

## SOME CONVEXITY AND SUBADDITIVITY PROPERTIES OF ENTROPY

BY ELLIOTT H. LIEB<sup>1</sup>

**I. Introduction.** Statistical mechanics is the science of explaining, predicting and understanding the gross, macroscopic attributes of matter (which may be taken to mean mechanical systems with essentially an infinite number of degrees of freedom) in terms of the elementary dynamical laws governing its atomic constituents. The problems that arise are sufficiently complex and intriguing, but at the same time sufficiently well posed, that the subject is nowadays as much a part of mathematics as of physics. The fields of information theory and ergodic theory had their genesis in statistical mechanical modes of thought and are now well established in the mathematics literature; there will be more to come.

Ludwig Boltzmann, who died in 1906, was one of the principal founders of statistical mechanics, and his monument in Vienna contains the following eloquent testimonial to his scientific creativity:

$$(1) \quad S = k \log W.$$

Surely, this hypothesis of Boltzmann [1] is one of the most important and daring in statistical mechanics, for it relates  $S$ , the macroscopic entropy of a system, to  $W$ , the number of microscopic states of the system which have the same, given macroscopic properties. The number  $k$  is a universal constant, called Boltzmann's constant, and, for our purposes, we can consider it to be 1.

In these lectures we shall explore some of the abstract properties of entropy, after first giving a precise formulation of it, and will include some recent results (with M. B. Ruskai) which extend formerly known facts about the strong subadditivity of entropy from the domain of classical mechanics to the quantum-mechanical domain. The presentation here will be sketchy and the reader is referred to the original papers [3], [4], [5] for more details.

---

An expanded version of an invited address delivered before the M.I.T. meeting of the Society on October 27, 1973 by invitation of the Committee to Select Hour Speakers for Eastern Sectional Meetings; received by the editors January 7, 1974.

*AMS (MOS) subject classifications* (1970). Primary 80A10, 81A81, 82A05, 82A15, 94A15; Secondary 15A45, 28A35, 28A65, 47A99.

*Key words and phrases.* Entropy, strong subadditivity, convexity, density matrix.

<sup>1</sup> Work supported by National Science Foundation Grant GP 31674 X.

**II. Definitions of entropy.** First we shall define abstractly what we mean by entropy in the *classical discrete case*. Let  $\rho$  denote a probability measure on an atomic probability space whose points are labelled by  $i \in \mathcal{N}$ . Hence  $\rho(i) \in [0, 1]$  denotes the probability that event  $i$  occurs and  $\sum_{i=1}^{\infty} \rho(i) = 1$ . The entropy of  $\rho$  is defined by

$$(2) \quad S(\rho) = -\sum_{i=1}^{\infty} \rho(i) \ln \rho(i)$$

with  $0 \ln 0 \equiv 0$ . As each term in the sum is nonpositive,  $S$  is well defined, although it may be  $+\infty$ . Obviously,  $S$  measures the extent to which  $\rho$  is "chaotic" or "spread out": If  $\rho$  is concentrated on one point (complete certainty) then  $S=0$ ; If  $\rho=1/W$  on  $W$  points and 0 otherwise, then  $S=\ln W$ . This last observation establishes the connection between (2) and (1). Clearly there are other functions besides  $\rho \rightarrow -\rho \ln \rho$  which have the same qualitative property, but  $-\rho \ln \rho$  alone has an important additivity property (additivity of entropy for independent systems) which we shall explain later (cf. equation (16)).

To establish contact with information theory we can define

$$(3) \quad I(\rho) = -S(\rho)$$

to be the information content of  $\rho$  (Shannon). The idea behind (3) is the following: Think of the index  $i \in \mathcal{N}$  as denoting possible states of a system which is in some definite state  $j$  unknown to us. Interpret  $\rho(i)$  as an assertion of a priori belief that the system is in the state  $i$ . Then, after we measure the system and find it to be in the state  $j$ , the new probability function is  $\tilde{\rho}(i) = \delta_{i,j}$ ,  $i \in \mathcal{N}$  (Kronecker delta) and  $S(\tilde{\rho}) = I(\tilde{\rho}) = 0$ . Thus, our knowledge (information) has increased by  $S(\rho)$  and the entropy of the system has decreased by  $S(\rho)$ . For this reason, it is sometimes said that information is negative entropy. While such an assertion is true by definition (3), it is a matter of dispute whether it has any true physical import.

A generalization of (2) is the *classical continuous case* in which the underlying measure space,  $\Omega$ , is not atomic and is equipped with a positive measure  $d\mu(x)$  (not necessarily finite) and  $\rho(x) \geq 0$  is a probability density. Thus  $\int \rho(x) d\mu(x) = 1$  and

$$(4) \quad S(\rho) = -\int \rho(x) \ln \rho(x) d\mu(x).$$

A typical example in statistical mechanics is an  $N$  particle system with a Hamiltonian function  $H(p, q)$ ,  $p \in \mathbf{R}^{3N}$ ,  $q \in \mathbf{R}^{3N}$ , and  $d\mu$  is Lebesgue measure on some subset  $\Omega$ , of  $\mathbf{R}^{3N} \times \mathbf{R}^{3N}$ . Then

$$\rho(p, q) = Z^{-1} \exp[-\beta H(p, q)],$$

where

$$(5) \quad Z = \int_{\Omega} \exp[-\beta H(p, q)] d\mu(p, q),$$

and  $\beta = (kT)^{-1}$  with  $T$  being the temperature.

Our third definition is the *quantum-mechanical case*. Instead of a measure space, one has a separable Hilbert space  $\mathcal{H}$  and  $\rho$  is a positive trace-class operator on  $\mathcal{H}$  (i.e.  $\rho$  is selfadjoint and  $(x, \rho x) \geq 0, \forall x \in \mathcal{H}$ ) with  $\text{Tr } \rho = 1$ , where  $\text{Tr}$  is the trace. Such an operator  $\rho$  is called a *density matrix*. Then

$$(6) \quad S(\rho) = -\text{Tr } \rho \ln \rho.$$

In a basis in which  $\rho$  is diagonal, (6) is seen to be identical to (2); the difference will manifest itself when we try to compare the entropies of two different  $\rho$ 's which do not commute with each other. In other words, equation (6) is the noncommutative version of equation (2). The typical statistical mechanics example is as in (5), except that  $H$  becomes a self-adjoint operator on  $\mathcal{H} = L^2(\mathbf{R}^{3N})$  and

$$(7) \quad \rho = Z^{-1} e^{-\beta H}, \quad Z = \text{Tr } e^{-\beta H}.$$

We remark in passing that entropy also plays a role in ergodic theory; given a measure preserving transformation  $T$  on a probability space  $\Omega$ , Kolmogorov and Sinai [2] have been able to define the entropy of  $T$  by making use of (2) in such a way that the entropy is invariant under isomorphism. There exists an analogous notion of a measure preserving transformation in a Hilbert space setting, but an unsolved problem is to define the analogue of the Kolmogorov-Sinai entropy. In other words, it is not clear how to give an unambiguous definition of the density matrix to use in (6). We shall say no more about ergodic theory in this lecture.

**III. Properties of entropy (one space).** We turn now to a study of some properties of  $S(\rho)$  that can be deduced from the definitions (2), (4) and (6). These properties are summarized in Table 1. The proofs we give will not only be sketchy but they will also assume that  $\mathcal{H}$  is finite dimensional in the quantum case. The proofs for the infinite dimensional case can be found in [4] and [5]. When we say that some property is false in some particular case, we mean, of course, that the property does not hold generally and we do not mean that the property never holds.

*Property A.*  $S(\rho) \geq 0$  (*positivity of entropy*). This property is easily seen to be true in the classical discrete and the quantum cases but is false in the classical continuous case. The difficulty cannot be mitigated by adding a positive constant to the right side of (4) because  $S(\rho)$  has no lower bound in the continuous case. Indeed, in the example (5) one sees that if  $H$  is nonconstant and has a unique minimum then as  $\beta \rightarrow \infty$ ,  $S \rightarrow -\infty$ . Thus,

if one believes, as Boltzmann did, that the entropy as we have defined it is the same (apart from an additive constant, possibly) as the physical entropy of a mechanical system, and if the latter is required to be positive, then classical mechanics cannot be valid at very low temperatures. Quantum mechanics must eventually be invoked.

*Property B.*  $S(\rho)$  is concave in  $\rho$ . By this is meant that if  $\rho = \alpha\rho' + (1-\alpha)\rho''$ ,  $0 \leq \alpha \leq 1$ , and  $\rho', \rho'' \geq 0$ , then

$$(8) \quad S(\rho) \geq \alpha S(\rho') + (1-\alpha)S(\rho'').$$

The physical interpretation is obvious: if two probability ensembles are mixed (which is not to be confused with the notion of mixing two systems, to be defined later) the entropy increases. This essential property of entropy is true in all three cases. In the classical cases the proof uses Jensen's inequality (i.e.  $\int e^{A+B} d\mu \geq \int e^A d\mu \exp\{\int B e^A d\mu / \int e^A d\mu\}$  for  $A$  and  $B$  real-valued functions). The analogous inequality in the Hilbert space case is the Peierls-Bogoliubov inequality:

$$(9) \quad \text{Tr } e^{A+B} \geq (\text{Tr } e^A) \exp\{\text{Tr } B e^A / \text{Tr } e^A\}$$

for  $A$  and  $B$  selfadjoint. To apply (9) assume that  $\rho'$  and  $\rho''$  are strictly positive and write

$$(10) \quad \begin{aligned} \Delta &\equiv -S(\rho) + \alpha S(\rho') + (1-\alpha)S(\rho'') = \alpha\Delta' + (1-\alpha)\Delta'', \\ \Delta' &\equiv \text{Tr } \rho' \{\ln \rho - \ln \rho'\}, \\ \Delta'' &\equiv \text{Tr } \rho'' \{\ln \rho - \ln \rho''\}. \end{aligned}$$

Then, using (9) with  $A = \ln \rho'$ ,  $B = \ln \rho - \ln \rho'$

$$e^{\Delta'} \leq \text{Tr } \exp\{\ln \rho' + \ln \rho - \ln \rho'\} = \text{Tr } \rho = 1.$$

Thus,  $\Delta' \leq 0$ . Likewise,  $\Delta'' \leq 0$ . Q.E.D. The case of  $\rho', \rho''$  semidefinite can be handled by a continuity argument, and the classical cases can be proved in the same way using Jensen's inequality.

**IV. Properties of entropy (two spaces).** The remaining properties refer to the "mixture of different systems". In the classical cases this means that the total probability space,  $\Omega_{12}$ , is taken to be the Cartesian product of two smaller spaces

$$(11) \quad \Omega_{12} = \Omega_1 \times \Omega_2$$

and the product measure is assumed. Given a positive function  $\rho_{12}$  on  $\Omega_{12}$  one can define a positive function  $\rho_1$  on  $\Omega_1$  by

$$(12) \quad \rho_1(x) = \int \rho_{12}(x, y) d\mu_2(y).$$

Similarly, one defines  $\rho_2$  on  $\Omega_2$ .

In the quantum case, the Hilbert space  $\mathcal{H}_{12}$  is taken to be the tensor product of two smaller spaces:

$$(13) \quad \mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2.$$

Given a positive, trace-class operator  $\rho_{12}$  on  $\mathcal{H}_{12}$ , one defines  $\rho_1$  to be a positive, trace class operator on  $\mathcal{H}_1$  by the partial trace operation:

$$(14) \quad \rho_1 \equiv \text{Tr}_2 \rho_{12}.$$

This means that

$$(15) \quad (x, \rho_1 y) \equiv \sum_{j=1}^{\infty} (x \otimes e_j, \rho_{12}(y \otimes e_j))$$

for all  $x, y \in \mathcal{H}_1$ , and where  $\{e_j\}_{j=1}^{\infty}$  is an orthonormal basis for  $\mathcal{H}_2$ . It is easy to prove that the right side of (15) is basis independent. In all cases, note that if  $\rho_{12}$  is normalized, then so is  $\rho_1$ .

Thus, for a mixed system one has three  $\rho$ 's to consider,  $\rho_1$ ,  $\rho_2$  and  $\rho_{12}$  on three spaces. Corresponding to these there are three entropies:  $S_1 = S(\rho_1)$ ,  $S_2 = S(\rho_2)$  and  $S_{12} = S(\rho_{12})$ . The two spaces,  $\Omega_1$  and  $\Omega_2$  (or  $\mathcal{H}_1$  and  $\mathcal{H}_2$ ), can have two different physical interpretations (or both at once):

(i)  $\Omega_1$  and  $\Omega_2$  refer to two different sets of degrees of freedom of one physical system (e.g. the  $p$  and  $q$  variables in (5), in which case  $\Omega_1 \subset \mathbb{R}^{3N}$  and  $\Omega_2 \subset \mathbb{R}^{3N}$ );

(ii)  $\Omega_1$  and  $\Omega_2$  refer to two different physical systems (e.g. the molecules of two different gases) which, under the product, are being thought of as one system.

One says that  $\Omega_1$  and  $\Omega_2$  (resp.  $\mathcal{H}_1$  and  $\mathcal{H}_2$ ) are independent if  $\rho_{12}(x, y) = \rho_1(x)\rho_2(y)$  (resp.  $\rho_{12} = \rho_1 \otimes \rho_2$ ). In this case it is easy to check the *additivity property*

$$(16) \quad S_{12} = S_1 + S_2.$$

*Property C.*  $S_{12} \geq S_1$  (*monotonicity of entropy*). This appealing statement is true only in the classical discrete case. It cannot be true generally in the classical continuous case because one could have  $\Omega_1$  and  $\Omega_2$  independent, so that (16) holds, and at the same time one could have  $S_2 < 0$  because of the failure of property A. In the quantum case we note that  $S(\rho) = 0$  if  $\rho$  is a *pure state*, i.e.  $\rho$  is a one-dimensional orthogonal projection on  $\mathcal{H}$ . (In the classical discrete case, the analogous statement is that  $\rho$  is concentrated on one point.) Conversely,  $S(\rho) = 0$  implies that  $\rho$  is a pure state. Take  $\rho_{12}$  to be pure, in which case  $\rho_1$  will not, in general, be pure in the quantum case. Then  $S_{12} = 0$  and  $S_1 > 0$ . The proof of C for the classical discrete case is as follows: Denote the function  $\rho_{12}$  on  $\mathcal{N} \times \mathcal{N}$  by  $\rho(i, j)$  and  $\rho_1$  on  $\mathcal{N}$  by  $\rho(i)$ ,  $i, j \in \mathcal{N}$ .  $\rho(i) = \sum_{j=1}^{\infty} \rho(i, j)$ . Define

$$\Delta \equiv S_1 - S_{12} = \sum_{i,j=1}^{\infty} \rho(i, j) (\ln \rho(i, j) - \ln \rho(i)).$$

Assume that  $\rho(i) > 0, \forall i$ . Then, by Jensen's inequality,

$$e^\Delta \leq \sum_{i,j} \frac{\rho(i,j)^2}{\rho(i)} \leq 1,$$

since  $\rho(i) \geq \rho(i,j), i, j \in \mathcal{N}$ . The general case follows by a continuity argument. Q.E.D.

The failure of C to hold presents a serious, but not insoluble problem for physics. It would mean, for example, that, under the second interpretation (ii) of  $\rho_{12}$ , the entropy of our planet could increase without limit while the entropy of the universe remains zero. Property E below, which holds in all cases, is partial compensation for the failure of C, but it is not enough. Instead, the resolution of the dilemma comes from further hypotheses about the kinds of  $\rho_{12}$ 's that actually occur in physical systems. In particular there are theorems that state that (in all three cases) when systems 1 and 2 are "large enough" then (16) is approximately true. In quantum mechanics  $S_2 > 0$ , and so the situation is saved—at least on the macroscopic level. More precisely, for macroscopic systems,  $S_1, S_2$ , and  $S_{12}$  are proportional to the volumes of the respective systems, whereas the error in (16),  $S_{12} - (S_1 + S_2)$ , is proportional to the area of the surface separating systems 1 and 2.

*Property D.*  $S_{12} \leq S_1 + S_2$  (*subadditivity of entropy*). This is one of the crucial facts about entropy and, fortunately, holds in all three cases. The proof is similar to that of property B: if  $\Delta \equiv S_{12} - S_1 - S_2$  then, from (9),  $e^\Delta \leq \text{Tr}_{12} \rho_1 \otimes \rho_2 = 1$ .

We now interject two technical lemmas that are relevant to the quantum case.

**LEMMA 1.** *Let  $\rho_{12}$  be a pure state on  $\mathcal{H}_{12}$ . Then the spectra of  $\rho_1$  and  $\rho_2$  on  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , including multiplicity, are identical except, possibly, for the point zero. In particular,  $S_1 = S_2$ .*

The proof of Lemma 1 is easy and can be found in [6], among other places.

**LEMMA 2.** *Given a positive, trace class operator  $\rho_1$  on a separable Hilbert space  $\mathcal{H}_1$ , there exist a separable Hilbert space  $\mathcal{H}_2$  and a pure state  $\rho_{12}$  on  $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$  such that  $\rho_1 = \text{Tr}_2 \rho_{12}$ .*

Again, the easy proof can be found in [6].

*Property E.*  $S_{12} \geq |S_1 - S_2|$  (*triangle inequality*). We call this the triangle inequality because, when it is combined with D one has

$$(17) \quad S_1 + S_2 \geq S_{12} \geq |S_1 - S_2|.$$

However, E is not true in the classical continuous case (because  $S_{12}$  can

be negative), but it is true in the other cases. To prove E in the quantum case we use Lemma 2 to find a Hilbert space  $\mathcal{H}_3$  and a pure state  $\rho_{123}$  on  $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$  such that  $\rho_{12} = \text{Tr}_3 \rho_{123}$ , and we define  $\rho_3 = \text{Tr}_{12} \rho_{123}$  (resp.  $\rho_{23} = \text{Tr}_1 \rho_{123}$ ) on  $\mathcal{H}_3$  (resp.  $\mathcal{H}_2 \otimes \mathcal{H}_3$ ). But then, by Lemma 1,  $S_{12} = S_3 \equiv S(\rho_3)$ ,  $S_1 = S_{23} \equiv S(\rho_{23})$  and, by property D,

$$(18) \quad -S_1 + S_{12} + S_2 = -S_{23} + S_3 + S_2 \geq 0 \quad \text{Q.E.D.}$$

Finally, while there is no natural analogue of Lemmas 1 and 2 in the classical discrete case, property E is nevertheless true there as well. This is so because  $S_{12} \geq S_1 - S_2$  holds in the special case that  $\rho_{12}$  commutes with  $\rho_1 \otimes I_2$  and  $I_1 \otimes \rho_2$ ; but this special case is precisely the classical discrete case, provided one thinks of the function  $\rho_{12}$  on  $\mathcal{N} \times \mathcal{N}$  in the obvious way as a diagonal matrix on  $I_1 \otimes I_1$ . There exists, in fact, a direct proof of property E for the classical discrete case similar to the proof of property C, but the foregoing detour through the quantum domain is more amusing.

**V. Properties of entropy (three spaces).** Up to this point we have been concerned with the product of two spaces. Now the plot thickens; we consider three spaces and the property of *strong subadditivity* and its variants. Given  $\rho_{123}$  on  $\Omega_1 \times \Omega_2 \times \Omega_3$  (or  $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ ) we can, by taking partial traces, define  $S_{123} = S(\rho_{123})$ ,  $S_{12} = S(\rho_{12})$ ,  $S_1 = S(\rho_1)$ , etc. We list three properties that are closely related (note that  $G$  refers to only two spaces).

*Property F.*  $S_{123} + S_2 \leq S_{12} + S_{23}$  (*strong subadditivity of entropy*).

*Property G.*  $S_{12} - S_1$  is concave in  $\rho_{12}$ .

*Property H.*  $S_1 + S_2 \leq S_{13} + S_{23}$ .

The significance of properties F, G and H will be discussed after Table 1 is completed. At this point we shall merely indicate the proofs of those three properties. Consider G first. In order to simplify the notation, we shall use  $\text{Tr}$  to mean  $\sum$  or  $\int$  in the classical cases. If we use the Jensen or Peierls-Bogoliubov inequality as in (10), we end up trying to prove that

$$f \equiv \alpha \text{Tr}_{12} \exp(K + \ln \rho'_1) + (1 - \alpha) \text{Tr}_{12} \exp(K + \ln \rho''_1) \leq 1,$$

with  $K = \ln \rho_{12} - \ln \rho_1$  and with an abuse of notation in which  $\rho_1$  stands for  $\rho_1 \otimes I_2$ . In the two classical cases  $f = 1$  since  $\rho_1 = \alpha \rho'_1 + (1 - \alpha) \rho''_1$ . For the quantum case we need a lemma [4]:

**LEMMA 3.** *Let  $K$  be any fixed selfadjoint  $n \times n$  matrix. Then the map from the positive  $n \times n$  matrices to the reals defined by  $A \rightarrow \text{Tr} \exp(K + \ln A)$  is concave.*

The proof [4] of Lemma 3 is lengthy, but we can use Lemma 3 to obtain

$$f \leq \text{Tr}_{12} \exp(K + \ln \rho_1) = 1.$$

Next we consider property H. Since  $\rho_{13}$  is linear in  $\rho_{123}$ , property G implies that  $S_{13}-S_1$  is concave in  $\rho_{123}$ . Likewise,  $S_{23}-S_2$  is concave in  $\rho_{123}$ . Thus  $\Delta \equiv (S_{13}-S_1)+(S_{23}-S_2)$  is concave in  $\rho_{123}$ . In the classical discrete and quantum cases,  $\Delta$  will have its minimum on extremal states, which are pure states. In that case (see Lemma 1)  $S_{13}=S_2$  and  $S_{23}=S_1$ . In the classical continuous case this argument fails and, in fact, H is false because one could take  $\rho_{123}=\rho_{12}\rho_3$  with  $S_3<0$ .

Finally, we consider property F. By applying the Jensen or Peierls-Bogoliubov inequality, we have to prove that

$$\Delta \equiv \text{Tr}_{123} \exp(\ln \rho_{12} + \ln \rho_{23} - \ln \rho_2) \leq 1.$$

Clearly,  $\Delta=1$  in the two classical cases. In the quantum case an involved argument using Lemma 3 (see [5]) shows that  $\Delta \leq 1$ . However, there is another way to prove F in the quantum case and this method displays the close connection between properties F and H. As in the proof of property E, introduce a fourth Hilbert space  $\mathcal{H}_4$  and  $\rho_{1234}$  pure on  $\mathcal{H}_{123} \otimes \mathcal{H}_4$  such that  $\text{Tr}_4 \rho_{1234} = \rho_{123}$ . Then  $S_{123}+S_2-S_{12}-S_{23}=S_4+S_2-S_{12}-S_{14}$  (by Lemma 1) and this is nonpositive by property H.

REMARK. For the quantum case we have proved  $G \Rightarrow H \Rightarrow F$ . It is also true, however, that  $F \Rightarrow G$  as may be seen by a special choice of  $\rho_{123}$  (see [5]).

To complete Table 1 we have to discuss properties I, J, K and L which, as Table 1 shows, are always false and therefore uninteresting. They are mentioned for two reasons: The first is that it occasionally occurs to someone that these properties may be true. Secondly, we shall need to know their falsity in discussing properties  $F''$  and  $G''$  of Table 3.

With the definition  $\Delta = S_{123} - S_{12} - S_{13} - S_{23} + S_1 + S_2 + S_3$ , property I states that  $\Delta \leq 0$ , while property J states that  $\Delta \geq 0$ . We shall prove that these statements are false in the classical discrete case and, *a fortiori*, they are also false in the other two cases. Take  $\rho_{123}(i, j, k) = \rho_{12}(i, j)\delta_{j, k}$ , whence  $S_{123} = S_{12}$ ,  $S_{13} = S_{12}$  and  $S_{23} = S_2 = S_3$ . Then  $\Delta = -S_{12} + S_1 + S_2$  and this can be positive since  $\rho_{12}$  is arbitrary. To demonstrate the falsity of property J an explicit example is required. Let  $\rho_{123}(1, 1, 2) = \frac{1}{3}$ ,  $\rho_{123}(1, 2, 1) = \frac{1}{3}$ ,  $\rho_{123}(2, 1, 1) = \frac{1}{3}$  and  $\rho_{123}(i, j, k) = 0$  otherwise for  $i, j, k \in \mathcal{N}$ . Then  $S_{123} = S_{12} = S_{13} = S_{23} = \ln 3$ ,  $S_1 = S_2 = S_3 = \ln 3 - \frac{2}{3} \ln 2$  and  $\Delta = \ln 3 - 2 \ln 2 < 0$ .

Property K (resp. L) states that  $\Delta \equiv S_{12} - S_1 - S_2$  is concave (resp. convex). For K consider  $\rho_{12}$  of the form  $\rho_{12}(i, j) = \rho_1(i)\delta_{i, j}$ . Then  $\Delta = -S_1 = -S(\rho_1)$ , but, as  $\rho_1$  is arbitrary, one can have  $-S(\rho_1') - S(\rho_1'') > -2S(\frac{1}{2}\rho_1' + \frac{1}{2}\rho_1'')$  (property B). For L, take  $\rho_{12}'(i, j) = \frac{1}{2}\delta_{i, 1}\delta_{j, 1} + \frac{1}{2}\delta_{i, 2}\delta_{j, 2}$  and  $\rho_{12}''(i, j) = \frac{1}{2}\delta_{i, 2}\delta_{j, 1} + \frac{1}{2}\delta_{i, 1}\delta_{j, 2}$ , so that  $\Delta(\rho_{12}') = \Delta(\rho_{12}'') = -\ln 2$  and  $\Delta(\frac{1}{2}\rho_{12}' + \frac{1}{2}\rho_{12}'') = 0$ .



TABLE 1

	<i>Classical discrete</i>	<i>Classical continuous</i>	<i>Quantum</i>
A. $S(\rho) \geq 0$	T	F	T
B. $S(\rho)$ concave in $\rho$	T	T	T
C. $S_{12} \geq S_1$	T	F	F
D. $S_{12} \leq S_1 + S_2$	T	T	T
E. $S_{12} \geq  S_1 - S_2 $	T	F	T
F. $S_{123} + S_2 \leq S_{12} + S_{23}$	T	T	T
G. $S_{12} - S_1$ concave in $\rho_{12}$	T	T	T
H. $S_1 + S_2 \leq S_{13} + S_{23}$	T	F	T
I. $S_{123} + S_1 + S_2 + S_3$ $\leq S_{12} + S_{13} + S_{23}$	F	F	F
J. $S_{123} + S_1 + S_2 + S_3$ $\geq S_{12} + S_{13} + S_{23}$	F	F	F
K. $S_{12} - S_1 - S_2$ concave in $\rho_{12}$	F	F	F
L. $S_{12} - S_1 - S_2$ convex in $\rho_{12}$	F	F	F

TABLE 1. Properties of entropy and their truth (T) or falsity (F) in the three cases.

Thus, Table 1 is complete and we now take up the interpretation of strong subadditivity. There are two. The first is simple but the second is involved and will lead us to the construction of two more tables.

**VI. First interpretation of strong subadditivity.** Property F can be viewed as a generalization of property D, i.e. subadditivity. For concreteness we use the measure space language, but the same idea is applicable in the Hilbert space setting. Suppose that to each Lebesgue measurable subset,  $\Lambda$ , of  $R^3$  we associate a probability space  $\Omega(\Lambda)$ , which can be thought of physically as the set of configurations of a "gas of particles" contained in  $\Lambda$ . We assume that when  $\Lambda_1$  and  $\Lambda_2$  are disjoint  $\Omega(\Lambda_1 \cup \Lambda_2) = \Omega(\Lambda_1) \times \Omega(\Lambda_2)$ . Given a probability density  $\rho_\Lambda$  on each  $\Omega(\Lambda)$  we then have an entropy  $S(\rho_\Lambda) \equiv S(\Lambda)$  for each  $\Lambda$ . Property D states that

$$(19) \quad S(\Lambda_1 \cup \Lambda_2) \leq S(\Lambda_1) + S(\Lambda_2)$$

when  $\Lambda_1$  and  $\Lambda_2$  are disjoint.

What if  $\Lambda_1$  and  $\Lambda_2$  are not disjoint? Define  $\Lambda_3 = \Lambda_1 \cap \Lambda_2$ ,  $\Lambda_4 = \Lambda_1 - \Lambda_3$  and  $\Lambda_5 = \Lambda_2 - \Lambda_3$ . Then, property F states that  $S_{345} + S_3 \leq S_{34} + S_{35}$ , or

$$(20) \quad S(\Lambda_1 \cup \Lambda_2) + S(\Lambda_1 \cap \Lambda_2) \leq S(\Lambda_1) + S(\Lambda_2),$$

which is an appropriate generalization of (19).

**VII. Second interpretation of strong subadditivity: relative entropy.**

Given the product of two spaces  $\mathcal{H}_1 \otimes \mathcal{H}_2$  (or  $\Omega_1 \times \Omega_2$ ) and a density matrix  $\rho_{12}$  on  $\mathcal{H}_{12}$  one can define

$$S(2 | 1) \equiv S_{12} - S_1$$

to be “the conditional entropy of 2 relative to 1”. The term is slightly misleading since  $S(2|1)$  depends upon  $\rho_{12}$  and not merely on  $\rho_2$ , but this is a minor flaw. The real question is whether or not the word entropy is justified in describing  $S(2|1)$ , which means that we have to check the whole list of properties A to H for  $S(2|1)$ . In order to do so, strong subadditivity will play a crucial role. It might appear at first glance that since property G involves three spaces and  $S(2|1)$  involves two spaces, we shall need a new theorem about four spaces and this will lead us into an endless hierarchy of more and more complex inequalities. This is not so. It turns out, remarkably enough, that we already have all the information we need.

There is, however, one complication. Properties C to H refer to two or more spaces and hence we can let either  $\mathcal{H}_2$  or  $\mathcal{H}_1$  be such a product. Thus, considering both possibilities we shall end up with two tables (Table 2 and Table 3 respectively). It will be seen that  $S(2|1)$  merits the appellation entropy in terms of expanding  $\mathcal{H}_2$  (Table 2) fairly well, but that in terms of expanding  $\mathcal{H}_1$  it fares poorly. In fact, in Table 3 there are two properties, C'' and H'', that are not only false but their opposites are

TABLE 2

	<i>Classical discrete</i>	<i>Classical continuous</i>	<i>Quantum</i>
A'. $S(2   1) \geq 0$	T	F	F
B'. $S(2   1)$ concave in $\rho_{12}$	T	T	T
C'. $S(23   1) \geq S(2   1)$	T	F	F
D'. $S(23   1) \leq S(2   1) + S(3   1)$	T	T	T
E'. $S(23   1) \geq  S(2   1) - S(3   1) $	T	F	F
F'. $S(234   1) + S(3   1) \leq S(23   1) + S(34   1)$	T	T	T
G'. $S(23   1) - S(2   1)$ concave in $\rho_{123}$	T	T	T
H'. $S(2   1) + S(3   1) \leq S(24   1) + S(34   1)$	T	F	F

TABLE 2. Properties of relative entropy,  $S(2|1)=S_{12}-S_1$ , with respect to the second space and their truth (T) or falsity (F).

true. These are indicated by  $\sim C''$  and  $\sim H''$ . The most important of these is  $\sim C''$ , namely for three spaces

$$(21) \quad S(2 | 13) \leq S(2 | 1).$$

The interpretation of inequality (21), and indeed of  $S(2|1)$  itself is obvious.  $S(2|1)$  is the incremental information gained by measuring the total system (12) as against merely measuring the subsystem (1). Inequality (21) states that this incremental information gain is less if one knows more to start with (namely (13) as against (1)).

TABLE 3

		<i>Classical discrete</i>	<i>Classical continuous</i>	<i>Quantum</i>
$\sim C''$ .	$S(2   13) \leq S(2   1)$	T	T	T
D''.	$S(2   13) \leq S(2   1) + S(2   3)$	T	F	T
E''.	$S(2   13)$ $\geq  S(2   1) - S(2   3) $	F	F	F
F''.	$S(2   134) + S(2   1)$ $\leq S(2   13) + S(2   14)$	F	F	F
$\sim F''$ .	$S(2   134) + S(2   1)$ $\geq S(2   13) + S(2   14)$	F	F	F
G''.	$S(2   13) - S(2   1)$ concave in $\rho_{123}$	F	F	F
$\sim G''$ .	$S(2   13) - S(2   1)$ convex in $\rho_{123}$	F	F	F
$\sim H''$ .	$S(2   1) + S(2   3)$ $\geq S(2   14) + S(2   34)$	T	T	T

TABLE 3. Properties of relative entropy,  $S(2|1) \equiv S_{12} - S_1$ , with respect to the first space and their truth (T) or falsity (F). The designation  $\sim$  indicates that a property is opposite to that in Table 1.

Having set forth Tables 2 and 3 we shall conclude with brief indications of how the entries in the two tables can be checked.

*Table 2.*

Properties A' and C' follow from property C.

Property B' follows from property G.

Property D' follows from property F.

Property E' is true in the classical discrete case because if we define  $\Delta \equiv -S_{123} - S_{13} + S_{12} + S_1$  then

$$e^\Delta \leq \sum_i \sum_j \sum_k \frac{\rho_{123}(i, j, k)^2 \rho_{13}(i, k)}{\rho_1(i) \rho_{12}(i, j)}.$$

But  $\rho_{13}(i, k) \leq \rho_1(i)$  and  $\rho_{123}(i, j, k) \leq \rho_{12}(i, j)$ , so  $e^\Delta \leq 1$ . In the quantum case, let  $\rho_{123}$  be a pure state so that  $S_{12} = S_3$  and  $\Delta = -S_{13} + S_1 + S_3$ . This can be positive (property D) since  $\rho_{13}$  is arbitrary. In the classical continuous case take  $\rho_{123} = \rho_1 \rho_2 \rho_3$  and  $S_3 < 0$ .

Property F' follows by applying property F to the three spaces  $\mathcal{H}_2, \mathcal{H}_1 \otimes \mathcal{H}_3$  and  $\mathcal{H}_4$  in place of  $\mathcal{H}_1, \mathcal{H}_2$  and  $\mathcal{H}_3$ .

Property G' follows from property G.

Property H' is true in the classical discrete case since  $S_{124} \geq S_{12}$  and  $S_{134} \geq S_{13}$ . It is false in the classical continuous case because one can take  $\rho_{1234} = \rho_1 \rho_2 \rho_3 \rho_4$  and  $S_4 < 0$ . In the quantum case, take  $\rho_{1234} = \rho_{14} \otimes \rho_2 \otimes \rho_3$  so that  $S_{12} + S_{13} - S_{124} - S_{134} = 2(S_1 - S_{14})$  and this can be positive (property C) since  $\rho_{14}$  is arbitrary.

Table 3.

Properties A'' and B'' (not shown) are respectively identical to A' and B'.

Property  $\sim C''$  follows from property F.

Property D'' is that  $\Delta = (S_{123} - S_{13} - S_{12} + S_1) - (S_{23} - S_3) \leq 0$ . This is true in the classical discrete case by properties F and C. In the classical continuous case it is false because one could take  $\rho_{123} = \rho_1 \rho_2 \rho_3$  with  $S_2 < 0$ . In the quantum case, use Lemmas 1 and 2: With  $\rho_{1234}$  pure on  $\mathcal{H}_{123} \otimes \mathcal{H}_4$ ,

$$\Delta = \frac{1}{2}(S_1 + S_4 - S_{13} - S_{34}) + \frac{1}{2}(S_1 + S_3 - S_{14} - S_{34}) + \frac{1}{2}(S_3 + S_4 - S_{14} - S_{13}) \leq 0$$

by property H.

Property E'' is that  $\Delta = (S_{123} - S_{13} - S_{12} + S_1) + (S_{23} - S_3) \geq 0$ . In the classical continuous case this is false for the same reason that property D'' is false. In the quantum case it is false because  $S(2|13)$  can be negative (see A'). In the classical discrete case take  $\rho_{123}(i, j, k) = \rho_{12}(i, j)\delta_{j,k}$ . Then  $S_{23} = S_3 = S_2$  and  $S_{123} = S_{13} = S_{12}$ , so  $\Delta = -S_{12} + S_1$  and this can be negative since  $\rho_{12}$  is arbitrary.

Property F'' and its contrary,  $\sim F''$ , are both false in all three cases. Take  $\rho_{1234} = \rho_1 \otimes \rho_{234}$  so that  $\Delta \equiv S(2|134) + S(2|1) - S(2|13) - S(2|14) = S_{234} - S_{34} - S_{23} - S_{24} + S_2 + S_3 + S_4$ . As  $\rho_{234}$  is arbitrary,  $\Delta$  can be positive or negative (by properties I and J of Table 1).

Property G'' is that  $\Delta \equiv S_{123} - S_{13} - S_{12} + S_1$  is a concave function of  $\rho_{123}$ . In the classical discrete case take  $\rho_{123}(i, j, k) = \rho_{12}(i, j)\delta_{j,k}$  so that  $S_{123} = S_{13}$  and  $\Delta = -S_{12} + S_1$ . If property G'' were true,  $-S_{12} + S_1$  would have to be a concave function of  $\rho_{12}$ , which is arbitrary, but the contrary is true (property G). To demonstrate the falsity of  $\sim G''$  in the classical discrete case, let  $\rho_{123} = \rho_1 \rho_{23}$  so that  $\Delta = S_{23} - S_2 - S_3$  would have to be convex in  $\rho_{23}$ . This is false by property L. Since properties G'' and  $\sim G''$

are false in the classical discrete case they are *a fortiori* false in the other two cases as well.

Our final task is to prove property  $\sim H''$ . This is easy to do since

$$\begin{aligned} S(2 | 1) + S(2 | 3) - S(2 | 14) - S(2 | 34) \\ = (-S_{124} + S_{12} + S_{14} - S_1) + (-S_{234} + S_{34} + S_{23} - S_3), \end{aligned}$$

and this is positive by property F.

ACKNOWLEDGEMENT. The author thanks Professor H. F. Weinberger for his careful reading of this manuscript and for suggesting several improvements and corrections.

#### REFERENCES

1. L. Boltzmann, *Ueber die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Waermetheorie und der Wahrscheinlichkeitsrechnung respektive den Saetzen ueber das Waermegleichgewicht*, Wiener Berichte **76** (1877), 373.
2. A. N. Kolmogorov, *A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces*, Dokl. Akad. Nauk SSSR **119** (1958), 861–864. (Russian) MR **21** #2035a; Ja. G. Sinai, *On the concept of entropy for a dynamic system*, Dokl. Akad. Nauk SSSR **124** (1959), 768–771. (Russian) MR **21** #2036a.
3. E. H. Lieb and M. B. Ruskai, *A fundamental property of quantum-mechanical entropy*, Phys. Rev. Lett. **30** (1973), 434–436.
4. E. H. Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Advances in Math. **11** (1973), 267–288; See also H. Epstein, *Remarks on two theorems of E. Lieb*, Comm. Math. Phys. **31** (1973), 317–325.
5. E. H. Lieb and M. B. Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, J. Mathematical Phys. **14** (1973), 1938–1941.
6. H. Araki and E. H. Lieb, *Entropy inequalities*, Comm. Math. Phys. **18** (1970), 160–170. MR **42** #1466.

*Current address:* Department of Mathematics and Physics, Princeton University, Princeton, New Jersey 08540

*Permanent address:* Department of Mathematics and Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139