

Microthermodynamic formalism

Hans Henrik Rugh

Department of Mathematics, University of Cergy-Pontoise, F-95302 Cergy-Pontoise Cedex, France

(Received 9 May 2001; published 15 October 2001)

We consider the microcanonical ensemble of a classical Hamiltonian dynamical system, the Hamiltonian being parameter dependent and in the possible presence of other first integrals. We describe a thermodynamic formalism in which a first law of thermodynamics, or fundamental relation, is based upon the bulk-entropy, S_Ω . Under an ergodic hypothesis, S_Ω is shown to be an adiabatic invariant. Expressions for derivatives and thermodynamic relations are derived within the microcanonical ensemble itself.

DOI: 10.1103/PhysRevE.64.055101

PACS number(s): 05.20.Gg, 05.45.-a, 05.70.-a, 02.40.Vh

Equilibrium properties of an isolated Hamiltonian dynamical system with many, say 10^{24} , degrees of freedom is probably best described using a thermodynamic formalism for a canonical ensemble at a fixed temperature, even though this means introducing fluctuations in an otherwise conserved quantity, the energy. For a more moderate number of degrees of freedom, say $10^3 - 10^8$, numerical simulations become feasible, and it is desirable to obtain a description in terms of the microcanonical ensemble itself where the values of the first integrals are fixed quantities. In such an approach, geometrical properties of the level surfaces reflect thermodynamic relations and, by invoking the ergodic hypothesis, also dynamical properties of the underlying system. In particular, when energy is the only first integral, measurements may then be done by time averaging (cf. Refs. [1,2]). The purpose of this Rapid Communication is to develop such a microthermodynamic formalism further, taking into account parameter dependency and the presence of other first integrals. Within this framework we will also (Sec. II) discuss a natural formulation of a first law of thermodynamics, or fundamental relation, based upon the bulk-entropy, S_Ω . We refer to Abraham and Marsden [3] as well as Landau and Lifshitz [4] for a general introduction to thermodynamic ensembles, to Lebowitz *et al.* [5] for an illustrative example of some differences between the ensembles and to Evans and Morriss [6] for practical calculations carried out in the micro canonical ensemble. We also refer to Jepps *et al.* [7] where the presence of other (approximative) first integrals is of relevance and to Otter [8] who studied reaction events using mixed ensemble averages.

I. MICROCANONICAL ENSEMBLES

For simplicity, we consider a Euclidean phase space, $\Omega = \mathbb{R}^{2d}$, $d \geq 1$, and a Hamiltonian function, $H: \Omega \rightarrow \mathbb{R}$, bounded from below and of sufficient rapid growth at infinity. The dynamics preserves the Liouville measure, here the Lebesgue measure, $m = d^d x d^d p$. There may be other first integrals, denoted $F = F_1, \dots, F_m$, $m \geq 0$. Note that in Sec. III we shall write $F_0 = H$ for the Hamiltonian which is then considered at the same footing as the other first integrals. All first integrals are assumed to be in involution. We also assume that all functions are known analytically and that the Hamiltonian depends smoothly on some external real param-

eters, denoted $\Lambda = \Lambda_1, \dots, \Lambda_n$. By contrast we do not allow the other first integrals to depend on Λ . This is for technical reasons (cf. below), though in some cases such a condition could be relaxed. For fixed values of parameters, $\Lambda = \lambda$, of first integrals, $F = I$, and of the energy, $H_\lambda = E$, the subspace, $A = A[E, I, \lambda] = \{\xi \in \mathbb{R}^{2d}: H_\lambda(\xi) = E, F(\xi) = I\}$, is invariant under the dynamics of H_λ . We will assume that values are chosen so that the differentials, dH, dF_1, \dots, dF_m , are all independent on $A[E, I, \lambda]$. This in particular implies that A is a smooth co-dimension $m + 1$ submanifold of our Euclidean space.

In the literature one will find (at least) two definitions (denoted bulk and surface) of a microcanonical entropy and temperature. It turns out that we shall need both. Thus we define

$$e^{S_\Omega(E, I, \lambda)} \equiv \int m \Theta(E - H_\lambda) \delta(I - F), \quad (1)$$

where Θ denotes the Heaviside function and

$$e^{S_\mu(E, I, \lambda)} \equiv \int m \delta(E - H_\lambda, I - F). \quad (2)$$

The bulk- and the surface-temperature are then given by

$$\frac{1}{T_\Omega} = \frac{\partial S_\Omega}{\partial E}, \quad \frac{1}{T_\mu} = \frac{\partial S_\mu}{\partial E}. \quad (3)$$

Derivatives with respect to other first integrals are considered in Sec. III. We also have generalized bulk- and surface-pressures

$$p_\Omega^i = T_\Omega \frac{\partial S_\Omega}{\partial \lambda_i}, \quad p_\mu^i = T_\mu \frac{\partial S_\mu}{\partial \lambda_i}, \quad i = 1, \dots, n. \quad (4)$$

Taking an average in the microcanonical ensemble will here mean taking the surface-average, i.e.,

$$\langle \phi \rangle_\mu \equiv \langle \phi | E, I, \lambda \rangle = \frac{\int m \delta(E - H_\lambda, I - F) \phi}{\int m \delta(E - H_\lambda, I - F)}. \quad (5)$$

In practice, $\langle \phi \rangle_\mu$, is often calculated by time-averaging (assuming ergodicity), thus giving a dynamical preference to the surface-average relative to other ensemble-averages.

The calculation of either of the two entropies may be difficult or even impossible when the number of degrees of freedom in the system is large. On the other hand, the associated temperatures and generalized pressures may be probed using time-averaging (cf. Sec. III). A bulk-pressure, p_Ω^i , may be calculated as follows:

$$p_\Omega^i = - \left\langle \frac{\partial H_\lambda}{\partial \lambda_i} \right\rangle_\mu = - \left\langle \frac{\partial H_\lambda}{\partial \lambda_i} \middle| E, I, \lambda \right\rangle. \quad (6)$$

To see this we note that the derivative of a Heaviside-function yields a δ -function. It follows that $1/T_\Omega = \partial S_\Omega / \partial E = e^{S_\mu} / e^{S_\Omega}$, and therefore,

$$p_\Omega^i = T_\Omega \frac{\partial S_\Omega}{\partial \lambda_i} = - \frac{\int m \delta(E - H_\lambda, I - F) (\partial H_\lambda / \partial \lambda_i)}{\int m \delta(E - H_\lambda, I - F)}. \quad (7)$$

By the very definition, there is always a first law of thermodynamics for the bulk-entropy:

$$T_\Omega dS_\Omega = dE + \sum_i p_\Omega^i d\lambda_i. \quad (8)$$

The same kind of relation is, of course, valid for the surface-entropy but, as we shall see, adiabatic invariance clearly gives a preference to the version involving the bulk-entropy.

II. ADIABATIC INVARIANCE

An adiabatic process is characterized by a slow and smooth change in parameters during which ‘thermal equilibrium’ is maintained. For example, slowly moving a piston of a cylinder containing a gas of particles, in order to avoid what is known as parametric resonance, we fix a smooth path in parameter space, $s \in [0, 1] \mapsto c(s) \in \mathbb{R}^n$, and traverse this path in rescaled time, i.e., $t \in [0, \Delta] \mapsto \lambda(t) = c(t/\Delta)$, for some $\Delta > 0$. A physical trajectory, $\xi(t)$, is evolved from an (almost) arbitrary point, $\xi(0) \in A[E, I, \lambda]$, using the time- and Δ -dependent Hamiltonian, $H_{\lambda(t)}$. The goal is to determine the microcanonical state, i.e., the values of (E, I) , of the system at time $t = \Delta$. The process is called adiabatic if this state has a well-defined limit as $\Delta \rightarrow \infty$, the limit depending on the initial state (E, I, λ) and the path, but not on the choice of initial point, cf. Arnold [9], section 52. Here again, we do not permit the other first integrals to depend on Δ . Being time-independent and commuting with the Hamiltonian, they therefore remain constant under the time-involution. The energy, $E(t) = H_{\lambda(t)}(\xi(t))$, being time-dependent through λ , is in general not constant in time. Instead we get by Hamilton’s equations: $dE/dt = \sum_i \partial H_{\lambda_i} / \partial \lambda_i (d\lambda_i/dt)$, and the ‘adiabatic’ problem is then to see if $E(\Delta)$ has a well-defined limit as $\Delta \rightarrow \infty$.

In order to attain the adiabatic limit, Δ should be very large, in particular larger than the time-scale, τ_{erg} , over

which ergodic averaging takes place. We may then look at the time-average of an observable over an intermediate time-scale, $\tau_{\text{erg}} \ll \delta t \ll \Delta$. This time-scale is short enough so that λ does not change significantly but long enough so that a dynamical average may be replaced by a microcanonical average. Under this ‘adiabatic ergodic hypothesis’ we get for the energy-change

$$\frac{1}{\delta t} \int_t^{t+\delta t} \frac{dH_{\lambda(t)} \circ \xi(t)}{dt} dt \approx \left\langle \frac{dH_{\lambda(t)}}{dt} \middle| E(t), I, \lambda(t) \right\rangle. \quad (9)$$

Over the time-scale, δt , the λ -derivative is almost constant (because of the scaling with Δ) and may therefore be taken outside the average:

$$\frac{E(t+\delta t) - E(t)}{\delta t} \approx \sum_i \left\langle \frac{\partial H}{\partial \lambda_i} \middle| E(t), I, \lambda(t) \right\rangle \frac{d\lambda_i}{dt}, \quad (10)$$

an expression which is correct to order τ_{erg}/Δ . Taking the $\Delta \rightarrow \infty$ limit we get an identity between differentials,

$$dE = \sum_i \left\langle \frac{\partial H}{\partial \lambda_i} \middle| E, I, \lambda \right\rangle d\lambda_i, \quad (11)$$

valid precisely in the adiabatic limit. By Eqs. (6) and (8), we see that the bulk-entropy, S_Ω , is indeed an adiabatic invariant. In 1+1 dimensions the bulk-entropy is just the action integral and the phenomena is well-known (adiabatic invariance of the action, cf. Arnold [9]). In higher dimensions a similar result was obtained by Kazuga [10], though in a different context.

Now, the bulk-entropy is *a fortiori* strictly increasing as a function of the energy (at given parameter- and first integral-values). One particular gratifying, though nontrivial consequence is that by traversing a loop adiabatically the energy must return to its original value.

It is easy to put the adiabatic ergodic hypothesis into a more rigorous form (in terms of decay of correlation functions). It is, however, virtually impossible to check analytically if such a condition really holds in a given situation. One serious problem is that critical slowing down (meaning that τ_{erg} diverges) occurs if one encounters an additional first integral along the traversed path. On the other hand, singularities in the energy surface are likely to pose problems only in low dimensions (notably 1+1).

Should one allow other first integrals to depend on the parameters, the bulk-entropy, as we have defined it, is in general no longer an adiabatic invariant. We have not been successful in finding a good replacement for the bulk-entropy (the reader is encouraged to try for himself) and consequently not allowed such a parameter dependence.

III. SURFACE DERIVATIVES

Close to equilibrium we may express response functions as derivatives of averages with respect to parameters and values of the first integrals (in this section this includes the Hamiltonian). When the Hamiltonian is the only first integral and there is no parameter dependence, it was shown in [1,2]

(cf. also [7,11]) that energy derivatives of a microcanonical average may themselves be calculated as averages. The general case turns out to be quite similar.

Our assumption on the first integrals being independent means that in a neighborhood of the integral surface, $A[E, I, \lambda]$, we may construct vector fields, $\mathbf{X}_0, \dots, \mathbf{X}_m$, for which

$$dF_i(\mathbf{X}_j) = \delta_{ij}, \quad i, j = 0, \dots, m. \quad (12)$$

Geometrically, the vector field, \mathbf{X}_i , is transversal to the $F_i = I_i$ surface but parallel to the other surfaces, $F_j = I_j$, $j \neq i$. We make the observation that if $f(x)$ is a (suitably smooth) function and \mathbf{V} is a vector field for which $df(\mathbf{V}) \equiv 0$, then the Lie-derivative, $L_{\mathbf{V}}\delta(f(x)) = (\mathbf{V} \cdot \nabla)\delta(f(x)) \equiv 0$, vanishes identically. This is clearly true if the δ -function were a smooth function and the claim then follows by approximation (a rigorous proof is quite lengthy). This applies to our Hamiltonian setup, since for all $i, j = 0, \dots, m$

$$\left(\frac{\partial}{\partial I_i} + (\mathbf{X}_i \cdot \nabla) \right) (I_j - F_j) = \delta_{ij} - \delta_{ij} = 0. \quad (13)$$

Hence, if ϕ is any smooth function,

$$\int m \phi \left(\frac{\partial}{\partial I_i} + (\mathbf{X}_i \cdot \nabla) \right) \delta(I - F) = 0. \quad (14)$$

Taking the I_i derivative outside the integral and carrying the Lie-derivative out by partial integration we obtain

$$\frac{\partial}{\partial I_i} \int m \delta(I - F) \phi = \int m \delta(I - F) \nabla \cdot (\phi \mathbf{X}_i). \quad (15)$$

The identity, $[\partial/\partial \lambda_k + \partial H_\lambda / \partial \lambda_k (\partial/\partial E)] \delta(I - F) \equiv 0$ (recall that only the Hamiltonian depends on Λ), also implies

$$\frac{\partial}{\partial \lambda_k} \int m \delta(I - F) \phi = - \int m \delta(I - F) \nabla \cdot \left(\phi \frac{\partial H_\lambda}{\partial \lambda_k} \mathbf{X}_0 \right). \quad (16)$$

The surface entropy is the logarithm of a surface integral. Hence, when taking derivatives a normalization factor appears which precisely turns the derivative into a microcanonical average. Thus, for the inverse ‘‘generalized’’ surface-temperatures, $\beta_\mu^i = \partial S_\mu / \partial I_i$ (with $\beta_\mu^0 = 1/T_\mu$), we get using Eq. (15) and setting $\phi = 1$,

$$\beta_\mu^i = \frac{\frac{\partial}{\partial I_i} \int m \delta(I - F) \phi}{\int m \delta(I - F) \phi} = \langle \nabla \cdot \mathbf{X}_i | E, I, \lambda \rangle. \quad (17)$$

Similarly for the ‘‘generalized’’ pressures, using Eq. (16),

$$p_\mu^k = \frac{\partial S_\mu}{\partial \lambda_k} = - \left\langle \nabla \cdot \left(\frac{\partial H_\lambda}{\partial \lambda_k} \mathbf{X}_0 \right) \right| E, I, \lambda \rangle. \quad (18)$$

When taking a derivative of an average and writing $\delta = \delta(I - F)$ we have

$$\frac{\partial}{\partial I_i} \langle \phi \rangle_\mu = \frac{\frac{\partial}{\partial I_i} \int m \delta \phi}{\int m \delta} - \frac{\frac{\partial}{\partial I_i} \int m \delta \phi}{\int m \delta} \frac{\frac{\partial}{\partial I_i} \int m \delta}{\int m \delta}, \quad (19)$$

which by the definition of averages and generalized temperatures reduces to

$$\frac{\partial}{\partial I_i} \langle \phi \rangle_\mu = \langle \nabla \cdot (\phi \mathbf{X}_i) \rangle_\mu - \beta_\mu^i \langle \phi \rangle_\mu. \quad (20)$$

We calculate in the same way,

$$\frac{\partial}{\partial \lambda_k} \langle \phi \rangle_\mu = - \left\langle \nabla \cdot \left(\phi \frac{\partial H_\lambda}{\partial \lambda_k} \mathbf{X}_0 \right) \right|_\mu - p_\mu^k \langle \phi \rangle_\mu. \quad (21)$$

These thermodynamic identities provide the natural generalizations of the results found in [1,2]. In the above formulas the other first integrals may, in fact, be allowed to be parameter dependent (essentially because there is no Heaviside function in the above). The straight-forward derivation of formulas in this case is no more difficult and left to the reader.

We also note that if the Liouville measure has the form $dm = \rho(\xi) d^{2d} \xi$ (for instance, in local coordinates on a symplectic manifold), the only change in the above formulas is to replace the divergence of a vector field, $\nabla \cdot \mathbf{V}$, by $(1/\rho) \nabla \cdot (\rho \mathbf{V})$ (cf. Ref. [2]).

IV. BULK DERIVATIVES

Should one wish to take derivatives in the bulk-ensemble the procedure is slightly different. Restricting our attention to the energy derivative, suppose that Y is a smooth vector field such that [compare with Eq. (12)]

$$\nabla \cdot \mathbf{Y} \equiv 1 \quad \text{and} \quad dF_i(\mathbf{Y}) \equiv 0, \quad 1 \leq i \leq m. \quad (22)$$

One may certainly find such \mathbf{Y} when there are no other first integrals present. In the general case it is less clear because here we need the vector field to be defined ‘‘smoothly’’ throughout the ‘‘bulk.’’ Assuming, however, that we have found such a vector field we note that

$$e^{S_\Omega} = \int m (\nabla \cdot \mathbf{Y}) \Theta(E - H) \delta(I - F), \quad (23)$$

which by partial integration yields

$$e^{S_\Omega} = \int m dH_\lambda(\mathbf{Y}) \delta(E - H, I - F). \quad (24)$$

In particular, we obtain the following formula (which is well-known, when energy is the only first integral) for the bulk-temperature:

$$T_\Omega = e^{S_\Omega - S_\mu} = \langle dH_\lambda(\mathbf{Y}) | E, I, \lambda \rangle. \quad (25)$$

For comparison we note that in the canonical ensemble where an integral of ϕ looks as follows:

$$\int m e^{-\beta H_\lambda} \delta(I-F) \phi, \quad (26)$$

the same calculation shows that

$$T_{\text{can}} = \beta^{-1} = \langle dH_\lambda(\mathbf{Y}) | \beta, I, \lambda \rangle. \quad (27)$$

In the literature $dH(\mathbf{Y})$ is therefore often used to “define” the microcanonical temperature (again measured by time-averaging). With no other first integrals present one may take \mathbf{Y} to be the canonical momenta divided by the number of degrees of freedom, i.e., $\mathbf{Y} = \mathbf{P}/d$. In that case, $dH(\mathbf{Y}) = (1/d)dH(\mathbf{P}) = 1/d \sum \dot{q}_i p_i$ is the normalized reduced action (proportional to the kinetic energy, when quadratic in momenta). By time-averaging, we obtain a quantity proportional to the reduced action-integral. This is extremal under variations of the trajectory preserving the energy (Maupertius principle, [9] Section 45). One would thus expect a finite time-average of $dH(\mathbf{P})$ to “probe” a larger neighborhood of the trajectory than time-averages of other observables. This suggests a faster ergodic averaging, whence stronger numerical stability, when calculating time-averages of $dH(\mathbf{P})$ relative to other observables of the same ensemble-variance.

V. AN EXAMPLE

Consider an ensemble of N particles moving on a 3-dimensional torus, $(\mathbb{R}/\mathbf{Z})^3$, under the influence of pair-potentials. Putting things on a torus compactifies the configurational space but otherwise does not affect our results. The Hamiltonian, $H = \sum_{i=0}^N \mathbf{p}_i^2 / 2m_i + \sum_{i,j} U(\mathbf{x}_i - \mathbf{x}_j)$, is translational invariant, hence the total momenta, $\mathbf{p}_{\text{tot}} = \sum \mathbf{p}_i$, provides three first integrals in addition to the Hamiltonian. If one uses the normalized total kinetic energy here as a measure of the temperature one runs into the following paradox: A configuration which minimizes the potential energy but has all particles moving at the same constant velocity, \mathbf{v}_0 , is stationary. With no “apparent activity” one should assign a temperature zero to this configuration, but this is clearly not what the total kinetic energy does.

One solution to this paradox suggests itself: Move to a frame where the average velocity vanishes. Such a procedure is, in fact, a natural consequence of our microcanonical formalism. To get the prefactors right we carry out the details: If $\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_N, 0, \dots, 0)$ denotes the full momentum vector in $\mathbb{R}^{3N} \times (\mathbb{R}/\mathbf{Z})^{3N}$ and $\mathbf{v} = \sum_i \mathbf{p}_i / \sum_i m_i$ is the average velocity, then the reduced momentum vector and kinetic energy is given by $\mathbf{Z} = (\mathbf{p}_1 - m_1 \mathbf{v}, \dots, \mathbf{p}_N - m_N \mathbf{v}, 0, \dots, 0)$ and $K_{\text{red}} = \sum_i (\mathbf{p}_i - m_i \mathbf{v})^2 / 2m_i$, respectively. Straight-forward calculations give $dH(\mathbf{Z}) = 2K_{\text{red}}$ and $d\mathbf{p}_{\text{tot}}(\mathbf{Z}) \equiv 0$. The latter is precisely the requirement of “parallelism” needed for both Eqs. (12) and (22). Furthermore, we have $\nabla \cdot \mathbf{Z} = 3(N-1)$, and $\nabla \cdot (\mathbf{Z}/dH(\mathbf{Z})) = (3(N-1)-2)/(2K_{\text{red}})$. After the appropriate normalization, formulas (17) and (25) yield the surface- and bulk-temperatures

$$\frac{1}{T_\mu} = \left\langle \frac{3(N-1)-2}{2K_{\text{red}}} \right\rangle_\mu \quad \text{and} \quad T_\Omega = \left\langle \frac{2K_{\text{red}}}{3(N-1)} \right\rangle_\mu. \quad (28)$$

In both cases the conservation of three momenta results in a subtraction of three degrees of freedom as compared to the unconstrained case (pleasing on physical grounds). It is interesting also to note that the above formulas involving the reduced kinetic energy are valid irrespective of the total momentum being preserved or not. When the total momentum is only approximately preserved one could use the above construction to define an approximate “instantaneous” and possibly “local” temperature. In Ref. [7], such a problem was considered and the solution suggested was to choose a vector field, \mathbf{Z} , depending on the configurational coordinates only. This automatically ensures “parallelism,” i.e., $d\mathbf{p}_{\text{tot}}(\mathbf{Z}) \equiv 0$, and leads to what is often denoted the “configurational temperature.” On the other hand, for numerical reasons the reduced kinetic energy may be preferable.

The case of a conserved angular momentum (an experiment involving an axial symmetry) follows the same straight-forward procedure and is left to the reader.

[1] H. H. Rugh, Phys. Rev. Lett. **78**, 772 (1997).

[2] H. H. Rugh, J. Phys. A **31**, 7761 (1998).

[3] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin/Cummings, New York, 1978).

[4] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1* (Oxford, Pergamon, 1980).

[5] J. L. Lebowitz, J. K. Percus, and L. Verlet, Phys. Rev. **153**, 250 (1967).

[6] D. E. Evans and G. P. Morriss, *Statistical Mechanics of Non-*

Equilibrium Liquids (Academic Press, New York, 1990).

[7] O. G. Jepps, G. Ayton, and D. J. Evans, Phys. Rev. E **62**, 4757 (2000).

[8] W. K. den Otter, J. Chem. Phys. **112**, 7283 (2000).

[9] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978).

[10] T. Kazuga, Proc. Jpn. Acad. **37**, 366 (1961).

[11] G. Rickayzen and J. G. Powles, J. Chem. Phys. **114**, 4333 (2001).