PHYSICAL REVIEW E **68**, 066126 (2003)

Structure of equations of macrophysics

V. L. Berdichevsky
Mechanical Engineering, Wayne State University, Detroit, Michigan 48202, USA

(Received 21 March 2003; published 31 December 2003)

The first and the second laws of thermodynamics form the constraints on the equations describing macrophysical phenomena. It is argued in this paper that there are some additional universal constraints. These constraints are caused by the Hamiltonian structure of microequations. Previously one feature of micromotion, its reversibility, was used by Onsager to explain the observed reciprocity relations. Hamiltonian structure is richer than reversibility and yields richer consequences. Some of these consequences are a nonlinear version of Onsager's relations, Hamiltonian structure of reversible equations of macrophysics, and quasi-Hamiltonian structure of irreversible equations.

DOI: 10.1103/PhysRevE.68.066126 PACS number(s): 05.70.Ln, 05.20.Dd, 05.40.−a, 05.45.−a

I. INTRODUCTION

It became common wisdom after Clausius [1] that macroworld equations must obey the first and second laws of thermodynamics. Are these two universal laws of Nature the only constraints which every macroscopic theory must obey? An indication that there might be some additional universal features of the macroworld stems from the wide experimental justifications of classical models of reversible processes in mechanics: all these models possess the Hamiltonian structure. The Hamiltonian structure reflects a very special way of reciprocal interactions between various degrees of freedom.

Equations describing the irreversible processes also have a special structure. The first proposition of this kind was made by Thompson for thermoelectric phenomena. Analyzing Thompson's observation and the similar facts gained in other branches of physics, Onsager recognized [2] that any linear macroscopic theory of irreversible processes must obey, in addition to the first and second laws of thermodynamics, some reciprocal relations, and the origin of these relations is the reversibility of micromotion.

To formulate the Onsager result in unambiguous terms, consider an isolated macroscopic system characterized by a finite number of macroscopic "kinematic" parameters $y=(y^1,\dots,y^m)$. The first and second laws of thermodynamics state the following.

(i) There are two additional characteristics of the system, energy E and entropy S.

(ii) There is an equilibrium state in which E, S, and y do not change; the system can stay in this state indefinitely.

(iii) The equilibrium values of E, S, and y are linked by the relation

$$E=E(S,y); \tag{1.1}$$

being solved with respect to S, this relation can also be written in the form

$$S=S(E,y). \tag{1.2}$$

The functions $E(S,y)$ or $S(E,y)$ characterize the equilibrium properties of the system.

(iv) In any process of an isolated system, $y(t)$, energy does not change.

(v) In any process of an isolated system, entropy does not decrease,

$$\frac{dS}{dt}\geqslant 0. \tag{1.3}$$

Note that the work of external force, dA, and the heat supply, dQ, are zero for an isolated system and the equation of the first law of thermodynamics, $dE=dA+dQ$, is satisfied automatically.

To describe the evolution of the system from some state to the equilibrium state, one has to set up the equations for $y(t)$ and $S(t)$ (E does not change). To narrow the subject, we focus, following Onsager, only on the case of the local-in-time dependence of dy/dt and dS/dt on $y(t)$ and $S(t)$. Then the general evolution equations can be written in the form of a system of ordinary differential equations,

$$\dot{y}^{\mu}=G^{\mu}(y,S), \tag{1.4}$$

$$\dot{S}=D(y,S). \tag{1.5}$$

Here and in what follows, Greek indices μ,ν,λ run through values $1,\dots,m$.

The first two laws of thermodynamics do not put constraints on the functions G^{μ} and D, except that the dissipation function D must be positive, and the solutions of Eqs. (1.4) and (1.5), $y(t)$ and $S(t)$, tend to the limit values linked by a known function (1.2) as $t\to\infty$. The question under consideration is as follows: Are there other constraints for functions G^{μ} and D?

In most cases, one can accept the hypothesis of local equilibrium: Eq. (1.2) holds not only at equilibrium but also at each instant in the path to equilibrium. Since the equilibrium properties of the system, expressed by the function (1.2), are supposed to be known, the local equilibrium hypothesis makes entropy a known function of y. One can say that the local equilibrium hypothesis eliminates one entry, the dissipation function D, which can be expressed in terms of G^{μ} and $F_{\mu}=\partial S(E,y)/\partial y^{\mu}$,

1063-651X/2003/68(6)/066126(26)/$20.00

$$D=G^{\mu}F_{\mu}\geqslant 0.$$

Summation over repeated indices is implied. Following the tradition, we call G^{μ} and F_{μ} thermodynamic fluxes and thermodynamic forces, respectively.

Further constraints depend on the type of y variables. First, let y variables be ‘‘coordinate-type’’ variables (this term will be explained later in Sec. II). Onsager considered the case of linear equations when G^{μ} and F_{μ} are linear functions of y.

Onsager showed that, due to reversibility of micromotion, G^{μ} must be expressed in terms of F_{μ} by the relations

$$G^{\mu}=D^{\mu\nu}F_{\nu}, \tag{1.6}$$

where the dissipation coefficients $D^{\mu\nu}$ are symmetric constants,

$$D^{\mu\nu}=D^{\nu\mu}. \tag{1.7}$$

Onsager’s relation (1.7) is sometimes called the third law of thermodynamics. Perhaps it is reasonable to use this term for the issues concerning all constraints on macroequations that are in addition to those of the first and second laws.

There were numerous attempts to generalize Onsager’s relations to nonlinear phenomena when G^{μ} are some nonlinear functions of F_{μ}. The most widely used is the proposition that there exists a scalar function $\Phi(F_{\mu})$ such that

$$G^{\mu}=\frac{\partial\Phi(F_{\mu})}{\partial F_{\mu}}. \tag{1.8}$$

If Φ is a quadratic function, $\Phi=\frac{1}{2}D^{\mu\nu}F_{\mu}F_{\nu}$, Eq. (1.8) transforms to Eqs. (1.6) and (1.7). Most models of continuum mechanics are based on potential relations (1.8). For example, such are the models of plasticity theory with a homogeneous function Φ of the first order. Potential relations (1.8) motivated various variational principles which, in turn, are being used often as a basis for mathematical modeling.

Relations (1.8) are very convenient for studying such issues as thermodynamic stability or correctness of the corresponding mathematical problems. The question is, however, whether Eq. (1.8) is ‘‘a universal law of Nature’’ or just a matter of mathematical convenience.

Some doubts as to the validity of Eq. (1.8) in general arose from the Onsager observation that Eq. (1.7) holds true only in the absence of magnetic field. If magnetic field $\mathbf{m}$ is presented, the coefficients $D^{\mu\nu}$ are not necessarily symmetric, and Eq. (1.7) must be replaced by the relation

$$D^{\mu\nu}(\mathbf{m})=D^{\nu\mu}(-\mathbf{m}). \tag{1.9}$$

Obviously, if $D^{\mu\nu}$ depend on $\mathbf{m}$, Eq. (1.8) does not hold even in the linear case.

The consequences of reversibility of micromotion for nonlinear phenomena have been studied by Stratanovich [6]. He obtained a series of constraints in a quite general setting and did not find a confirmation of Eq. (1.8). He noted that all attempts to prove Eq. (1.8) failed. Moreover, there are examples when Eq. (1.8) does not hold.

Reviewing the situation with nonlinear generalizations of Onsager’s relations, Mazur wrote in 1996 [7] ‘‘ . . . Another issue of interest has been whether a generalization existed of the Onsager reciprocity which would hold for *nonlinear* laws describing irreversible processes. At the first IUPAP International Conference on statistical mechanics, held in 1949 in Florence, Casimir presented a paper ‘On some aspects of Onsager’ theory . . . ’ in which he remarks ‘ . . . in its present form Onsager’s theory applies only to equations of [linear] type.’ Onsager, who was present, offered the following comment, ‘Linear relations between rates of flow and driving forces are assumed in my derivation of reciprocal relations. The possibilities of useful generalizations have not been fully explored; none have been found so far.’ It would seem that Onsager’s comment has retained its actuality to this day.’’

It is argued in this paper that a nonlinear version of Onsager’s relations has the form

$$G^{\mu}=D^{\mu\nu}(F_{\lambda})F_{\nu} \tag{1.10}$$

with the symmetric (in the absence of magnetic field) dissipation coefficients $D^{\mu\nu}$,

$$D^{\mu\nu}(F_{\lambda})=D^{\nu\mu}(F_{\lambda}). \tag{1.11}$$

The function $D^{\mu\nu}F_{\mu}F_{\nu}$ must be non-negative,

$$D^{\mu\nu}(F_{\lambda})F_{\mu}F_{\nu}\geqslant 0 \quad \text{for any } F_{\mu},$$

in order to comply with the second law of thermodynamics.

If $D^{\mu\nu}$ depend on magnetic field, Eq. (1.11) is replaced by a more complex relation discussed in Sec. V. The dissipation coefficients may depend on the parameter E.

It will be seen in Sec. II that Eqs. (1.10) and (1.11) can be obtained from a little known paper by Kolmogorov [8]. Equations (1.10) and (1.11) are, to some extent, misleading: in fact, they do not put constraints on the functional dependence of G^{μ} on F_{μ} beyond the point $F_{\mu}=0$ (see Sec. IX). Nevertheless, there are serious reasons to write this dependence in the form (1.10) and (1.11); they are discussed in Sec. IX.

Nonlinear Onsager’s relations (1.10) and (1.11) are a byproduct of consideration of the following questions: Micromotion possesses a much richer peculiarity than reversibility; it is governed by equations with the Hamiltonian structure. The Hamiltonian structure of microequations should yield richer consequences than just reversibility. What are the constraints on the macroequations caused by the Hamiltonian structure of microequations? We show that no further constraints appear for the coordinate-type y variables. However, if y variables are canonical, i.e., if they are some of the coordinates and momenta of the underlying Hamiltonian system, then the equations take a special quasi-Hamiltonian form: there exist an effective Hamiltonian $H_{\mathrm{eff}}(S,y)$ and dissipative coefficients $D^{\mu\nu}(S,y)$ such that

$$\frac{dy^{\mu}}{dt}=\omega^{\mu\nu}\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial y^{\nu}}-\frac{D^{\mu\nu}(S,y)}{T(S,y)}\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial y^{\nu}}, \tag{1.12}$$

$$T=\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial S}, \tag{1.13}$$

$$\frac{dS}{dt}=\frac{D^{\mu\nu}(S,y)}{T^2(S,y)}\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial y^{\mu}}\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial y^{\nu}}. \tag{1.14}$$

Here $\omega^{\mu\nu}$ is a constant antisymmetric tensor defining the Hamiltonian structure [for exact definition, see Eq. (2.18)], and T is temperature.

The origin of the link between microequations and macroequations was first revealed by Boltzmann: the time scale of macrovariables is much bigger than that of microvariables. One can say that macroequations can be obtained by the elimination of fast variables from microequations. Thermodynamics is a theory of slow variables for microdynamics. Boltzmann's observation enabled Hertz [9] to derive the laws of equilibrium thermodynamics from Hamiltonian mechanics (a modern exposition of Hertz's results, including the case of low-dimensional systems, can be found in Chap. 1 of [10]). Hertz's paper made it clear that the reason why the laws of equilibrium thermodynamics are true is the Hamiltonian structure of microequations.

Note that the term "Hamiltonian structure" is unambiguous if the phase space of the dynamical model of microworld has been fixed; otherwise any system of ordinary differential equations can be put in Hamiltonian form (see Appendix A). The choice of the phase space and the Hamiltonian is determined by physics of the phenomena under consideration.

One may wonder whether the Hamiltonian structure of microdynamics is necessary to observe the classical equilibrium thermodynamics on the macrolevel. This issue is discussed in Appendix A under some assumptions which do not seem physically constraining. It turns out that microdynamic equations may have slightly more general structure than the standard Hamiltonian equations and still yield the laws of equilibrium thermodynamics.

Hertz considered the classical (nonquantum) microworld, as we will do throughout this paper. The quantum nature of the microworld affects the macroscopic laws, and a consistent theory should start from the laws of quantum mechanics. We assume, however, that in the transition micro→macro there is a level of description at which classical Hamiltonian mechanics provides an adequate picture of microdynamics. Such an assumption restricts the range of admissible values of macroparameters.

There was a long and difficult way to recognize that the laws of irreversible nonequilibrium thermodynamics do not contradict the underlying reversible Hamiltonian dynamics (see, e.g., the review [11] and pp. 77–79 of [10]). A clear understanding of how the laws of nonequilibrium thermodynamics may be obtained from Hamiltonian mechanics was achieved by Kubo [4,5] (see also Zwanzig [12]). Kubo considered a special case of open systems, the driven systems: a parameter of a Hamiltonian system, say y, is changed slowly in a given way, $y(t)$, and one wonders how the force causing this change depends on $y(t)$. In the linear case, Kubo found this dependence explicitly from an asymptotic analysis of Liouville's equation and established the relation between the dissipative coefficients and a correlation of thermodynamic fluxes in equilibrium.

Another line of thought stemmed from an observation that Onsager's reasoning can be put on mathematical grounds if one assumes that macrodynamics is a Markov process governed by linear ordinary stochastic differential equations [3,13]. Time reversibility yields then Onsager's reciprocal relations. The criterion of reversibility for Markov's continuous processes in the nonlinear case was obtained by Kolmogorov [8] and, in a slightly generalized form, by Yaglom [14]. Stratanovich [6] studied the consequences of reversibility for the general Markov noncontinuous and non-Markov stochastic processes.

An attempt to take into account the Hamiltonian structure of microdynamics was made in [15] for the case of canonical slow variables. It was suggested that, at equilibrium, the probability flux of Markov's process coincides with that of the underlying ergodic Hamiltonian system. It turns out that the probability flux hypothesis selects a quasi-Hamiltonian form of equations of the slow evolution (1.12) and (1.14). Originally, these equations were obtained in [15] in a different equivalent form (2.30).

The aims of this paper are as follows: to derive a nonlinear version of Onsager's relation for coordinate-type y variables (1.10) and (1.11); to justify the probability flux hypothesis for canonical variables by asymptotic analysis of Liouville's equation; to obtain the corresponding consequences for constitutive equations of solids and dynamical equations for defects in a crystal lattice; and to extend the results for interacting systems with one noncanonical slow variable—energy—and to conduct a limited transition to continuum theory to obtain the equations of nonlinear heat conductivity.

Restraining the consideration by isolated systems does not affect the results on the structure of the governing equations for local-in-space systems: for such systems, isolation means just a special choice of the boundary conditions.

Note that the limit transition to continuum theory is a nontrivial issue. The general form of continuum equations compatible with the Hamiltonian structure of microequations has yet to be established. Some observation on the structure of equations used in continuum mechanics can be found in [26,27]. It remains to be seen whether the Hamiltonian structure of microequations imposes additional constraints.

The above-discussed special structure of macroequations pertains to the case when the system has two characteristic well-separated time scales, namely fast time and slow time. Elimination of the fast variables yields classical thermodynamics, which we will call also "primary thermodynamics." There are situations with three well-separated time scales. In such cases, the system is characterized by three types of variables: fastest, intermediate-fast, and slow. Elimination of the fastest variables yields the equations for the intermediate-fast and slow variables. The corresponding laws are the laws of primary thermodynamics. Equations of primary thermodynamics possess the special structure discussed above. Elimination of the intermediate-fast variables from the equations of primary thermodynamics gives the equations for the slow variables. The theory of slow variables can be called "secondary thermodynamics" or "thermodynamics of attractors" because the existence of attractors is a characteristic feature of the equations of primary thermodynamics. An important example is a turbulent motion of fluids. The fastest time is a characteristic time of molecule collisions. Primary thermodynamics brings the Navier-Stokes equations. These equations have two characteristic times, namely the characteristic time of the change of velocity at a space point and the characteristic time of the change of average flow parameters (such as Reynolds' stresses). Turbulence theory, which studies the averaged characteristics of the flow, is, in fact, thermodynamics of attractors of Navier-Stokes equations. Thermodynamics of attractors is at an embryonic stage. A brief account of it is given in Sec. X.

The paper composition is as follows: In the next section, we review the consequences of reversibility of Markov's processes, formulate a nonlinear version of Onsager's relations, briefly discuss the probability flux hypothesis, and outline the derivation of the quasi-Hamiltonian structure from the probability flux hypothesis. In Secs. III–V, the validity of the probability flux hypothesis is justified by the asymptotic analysis of Liouville's equation. The quasi-Hamiltonian structure is used to establish a general form of constitutive equations for solids in Sec. VI. Nonlinear heat conductivity is considered in Sec. VII. The quasi-Hamiltonian structure for dynamics of defects in solids is obtained in Sec. VIII. The features of the nonlinear version of Onsager's relations (1.10) and (1.11) are discussed in Sec. IX. Secondary thermodynamics is reviewed in Sec. X. This is followed by a discussion of the term "Hamiltonian structure" and a description of microdynamic equations which are compatible with equilibrium thermodynamics in Appendix A, and some auxiliary estimates in Appendix B.

II. EVOLUTION TO EQUILIBRIUM AS A STOCHASTIC PROCESS

A. Markov's stochastic process

Macroparameters are always fluctuating due to interaction with fast degrees of freedom, and thus they can be considered as some random variables. Denote macroparameters by $y^\mu, \mu=1,\dots,m$ and their probability density at instant t by $f(t,y)$. It will be derived in the subsequent sections that, under some assumptions, the function $f(t,y)$ obeys the Fokker-Planck equation

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial J^\mu}{\partial y^\mu}=0,$$

$$J^\mu=V^\mu(y)f-D^{\mu\nu}(y)\frac{\partial f}{\partial y^\nu}, \tag{2.1}$$

where J^μ is the probability flux and $D^{\mu\nu}$ are the diffusion coefficients. The diffusion coefficients are symmetric, $D^{\mu\nu}=D^{\nu\mu}$, and positive, $D^{\mu\nu}x_\mu x_\nu\geqslant 0$ for any x_ν. The probability density of y under condition that y takes the value y_0 at instant t_0, $f(t,y|t_0,y_0)$, also satisfies Eq. (2.1). This indicates that $y(t)$ may be considered as a stationary Markov stochastic process obeying some ordinary stochastic differential equations,

$$\frac{dy^\mu}{dt}=F^\mu(y)+\sigma^\mu_a(y)\dot{w}^a, \tag{2.2}$$

where $\dot{w}^a$ $(a=1,\dots,s,\ s\geqslant m)$ are some independent white noises. There is a link between F^μ,σ^μ_a and $V^\mu,D^{\mu\nu}$. The link depends on the sense of time derivative in Eq. (2.2). If Eq. (2.2) is understood in Stratanovich's sense, then

$$F^\mu=V^\mu+\frac{1}{2}\sigma^\mu_a\frac{\partial\sigma^\nu{}_a}{\partial y^\nu},\quad D^{\mu\nu}=\frac{1}{2}\sigma^\mu_a\sigma^\nu_a. \tag{2.3}$$

The diffusion coefficients $D^{\mu\nu}$ are very small [the scale of smallness is established by Eq. (2.32) below]. Therefore, according to Eq. (2.1), $f(t,y)$ in the first approximation is the δ function, $f(t,y)=\delta(y-\overline{y(t)})$, where $\overline{y(t)}$ is a solution of the ordinary differential equations

$$\frac{d\overline{y^\mu}}{dt}=V^\mu(\overline{y}). \tag{2.4}$$

Equations (2.4) hold true also for the average value of y^μ, $\int y^\mu f(t,y)dy$, in the first approximation. This can be obtained by averaging Eq. (2.2). In addition, one can find from the averaged equation (2.2) the corrections to the right-hand side of Eq. (2.4). They turn out to be of the order $D^{\nu\lambda}\partial^2V^\mu/\partial y^\nu\partial y^\lambda$.

Equations (2.4) for the average values of y are exact in the linear case when V^μ are linear functions of y and $D^{\mu\nu}$ are constants. Markov's character of macrovariables has been observed in many physical systems [16].

B. Reversibility of Markov's process

Denote by $f(t_1,y_1;t_2,y_2)$ the two-point probability density function of the process $y(t)$: probability of the event that $y(t_1)$ belongs to a small vicinity of the point y_1, while $y(t_2)$, in a small vicinity of the point y_2, is equal to $f(t_1,y_1;t_2,y_2)\triangle V_1\triangle V_2$, $\triangle V_1$ and $\triangle V_2$ being the volumes of the vicinities. It is seen from the definition that, for any process, $f(t_1,y_1;t_2,y_2)$ is a symmetric function,

$$f(t_1,y_1;t_2,y_2)=f(t_2,y_2;t_1,y_1). \tag{2.5}$$

A stochastic process is called reversible if the two-point probability density possesses the following property: for any t_1,t_2,

$$f(t_1,y_1;t_2,y_2)=f(t_2,y_1;t_1,y_2). \tag{2.6}$$

In accordance with Eq. (2.5), the definition of reversibility can be also written as

$$f(t_1,y_1;t_2,y_2)=f(t_1,y_2;t_2,y_1). \tag{2.7}$$

The two-point probability density $f(t_1,y_1;t_2,y_2)$ can be expressed in terms of the conditional probability $f(t_1,y_1|t_2,y_2)$ by the relation

$$f(t_1,y_1;t_2,y_2)=f(t_1,y_1|t_2,y_2)f(t_2,y_2).$$

In terms of the conditional probability, the reversibility condition (2.7) takes the form

$$f(t_1,y_1|t_2,y_2)f(t_2,y_2)=f(t_1,y_2|t_2,y_1)f(t_2,y_1). \tag{2.8}$$

Assume that there is an equilibrium state with an equilibrium distribution $f_\infty(y)$, and the stochastic process is stationary, thus the conditional probability $f(t_1,y_1|t_2,y_2)$ depends only on the time shift $t=t_1-t_2$: $f(t_1,y_1|t_2,y_2)=f(t,y_1|y_2)$. Then reversibility of the stochastic process *at equilibrium*, according to Eq. (2.8), means that the conditional probability must satisfy the equation

$$f(t,y_1|y_2)f_\infty(y_2)=f(t,y_2|y_1)f_\infty(y_1). \tag{2.9}$$

Equation (2.9) puts some strong constraints on the admissible values of drifts, V^μ, and diffusion coefficients, $D^{\mu\nu}$. These constraints were first found by Kolmogorov in 1937 [8]: in order for Eq. (2.9) to be true, it is necessary and sufficient that probability flux J^μ vanishes on the equilibrium distribution,

$$J^\mu=V^\mu f_\infty-D^{\mu\nu}\frac{\partial f_\infty}{\partial y^\nu}=0, \tag{2.10}$$

or, equivalently,

$$V^\mu=D^{\mu\nu}\frac{\partial \ln f_\infty}{\partial y^\nu}. \tag{2.11}$$

Kolmogorov obtained Eq. (2.11) in a slightly different form linked to the technicalities of his proof.

Reversibility condition (2.6) is written for the case when y variables are coordinate-type variables, i.e., they do not change sign for a reversed process. If some of the y variables are velocity-type variables, i.e., they change sign for a reversed process, then Eq. (2.6) must be changed accordingly to incorporate the sign change. The constraints imposed by such a modified reversibility condition on drifts and diffusion coefficients were obtained by Yaglom [14]. Later, Kolmogorov-Yaglom relations were rediscovered in a number of papers (see [17]). Condition (2.6) is sometimes called the detailed balance.

C. Einstein's formula

Let y be slow variables of an ergodic Hamiltonian system. The system is isolated and has energy E. Thermodynamic properties of the system are characterized by entropy $S(E,y)$. Probability distribution of y variables at equilibrium, $f_\infty(y)$, is given by Einstein's formula,

$$f_\infty(y)=ce^{S(E,y)}, \tag{2.12}$$

where c is a normalizing constant.

Einstein's formula has an asymptotic character: it is valid in the limit of an infinite number of fast degrees of freedom. Regarding the derivation of Einstein's formula "from mechanics" and its generalization for a finite number of fast degrees of freedom, see [18,10].

D. Nonlinear Onsager relations

Combining Kolmogorov's criterion of reversibility with Einstein's formula (2.12), we arrive at the expression for the drift in terms of diffusion coefficients and entropy,

$$V^\mu=D^{\mu\nu}(y)\frac{\partial S}{\partial y^\nu}. \tag{2.13}$$

Thus, the evolution equations for y take the form

$$\frac{dy^\mu}{dt}=D^{\mu\nu}(y)\frac{\partial S}{\partial y^\nu}. \tag{2.14}$$

Here we dropped the bar from y since the actual values of y deviate only slightly from the average values.

Entropy is proportional to a large number, the number of fast degrees of freedom. Therefore, the right-hand side of Eq. (2.14) may be finite even though $D^{\mu\nu}$ are small.

Entropy grows according to the equation

$$\frac{dS(E,y)}{dt}=D^{\mu\nu}(y)\frac{\partial S}{\partial y^\mu}\frac{\partial S}{\partial y^\nu}. \tag{2.15}$$

If entropy has the only point of maximum, $\hat{y}$, then the system goes to this point.

Thermodynamic fluxes and thermodynamic forces in the case under consideration are the drift components V^μ and the derivatives of entropy, respectively. In the linear approximation, when $D^{\mu\nu}$ are constants and S is a quadratic function, Eqs. (2.13) coincide with Onsager's statement. Thus, it is sensible to consider Eq. (2.13) as a nonlinear version of Onsager's relations. The Fokker-Planck equation (2.1) motivates the term "diffusion coefficients" for $D^{\mu\nu}$ while the entropy equation (2.15) justifies the term "dissipation coefficients." We use both terms for $D^{\mu\nu}$ depending on the context. This duality is the essence of fluctuation-dissipation theorem.

E. Ergodic Hamiltonian systems and the probability flux hypothesis

Now let the dynamics of the system be governed by Hamiltonian equations, and the slow variables y be some variables of these equations. In this case, we call the y variables canonical variables. Denote the fast variables of the system by $x=(x^1,\dots,x^{2n})$ and the Hamiltonian by $H(x,y)$. We identify the first n coordinates of point x with generalized coordinates of the system, $q=(q^1,\dots,q^n)$, and the last n coordinates with generalized momenta $p=(p_1,\dots,p_n)$. Similarly, $y=(Q^1,\dots,Q^M,\ P_1\cdots P_M)$, $2M=m$. The Hamiltonian equations are

$$\frac{dq}{dt}=\frac{\partial H}{\partial p},\quad \frac{dp}{dt}=-\frac{\partial H}{\partial q},$$

$$\frac{dQ}{dt}=\frac{\partial H}{\partial P},\quad \frac{dP}{dt}=-\frac{\partial H}{\partial Q}.$$

In index notation they take a shorter form,

$$\frac{dx^i}{dt}=\omega^{ij}\frac{\partial H}{\partial x^j}, \tag{2.16}$$

$$\frac{dy^\mu}{dt}=\omega^{\mu\nu}\frac{\partial H}{\partial y^\nu}, \tag{2.17}$$

where Latin indices $i,j,\dots$, run through values $1,\dots,2n$, and $\omega^{ij},\omega^{\mu\nu}$ are constant antisymmetric tensors:

$$\omega^{ij}=1\ \text{ if }\ i\leqslant n,\ j=i+n;\quad \omega^{ij}=-1\ \text{ if }\ j\leqslant n,\ i=j+n;\quad \omega^{ij}=0\ \text{ otherwise}; \tag{2.18}$$
$$\omega^{\mu\nu}=1\ \text{ if }\ \mu\leqslant M,\ \nu=\mu+M;\quad \omega^{\mu\nu}=-1\ \text{ if }\ \nu\leqslant M,\ \mu=\nu+M;\quad \omega^{\mu\nu}=0\ \text{ otherwise.}$$

We assume that in the (x,y)-phase space the surfaces $H(x,y)=E$ are compact and bound the regions with finite volume $\Gamma(E)$,

$$\Gamma(E)=\int\theta(E-H(x,y))dxdy. \tag{2.19}$$

$\theta(E)$ is the step function: $\theta(E)=1$ for $E>0$ and $\theta(E)=0$ for $E\leqslant 0$. In the x-phase space, for each fixed y, the surfaces $H(x,y)=E$ are compact and bound the regions with a finite volume $\Gamma(E,y)$,

$$\Gamma(E,y)=\int\theta(E-H(x,y))dx. \tag{2.20}$$

Motion of the system (2.16) and (2.17) is assumed to be ergodic on the energy surfaces $H(x,y)=E$ in (x,y)-phase space. For each fixed y, motion of the system (2.16) is also assumed ergodic on the energy surfaces $H(x,y)=E$ in x-phase space. Denote entropy of the system (2.16) by $S(E,y)$. One can show [9] that

$$S(E,y)=\ln\Gamma(E,y).$$

Besides, Einstein's formula (2.12) holds true (in the limit of large n) [10,18].

In what follows, it is convenient to use instead of Eq. (2.12) the exact formula [10,18]

$$f(\infty,y)=\frac{1}{\Gamma_E(E)}\frac{\partial}{\partial E}e^{S(E,y)}$$
$$=\frac{1}{T(E,y)\Gamma_E(E)}e^{S(E,y)}, \tag{2.21}$$

where $T=\partial S/\partial E$ is the absolute temperature and $\Gamma_E(E)\equiv d\Gamma(E)/dE$.

The probability density $f(t,x,y)$ of (x,y) variables obeys Liouville's equation,

$$\frac{\partial f}{\partial t}+\omega^{ij}\frac{\partial H}{\partial x^j}\frac{\partial f}{\partial x^i}+\omega^{\mu\nu}\frac{\partial H}{\partial y^\nu}\frac{\partial f}{\partial y^\mu}=0. \tag{2.22}$$

Integrating this equation over x, one obtains the equation for the probability density of y variables, $f(t,y)$,

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial J^\mu}{\partial y^\mu}=0,\quad f(t,y)=\int f(t,x,y)dx,$$

$$J^\mu=\int\omega^{\mu\nu}\frac{\partial H}{\partial y^\nu}f(t,x,y)dx. \tag{2.23}$$

The probability density at equilibrium, $f(\infty,x,y)$, is equal to const$\times\,\delta(E-H(x,y))$. The probability flux at equilibrium, can be found explicitly [15],

$$J^\mu=-\frac{1}{\Gamma_E(E)}\omega^{\mu\nu}\frac{\partial e^{S(E,y)}}{\partial y^\nu}. \tag{2.24}$$

Let the evolution of y variables to equilibrium be modeled as a Markov stochastic continuous process. Then the probability density function of y variables, $f(t,y)$, obeys the Fokker-Planck equation

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial}{\partial y^\mu}J^\mu_{(M)}=0,$$

$$J^\mu_{(M)}=V^\mu f(t,y)-D^{\mu\nu}\frac{\partial f(t,y)}{\partial y^\nu}. \tag{2.25}$$

Here we add index (M) to the probability flux of Markov's process to distinguish it from the probability flux in the integrated Liouville equation (2.23). Since the original Hamiltonian system is autonomous, the stochastic process must be stationary. Therefore, V^μ and $D^{\mu\nu}$ are the functions of y only and do not depend on time.

Function $f(t,y)$ is the same in the integrated Liouville equation (2.23) and the Fokker-Planck equation (2.25). Thus, divergences of the probability fluxes J^μ and $J^\mu_{(M)}$ must coincide,

$$\frac{\partial J^\mu_{(M)}}{\partial y^\mu} = \frac{\partial J^\mu}{\partial y^\mu}.$$

We accept a stronger condition (the probability flux hypothesis): *at equilibrium*,

$$J^\mu_{(M)} = J^\mu. \tag{2.26}$$

From Eq. (2.26) and the expressions for probability fluxes of Markov's process (2.25) and Hamiltonian system (2.24), we have

$$V^\mu \frac{e^S}{\Gamma_E T} - D^{\mu\nu} \frac{\partial}{\partial y^\nu} \frac{e^S}{\Gamma_E T} = -\omega^{\mu\nu} \frac{1}{\Gamma_E(E)} \frac{\partial}{\partial y^\nu} e^S. \tag{2.27}$$

Taking into account that

$$\frac{\partial}{\partial y^\mu} \frac{e^S}{T} = \frac{e^S}{T}\left(\frac{\partial S}{\partial y^\mu} - \frac{1}{T}\frac{\partial T}{\partial y^\mu}\right) \tag{2.28}$$

and that the last term in Eq. (2.28) can be neglected compared with the preceding one (since entropy S is proportional to a large number of fast degrees of freedom, $2n$, while temperature T does not depend on n), one finds from Eqs. (2.27) and (2.28) the drift

$$V^\mu = -\omega^{\mu\nu} T \frac{\partial S}{\partial y^\nu} + D^{\mu\nu} \frac{\partial S}{\partial y^\nu}. \tag{2.29}$$

The corresponding equations describing the evolution of the system to equilibrium are

$$\frac{dy^\mu}{dt} = -\omega^{\mu\nu} T \frac{\partial S(E,y)}{\partial y^\nu} + D^{\mu\nu}(y) \frac{\partial S(E,y)}{\partial y^\nu}. \tag{2.30}$$

Equations (2.30) augmented by the definition of temperature

$$\frac{1}{T} = \frac{\partial S(E,y)}{\partial E} \tag{2.31}$$

form a closed system of equations.

The system of Equations (2.30) has quite a special form. It involves two "entries": equilibrium entropy $S(E,y)$, and a symmetric tensor of diffusion coefficients, $D^{\mu\nu}(y)$. To specify the system, one has to prescribe functions $S(E,y)$ and $D^{\mu\nu}(y)$. The symmetry of tensor $D^{\mu\nu}$ is caused by Markov's nature of the process. Onsager's relations are hidden in the assumption on Markov's character of the process and the probability flux hypothesis.

Note that temperature T is small compared to the total energy of the system E: $T \sim E/n$, the total energy being considered finite. Entropy is proportional to n. Therefore, for large n, the first term on the right-hand side of Eq. (2.30) is finite. The second term is in the order of the first one if $D^{\mu\nu}$ are small,

$$D^{\mu\nu} \sim T. \tag{2.32}$$

Various components of $D^{\mu\nu}$ have different dimensions, and, in a particular problem, (2.32) is to be rectified depending on the physical meaning of coordinates y.

F. Effective Hamiltonian

Let the Hamiltonian of the system have the form

$$H(x,y) = H(y) + H_0(x,y).$$

Denote by S_0 the entropy of the "fast" Hamiltonian system with Hamiltonian $H_0(x,y)$ and fixed values of y. Motion occurs on the energy surfaces $H_0(x,y) = U = \text{const}$. Entropy S_0 is a function of U and y, $S_0 = S_0(U,y)$. It is easy to find that entropy of the system with Hamiltonian $H(x,y)$, $S(E,y)$, is (see, e.g. [10,18])

$$S(E,y) = S_0(E - H_0(y), y), \tag{2.33}$$

E being the total energy of the system.

For a fixed value S of function $S_0(U,y)$, U is a function of y and S determined from the equation

$$S_0(U,y) = S. \tag{2.34}$$

Denote this function by $U(S,y)$. It obeys the identity

$$S_0(U(S,y),y) = S \quad \text{for any } S \text{ and } y. \tag{2.35}$$

Differentiating Eq. (2.35) with respect to S and y, we have

$$\frac{\partial S_0(U,y)}{\partial U} \frac{\partial U(S,y)}{\partial S} = 1,$$

$$\frac{\partial S_0(U,y)}{\partial U} \frac{\partial U(S,y)}{\partial y^\mu} + \frac{\partial S_0(U,y)}{\partial y^\mu} = 0. \tag{2.36}$$

If one sets U in Eq. (2.34) to be equal to $E - H_0(y)$, then S coincides with the total entropy. Since, from Eq. (2.33),

$$\left.\frac{\partial S_0}{\partial U}\right|_{U=E-H_0(y)} = \frac{\partial S}{\partial E} = \frac{1}{T},$$

Eqs. (2.36) take the form

$$\frac{\partial U(S,y)}{\partial S} = T,$$

$$\frac{\partial U(S,y)}{\partial y^\mu} = -T \left.\frac{\partial S_0(U,y)}{\partial y^\mu}\right|_{U=E-H_0(y)}. \tag{2.37}$$

We define the effective Hamiltonian $H_{\text{eff}}(S,y)$ by the formula

$$H_{\text{eff}}(S,y) = H_0(y) + U(S,y). \tag{2.38}$$

For derivative of the effective Hamiltonian with respect to y at a fixed value of entropy, we have from Eqs. (2.33), (2.37), and (2.38)

$$\begin{aligned}\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial y^{\mu}} &= \frac{\partial H_0(y)}{\partial y^{\mu}} + \frac{\partial U(S,y)}{\partial y^{\mu}} \\ &= \frac{\partial H_0(y)}{\partial y^{\mu}} - T\frac{\partial S_0(U,y)}{\partial y^{\mu}}\bigg|_{U=E-H_0(y)} \\ &= -T\frac{\partial S_0(E-H_0(y),y)}{\partial y^{\mu}} \\ &= -T\frac{\partial S(E,y)}{\partial y^{\mu}}.\end{aligned}$$

Besides, in accordance with Eq. (2.37),

$$\frac{\partial H_{\mathrm{eff}}(S,y)}{\partial S} = T.$$

In terms of derivatives of the effective Hamiltonian, the evolution equations (2.30) take the form (1.12). For a given effective Hamiltonian and given diffusion coefficients, Eqs. (1.12) are not closed: they involve an unknown function $S(t)$. The additional equation for $S(t)$ can be obtained either from Eq. (2.30) or from the condition that the total energy of the system, $H_{\mathrm{eff}}(S,y)$, is conserved. This yields Eq. (1.14). Equations (1.12)–(1.14) form a closed system of equations.

We call the form of Eqs. (1.12)–(1.14), or the equivalent Eqs. (2.30), quasi-Hamiltonian because these equations become Hamiltonian if the diffusion coefficients are zero.

Deviations from the standard Hamiltonian form are caused by the dependence of the effective Hamiltonian on entropy, the determination of entropy from the additional equation (1.14), and by the dissipative term $T^{-1}D^{\mu\nu}\partial H_{\mathrm{eff}}/\partial y^{\nu}$ in Eq. (1.12).

Our next goal is to justify Markov's character of slow variables, the probability flux hypothesis (2.26), and the quasi-Hamiltonian form of equations for slow variables by asymptotic analysis of Liouville's equations.

III. ASYMPTOTIC ANALYSIS OF LIOUVILLE'S EQUATION IN THE CASE OF CANONICAL SLOW VARIABLES

Consider the Hamiltonian system (2.16) and (2.17). Variables y are assumed to be slow compared to variables x. We do not introduce a small time parameter explicitly since we focus only on the first approximation and do not construct the full asymptotic expansion. We set $\partial H/\partial x^i$ in the order of unity and assume that $\partial H/\partial y^{\mu}$ are much smaller than $\partial H/\partial x^i$.

The system of ordinary differential equations (2.16) and (2.17) is equivalent to the partial differential equation of the first order, Liouville, equation,

$$\frac{\partial f(t,x,y)}{\partial t} + \omega^{ij}\frac{\partial H(x,y)}{\partial x^j}\frac{\partial f(t,x,y)}{\partial x^i} + \omega^{\mu\nu}\frac{\partial H(x,y)}{\partial y^{\nu}}\frac{\partial f(t,x,y)}{\partial y^{\mu}} = 0. \tag{3.1}$$

The equilibrium state of the ergodic Hamiltonian system corresponds to the steady solution

$$f_{\infty}(x,y) = \frac{1}{\Gamma_E(E)}\delta(E-H(x,y)). \tag{3.2}$$

We expect that function $f(t,x,y)$ tends to function (3.2) as $t\to\infty$ if initially y have some prescribed value $y_0 = (y_0^1, \ldots, y_0^m)$ while x have equilibrium distribution:

$$f(0,x,y) = \frac{1}{\Gamma_E(E,y_0)}\delta(E-H(x,y_0))\delta(y-y_0), \tag{3.3}$$

where $\delta(y-y_0)\equiv\delta(y^1-y_0^1)\cdots\delta(y^m-y_0^m)$ and $\Gamma_E(E,y) \equiv \partial\Gamma(E,y)/\partial E$. Obviously, the normalization condition for probability density $f(t,x,y)$ is satisfied at the initial instant,

$$\begin{aligned}\int f(0,x,y)dxdy &= \frac{1}{\Gamma_E(E,y_0)}\int \delta(E-H(x,y_0))dx \\ &= \frac{1}{\Gamma_E(E,y_0)}\frac{\partial}{\partial E}\int \theta(E-H(x,y_0))dx \\ &= \frac{1}{\Gamma_E(E,y_0)}\frac{\partial}{\partial E}\Gamma(E,y_0) \\ &= 1.\end{aligned}$$

We use the abbreviations $dx = dx^1\cdots dx^n$, $dy = dy^1\cdots dy^m$. The normalization condition is maintained in the course of motion since for any solution of Liouville's equation, as follows from Eq. (3.1),

$$\frac{d}{dt}\int f(t,x,y)dxdy = 0.$$

Evolution of the initial distribution (3.3) to the equilibrium distribution (3.2) should be understood in a weak sense which mimics a coarse graining. This is a subject for separate consideration. We describe here only the formal procedure. Note, however, that the convergence of $f(t,x,y)$ to $f_{\infty}(x,y)$ may take place only for systems with some mixing properties. This is in contrast to the relations of equilibrium thermodynamics, which require only ergodicity.

So, our task is to find the asymptotics of the solution of Liouville's equation (3.1) with initial conditions (3.3) assuming that the last term in Eq. (3.1) is small. In fact, we aim to find the governing equation for the probability density function of y variables,

$$f(t,y) = \int f(t,x,y)dx. \tag{3.4}$$

A general form of such an equation follows from integration of Liouville's equation over x,

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial J^{\mu}}{\partial y^{\mu}}=0, \tag{3.5}$$

where the probability flux is

$$J^{\mu}=\omega^{\mu\nu}\int\frac{\partial H}{\partial y^{\nu}}f(t,x,y)dx. \tag{3.6}$$

We use asymptotic reasonings to link J^{μ} and $f(t,y)$.

Due to ergodicity, the Hamiltonian system has the only integral $H(x,y)=E$. Thus, it is worthwhile to make a change of unknown function in Liouville's equation,

$$f(t,x,y)=\delta(E-H(x,y))\Phi(t,x,y). \tag{3.7}$$

At the initial instant, $\Phi(t,x,y)$ obeys the condition

$$\Phi(0,x,y)=\frac{1}{\Gamma_E(E,y_0)}\delta(y-y_0). \tag{3.8}$$

Denote by $\Phi(t,y)$ the average value of Φ over the energy surface

$$\begin{aligned}\Phi(t,y)&=\int\Phi(t,x,y)\,\delta(E-H(x,y))dx\bigg/\int\delta(E-H(x,y))dx\\&=\frac{1}{\Gamma_E(E,y)}\int\Phi(t,x,y)\,\delta(E-H(x,y))dx.\end{aligned} \tag{3.9}$$

Without loss of generality, the function $\Phi(t,x,y)$ can be presented in the form

$$\Phi(t,x,y)=\Phi(t,y)+\Phi'(t,x,y), \tag{3.10}$$

where $\Phi'(t,x,y)$ satisfy the constraint

$$\int\Phi'(t,x,y)\,\delta(E-H(x,y))dx=0. \tag{3.11}$$

Function $\Phi(t,y)$ is linked to the probability density function of y variables by a simple relation following from Eqs. (3.7) and (3.9),

$$f(t,y)=\Phi(t,y)\Gamma_E(E,y). \tag{3.12}$$

Let us show that, in terms of $f(t,y)$ and Φ', the probability flux takes a simple form

$$J^{\mu}=\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle f(t,y)+\omega^{\mu\nu}\int H'_{,\nu}\delta(E-H(x,y))\Phi'dx, \tag{3.13}$$

where $\langle\partial H/\partial y^{\mu}\rangle$ are the average values of $\partial H/\partial y^{\mu}$ over the energy surface and $H'_{,\nu}$ are the fluctuations of $\partial H/\partial y^{\mu}$,

$$H'_{,\nu}=\frac{\partial H}{\partial y^{\nu}}-\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle. \tag{3.14}$$

Indeed, in terms of $\Phi(t,y)$ and Φ', the probability flux can be written as

$$\begin{aligned}J^{\mu}=&\,\omega^{\mu\nu}\int\frac{\partial H}{\partial y^{\nu}}\delta(E-H(x,y))dx\Phi(t,y)\\&+\omega^{\mu\nu}\int\frac{\partial H}{\partial y^{\nu}}\delta(E-H(x,y))\Phi'dx.\end{aligned} \tag{3.15}$$

Since

$$\begin{aligned}\int\frac{\partial H}{\partial y^{\nu}}\delta(E-H(x,y))dx&=-\int\frac{\partial\theta(E-H(x,y))}{\partial y^{\nu}}dx\\&=-\frac{\partial}{\partial y^{\nu}}\int\theta(E-H(x,y))dx\\&=-\frac{\partial\Gamma(E,y)}{\partial y^{\nu}}\\&=-e^{S}\frac{\partial S(E,y)}{\partial y^{\nu}},\end{aligned}$$

we have for the average value of $\partial H/\partial y^{\nu}$ over the energy surface

$$\begin{aligned}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle&=\int\frac{\partial H}{\partial y^{\nu}}\delta(E-H(x,y))dx\bigg/\int\delta(E-H(x,y))dx\\&=-\frac{\partial\Gamma(E,y)}{\partial y^{\nu}}\bigg/\frac{\partial\Gamma(E,y)}{\partial E}\\&=-\frac{\partial S(E,y)}{\partial y^{\nu}}\bigg/\frac{\partial S(E,y)}{\partial E}.\end{aligned} \tag{3.16}$$

Equation (3.13) follows from Eqs. (3.15), (3.16), and (3.14).

If we find Φ' in terms of the probability density of slow variables $f(t,y)$ and, using Eq. (3.13), express J^{μ} in terms of $f(t,y)$, Eq. (3.5) will control the evolution of $f(t,y)$.

To find Φ', we assume that Φ' is small compared with $\Phi(t,y)$ and we use Liouville's equation to determine Φ'. Then we will check that Φ' obtained from this assumption is small indeed.

To write Liouville's equation in terms of function Φ', first we have to eliminate the degeneracy contained in formula (3.7): for given $f(t,x,y)$, function $\Phi(t,x,y)$ may contain an arbitrary dependence of $H(x,y)$ which is nulled by the factor $\delta(E-H(x,y))$. To do that, we introduce in x-phase space some curvilinear coordinates $\zeta^{\alpha}(x)$, $\alpha=1,\dots,2n-1$, and $h=H(x,y)$. Note that functions $\zeta^{\alpha}(x)$ do not depend on y. Dynamics in ζ^{α} variables is governed by the equations

$$\frac{d\zeta^{\alpha}}{dt}=v^{\alpha}(\zeta,y),\quad v^{\alpha}\equiv\frac{\partial\zeta^{\alpha}(x)}{\partial x^{i}}\omega^{ij}\frac{\partial H(x,y)}{\partial x^{j}}.$$

Liouville's equation for function $f(t,\zeta,h,y)$ takes the form

$$\frac{\partial f(t,\zeta,h,y)}{\partial t}+v^{\alpha}\frac{\partial f}{\partial \zeta^{\alpha}}+\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\frac{\partial f}{\partial y^{\nu}}=0. \qquad (3.17)$$

We see that the derivative with respect to h disappeared. Now we can rectify formula (3.7) requiring that function Φ in (ζ,h,y) variables does not depend on h,

$$f(t,\zeta,h,y)=\delta(E-h)\Phi(t,\zeta,y). \qquad (3.18)$$

This establishes a one-to-one correspondence between f and Φ.

Note that $f(t,\zeta,h,y)$ is not a probability density of ζ and y: the latter includes the Jacobian $\Delta(\zeta,h,y)$ of the transformation $(\zeta,h)\rightarrow x$,

$$dx=\Delta(\zeta,h,y)d\zeta dh.$$

Here $d\zeta=d\zeta^{1}\cdots d\zeta^{2n-1}$.

Probability density of (ζ,y) variables is $f\Delta$. To write Liouville's equation in terms of $f\Delta$, we note the identities

$$\frac{\partial}{\partial \zeta^{\alpha}}(v^{\alpha}\Delta)=0,\quad \frac{\partial}{\partial y^{\mu}}\left(\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\Delta\right)=0. \qquad (3.19)$$

We have from Eqs. (3.17) and (3.19)

$$\frac{\partial}{\partial t}f\Delta+\frac{\partial}{\partial \zeta^{\alpha}}(v^{\alpha}f\Delta)+\frac{\partial}{\partial y^{\alpha}}\left(\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}f\Delta\right)=0. \qquad (3.20)$$

Plugging Eq. (3.18) in Eq. (3.20), we obtain the equation for $\Phi(t,\zeta,y)$,

$$\frac{\partial}{\partial t}\Phi(t,\zeta,y)\Delta+\frac{\partial}{\partial \zeta^{\alpha}}(v^{\alpha}\Phi(t,\zeta,y)\Delta)+\frac{\partial}{\partial y^{\nu}}\left(\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\Phi(t,\zeta,y)\Delta\right)=0 \qquad (3.21)$$

and, from Eq. (3.8), the initial condition

$$\Phi(0,\zeta,y)=\frac{1}{\Gamma_{E}(E,y_{0})}\delta(y-y_{0}). \qquad (3.22)$$

Again, due to identities (3.19), Eq. (3.21) can be written also as

$$\frac{\partial \Phi(t,\zeta,y)}{\partial t}+v^{\alpha}\frac{\partial \Phi(t,\zeta,y)}{\partial \zeta^{\alpha}}+\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\frac{\partial \Phi(t,\zeta,y)}{\partial y^{\mu}}=0. \qquad (3.23)$$

In terms of ζ variables, some previous formulas take a simple form,

$$f(t,y)=\int \Phi(t,\zeta,y)\Delta d\zeta,$$

$$\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle=\int \frac{\partial H}{\partial y^{\nu}}\Delta d\zeta\Big/\int \Delta d\zeta,$$

$$\Gamma_{E}(E)=\int \Delta(\zeta,E,y)d\zeta dy,\quad \Gamma_{E}(E,y)=\int \Delta(\zeta,E,y)d\zeta,$$

$$\begin{aligned}J^{\mu}&=\omega^{\mu\nu}\int \frac{\partial H}{\partial y^{\nu}}\Phi(t,\zeta,y)\Delta d\zeta bf\\&=\omega^{\mu\nu}\int \frac{\partial H}{\partial y^{\nu}}\Delta d\zeta\Phi(t,y)+\omega^{\mu\nu}\int \frac{\partial H}{\partial y^{\nu}}\Phi'(t,\zeta,y)\Delta d\zeta\\&=\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle f(t,y)+\omega^{\mu\nu}\int H'_{,\nu}\Phi'(t,\zeta,y)\Delta d\zeta. \qquad (3.24)\end{aligned}$$

Measure $\Delta(\zeta,h,y)d\zeta$ has the sense of the invariant measure on the energy surfaces in x-phase space.

We seek a solution of Eq. (3.23) of the form $\Phi=\Phi(t,y)+\Phi'(t,\zeta,y)$. Let us impose on $\Phi(t,y)$ the initial condition

$$\Phi(t,y)=\frac{1}{\Gamma_{E}(E,y_{0})}\delta(y-y_{0}). \qquad (3.25)$$

Then $\Phi'(t,\zeta,y)$ is zero at the initial instant,

$$\Phi'(0,\zeta,y)=0. \qquad (3.26)$$

Function $\Phi'(t,\zeta,y)$ satisfies the equation

$$\begin{aligned}\frac{\partial \Phi'}{\partial t}+v^{\alpha}\frac{\partial \Phi'}{\partial \zeta^{\alpha}}=&-\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\frac{\partial \Phi(t,y)}{\partial y^{\mu}}-\frac{\partial \Phi(t,y)}{\partial t}\\&-\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\frac{\partial \Phi'}{\partial y^{\nu}}. \qquad (3.27)\end{aligned}$$

Let us show that Eq. (3.27) can be written as

$$\begin{aligned}\frac{\partial \Phi'}{\partial t}+v^{\alpha}\frac{\partial \Phi'}{\partial \zeta^{\alpha}}=&-\omega^{\mu\nu}H'_{,\nu}\frac{\partial \Phi(t,y)}{\partial y^{\mu}}-\omega^{\mu\nu}\left(\frac{\partial H}{\partial y^{\nu}}\frac{\partial \Phi'}{\partial y^{\mu}}\right.\\&\left.-\frac{1}{\Gamma_{E}(E,y)}\frac{\partial}{\partial y^{\mu}}\int H_{,\nu}\Phi'\Delta d\zeta\right). \qquad (3.28)\end{aligned}$$

Indeed, from Eqs. (3.5), (3.12), (3.13), and (3.16),

$$\begin{aligned}\frac{\partial \Phi(t,y)}{\partial t}\Gamma_{E}(E,y)+\frac{\partial}{\partial y^{\mu}}\Bigg[&-\omega^{\mu\nu}\frac{\partial \Gamma(E,y)}{\partial y^{\nu}}\Phi(t,y)\\&+\int \omega^{\mu\nu}H'_{,\nu}\Phi'\Delta d\zeta\Bigg]=0. \qquad (3.29)\end{aligned}$$

The second term can be transformed to $-\omega^{\mu\nu}\partial\Gamma(E,y)/\partial y^{\nu}\partial\Phi(t,y)/\partial y^{\mu}$.

Therefore, due to Eq. (3.16),

$$\frac{\partial\Phi(t,y)}{\partial t}=-\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle\frac{\partial\Phi(t,y)}{\partial y^{\mu}}-\frac{1}{\Gamma_E(E,y)}\frac{\partial}{\partial y^{\mu}}\times\int\omega^{\mu\nu}H'_{,\nu}\Phi'\Delta d\zeta. \tag{3.30}$$

Plugging Eq. (3.30) in Eq. (3.27), we obtain Eq. (3.28).

Note that the average value of the right-hand side of Eq. (3.28) over ζ (with the weight Δ) is zero: vanishing of the integral of the first term follows from a definition of $H'_{,\mu}$ (3.14) while integral of the last two terms vanishes due to the second Eq. (3.19),

$$\begin{aligned}&\int\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\frac{\partial\Phi'}{\partial y^{\mu}}\Delta d\zeta\\&\quad-\int\frac{\Delta d\zeta}{\Gamma_E(E,y)}\frac{\partial}{\partial y^{\mu}}\int\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\Phi'\Delta d\zeta\\&=\int\frac{\partial}{\partial y^{\mu}}\left(\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\Delta\Phi'\right)d\zeta\\&\quad-\frac{\partial}{\partial y^{\mu}}\int\omega^{\mu\nu}\frac{\partial H}{\partial y^{\nu}}\Phi'\Delta d\zeta\\&=0.\end{aligned}$$

Let us assume that the last two terms on the right-hand side of Eq. (3.28) are much smaller than the first one. Then we arrive at the initial value problem

$$\frac{\partial\Phi'}{\partial t}+v^{\alpha}\frac{\partial\Phi'}{\partial\zeta^{\alpha}}=-\omega^{\mu\nu}H'_{,\nu}\frac{\partial\Phi(t,y)}{\partial y^{\mu}},\quad\Phi'(0,\zeta,y)=0. \tag{3.31}$$

We will find Φ' from this problem and then, in Appendix B, check under which conditions such simplification makes sense.

To write the solution of Eq. (3.31) in an explicit form, we introduce a system of ordinary differential equations,

$$\frac{d\zeta^{\alpha}}{dt}=v^{\alpha}(\zeta,y). \tag{3.32}$$

This system determines the mapping of values of ζ at some instant t',ζ' to the values at instant t,

$$\zeta=\chi(t;t',\zeta';y),\qquad\frac{\partial\chi(t;t',\zeta';y)}{\partial t}=v^{\alpha}(\chi,y). \tag{3.33}$$

The last argument in Eqs. (3.33) emphasizes the dependence of the mapping on parameters y. In Eq. (3.33), t' may be either less than or greater than t. In the latter case, the mapping is determined by the differential equations (3.33) with the final condition

$$\chi(t;t',\zeta';y)\to\zeta'\quad\text{if}\quad t\to t'-0.$$

The solution of the problem (3.31) can be written in terms of mapping (3.33) as

$$\Phi'(t,\zeta,y)=-\int_0^t\omega^{\mu\nu}H'_{,\nu}(\chi(t';t,\zeta;y),y)\frac{\partial\Phi(t',y)}{\partial y^{\mu}}dt'. \tag{3.34}$$

Plugging Eq. (3.34) in Eq. (3.24), we find the probability flux

$$\begin{aligned}J^{\mu}&=\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle f(t,y)\\&\quad-\omega^{\mu\kappa}\int H'_{,\kappa}(\zeta,y)\omega^{\nu\lambda}H'_{,\lambda}(\chi(\tilde{t};t,\zeta;y)y)\\&\quad\times\frac{\partial\Phi(\tilde{t},y)}{\partial y^{\nu}}d\tilde{t}\Delta d\zeta\\&=\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle f(t,y)-\int_0^t K^{\mu\nu}(\tilde{t},t;y)\frac{\partial\Phi(\tilde{t},y)}{\partial y^{\nu}}\Gamma_E(E,y)d\tilde{t},\end{aligned} \tag{3.35}$$

where

$$\begin{aligned}&K^{\mu\nu}(\tilde{t},t;y)\\&\quad=\omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\zeta,y)H'_{,\lambda}(\chi(\tilde{t};t,\zeta;y),y)\frac{\Delta(\zeta,y)d\zeta}{\Gamma_E(E,y)}\end{aligned} \tag{3.36}$$

or

$$\begin{aligned}J^{\mu}&=\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle f(t,y)+\int_0^t K^{\mu\nu}(\tilde{t},t;y)f(\tilde{t},y)d\tilde{t}\frac{\partial\ln\Gamma_E}{\partial y^{\nu}}\\&\quad-\int_0^t K^{\mu\nu}(\tilde{t},t;y)\frac{\partial f(\tilde{t},y)}{\partial y^{\nu}}d\tilde{t}.\end{aligned} \tag{3.37}$$

Note that $K^{\mu\nu}(\tilde{t},t;y)$ are defined both for $\tilde{t}>t$ and $\tilde{t}<t$. If $K^{\mu\nu}(\tilde{t},t;y)$ decay as $t-\tilde{t}$ increases so fast that $K^{\mu\nu}\approx0$ for $t-\tilde{t}>\tau$, τ is the correlation time, and $f(t,y)$ do not change appreciably on the times of order τ, then

$$\begin{aligned}J^{\mu}&=\left(\omega^{\mu\nu}\left\langle\frac{\partial H}{\partial y^{\nu}}\right\rangle+D^{\mu\nu}\frac{\partial\ln\Gamma_E(E,y)}{\partial y^{\nu}}\right)f(t,y)\\&\quad-D^{\mu\nu}(y)\frac{\partial f(t,y)}{\partial y^{\nu}},\end{aligned} \tag{3.38}$$

where

$$D^{\mu\nu}(y)=\int_0^t K^{\mu\nu}(\tilde{t},t;y)d\tilde{t}=\int_{-\infty}^t K^{\mu\nu}(\tilde{t},t;y)d\tilde{t}. \tag{3.39}$$

Since $\partial H/\partial y^{\nu}=-T\partial S/\partial y^{\nu}$ and

$$\begin{aligned}\frac{\partial \ln \Gamma_E(E,y)}{\partial y^\nu} &= \frac{1}{\Gamma_E}\frac{\partial^2 e^S}{\partial E \partial y^\nu}\\ &= e^{-S}\frac{1}{S_E}\frac{\partial}{\partial y^\nu}(e^S S_E)\\ &= \frac{1}{S_E}\left(S_{,\nu}S_E + \frac{\partial}{\partial y^\nu}S_E\right)\\ &= \frac{\partial S}{\partial y^\nu} - \frac{1}{T}\frac{\partial T}{\partial y^\nu},\end{aligned}$$

the probability flux (3.38) coincides with Eqs. (2.25) and (2.29) if one identifies $D^{\mu\nu}$ in Eqs. (2.25) and (2.29) with the tensor (3.39). The latter can be done if tensor $D^{\mu\nu}$ (3.39) is symmetric and positive. Symmetry features of tensor $D^{\mu\nu}$ are considered in the next section.

IV. SYMMETRY OF DIFFUSION COEFFICIENTS AND REVERSIBILITY OF MOTION

Consider mapping Eq. (3.33) in more detail. First, since Eqs. (3.32) are autonomous, mapping depends only on the difference $\tau = t - t'$, and

$$\zeta = \chi(\tau, \zeta'). \tag{4.1}$$

For brevity, we drop the dependence of all functions on parameters y in the formulas of this section; it can be restored in the final relations.

Time shift τ in Eq. (4.1) can be positive and negative. The inverse mapping is

$$\zeta' = \chi(-\tau, \zeta). \tag{4.2}$$

Tensor $K^{\mu\nu}$ (3.36) depends on $\tau = \tilde{t} - t$. It enters in all relations at negative values of $\tau(\tilde{t} < t)$. Nevertheless, it is convenient to consider formula (3.36) as the definition of $K^{\mu\nu}$ for positive values of τ as well. Then the following property of $K^{\mu\nu}$ holds:

$$K^{\mu\nu}(\tau) = K^{\nu\mu}(-\tau). \tag{4.3}$$

Indeed,

$$K^{\mu\nu}(\tau) = \omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\zeta)H'_{,\lambda}(\chi(\tau,\zeta))\frac{\Delta(\zeta)d\zeta}{\Gamma_E}. \tag{4.4}$$

Let us change the variable of integration ζ by $\zeta' = \chi(\tau, \zeta)$. Due to Eq. (3.19), measure $\Delta(\zeta)d\zeta$ is conserved by the flow (3.32), therefore

$$\Delta(\zeta)d\zeta = \Delta(\zeta')d\zeta'.$$

Using also that $\zeta = \chi(-\tau, \zeta')$, we have

$$\begin{aligned}&K^{\mu\nu}(\tau)\\ &\quad = \omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\chi(-\tau,\zeta'))H_{,\lambda}(\zeta')\frac{\Delta(\zeta')d\zeta'}{\Gamma_E}.\end{aligned} \tag{4.5}$$

The right-hand side of Eq. (4.5) is $K^{\nu\mu}(-\tau)$, as claimed.

Additional symmetry properties may be warranted if micromotion is reversible. That means the following. Denote by (p,q) the generalized momenta and coordinates of micromotion, $x = (p,q)$, and by x^* the point $x^* = (-p,q)$. If

$$H(x,y) = H(x^*,y) \tag{4.6}$$

and the phase flow moves a point x to the point x_1, then, as is easy to check, for the same time interval it moves also the point x_1^* to the point x^*. In terms of mapping (4.1), it can be written as

$$\zeta_1 = \chi(\tau,\zeta) \;\Rightarrow\; \zeta^* = \chi(\tau,\zeta_1^*) \quad \text{or} \quad \zeta_1^* = \chi(-\tau,\zeta^*) \quad \text{or}$$

$$\chi^*(\tau,\zeta) = \chi(-\tau,\zeta^*), \tag{4.7}$$

where ζ and ζ_1 are ζ coordinates of the points x and x_1.

This property yields the symmetry of tensor $K^{\mu\nu}(\tau)$:

$$K^{\mu\nu}(\tau) = K^{\nu\mu}(\tau). \tag{4.8}$$

Indeed, from Eq. (4.6), $\partial H/\partial y^\mu|_\zeta = \partial H/\partial y^\mu|_{\zeta^*}$. Choosing also coordinates ζ in such a way that $\Delta(\zeta)d\zeta = \Delta(\zeta^*)d\zeta^*$, we have

$$\begin{aligned}K^{\mu\nu}(\tau) &= \omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\zeta)H'_{,\lambda}(\chi(\tau,\zeta))\frac{\Delta(\zeta)d\zeta}{\Gamma_E}\\ &= \omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\zeta)H'_{,\lambda}(\chi^*(\tau,\zeta))\frac{\Delta(\zeta)d\zeta}{\Gamma_E}\\ &= \omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\zeta)H'_{,\lambda}(\chi(-\tau,\zeta^*))\frac{\Delta(\zeta)d\zeta}{\Gamma_E}\\ &= \omega^{\mu\kappa}\omega^{\nu\lambda}\int H'_{,\kappa}(\zeta^*)H'_{,\lambda}(\chi(-\tau,\zeta^*))\frac{\Delta(\zeta^*)d\zeta^*}{\Gamma_E}\\ &= K^{\mu\nu}(-\tau).\end{aligned} \tag{4.9}$$

Formula (4.8) follows from Eqs. (4.9) and (4.3). In contrast to Eq. (4.8), formula (4.3) is universal, it does not depend on the symmetry properties of the Hamiltonian.

In the presence of magnetic field, **m**, relation (4.6) must be modified,

$$H(x,y,\mathbf{m}) = H(x^*,y,-\mathbf{m}), \tag{4.10}$$

and, as is easy to check, Eq. (4.9) is replaced by

$$K^{\mu\nu}(\tau,\mathbf{m}) = K^{\mu\nu}(-\tau,-\mathbf{m}). \tag{4.11}$$

Thus, from Eqs. (4.11) and (4.3), we have

$$K^{\mu\nu}(\tau,\mathbf{m}) = K^{\nu\mu}(\tau,-\mathbf{m}). \tag{4.12}$$

Symmetry of $K^{\mu\nu}$ yields the symmetry of the diffusion coefficients.

V. DYNAMICS OF SLOW VARIABLES AS A MARKOV PROCESS

Approximation of the probability flux (3.37) by the expression (3.38) yields the equation for probability density of y variables, $f(t,y)$,

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial}{\partial y^{\mu}}\left[\left(-\omega^{\mu\nu}T\frac{\partial S}{\partial y^{\nu}}+D^{\mu\nu}\frac{\partial S}{\partial y^{\nu}}\right)f(t,y)-D^{\mu\nu}\frac{\partial f(t,y)}{\partial y^{\nu}}\right]=0. \tag{5.1}$$

If $D^{\mu\nu}$ is a positive symmetric tensor, then Eq. (5.1) may be interpreted as a Fokker-Planck equation for a Markov process $y(t)$.

If $D^{\mu\nu}$ are not symmetric, as may be the case in the presence of magnetic field, such an interpretation of Eq. (5.1) fails. However, one can make the following transformation: let us present $D^{\mu\nu}$ as a sum of its symmetric part, $D^{(\mu\nu)}$, and its antisymmetric part, $D^{[\mu\nu]}$,

$$D^{\mu\nu}=D^{(\mu\nu)}+D^{[\mu\nu]},\quad (D^{(\mu\nu)}=D^{(\nu\mu)},\quad D^{[\mu\nu]}=-D^{[\nu\mu]}),$$

and use the identity

$$\frac{\partial}{\partial y^{\mu}}\left(D^{[\mu\nu]}\frac{\partial f(t,y)}{\partial y^{\nu}}\right)=\frac{\partial D^{[\mu\nu]}}{\partial y^{\mu}}\frac{\partial f(t,y)}{\partial y^{\nu}}=\frac{\partial}{\partial y^{\nu}}\left(\frac{\partial D^{[\mu\nu]}}{\partial y^{\mu}}f(t,y)\right).$$

Then Eq. (5.1) can be written in the form

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial}{\partial y^{\mu}}\left[\left(-\omega^{\mu\nu}T\frac{\partial S}{\partial y^{\nu}}+D^{\mu\nu}\frac{\partial S}{\partial y^{\nu}}-\frac{\partial D^{[\mu\nu]}}{\partial y^{\nu}}\right)\times f(t,y)-D^{(\mu\nu)}\frac{\partial f(t,y)}{\partial y^{\nu}}\right]=0. \tag{5.2}$$

We see that y variables can be interpreted as a Markov process with the coefficients $D^{(\mu\nu)}$ and the drift

$$V^{\mu}=-\omega^{\mu\nu}T\frac{\partial S}{\partial y^{\nu}}+D^{\mu\nu}\frac{\partial S}{\partial y^{\nu}}-\frac{\partial D^{[\mu\nu]}}{\partial y^{\nu}}. \tag{5.3}$$

The difference from the case of symmetric diffusion coefficients is in the second term, where the nonsymmetric coefficients stand, and in the additional third term. Since diffusion coefficients $D^{\mu\nu}$ are in the order of T, and S is proportional to the number of degrees of freedom, the contribution of the third term in Eq. (5.3) is small unless the derivative of $D^{[\mu\nu]}$ with respect to y brings a large factor. For example, one cannot rule out a dependence $D^{[\mu\nu]}$ on y of the form const $e^{S(y)}$. In the typical cases when derivatives of $D^{\mu\nu}$ are in the order of $D^{\mu\nu}$, the third term can be neglected.

VI. CONSTITUTIVE EQUATIONS FOR SOLIDS

In this section, we use the quasi-Hamiltonian structure of macroequations to establish a general form of the constitutive equations for solids. We are going to show that, if stresses $\sigma=(\sigma^{ij})$, $i,j=1,2,3$, depend on strains $\varepsilon=(\varepsilon_{ij})$ and strain rates $\dot{\varepsilon}=(\dot{\varepsilon}_{ij})$ only, then the general form of such a dependence is

$$\sigma^{ij}=-T\frac{\partial\eta(U,\varepsilon_{ij})}{\partial\varepsilon_{ij}}+\frac{1}{T}D^{ijkl}(T,\varepsilon,\dot{\varepsilon})\dot{\varepsilon}_{kl}, \tag{6.1}$$

where $\eta(U,\varepsilon)$ is the equilibrium entropy per unit volume and the viscosity tensor, D^{ijkl} has the symmetry of elastic moduli tensor,

$$D^{ijkl}=D^{jikl}=D^{ijlk}=D^{klij}. \tag{6.2}$$

The dissipative part of the constitutive equations is obviously not potential. Consider, for example, the case of isotropic body,

$$D^{ijkl}\dot{\varepsilon}_{kl}=A(\varepsilon,\dot{\varepsilon})\delta^{ij}\dot{\varepsilon}^{k}_{k}+B(\varepsilon,\dot{\varepsilon})\dot{\varepsilon}^{ij}.$$

Let us assume that the coefficients A,B depend only on the first two invariants of the strain rate tensor: $e_1=\dot{\varepsilon}^{k}_{k}$ and $e_2=\frac{1}{2}\dot{\varepsilon}_{ij}\dot{\varepsilon}^{ij}$. Then one can easily check that potentiality takes place if and only if

$$\frac{\partial}{\partial e_2}(e_1A)=\frac{\partial B}{\partial e_1}.$$

No reasons are seen why this condition might be satisfied in the general case.

To obtain Eqs. (6.1) and (6.2), consider a piece of solid deformed homogeneously. At the boundary of this piece the stresses σ^{ij} are applied. Microscopically this means that there is a force acting on the boundary particles of the form $\sigma^{ij}n_j\alpha$, where σ^{ij} are constants depending on time, n_j are the components of the unit normal vector to the boundary, and α is the surface area per one particle. The Hamiltonian of the system has the form

$$H=H(x)-\alpha\sum_{a\in\partial V}\sigma^{ij}n_j q_{i(a)}, \tag{6.3}$$

where $q_{i(a)}$ are the coordinates of the ath particle and ∂V is the boundary of V. We define the strain tensor as

$$\varepsilon_{ij}=\frac{\alpha}{|V|}\sum_{a\in\partial V}\frac{1}{2}(n_j q_{i(a)}+n_i q_{j(a)}), \tag{6.4}$$

where $|V|$ is the specimen volume.

One can make a change of variables in the Hamiltonian system choosing the components of the strain tensor (6.4) as some of the coordinates of the system. Denote the corresponding momenta by P^{ij}. We assume that ε_{ij} and P^{ij} are the only slow coordinates of the system and denote all other coordinates by x. The Hamiltonian can be written in the form

$$H=\frac{1}{2}I_{ijkl}P^{ij}P^{kl}+H_0(x,\varepsilon)-\sigma^{ij}\varepsilon_{ij}|V|. \tag{6.5}$$

If $|V|\to 0$, the momenta and the coefficients I_{ijkl} have the orders $P^{ij}\sim\dot{\varepsilon}_{ij}|V||V|^{2/3}$, $I_{ijkl}\sim|V|^{-5/3}$. The possibility of separating the inertia of homogeneous deformation may be established, for example, by Kunin's quasicontinuum technique [19]. Denote by $S_0(U,\varepsilon)$ the entropy of the system with Hamiltonian $H_0(x,\varepsilon)$ and the value of energy U, and by η the entropy per unit volume, $S_0/|V|=\eta(U/|V|,\varepsilon)$. Then the entropy of the system under consideration is

$$S=|V|\,\eta\left(|V|^{-1}E-\frac{1}{2}|V|^{-1}I_{ijkl}P^{ij}P^{kl}+\sigma^{ij}\varepsilon_{ij},\varepsilon_{ij}\right). \tag{6.6}$$

Equations (2.30) take the form

$$\dot{\varepsilon}_{ij}=I_{ijkl}P^{kl},$$

$$\dot{P}^{ij}=-\left(\sigma^{ij}+T\frac{\partial\eta}{\partial\varepsilon_{ij}}\right)|V|+\frac{1}{T}D^{ijkl}I_{klmn}P^{mn}. \tag{6.7}$$

For $|V|\to 0$ and $\dot{\varepsilon}_{ij}$ finite, $P^{ij}\sim|V|^{5/3}$ and the left-hand side of Eq. (6.7) can be neglected. Then Eq. (6.7) transforms to Eq. (6.1). Symmetry properties of D^{ijkl} follow from the symmetry properties of $D^{\mu\nu}$ in Eq. (2.30).

VII. HEAT CONDUCTIVITY

The simplest example of a noncanonical variable is energy. Energy appears as an additional slow variable if one considers an interaction of a number of systems. Total energy is conserved while energy of each subsystem changes slowly due to "heat (energy) transfer" between the subsystems. We are going to derive in this section the equations of nonlinear heat conduction. This involves, in addition to the asymptotic analysis of Liouville's equation, a limit transition from a finite number of subsystems to continuum. The resulting equations are as follows.

Let $\xi=(\xi^1,\xi^2,\xi^3)=(\xi^\mu)$, $\mu=1,2,3$ be a point of three-dimensional continuum, and $U(t,\xi)$ and $\eta(U,\xi)$ be energy and entropy densities per unit volume, respectively. The system is isolated. Let the system be deviated from the equilibrium state initially. Then the evolution to equilibrium is governed by the equations

$$\frac{\partial U}{\partial t}=-\frac{\partial q^\mu}{\partial\xi^\mu}, \tag{7.1}$$

$$q^\mu=\int D^{\mu\nu}(\xi,\xi')\frac{\partial}{\partial\xi'^{\nu}}\frac{1}{T(t,\xi')}d^3\xi',$$

$$\frac{1}{T}=\frac{\partial\eta(U,\xi)}{\partial U}, \tag{7.2}$$

where $D^{\mu\nu}$ is a symmetric positive kernel

$$D^{\mu\nu}(\xi,\xi')=D^{\nu\mu}(\xi',\xi), \tag{7.3}$$

$$\int D^{\mu\nu}(\xi,\xi')\varphi_\mu(\xi)\varphi_\nu(\xi')d^3\xi d^3\xi'\geqslant 0 \tag{7.4}$$

for any $\varphi_\mu(\xi)$.

The kernel might be a functional of the temperature field. In the case of local dependence of heat flux q^μ only on temperature and temperature gradient,

$$q^\mu=D^{\mu\nu}(\xi,T,\nabla T)\frac{\partial}{\partial\xi^\nu}\frac{1}{T} \tag{7.5}$$

with a symmetric positive tensor $D^{\mu\nu}(\xi,T,\nabla T)$.

Derivation of Eqs. (7.1)–(7.5) involves a number of assumptions formulated below as some necessary notations are introduced.

A. Hamiltonian system

Let us have N subsystems and let x_a be generalized momenta and coordinates of the ath subsystem, $a=1,\ldots,N$; $x_a=(x_a^1,\ldots,x_a^{2n})$. The dimension of x_a-phase space may change from one subsystem to another, but we do not emphasize this in the notation. There is no mass exchange between subsystems, and each subsystem consists of the same particles all the time. The Hamiltonian of the system is taken in the form

$$H=\sum_{a=1}^{N}H_a(x_a)+\tilde{H}(x_1,\ldots,x_N), \tag{7.6}$$

where $H_a(x_a)$ is the Hamiltonian of the isolated ath subsystem and $\tilde{H}(x_1,\ldots,x_N)$ is the interaction Hamiltonian. The sense in the partition of the Hamiltonian into the sum (7.6) is introduced by the assumption

$$\tilde{H}(x_1,\ldots,x_N)\ll\sum_{a=1}^{N}H_a(x_a). \tag{7.7}$$

Again, though we are going to use some asymptotic reasoning, we do not need to introduce a formal small parameter via Eq. (7.7) since we are interested only in the leading terms of the asymptotic expansion.

We assume that each subsystem is ergodic on the energy surfaces $H_a(x_a)=\text{const}$, and the entire system with Hamiltonian (7.6) is also ergodic on the energy surfaces $H(x_1,\ldots,x_N)=E=\text{const}$ in $(x_1,\ldots,x_N)$-phase space. We are going to study the dynamics of slow variables $h_a=H_a(x_a)$.

B. Equilibrium distribution

In thermodynamical equilibrium, the probability density of variables $h=(h_1,\ldots,h_N)$, $f_\infty(h)$, is given by the formula (see, e.g., [10], Section 2.2)

$$f_\infty(h)=\frac{1}{\Gamma_E(E)}\Gamma_{Eh},$$

$$\Gamma_{Eh}\equiv\frac{\partial}{\partial E\,\partial h_1\cdots\partial h_N}\Gamma(E,h), \tag{7.8}$$

where

$$\Gamma(E,h)=\int\theta(E-H(x))\theta(h_1-H_1(x_1))\cdots\theta \\ \times(h_N-H_N(x_N))dx_1\cdots dx_N \tag{7.9}$$

and

$$\Gamma(E)=\int\theta(E-H(x))dx_1\cdots dx_N.$$

Obviously,

$$\Gamma_{Eh}=\int\delta\left(E-\sum_{a=1}^{N}H_a-\tilde{H}\right)\delta(h_1-H_1(x_1))\cdots \\ \delta(h_N-H_N(x_N))dx_1\cdots dx_N. \tag{7.10}$$

Denote by $\Gamma_a(h_a)$ the phase volume of the ath subsystem,

$$\Gamma_a(h_a)=\int\theta(h_a-H_a(x_a))dx_a,$$

and by $S_a(h_a)$ its entropy

$$S_a(h_a)=\ln\Gamma_a(h_a). \tag{7.11}$$

Assume that the number of degrees of freedom in each subsystem is huge. Then the formula for entropy (7.11) is asymptotically equivalent to the formula

$$S_a(h_a)=\ln\frac{d\Gamma_a(h_a)}{dh_a}.$$

Neglecting the interaction energy $\tilde{H}$ compared to $\Sigma_{a=1}^{N}H_a$ in Eq. (7.10), we obtain

$$\Gamma_{Eh}=\delta\left(E-\sum_{a=1}^{N}h_a\right)\frac{d\Gamma_1}{dh_1}\cdots\frac{d\Gamma_N}{dh_N} \\ =\delta\left(E-\sum_{a=1}^{N}h_a\right)e^{S_1(h_1)+\cdots+S_N(h_N)} \tag{7.12}$$

and, for probability density,

$$f_\infty(h)=\frac{1}{\Gamma_E(E)}\delta\left(E-\sum_{a=1}^{N}h_a\right)e^{S(h_1)+\cdots+S_N(h_N)}. \tag{7.13}$$

We see from Eq. (7.13) that the most probable state corresponds to the maximum value of total entropy $S_1(h_1)+\cdots+S_N(h_N)$, under the condition that the total energy is fixed, $\Sigma_{a=1}^{N}h_a=E$ (cf. [20]). This is obviously the state with equal temperatures of the subsystems, $T_a=[dS_a(h_a)/dh_a]^{-1}$.

In what follows, we need a more precice formula than Eq. (7.12) in which the δ function is replaced by a more exact expression. Namely, let us show that

$$\Gamma_{Eh}=\text{const}\,e^{S_1(h_1)+\cdots+S_N(h_N)-(E-h_1-\cdots-h_N)^2/2\varepsilon^2} \tag{7.14}$$

with some small constant ε depending on interaction energy. Indeed, using the Fourier presentation of the δ function,

$$\delta(E)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{iEz}dz,$$

one can write Eq. (7.10) as

$$\Gamma_{Eh}=\frac{1}{2\pi}\int_{\infty}^{+\infty}e^{iz(E-\Sigma h_a)}e^{-iz\tilde{H}} \\ \times\prod_{a=1}^{N}\delta(h_a-H_a(x_a))dx_1\cdots dx_N. \tag{7.15}$$

Keeping in the expansion of $\exp(-iz\tilde{H})$,

$$e^{-iz\tilde{H}}=1-iz\tilde{H}-\frac{z^2}{2}\tilde{H}^2+\cdots,$$

only the first three terms, we get

$$\Gamma_{Eh}=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{iz(E-\Sigma h_a)-izA-(z^2/2)\varepsilon^2}dz\frac{d\Gamma_1}{dh_1}\cdots\frac{d\Gamma_N}{dh_N}, \tag{7.16}$$

where we introduced the notations

$$A=\int\tilde{H}\delta(h_1-H_1)\cdots\delta(h_N-H_N)dx_1\cdots dx_N,$$

$$\varepsilon^2=\int\tilde{H}^2\delta(h_1-H_1)\cdots\delta(h_N-H_N)dx_1\cdots dx_N-A^2.$$

From Eq. (7.16), computing the integral over z, we obtain

$$\Gamma_{Eh}=\text{const}\,e^{-(E-\Sigma h_a-A)^2/2\varepsilon^2}\frac{d\Gamma_1}{dh_1}\cdots\frac{d\Gamma_N}{dh_N}. \tag{7.17}$$

Neglecting A compared to Σh_a, we arrive at Eq. (7.14).

An interesting consequence of Eq. (7.14) is that (for constant ε) the most probable state satisfies the equations

$$\frac{\partial S_a}{\partial h_a}+\frac{E-h_1-\cdots-h_N}{\varepsilon^2}=0.$$

Therefore, at the point of equilibrium, $E-h_1-\cdots-h_N$ is a small negative constant,

$$E-h_1-\cdots-h_N=-\frac{\varepsilon^2}{T}. \tag{7.18}$$

C. Evolution to equilibrium

First, we are going to show that evolution to equilibrium is described by a system of ordinary differential equations,

$$\frac{dh_a}{dt}=D_{ab}\frac{\partial S_b}{\partial h_b}, \tag{7.19}$$

where D_{ab} are some functions on energies of all subsystems which obey the conditions

$$D_{ab}=D_{ba}, \qquad \sum_b D_{ab}=0. \tag{7.20}$$

The physical properties of the system are described by equilibrium entropies of subsystems $S_b(h_b)$. The total energy of the system $h_1+\cdots+h_N$ is constant in the process of evolution as follows from Eqs. (7.19) and (7.20). The total entropy $S=S_1(h_1)+\cdots+S_N(h_N)$ changes in accordance with the equation

$$\frac{dS}{dt}=D_{ba}\frac{\partial S_a}{\partial h_a}\frac{\partial S_b}{\partial h_b}=D_{ab}\frac{\partial S}{\partial h_a}\frac{\partial S}{\partial h_b}. \tag{7.21}$$

Positivity of the quadratic form

$$D_{ab}\frac{\partial S}{\partial h_a}\frac{\partial S}{\partial h_b}\geqslant 0 \tag{7.22}$$

warrants the growth of entropy.

Evolution equations (7.19) do not have a Hamiltonian part; they are pure dissipative.

D. Cauchy's problem for Liouville's equation

The dynamics of the system is described by Liouville's equation for probability density $f(t,x)$,

$$\frac{\partial f(t,x)}{\partial t}+\omega^{ij}\frac{\partial H}{\partial x_a^j}\frac{\partial f(t,x)}{\partial x_a^i}=0. \tag{7.23}$$

We need to introduce explicitly the slow variables $h_a=H_a(x_a)$. To this end, we choose some curvilinear coordinates $\zeta_a=(\zeta_a^\alpha)$, $\alpha=1,\ldots,2n-1$, on the surfaces $H_a(x_a)=h_a=\text{const}$ in the x_a-phase spaces and write Liouville's equation in coordinates $\zeta=(\zeta_1,\ldots,\zeta_N)$, $h=(h_1,\ldots,h_N)$,

$$\frac{\partial f(t,\zeta,h)}{\partial t}+v_a^\alpha\frac{\partial f(t,\zeta,h)}{\partial \zeta_a^\alpha}+v_a\frac{\partial f(t,\zeta,h)}{\partial h_a}=0. \tag{7.24}$$

Here the following notations are used:

$$v_a^\alpha\equiv\frac{\partial \zeta_a^\alpha}{\partial x_a^i}\omega^{ij}\frac{\partial H}{\partial x_a^j}=v_a^\alpha(\zeta,h),$$

$$v_a\equiv\frac{\partial H_a}{\partial x_a^i}\omega^{ij}\frac{\partial H}{\partial x_a^j}=v_a(\zeta,h). \tag{7.25}$$

Denote by Δ_a the Jacobian of transformation $x_a\to(\zeta_a,h_a)$,

$$dx_a=\Delta_a(\zeta_a,h_a)d\zeta_a dh_a,$$

and by the Δ the product of Jacobians,

$$\Delta=\Delta_1\Delta_2\cdots\Delta_N.$$

The function $f(t,\zeta,h)\Delta(\zeta,h)$ is the probability distribution of variables ζ and h at the instant t.

Initially let $f(t,\zeta,h)$ have the value

$$f(0,\zeta,h)=c\,\delta(E-H(x(\zeta,h)))\times\delta(\mathring{h}_1-h_1)\cdots\delta(\mathring{h}_N-h_N). \tag{7.26}$$

The constant c is determined from the normalization condition

$$\int f(0,\zeta,h)\Delta\, d\zeta dh=1. \tag{7.27}$$

Our task is to study the solution of Liouville's equation with the initial data (7.26). We expect that probability density of h variables evolves to the function (7.13) [or, more precisely, (7.14)], while probability density of ζ and h approaches the function

$$f(\infty,\zeta,h)=c_\infty\delta(E-H(x(\zeta,h))). \tag{7.28}$$

E. Liouville's equation and evolution equations (7.19)

First, let us obtain the equation for the probability function of variables h,

$$f(t,h)=\int f(t,\zeta,h)\Delta\, d\zeta. \tag{7.29}$$

Note the identity

$$\frac{\partial(v_a^\alpha\Delta)}{\partial\zeta_a^\alpha}+\frac{\partial(v_a\Delta)}{\partial h_a}=0. \tag{7.30}$$

According to Eq. (7.30), Liouville's equation (7.24) can be written also in the divergence form,

$$\frac{\partial f(t,\zeta,h)\Delta}{\partial t}+\frac{\partial(v_a^\alpha f(t,\zeta,h)\Delta)}{\partial\zeta_a^\alpha}+\frac{\partial(v_a f(t,\zeta,h)\Delta)}{\partial h_a}=0. \tag{7.31}$$

Integrating Eq. (7.28) over ζ, we find

$$\frac{\partial f(t,h)}{\partial t}+\frac{\partial J_a}{\partial h_a}=0,$$

$$J_a=\int v_a f(t,\zeta,h)\Delta d\zeta. \tag{7.32}$$

Our goal is to show that the probability flux J_a in Eq. (7.32) in the first approximation has the form

$$J_a=D_{ab}\frac{\partial S_b}{\partial h_b}f(t,h)-D_{ab}\frac{\partial f(t,h)}{\partial h_b}. \tag{7.33}$$

Then the evolution equations (7.19) follow from Eqs. (7.32) and (7.33) and the smallness of the diffusion coefficients D_{ab}. Note that the probability flux vanishes at equilibrium. Therefore, Eq. (7.33) can be obtained also from reversibility of micromotion and the hypothesis on Markov's character of approaching to equilibrium.

F. Solution of Liouville's equation in the first approximation

We seek a solution of Liouville's equation (7.24) which corresponds to the state of local equilibrium: in the first approximation,

$$f(t,\zeta,h)=\delta(E-H(x(\zeta,h)))\Phi(t,h), \tag{7.34}$$

i.e., for given h, ζ are distributed over energy surface in accordance with "the ergodic thermodynamic equilibrium." Therefore, we seek a solution of Liouville's equation in the form

$$f(t,\zeta,h)=\delta(E-H(x(\zeta,h)))\times[\Phi(t,h)+\Phi'(t,\zeta,h)], \tag{7.35}$$

where $\Phi'\ll\Phi$. Plugging Eq. (7.35) into Eq. (7.24), we have

$$\frac{\partial\Phi'}{\partial t}+v_a^{\alpha}\frac{\partial\Phi'}{\partial\zeta_a^{\alpha}}=-v_a\frac{\partial\Phi}{\partial h_a}-\frac{\partial\overline{\Phi}}{\partial t}-v_a\frac{\partial\Phi'}{\partial h_a}. \tag{7.36}$$

Note that coefficients v_a of Liouville's equation (7.24) are much smaller than coefficients v_a^{α}: this corresponds to a slow change of energies h_a. The further analysis is quite similar to that of Sec. III. We set up for the function Φ the initial data

$$\Phi(0,h)=c_0\delta(h-h_0). \tag{7.37}$$

Then

$$\Phi'(0,\zeta,h)=0. \tag{7.38}$$

Without loss of generality, we may impose the constraint

$$\int\delta(E-H)\Phi'(t,\zeta,h)\Delta d\zeta=0, \tag{7.39}$$

redefining, if necessary, function $\Phi(t,h)$.

We drop the last two terms on the right-hand side of Eq. (7.36) and find Φ' explicitly in terms of solutions of the system of ordinary differential equations

$$\frac{d\zeta_a^{\alpha}(t,\zeta)}{dt}=v_a^{\alpha}(t,\zeta,h).$$

The system defines the mapping

$$\zeta=\chi(t;t',\zeta';h)$$

of the points ζ' at the instant t' to the points ζ at the instant t.

We have

$$\Phi'(t,\zeta,h)=-\int_0^t v_a(\chi(t';t,\zeta;h),h)\frac{\partial\Phi(t',h)}{\partial h_a}dt'. \tag{7.40}$$

From Eqs. (7.35) and (7.32),

$$\begin{aligned}J_a&=\int v_a\Delta f(t,\zeta,h)d\zeta=\int v_a(\zeta,h)\delta(E-H)\Delta d\zeta\Phi(t,h)\\&\quad-\int_0^t\int v_a(\zeta,h)v_a(\chi(t';t,\zeta;h),h)\frac{\partial\Phi(t',h)}{\partial h_b}dt'\\&\quad\times\delta(E-H)\Delta d\zeta.\end{aligned} \tag{7.41}$$

Let us show that the first term on the right-hand side of Eq. (7.41) is zero. Indeed, according to Eq. (7.25), we have

$$\begin{aligned}&\int v_a\delta(E-H)\Delta d\zeta\\&\quad=\int\frac{\partial H_a}{\partial x_a^i}\omega^{ij}\frac{\partial H}{\partial x_a^j}\delta(E-H)\delta(h_1-H_1)\cdots\delta(h_N-H_N)\\&\qquad\times dx_1\cdots dx_N\\&\quad=\int\frac{\partial\theta(h_a-H_a)}{\partial x_a^i}\omega^{ij}\frac{\partial\theta(E-H)}{\partial x_a^j}\\&\qquad\times\prod_{b\neq a}\delta(h_b-H_b)dx_1\cdots dx_N\\&\quad=\int\frac{\partial}{\partial x_a^j}\left(\frac{\partial\theta(h_a-H_a)}{\partial x_a^i}\omega^{ij}\theta(E-H)\right)\\&\qquad\times\prod_{b\neq a}\delta(h_b-H_b)dx_1\cdots dx_N=0\end{aligned} \tag{7.42}$$

Here we used the divergence theorem and the fact that $\theta(E-H)$ is equal to zero for sufficiently large x due to the compactness of the surface $H=E$.

Equation (7.42) means that the average value of v_a over the energy surface in any local equilibrium is zero. This is why the Hamiltonian part does not appear in the evolution equations (7.19).

So, v_a has a character of fluctuations. As before, we introduce the correlation tensor

$$K_{ab}(t,t';h)=\int v_a(\zeta,h)v_b(\chi(t';t,\zeta),h)\,\delta(E-H)\times\Delta d\zeta/\Gamma_{Eh}\,. \tag{7.43}$$

For $t'\leqslant t$, K_{ab} depend only on the difference $t-t'$. We assume that $K_{ab}\neq 0$ only for $|t'-t|\leqslant\tau$, and that $\partial\Phi/\partial h_b$ do not change considerably on the time intervals in the order of the correlation time τ. Then

$$J_a=-D_{ab}(h)\frac{\partial\Phi}{\partial h_b}\Gamma_{Eh}\,, \tag{7.44}$$

where

$$D_{ab}(h)=\int_{-\infty}^{t}K_{ab}(t-t',h)dt.$$

The first constraint on D_{ab} (7.20) follows from reversibility of motion in the same way as in Sec. IV. The second constraint (7.20) holds due to conservation of total energy.

Now note that the probability distribution of $h, f(t,h)$, is linked to $\Phi(t,h)$, according to Eqs. (7.29), (7.35), and (7.10), by the relation

$$f(t,h)=\Phi(t,h)\Gamma_{Eh}\,. \tag{7.45}$$

From Eqs. (7.44) and (7.45),

$$J_a=-D_{ab}\frac{\partial f(t,h)}{\partial h_b}+D_{ab}\frac{\partial\ln\Gamma_{Eh}}{\partial h_b}. \tag{7.46}$$

Equation (7.19) follows from Eqs. (7.46), (7.14), and (7.20) and the smallness of the diffusion coefficients. The magnitude of the neglected terms in Eq. (7.36) may be estimated in the same way as for the canonical variables in Appendix B.

G. Transition to continuum

Let the subsystems be small pieces of some body. We identify the number of the piece, *a*, with the coordinates of the center of that piece, ξ. Evolution to equilibrium is described by the field $U(t,\xi)$ while $h(t,\xi)=U(t,\xi)d^3\xi$, $S=\eta d^3\xi$, and D_{ab} becomes a function of two variables $D(\xi,\xi')$. The function $D(\xi,\xi')$ is, in fact, a functional of the field $h(t,\xi)$ [or, equivalently, $T(t,\xi)$] but we do not emphasize this in the notation. We assume that $D(\xi,\xi')=\hat{D}(\xi,\xi')d^3\xi d^3\xi'$. Then Eq. (7.19) transforms to the equation

$$\frac{\partial U(t,\xi)}{\partial t}=\int\hat{D}(\xi,\xi')\frac{1}{T(t,\xi')}d^3\xi'. \tag{7.47}$$

The second law requires positiveness of the kernel: for any function $\varphi(\xi)$,

$$\int\hat{D}(\xi,\xi')\varphi(\xi)\varphi(\xi')d^3\xi d^3\xi'\geqslant 0. \tag{7.48}$$

Conditions (7.20) take the form

$$\hat{D}(\xi,\xi')=\hat{D}(\xi',\xi),\qquad\int\hat{D}(\xi,\xi')d^3\xi'=0. \tag{7.49}$$

For simplicity, we assume the body is unbounded. Therefore, the integrals in Eqs. (7.47)–(7.49) are taken over the entire three-dimensional space. Interactions of remote parts of the body decay with distance, and $\hat{D}(\xi,\xi')\to 0$ as $|\xi-\xi'|\to\infty$. The rate of decay is assumed to be fast enough for the convergence of all integrals involved.

We are going to show now that the kernel has the form

$$\hat{D}(\xi,\xi')=\frac{\partial^2}{\partial\xi^\mu\partial\xi'^\nu}D^{\mu\nu}(\xi,\xi'), \tag{7.50}$$

where $D^{\mu\nu}(\xi,\xi')$ have the symmetry

$$D^{\mu\nu}(\xi,\xi')=D^{\nu\mu}(\xi',\xi). \tag{7.51}$$

Indeed, consider the Fourier transform of the kernel

$$D(k,k')=\int\hat{D}(\xi,\xi')e^{i(k\xi+k'\xi')}d^3\xi d^3\xi'. \tag{7.52}$$

According to the first relation (7.49), $D(k,k')$ is a symmetric function

$$D(k,k')=D(k',k). \tag{7.53}$$

Putting $k'=0$ in Eq. (7.52), we find using the second relation (7.49)

$$D(k,0)=0. \tag{7.54}$$

Due to the symmetry of $D(k,k')$, we have also

$$D(0,k')=0. \tag{7.55}$$

Assume that $D(k,k')$ is an analytic function of k and k' at the point $k=k'=0$. Then, from Eq. (7.55), $D(k,k')$ can always be presented in the form

$$D(k,k')=k_\mu D^\mu(k,k'),$$

where $D^\mu(k,k')$ are some analytic functions of k and k' (see, for example, the lemma in the proof of Morse's theorem [21]). According to Eq. (7.54), $D^\mu(k,0)=0$. For the same reason, $D^\mu(k,k')$ can be presented in the form $D^\mu(k,k')=D^{\mu\nu}(k,k')k'_\nu$ and

$$D(k,k')=D^{\mu\nu}(k,k')k_\mu k'_\nu \tag{7.56}$$

with the symmetry $D^{\mu\nu}(k,k')=D^{\nu\mu}(k',k)$.

Plugging Eq. (7.56) in Eq. (7.52) and performing the inversed Fourier transform, we arrive at Eqs. (7.50) and (7.51), where $D^{\mu\nu}(\xi,\xi')$ is the inverse Fourier transform of

$-D^{\mu\nu}(k,k')$. Equation (7.2) follows from Eqs. (7.47) and (7.50). Equation (7.3) is a consequence of Eq. (7.2) and the locality condition.

H. Other nonlinear effects in continua

The above analysis can be extended to include viscosity and other nonlinear effects in continua in the case when there is no mass exchange between the continuum particles (subsystems). The corresponding analysis is cumbersome, however, because the additional integrals of motion, momentum, and momentum momentum, must be taken into account. The final equations are similar to the usual equations of continuum media [3,22] in which Onsager's relations must be understood in their nonlinear form (1.10) and (1.11). Consider as an example the following nonlinear phenomenon. In nonlinear elastic body entropy, density is a function of internal energy U and Lagrangian coordinates of the strain tensor, ε_{ab}, $a,b=1,2,3$. In many cases, one can accept that

$$\eta(U,\varepsilon_{ab})=\eta_0(U)-\frac{1}{2}C^{abcd}(U)\varepsilon_{ab}\varepsilon_{cd},$$

where $\eta_0(U)$ determines the heat capacity of the body at zero strains. Functions $C^{abcd}(U)$ describe the dependence of Young's moduli on temperature. Temperature is determined by the equation

$$\frac{1}{T}=\frac{\partial\eta(U,\varepsilon_{ab})}{\partial U}=\frac{d\eta_0(U)}{dU}-\frac{1}{2}\frac{dC^{abcd}(U)}{dU}\varepsilon_{ab}\varepsilon_{cd}. \tag{7.57}$$

The last term in Eq. (7.57) shows that the dependence of Young's moduli on temperature causes a reciprocal influence of strains on temperature. This is a pure equilibrium effect. The nonlinear heat conduction equation (7.5) means that the heat flux depends on the gradient of temperature (7.57), which must include the gradient of strains [the gradient of the last term in Eq. (7.57)]. Being neglected in linear heat conductivity, the dependence of heat flux on strain gradient might be important for some materials.

VIII. MOTION OF DEFECTS IN A CRYSTAL LATTICE

Probability flux in heat conduction vanishes at equilibrium. There is another interesting case where this occurs: motion of defects in a crystal lattice. Consider first a general setting when the Hamiltonian system possesses some slow variables, Φ^{μ}, which are functions of canonical variables, x: $\Phi^{\mu}=\Phi^{\mu}(x)$. That means that the derivatives

$$\frac{d\Phi^{\mu}}{dt}=\frac{\partial\Phi^{\mu}}{\partial x^i}\omega^{ij}\frac{\partial H}{\partial x^j}$$

are small compared to $\partial H/\partial x^i$. Denote the probability of Φ^{μ} by $f(t,y)$,

$$f(t,y)=\int f(t,x)\,\delta(y-\Phi(x))dx,$$

where $\delta(y-\Phi(x))=\delta(y^1-\Phi^1(x))\cdots\delta(y^m-\Phi^m(x))$. Function $f(t,y)$ obeys the equation

$$\frac{\partial f(t,y)}{\partial t}+\frac{\partial J^{\mu}}{\partial y^{\mu}}=0,$$

$$J^{\mu}=\int\omega^{ij}\frac{\partial H}{\partial x^j}\frac{\partial\Phi^{\mu}}{\partial x^i}\delta(y-\Phi(x))f(t,x)dx, \tag{8.1}$$

which follows from Liouville's equation and finiteness of the support of function $f(t,x)$ [remember that energy surfaces $H(x)=E$ are compact],

$$\begin{aligned}\frac{\partial f(t,y)}{\partial t}&=\int\frac{\partial f(t,x)}{\partial t}\delta(y-\Phi(x))dx\\&=-\int\frac{\partial}{\partial x^i}\left(\omega^{ij}\frac{\partial H}{\partial x^j}f(t,x)\right)\delta(y-\Phi(x))dx\\&=\int\omega^{ij}\frac{\partial H}{\partial x^j}f(t,x)\frac{\partial}{\partial x^i}\delta(y-\Phi(x))dx\\&=-\int\omega^{ij}\frac{\partial H}{\partial x^j}\frac{\partial\Phi}{\partial x^i}f(t,x)\,\delta'(y-\Phi(x))dx\\&=-\frac{\partial}{\partial y^{\mu}}\int\omega^{ij}\frac{\partial H}{\partial x^j}\frac{\partial\Phi}{\partial x^i}f(t,x)\,\delta(y-\Phi(x))dx.\end{aligned}$$

Let the motion of the Hamiltonian system with the constraints $\Phi^{\mu}(x)=y^{\mu}$ be ergodic and mixing on the surfaces $H(x)=E$, $\Phi^{\mu}(x)=y^{\mu}$. Then one may expect that the asymptotics of the $f(t,x)$ has the form

$$f(t,x)=\delta(E-H(x))[\overline{f}(t,\Phi(x))+f'(t,x)]$$

and, from the same chain of reasoning as in Sec. III, one can derive that y^{μ} is a Markov process.

Let function $\Phi^{\mu}(x)$ depend only on generalized coordinates q, $\Phi^{\mu}=\Phi^{\mu}(q)$, and let $H(p,q)$ be an even function of generalized momenta p. Then the probability flux

$$J^{\mu}=\int\frac{\partial H}{\partial p}\frac{\partial\Phi^{\mu}(q)}{\partial q}\delta(y-\Phi(q))f(t,p,q)dp\,dq \tag{8.2}$$

is zero at equilibrium when $f(t,p,q)=c\,\delta(E-H(p,q))$ since the integrand in Eq. (8.2) is an odd function of p. Vanishing of the probability flux along with Markov's property determines the drift.

Now we apply this reasoning to the motion of vacancies. Consider a vacancy in an otherwise perfect crystal lattice. Vacancy coordinates r can be viewed as slow variables. Vacancy coordinates may be thought of as some functions of positions of atoms, q. For example, one can define r as a point where the function $\Sigma_{a=1}^{N}\varphi(|r-q_a|)$ (a is the atom number, φ is a growing function) reaches its maximum. For Markov's process of the vacancy diffusion, the probability flux is

$$J^{\mu}=V^{\mu}f(t,r)-D^{\mu\nu}\frac{\partial f(t,r)}{\partial r^{\nu}}. \tag{8.3}$$

Equilibrium distribution of $f(t,r)$ is

$$f(\infty,r)=ce^{-H(r)/T}, \tag{8.4}$$

where $H(r)$ is the energy value when the vacancy is positioned at the point r. Since the probability flux vanishes at equilibrium, we find the drift from Eqs. (8.3) and (8.4),

$$V^{\mu}=-\frac{D^{\mu\nu}}{T}\frac{\partial H(r)}{\partial r^{\nu}}. \tag{8.5}$$

Perhaps similar reasoning can be applied to the slow motion of dislocations. One may speculate that the dislocation position depends only on the positions of atoms (though to write down this dependence explicitly is not a simple task). Therefore, the probability flux must be zero.

The dislocation line can be discretized and approximated by a set of vectors, $\mathbf{r}_a=(r_a^i)$, $i,j=1,2,3$, $a=1,\ldots,N$, $\mathbf{r}=(\mathbf{r}_1,\ldots,\mathbf{r}_N)$. Then

$$V_a^i=-\sum_b\frac{D_{ab}^{ij}}{T}\frac{\partial H}{\partial r_b^j}.$$

Extrapolating the set of vectors $\mathbf{r}_a(t)$ by a continuum curve $\mathbf{r}=\mathbf{r}(t,\sigma)$, we have

$$V^i(\sigma)=-\int\frac{D^{ij}(\sigma,\sigma')}{T}\frac{\delta H}{\delta r^j(t,\sigma')}d\sigma',$$

where $H(\mathbf{r}(t,\sigma))$ is the energy functional of the dislocation position $\mathbf{r}(t,\sigma)$, and $\delta H/\delta r^j$ is its variational derivative. Obviously, $D^{ij}(\sigma,\sigma')=D^{ji}(\sigma',\sigma)$ and $D^{ij}(\sigma,\sigma')dr^j/d\sigma=0$. The equation of dislocation motion is

$$\frac{dr^i(t,\sigma)}{dt}=-\int\frac{D^{ij}(\sigma,\sigma')}{T}\frac{\delta H}{\delta r^j(\sigma')}d\sigma'.$$

If $D^{ij}(\sigma,\sigma')$ can be approximated by a δ function, $D^{ij}(\sigma,\sigma')=\delta(\sigma-\sigma')D^{ij}$, the dynamical equation of dislocation takes the form

$$\frac{dr^i(t,\sigma)}{dt}=-\frac{D^{ij}}{T}\frac{\delta H}{\delta r^j(t,\sigma)}.$$

This equation must be valid if the dislocation velocity is much less than the characteristic velocity of micromotion, i.e., the speed of sound.

IX. FEATURES OF NONLINEAR ONSAGER'S RELATIONS

In this section, we will argue that the nonlinear version of Onsager's relations (1.10) and (1.11) does not put constraints on the functional dependence of G^{μ} on F_{μ} beyond the point $F_{\mu}=0$. If G^{μ} are analytical functions of F_{μ}, then these constraints are just Onsager's relations

$$D^{\mu\nu}(0)=D^{\nu\mu}(0).$$

If G^{μ} have a singularity at $F_{\lambda}=0$, for example, as in plasticity theory, $D^{\mu\nu}(F_{\lambda})=d^{\mu\nu}(F_{\lambda})/\sqrt{F_{\sigma}F_{\sigma}}$, then Eq. (1.11) yields the constraint

$$d^{\mu\nu}(0)=d^{\nu\mu}(0).$$

However, as follows from further consideration, Eqs. (1.10) and (1.11) do not seem to be constraining the functional dependence of G^{μ} on F_{μ} beyond the point $F_{\mu}=0$.

Such a conclusion leads to a natural question: Why should one bother about $m(m+1)/2$ functions $D^{\mu\nu}(F_{\lambda})$ if the equations contain only m functions $G^{\mu}(F_{\lambda})$? To close the system of equations, one can prescribe m functions $G^{\mu}(F_{\lambda})$ which possess the following two properties: $\partial G^{\mu}/\partial F_{\lambda}|_{F_{\lambda}=0}$ is a symmetric tensor and $G^{\mu}(F_{\lambda})F_{\mu}\geqslant 0$. There is an important reason, however, to write the dependence G^{μ} on F_{λ} in the form (1.10) and (1.11) with a provision that the closing of the system of equations assumes prescribing $m(m+1)/2$ functions $D^{\mu\nu}(F_{\lambda})$. As was shown by Kubo in the linear case and as we have seen in Secs. II and III in the nonlinear case, the dissipation coefficients $D^{\mu\nu}$ have an additional peculiarity: they characterize the correlations of fluctuations of thermodynamic fluxes. Thus, the dissipation coefficients can be determined, in principle, from the experiments which are independent of the direct measurements of G^{μ} and F_{μ}. We do not know all the nonequilibrium properties of the system if we do not know the dissipation coefficients. In mathematical modeling, the dissipation coefficients should be prescribed along with the thermodynamic functions characterizing the equilibrium properties of the system.

After these general comments, we proceed to the derivation of the statement made at the beginning of this section.

Let $G^{\mu}(F_{\lambda})$ be analytical functions at the point $F_{\lambda}=0$ vanishing at this point. We expand G_{μ} in Taylor's series in a vicinity of zero,

$$G^{\mu}=G_0^{\mu\nu}F_{\nu}+G^{\mu\nu\lambda}F_{\nu}F_{\lambda}+G^{\mu\nu\lambda\kappa}F_{\nu}F_{\lambda}F_{\kappa}+\cdots. \tag{9.1}$$

Tensor $G^{\mu\nu\lambda}$ is, obviously, symmetric over indices ν,λ, tensor $G^{\mu\nu\lambda\kappa}$ over indices $\nu\lambda\kappa$, etc. We can rewrite the expansion (9.1) in the form

$$G^{\mu}=G^{\mu\nu}(F_{\lambda})F_{\nu},$$

$$G^{\mu\nu}(F_{\lambda})=G_0^{\mu\nu}+G^{\mu\nu\lambda}F_{\lambda}+G^{\mu\nu\lambda\kappa}F_{\lambda}F_{\kappa}+\cdots. \tag{9.2}$$

Tensor $G^{\mu\nu}(F_{\lambda})$ is not necessarily symmetric. The question is, could one find a symmetric tensor $D^{\mu\nu}(F_{\lambda})=D^{\nu\mu}(F_{\lambda})$ such that the expansion (9.1) can be written as

$$G^{\mu}=D^{\mu\nu}(F_{\lambda})F_{\nu},$$

$$D^{\mu\nu}(F_{\lambda})=D_0^{\mu\nu}+D^{\mu\nu\lambda}F_{\lambda}+\cdots \tag{9.3}$$

with some symmetric over $\mu\nu$ tensors $D_0^{\mu\nu},D^{\mu\nu\lambda},\ldots$. We are going to show that this is possible.

Let $G^{\mu\nu}(F_\lambda)$ be given. We are going to find $D^{\mu\nu}(F_\lambda)$. Compare the terms of the same power in Eqs. (9.2) and (9.3). Obviously, $D_0^{\mu\nu}=G_0^{\mu\nu}$. Equate now the quadratic terms of both expansions,

$$G^{\mu\nu\lambda}F_\nu F_\lambda = D^{\mu\nu\lambda}F_\nu F_\lambda . \tag{9.4}$$

Tensor $D^{\mu\nu\lambda}$ is symmetric with respect to μ,ν but not necessarily symmetric over $\nu\lambda$. Tensor $G^{\mu\nu\lambda}$ is symmetric with respect to $\nu\lambda$ and not necessarily symmetric over $\mu\nu$. From Eq. (9.4),

$$\frac{1}{2}(D^{\mu\nu\lambda}+D^{\mu\lambda\nu})=G^{\mu\nu\lambda}. \tag{9.5}$$

For a given tensor $G^{\mu\nu\lambda}$, Eqs. (9.5) may be considered as a system of linear equations for $D^{\mu\nu\lambda}$. This system admits an explicit unique solution

$$D^{\mu\nu\lambda}=\frac{1}{2}(G^{\mu\nu\lambda}+G^{\nu\mu\lambda}-G^{\lambda\mu\nu}). \tag{9.6}$$

Formula (9.6) can be checked by plugging Eq. (9.6) in Eq. (9.5).

So, if $G^{\mu}(F_\lambda)$ is a quadratic function, it can be presented in the form (9.3).

Consider now the terms of some power s, $s>2$. The coefficients $D^{\mu\nu_1\cdots\nu_s}$ and $G^{\mu\nu_1\cdots s}$ are linked by the equations

$$\frac{1}{s}(D^{\mu\nu_1\cdots\nu_s}+D^{\mu\nu_2\cdots\nu_s\nu_1}+\cdots+D^{\mu\nu_s\nu_1\cdots\nu_{s-1}}) = G^{\mu\nu_1\cdots\nu_s}. \tag{9.7}$$

We consider Eq. (9.7) as a system of linear algebraic equations with respect to $D^{\mu\nu_1\cdots\nu_s}$ assuming that $G^{\mu\nu_1\cdots\nu_s}$ are known. Let us compute the number of equations (9.7) identifying the equations obtained by permutation of indices $\nu_1,\ldots,\nu_s$. The number of equations is equal to the number of independent components of an $(s+1)$-rank tensor which is symmetric over s indices. This is equal to the space dimension m times the number of independent component of a symmetric tensor of rank s. The latter is C^s_{m+s-1} [29]. Thus, the number of equations is

$$mC^s_{m+s-1}=\frac{m(m+s-1)!}{s!(m-1)!}.$$

The number of unknowns is equal to the number of independent components of a symmetric tensor of second order times the number of independent components of a symmetric tensor of rank $s-1$,

$$\frac{m(m+1)}{2}C^{s-1}_{m+s-2}=\frac{m(m+1)(m+s-2)!}{2(s-1)!(m-1)!}.$$

The difference between the number of unknowns and the number of equations is

$$\begin{aligned}&\frac{m(m+1)(m+s-2)!}{2(s-1)!(m-1)!}-\frac{m(m+s-1)!}{s!(m-1)}\\ &=\frac{m(m+s-1)!}{s!(m-1)!}\left(\frac{(m+1)s}{2(m+s-1)}-1\right)\\ &=\frac{m(m+s-1)!}{s!(m-1)!}\,\frac{(m-1)(s-2)}{2(m-1+s)}\\ &=\frac{(s-2)m(m-2+s)!}{2(m-2)!s!}.\end{aligned}$$

It is seen that the number of unknowns is equal to the number of equations for $s=2$ (and, as we have seen, there is a unique solution in this case), but for $s>2$ and $m\geqslant 2$, the number of unknowns is larger than the number of equations.

The only obstacle for the solvability of the system of Eqs. (9.7) might be the linear dependence of the left-hand sides of Eqs. (9.7). Let us show that they are, in fact, linearly independent. Assume the opposite: there is some nonzero tensor $a_{\mu\nu_1\cdots\nu_s}$ symmetric over $\nu_1\cdots\nu_s$ such that for any $D^{\mu\nu_1\cdots\nu_s}$,

$$a_{\mu\nu_1\cdots\nu_s}(D^{\mu\nu_1\cdots\nu_s}+\cdots+D^{\mu\nu_s\nu_1\cdots\nu_{s-1}})=0. \tag{9.8}$$

Equation (9.8) reduces to

$$a_{\mu\nu_1\cdots\nu_s}D^{\mu\nu_1\cdots\nu_s}=0.$$

The arbitrariness of $D^{\mu\nu_1\cdots\nu_s}$ and their symmetry over the first two indices yield the system of linear equations for $a_{\mu\nu_1\cdots\nu_s}$,

$$a_{\mu\nu_1\nu_2\cdots\nu_s}+a_{\nu_1\mu\nu_2\cdots\nu s}=0. \tag{9.9}$$

This system has only a zero solution. Indeed, changing indices in Eq. (9.9), one can also write

$$a_{\nu_2\mu\nu_1\nu_3\cdots\nu_s}+a_{\mu\nu_2\nu_1\nu_3\cdots\nu s}=0 \tag{9.10}$$

and

$$a_{\nu_2\nu_1\mu\nu_3\cdots\nu s}+a_{\nu_1\nu_2\mu\nu_3\cdots\nu_s}=0. \tag{9.11}$$

Using Eqs. (9.9) and (9.11) and deducting Eq. (9.10), we obtain

$$2a_{\nu_1\mu\nu_2\cdots\nu s}=0.$$

Thus, the left-hand sides of Eq. (9.7) are linearly independent and Eqs. (9.7) are solvable.

So, without loss of generality, $D^{\mu\nu}(F_\lambda)$ can be chosen symmetric, as claimed.

X. SECONDARY THERMODYNAMICS (THERMODYNAMICS OF ATTRACTORS)

The special structure of macrophysical equations pertains to the "first level of averaging" when the system has only two characteristic time scales. In many cases, there are "intermediate-fast variables" between the macrolevel and the microlevel. In such cases, the quasi-Hamiltonian structure is characteristic for the equations governing the intermediate-fast and slow variables. Equations for slow variables are obtained by elimination of the intermediate-fast variables from the quasi-Hamiltonian equations. The corresponding macrotheory may be called secondary thermodynamics or thermodynamics of attractors since the existence of attractors is a characteristic feature of the quasi-Hamiltonian equations. The examples of the intermediate-fast variables are the coordinates of defects in solids (vacancies, dislocations, etc.). The entire realm of plasticity is a subject of secondary thermodynamics, as well as turbulence theory.

There might be many "thermodynamics." The number of thermodynamics is equal to the number of the well-separated time scales. Even the secondary thermodynamics is at embrionic stage. Some statements of secondary thermodynamics can be found in the book [10]. In particular, it was shown [10,23] that the relations of secondary thermodynamics of vibrating systems are potential in the limit of vanishing friction. Perhaps, similar facts hold for turbulent flows with high Reynolds number. Note also Ruelle's paper [24] on attractor's response to slow excitation.

APPENDIX A: A GENERAL FORM OF MICRODYNAMIC EQUATIONS CONSISTENT WITH EQUILIBRIUM THERMODYNAMICS

The term widely used in this paper, the Hamiltonian structure, needs to be made more precice because any system of differential equations can be written in a Hamiltonian form [28]

$$\frac{dq^i}{dt}=\frac{\partial H(p,q)}{\partial p^i}, \quad \frac{dp^i}{dt}=-\frac{\partial H(p,q)}{\partial q^i}. \tag{A1}$$

Indeed, let us have a system of ordinary differential equations

$$\frac{dq^i}{dt}=Q^i(q). \tag{A2}$$

One can embed it in a Hamiltonian system of equations by introducing the additional variables $p_i(t)$ and determining p_i from the equations

$$\frac{dp_i}{dt}=-p_k\frac{\partial Q^k(q)}{\partial q^i}. \tag{A3}$$

Equations (A2) and (A3) form a Hamiltonian system with the Hamiltonian

$$H(p,q)=p_kQ^k(q). \tag{A4}$$

So, we have to rectify what is understood under the term "Hamiltonian structure."

Modeling of any system begins with a description of its kinematics. Therefore, the starting point in the description of a mechanical system is a choice of the phase space. After this choice is made, one has to specify the interactions in the systems, i.e., to define the Hamiltonian $H(p,q)$. To be consistent with equilibrium thermodynamics, the Hamiltonian must possess the following two properties: energy surfaces, $H(p,q)=\text{const}$, bound compact regions with finite volumes in the phase space; and motion is ergodic on the energy surfaces. These properties eliminate a pathological construction (A2)–(A4) [obviously, the energy surfaces of the Hamiltonian (A4) do not bound the regions with finite volumes].

The above-mentioned features of microdynamics are key to deriving equilibrium thermodynamics from the underlying Hamiltonian mechanics. One may ask to what extent these features define microdynamics, or, in other words, do microdynamic equations need to possess the Hamiltonian structure in order to be consistent with equilibrium thermodynamics?

We describe in this appendix, under natural assumptions, the microdynamic equations consistent with equilibrium thermodynamics. They turn out to be slightly more general than the standard Hamiltonian equations.

Let the system be described by fast microvariables $x=(x^1,\dots,x^n)$ and slow macrovariables $y=(y^1,\dots,y^m)$. We consider the driven systems when y variables are changed in a prescribed way, $y(t)$. Microdynamics is governed by a system of ordinary differential equations,

$$\frac{dx^i}{dt}=F^i(x,y). \tag{A5}$$

We make the following assumptions.

(i) The x-phase space is split in a one-parametric set of hypersurfaces described by the equation $H(x,y)=\text{const}$ with a smooth function $H(x,y)$. For any fixed y, a trajectory starting on a surface $H(x,y)=\text{const}$ remains on the surface all the time, i.e., $H(x,y)$ is an integral of motion,

$$\frac{\partial H(x,y)}{\partial x^i}F^i(x,y)=0. \tag{A6}$$

Obviously, any vector F^i satisfying Eq. (A6) can be presented in a form

$$F^i(x,y)=\omega^{ij}(x,y)\frac{\partial H(x,y)}{\partial x^j}, \tag{A7}$$

where $\omega^{ij}(x,y)$ is an antisymmetric tensor field.

(ii) For each fixed y motion is ergodic on the surfaces $H(x,y)=\text{const}$.

(iii) Tensor ω^{ij} in Eq. (A7) is nondegenerate,

$$\omega=\det\|\omega^{ij}\|\neq 0. \tag{A8}$$

An immediate consequence of Eq. (A8) is evenness of the dimension of the x-phase space (remember that any antisymmetric matrix of an odd order has zero determinant).

(iv) Vector field $F^i(x,y)$ is incompressible. Note that we cannot write the incompressibility condition simply as

$$\frac{\partial F^i(x,y)}{\partial x^i}=0 \tag{A9}$$

because Equation (A9) is not invariant under coordinate transformations. To put Eq. (A9) in an invariant form, one has to include F^i in Eq. (A9) with a factor which is a relative scalar of the weight 1/2. The only scalar of this type which one can form from the above entries is $1/\sqrt{\omega}$. Thus, we specify the assumption D by putting

$$\frac{\partial}{\partial x^i}\frac{1}{\sqrt{\omega}}F^i=0. \tag{A10}$$

(v) Tensor ω^{ij} characterizes the geometry of the x-phase space and does not depend on specific interactions in the system. In other words, ω^{ij} depend only on x and do not depend on the choice of Hamiltonian. Then, from Eqs. (A7) and (A10), the incompressibility constraint is a constraint on $\omega^{ij}(x)$ only,

$$\frac{\partial}{\partial x^i}\left(\frac{1}{\sqrt{\omega}}\omega^{ij}(x)\right)=0. \tag{A11}$$

The systems possessing the properties (i)–(v) can be written in the form

$$\frac{dx^i}{dt}=\omega^{ij}(x)\frac{\partial H(x,y)}{\partial x^j}, \tag{A12}$$

where tensor ω^{ij} satisfies Eq. (A11).

The system (A12) is of a more general form than that describing Hamiltonian flow on a symplectic manifold [30]. The latter has the form (A12) with a provision that the 2-form $\Omega_{ij}dx^i\wedge dx^j$, Ω_{ij} being the inverse tensor for ω^{ij}, must be closed, or, equivalently, there is a covector field $P_i(x)$ such that

$$\Omega_{ij}(x)=\frac{\partial P_i(x)}{\partial x^j}-\frac{\partial P_j(x)}{\partial x^i}. \tag{A13}$$

Hamiltonian flows on symplectic manifolds obey the incompressibility condition (A11). Indeed,

$$\begin{aligned}
\frac{\partial}{\partial x^i}\left(\frac{1}{\sqrt{\omega}}\omega^{ij}\right)&=\frac{1}{\sqrt{\omega}}\frac{\partial\omega^{ij}}{\partial\Omega_{kl}}\frac{\partial\Omega_{kl}}{\partial x^i}-\frac{1}{2\,\omega^{3/2}}\omega^{ij}\frac{\partial\omega}{\partial\omega^{kl}}\frac{\partial\omega^{kl}}{\partial\Omega_{mn}}\frac{\partial\Omega_{mn}}{\partial x^i}\\
&=-\frac{1}{\sqrt{\omega}}\omega^{il}\omega^{kj}\left(\frac{\partial^2P_k}{\partial x^l\partial x^i}-\frac{\partial^2P_l}{\partial x^k\partial x^i}\right)-\frac{1}{2\sqrt{\omega}}\omega^{ij}\Omega_{kl}(-\omega^{kn}\omega^{ml})\left(\frac{\partial^2P_m}{\partial x^n\partial x^i}-\frac{\partial^2P^n}{\partial x^m\partial x^i}\right)\\
&=\frac{1}{\sqrt{\omega}}\omega^{kj}\omega^{il}\frac{\partial^2P_l}{\partial x^i\partial x^k}+\frac{1}{2\sqrt{\omega}}\omega^{ij}\omega^{mn}\left(\frac{\partial^2P_m}{\partial x^n\partial x^i}-\frac{\partial^2P_n}{\partial x^m\partial x^i}\right)\\
&=0.
\end{aligned}$$

Here we used the definition of the inverse tensor, $\omega^{ij}\Omega_{kj}=\delta^i_k$, and its consequences: $\omega^{ij}\Omega_{ik}=\delta^j_k$, $\partial\omega/\partial\omega^{ij}=\omega\Omega_{ij}$, and $\partial\omega^{ij}/\partial\Omega_{mn}=-\omega^{in}\omega^{mj}$.

The inverse statement is not true: incompressibility condition (A11) does not necessarily yield closedness of the form $\Omega_{ij}dx^i\wedge dx^j$ since it imposes only n constraints on $n(n-1)/2$ functions $\omega^{ij}(x)$. This number of constraints does not seem enough to reduce $n(n-1)/2$ functions $\omega^{ij}(x)$ to n functions $P_i(x)$. Therefore, the systems (A11) and (A12) form a wider class than the class of Hamiltonian flows on symplectic manifolds. Note that Hamiltonian systems on symplectic manifolds can be put locally in the standard form (2.16) and (2.18) by a coordinate transformation, because any differential form $P_1dx^1+\cdots+P_ndx^n$ (n iseven, $n=2k$) can be transformed locally to $p_1dq^1+\cdots+p_kdq^k$ by a change of coordinates.

We are going to show that Eqs. (A12) yield the classical relations of equilibrium thermodynamics.

Consider the stationary Liouville equation

$$\frac{\partial}{\partial x^i}\left(\omega^{ij}(x)\frac{\partial H}{\partial x^j}\rho\right)=0.$$

Due to Eq. (A11), it can be written also as

$$\omega^{ij}\frac{\partial H}{\partial x^j}\frac{\partial}{\partial x^i}(\rho\sqrt{\omega})=0. \tag{A14}$$

Therefore, it has a solution $\rho=1/\sqrt{\omega}$. Introducing, as in Sec. III, some curvilinear coordinates in the x-phase space, $\{\zeta^\alpha,h\}$, $h=H(x,y)$, $\zeta^\alpha=\zeta^\alpha(x)$, and denoting by Δ the Jacobian of transformation $\zeta,h\rightarrow x$: $dx=\Delta\alpha\zeta dh$, we get the invariant measure on the energy surfaces $\rho\Delta d\zeta$.

For any function $\varphi(x)$, the time average over the trajectory,

$$\langle\varphi\rangle=\lim_{\theta\to\infty}\frac{1}{\theta}\int_0^\theta\varphi(x(t))dt,$$

can be computed by means of integration over the energy surface,

$$\langle \varphi \rangle = \frac{\int \varphi \rho \Delta d\zeta}{\int \rho \Delta d\zeta}. \tag{A15}$$

Equation (A15) can be written also in the form

$$\langle \varphi \rangle = \int \varphi \rho \delta(E - H(x,y)) dx \bigg/ \int \rho \delta(E - H(x,y)) dx. \tag{A16}$$

The denominator in Eq. (A16) can be expressed in terms of an invariant phase volume,

$$\hat{\Gamma}(E,y) = \int \rho \theta(E - H(x,y)) dx.$$

We have

$$\int \rho \delta(E - H(x,y)) dx = \frac{\partial \hat{\Gamma}(E,y)}{\partial E} \equiv \hat{\Gamma}_E.$$

Let us introduce also the "standard" phase volume

$$\Gamma(E,y) = \int \theta(E - H(x,y)) dx$$

and the "average value" of ρ

$$\hat{\rho}(E,y) = \int \rho \theta(E - H(x,y)) dx \bigg/ \int \theta(E - H(x,y)) dx = \frac{\hat{\Gamma}(E,y)}{\Gamma(E,y)}.$$

Then the following relation holds:

$$\left\langle \frac{\hat{\rho}}{\rho} x^i \frac{\partial H}{\partial x^j} \right\rangle = \delta^i_j \frac{\hat{\Gamma}}{\hat{\Gamma}_E}. \tag{A17}$$

Indeed,

$$\begin{aligned} \left\langle \frac{1}{\rho} x^i \frac{\partial H}{\partial x^j} \right\rangle &= \int x^i \frac{\partial H}{\partial x^j} \delta(E - H(x,y)) dx \bigg/ \hat{\Gamma}_E \\ &= -\int x^i \frac{\partial \theta(E - H(x,y))}{\partial x^j} dx \bigg/ \hat{\Gamma}_E \\ &= \int \frac{\partial x^i}{\partial x^j} \theta(E - H(x,y)) dx \bigg/ \hat{\Gamma}_E \\ &= \delta^i_j \frac{\Gamma}{\hat{\Gamma}_E}. \end{aligned} \tag{A18}$$

Multiplying Eq. (A18) by $\hat{\rho}$, we get Eq. (A17).

We interpret Eq. (A17) as an equipartition law, and thus we introduce temperature by the formula

$$T = \frac{\hat{\Gamma}}{\hat{\Gamma}_E}. \tag{A19}$$

This formula is consistent with the thermodynamic relation

$$\frac{1}{T} = \frac{\partial S(E,y)}{\partial E} \tag{A20}$$

if we introduce entropy as a logarithm of the invariant phase volume,

$$S(E,y) = \ln \hat{\Gamma}(E,y), \tag{A21}$$

and identify the value of $H(x,y)$ with energy E.

Then the thermodynamic constitutive equation for the forces, $\langle \partial H / \partial y^\mu \rangle$, is also true,

$$\left\langle \frac{\partial H}{\partial y^\mu} \right\rangle = -T \frac{\partial S(E,y)}{\partial y^\mu}. \tag{A22}$$

It follows from Eqs. (A16), (A20), and (A21) that

$$\begin{aligned} \left\langle \frac{\partial H}{\partial y^\mu} \right\rangle &= \int \rho \frac{\partial H}{\partial y^\mu} \delta(E - H(x,y)) dx / \hat{\Gamma}_E \\ &= -\int \rho \frac{\partial \theta(E - H(x,y))}{\partial y^\mu} dx \bigg/ \hat{\Gamma}_E \\ &= -\frac{\partial \hat{\Gamma}}{\partial y^\mu} \bigg/ \frac{\partial \hat{\Gamma}}{\partial E} \\ &= -T \frac{\partial S}{\partial y^\mu}. \end{aligned}$$

It remains to show that S is an adiabatic invariant of the system. Consider the energy equation

$$\frac{dE}{dt} = \frac{\partial H}{\partial y^\mu} \frac{\partial y^\mu}{dt}. \tag{A23}$$

Energy is a slow variable because dy^μ/dt are small. Averaging Eq. (A23) over time, we have

$$\frac{dE}{dt} = \left\langle \frac{\partial H}{\partial y^\mu} \right\rangle \frac{\partial y^\mu}{dt}. \tag{A24}$$

From Eqs. (A22) and (A24),

$$\frac{dE}{dt} = -\frac{1}{\partial S / \partial E} \frac{\partial S}{\partial y^\mu} \frac{\partial y^\mu}{dt}$$

or

$$\frac{d}{dt} S(E,y) = 0$$

as claimed. This completes the justification of the statements made.

APPENDIX B: SOME ESTIMATES

System (3.32) is ergodic by our assumption. Therefore, for any function $\varphi(\zeta)$ and almost all ζ,

$$\frac{1}{t}\int_0^t \varphi(\chi(\tilde{t};0,\zeta))d\tilde{t} \xrightarrow[t\to\infty]{} \langle\varphi\rangle = \int \varphi(\zeta)\frac{d\zeta}{\Gamma_E},$$

$$\frac{1}{|t'|}\int_{t'}^0 \varphi(\chi(\tilde{t};0,\zeta))d\tilde{t} \xrightarrow[t'\to-\infty]{} \langle\varphi\rangle = \int \varphi(\zeta)\frac{d\zeta}{\Gamma_E}.$$

If $\langle\varphi\rangle = 0$, then the integral

$$I(\varphi) = \int_0^t \varphi(\chi(\tilde{t};0,\zeta))d\tilde{t}$$

may tend to infinity as $t\to\infty$. It must grow, however, more slowly than t, so that the ratio $I(\varphi)/t$ tends to zero.

We need this integral, in fact, to go to zero as $t\to\infty$ at least for functions $\varphi = H'_{,\nu}$ and for most initial points. The assumption of mixing brings us closer to such a feature: for a mixing system, the following property holds [25]: for any square integrable functions $\varphi(\zeta)$ and $\psi(\zeta)$,

$$\int \varphi(\chi(t;0,\zeta))\psi(\zeta)\frac{d\zeta}{\Gamma_E} \xrightarrow[\text{as } t\to\pm\infty]{} \int \varphi(\zeta)\frac{d\zeta}{\Gamma_E}\int \psi(\zeta)\frac{d\zeta}{\Gamma_E}.$$

If average values of φ and ψ are zero, then

$$I_{\varphi\psi}(t) = \int \varphi(\chi(t;0,\zeta))\psi(\zeta)\frac{d\zeta}{\Gamma_E} \to 0 \quad \text{as } t\to\pm\infty.$$

We accept a stricter condition of mixing,

$$\int_0^\infty |I_{\varphi\psi}(t)|dt \quad \text{is finite.}$$

This allows one to introduce a characteristic time of mixing of functions φ and ψ, for example, as

$$\tau_{\varphi\psi} = \int_0^\infty |I_{\varphi\psi}|dt \Big/ \sqrt{\int \varphi^2(\zeta)\frac{d\zeta}{\Gamma_E}\int \psi^2(\zeta)\frac{d\zeta}{\Gamma_E}}.$$

To obtain estimates justifying our approximations, the following two conditions are, perhaps, sufficient.

(i) There is a finite correlation time

$$\tau = \sup_{\varphi,\psi\in L_2}\tau_{\varphi\psi}.$$

(ii) The correlation time τ is much smaller than the characteristic time of change of $\Phi(t,y)$ in t.

We give here just some rough estimates. Let θ and Δy be characteristic scales for function $\Phi(t,y)$ with respect to t and y, respectively, and $H_{,\nu} \sim a$. Then from Eq. (3.34), $\Phi' \sim \tau a\,\partial\Phi(t,y)/\partial y$, and the term neglected in Eq. (3.28) has the order

$$a^2\tau\frac{1}{\Delta y}\frac{\partial\Phi(t,y)}{\partial y} + \frac{1}{\Delta y}a^2\tau\frac{\partial\Phi(t,y)}{\partial y}.$$

This is much smaller than the kept term [which is of order $a\,\partial\Phi(t,y)/\partial y$] if

$$\tau a \ll \Delta y.$$

The function $f(t,y)$ [and, thus, $\Phi(t,y)$], as is seen from Fokker-Planck's equation, changes on scales Δy and θ such that

$$\frac{\Delta y}{\theta} \sim a^2\tau.$$

Thus, our approximation is valid if

$$\frac{\tau}{\theta} \ll 1.$$

---

[1] R. Clausius, *Die Mechanische Wärmtheorie*, 2^{te} Aufgabe, Braunschweig, Bd. I-III, 1876–1891.

[2] L. Onsager, Phys. Rev. **37**, 405 (1931); **38**, 2265 (1931).

[3] S. R. de Groot and P. Mazur, *Non-equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962).

[4] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).

[5] R. Kubo, Lect. Theor. Phys. **1**, 120 (1959).

[6] R. L. Stratanovich, *Nonlinear Nonequilibrium Thermodynamics* (Springer, Berlin, 1992), Vols. 1 and 2.

[7] P. Mazur, in *The Collected Works of L. Onsager*, edited by P.C. Hemmer, H. Holden, and S.K. Ratkje (World Scientific, Singapore, 1996), p. 69.

[8] A.N. Kolmogorov, Math. Ann. **113**, 766 (1937).

[9] P. Hertz, Ann. Phys. (Leipzig) **33**, 225 (1910); **33**, 537 (1910).

[10] V. Berdichevsky, *Thermodynamics of Chaos and Order* (Addison-Wesley Longman, 1997).

[11] J.L. Leibowitz, Physica A **263**, 516 (1999).

[12] R. Zwanzig, Phys. Rev. **124**, 983 (1961).

[13] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).

[14] A.M. Yaglom, Dokl. Akad. Nauk SSSR **56**, 347 (1947); Mat. Sb. **24**, 457 (1949).

[15] V. Berdichevsky, Int. J. Eng. Sci. **28**, 697 (1990).

[16] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer, Berlin, 1991).

[17] C.M. Gardiner, *Handbook on Stochastic Methods* (Springer, Berlin, 1983).

[18] V. Berdichevsky, J. Appl. Math. Mech. **52**, 738 (1988).

[19] I. Kunin, *Theory of Elastic Media with Microstructure* (Springer, Berlin, 1983).

[20] L.D. Landau and E.M. Lifshits, *Statistical Physics* (Pergamon, London, 1975), Sec. 9.

[21] T. Poston, and I. Stewart, *Catastrophe Theory and its*

Applications (Pitman, London, 1978), Sec. 42.

[22] L.I. Sedov, *Continuum Mechanics* (World Scientific, Singapore, 1997).

[23] V. Berdichevsky, Int. J. Eng. Sci. **31**, 12 (1993).

[24] D. Ruelle, Phys. Lett. A **245**, 220 (1998).

[25] I.P. Kornfeld, J.G. Sinai, and S.V. Fomin, *Ergodic Theory* (Nauka, Moscow, 1980).

[26] M. Grmela and H.C. Ottinger, Phys. Rev. E **56**, 6620 (1997); **56**, 6633 (1997).

[27] A.N. Beris and B.J. Edwards, *Thermodynamics of Flowing Systems with Internal Microstructure* (Oxford University Press, Oxford, 1994).

[28] I. Kunin (private communication).

[29] A. Kaufmann, *Introduction a la Combinatorique en Vue des Applications* (Dunod, Paris, 1968).

[30] V.I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1979).