PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 93, Number 3, March 1985

ENTROPY INCREASE AS A CONSEQUENCE OF MEASURE INVARIANCE

C. C. BROWN

ABSTRACT. An inequality, used in statistical mechanics for proving that entropy does not decrease, is shown to hold for general σ-finite measure spaces. We comment briefly on the corresponding Hilbert space result.

Let $(\Omega, \mathfrak{A}, \mu)$ be a σ-finite measure space, $\mathscr{F}_\mu$ the family of μ-absolutely continuous measures on $\mathfrak{A}$, and $T\colon \mathscr{F}_\mu \to \mathscr{F}_\mu$ a positive linear and monotone continuous mapping; i.e. $\sup_i T(\nu_i) = T(\sup_i \nu_i)$ for every monotone sequence, $\nu_1 \leqslant \nu_2 \leqslant \cdots$, of elements of $\mathscr{F}_\mu$. If T preserves total measure, then every probability measure in $\mathscr{F}_\mu$ maps into a probability measure under T. If ρ is a probability density with respect to $(\Omega, \mathfrak{A}, \mu)$, then the image of the corresponding probability distribution under T will have a density ρ'. A question of some interest in statistical mechanics concerns the behavior of integrals of the form $\int \phi(\rho(\omega))\, \mu(d\omega)$. Suppose ϕ is a convex function defined in the nonnegative reals. Under what conditions is the inequality

$$\int \phi(\rho(\omega))\, \mu(d\omega) \geqslant \int \phi(\rho'(\omega))\, \mu(d\omega)$$

valid? The main case of interest for statistical mechanics is where $\phi(x) = x \log x$ $(x \geqslant 0)$, in which the integral corresponds to the Gibbs entropy [**13**].

A sufficient condition for the validity of the inequality is essentially the invariance of μ under T. For finite μ on a countable space Ω, this result can be found in the book of Penrose [**12**]. The result appears to be attributable to M. J. Klein [**10**]. A similar result has also been proved for finite μ on a compact Ω by J. Voigt [**15**, Lemma 1.4].

Using an argument that is well known in the theory of probability in connection with the representation of conditional expectations, it is possible to prove the inequality for general σ-finite spaces. Slightly more generally, let $(\Omega, \mathfrak{A}, \mu)$ and $(\overline{\Omega}, \overline{\mathfrak{A}}, \overline{\mu})$ be σ-finite measure spaces, and T a monotone continuous positive linear mapping from $\mathscr{F}_\mu$ into the set of measures on $\overline{\mathfrak{A}}$. If $T\mu = \overline{\mu}$, then $T\nu$ is $\overline{\mu}$-absolutely continuous for every $\nu \in \mathscr{F}_\mu$. If $T\nu(\overline{\Omega}) = \nu(\Omega)$ for every $\nu \in \mathscr{F}_\mu$, then let ρ be a probability density for ν and ρ' a probability density for $T\nu$.

Received by the editors January 5, 1984 and, in revised form, May 1, 1984.
1980 *Mathematics Subject Classification*. Primary 28D15; Secondary 82A05, 82A15.
Key words and phrases. Entropy, measure preserving maps.

THEOREM. *For the mapping T as above, let ρ and ρ' be the probability densities just defined, and let ϕ be a finite-valued convex function defined on* $[0, \infty)$. *If the integral, $H(\rho) := \int \phi(\rho(\omega))\, \mu(d\omega)$, exists and is less than infinity, then the integral, $\overline{H}(\rho') := \int \phi(\rho'(\overline{\omega}))\, \overline{\mu}(d\overline{\omega})$, exists and is less than infinity. If both integrals exist in the extended sense, then $H(\rho) \geqslant \overline{H}(\rho')$.*

PROOF. Let $\mathscr{L}$ be the set of extended valued nonnegative measurable functions defined with respect to $(\Omega, \mathfrak{A})$, and define $\overline{\mathscr{L}}$ correspondingly for $(\overline{\Omega}, \overline{\mathfrak{A}})$. For $g \in \mathscr{L}$ let $\nu_g \in \mathscr{F}_\mu$ be the measure on $\mathfrak{A}$ having the density g with respect to μ. Denoting by $\mathscr{L}_\mu$ the equivalence classes mod μ of elements of $\mathscr{L}$, the correspondence $g \to dT(\nu_g)/d\overline{\mu}$ defines a positive linear and monotone continuous mapping T' from $\mathscr{L}$ into $\overline{\mathscr{L}}_{\overline{\mu}}$, the set of $\overline{\mu}$-equivalence classes in $\overline{\mathscr{L}}$. T' is also a mapping of the same type from $\mathscr{L}_\mu$ into $\overline{\mathscr{L}}_{\overline{\mu}}$. For $E \in \mathfrak{A}$ let I_E be the indicator function of E. $T'(I_E)$ is represented by a measurable function of $\overline{\omega} \in \overline{\Omega}$, and $I_{\overline{\Omega}} \in T'(I_\Omega)$. Using this fact and an argument that is nearly the same as that in Breiman [**2**, Chapter 4] or Doob [**3**, Chapter I] for proving the existence of conditional probability kernels, if f is a measurable mapping from Ω into a Borel space $(B, \mathfrak{B})$, then we can construct a stochastic kernel K_f from $(\overline{\Omega}, \overline{\mathfrak{A}})$ to $(B, \mathfrak{B})$ with the property

$$K_f(F, \cdot) \in T'(I_F \circ f) \qquad (F \in \mathfrak{B}).$$

As in the conditional expectation case, one has

$$\int \phi(y) K_f(dy, \cdot) \in T'(\phi \circ f)$$

for every nonnegative extended valued measurable function ϕ defined on $(B, \mathfrak{B})$. Consequently,

$$\rho' := \int |y| K_\rho(dy, \cdot)$$

is a density for $T\nu_\rho$. Since ρ' is $\overline{\mu}$-almost everywhere finite, we can suppose that K_ρ has been so chosen that ρ' is everywhere finite. Thus $\int |y|\, K_\rho(dy, \cdot\,)$ is finite valued and the conditions for an application of Jensen's inequality hold. For the convex function $\phi_+(y) := \max[\phi(y), 0]$ $(y \in [0, \infty))$, one has

$$\phi_+(\rho') = \phi_+\left(\int y K_\rho(dy, \cdot)\right) \leqslant \int \phi_+(y) K_\rho(dy, \cdot)$$

and

$$\int \phi_+(\rho')\, \overline{\mu}(d\overline{\omega}) \leqslant \int \left\{ \int \phi_+(y) K_\rho(dy, \overline{\omega}) \right\} \overline{\mu}(d\overline{\omega})$$

$$= \int T'(\phi_+ \circ \rho)\, \overline{\mu}(d\overline{\omega}) = \int \phi_+(\rho)\, \mu(d\omega).$$

Repeating this argument with ϕ in place of ϕ_+, a proof of the theorem as stated is only a matter of technical details.

For a classical mechanical system, let $(\Omega, \mathfrak{A}, \mu)$ be the phase space with the μ-preserving family $\{T_t\}_{t \in R}$ of one-to-one surjective transformations $T_t\colon \Omega \to \Omega$

representing the system flow in Ω. It is often possible to assume that μ is σ-finite on the σ-algebra $\mathscr{T}$ of invariant measurable sets in Ω (see, for example, [**9**]). It follows [**9**] that every initial probability density ρ_0 in $(\Omega, \mathfrak{A}, \mu)$ has a time average which, in any reasonable definition of approach to equilibrium with time, is equal almost everywhere to the equilibrium density ρ_∞ corresponding to ρ_0. The density ρ_∞ is a conditional expectation $E(\rho_0|\mathscr{T})$ of ρ_0 under μ with respect to $\mathscr{T}$. Because μ is σ-finite on $\mathscr{T}$, $E(f|\mathscr{T})$ is also defined for every nonnegative measurable extended valued function f on Ω. The time average is therefore a restriction of the mapping T' from $\mathscr{L} \to \mathscr{L}_\mu$ which derives from the mapping T: $\mathscr{F}_\mu \to \mathscr{F}_\mu$ given by $\nu_g \to \nu_{E(g|\mathscr{T})}$ $(g \in \mathscr{L})$. This mapping T has the properties demanded by the theorem and is the only such mapping specializing to the average value map.

For quantum mechanical systems, a corresponding result has been proved by G. Lindblad [**11**]. The following easy consequence of Jensen's inequality seems to simplify Lindblad's proof considerably.

THEOREM. *Let $U \subset R$ be an interval, ϕ a bounded convex function on U, $A_1, A_2, \ldots$ nonegative selfadjoint operators in a separable Hilbert space $\mathscr{H}$, with $A_1 + A_2 + \cdots = I =$ Identity operator, and $x_1, x_2, \ldots$ elements of U. If $\sum_\alpha |x_\alpha| \operatorname{tr} A_\alpha < \infty$ then $\phi(\sum_\alpha x_\alpha A_\alpha)$ is defined as a bounded self adjoint operator in $\mathscr{H}$. If, furthermore,* $\operatorname{tr}[\sum_\alpha \phi(x_\alpha) A_\alpha]$ *exists less than infinity, then*

$$\operatorname{tr}\left[\phi\left(\sum_\alpha x_\alpha A_\alpha\right)\right] \leqslant \operatorname{tr}\left[\sum_\alpha \phi(x_\alpha) A_\alpha\right].$$

REFERENCES

1. H. Bauer, *Wahrscheinlichkeitstheorie und Grudnzüge der Masstheorie*, De Gruyter, Berlin, 1978.

2. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass., 1968.

3. J. L. Doob, *Stochastic processes*, Wiley, London, 1953.

4. P. und T. Ehrenfest, *Zur Theorie der Entropiezunahme in der Statistischen Mechanik von Gibbs*, Wien. Ber. **115** (1906).

5. H. Grad, *Principles of the kinetic theory of gases*, Handbuch der Physik, Springer, Berlin and New York, 1958.

6. K. Jacobs, Lecture notes on ergodic theory, Aarhus Univ., Aarhus, 1963.

7. M. Kac, et al., *Probability and related topics in physical sciences*, Lectures in Appl. Math. (Proc. Summer Seminar, Boulder, Colo., 1957), Vol. I, Interscience, London and New York, 1959.

8. ______, *Foundations of kinetic theory*, Proc. Third Berkeley Sympos., Univ. of California Press, Berkeley, Vol. 3, 1954.

9. A. I. Khinchin, *Mathematical foundations of statistical mechanics*, Dover, New York, 1949.

10. M. J. Klein, *Entropy and the Ehrenfest urn model*, Physica **22** (1956), 569–575.

11. G. Lindblad, *Expectations and entropy inequalities for quantum systems*, Comm. Math. Phys. **39** (1974), 111–120.

12. O. Penrose, *Foundations of statistical mechanics*, Pergamon Press, London, 1970.

13. R. C. Tolman, *The principles of statistical mechanics*, Oxford Univ. Press, 1938.

14. G. E. Uhlenbeck and G. W. Ford, *Lectures in statistical mechanics*, Lectures in Appl. Math., Vol. 1, Amer. Math. Soc., Providence, R. I., 1963.

15. J. I. Voigt, *Stochastic operators, information and entropy*, Comm. Math. Phys. **81** (1981), 31–38.

16. A. Wehrl, *General properties of entropy*, Rev. Modern Phys. **50** (1978), 221–260.

MATHEMATICS INSTITUT, FREIE UNIVERSITÄT BERLIN, ARNIMALLEE 2-6, BERLIN 33, WEST GERMANY