\rangle_t \qquad (B5)$$

Using Eqs. (13,14) we obtain:

$$\frac{\partial S}{\partial E} = \frac{1}{T} \qquad (B6)$$

$$\frac{\partial S}{\partial V} = \frac{P}{T} \qquad (B7)$$

APPENDIX C: PROOF OF THEOREM 2

The proof of Theorem 2 makes use of the multidimensional version of Eq. (B1), that is:

$$\Omega_N = \frac{\partial \Phi_N}{\partial E} \qquad (C1)$$

This can be proved, in a similar way, by expressing Φ_N as $\int d\mathbf{q} d\mathbf{p} \theta(E - H(\mathbf{q}, \mathbf{p}; V))$ and using the relation $\delta(x) = d\theta(x)/dx$. The equipartition theorem [17]

$$\frac{2 \langle K \rangle_\mu}{3N} = \frac{\Phi_N}{\Omega_N} \qquad (C2)$$

is the generalization of Eq. (B3) to many dimensions. Using (C2) and (C1) with Eq. (37) we get:

$$\frac{\partial S_N}{\partial E} = \frac{3N k_B}{2 \langle K_N \rangle_\mu} \qquad (C3)$$

In a similar way we also get:

$$\frac{\partial S_N}{\partial V} = -\frac{3N k_B}{2 \langle K_N \rangle_\mu} \left\langle \frac{\partial \varphi_N}{\partial V} \right\rangle_\mu \qquad (C4)$$

Using Eqs. (34,35) with the ergodic hypothesis, we finally arrive at:

$$\frac{\partial S_N}{\partial E} = \frac{1}{T_N} \qquad (C5)$$

$$\frac{\partial S_N}{\partial V} = \frac{P_N}{T_N} \qquad (C6)$$

APPENDIX D: PROOF OF ADIABATIC INVARIANCE OF Φ_N

We consider the time-dependent Hamiltonian

$$H_N(\mathbf{q}, \mathbf{p}; V(t)) = K(\mathbf{p}) + \varphi(\mathbf{q}; V(t)). \qquad (D1)$$

To prove that Φ_N in Eq. (38) is an adiabatic invariant we first take the total time derivative of the Hamiltonian $H_N(\mathbf{q}, \mathbf{p}; V(t))$ in Eq. (D1),

$$\frac{dH_N(\mathbf{q}, \mathbf{p}; V)}{dt} = \frac{\partial H_N(\mathbf{q}, \mathbf{p}; V)}{\partial V} \frac{dV}{dt} \qquad (D2)$$

where the terms involving $\dot{\mathbf{q}}$ and $\dot{\mathbf{p}}$ cancel by Hamilton's equations.[36] The derivative dV/dt changes slowly in time, but dH_N/dt and $\partial H_N/\partial V$ can change rapidly because of their dependence on $\mathbf{q}(t)$ and $\mathbf{p}(t)$. To eliminate the fast variables $\mathbf{q}, \mathbf{p}$ we take the average of Eq. (D2) with respect to the microcanonical ensemble, which gives

$$\left\langle \frac{dH_N}{dt} \right\rangle_\mu = \left\langle \frac{\partial H_N}{\partial V} \right\rangle_\mu \frac{dV}{dt}. \tag{D3}$$

By Liouville's theorem [36] the average on the left-hand side of Eq. (D3) is

$$\left\langle \frac{dH_N}{dt} \right\rangle_\mu = \frac{dE}{dt}. \tag{D4}$$

The microcanonical average in Eq. (33) on the right-hand side of Eq. (D3) is

$$\begin{aligned} \left\langle \frac{dH_N}{dV} \right\rangle_\mu &= \int \cdots \int d\mathbf{q}d\mathbf{p} \; \rho_N(\mathbf{q}, \mathbf{p}, E, V) \frac{\partial H_N(\mathbf{q}, \mathbf{p}, V)}{\partial V} \\ &= -\frac{1}{\Omega_N} \frac{\partial \Phi_N}{\partial V}, \end{aligned} \tag{D5}$$

where Φ_N and Ω_N are given in Eqs. (32) and (38), respectively. Substituting Eqs. (D4) and (D5) into Eq. (D3) and using $\Omega_N = \partial \Phi_N/\partial E$, we obtain

$$\frac{d\Phi_N}{dt} \equiv \frac{\partial \Phi_N}{\partial E}\frac{dE}{dt} + \frac{\partial \Phi_N}{\partial V}\frac{dV}{dt} = 0, \tag{D6}$$

which shows that Φ_N is constant and therefore an adiabatic invariant.

APPENDIX E: PROBLEMS

Consider the following Hamiltonian of a one dimensional harmonic oscillator with angular frequency V (see Fig. 1)

$$H(x, p; V) = \frac{p^2}{2m} + \frac{mV^2x^2}{2} \tag{E1}$$

(a) Calculate the area $\Phi(E, V)$ enclosed by the trajectory of energy E and angular frequency V. Using Eq. (B1) check that the period of the orbit is, as expected, given by $\tau(E, V) = 2\pi/V$.

(b) Using Eqs. (13, 14) show that $k_B T(E, V) = E$, $P(E, V) = -E/V$

(c) Show that the differential form $dE + PdV$, with $P(E, V)$ as in (b) is not exact. (Hint: show that Eq. (A8) is not satisfied.) Show that the integral of $dE + PdV$ over the rectangular path with corners $(E_0, V_0), (E_0, V_1), (E_1, V_1), (E_1, V_0)$, and $E_0 \neq E_1$ $V_0 \neq V_1$, is not null.

(d) Consider the differential form $\omega = (1/T)dE + (P/T)dV$, with $P(E, V), T(E, V)$ as in (b). Find a primitive function $S(E, V)$ for ω. Show that, apart from an additive constant, it is $S(E, V) = \log \Phi(E, V)$, as dictated by Theorem 1. Check that Eq. (A8) is satisfied.

[1] A. Khinchin, *Mathematical foundations of statistical mechanics.* (New York: Dover, 1949). See page 142.

[2] M. Campisi, "On the mechanical foundations of thermodynamics: The generalized Helmholtz theorem," *Stud. Hist. Phil. Mod. Phys.* **36**, 275–290 (2005).

[3] H. Helmholtz, "Principien der Statik monocyklischer Systeme". In *Wissenshafltliche Abhandlungen,* vol III, p.142-162 and p.179-202; "Studien zur Statik monocyklischer Systeme," *in Wissenshafltliche Abhandlungen,* vol III, p.163-172 and p.173-178. 1895.

[4] L. Boltzmann, "Über die Eigenschaften monocyklischer und anderer damit verwandter Systeme," *Crelle's Journal* **98**, 68–94, (1884). Reprinted in Hasenöhrl (ed.), *Wissenschaftliche Abhandlungen,* vol. 3, pp. 122-152. New York: Chelsea.

[5] M. Campisi, "Statistical mechanical proof of the second law of thermodynamics based on volume entropy," *Stud. Hist. Phil. Mod. Phys.* **39**, 181–194 (2008).

[6] M. Campisi, "Increase of Boltzmann entropy in a quantum forced harmonic oscillator," *Phys. Rev. E* **78**, 051123 (2008).

[7] J. Uffink, "Bluff your way in the second law of thermodynamics," *Stud. Hist. Phil. Mod. Phys.* **32**(3), 305–394 (2001).

[8] H. B. Callen, *Thermodynamics.* (New York: Wiley, 1960).

[9] D. Chandler, *Introduction to Modern Statistical Mechanics.* (Oxford University Press, 1987).

[10] G. Gallavotti, *Statistical mechanics. A short treatise.* (Berlin: Springer Verlag, 1995).

[11] A path in the state variables' space corresponds to a transformation that leads the system from A to B through a sequence of almost equilibrium states. Accordingly, the change δQ has to be understood as a quasi-static change.

[12] For the common case of work due to expansion and compression the external parameter is the volume and its conjugate force is the pressure (for this reason the notation V P is adopted). In general V and P stand for any conjugate pair of "displacement" and "force" depending on the specific nature of the work done on the system (e.g., magnetic field and magnetization).

[13] To use the language of vector fields, this says that the field $\vec{\mathcal{F}} := (1/T, F/T)$ is a *conservative* vector field. That is, there exists a *potential* function $S(E, V)$ such that $\vec{\mathcal{F}} = \vec{\nabla}S$ where $\vec{\nabla}$ is the gradient operator in the space (E, V). For this reason the entropy can be understood as

a *thermodynamic potential.*

[14] The variables (E, V) should be understood as the system's "thermodynamic" state variables, and should not be confused with the system's "mechanical" state variables (p, x).

[15] L. Landau and E. Lifshitz, *Mechanics.* (Oxford: Pergamon, 1960).

[16] V. L. Berdichevsky, *Thermodynamics of chaos and order.* (Essex: Addison Wesley Longman, 1997).

[17] K. Huang, *Statistical mechanics.* (Singapore: John Wiley & Sons, 2nd ed., 1963).

[18] L. Landau and E. Lifshitz, *Statistical Physics*, 2nd ed. (Oxford: Pergamon, 1969).

[19] S. Hilbert and J. Dunkel, "Nonanalytic microscopic phase transitions and temperature oscillations in the microcanonical ensemble: An exactly solvable 1d-model for evaporation," *Phys. Rev. E* **74**, 011120 (2006).

[20] A. Adib, "Does the Boltzmann principle need a dynamical correction?," *J. Stat. Phys.* **117**, 581–597 (2004).

[21] P. Hertz, "Über die mechanischen Grundlagen der Thermodynamik," *Ann. Phys. (Leipzig)* **33**, 225–274 (1910).

[22] P. Hertz, "Über die mechanischen Grundlagen der Thermodynamik," *Ann. Phys. (Leipzig)* **33**, 537–552 (1910).

[23] A. Münster, *Statistical thermodynamics*, vol. 1. (Berlin: Springer Verlag, 1969).

[24] H. H. Rugh, "Microthermodynamic formalism," *Phys. Rev. E* **64**(5), 055101 (2001).

[25] A. Schlüter, "Zur Statistik klassischer Gesamtheiten," *Z. Naturforschg.* **3A**, 350–360 (1948). An english translation of this article is found at http://people.chem.byu.edu/rbshirts/research/schluter1948translation.doc.

[26] E. M. Pearson, T. Halicioglu, and W. A. Tiller, "Laplace-transform technique for deriving thermodynamic equations from the classical microcanonical ensemble," *Phys. Rev. A* **32**(5), 3030–3039 (1985).

[27] P. Talkner, P. Hänggi, and M. Morillo "Microcanonical quantum fluctuation theorems," *Phys. Rev. E* **77**, 051131 (2008)

[28] The conditions $\Phi \geq 0, \partial_E \Phi = \Omega \geq 0$ ensure that the temperature derived from the Hertz entropy, i.e. $T = (\partial_E \log \Phi)^{-1} = \Phi/\Omega$ be definite positive.

[29] J. Gibbs, *Elementary principles in statistical mechanics.* (Yale: Yale University Press, 1902. Reprinted by Dover, New York, 1960).

[30] J. L. Lebowitz and O. Penrose, "Modern ergodic theory," *Physics Today*, **26**(2) 23–29 (1973).

[31] Ya. G. Sinai, "On the foundation of the ergodic hypothesis for a dynamical system of statistical mechanics," *Sov. Math. Dokl.* **4** 1818-1822 (1963).

[32] J. Uffink, "Compendium of the foundations of classical statistical physics". In J. Butterfield and J. Earman (Eds.) *Philosophy of Physics.* (Amsterdam; Boston: Elsevier, 2007). Preprint available at http://philsci-archive.pitt.edu/archive/00002691/

[33] M. Campisi, P. Talkner and P. Hänggi, "Finite Bath Fluctuation Theorem," *Phys. Rev. E* **80**, 031145 (2009).

[34] The simplest example of a non-ergodic system that satisfies this weaker condition is that of a 1-d particle in a symmetric double well potential. For energies below a certain critical value E_c the curve of constant energy splits into two disjoint curves γ_l γ_r, with the phase space trajectory covering only one of them (for energy above E_c, there is only one trajectory and ergodicity holds). Since γ_l and γ_r are one the mirror image of the other, the time averages of even functions of x, such as K and $-\partial\varphi/\partial V$ are not affected by which of the two curves the motion actually takes place on. It follows then, that such averages can be calculated as microcanonical averages (that is over the curve $\gamma_l \cup \gamma_r$).

[35] In the main text the role of (x_1, x_2) is played by the state variables (E, V).

[36] H. Goldstein, *Classical mechanics.* (Addison-Wesley series in physics, Reading, Mass.: Addison-Wesley Pub. Co., 2nd ed., 1980).