

# Derivation of the Boltzmann principle

Michele Campisi<sup>a)</sup>

*Institute of Physics, University of Augsburg, Universitätsstrasse 1, D-86135 Augsburg, Germany*

Donald H. Kobe<sup>b)</sup>

*Department of Physics, University of North Texas, P.O. Box 311427, Denton, Texas 76203-1427*

(Received 22 September 2009; accepted 4 January 2010)

We derive the 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 19th century in the works of Helmholtz and Boltzmann. Despite its simplicity, this argument has remained almost unknown. We present it in a contemporary, self-contained, and accessible form. The approach constitutes an important link between classical mechanics and statistical mechanics. © 2010 American Association of Physics Teachers.

[DOI: 10.1119/1.3298372]

## I. INTRODUCTION

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

$$S_B = k_B \ln \mathcal{W}, \quad (1)$$

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  $\mathcal{W}$  of available microscopic states, its theoretical justification remains obscure and vague in most statistical mechanics textbooks. In this respect Khinchin has commented:<sup>1</sup> "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." The lack of a crystal-clear proof of Boltzmann's principle puts physics students, teachers, and all physicists in the uncomfortable position of being forced to accept the relation as a postulate which is necessary to link thermodynamic entropy to microscopic dynamics.

Recent studies in the history and foundations of statistical mechanics<sup>2</sup> have drawn attention to the fact that a similar relation,

$$S = k_B \ln \Phi, \quad (2)$$

emerges naturally from classical mechanics if the ergodic hypothesis is made, the properties that entropy should satisfy are appropriately set, and the basic quantities are consistently defined. The quantity  $\Phi$  is the volume in phase space enclosed by a hypersurface of constant energy  $E$ .

Equation (2) is valid for both large and small systems and coincides with the Boltzmann formula for large systems. Hence, the derivation of Eq. (2) provides the missing link for Eq. (1). The basic argument underlying the derivation of Eq. (2) can be traced to as early as the second half of the 19th century in the work of Helmholtz and Boltzmann.<sup>3,4</sup>

The purpose of this article is to provide an accessible and contemporary presentation of the original argument of Helmholtz and Boltzmann<sup>3,4</sup> and of its recent developments.<sup>2</sup> We derive Boltzmann's principle from classical mechanics with one simple guiding principle—the heat theorem (see statement 1) and one central assumption, namely, the ergodic hypothesis.

In Sec. II we briefly review the basics of thermodynamics. We give concise formulations of the first and second laws of thermodynamics and introduce the heat theorem. We con-

struct a one-dimensional mechanical model of thermodynamics in Sec. III according to the work of Helmholtz.<sup>3</sup> The concepts of ergodicity and microcanonical probability distribution emerge naturally in this model and are introduced in Sec. IV. In Sec. V we generalize the one-dimensional model to more realistic Hamiltonian systems of  $N$ -particles in three dimensions. At this stage the ergodic hypothesis is assumed 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 in entropy have been treated elsewhere.<sup>5,6</sup> 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.

We present the line of reasoning and main results in the text, while proofs and problems are provided in Appendices B–E.

## II. CLAUDIUS ENTROPY

We first review the first and second laws of thermodynamics in the formulation given by Clausius,<sup>7</sup> which gives the definition of thermodynamic entropy.

The differential form of the first law of thermodynamics is<sup>8</sup>

$$dE = \delta Q + \delta W, \quad (3)$$

where  $dE$  is the change in internal energy,  $\delta Q$  is the energy added to the system due to heating, and  $\delta W$  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 heating. It is essential to realize that  $\delta Q$  and  $\delta W$  are inexact differentials, whereas  $dE$  is an exact differential. The internal energy  $E$  is a state variable, a quantity that characterizes the thermodynamic equilibrium state of the system. In contrast,  $W$  and  $Q$  characterize thermodynamic energy transfers only and are not properties of the state of the system.<sup>9</sup>

A differential is exact if and only if the integral along a path in the system's state space depends only on the end points. In contrast, the corresponding integral of an inexact differential with the same end points depends on the path taken. A summary of the formal definition of differential forms (more precisely, 1-form) and their major properties is given in Appendix A.

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

*Statement 1.* The differential  $\delta Q/T$  is exact. This statement is one of the most important statements of thermodynamics. Although  $\delta Q$  is not an exact differential, an exact differential is obtained when it is divided by the absolute temperature  $T$ . An equivalent statement is that there exists a state function  $S$  such that

$$\frac{\delta Q}{T} = dS. \quad (4)$$

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

Statement 1 can also be stated in an equivalent integral form. The integral of  $\delta Q/T$  along a path connecting a state  $A$  to a state  $B$  in the system's state space does not depend on the path, but only on its end points  $A$  and  $B$ .<sup>11</sup> This statement means that there exists a state function  $S$  (the thermodynamic entropy) such that

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

From Eq. (3), the heat transfer is  $\delta Q = dE - \delta W$ . In general, work can be performed by changing external parameters  $\lambda_i$ , such as the volume, magnetic field, and the electric field. The work  $\delta W$  done is given by  $-\sum_i F_i d\lambda_i$ , where  $F_i$  denotes the corresponding conjugate forces, pressure, magnetization, and electric polarization, respectively. Therefore, the first law can be written as

$$\delta Q = dE + \sum_i F_i d\lambda_i. \quad (6)$$

Without loss of generality, we will restrict ourselves in the following to only one external parameter  $V$  with conjugate force  $P$ :<sup>12</sup>

$$\delta Q = dE + PdV. \quad (7)$$

In this case, statement 1 can be expressed as

$$(dE + PdV)/T = \text{exact differential} = dS. \quad (8)$$

Equation (8) is known in literature as the heat theorem.<sup>10</sup> Because by definition  $dS = (\partial S/\partial E)dE + (\partial S/\partial V)dV$ , the heat theorem can be expressed in equivalent terms as there exists a function  $S(E, V)$  such that<sup>13</sup>

$$\frac{\partial S}{\partial E} = \frac{1}{T}, \quad \frac{\partial S}{\partial V} = \frac{P}{T}. \quad (9)$$

Any inexact differential, such as  $\delta Q = dE + PdV$ , does not enjoy the same property: It is impossible 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,

$$\Delta S = 0. \quad (10)$$

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

$$\Delta S \geq 0. \quad (11)$$

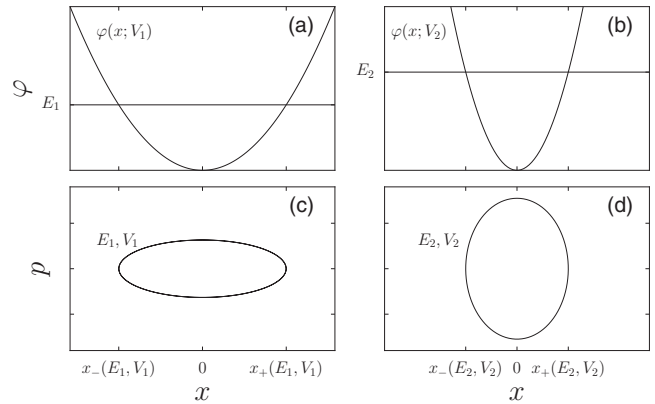


Fig. 1. Point particle in the U-shaped potential  $\varphi(x; V) = mV^2 x^2/2$ . (a) Shape of the potential for a  $V = V_1$ . (b) Shape of the potential for a  $V = V_2$ . (c) Phase space orbit corresponding to the potential  $\varphi(x; V_1)$  at energy  $E_1$ . (d) Phase space orbit corresponding to the potential  $\varphi(x; V_2)$  at energy  $E_2$ . The two quantities  $E, V$ , uniquely determine one “state,” that is, one closed orbit in phase space.

A crucial point in statements 2 and 3 is that they are for thermally isolated systems. Thermally isolated systems are those that are not in contact with a thermal bath, by means of which either their temperature or energy can be controlled. Thus, these processes are those in which only the external parameter  $V$  is varied in a controlled way and the variable  $E$  is uncontrolled. In statement 2 a quasi-static process is one in which the change in the parameter  $V$  is so slow that at any instant of time the system is almost in equilibrium. In statement 3, this requirement is relaxed.

### III. ONE-DIMENSIONAL MECHANICAL MODELS OF THERMODYNAMICS

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

Consider a particle of mass  $m$  and coordinate  $x$  moving in a U-shaped potential  $\varphi(x)$ , as illustrated in Fig. 1. To allow for the possibility of doing work on the particle by means of an external intervention, we assume that the potential  $\varphi$  depends on some externally controllable parameter  $V$  so that  $\varphi = \varphi(x; V)$ . The Hamiltonian of this system is

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

where  $K(p) = p^2/2m$  is the kinetic energy and  $p$  is the momentum. Given this mechanical system, we need to specify its internal energy  $E$ , “temperature”  $T$ , and the “force”  $P$  conjugate to the external parameter  $V$ .

For the internal energy, we take the energy  $E$  given by the Hamiltonian. For a fixed  $V$ , the particle's energy  $E$  is a constant of motion. For simplicity, we define the zero of energy 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 of the particle in phase space is fully determined. We call  $E$  and  $V$  the system's “thermodynamic” state variables.<sup>14</sup> The particle moves between the two turning points  $x_{\pm}(E, V)$  and forms closed orbits in phase space with a 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 that temperature for classical systems is a measure of the kinetic energy of the particles, we take the temperature to be proportional to the kinetic energy averaged over a period

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

The Boltzmann constant  $k_B$  ensures the correct dimensions for the temperature  $T$ . For the conjugate force, we take the time average of  $-\partial\varphi/\partial V$ ,<sup>15</sup>

$$P(E, V) \equiv -\frac{1}{\tau(E, V)} \int_0^{\tau(E, V)} \frac{\partial\varphi(x(t; E, V); V)}{\partial V} dt. \quad (14)$$

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

Having identified the mechanical analogs of internal energy, external parameter, temperature, and conjugate force with the quantities  $E$ ,  $V$ ,  $T$ , and  $P$ , respectively, we can now ask whether a mechanical analog of the entropy  $S$  exists. To answer this question, we must find, according to statement 1, as expressed in Eq. (9), a function  $S(E, V)$  such that

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

That such a function exists is given by the following theorem.

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

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

The proof of the theorem is given in Appendix B and Ref. 10, pp. 45–46. The entropy  $S(E, V)$  can be rewritten compactly as

$$S(E, V) = k_B \log \oint p dx, \quad (17)$$

where  $p = \sqrt{2m(E - \varphi(x; V))}$  is the momentum of the particle when it is located at  $x$ : The integral  $\oint p dx$  is called the reduced action.<sup>15</sup> It is the area  $\Phi$  in phase space enclosed by the orbit of energy  $E$  and parameter  $V$ ,

$$S(E, V) = k_B \log \Phi(E, V), \quad (18)$$

where

$$\Phi(E, V) = \int_{H(x, p; V) \leq E} dp dx. \quad (19)$$

Helmholtz's theorem shows that there exists a mechanical counterpart of the entropy given by the logarithm of the phase space volume enclosed by the curve of constant energy  $H(x, p; V) = E$ .

That there exists a function  $S$  satisfying Eq. (15) is a remarkable result that shows there is a consistent one-dimensional mechanical model of thermodynamics. Although this model suggests the deep connection between classical mechanics and thermodynamic entropy, it is too

simple to model a real thermodynamic system composed of as many as  $10^{23}$  particles. Thus, it is necessary to generalize the Helmholtz theorem to multidimensional systems.

#### IV. ERGODICITY AND THE MICROCANONICAL ENSEMBLE

To generalize our model to more degrees of freedom, we need ergodicity and the microcanonical ensemble. The one-dimensional example of Sec. III can be used to introduce these important concepts.

The time average of a phase function  $f(x, p)$  over the orbit specified by  $E$  and  $V$  is

$$\langle f \rangle_t \equiv \frac{1}{\tau} \int_0^{\tau} f(x(t), p(t)) dt. \quad (20)$$

For simplicity, we have dropped the explicit dependence on  $E$ ,  $V$  of  $\tau$ ,  $x(t)$ ,  $p(t)$ , and  $\langle f \rangle_t$ . Because  $p = mv = mdx/dt$ , the differential  $dt$  is

$$dt = m \frac{dx}{p(x)}. \quad (21)$$

Using this differential, we obtain

$$\langle f \rangle_t = \frac{2m}{\tau} \int_{x_-}^{x_+} \frac{dx}{p(x)} f(x, p(x)). \quad (22)$$

The factor 2 is due to the particle going from  $x_-$  to  $x_+$  in a half period  $\tau/2$ . Now consider the integral

$$\int dp \delta(p^2/2m + \varphi(x; V) - E), \quad (23)$$

where  $\delta$  denotes the Dirac delta function. If we use the expression  $\delta(f(p)) = \sum_i \delta(p - p_i) / |f'(p_i)|$ , where the  $p_i$ 's are the zeroes of  $f(p)$  and  $\int \delta(p - p_i) dp = 1$ , we obtain for the integral,

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

Then, Eq. (22) becomes

$$\langle f \rangle_t = \frac{1}{\tau} \int dx \int dp \delta(p^2/2m + \varphi(x; V) - E) f(p, x), \quad (25)$$

where it is unnecessary to specify the integration limits  $x_{\pm}$  because they are implied by the Dirac delta function. The period  $\tau$  is given by

$$\begin{aligned} \tau &= \int_0^{\tau} dt = 2 \int_{x_-}^{x_+} \frac{dx}{p(x)} \\ &= \int dx \int dp \delta(p^2/2m + \varphi(x; V) - E). \end{aligned} \quad (26)$$

Hence, we arrive at

$$\langle f \rangle_t = \int dx \int dp \rho(x, p; E, V) f(p, x), \quad (27)$$

where we have introduced the phase space probability density function

$$\rho(x,p;E,V) = \frac{1}{\tau(E,V)} \delta(p^2/2m + \varphi(x;V) - E). \quad (28)$$

From Eq. (26), it is clear that  $\rho$  is properly normalized. The function  $\rho(x,p;E,V)$  is the microcanonical distribution. Equation (27) shows that the time average of a phase space function  $f(x,p)$  over a period is equal to its microcanonical average. This property is called ergodicity. All one-dimensional systems with a U-shaped potential are ergodic.

## V. MULTIPARTICLE MECHANICAL MODELS OF THERMODYNAMICS

We now extend the treatment for one degree of freedom to systems of  $N$ -particles in three dimensions with  $3N$  degrees of freedom. The Hamiltonian for an interacting system of  $N$ -particles of mass  $m$  is

$$H_N(\mathbf{q}, \mathbf{p}; V) = K_N(\mathbf{p}) + \varphi_N(\mathbf{q}; V), \quad (29)$$

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

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

$$\rho_N(\mathbf{q}, \mathbf{p}; E, V) = \frac{1}{\Omega_N(E, V)} \delta(E - H_N(\mathbf{q}, \mathbf{p}; V)), \quad (30)$$

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

$$\Omega_N(E, V) = \int \cdots \int \delta(E - H_N(\mathbf{q}, \mathbf{p}; V)) d\mathbf{q} d\mathbf{p}. \quad (31)$$

The microcanonical average  $\langle f \rangle_\mu$  of a phase function  $f(\mathbf{q}, \mathbf{p})$  is defined as

$$\langle f \rangle_\mu \equiv \int \cdots \int \rho_N(\mathbf{q}, \mathbf{p}; E, V) f(\mathbf{q}, \mathbf{p}) d\mathbf{q} d\mathbf{p}. \quad (32)$$

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

*Ergodic hypothesis.* For a 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  $\langle f \rangle_\mu$ ,

$$\langle f \rangle_t = \langle f \rangle_\mu. \quad (33)$$

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

$$T_N(E, V) \equiv \frac{2}{3Nk_B} \langle K_N \rangle_t, \quad (34)$$

which is twice the average kinetic energy per degree of freedom. The conjugate force is defined as

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

We ask, in agreement with statement 1 as expressed in Eq. (9), does there exist a function  $S_N(E, V)$  such that

$$\frac{\partial S(E, V)_N}{\partial E} = \frac{1}{T_N(E, V)}, \quad \frac{\partial S(E, V)_N}{\partial V} = \frac{P_N(E, V)}{T_N(E, V)}? \quad (36)$$

The following theorem gives the answer.

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

$$S_N(E, V) = k_B \log \Phi_N(E, V), \quad (37)$$

where

$$\Phi_N(E, V) \equiv \int \cdots \int_{H_N(\mathbf{q}, \mathbf{p}) \leq E} d\mathbf{q} d\mathbf{p}. \quad (38)$$

The proof, which is based on the equipartition theorem, is given in Appendix C and Ref. 2. Unlike the temperature and conjugate force,  $S_N$  is not a time average of some phase function  $f(\mathbf{q}, \mathbf{p})$ .

This theorem shows that ergodic systems constitute ideal mechanical models of thermodynamics. The state variables  $E$  and  $V$  define their thermodynamic state. In addition, their temperature and conjugate force can be defined 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 the generalized Helmholtz theorem,  $S_N$  is consistent with the first law of thermodynamics and statement 1 of the second law of thermodynamics. Is this construction also consistent with statements 2 and 3 of the second law of thermodynamics?

For  $S_N$  to be consistent with statement 2, it is necessary for it to satisfy the following condition. If the parameter  $V$  is varied very slowly in time from a  $V_0 = V(t_0)$  to  $V_f = V(t_f)$ , the corresponding change in the entropy  $S_N$  is zero. The time of variation must be much slower than any time scale of the system's dynamics. By allowing for a time dependence of  $V$ , the Hamiltonian function of the system is also time dependent, and hence energy is not conserved. Consider the initial system at  $t = t_0$  with phase space point  $\mathbf{q}_0, \mathbf{p}_0$ . The system then evolves under the time-dependent Hamiltonian

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

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 have made explicit the dependence on the initial condition of the evolved phase space point. 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<sup>16</sup> that for ergodic systems the energy  $E_f$  reached at the end of a very slow change depends only on the initial energy  $E_0 = H_N(\mathbf{q}_0, \mathbf{p}_0; V(t_0))$ . The final energy  $E_f$  is determined by solving the equation

$$\Phi_N(E_0, V_0) = \Phi_N(E_f, V_f). \quad (40)$$

Thus, the quantity  $\Phi_N(E, V)$  does not change when  $V$  is varied infinitely slowly in time. In classical mechanics this property is called adiabatic invariance. Because the entropy 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 also an adiabatic invariant. Therefore, the entropy does not change if  $V$  is varied very slowly in time and thus  $S_N$  complies with statement 2. A proof of adiabatic invariance of  $\Phi_N$  is given in Appendix D and Ref. 16, pp. 27–30.



It is also possible to prove that, in an averaged sense,  $S_N$  complies with statement 3 as well.<sup>5,6</sup> In this case the average change in entropy must be considered because the final energy is not uniquely determined by the initial energy for fast transformations. Depending on the initial conditions, different final energies can be obtained and hence different final entropies.

## VII. THE BOLTZMANN PRINCIPLE

For a system composed of a very large number  $N$  of particles that interact via short-range forces, the phase space volume  $\Phi_N(E)$  approaches an exponential dependence on  $E$ ,  $\Phi_N(E) \propto e^E$ . Because  $\Omega_N = \partial\Phi_N/\partial E$  [see Eq. (31)], we also have  $\Phi_N \propto \Omega_N$  (see Ref. 17, p. 148).

The quantity  $\Omega_N$  in Eq. (31) is the measure of the shell of constant energy  $H_N(\mathbf{q}, \mathbf{p}; V) = E$ . As such, it is proportional to the number  $\mathcal{W}$  of microstates compatible with a given energy  $E$ . According to semiclassical theory each microstate occupies a volume  $h^{3N}$  of phase space, where  $h$  is Planck's constant.<sup>18</sup> By introducing an arbitrary energy scale  $\Delta E$ , the number of microstates  $\mathcal{W}$  is given by  $\mathcal{W} = \Omega_N \Delta E / h^{3N}$ . Thus, for very large  $N$  we obtain the proportionalities,

$$\Phi_N \propto \Omega_N \propto \mathcal{W} \quad (N \gg 1). \quad (41)$$

We take the logarithm and obtain

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

except for an irrelevant constant. Equation (42) shows that for large ergodic systems composed of many particles interacting via short range forces, the differential of the Boltzmann entropy is equal to the differential  $\delta Q/T$ . Hence, it can be identified with the Clausius entropy, which gives a proof of the 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, we can ask whether, as prescribed by the heat theorem, the combination

$$\frac{dE + P_N dV}{T_N} \quad (43)$$

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

The entropy in Eq. (37) is sometimes referred to as the Hertz entropy.<sup>19,20</sup> Hertz<sup>21,22</sup> derived it from the requirement of adiabatic invariance (see also Refs. 16, 23, and 24), whereas we have derived it from the heat theorem. The fundamental character of the entropy in Eq. (37) 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.<sup>27,28</sup> The entropy in Eq. (37) also appears in Gibbs seminal book.<sup>29</sup> However, its connection with the heat theorem has not been recognized previously.

The most crucial point in the derivation of the Boltzmann principle is the introduction of the ergodic hypothesis. Although this hypothesis is generally believed to hold for real macroscopic systems, its proof is a formidable challenge which has been achieved only in few special cases.<sup>30</sup> A proof that a gas of elastically colliding hard spheres is ergodic was announced in 1963 by Sinai.<sup>31</sup> However, the full proof was not published and the problem is still open (see Ref. 32 for a more detailed discussion). Nevertheless, the ergodicity of hard spheres systems is plausible as indicated by recent numerical simulations (see Sec. IV of Ref. 33).

In regard to these difficulties, we note that the present derivation of the Boltzmann principle does not use the fact that the average of any arbitrary phase function is equal to its microcanonical average, as required by the ergodic hypothesis. It only uses the fact that the time average of  $K$  and  $-\partial\phi/\partial V$  is equal to their microcanonical averages (see the proof of the generalized Helmholtz theorem in Appendix C). Thus, the ergodic hypothesis can be replaced by the following hypothesis.

*Weak ergodic hypothesis.* For given  $E$  and  $V$ , the time averages  $\langle K \rangle_t$  and  $\langle \partial\phi/\partial V \rangle_t$  are uniquely determined and are equal to their microcanonical averages, that is,

$$\langle K \rangle_t = \langle K \rangle_\mu, \quad (44)$$

$$\langle \partial\phi/\partial V \rangle_t = \langle \partial\phi/\partial V \rangle_\mu. \quad (45)$$

In summary, ergodicity is too stringent a condition and the weaker ergodic condition for only the temperature and pressure is sufficient to obtain the Boltzmann principle.<sup>34</sup>

## ACKNOWLEDGMENTS

The authors wish to thank the Texas Section of the American Physical Society for the Robert S. Hyer Recognition for Exceptional Research award presented at its Fall 2008 meeting. They also thank Cosimo Gorini for reading the manuscript. They are grateful to Professor Randall B. Shirts for providing the translation of Ref. 25 and for his comments. Valuable remarks from anonymous referees are also gratefully acknowledged.

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

A differential form  $\omega$  (more precisely, a 1-form) in a connected subset  $\mathcal{A}$  of  $\mathbb{R}^2$  is formally written as

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

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  $\mathbb{R}^2$ .<sup>35</sup>

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

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

the integral of  $\omega$  along the curve  $\psi$  is defined as

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

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

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

$$\omega = dG, \quad (\text{A4})$$

that is,

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

where  $G$  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 two points  $A$  and  $B$  in  $\mathcal{A}$  and curves  $\psi$  and  $\phi$  in  $\Sigma(A, B)$  it satisfies

$$\int_{\psi} \omega = \int_{\phi} \omega. \quad (\text{A6})$$

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

$$\int_{\psi} \omega = \int_{\psi} dG = G(B) - G(A). \quad (\text{A7})$$

The following statement also holds: A differential form is exact if and only if its integral along any closed curve is zero. If the functions  $\mathcal{M}$  and  $\mathcal{N}$  are of class  $C^1$  (that is, they are differentiable), then a necessary condition for the form  $\omega$  to be exact is that

$$\frac{\partial \mathcal{M}}{\partial x_2} = \frac{\partial \mathcal{N}}{\partial x_1}. \quad (\text{A8})$$

In this case the differential form  $\omega$  is said to be *closed*.

## APPENDIX B: PROOF OF HELMHOLTZ'S THEOREM

For a one-dimensional system confined in a U-shaped potential the period  $\tau$  of the orbit is equal to the derivative with respect to energy of the phase space area  $\Phi$  enclosed by the orbit<sup>15</sup>

$$\tau = \frac{\partial \Phi}{\partial E}. \quad (\text{B1})$$

A simple way to prove this relation is by expressing the area as  $\Phi(E, V) = \int \theta(E - H(x, p; V)) dp dx$ , where  $\theta(x)$  is the Heaviside step function [ $\theta(x) = 1$  if  $x \geq 0$ ,  $\theta(x) = 0$  if  $x < 0$ ]. If we take the derivative with respect to  $E$  and use the relation  $\delta(x) = d\theta(x)/dx$ , we obtain  $\tau$  [see Eq. (26)]. By using Eqs. (B1) and (18), we have

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

From Eq. (22), we obtain the relation

$$2\langle K \rangle_t = \frac{\Phi}{\tau}, \quad (\text{B3})$$

from which we obtain

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

using Eq. (B2). Similarly, we also obtain

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

From Eqs. (13) and (14), we obtain

$$\frac{\partial S}{\partial E} = \frac{1}{T}, \quad (\text{B6})$$

$$\frac{\partial S}{\partial V} = \frac{P}{T}. \quad (\text{B7})$$

## APPENDIX C: PROOF OF THE GENERALIZED HELMHOLTZ THEOREM

The proof of generalized Helmholtz theorem makes use of the multidimensional version of Eq. (B1),

$$\Omega_N = \frac{\partial \Phi_N}{\partial E}, \quad (\text{C1})$$

which can be proved, in a way similar to Eq. (B1), by expressing  $\Phi_N$  as  $\int \theta(E - H(\mathbf{q}, \mathbf{p}; V)) d\mathbf{q} d\mathbf{p}$  and using the relation  $\delta(x) = d\theta(x)/dx$ . The equipartition theorem,<sup>17</sup>

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

is the generalization of Eq. (B3) to many dimensions. We use Eqs. (C2) and (C1) with Eq. (37) and obtain

$$\frac{\partial S_N}{\partial E} = \frac{3Nk_B}{2\langle K_N \rangle_{\mu}}. \quad (\text{C3})$$

In a similar way, we obtain

$$\frac{\partial S_N}{\partial V} = -\frac{3Nk_B}{2\langle K_N \rangle_{\mu}} \left\langle \frac{\partial \varphi_N}{\partial V} \right\rangle_{\mu}. \quad (\text{C4})$$

By using Eqs. (34) and (35) with the ergodic hypothesis, we finally arrive at

$$\frac{\partial S_N}{\partial E} = \frac{1}{T_N}, \quad (\text{C5})$$

$$\frac{\partial S_N}{\partial V} = \frac{P_N}{T_N}. \quad (\text{C6})$$

## APPENDIX D: PROOF OF ADIABATIC INVARIANCE OF $\Phi_N$

To prove that  $\Phi_N$  in Eq. (38) is an adiabatic invariant, we use the time-dependent Hamiltonian

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

We take the total time derivative of the Hamiltonian  $H_N(\mathbf{q}, \mathbf{p}; V(t))$  in Eq. (D1),

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

where the terms involving  $\dot{\mathbf{q}}$  and  $\dot{\mathbf{p}}$  cancel by Hamilton's equations.<sup>36</sup> 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

$$\left\langle \frac{dH_N}{dt} \right\rangle_\mu = \left\langle \frac{\partial H_N}{\partial V} \right\rangle_\mu \frac{dV}{dt}. \quad (\text{D3})$$

By Liouville's theorem<sup>36</sup> the average on the left-hand side of Eq. (D3) is

$$\left\langle \frac{dH_N}{dt} \right\rangle_\mu = \frac{dE}{dt}. \quad (\text{D4})$$

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

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

where  $\Phi_N$  and  $\Omega_N$  are given in Eqs. (31) and (38), respectively. If we substitute Eqs. (D4) and (D5) into Eq. (D3) and use  $\Omega_N = \partial \Phi_N / \partial E$ , we obtain

$$\frac{d\Phi_N}{dt} \equiv \frac{\partial \Phi_N}{\partial E} \frac{dE}{dt} + \frac{\partial \Phi_N}{\partial V} \frac{dV}{dt} = 0, \quad (\text{D6})$$

which shows that  $\Phi_N$  is constant and therefore an adiabatic invariant.

## APPENDIX E: SUGGESTED PROBLEM

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

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

- Calculate the area  $\Phi(E, V)$  enclosed by the trajectory of energy  $E$  and angular frequency  $V$ . Use Eq. (B1) and check that the period of the orbit is, as expected, given by  $\tau(E, V) = 2\pi/V$ .
- Use Eqs. (13) and (14) to show that  $k_B T(E, V) = E$  and  $P(E, V) = -E/V$ .
- Show that the differential form  $dE + PdV$ , with  $P(E, V)$  as in part (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 zero.
- Consider the differential form  $\omega = (1/T)dE + (P/T)dV$ , with  $P(E, V)$  and  $T(E, V)$  as in part (b). Find a primitive function  $S(E, V)$  for  $\omega$ . 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.

<sup>a</sup>Electronic mail: michele.campisi@physik.uni-augsburg.de

<sup>b</sup>Electronic mail: kobe@unt.edu

<sup>1</sup>A. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949), p. 142.

<sup>2</sup>M. Campisi, "On the mechanical foundations of thermodynamics: The generalized Helmholtz theorem," *Stud. Hist. Philos. Mod. Phys.* **36**, 275–290 (2005).

<sup>3</sup>H. Helmholtz, "Prinzipien der statik monocyclischer Systeme," *Borchardt-Crelles Journal für die reine und angewandte Mathematik* **97**, 111–140 (1984); also in *Wissenschaftliche Abhandlungen*, edited by G. Wiedemann (Johann Ambrosius Barth, Leipzig, 1985), Vol. 3, pp. 142–162, 179–202.

<sup>4</sup>L. Boltzmann, "Über die Eigenschaften monocyclischer und anderer damit verwandter Systeme," *Crelles Journal* **98**, 68–94 (1884); also L. Boltzmann, in *Wissenschaftliche Abhandlungen*, edited by F. Hasenohrl (J. A. Barth, Leipzig, 1909), Vol. 3, pp. 122–152, reissued Chelsea, New York, 1969.

<sup>5</sup>M. Campisi, "Statistical mechanical proof of the second law of thermodynamics based on volume entropy," *Stud. Hist. Philos. Mod. Phys.* **39**, 181–194 (2008).

<sup>6</sup>M. Campisi, "Increase of Boltzmann entropy in a quantum forced harmonic oscillator," *Phys. Rev. E* **78**, 051123–1–10 (2008).

<sup>7</sup>J. Uffink, "Bluff your way in the second law of thermodynamics," *Stud. Hist. Philos. Mod. Phys.* **32** (3), 305–394 (2001).

<sup>8</sup>H. B. Callen, *Thermodynamics* (Wiley, New York, 1960).

<sup>9</sup>D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford U. P., Oxford, 1987).

<sup>10</sup>G. Gallavotti, *Statistical Mechanics: A Short Treatise* (Springer-Verlag, Berlin, 1999).

<sup>11</sup>A path in the space of state variables corresponds to a transformation that leads the system from  $A$  to  $B$  through a sequence of almost equilibrium states. Consequently, the change  $\delta Q$  is understood to be a quasi-static change.

<sup>12</sup>For the common case of work due to expansion or compression the external parameter  $V$  is the volume and its conjugate force  $P$  is the pressure. In general,  $V$  and  $P$  stand for any conjugate pair of "displacement" and force depending on the nature of the work done on the system (for example, magnetic field and magnetization).

<sup>13</sup>In other words, the field  $\vec{F} \equiv (1/T, P/T)$  is a conservative vector field. That is, there exists a potential function  $S(E, V)$  such that  $\vec{F} = \vec{\nabla} S$ , where  $\vec{\nabla}$  is the gradient operator in the space  $(E, V)$ . For this reason, the entropy can be understood as a thermodynamic potential.

<sup>14</sup>The thermodynamic state variables  $(E, V)$  should not be confused with the mechanical state variables  $(p, x)$ .

<sup>15</sup>L. Landau and E. Lifshitz, *Mechanics* (Pergamon, Oxford, 1960).

<sup>16</sup>V. L. Berdichevsky, *Thermodynamics of Chaos and Order* (Addison-Wesley Longman, Essex, 1997).

<sup>17</sup>K. Huang, *Statistical Mechanics*, 2nd ed. (Wiley, New York, 1987).

<sup>18</sup>L. Landau and E. Lifshitz, *Statistical Physics*, 2nd ed. (Pergamon, Oxford, 1969).

<sup>19</sup>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**, 011120–1–7 (2006).

<sup>20</sup>A. Dib, "Does the Boltzmann principle need a dynamical correction?," *J. Stat. Phys.* **117**, 581–597 (2004).

<sup>21</sup>P. Hertz, "Über die mechanischen Grundlagen der Thermodynamik," *Ann. Phys.* **33**, 225–274 (1910).

<sup>22</sup>P. Hertz, "Über die mechanischen Grundlagen der Thermodynamik," *Ann. Phys.* **33**, 537–552 (1910).

<sup>23</sup>A. Münster, *Statistical Thermodynamics* (Springer-Verlag, Berlin, 1969), Vol. 1.

<sup>24</sup>H. H. Rugh, "Microthermodynamic formalism," *Phys. Rev. E* **64** (5), 055101–1–4 (2001).

<sup>25</sup>A. Schlüter, "Zur Statistik klassischer Gesamtheiten," *Z. Naturforsch. A* **3A**, 350–360 (1948). An English translation is available at [people.chem.byu.edu/rbshirts/research/schluter1948translation.doc](http://people.chem.byu.edu/rbshirts/research/schluter1948translation.doc).

<sup>26</sup>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).

<sup>27</sup>P. Talkner, P. Hänggi, and M. Morillo "Microcanonical quantum fluctuation theorems," *Phys. Rev. E* **77**, 051131–1–6 (2008).

<sup>28</sup>The conditions  $\Phi \geq 0$ ,  $\partial_E \Phi = \Omega \geq 0$  ensure that the temperature derived from the Hertz entropy, that is,  $T = (\partial_E \log \Phi)^{-1} = \Phi/\Omega$  is positive definite.

<sup>29</sup>J. Gibbs, *Elementary Principles in Statistical Mechanics* (Yale U. P., New Haven, 1902), reprinted by Dover, New York, 1960.

- <sup>30</sup>J. L. Lebowitz and O. Penrose, "Modern ergodic theory," *Phys. Today* **26** (2), 23–29 (1973).
- <sup>31</sup>Ya. G. Sinai, "On the foundation of the ergodic hypothesis for a dynamical system of statistical mechanics," *Sov. Math. Dokl.* **4**, 1818–1822 (1963).
- <sup>32</sup>J. Uffink, "Compendium of the foundations of classical statistical physics," in *Philosophy of Physics*, edited by J. Butterfield and J. Earman (Elsevier, Amsterdam, 2007) ([philsci-archive.pitt.edu/archive/00002691/](http://philsci-archive.pitt.edu/archive/00002691/)).
- <sup>33</sup>M. Campisi, P. Talkner, and P. Hänggi, "Finite bath fluctuation theorem," *Phys. Rev. E* **80**, 031145-1–9 (2009).
- <sup>34</sup>The simplest example of a nonergodic system that satisfies this weaker

condition is a one-dimensional 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  $\gamma_l$  and  $\gamma_r$ , with the phase space trajectory covering only one of them. (For energy above  $E_c$  there is only one trajectory and ergodicity holds.) Because  $\gamma_l$  and  $\gamma_r$  are mirror images of each other, the time averages of even functions of  $x$ , such as  $K$  and  $-\partial\phi/\partial V$ , are not affected by which of the two curves the motion follows. Thus, such averages can be calculated as microcanonical averages over the curve  $\gamma_l \cup \gamma_r$ .

- <sup>35</sup>In the main text the role of  $(x_1, x_2)$  is played by the state variables  $(E, V)$ .
- <sup>36</sup>H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, San Francisco, 2002), pp. 337, 419–421.



Dribble Glasses. More properly, these two glasses in the collection of Union College in Schenectady, New York, are examples of Tantalus cup or the automatic siphon. When the water level reaches the bend in the internal glass tube, the siphon starts and the glass empties itself of water. (Photograph and Notes by Thomas B. Greenslade, Jr., Kenyon College)