Journal of Statistical Physics, Vol. 49, Nos. 1/2, 1987

Free Energy and the Relative Entropy

Matthew J. Donald[1]

Received October 29, 1986; revision received February 16, 1987

Under the assumption of an identity determining the free energy of a state of a statistical mechanical system relative to a given equilibrium state by means of the relative entropy, it is shown: first, that there is in any physically definable convex set of states a unique state of minimum free energy measured relative to a given equilibrium state; second, that if a state has finite free energy relative to an equilibrium state, then the set of its time translates is a weakly relatively compact set; and third, that a unique perturbed equilibrium state exists following a change in Hamiltonian that is bounded below.

KEY WORDS: Free energy; relative entropy; relative Hamiltonian.

1. INTRODUCTION AND MATHEMATICAL FRAMEWORK

The purpose of this paper is the exposition of three applications to statistical mechanics of some results on the relative entropy. These results were originally developed in the quite different context of the study of the foundations of quantum mechanics. The applications are very general, abstract results with all the ensuing advantages and disadvantages. They are; a theorem that for general systems at finite temperature there is in any closed convex set of states with nonempty interior a unique state of minimum free energy measured relative to a given equilibrium state; a theorem that if a state has finite free energy relative to an equilibrium state, then the set of its time translates can never be a very big set—more precisely, any sequence of time translates of a state of finite free energy has a subsequence weakly convergent to a normal state; and a theorem showing that if the internal energy of a system changes in such a way that for no state does the internal energy decrease by an infinite amount, then there is a unique state of minimum free energy for the new system

[1] The Cavendish Laboratory, Cambridge CB3 OHE, Great Britain.

0022-4715/87/1000-0081$05.00/0

measured relative to a given equilibrium state for the old. While I have tried to give a self-contained explanation of the results, the reader will have to go to Ref. 1 and/or Ref. 2 for the historical background, for complete references, and for general mathematical definitions of the relative entropy, and to Ref. 3 for the proof of the result of Section 2.

The relative entropy and its relation to free energy are best introduced by an example. This example is elementary in mathematical structure, but it is, of course, not sufficient to encompass general statistical mechanical systems.[2] Let σ be a normal state, i.e., a quantum mechanical density matrix, on a Hilbert space $\mathscr{H}$, with Hamiltonian H and inverse temperature $\beta = 1/kT$. Let ω be the Gibbs equilibrium state $\omega = e^{-\beta H}/Z$, where $Z = \mathrm{tr}(e^{-\beta H})$—assuming that this definition makes sense. The free energy of σ is defined, using the thermodynamic definition $F = U - TS$, as

$$F(\sigma) = \mathrm{tr}(\sigma H) - \beta^{-1}\,\mathrm{tr}(-\sigma \log \sigma)$$

This gives $F(\omega) = -\beta^{-1} \log Z$ and

$$F(\sigma) - F(\omega) = -\beta^{-1}\,\mathrm{tr}(-\sigma \log \sigma + \sigma \log \omega) \tag{1.1}$$

Apart from the factor, this agrees with the definition of the so-called relative entropy[1,2] of σ with respect to ω on the algebra $\mathscr{B}(\mathscr{H})$ of all bounded operators on $\mathscr{H}$—written $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma \mid \omega)$—so that we have

$$F(\sigma) - F(\omega) = -\beta^{-1}\,\mathrm{ent}_{\mathscr{B}(\mathscr{H})}(\sigma \mid \omega)$$

General statistical mechanical systems (both quantum and classical) are described by algebras $\mathscr{A}$ (generalizing $\mathscr{B}(\mathscr{H})$) and we may choose, by analogy with (1.1), to define the free energy of an arbitrary state σ relative to a given equilibrium state ω for $0 < \beta < \infty$ by

$$F(\sigma) - F(\omega) = -\beta^{-1}\,\mathrm{ent}_{\mathscr{A}}(\sigma \mid \omega) \tag{1.2}$$

This definition can be made for a very general class of models since the relative entropy can be defined for arbitrary von Neumann algebras and states[1,2]. The generalization is much harder to handle mathematically than the simple formula (1.1). In order to use the comparatively simple definition of $\mathrm{ent}_{\mathscr{A}}(\sigma \mid \omega)$ given in Ref. 1 and the results of Ref. 3, we shall assume throughout that $\mathscr{A}$ is an injective von Neumann algebra. For physical applications this is no loss of generality. Whether in each case (1.2) really does correspond to the physical concept of a free energy difference can only be fully answered by detailed study of individual models. References 1 and 3 show that $\mathrm{ent}_{\mathscr{A}}$ is the "natural" generalization in various ways of $\mathrm{ent}_{\mathscr{B}(\mathscr{H})}$. This paper presents some general properties that both give us physical results when we can justify the correspondence and encourage us to make it in other cases. For quantum spin models there is good reason to believe that the correspondence is correct, and the definition has been used in such models for some time in order to characterize equilibrium states (see Ref. 2, Section 6.2.3, and references therein). In particular, Araki[4] has shown that if σ is a second equilibrium state, then $F(\sigma)-F(\omega)=\infty$. Physically this is the statement that an infinite amount of work is required to move a system from one equilibrium to another in the infinite-volume thermodynamic limit.

Here we are interested in finite perturbations of our system, so we assume a given equilibrium state ω in which our system starts, or which is taken to define the boundary conditions. We are not interested in finiteness of the free energy per unit volume, but in its global finiteness; ω is the state from which we measure free energy. The only mathematical constraint on the state ω is that it should be faithful; for nonzero temperatures, this property is certainly possessed by quantum equilibrium states (Ref. 2, Theorem 5.3.10) and its failure for any system would be physically absurd.

2. THE UNIQUENESS OF STATES OF MINIMUM FREE ENERGY

Theorem. Let K be a convex set of states which is closed and has nonempty interior in the w^* topology. Then, there is a unique state $\hat{\sigma}(K,\omega)$ in K minimizing $F(\sigma)-F(\omega)$, and $\hat{\sigma}(K,\omega)$ is normal.

This is a direct translation of Theorem 4.4 of Ref. 3, where a full proof may be found. It is my belief that this result should be of sufficient interest to statistical physicists to be worth repeating here in their language. In Ref. 3 it is discussed in the context of the mathematical theory of relative entropy and its use as a tool for quantum statistical inference.

The essential step in the proof is the demonstration of the strict concavity of the relative entropy in its first variable, a result that follows from the nontriviality of the relative entropy and the identity

$$\begin{aligned} x_1 \operatorname{ent}_{\mathscr{A}}(\sigma_1 \mid \omega) + x_2 \operatorname{ent}_{\mathscr{A}}(\sigma_2 \mid \omega) \\ = \operatorname{ent}_{\mathscr{A}}(\sigma \mid \omega) + x_1 \operatorname{ent}_{\mathscr{A}}(\sigma_1 \mid \sigma) + x_2 \operatorname{ent}_{\mathscr{A}}(\sigma_2 \mid \sigma) \end{aligned}$$

which holds for all states σ_1, σ_2, and ω, and for all $x_1 \in [0,1]$ with $x_1 + x_2 = 1$ and $\sigma = x_1\sigma_1 + x_2\sigma_2$. Since Ref. 3 was submitted, Petz[5] has published an independently derived and quite different proof of this strict concavity result.

As an example, consider a spin system with an equilibrium state with all spins pointing south. Take a finite number of the spins and constrain them to point west. Use a real physical system, so that the west-pointing spins point west only to a given approximation (which can be arbitrarily small, but which must not vanish). Then the set K is all possible states with these particular spins pointing west to within the given approximation, and the result says that there is a unique state in K of minimum free energy for the system relative to the south-pointing equilibrium state with which we started.

This is mainly of interest as a fact about the spins away from the constrained ones. Indeed, this example should indicate that the physical relevance of the theorem depends on the extent to which we can limit the set of states available to the system without essentially altering it. If the constraints can be seen instead as a change in Hamiltonian, then the result in Section 4 will be more relevant.

As a more general example, suppose that after a finite perturbation of an arbitrary system from an equilibrium state ω, a number of measurements have revealed that the state of the system has been moved into some set K for a time long compared with the microscopic relaxation times of the system. Then $\hat{\sigma}(K, \omega)$ is the most plausible state to assign to the system. It may be that this provides an interesting way of considering metastability: take K to be the set of all states that in a finite region are on the "wrong" side of some activation barrier—the local metastable state should then be the state of minimum free energy in K.

Finally, note that the result given does need proof—for example, states of minimum energy in a set of all states with energy not less than some fixed amount certainly need not be unique.

3. STATES OF FINITE FREE ENERGY CANNOT WANDER

One goal of statistical mechanics is to discuss conditions for the return to equilibrium of perturbed systems (see, for example, Ref. 2, Chapter 5.4). In this section we give a very general result, which might be a useful preliminary to such a discussion. As above, we take a given an equilibrium state ω of a statistical mechanical system on an injective von Neumann algebra $\mathscr{A}$. We suppose that this system is closed and so time translation corresponds to a family τ_t of *-automorphisms of $\mathscr{A}$ defined for t in some set of times $T \subset \mathbb{R}$. The equilibrium nature of ω is used to yield the hypothesis that ω is invariant in time, i.e., $\omega \circ \tau_t = \omega$.

Theorem. Let $\mathscr{A}$ be an injective von Neumann algebra and σ and ω be states on $\mathscr{A}$ with ω normal. Let $\{\tau_t : t \in T\}$ be a set of *-automorphisms of $\mathscr{A}$ under which ω is invariant. Suppose that $F(\sigma) - F(\omega) = -\beta^{-1} \operatorname{ent}_{\mathscr{A}}(\sigma \mid \omega)$ is finite. Then $\{\sigma \circ \tau_t : t \in T\}$ is a weakly relatively compact set of normal states. In particular, all the limit points of this set are normal.

Proof. The invariance of ω and Theorem 8.6 of Ref. 1 give, for all $t \in T$,

$$\mathrm{ent}_{\mathscr{A}}(\sigma \circ \tau_t \,|\, \omega) = \mathrm{ent}_{\mathscr{A}}(\sigma \circ \tau_t \,|\, \omega \circ \tau_t) = \mathrm{ent}_{\mathscr{A}}(\sigma \,|\, \omega)$$

The result is now a consequence of the following:

Proposition. Let ω be a normal state on an injective von Neumann algebra $\mathscr{A}$. Then, for all finite α, $\{\rho: \mathrm{ent}_{\mathscr{A}}(\rho \,|\, \omega) \geqslant \alpha\}$ is a weakly compact set of normal states on $\mathscr{A}$.

Proof. Let $\Sigma(\alpha, \omega) = \{\rho: \mathrm{ent}_{\mathscr{A}}(\rho \,|\, \omega) \geqslant \alpha\}$. By Ref. 3, Lemma 4.3, each $\rho \in \Sigma(\alpha, \omega)$ is normal. By w^* upper semicontinuity of ent (Ref. 1, property c) $\Sigma(\alpha, \omega)$ is closed and hence compact in the w^* topology on the space of all states, and this is the weak topology [i.e., the $\sigma(\mathscr{A}_*, \mathscr{A})$ topology] on the normal states. ▌

Remarks. 1. Translating this into more physical language, it says simply that in a closed system at fixed, finite temperature the free energy of a state does not change in time—free energy is available work, but that work could only be done outside the system—and that the set of all states with free energy bounded by a given amount is weakly compact.

2. The result should be contrasted with the situation in quantum mechanical scattering theory. There we have a Hamiltonian H and a pure state $|\psi\rangle\langle\psi|$, where ψ is a Hilbert space vector in the absolutely continuous subspace of H. Setting $\sigma_t = e^{-itH}\,|\psi\rangle\langle\psi|\, e^{itH}$, it is a consequence of Ref. 6, §XI.3, Lemma 2, p. 24, that $\{\sigma_t: t \in [0, \infty)\}$ is never weakly relatively compact. Scattering states wander, states of finite free energy do not.

3. Extensions to open systems might involve time translations λ_t that are normal, completely positive maps under which ω is invariant. Then the Uhlmann–Lindblad inequality (property h of Ref. 1) shows that the free energy relative to ω cannot increase with time. Concavity of the relative entropy (property b of Ref. 1) shows that the free energy also cannot increase under "coarse graining"—taking convex combinations of the $\sigma \circ \tau_t$ at different times.

4. The hypotheses on τ_t and ω are satisfied by taking τ_t to be the identity map and ω to be any normal state, so certainly we have no more than a preliminary to a theory of the return to equilibrium.

4. PERTURBED EQUILIBRIUM

In the elementary example given in the introduction, suppose that the Hamiltonian of the system changes from H to $H + h$. The free energy of state σ changes to $F(\sigma)+\sigma(h)$, so the state of minimum free energy under the new Hamiltonian, which should be the equilibrium state of the perturbed system, will be the state ω^h that minimizes the function $\sigma \mapsto F(\sigma) - F(\omega) + \sigma(h)$. Existence and uniqueness of such states for general systems is proved in the following theorem under the condition that h is bounded below, i.e., for no state does the internal energy decrease by an infinite amount.

Theorem. Let ω be a normal state on an injective von Neumann algebra $\mathscr{A}$, $\beta > 0$, and h be a self-adjoint operator affiliated with $\mathscr{A}$ that is bounded below. Suppose that there is some state ρ with

$$F(\rho) - F(\omega) + \rho(h) = -\beta^{-1}[\mathrm{ent}_{\mathscr{A}}(\rho \mid \omega) - \beta\rho(h)]$$

finite. Then there exists a unique state ω^h minimizing the function $\sigma \mapsto F(\sigma) - F(\omega) + \sigma(h)$, and this state is normal.

Proof. For $n = 1, 2, \ldots$, let P_n be the spectral projection of h in $(-\infty, n]$. Then $\sigma(h)$ is defined by $\sigma(h) = \lim_{n\to\infty} \sigma(hP_n)$. As the limit of a decreasing sequence of bounded linear functionals, $-\beta\sigma(h)$ is w^* upper semicontinuous and concave. Now, by the method of Theorem 4.4 of Ref. 3, $\mathrm{ent}_{\mathscr{A}}(\sigma \mid \omega) - \beta\sigma(h)$ is finite at ρ, is w^* upper semicontinuous, strictly concave, and finite only for normal states σ. The result follows. ▮

In the case that h is bounded above as well as below, these states ω^h are the same as those constructed by Araki[7] by a very different method (see Ref. 2, Theorem 5.4.4 and Proposition 6.2.32). For bounded h, ω^h always exists, since $\mathrm{ent}_{\mathscr{A}}(\omega \mid \omega) - \beta\omega(h) = -\beta\omega(h)$, which is automatically finite. The relation between ω^h for bounded and semibounded h is made more precise by the next result.

Proposition. With the notation and conditions of the preceding theorem, set $h_n = hP_n$. Then, ω^{h_n} converges weakly to ω^h.

Proof. Take $c \in (-\infty, 0]$ such that $h \geqslant c1$. Choose $m \geqslant n$. Then $h \geqslant h_m \geqslant h_n \geqslant c1$, and so

$$\begin{aligned}
-\infty &< \mathrm{ent}_{\mathscr{A}}(\omega^h \mid \omega) - \beta\omega^h(h) \\
&\leqslant \mathrm{ent}_{\mathscr{A}}(\omega^h \mid \omega) - \beta\omega^h(h_m) \\
&\leqslant \mathrm{ent}_{\mathscr{A}}(\omega^{h_m} \mid \omega) - \beta\omega^{h_m}(h_m) \qquad \text{(by definition of } \omega^{h_m}) \\
&\leqslant \mathrm{ent}_{\mathscr{A}}(\omega^{h_m} \mid \omega) - \beta\omega^{h_m}(h_n) \\
&\leqslant \mathrm{ent}_{\mathscr{A}}(\omega^{h_n} \mid \omega) - \beta\omega^{h_n}(h_n) \\
&\leqslant \mathrm{ent}_{\mathscr{A}}(\omega^{h_n} \mid \omega) - \beta c
\end{aligned}$$

Thus,

$$\mathrm{ent}_{\mathscr{A}}(\omega^{h_n} \mid \omega) \geqslant \mathrm{ent}_{\mathscr{A}}(\omega^h \mid \omega) - \beta\omega^h(h) + \beta c$$

and by the proposition of Section 3 it follows that $\{\omega^{h_n}: n = 1, 2,...\}$ is relatively weakly compact. Let $(\omega^{h_\alpha})_{\alpha \in I}$ be a weakly convergent subnet of $(\omega^{h_n})_{n \geqslant 1}$ converging weakly to ω''. By w^* upper semicontinuity and the inequalities above, for each n,

$$\begin{aligned}\mathrm{ent}_{\mathscr{A}}(\omega'' \mid \omega) - \beta\omega''(h_n) &\geqslant \limsup_{\alpha \in I} [\mathrm{ent}_{\mathscr{A}}(\omega^{h_\alpha} \mid \omega) - \beta\omega^{h_\alpha}(h_n)] \\ &\geqslant \mathrm{ent}_{\mathscr{A}}(\omega^h \mid \omega) - \beta\omega^h(h)\end{aligned}$$

so, as $\omega''(h_n) \to \omega''(h)$,

$$\mathrm{ent}_{\mathscr{A}}(\omega'' \mid \omega) - \beta\omega''(h) \geqslant \mathrm{ent}_{\mathscr{A}}(\omega^h \mid \omega) - \beta\omega^h(h)$$

By uniqueness of ω^h, $\omega^h = \omega''$ and the result is proved. ∎

ACKNOWLEDGMENTS

I am grateful to S. F. Edwards and G. L. Sewell for useful conversations, and to the latter in particular for challenging me to produce a physical application of Ref. 3 other than to my somewhat unconventional work on the interpretation of quantum theory. I am also grateful to the Leverhulme Trust for financial support.

REFERENCES

1. M. J. Donald, *Commun. Math. Phys.* **105**:13 (1986).
2. O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol. II (Springer-Verlag, New York, 1981).
3. M. J. Donald, *Math. Proc. Camb. Phil. Soc.* **101**:363 (1987).
4. H. Araki, *Commun. Math. Phys.* **44**:1 (1975).
5. D. Petz, *Commun. Math. Phys.* **105**:123 (1986).
6. M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Volume III, *Scattering Theory* (Academic Press, New York, 1979).
7. H. Araki, *Publ. Res. Inst. Math. Sci. (Kyoto)* **9**:165 (1973).