On the stability of equilibrium states of finite classical systems

Joel L. Lebowitz*

Belfer Graduate School of Science, Yeshiva University, New York, New York

Michael Aizemann†

Courant Institute, New York University, New York, New York

Sheldon Goldstein‡

Institute for Advanced Study, Princeton, New Jersey
(Received 2 January 1975)

The state of a system is characterized, in statistical mechanics, by a measure ω on Γ, the phase space of the system (i.e., by an ensemble). To represent an equilibrium state, the measure must be stationary under the time evolution induced by the systems Hamiltonian $H(x)$, $x \in \Gamma$. An example of such a measure is $\omega(dx) = f(H)dx$; dx is the Liouville (Lebesgue) measure and $f(H(x))$ is the ensemble density. For "nonergodic" systems there are also other stationary measures with ensemble densities, e.g., for integrable dynamical systems the density can be a function of any of the constants of the motion. We show, however, that the requirement that the equilibrium measure have a certain type of "stability" singles out, in the typical case, densities which depend only on H.

1. INTRODUCTION

The macroscopic description of a physical system is assumed in statistical mechanics to be given by a probability measure ω on the phase space Γ of the system[1,2]: If A is a region of the phase space, $A \subset \Gamma$, then $\omega(A)$ is the probability that the phase point of the system will be found in A. Equivalently, $\omega(A)$ is the fraction of systems in the ensemble in the region A. To describe a system in equilibrium the measure must be stationary under the time evolution. Since the energy (Hamiltonian) H of a finite system of particles is always a constant of the motion, a measure given by a function of the energy (times Lebesgue measure), $\omega(A) = \int_A f(H)\,dx$, will always be stationary. Conversely, if the time evolution is ergodic on all the energy surfaces S_E [specified by $H(x) = E$] equipped with their natural microcanonical measures, then every stationary measure ω given by a density ρ, i.e., $\omega(A) = \int_A \rho(x)\,dx$ so that $\omega(A) = 0$ if $\int_A dx = 0$, will be of this form.[2,3] If, in the other extreme, the system is integrable,[2,4] so that there are in addition to H other "smooth" constants of the motion, then there will also be stationary states whose densities are functions of these constants of the motion.

Consider, for example, an ideal gas consisting of n particles moving in a unit box with periodic boundary conditions—the unit torus T^3. The phase space of this system is $T^{3n} \times R^{3n}$ and the time evolution T_t, is given by

$$T_t(q_1, \ldots q_{3n}, p_1, \ldots p_{3n}) = T_t(\mathbf{q}, \mathbf{p}) = (\mathbf{q} + \mathbf{p}t, \mathbf{p}),$$

where $(\mathbf{q}, \mathbf{p}) = x \in T^{3n} \times R^{3n} = \Gamma$ and the addition is modulo 1. This evolution comes from the ideal gas Hamiltonian

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{3n} \tfrac{1}{2} p_i^2.$$

The p_i, as well as H, are constants of the motion and thus any ensemble density which is a function of $\mathbf{p}$ only will be stationary under the time evolution.

Nevertheless, the equilibrium properties of finite systems, even those which are not ergodic, are usually assumed to be determined by a density which is a function of H only.[1,2] In this note we shall not discuss any specific form of this function but consider justifications of the assumption that the ensemble density is a function of H only even when there are also other constants of the motion present.[5] (It is known that all reasonable functions of H lead to the same results for local quantities in the thermodynamic limit.[1])

There may, of course, exist singular stationary measures (not given by a density) which are not "constant" on energy surfaces even if the system is ergodic on all such surfaces.[2] It may be argued, however, that these measures, which assign a finite probability to the system being found in a region of the phase space A which has zero (Liouville—Lebesgue) volume, i.e., $\int_A d\mathbf{q}\,d\mathbf{p} = 0$, should be irrelevant for explaining experimentally observed behavior. Experimental results depend on reproducibility and it seems plausible to assume that there will be a vanishing "probability" for "preparing" a physical system in such a region.[1a] We may then regard as physically reasonable only those measures which are absolutely continuous, i.e., have a density, with respect to Lebesgue measure. We shall adopt this attitude here and only worry about the justification of assuming $\rho(x)$ to depend on H only. [The microcanonical ensemble, at a fixed energy E, is itself singular with respect to Lebesgue measure $d\mathbf{q}\,d\mathbf{p}$; it may, however, be regarded[1,2] as the limit when $\Delta E \to 0$, of measures concentrated, with uniform density, on the energy shell $(E, E + \Delta E)$, i.e., $\rho(x) = f(H) = \text{const}$ for $E \le H \le E + \Delta E$, and is zero outside this shell. As already noted the results, for large systems, are independent of this limit.]

While we shall be concerned here exclusively with finite systems similar problems arise for infinite systems. In the case of infinite quantum systems, Haag, Kastler, and Trych-Pohlmeyer[6] (HKP) have shown that a condition of stability under local perturbations of the time evolution is useful for the characterization of equilibrium states, i.e., under certain reasonable assumptions the only stable states are KMS states. Their argument may be adapted to prove a similar result for infinite classical systems.[7] In this note we wish to consider the extent to which "stability" may be useful for the characterization of equilibrium ensembles for finite systems.

2. FORMULATION OF PROBLEM

The notion of stability which we wish to use is similar to that used by HKP and may be described roughly as follows: Let ω be the stationary state given by the function $f=f(H)$. If we perturb H slightly to obtain a new Hamiltonian $H_\lambda = H + \lambda h$, we obtain a new time evolution $T_t^{\lambda h}$ ($\equiv T_t^\lambda$) such that there exists a measure $\omega^{\lambda h}$ ($\equiv \omega^\lambda$) [given by the function $f(H+\lambda h)$] which is (a) stationary under T_t^λ and (b) "close" to ω. We will say that a state ω stationary under T_t is stable if there exists such a family $\omega^{\lambda h}$ which is close to ω for all (sufficiently nice) perturbations h. A state ω which fails to be stable in this sense should not be regarded as "physical" because an arbitrarily small error in our knowledge of H could imply that ω does not even approximate a state stationary under the actual Hamiltonian time evolution.

To obtain a precise formulation of stability we must decide exactly how ω^λ is to be close to ω. Since the only use of the measure (or ensemble) is to obtain expectation values of physical observables, i.e., of functions $A(x)$, which (by the very nature of physical observations) may be assumed to be smooth functions of x, $x \in \Gamma$, closeness should refer to such expectation values. We shall write $\omega(A)$ and $\omega^\lambda(A)$ for the expectation value[8] of A, with respect to the measures ω and ω^λ, and will assume throughout that H and all perturbations are $\in C^2(\Gamma)$ and that h is bounded. Some possibilities are:

(i) $\omega^\lambda \to \omega$ in norm, i.e.,

$$|\omega^\lambda(A) - \omega(A)| \le \epsilon(\lambda) \, \| A \|,$$

where $\lim_{\lambda\to 0} \epsilon(\lambda) = 0$, $A \in C(\Gamma)$, the bounded continuous functions on the phase space Γ of the finite system, and $\|A\| = \sup_{x\in\Gamma} |A(x)|$;

(ii) $\omega^\lambda \xrightarrow[\lambda\to 0]{} \omega$ weakly, i.e., $\omega^\lambda(A) \xrightarrow[\lambda\to 0]{} \omega(A)$ for all $A \in C(\Gamma)$.

Clearly, (i) implies (ii). It is also worth noting that there is a natural dynamical formulation of stability which is equivalent to (i).

(i′) $T_t^{\lambda h}\omega$ remains close (in norm) to ω uniformly in t, for any perturbation h, when λ is sufficiently small, i.e.,

$$|\omega(T_t^\lambda A) - \omega(A)| < \epsilon(\lambda) \, \|A\|$$

for all $A \in C(\Gamma)$ and all t.

To prove equivalence we note that (i′) follows from (i) because

$$|\omega(T_t^\lambda A) - \omega(A)| \le |\omega(T_t^\lambda A) - \omega^\lambda(T_t^\lambda A)| + |\omega^\lambda(A) - \omega(A)| \le 2\epsilon(\lambda) \, \|A\|,$$

since $\omega^\lambda(T_t^\lambda A) = \omega^\lambda(A)$ by the stationarity of ω^λ under the perturbed evolution and $\|T_t^\lambda A\| = \|A\|$. Conversely, if (i′) holds we may construct ω^λ norm close to ω as a weak limit point of the time averages $\overline{\omega}_T^\lambda$ of the measures

$$T_t^\lambda \omega \quad (\overline{\omega}_T^\lambda = 1/T \int_0^T dt \, T_t^\lambda \omega).$$

Condition (i′) may be called dynamical stability: Suppose a perturbation λh is added to H at some time, say $t=0$; then ω will change with time for $t>0$. If, however, ω satisfies (i′) and λ is small then the expectation values of physical observables will also be changed only slightly even after very long times. (This remains true also if the initial state is not exactly ω but some state ω' which is close to ω in norm.)

These conditions have quantum counterparts: one replaces $C(\Gamma)$ in the above by the C^*-algebra $B(\mathcal{H})$ of bounded operators on the Hilbert space $\mathcal{H}$ corresponding to the finite quantum system—of a finite number of particles in a finite volume. ω and ω^λ correspond to normal states on $B(\mathcal{H})$ [i.e., positive linear functionals $\tilde{\omega}$ of the form $A \to \mathrm{tr}(A\rho)$, $A \in B(\mathcal{H})$, where $\rho \in B(\mathcal{H})$ is positive and $\mathrm{tr}(\rho)=1$] which are invariant under the one-parameter groups T_t and $T_t^{\lambda h}$ generated by the Hamiltonians H and $H+\lambda h$, $h \in B(\mathcal{H})$, respectively. For finite systems H has discrete spectrum and corresponding to states of the form $f(H)$ for classical systems one has the invariant states given by $\rho = f(H)$ (e.g., $\rho = e^{-\beta H}/\mathrm{Tr}e^{-\beta H}$) for quantum systems.

In both the classical and qunatum situations, a state given by a (reasonable) function $f(H)$ will satisfy (i) and (ii) and thus, also (i′). In the quantum case a state is stationary if and only if $[\rho, H]$ ($\equiv \rho H - H\rho$) $= 0$, so that if H has nondegenerate spectrum, ρ must clearly be of the desired form. Even if H is degenerate the restriction of ρ to each energy level must still be the identity if (ii) is to be satisfied, since any splitting of an energy level may be achieved by the appropriate choice of perturbation.[6] In the classical situation we need stronger conditions that (i) and (ii) to obtain a general result. Before introducing such a condition, in Sec. 4, we shall, in the next section, investigate some consequence which follow solely from the "weak stability" condition (ii).

3. SOME CONSEQUENCES OF WEAK STABILITY

Proposition 1: Let ω be weakly stable under the perturbation h as in (ii), i.e., there exists a collection $\omega^{\lambda h}$ of states invariant under the dynamics generated by $H+\lambda h$ which converge weakly to ω. Then $\omega^{\lambda h}(Q)$ is differentiable at $\lambda=0$ on observables of the form $Q=\{H,B\}$ [the Poisson bracket (P.B.) of H with B] for some $B \in C_0^1(\Gamma)$ (C^1 functions of compact support)[9] and

$$\frac{d}{d\lambda}\, \omega^{\lambda h}(\{H,B\})_{\lambda=0} = -\,\omega(\{h,B\}). \tag{3.1}$$

In particular, if B is a constant of the motion $\{H,B\}=0$, then

$$\omega(\{h,B\}) = 0. \tag{3.2}$$

Proof: For any $B \in C_0^1(\Gamma)$ the perturbed states satisfy

$$0 = \frac{d}{dt}\, \omega^{\lambda h}(T_t^\lambda B)_{t=0} = \omega^{\lambda h}(\{H+\lambda h, B\}),$$

or

$$\frac{1}{\lambda}\,\omega^{\lambda h}(\{H,B\}) = -\,\omega^{\lambda h}(\{h,B\}).$$

The weak continuity of $\omega^{\lambda h}$ at $\lambda=0$ implies therefore the existence of the limit

$$\lim_{\lambda\to 0}\frac{1}{\lambda}\,\omega^{\lambda h}(\{H,B\}) = -\lim_{\lambda\to 0}\omega^{\lambda h}(\{h,B\})$$
$$= -\,\omega(\{h,B\}).$$

Since, by stationarity,

$$\omega(\{H,B\}) = 0,$$

the above limit is the weak derivative of $\omega^{\lambda h}$ on $Q=\{H,B\}$.

Proposition 2: If ω satisfies stability (ii) and is given by a $C^1(\Gamma)$ density ρ, then

$$\{\rho, B\} = 0 \tag{3.3}$$

for any $B\in C_0^1(\Gamma)$ such that $\{H,B\}=0$.

Proof: By Proposition 1 $\{H,B\}=0$ implies $\omega(\{h,B\})=0$ for any $h\in C_0^2$. In terms of ρ we thus have, using well-known properties of the P. B.,

$$0=\int dx\,\rho\{h,B\}$$
$$=\int dx\,\{\rho h,B\}-\int dx\,h\{\rho,B\}$$
$$=-\int dx\,h\{\rho,B\}.$$

Since h is arbitrary this implies (3.3.).

We have thus obtained a simple condition on ω, (3.2) and (3.3), necessary for stability (ii).

The above arguments can be reproduced for quantum systems, with the understanding that $\{\ ,\ \}$ stands for the commutator. According to (3.3) a state of a quantum system, given by a density operator ρ, is stable (ii) only if ρ commutes with all operators which commute with the Hamiltonian H. Since H has discrete spectrum it follows simply that ρ is a function of H.

No such general result can be expected for classical systems as may be seen by considering integrable systems for which the Kolmogorov—Arnold—Moser (KAM) theorem[2,4] is applicable. It can be shown, see remark at end of Sec. 4, that for such systems even the stronger stability condition (i) is not sufficient to insure the desired result $\rho=f(H)$.

The difference between classical and quantum systems appears to be due to the lack of a sufficient number of global constants of the motion in the classical case. This prevents fuller exploitation of Proposition 2 whose usefulness depends on the existence of an abundance of invariants. Even integrable systems, if they satisfy the conditions of the KAM theorem, have only a "limited" number of such constant (i.e., n constants when Γ is a $2n$-dimensional space). This shows up in the requirements for KAM theorem that the frequencies be incommensurable[4]; this reduces the number of smooth invariants; e.g., for two uncoupled oscillators there exists a function of the two phases which is a (smooth) invariant iff the frequencies are commensurable. Indeed, we shall now prove that in the extreme case of a periodic system weak stability alone implies that $\rho=f(H)$. We shall consider this case explicitly, despite its limited applicability, to illustrate the method used in the next section for more "typical" systems.

Proposition 3: Let ω be a state of a periodic system, given by a $C^1(\Gamma)$ density ρ. If ω is weakly stable [i.e., satisfies stability (ii)], then locally ρ is a function of H, i.e.,

$$\text{grad}\rho \text{ is parallel to grad}H. \tag{3.4}$$

Proof: Denote by τ the period of the system. Then, for any $A\in C_0^1$,

$$\bar{A}(x)=\int_0^\tau dt\,A(T_t x)$$

is a constant of the motion. Proposition 2 now implies that

$$0=\{\rho,\bar{A}\}=\{\rho,\int_0^\tau dt\,T_tA\}=\int_0^\tau dt\{\rho,T_tA\}$$
$$=\int_0^\tau dt\,T_t\{\rho,A\}, \tag{3.5}$$

where we have used the invariance of ρ under T_t. Assume now that gradρ is not parallel to grad$H$ at some point x. One could then find on observable A, with support in a neighborhood of x, in which $\{\rho,A\}>0$ along the orbit of x. This would contradict (3.5).

The typical (generic) integrable system is not periodic. Nevertheless its periodic points are dense in the phase space.[4] In the next section we show how to obtain a positive result for such systems at the price of imposing a somewhat stronger, and not so physical, requirement of stability on the equilibrium states.

4. A STRONGER STABILITY CONDITION

As we have seen in Propositions 1 and 2, the weak stability of a state ω enables one to define, for each smooth perturbation h of compact support, a functional L_h, whose domain are observables of the form $Q=\{H,B\}$, by

$$L_h(\{H,B\})=-\,\omega(\{h,B\}).$$

L_h was shown there to be the weak derivative of the perturbed states $\omega^{\lambda h}$.

Definition: A state ω satisfies stability (iii) if it is weakly stable and if, for each $h\in C_0^2$, the functional L_h is given by a $C^2(\Gamma)$ function f_h, i.e.,

$$L_h(\{H,B\})=\int dx\,f_h(x)\{H,B\}.$$

When ω has a density ρ

$$\int dx\,f_h\{H,B\}=-\int dx\,\rho\{h,B\}.$$

This gives after integration by parts, assuming $\rho\in C^1(\Gamma)$,

$$-\int dx\,B\{H,f_h\}=\int dx\,B\{h,\rho\}.$$

Since this holds for, essentially, any B it implies

$$-\{H,f_h\}=\{h,\rho\}. \tag{4.1}$$

Thus, for states given by a density, stability (iii) implies that for each perturbation h there exists a $C^1(\Gamma)$ function f_h which satisfies (4.1). This condition is satisfied by ρ of the desired form, i.e., $\rho=f(H)$, $f\in C^1$, since

$$\{h, \rho\} = \{h, f(H)\} = f'(H)\{h, H\} = \{f'(H)h, H\}$$

and one may choose $f_h = f'(H)h$.

We will now show that in the generic case, the converse of the above statement is also true.

Proposition 4: Let ω satisfy stability (iii) and be given by a C^1 density ρ. If periodic orbits (under T_t) are dense in Γ and if the energy surfaces S_E are connected then ρ is a function of H.

Proof: Let $y \in \Gamma$ be a periodic point with period τ. By stability (iii), there corresponds to each $h \in C_0^2$ a $C^1(\Gamma)$ function f_h such that

$$\{\rho, h\} = \{H, f_h\}.$$

Therefore, using the periodicity of the orbit through y, we obtain

$$\int_0^\tau du\,\{\rho, h\}(T_u y) = \int_0^\tau du\,\{H, f_h\}(T_u y) = \int_0^\tau du\,\frac{d}{du} f_h(T_u y) = f_h(T_u y) - f_h(y) = 0$$

for any $h \in C_0^2$. By the same argument as in the proof of Proposition 3, we conclude that gradρ is parallel to grad$H$ at y.

Since the periodic points are dense the gradients of ρ and of H are parallel everywhere. The connectedness of energy surfaces now implies that ρ is a function of H.

Remark: The assumptions made in Proposition 4 cannot easily be weakened as may be seen by considering stability in integrable systems to which the KAM theorem is applicable.[4] (The ideal gas in a torus is such a system.) In these systems the phase space is decomposable into invariant (under T_t) tori "most" of which are stable under small (sufficiently smooth) perturbations h: That is, except for a family of tori of total measure $\epsilon(\lambda)$, there corresponds to each T_t-invariant torus M a uniformly close $T_t^{\lambda h}$-invariant torus M^λ (on which the $T_t^{\lambda h}$ time evolution uniformly approximates the T_t evolution on M). Here $\epsilon(\lambda) \to 0$ as $\lambda \to 0$ and M^λ is "differentiably close" to M. Hence for any T_t-stationary measure which is given by a smooth "function of the invariant tori" (i.e., a function of the "action variables" parameterizing the tori) one may use the correspondence $M \longleftrightarrow M^\lambda$ to construct a T_t^λ-stationary measure ω^λ which is norm close to ω and even differentiably close. Thus, unless the use of perturbations to which KAM does not apply is allowed—in our argument h could be arbitrarily smooth—the proposition will not hold if we replace in it stability (iii) by stability (ii) or even stability (i). Stability (iii), on the other hand, will rule out these cases because the derivative of ω^λ at $\lambda = 0$ may fail to be even a function and will certainly not be C^1. A positive result may, however, be possible if the ω^λ are required to be given by smooth functions, since this is almost certainly not the case for the ω^λ which can be constructed by the use of the KAM theorem.

ACKNOWLEDGMENTS

We would like to thank Rudolf Haag for informing us of (and explaining to us) the results of Ref. 6 prior to publication, and for many very useful comments. We have also benefited greatly from discussions with Oscar Lanford. [The proof of the equivalence of (i) and (i′) is due to him.] Finally, we would like to thank the Battelle Foundation for the opportunity to participate in the Rencontres on Dynamical Systems (Seattle, 1974) where we had valuable discussions with Jurgen Moser and learned something about the KAM theorem.

*Supported in part by USAFOSR Grant #73-2430B.
†Supported in part by NSF Grant GP-37069X.
‡Supported in part by NSF Grant GP-16147A#1.

[1]See, for example, (a) O. Penrose, *Foundations of Statistical Mechanics* (Pergamon, Oxford, 1970); (b) Arthur Hobson, *Concepts in Statistical Mechanics* (Gordon and Breach, New York, 1971); (c) D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1968).

[2]J. L. Lebowitz and O. Penrose, Physics Today 26, N. 2 (1973) and references cited there.

[3]We take our system to be confined to a box, or torus, with bounded energy surfaces.

[4]See V.I. Arnold and A. Avez, *Ergodic Problems of Statistical Mechanics* (Benjamin, New York, 1968).

[5]There are of course special situations; c.f. Ref. 1, where one considers explicitly equilibrium ensemble densities which depend also on the total linear and/or angular momentum of the system. We do not discuss these here.

[6]R. Haag, D. Kastler, and E. Trych-Pohlymeyer, Commun. Math. Phys. 38, 173 (1974).

[7]M. Aizenman, G. Gallavotti, S. Goldstein, and J. L. Lebowitz (to be published).

[8]This is consistent with our definition of $\omega(A)$ in Sec. 1 if we think there of $A(x)$ as the characteristic function of the set $A \subset \Gamma$, $A(x) = 1$ if $x \in A$, $A(x) = 0$ for $x \notin A$.

[9]This collection includes observables of the form $T_t A - A$, for $t \in \mathbb{R}$ and $A \in C_0^1(\Gamma)$ [since $T_t A - A = \{H, \int_0^t T_u(A)\,du\}$] and is, therefore, dense in the orthogonal complement [in $L_2(\omega)$] of the constants of the motion.