General properties of entropy

Alfred Wehrl*

Institute for Theoretical Physics, University of Vienna, Vienna, Austria

It is rather paradoxical that, although entropy is one of the most important quantities in physics, its main properties are rarely listed in the usual textbooks on statistical mechanics. In this paper we try to fill this gap by discussing these properties, as, for instance, invariance, additivity, concavity, subadditivity, strong subadditivity, continuity, etc., in detail, with reference to their implications in statistical mechanics. In addition, we consider related concepts such as relative entropy, skew entropy, dynamical entropy, etc. Taking into account that statistical mechanics deals with large, essentially infinite systems, we finally will get a glimpse of systems with infinitely many degrees of freedom.

CONTENTS

INTRODUCTION

Entropy is the crucial concept of thermodynamics and statistical mechanics. Its sovereign role regarding the behavior of macroscopic systems was recognized about one century ago by Clausius, Kelvin, Maxwell, Boltzmann, and many others. Therefore it is very astonishing to note that most of the main features of entropy are widely unknown to physicists and that many problems in connection with entropy are still open or have been solved only in the last few years.

There may be two reasons for this. The first is that physicists usually are more interested in determining the entropy of a given specific system as a function of some parameters such as temperature, density, etc., rather than in considering general properties of entropy. The second is that the proofs of many of those properties are generally rather tricky and involved and require some, so-to-say, unusual mathematics, at least in the quantum-mechanical case.

Traditionally entropy is derived from phenomenological thermodynamical considerations based upon the second law of thermodynamics. This method does not seem to be very appropriate for a profound understanding

---

*Work supported in part by "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich," Projekt Nr. 3016.

of entropy and may frequently lead to rather obscure conceptions and to very speculative or even mystical ideas. (An example is the famous heat death.) However, it has to be stressed that the concept of entropy is not at all unclear but a very well defined one. Of course, a correct definition is only possible in the framework of quantum mechanics, whereas in classical mechanics entropy can only be introduced in a somewhat limited and artificial manner.

Admittedly entropy has an exceptional position among the physical quantities. For instance, it does not show up in the fundamental equations of motion, such as the Schrödinger equation. Its nature is rather, roughly speaking, a statistical or probabilistic one; entropy can be interpreted as a measure of the amount of chaos within a quantum-mechanical mixed state. However, entropy by no means has to be considered as an entirely new quantity going beyond the concepts of classical or quantum mechanics. This idea has been discussed frequently in the past and, from time to time, is even found in the present-day literature. Let me emphasize that for a description of entropy the usual concepts of quantum mechanics such as Hilbert space, wave function, observables, and density matrices are absolutely sufficient (Sec. 1.A).

Entropy relates macroscopic and microscopic aspects of nature and determines the behavior of macroscopic systems, i.e., real matter, in equilibrium (or close to equilibrium). Why this is true unfortunately is not yet understood in full detail, in spite of a century's efforts of thousands and thousands of physicists. There are many opinions and proposals for a solution to this problem; however, none of them seems to be completely satisfactory. Since there is an abundant literature on this topic, I will not, in this review, try to take account of all the results obtained so far, but will restrict myself to a few remarks only (Sec. I.B.).

What I rather want to do is to give a survey of the general properties of entropy, i.e., those properties that do not depend on certain specific systems but are generally true. This is the main content of Sec. II. Certainly some of these properties are well known whereas others seem to have escaped general attention, as, for instance, strong subadditivity. But all of them are very important and indispensable for, say, a correct treatment of the thermodynamic limit and various other problems. I have tried to indicate in several places what these properties are good for in physics, however. Sometimes this will be rather sketchy and I will outline the main ideas only and will have to refer to the original papers for a detailed treatment.

Besides entropy itself there are many other quantities related to it that are of interest, as, for instance, the relative entropy and several other concepts. They will be treated in Secs. III and IV. One thing should be said in this connection: there is a tremendous variety of entropylike quantities, especially in the classical case, and perhaps every month somebody invents a new one. Among all these "entropies" I have tried to select those that, in my opinion, are of some physical significance. Maybe my choice will be felt to be subjective.

Since statements in statistical mechanics are frequently true in the infinite limit only, one cannot dispense with an *ab initio* description of infinite systems. This will be done in the last section, but again, I can present only a very sketchy treatment because there are severe mathematical obstacles that require extensive studies and go beyond the scope of this review. But after all one cannot avoid this approach because such important properties as ergodicity, mixing, stability, etc., can (quantum-mechanically) only hold in *strictly* infinite systems.

As already mentioned, entropy can be considered as a measure of the amount of chaos, or, to what extent a density matrix can be considered as "mixed." In Sec. II.C an elaborate version of this concept of "mixedness" of a density matrix is presented. Since, on the other hand, entropy can also be regarded as a measure of the lack of information about a system (this is just another point of view of the preceding statement), it is also necessary to comment on the relation between (physical) entropy and information theory (Sec. II.G).

Of course, a few words also have to be said about the classical ensembles of statistical mechanics (Sec. I.C) as well as about the history of the subject (Sec. I.D). Again, this will be rather cursory because there exists a rich and excellent literature about all that.

I hope that the physics will not be hidden behind mathematical technicalities. At least I have tried to avoid this.

I. GENERALITIES

A. Definition of entropy

As already discussed in the introduction, entropy is different from most physical quantities. In quantum mechanics one has to distinguish between *observables* and *states*. Observables, like position, momentum, angular momentum, etc., are mathematically described by self-adjoint operators in Hilbert space. States—which generally are mixed—are characterized by a density matrix, say, ρ, i.e., a Hermitian operator, $\geqslant 0$, with trace $= 1$. The expectation value of an observable A in the state ρ is $\langle A\rangle = \mathrm{Tr}\rho A$.

Now entropy is not an observable; that means that there does not exist an operator with the property that its expectation value in some state would be its entropy. It is rather a function of a state. If the state is described by the density matrix ρ, its entropy is defined by

$$S(\rho) = -k_B \mathrm{Tr}\rho \ln\rho . \tag{1.1}$$

This formula is due to von Neumann (1927) and generalizes the classical expression of Boltzmann and Gibbs to quantum mechanics. [von Neumann's derivation is based on earlier arguments by Einstein (1914) and Szilard (1925)]. k_B is Boltzmann's constant $= 1.38 \times 10^{-16}$ erg/K. In what follows we will put it equal to 1 which corresponds to measuring the temperature in ergs instead of Kelvin; thus entropy becomes dimensionless. (Occasionally we will insert in the formula for $S(\rho)$ an arbitrary, compact, positive operator rather than a density matrix. The quantity thus obtained has, of course, no direct physical meaning.)[1]

Entropy is a well defined quantity, no matter what the kind or size of the system under consideration is. (This statement, however, has nothing to do with the question to what extent entropy is a *useful* quantity in physics.) It is always $\geqslant 0$, and, as we will see immediately, $=0$ exactly for the pure states, possibly $=+\infty$. (In a certain sense this latter possibility happens to be the usual case. Fortunately, this has no serious consequences in physics, cf. Sec. II.D.) It is another question how well it can be *measured* (cf. Sec. IV.B). Admittedly in most cases one is not able to perform sufficiently many measurements in order to determine the density matrix ρ, and thus $S(\rho)$, completely. But this problem does not concern entropy specifically, only quite generally the quantum-mechanical concepts of density matrices and wave function. However, it is true that even if one knows ρ completely, it may be extremely hard to calculate $S(\rho)$, although, of course, this can be done in principle, because one would have to diagonalize an infinite matrix in order to compute the trace of a function of it, namely, $-\rho \ln \rho$.

1. Various interpretations of the expression for the entropy

Before trying to clarify the relation between the expression $S(\rho)$ and physical reality, I want to mention a few interpretations of von Neumann's formula.

Ludwig Boltzmann's great discovery was the celebrated formula

$$S = k \ln W$$

which appeared[2] in a paper in 1877 and established the connection between the variable of state, "entropy," which had been derived from phenomenological considerations, and the "amount of chaos" (or disorder) of a system, which, more precisely, means the number of microstates which have the same prescribed macroscopic properties. (This number has been denoted as "thermodynamical probability," in German "thermodynamische Wahrscheinlichkeit"—hence the letter W.) Of course, Boltzmann's treatment was a purely classical one. Since the "number of microstates" does not literally make sense in classical mechanics he took it as the available volume in phase space divided by the volume of an (at first arbitrarily chosen) "unit cell."

In quantum mechanics, however, there is no ambiguity at all; the "number of microstates" may be interpreted as the number of pure states with some prescribed expectation values. Let us assume that in a certain system there are W different pure states, each of them occurring with the same probability. Then the entropy is $S = \ln W$ (remember that we have put $k_B = 1$). The density matrix of this system is $\rho = (1/W)P$, P being a W-dimensional projection. One easily can see that $\ln W = -\mathrm{Tr}\,\rho \ln \rho$.

If ρ is of a more general type, then one has to look for an expression that interpolates between density matrices of the form $1/W$ times a W-dimensional projection. Of course, this is done by $S(\rho) = -\mathrm{Tr}\,\rho \ln \rho$, but there are many more expressions which do the same (for instance, $-\ln \mathrm{Tr}\,\rho^2$; cf. Sec. IV.B). However, $S(\rho) = -\mathrm{Tr}\,\rho \ln \rho$ is the only possibility with *reasonable* properties (such as additivity and subadditivity, cf. Sec. II.E and F. Furthermore, the latter expression enjoys nice "mixing properties" that are very desirable from the point of view of physics; cf. Sec. II.B).

It is rather instructive to pay attention to the combinatorial aspects of von Neumann's formula. Each density matrix can be diagonalized: $\rho = \sum p_k |k\rangle\langle k|$ [where $|k\rangle$ = normed eigenvector corresponding to the eigenvalue p_k, $|k\rangle\langle k|$ = projection onto $|k\rangle$, $p_k \geqslant 0$, $\sum p_k = 1$]. $S(\rho) = -\sum p_k \ln p_k$ (we understand that $0 \ln 0 = 0$). p_k is the probability of finding the system in a pure state $|k\rangle$. If one performs N measurements, one will obtain as a result that (at least for large N) the system is found $p_1 \cdot N$ times in the state $|1\rangle$, $p_2 \cdot N$ times in the state $|2\rangle$, etc. (Of course, these quantities need to be integers, but this is only a minor point which easily can be corrected.) Now the density matrix does not contain any information about the order in which one will find the states $|1\rangle, |2\rangle, \ldots$, etc. There are $N!/(p_1N)!(p_2N)!\ldots$ possibilities for this; and for $N \to \infty$ we find (by virtue of Stirling's formula) that $1/N$ times the logarithm of this number of possibilities converges to S.

One may likewise interpret this fact in the following manner: consider N copies of the same system (represented by the Hilbert space $\mathsf{H} \otimes \mathsf{H} \otimes \cdots \otimes \mathsf{H}$, H = Hilbert space of the original system). In this new system there are microstates of the form $|1\rangle \otimes |2\rangle \ldots$, etc., where $|1\rangle$ occurs p_1N times, $|2\rangle$ p_2N times, and so on. All these microstates have the same weight. According to Boltzmann one obtains for the entropy $\ln W_N$ (with $W_N = N!/(p_1N)!(p_2N)! \cdots$). The corresponding portion for one system is $(1/N)\ln W_N$, which goes to $S(\rho)$ as $N \to \infty$.

2. Entropy and information theory

As already explained, entropy is a measure of the "amount of chaos" or of the lack of information about a system. If one has complete information, i.e., if one is concerned with a pure state, entropy = 0. Otherwise it is >0, and it is bigger the more microstates exist and the smaller their statistical weight is. [One easily checks the inequality $S(\rho) \geqslant \ln(1/p_1)$, p_1 being the biggest eigenvalue (= operator norm) of ρ.] This principle, namely, that entropy is a measure of our ignorance about a system, described by a density matrix, or, in the classical case, by a probability distribution, enables one to apply results of mathematical information theory to physics (Sec. II.G). Also the formal correspondence between the expression $-\sum p_k \ln p_k$ and Shan-

---

[1] Sometimes, mainly in the mathematical literature, one uses the letter H instead of S for entropy. It is claimed that the H should be a capital "eta"; however, this is not so certain. In any case, the letter H was introduced by Burbury in only 1890, whereas Boltzmann himself originally used "E." In physics, H is not a very good notation because of the risk of confusion with the Hamiltonian. The name "entropy" is due to Clausius (1865) and means transformation (τροπη). The prefix "en" was chosen to have a resemblance to the word "energy."

[2] Not quite in this form, which is due to Planck (1906).

non's expression for the information content of a discrete probability distribution suggests such a procedure. We will discuss it in detail in Sec. II.G.

3. The classical approximation

The "classical limit" of the expression for the entropy is obtained by the usual prescription (we first consider the case of one degree of freedom only)

density matrix → probability distribution in phase space

$$\text{trace} \to \int \frac{dp\,dq}{2\pi\hbar}\,.$$

This can be justified mathematically by means of *coherent states*.

Coherent states were introduced by Schrödinger in 1927. [A detailed treatment is presented in the book of Klauder and Sudarshan (1968).] They are functions of the form $U(p,q)|0,0\rangle \equiv |p,q\rangle, U(p,q) \equiv e^{(i/\hbar)(pQ-qP)}$; thereby p,q = numbers, P, Q = momentum or position operator, respectively; $|0,0\rangle$ = the wave function in configuration space $\pi^{-1/4}\hbar^{-1/4}e^{-x^2/2\hbar}$. We have $\langle p,q|P(Q)|p,q\rangle = p(q)$. The $|p,q\rangle$ are Gaussian wave packets with minimal uncertainty.

One can prove the following important relation:

$$\mathrm{Tr}A = \int \frac{dp\,dq}{2\pi\hbar}\,\langle p,q|A|p,q\rangle\,.$$

One should bear in mind that the $|p,q\rangle$ are normed but not pairwise orthogonal, in fact, using the abbreviation $z = (q+ip)/(2\hbar)^{1/2}, |z\rangle \equiv |p,q\rangle$ one finds

$$\langle z|z'\rangle = \exp[-\tfrac{1}{2}|z|^2 + z^*z' - \tfrac{1}{2}|z'|^2]\,.$$

For $\hbar \to 0$ the wave packets $|p,q\rangle$ become more and more concentrated around (p,q), in the sense that $\langle p,q|(P-p)^2|p,q\rangle \to 0$ and $\langle p,q|(Q-q)^2|p,q\rangle \to 0$. It is possible to incorporate a factor ω in the definition of coherent states: let

$$|0,0\rangle = (\omega/\pi\hbar)^{1/4}e^{-\omega x^2/2\hbar}$$

$$|p,q\rangle = \exp(i/\hbar)(\omega^{-1/2}pQ - \omega^{1/2}qP)|0,0\rangle\,.$$

Our following considerations are equally valid for these kinds of coherent states.

If one defines

$$\rho^{\mathrm{cl}}(p,q) \equiv \langle p,q|\rho|p,q\rangle \tag{1.2}$$

as the classical probability distribution in phase space, then

$$\int \frac{dp\,dq}{h}\,\rho^{\mathrm{cl}}(p,q) = \mathrm{Tr}\rho = 1$$

and

$$S = -\int \frac{dp\,dq}{h}\,\langle p,q|\rho\ln\rho|p,q\rangle\,. \tag{1.3}$$

The classical approximation consists in replacing $\langle p,q|\rho\ln\rho|p,q\rangle$ by $\rho^{\mathrm{cl}}\ln\rho^{\mathrm{cl}}$:

$$S^{\mathrm{cl}} \equiv -\int \frac{dp\,dq}{h}\,\rho^{\mathrm{cl}}\ln\rho^{\mathrm{cl}}.. \tag{1.4}$$

Since $-x\ln x$ is a concave function, $-\langle p,q|\rho\ln\rho|p,q\rangle \leq -\rho^{\mathrm{cl}}(p,q)\ln\rho^{\mathrm{cl}}(p,q)$ and

$$S \leq S^{\mathrm{cl}}\,. \tag{1.5}$$

(In the rest of this article, we simply will write S, ρ instead of $S^{\mathrm{cl}}, \rho^{\mathrm{cl}}$, if there is no risk of confusion.)

The above inequality is a consequence of the following inequality for matrix elements: let f be a convex (concave) function, A be a self-adjoint operator and φ be a normed vector. Then $\langle\phi|f(A)|\phi\rangle \geq (\leq) f(\langle\phi|A|\phi\rangle)$. For the proof let us (for the sake of simplicity only) assume that A has a pure point spectrum: $A = \sum \alpha_k|k\rangle\langle k|$. Then $\phi = \sum c_k|k\rangle, \sum|c_k|^2 = 1$. $\langle\phi|f(A)|\phi\rangle = \sum|c_k|^2 f(\alpha_k) \geq (\leq) f(\sum|c_k|^2\alpha_k) = f(\langle\phi|A|\phi\rangle)$.

For many density matrices, the error due to the replacement of $\langle p,q|\rho\ln\rho|p,q\rangle$ by $\rho^{\mathrm{cl}}\ln\rho^{\mathrm{cl}}$ will be negligibly small. It turns out that the classical approximation is good as long as $\rho^{\mathrm{cl}}(p,q)$ is a smooth function spread over a volume in phase space that is $\gg h$ (Wehrl, 1977). If there are small distance fluctuations or if ρ^{cl} is concentrated on small regions of phase space, then the classical approximation can be very bad. (For an estimate of this error in typical situations, cf. Sec. C.)

There is a striking paradox since quantum-mechanically one always has $S(\rho) \geq 0$, because $S(\rho) = -\sum p_k \ln p_k$, and, since, $p_k \geq 0, \sum p_k = 1, p_k \leq 1$, one has $-p_k\ln p_k \geq 0$ (=0 if, and only if, $p_k = 0$ or 1; hence if one $p_k = 1$, all the others must be =0: therefore ρ is a one-dimensional projection, i.e., a pure state). Thus $S(\rho) \geq 0$ (cf., also Sec. II.A). The *conventional* classical entropy, however, may very well be <0, even $-\infty$, in spite of the inequality $S \leq S^{\mathrm{cl}}$. How can this happen? The reason is that usually the classical entropy is introduced in a less critical manner. Namely, it is defined by *every* probability distribution $f(p,q)$ (i.e., every function with $f \geq 0, \int f = 1$), no matter whether there is a density matrix ρ such that $f(p,q) = \rho^{\mathrm{cl}}(p,q)$ or not. Thus one does not suppose that always $f \leq 1$ and, consequently, the "classical entropy" of f,

$$S(f) \equiv -\int \frac{dp\,dq}{h}\,f\ln f\,,$$

can become negative (see Fig. 1).

Suppose that $S(f) < 0$. Because $\int(dp\,dq/h)f = 1$, the extent of the region, where $f > 1$, must be $<h$. Hence a negative classical entropy arises if one tries to localize a particle in phase space in a region $<h$, i.e., if the uncertainty relation is violated. Therefore in applying the conventional classical expression one has to keep in

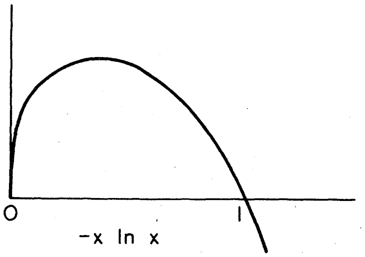

FIG. 1 Graph of $f(x) = -x\ln x$.

mind that not every classical probability distribution can be observed in nature.

Although the conventional classical entropy has some other inconvenient features (for instance, it is not monotonic as the "true" classical entropy is; cf. Sec. II.F), we will nevertheless, for the rest of this paper, always understand by "classical entropy" the conventional one, if not otherwise stated, in order to avoid any confusion.

At this place it should also be remarked—which, of course, is well known to everybody—that in purely classical reasoning the expression for the entropy can only be derived up to an additive constant. For $dp\,dq$ has the dimension of an action, hence in order to obtain something dimensionless in the normalization condition "$\int \rho^{\text{cl}} = 1$" one has to divide $dp\,dq$ by some quantity of the dimension of an action (=volume of a "unit cell"). The right quantity, as we have just seen, is Planck's constant h (not $\hbar$!). If one takes some other quantity, say h' (in a classical theory h cannot occur), we obtain the normalization condition $\int (dp\,dq/h')\rho^{\text{cl}} = 1$ and for the entropy

$$\begin{aligned} S^{\text{cl}} &= -\int \frac{dp\,dq}{h'} \rho^{\text{cl}} \ln \rho^{\text{cl}} \\ &= -\int \frac{dp\,dq}{h} \rho' \ln \rho' - \ln \frac{h'}{h} \\ &= \text{correct classical entropy} - \ln \frac{h'}{h} \left(\rho' \equiv \rho \frac{h}{h'} \right). \end{aligned}$$

If, in particular, $\rho = \text{const}$ in a certain region ("phase volume") of the (p, q) plane, otherwise $= 0$, then S^{cl} $=$ logarithm of the phase volume measured in units of h = logarithm of the number of "cells." If the size of the cells is changed, then of course the expression for the entropy changes too. (In classical statistical mechanics this problem is partly overcome by the *ad hoc* postulate of the third law of thermodynamics.)

4. The classical discrete approximation

In approximating the expression for the entropy one can go one step further and discretize the classical probability distribution ρ^{cl}. That means that one partitions the phase space into cells of size h (enumerated by some index i) and replaces ρ^{cl} in each cell by its average, which we will denote by $\bar{\rho}_i$, i.e.,

$$\bar{\rho}_i = \int_{\text{cell}\,``i"} \rho^{\text{cl}} \frac{dp\,dq}{h} . \tag{1.6}$$

Then $\sum \bar{\rho}_i = 1$. The classical discrete entropy is defined by

$$S^{\text{cl},d} = -\sum \bar{\rho}_i \ln \bar{\rho}_i . \tag{1.7}$$

Because of the inequality

$$x(\ln x - \ln y) \geq x - y , \tag{1.8}$$

one obtains

$$S^{\text{cl},d} \geq S^{\text{cl}} .$$

Like the classical (continuous) approximation, the classical discrete approximation may be sufficiently good for many purposes.

It should be noted that the same formal structure arises if all density matrices under consideration in a certain problem commute. In this case, there exists a common set of eigenvectors $|i\rangle$ such that $\rho^{(\alpha)}|i\rangle = p_i^{(\alpha)}|i\rangle$ (α labels the density matrices) and $S(\rho^{\alpha}) = -\sum p_i^{(\alpha)} \ln p_i^{(\alpha)}$. This shows that every general theorem that is true in quantum mechanics also must be true in the classical discrete case, or, vice versa, if a theorem is not true in the classical discrete case, it also cannot be true in the general quantum-mechanical case.

5. Hilbert spaces for statistical mechanics

Although the expression for the entropy does not refer to any special structure of a system, there are some particular features of many-body systems. Let me begin with a short review of the Hilbert spaces of those systems that are of primary interest in statistical mechanics. For a careful presentation, see Ruelle, 1969.

One-particle systems

The Hilbert space of a particle moving in a subvolume V of $\mathbb{R}^d$ (d = space dimension) is $L^2(V)$ = space of square-integrable functions $\psi(x)$ $(x \in V)$. Here and throughout the rest of this paper we neglect spin since our treatment will be a nonrelativistic one only.

Many-particle systems (Maxwell-Boltzmann statistics)

Here the Hilbert space is the tensor product of N copies of $L^2(\mathbb{R}^d)$, thus the particles are supposed to be distinguishable. Since in nature there are only very few distinguishable particles, Maxwell-Boltzmann statistics is not very well suited for purposes of statistical mechanics.

Bose-Einstein statistics

The Hilbert space of N identical particles obeying B-E statistics is the symmetric tensor product of N copies of $L^2(V)$:

$$\mathrm{H}_N^+(V) \equiv L^2(V) \underset{s}{\otimes} \cdots \underset{s}{\otimes} L^2(V)$$

which equals space of square-integrable functions $\psi(x_1, \ldots, x_N)$ $(x_i \in V)$, that are symmetric in $x_1, \ldots, x_N$.

Fermi-Dirac statistics

Like the B-E case, but "symmetric" being replaced by "antisymmetric";

$$\mathrm{H}_N^-(V) \equiv L^2(V) \underset{a}{\otimes} \cdots \underset{a}{\otimes} L^2(V) .$$

Fock space

If the number of particles is not kept fixed but if one rather wants to take into account the possibility of a variable number of particles, one considers Fock space

$$\mathrm{H}^{\pm}(V) = \bigoplus_{N=0}^{\infty} \mathrm{H}_N^{\pm}(V) .$$

[$\mathrm{H}_0^{\pm} \equiv C$ (one-dimensional space = vacuum).] If the measure of the intersection of two volumina V_1 and V_2 is zero (by abuse of language we will always write $V_1 \cap V_2$

$= \phi$ in that case), then

$$\mathrm{H}^{\pm}(V_1 \cup V_2) = \mathrm{H}^{\pm}(V_1) \otimes (V_2) . \qquad (1.9)$$

Lattice systems

Mainly for the study of models (such as lattice gases, ferromagnets, binary alloys, etc.) one is interested in lattice systems. There, to each point x of a lattice Z^d of dimension d, one assigns a Hilbert space H_x of fixed finite dimension. Let $V \subset Z^d$ be a subset. Then

$$\mathrm{H}(V) = \bigotimes_{x \in V} \mathrm{H}_x .$$

Again, $\mathrm{H}(V_1 \cup V_2) = \mathrm{H}(V_1) \otimes (V_2)$ if $V_1 \cap V_2 = \phi$.

Several species of particles

We will not consider this case since it can be treated in an obvious manner once the results for identical particles are established.

After having sketched the various types of Hilbert spaces that we are interested in let us discuss some aspects of Eq. (1.1) for the entropy.

Independent particles

Consider N identical particles (fermions or bosons). Let ρ be the one-particle density matrix. The N-particle density matrix might be expected to be $\rho \otimes \rho \otimes \cdots \otimes \rho$. However, the trace of this operator, restricted to the Hilbert space H_N^+ or H_N^-, is not $=1$. Take, for instance, $N=2$. Then

$$\mathrm{Tr}\rho \otimes \rho = \begin{cases} \frac{1}{2}(1 - \mathrm{Tr}\rho^2) \text{ for fermions} . \\ \frac{1}{2}(1 + \mathrm{Tr}\rho^2) \text{ for bosons} . \end{cases} \qquad (1.10)$$

If $\mathrm{Tr}\rho^2 \ll 1$ [and, consequently, $S(\rho) \gg 0$], $\mathrm{Tr}\rho \otimes \rho \simeq \frac{1}{2}$, and, similarly, for N particles, $\mathrm{Tr}\rho \otimes \ldots \otimes \rho \simeq 1/N!$. The entropy is then $\simeq -\mathrm{Tr}N!(\rho \otimes \cdots \otimes \rho)\ln(N!\rho \otimes \cdots \otimes \rho) \simeq -\ln N! + NS(\rho)$. (The necessity of subtracting a term $-\ln N!$ was first demonstrated by Gibbs' Paradox.)

Let us illustrate this by the simple example of a "microcanonical" density matrix $\rho = (1/W)P$ (P = projection of dimension W and $W \gg N$). Here S = logarithm of the number of microstates $= \ln\binom{W}{N}$ for fermions, $= \ln\binom{W+N-1}{N}$ for bosons. In either case $S \simeq -\ln N! + N \ln W$.

The term $-\ln N!$ that appears in these calculations can be derived from a rule known as correct Boltzmann counting: microstates of the type, say, $|1\rangle \otimes |2\rangle \otimes \cdots$ and $|2\rangle \otimes |1\rangle \otimes \cdots$ are to be identified (which clearly is a consequence of the identity of the particles), whereas the contribution of the states of the form $|1\rangle \otimes |1\rangle \otimes \cdots$ can be neglected.

Similar statements, of course, hold for the classical approximation. Up to now we have considered one particle and one degree of freedom only. If there are d degrees of freedom, the classical probability distribution $\rho^{\mathrm{cl}}(\mathbf{p}, \mathbf{q})$ is obtained in a straightforward manner. However, for N identical particles one must not—either in the normalization condition or in the expression for the classical entropy—take the integral (we have put $d=3$)

$$\int \frac{d^3p_1 \ldots d^3q_N}{h^{3N}} \cdots$$

but rather

$$\int \frac{d^3p_1 \ldots d^3q_N}{h^{3N}N!} \cdots$$

in order to correct for the fact that, for instance, the points in phase space $(\mathbf{p}_1, \mathbf{q}_1, \mathbf{p}_2, \mathbf{q}_2, \ldots)$ and $(\mathbf{p}_2, \mathbf{q}_2, \mathbf{p}_1, \mathbf{q}_1, \ldots)$ cannot be distinguished.

6. The generalized Boltzmann-Gibbs-Shannon entropy

Let me conclude this section with a remark on a purely formal level. From the mathematical point of view the expression for the classical entropy can be considered as a special case of the so-called "generalized Boltzmann–Gibbs–Shannon (BGS)" entropy (cf., Ochs, 1976). Its definition is: Let $(\Omega, \mathrm{A}, \mu)$ be a σ-finite measure space, ν be a probability measure that is absolutely continuous with respect to μ (hence its Radon–Nikodym derivative $d\nu/d\mu$ exists). Then the generalized BGS entropy is

$$S \equiv -\int \frac{d\nu}{d\mu} \ln \frac{d\nu}{d\mu}\, d\mu \ \left(\text{if } \frac{d\nu}{d\mu} \ln \frac{d\nu}{d\mu} \text{ is integrable}\right). \qquad (1.11)$$

Important examples are the following ones:

1. *Boltzmann–Gibbs (Classical) Entropy:*

$d\mu = d^{3N}p\, d^{3N}q/h^{3N}$ (or $\cdots/h^{3N}N!$, respectively),

$d\nu = \rho^{\mathrm{cl}}\, d\mu$.

2. *Shannon Entropy of Information Theory:*

$\Omega = \{1, 2, \ldots\}$, $\mu(\{1\}) = \mu(\{2\}) = \cdots = 1$, $\nu(\{i\}) = p_i$.

Then

$$S = -\sum p_i \ln p_i .$$

3. *Relative Entropy.* $d\mu = \sigma d^{3N}p\, d^{3N}q/h^{3N}(N!, d\nu = \rho d^3p\, d^3q/h^3(N!), \sigma, \rho$ being probability distributions. In this case the generalized BGS entropy is $-\int \rho (\ln\rho - \ln\sigma) d^3p\, d^3q/h^{3N}(N!)$. With the + sign in front, this quantity is called "relative entropy" and plays an important role as we will see very soon. [In quantum mechanics, one defines the relative entropy between two density matrices σ, ρ as $S(\sigma|\rho) = \mathrm{Tr}\rho(\ln\rho - \ln\sigma)$. We will study this concept in detail in Sec. III.B.]

4. *Renyi's Information Gain.* This is a discrete version of relative entropy. $\Omega = \{1, 2, \ldots\}$ as in example 2, but $\mu(\{i\}) = q_i$ instead of $1, \sum q_i = 1$. Then $S = -\sum p_i (\ln p_i - \ln q_i)$. (See Renyi, 1966.)

The most general concept in this direction is the Segal entropy (Segal, 1960). It covers both the classical generalized BGS entropy and quantum-mechanical entropy. (cf. Sec. IV.C.)

B. Entropy and physics

The relation between entropy and physics is established by an empirical principle, namely, the second law of thermodynamics. There are several formulations of this law, of varying degrees of validity, which we will now briefly discuss. However, as mentioned in the introduction, the problem of the second law of thermodynamics does not appear to be fully understood yet.

1. A paradox

A very common formulation of the second law of thermodynamics reads as follows: the entropy of a closed system never decreases; it can only remain constant or increase. A less sharp formulation is the following (maximum entropy principle): the entropy of a closed system in equilibrium always takes the maximal possible value. (Of course, both formulations are a little bit vague and have to be specified in concrete instances.)

These statements are, however, in striking contradiction to the fact that the entropy of a system obeying the Schrödinger equation (with a time-independent Hamiltonian) *always remains constant*. For the density matrix at time t—let us denote it by $\rho(t)$—is obtained from the density matrix at time 0—ρ—by the formula

$$\rho(t)=e^{-iHt}\rho\, e^{iHt}\,. \tag{1.12}$$

Since e^{iHt} is a unitary operator, the eigenvalues of $\rho(t)$ are the same as the eigenvalues of ρ. But the expression for the entropy only involves the eigenvalues of the density matrix, hence $S(\rho(t))=S(\rho)$. (In the classical case, the analogous statement is a consequence of Liouville's theorem.)

This result seems to be absurd since one knows by experimental experience that the second law is something very sensible and very useful. There is one way out of this dilemma; that is, that the time evolution of a system is not described by the Schrödinger equation but by some other equation. In fact, in statistical mechanics one uses, with great success, equations like the Boltzmann equation, the master equation, and other equations.

2. The Boltzmann equation

To begin with, let us look at the classical Boltzmann equation (Boltzmann, 1872). In historical development, this equation was the first one to describe an irreversible behavior of a system in a rigorous way. Yet this equation is still the best known to most physicists. Many of its features are characteristic of all equations that aim at overcoming the difficulty that microscopic description and irreversibility do not fit together. (See the article by Grad, 1958, or Cohen and Thirring, Eds., *The Boltzmann Equation*, 1972. Of course, the Boltzmann equation is also discussed in all textbooks on statistical mechanics.) Perhaps the reader should be warned that although usual macroscopic equations, such as the Navier–Stokes equations, can be derived from the Boltzmann equation by means of further approximations, the Boltzmann equation also has to be considered as a *macroscopic* equation because it provides a description of the system in which the original number of degrees of freedom, 10^{23}, is reduced to 6.

The Boltzmann equation is by no means an immediate consequence of the laws of classical mechanics, i.e., the Hamiltonian equations. Rather it is based on several assumptions, such as, for instance, the molecular chaos, or the "Stosszahlansatz," and upon the fact that one considers the one-particle correlation function only, instead of taking the whole probability distribution in phase space. It turns out that, although the time evolution of the total system is given by the Hamiltonian dynamics, under certain conditions the time evolution of the first correlation function can be described, in fairly good approximation, by an irreversible equation.

The correlation function of a system of N identical particles (with mass = 1) is defined by

$$F(p,q,t)d^3p\ \ d^3q=\text{number of particles in the volume } d^3p\ d^3q \text{ at time } t \tag{1.13}$$

[hence $(1/N)F$ = probability of finding one particle in d^3p d^3q, irrespective of where the others are]. From the definition,

$$\int F d^3p\ d^3q=N\,. \tag{1.14}$$

F is obtained from $\rho^{\mathrm{cl}}(p_1,q_1,\ldots)$ by the formula

$$F(p,q,t)=\int\frac{d^3p_2\,d^3q_2\cdots d^3p_N\,d^3q_N}{h^{3N}(N-1)!}\times\rho^{\mathrm{cl}}(p,q,p_2,q_2,\ldots,t)\,. \tag{1.15}$$

(Because of the symmetry of ρ the exceptional position of the first particle is only fictitious.)

The assumption of molecular chaos states that the number of pairs of particles in the element d^3q in configuration space, with momenta in d^3p_1, or d^3p_2, respectively, equals $[F(p_1,q,t)d^3p_1\ d^3q][F(p_2,q,t)d^3p_2\ d^3q]$.

From it one derives (we consider the simplest case: no external forces, no internal degrees of freedom, etc.) the Boltzmann equation

$$\left(\frac{\partial}{\partial t}+\mathrm{p}_1\nabla_q\right)F_1=\int d\Omega\ d^3p_2\ \sigma(\Omega)\left|p_1-p_2\right|(F_2'F_1'-F_2F_1)\,, \tag{1.16}$$

where $\sigma(\Omega)$ is the differential cross section for a collision $(p_1,p_2)\to(p_1',p_2')$ (Ω = solid angle), $F_i\equiv F(p_i,q,t)$, $F_i'\equiv F(p_i',q,t)(i=1,2)$.

The Boltzmann equation implies the H Theorem the function

$$H(t)=-\int d^3p\ d^3q\ F\ln F \tag{1.17}$$

is nondecreasing in time. The following remarks apply:

(1) H, as defined by Eq. (1.17), does not coincide with the classical entropy in general. This is the case only if ρ^{cl} factorizes: $\rho^{\mathrm{cl}}(w_1,w_2,\ldots,w_N)=N!\,\rho_1^{\mathrm{cl}}(w_1)\cdots\rho_1^{\mathrm{cl}}(w_N)$ $(w_i\equiv(p_i,q_i))$, we then have

$$S^{\mathrm{cl}}=NS(\rho_1^{\mathrm{cl}})-\ln N! \tag{1.18}$$

(cf. Sec. A). Otherwise, $H\gtrsim S^{\mathrm{cl}}$ (cf. Sec. II.F).

(2) The correlation function F is obtained from the

"true" distribution in phase space by some sort of averaging.

(3) The assumption of molecular chaos cannot be justified from first principles. It may be probable to a more or less high extent, but it certainly is neither necessary nor true for all time.

From our discussion up to now we have learned that the mechanism of nondecrease of entropy is based upon averaging and probability assumptions. (We will recognize this in a somewhat clearer fashion in the example of the "master equation.") However, it should be mentioned that there is a *rigorous* derivation of the Boltzmann equation (but only for small times) for a gas of hard spheres of diameter d in the limit $d \to 0$, $n^2 d$ kept fixed, where n = number of particles/cm^3, by Lanford (1975). In this limit, the system consists of infinitely many particles. This is one of the hints that rigorous versions of irreversibility, and quite generally thermodynamical behavior, are to be expected for *infinite* systems (and possibly for a restricted class of initial states) only [cf. Sec. IV.C).

Let us return to finite systems and proceed by discussing ergodicity and mixing properties of classical systems.

3. Ergodicity and mixing

I want to start with the concept of energy shell. Let $\rho(t)$ be the time evolution of a density matrix ρ, i.e., $\rho(t) = e^{-iHt}\rho\, e^{iHt}$, and let $|n\rangle$ be the eigenvectors of the Hamiltonian, thus $H|n\rangle = E_n|n\rangle$. Then the matrix elements of $\rho(t)$ are

$$(\rho(t))_{nm} \equiv \langle n|\rho(t)|m\rangle = e^{(i/\hbar)(E_m - E_n t)}\langle n|\rho|m\rangle\,. \tag{1.19}$$

We may classify them as follows:

(a) Matrix elements that change in a significant way only during *macroscopic* time intervals (say, 10^{-9} sec). They are connected with extremely small (unmeasurable) energy differences $E_m - E_n$ ($<10^{-18}$ erg).

(b) If the difference $E_m - E_n$ is bigger, then $(\rho(t))_{nm}$ is a very rapidly oscillating function of t. Since macroscopic measurements last rather long compared with the frequency of these oscillations, they in fact will average over $(\rho(t))_{nm}$. The mean value of these matrix elements being of order of magnitude $\hbar/(E_m - E_n)\Delta t$ (Δt = period of the measurement), one can neglect them, or, expressed in other terms, those fluctuations are too rapid to be observed.

Now let $E = \mathrm{Tr}\rho(t)H$ be the expectation value of the energy. Of course, this is a constant of motion since $\mathrm{Tr}\rho(t)H = \mathrm{Tr}e^{-iHt}\rho\, e^{iHt}H = \mathrm{Tr}e^{-iHt}\rho H\, e^{iHt} = \mathrm{Tr}\rho H$. On the other hand, $E = \sum \rho(t)_{nn}E_n$. Due to our foregoing considerations for a description of macroscopic changes of the system, one only has to take into account the matrix elements of class (a). Hence we can restrict ourselves to that subspace of the Hilbert space that is spanned by those energy eigenvectors $|n\rangle$ for which $|E - E_n| < \epsilon$ with ϵ being sufficiently small. We will call this Hilbert space the *energy shell*. (In our considerations we always have assumed that the Hamiltonian has a pure point spectrum. However, it is a simple matter to generalize our arguments to Hamiltonians with a continuous spectrum.)

In the classical case (cf. Sec. A) we formally can put $\hbar = 0$ and therefore we are allowed to choose the energy shell infinitesimally thin. Of course, here "energy shell" no longer means a subspace of Hilbert space $L^2(\mathrm{R}^{3N})$, but rather a subset of R^{6N}. We will denote the classical energy shell by Ω_E (or simply Ω):

$$\Omega_E = \{(p_1, \ldots, q_N) : H(p_1, \ldots, q_N) = E\}\,, \tag{1.20}$$

[$H(\ldots)$ = classical Hamiltonian]. In the following we want to use the abbreviation $w \equiv (p_1, \ldots, q_N)$.

The restriction of the measure $d^3p_1 \cdots d^3q_N/N!$ (as always we assume the particles to be identical) to the energy shell Ω_E, formally given by

$$\delta(E - H(p_1, \ldots, q_N))d^3p_1 \cdots d^3q_N/N!\,, \tag{1.21}$$

defines a measure dw. (For a more precise definition see, for example, Reed and Simon (1972) or Arnold and Avez (1969), and other textbooks on ergodic theory.) By virtue of Liouville's theorem this measure is time invariant, i.e., $dw = dw(t)$. Let me denote by $W(\Omega)$ (or simply W) the measure of all of Ω, by $W(A)$ the measure of a subset $A \subset \Omega$.

In classical statistical mechanics the concept of ergodicity has been introduced by Boltzmann in order to justify the microcanonical ensemble. A "microcanonical ensemble" means a uniform probability distribution over the energy shell, i.e.,

$$\rho_{\mathrm{mc}}(w) = \frac{1}{W(\Omega)}\,. \tag{1.22}$$

[We write ρ_{mc} instead of the more precise notation $\rho_{\mathrm{mc}}^{\mathrm{cl}}$, recalling our remark following Eq. (1.5)]. However, ergodicity certainly is too weak a property to establish that every probability distribution tends (at least in a certain sense) to the microcanonical one. Therefore one has to introduce a stronger notion: mixing. (This concept is due to Hopf, 1932.)

A system is called "mixing" if the following is true: let A_t be the time evolution of a subset $A \subset \Omega$ (i.e., $A_t = \{w(t) \colon w(0) \in A\}$). Then, for any two sets $A, B \subset \Omega$, always

$$\lim_{t\to\infty} W(A_t \cap B) = W(A)\frac{W(B)}{W(\Omega)}\,. \tag{1.23}$$

Ergodicity only would state that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t W(A_{t'} \cap B)dt' = W(A)\frac{W(B)}{W(\Omega)}\,. \tag{1.24}$$

There is no direction of time favored in this definition, which perhaps is not easy to recognize at first, because

$$\lim_{t\to-\infty} W(A_t \cap B) = \lim_{t\to-\infty} W(A \cap B_{-t}) = \lim_{-t\to\infty} W(A \cap B_{-t}) = \frac{W(A)W(B)}{W(\Omega)}\,.$$

One can think of such a system as a flow with strongly turbulent aspects. After sufficiently large times every set A is so ragged that its relative portion in every fixed part B of Ω is just $W(A)/W(\Omega)$ (see Fig. 2). It should be kept in mind, however, that such a behavior has *nothing* to do with irreversibility. Given A_t, for $t > 0$, one can reconstruct A for $t = 0$. In fact, this sometimes even can be done

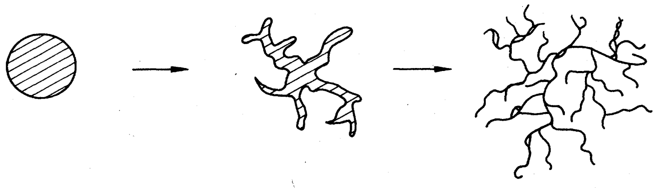

FIG. 2 Time evolution of the set A.

experimentally, for instance, in the spin echo experiment (cf. Blau, 1959; Mayer, 1961).

Of course, for a mixing system the entropy also remains constant. Nevertheless such a system gives the impression that if one does not look at it in a highly accurate manner, every set A after a certain time appears to be distributed uniformly over Ω. To make these feelings precise, let us divide Ω into "cells" (not to be confused with the cells of section A) of finite size: $\Omega=\Omega_1 \cup \Omega_2 \cup \cdots$ $(\Omega_i \cap \Omega_k=\phi$ if $i\neq k)$. The idea behind this is that macroscopic measuring apparatuses have only a restricted precision and are not able to distinguish between points inside one cell. Then also $\rho(w)$ cannot be measured by them exactly, but rather only its mean value over the cells. (Of course, there is a certain arbitrariness with this concept because there is no canonical way of defining these cells.) Let us define the coarse-grained density as follows:

$$\rho_{cg}(w)=\frac{1}{W(\Omega_i)}\int_{\Omega_i}\rho(w')\,dw' \tag{1.25}$$

if $w\in\Omega_i$. (That coarse-graining is essential in statistical mechanics was first pointed out by the Ehrenfests, 1911.) This corresponds to replacing ρ by a distribution that is uniform inside the cells. As discussed before, one cannot distinguish ρ and ρ_{cg} by macroscopic measurements. The coarse-grained entropy is

$$\begin{aligned} &S_{cg}(\rho)=S(\rho_{cg})=-\sum P_i \ln P_i/W(\Omega_i)\,,\\ &P_i\equiv\int_{\Omega_i}\rho(w)\,dw\,. \end{aligned} \tag{1.26}$$

Of course, $S_{cg}\geq S$, since we have lost information. (A proof is easily obtained by means of the inequality below.)

Now mixing implies that

$$\lim_{t\to\pm\infty}\int\rho(w(t))\,dw=\frac{W(\Omega_i)}{W(\Omega)}\,, \tag{1.27}$$

because this is true for distributions of the form $\rho(w)=\chi_A(w)/W(A)$ [$\chi_A(w)=$ characteristic function of A], hence for their convex combinations and eventually for the limits of them, i.e., all density distributions. Therefore the caorse-grained density tends towards the microcanonical one:

$$\rho_{cg}(w(t))\to\frac{1}{W(\Omega)} \tag{1.28}$$

and $S_{cg}\to\ln W(\Omega)$. Note that the convergence need not be monotonic.

Here $\ln W(\Omega)$ is the maximal possible value of the entropy:

$$-\int\rho\ln\rho\,dw\leq\ln W(\Omega)\,. \tag{1.29}$$

For the proof we utilize the concavity of $s(x)\equiv -x\ln x$. It implies that $s(y)-s(x)\leq s'(x)(y-x)$, hence $y(\ln y-\ln x)\geq y-x$. Putting $y=\rho(w)$, $x=1/W(\Omega)$ one finds $-\rho(w)\ln\rho(w)\leq-\rho(w)\ln[1/W(\Omega)]+[1/W(\Omega)-\rho(w)]$, and, after integration, $S(\rho)\leq\ln W(\Omega)$. If one inserts another distribution $\sigma(w)$ instead of $1/W(\Omega)$, one arrives at the inequality for the relative entropy $\int\rho(\ln\rho-\ln\sigma)\geq0$.

4. The master equation

The numbers P_i as introduced above may be interpreted as the probability of finding the system in the cell i. They do not obey simple differential equations; in order to compute $P_i(t>0)$ it is not sufficient to know all $P_j(t=0)$ and perhaps their derivatives of low order.

Under some simplifying assumptions, however, it is possible to derive a simple differential equation (Pauli, 1928), which, of course, is of restricted validity but may be suitable for practical purposes.

Take a distribution that is constant in cell i and $=0$ in all other cells. By Hamilton's equations one obtains from it the density distribution at time t: $\rho(t)$, and the probabilities $P_j(t)$. If $\rho(t=0)$ were concentrated in the cell i, but not constant, one would obtain another density distribution $\bar{\rho}(t>0)$ and other probabilities $\bar{P}_j(t)$. Now if the cells are not too small one can find arguments that in the overwhelming majority of possible cases $\bar{P}_j(t)=P_j(t)$. Starting with arbitrary distributions $\rho(t=0)$, no longer necessarily concentrated within one cell, one concludes that "almost always" $P_j(t)$ can be calculated from the $P_j(0)$:

$$P_j(t)=\sum_i T_{ji}(t)P_i(0)\,. \tag{1.30}$$

On the other hand, $\rho(t+t')$ can be calculated from $\rho(t)$, hence, similarly,

$$P_j(t+t')=\sum T_{ji}(t')P_i(t)\,.$$

For simplicity and mathematical convenience one may impose the Markov property on the T's:

$$T_{ji}(t+t')=\sum_k T_{jk}(t')T_{ki}(t) \tag{1.31}$$

(Chapman–Kolmogorov equation). The differential form of this equation is obtained by inserting $t'=dt$. Then $T_{jk}(dt)$ must be of the form

$$\delta_{jk}\Big(1-dt\sum_l W_{lk}\Big)+dtW_{jk}\,.$$

The invariance properties of the Hamiltonian equations imply that $W_{jk}W(\Omega_k)=W_{kj}W(\Omega_j)$ (microscopic reversibility, detailed balancing). Also $T_{jk}\geq0$ in order that $P_i(t)\geq0$ if $P_i(0)\geq0$, hence $W_{jk}\geq0$. (Note that the diagonal terms W_{kk} cancel in the above formula.) We thus arrive at the (classical) master equation

$$\dot{P}_j=\sum_k(W_{jk}P_k-W_{kj}P_j)\,. \tag{1.32}$$

From it one can derive various macroscopic or phenomenological equations for which I want to refer to the literature only. (For a good bibliography, see Reif, 1965.)

The considerations presented above are not intended to give any "proof" of the second law of thermodynamics. I rather wanted to draw attention to those assumptions that are necessary in order to "produce" an irreversible behavior, or to derive equations predicting approach to equilibrium. Let me single out once more the main features:

(1) Some averaging procedure is needed. Concerning the Boltzmann equation, it consisted in considering the first correlation function instead of the complete distribution in phase space. For closed systems, this leads to a nonlinear equation. For open systems in free space (which we did not discuss) this leads to a *linear* transport equation. The treatment of open systems in bounded regions (systems in a heat bath) leads to master equations of the type mentioned above or, more generally, to dynamical semigroups (Kossakowski, 1972; Gorini, Frigerio, Verri, Kossakowski, and Sudarshan, 1976; Davies, 1976).

In our last discussion, we introduced coarse-graining. This may seem to be somewhat artificial. However, after all, one can regard it in another way. Suppose we are dealing with a system consisting of two subsystems, the phase space of one of them being discrete, $\Omega_1 = \{1, 2, \dots\}$, the phase space of the second one being continuous, and the phase space of the whole system being $\Omega = \Omega_1 \times \Omega_2$. In addition, assume that the composite system is mixing. Let $\rho = \rho(i, w)(i \in \Omega_1, w \in \Omega_2)$ be a probability distribution. The corresponding density distribution of the two subsystems obviously is to be taken as

$$\begin{aligned} \rho_1(i) &= \int_{\Omega_2} dw\, \rho(i, w) \\ \rho_2(w) &= \sum_{\Omega_1} \rho(i, w) \end{aligned} \tag{1.33}$$

(cf. Sec. II.F for this concept); so that the entropy of the first subsystem is

$$-\sum \rho_1(i) \ln\rho_1(i),$$

which is just the coarse-grained entropy with respect to the partition $\{\Omega_i\}(\Omega_i = \{i\} \times \Omega_2)$, except for the term $\ln W(\Omega_2)$. Thus the entropy of the first system approaches its maximal value. We will see later on (in detail in Sec. IV.C) that exactly this mechanism is responsible for the possible "increase" of entropy in quantum systems.

(2) In order to achieve approach to equilibrium, some ergodicity properties are definitely needed, and one can expect that the better the ergodicity properties of the system are, the more an arbitrary density distribution will tend to a stable one.

(3) The derivations of the Boltzmann equation or the master equation depend on randomness assumptions that are supposed to hold at any time (molecular chaos, or replacing $\rho(t)$ by a distribution that is uniform inside the cells, respectively). These assumptions may be likely to hold, but certainly cannot be proven at all. Nevertheless one can say that these equations are the best one can expect because in a realistic situation there are always many uncontrollable perturbations that will have the effect that the "true" dynamics of the system is spoiled and that the time evolution of a point in phase space no longer obeys the Hamiltonian equations but rather behaves in a stochastic manner.

Properties of the solutions of the master equation were first discussed by von Mises (1931), Fréchet (1938), and Feller (1950), just to mention the earliest treatments. In particular the master equation implies a monotone increase (or nondecrease) of the coarse-grained entropy

$$S_{\text{cg}} = -\sum P_i \ln \frac{P_i}{W(\Omega_i)}.$$

A more general result in this direction is that for density matrices $\rho(t)$, whose time evolution is given by a dynamical semigroup, the relative entropy $S(\rho_0|\rho(t))$ decreases. [See Eq. (1.41) below. ρ_0 = the stationary state.] The entropy production

$$-\frac{d}{dt} S(\rho_0|\rho(t))$$

then is positive and $S(\rho_0|\rho(t))$ is convex (Spohn, 1977). (The latter fact relies on Lieb's theorem; cf. Sec. III.) For the proof of our first statement, let us, for simplicity, assume that all cells have the same size, $W(\Omega_i) = \omega$. Then

$$S_{\text{cg}} = \ln\omega - \sum P_i \ln P_i .$$

Now we arrange the P_i in decreasing order: $P_1 \geqslant P_2 \geqslant \cdots$. For the sum of the biggest nP_i's we find

$$\frac{d}{dt} \sum_{i=1}^{n} P_i = \sum_{i=1}^{n} \sum_{k} (W_{ik} P_k - W_{ki} P_i) \leqslant 0 .$$

Therefore, the sum of the biggest nP_i's at time t_1 is $\leqslant$ the sum of the biggest nP_i's at time $t_2 < t_1$. (Note that the *indices* of the P_i may change with time.)

Now we use the lemma: Consider two decreasing sequences of numbers $\alpha_1 \geqslant \alpha_2 \geqslant \cdots$, $\beta_1 \geqslant \beta_2 \geqslant \cdots$ such that $\sum \alpha_i = \sum \beta_i = 1$, for which the following relations hold: $\alpha_1 \leqslant \beta_1$, $\alpha_1 + \alpha_2 \leqslant \beta_1 + \beta_2, \dots, \alpha_1 + \alpha_2 + \cdots + \alpha_r \leqslant \beta_1 + \beta_2 + \cdots + \beta_r, \dots$. Then, $-\sum \alpha_i \ln\alpha_i \geqslant -\sum \beta_i \ln\beta_i$. The proof is based on a discrete version of the inequality for the relative entropy of Sec. A: $\sum \beta_i(\ln\beta_i - \ln\alpha_i) \geqslant 0$. By assumption, $\sum \alpha_i \ln\alpha_i = \alpha_1(\ln\alpha_1 - \ln\alpha_2) + (\alpha_1 + \alpha_2)(\ln\alpha_2 - \ln\alpha_3) + (\alpha_1 + \alpha_2 + \alpha_3)(\ln\alpha_3 - \ln\alpha_4) + \cdots \leqslant \beta_1(\ln\alpha_1 - \ln\alpha_2) + (\beta_1 + \beta_2)(\ln\alpha_2 - \ln\alpha_3) + \cdots$, i.e., $\sum \alpha_i \ln\alpha_i \leqslant \sum \beta_i \ln\alpha_i \leqslant \sum \beta_i \ln\beta_i$.

Thus, as far as the master equation is concerned, not only does the entropy never decrease, but also the sum of the n biggest eigenvalues never increases. This property is referred to as *mixing-enhancing* (cf. Sec. II.C. The English translation of Uhlmann's original appelation "mischungsverstärkend" is due to C. Fellbaum). Its meaning is that *à la longue* the large P's will get smaller, the small P's will get larger, until eventually all P's are equal. It turns out that mixing-enhancement not only implies that entropy does not decrease but also that, for any concave (or convex, respectively) function f, $\sum f(P_i)$ is nondecreasing (or nonincreasing, respectively). This result follows from a simple modification of our last proof. Let f be concave. Then $f(\beta_i) - f(\alpha_i) \geqslant (\beta_i - \alpha_i) f'(\beta_i)$. $\sum \beta_i f'(\beta_i) = \beta_1(f'(\beta_1) - f'(\beta_2)) + (\beta_1 + \beta_2)(f'(\beta_2) - f'(\beta_3)) + \cdots \geqslant \alpha_1(f'(\beta_1)$

$-f'(\beta_2)) + (\alpha_1+\alpha_2)(f'(\beta_2)-f'(\beta_3)) + \cdots = \sum \alpha_i f'(\beta_i)$ [remember that f' is decreasing, because f is concave, i.e., $(f'(\beta_1)-f'(\beta_2)) \leqslant 0$]. Hence $\sum [f(\beta_i)-f(\alpha_i)] \geqslant 0$ (cf. also Hardy, Littlewood, and Polya, 1934; Polya, 1950).

There is a close connection between mixing-enhancement and nondecrease of entropy. Suppose that there is a linear connection between $P_i(t)$ and $P_i(0)$ (not necessarily originating from a master equation):

$$P_i(t) = \sum M_{ik} P_i(0).$$

If, for all probability distributions, $-\sum P_i(t)\ln P_i(t) \geqslant -\sum P_i(0)\ln P_i(0)$, then the mapping $(P_i(0)) \to (P_i(t))$ is mixing-enhancing (Uhlmann, 1977). Let me sketch a simplified version of the proof, and thus suppose that the probability distribution is finite: $(P_1, \ldots, P_n)$. The matrix M_{ik} must be stochastic, i.e., $\sum_i M_{ik} = 1$ and $M_{ik} \geqslant 0$ in order to guarantee that $(P_1(t), \ldots, P_n(t))$ is a probability distribution. On the other hand, for $P_i(0) = 1/n$ (for all i), also $P_i(t) = 1/n$, because otherwise $-\sum P_i(t)\ln P_i(t)$ would be strictly $< \ln n$. From this, one finds that $\sum_k M_{ik} = 1$. Therefore M is a doubly stochastic matrix, and, by Birkhoff's theorem (see, for example, Glasman and Gubich, 1969), a convex combination of permutation matrices. This immediately implies mixing-enhancement.

If we try to adapt our previous considerations to quantum mechanics, we are immediately faced with the problem that a perfect analogy cannot exist. According to the usual "dictionary" one would expect that one had to replace

subset of phase space	by	projection
measure of a set	by	trace (= dimension of the corresponding projection)
density distribution	by	density matrix.

Thus "mixing" in quantum mechanics should mean that, for any two projections (in the finite-dimensional Hilbert space describing the "energy shell"), say P and Q,

$$\mathrm{Tr} P_t Q \to \mathrm{Tr} P \, \mathrm{Tr} Q / W \tag{1.34}$$

where $P_t \equiv e^{iHt} P \, e^{-iHt}$ (time evolution in the Heisenberg picture) and W = dimension of the Hilbert space. With the notation $\omega(P) = \mathrm{Tr}P/W$,

$$\omega(P_t Q) \to \omega(P)\omega(Q). \tag{1.35}$$

We may even take ω to be any arbitrary invariant state, i.e., $\omega(\cdot) = \mathrm{Tr}\rho\,\cdot$ in a Hilbert space of arbitrary, possibly infinite, dimension and $\omega(P_t) = \omega(P)$ for all projections. This implies that $\mathrm{Tr}\rho P_t = \mathrm{Tr}\rho\, e^{iHt} P e^{-iHt} = \mathrm{Tr}\, e^{-iHt}\rho\, e^{iHt} P = \mathrm{Tr}\rho P$, and therefore, $e^{-iHt}\rho\, e^{iHt} = \rho$. Writing ρ in the form $\rho = \sum p_k |k\rangle\langle k|$, $(\sum p_k = 1)$, it would follow that $e^{-iHt}|k\rangle\langle k| e^{iHt} = |k\rangle\langle k|$ for all $|k\rangle$, and

$$\omega(P_t Q) = \sum p_k \langle k| e^{iHt} P \, e^{-iHt} Q |k\rangle.$$

The limit as $t \to \infty$ should be equal to

$$\sum p_k \langle k|P|k\rangle \sum p_j \langle j|Q|j\rangle.$$

In particular, let $P = Q = |l\rangle\langle l|$. It follows that $p_l^2 = p_l$, i.e., $p_l = 0$ or 1. This shows that Eq. (1.35) can only hold if the Hilbert space is one-dimensional, which, of course, is of no interest at all because then a nontrivial time evolution does not exist. (A similar consideration applies to the quantum analog of ergodicity.)

Nevertheless it is not excluded that for certain projections Q (or even for certain partitions Q_i, by which is meant a family of pairwise orthogonal projections with $\sum Q_i = 1$), one always has

$$\omega(PQ_i) \to \omega(P)\omega(Q_i),$$

at least, if W is large enough, with arbitrarily good accuracy.

An early result in this direction, referring to ergodicity rather than to mixing, however, was obtained by von Neumann, 1929 ("Proof of the ergodic theorem and the H-theorem in the new mechanics") under certain assumptions on the spectrum of H (no degeneracies, no resonances), in the limit $W \to \infty$ and $\mathrm{Tr}Q_i/W$ kept fixed, for "almost all" partitions Q_i,

$$\lim \frac{1}{t}\int_0^t \omega(P_{t'} Q_i) dt' = \omega(P)\omega(Q_i).$$

It is very remarkable that von Neumann's paper appeared even *before* modern classical ergodic theory was initiated. The latter started in 1931 only with the work of Koopman, von Neumann, and Birkhoff.

If, for some partition Q_i, $\omega(P_t Q) \to \omega(P)\omega(Q)$, then, similarly to the classical case,

$$\mathrm{Tr}\rho_t Q_i \to \omega(Q_i) \tag{1.36}$$

$(\rho_t = e^{-iHt}\rho e^{iHt})$, and the (quantum-mechanical) coarse-grained entropy

$$S_{cg}(\rho) \equiv S(\rho_{cg}),$$

with

$$\rho_{cg} \equiv \sum Q_i \frac{\mathrm{Tr}\rho Q_i}{\mathrm{Tr}Q_i}, \tag{1.37}$$

tends to the maximal value $\ln W$. (That this is in fact the maximal value follows from Klein's inequality below.) Setting $\mathrm{Tr}\rho Q_i \equiv P_i \equiv W_i$

$$S_{cg}(\rho) = -\sum P_i \ln \frac{P_i}{W_i}, \tag{1.38}$$

in close analogy to Eq. (1.26). Pictorially, the coarse-grained density matrix ρ_{cg} arises from the true one by cutting out "blocks" and then replacing every block by a matrix which is a constant multiple of the unit matrix with the same trace (see Fig. 3). If we meet such a situation, then an obvious modification of our previous arguments will lead us to the quantum-mechanical master equation for the P_i's which of course looks like the

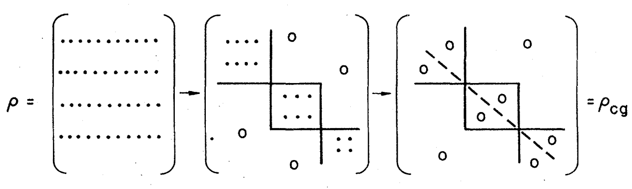

FIG. 3. Construction of the coarse-grained density matrix.

classical one. Also the same remarks apply as in the classical case concerning its validity.

In principle, however, all our previous statements about quantum-mechanical mixing are false on a formal basis because of the recurrence paradox ("Wiederkehreinwand"). In our case it states that, if the Hamiltonian has a discrete spectrum, the function $\mathrm{Tr}\rho_t Q$ is almost periodic in t. The way out is well known: the time it would take until the system gets close to the original state again is tremendously large for macroscopic systems and beyond any sensible imagination. Thus, if $t=\infty$ means something like "t = age of the universe," things are certainly okay. To correct for our above considerations in a mathematically incontestable way we have to deal with strictly infinite systems. We will do this in the last section only because of the mathematical technicalities that are involved. At this point let me just mention that ergodicity and mixing make perfect sense in the infinite case.

So far a few remarks (admittedly rather superficial) have been given on the problem of approach to equilibrium. For a more careful and detailed discussion I have to refer the reader to the literature, as announced in the introduction.

In the rest of this section I would like to comment on some properties of equilibrium states.

It is often argued on philosophical grounds that the microcanonical state is the equilibrium state (if the energy is fixed), because, after all, there is no physical principle which would distinguish between the different energy eigenstates of the energy shell and therefore any of them must occur with the same probability. However, it is not obviously certain that this application of Laplace's principle of insufficient reason to physical systems is really legitimate; one definitely has to elaborate those *physical* laws which are responsible for the validity of this principle for real matter.

Equilibrium states, and only they, also enjoy remarkable stability properties which roughly may be characterized as follows: small local perturbations of the dynamics (which, of course, never can be avoided) only lead to a local, but not to a global, change of state. In contrast, if a state is a time-invariant but not an equilibrium state, an arbitrarily small perturbation may be sufficient to produce a transition to an entirely different state. Again, these phenomena can be described rigorously in the infinite case only, so that we will come back to them in Sec. IV.D.

It is, of course, best known how the other classical ensembles are obtained from the microcanonical one, and we will consider them in somewhat more detail in the next section. What we will learn and what is of relevance for the second law of thermodynamics are the following facts:

(1) The classical ensembles obey the maximum entropy principle, i.e., the density matrix has the biggest entropy among all density matrices with the same prescribed expectation value. In our presentation this principle appears *a posteriori* only. It does not seem to be quite clear whether this principle can be justified *a priori*. Of course, there are arguments in favor of it invoking Laplace's principle or using the fact that, based on the interpretation of the formula for the entropy of the beginning of this chapter, the density matrix with maximal entropy is the most probable one.

(2) Under certain conditions to be discussed in the next section, the classical ensembles are equivalent in the sense that in the thermodynamic limit they give the same expressions for the intensive thermodynamical quantities. This shows that for large systems it is not really necessary that they be in the state with maximal possible entropy but that deviations from this state that are not too big do not change the thermodynamics. Thus "not too big" means that, for instance, a difference in entropy to the maximal value of, say, order $\sqrt{N}$ does not at all matter and can be neglected. Before further discussing the classical ensembles let me state some mathematical aspects of states with maximal entropy.

5. States with maximal entropy

We study the following problem: given $E=\mathrm{Tr}\rho H$ (H being a fixed Hamiltonian), what does the density matrix with maximal entropy look like? The answer is well known: it is

$$\begin{aligned}&\sigma_\beta \equiv e^{-\beta H}/\mathrm{Tr}e^{-\beta H} = Z^{-1}e^{-\beta H}\,,\\ &Z\equiv \mathrm{Tr}e^{-\beta H}(\text{partition function})\end{aligned}\tag{1.39}$$

(Gibbs state), where β is chosen such that $\mathrm{Tr}\sigma_\beta H=E$.

The proof is based on Klein's inequality (see, for example, Ruelle, 1969): Let f be a convex (concave) function. Then

$$\mathrm{Tr}[f(B)-f(A)] \underset{(\leqslant)}{\geqslant} \mathrm{Tr}(B-A)f'(A)\,. \tag{1.40}$$

This inequality is rather powerful and we will frequently make use of it. Let us state some important special cases.

(1) Take $f(x)=-x\ln x$. Then

$$\mathrm{Tr}A(\ln A-\ln B)\geqslant \mathrm{Tr}(A-B)\,.$$

If, in particular, A and B are density matrices ρ, σ, then we find for the relative entropy $S(\sigma|\rho)\equiv \mathrm{Tr}\rho(\ln\rho-\ln\sigma)$

$$S(\sigma|\rho)\geqslant 0\,. \tag{1.41}$$

(2) Let A be the "diagonal" of B. By this we mean: let ϕ_i be an orthonormal basis, and define A by $\langle\phi_i|A|\phi_k\rangle=\langle\phi_i|B|\phi_i\rangle\delta_{ik}$. Then, f being convex (concave),

$$\mathrm{Tr}f(A) \underset{(\geqslant)}{\leqslant} \mathrm{Tr}f(B)\,, \tag{1.42}$$

i.e.,

$$f(\langle\phi_i|B|\phi_i\rangle) \underset{(\geqslant)}{\leqslant} \mathrm{Tr}f(B)$$

(Peierls' inequality, Peierls, 1936). This also follows, of course, directly from the concavity inequality of Sec. A.

(3) Replacing B by $A+B$, A by $A+\langle B\rangle$, with $\langle B\rangle\equiv \mathrm{Tr}Be^A/\mathrm{Tr}e^A$, one obtains

$$\mathrm{Tr}e^{A+B}-\mathrm{Tr}e^{A+\langle B\rangle}\geqslant \mathrm{Tr}(B-\langle B\rangle)e^{A+\langle B\rangle}=0\,,$$

i.e.,

$$\mathrm{Tr}e^{A+B}\geqslant \mathrm{Tr}e^{A+\langle B\rangle} \tag{1.43}$$

(Peierls–Bogoliubov inequality). The analogous classical inequality

$$\int e^{f(x)+g(x)}dx \geqslant \int e^{f(x)+\bar{g}}dx \qquad (1.44)$$
$$\bar{g} \equiv \int e^{f(x)}g(x)dx \Big/ \int e^{f(x)}dx$$

is called Jensen's inequality. We now have to supply the proof of inequality (1.40). We recall the inequality

$$\langle\phi|f(B)|\phi\rangle \underset{(\leqslant)}{\geqslant} f(\langle\phi|B|\phi\rangle .$$

Hence, since convexity (concavity) implies $f(y)-f(x) \underset{(\leqslant)}{\geqslant} (y-x)f'(x)$, for the eigenvectors ϕ_i of A, belonging to the eigenvalues α_i ,

$$\mathrm{Tr}[f(B)-f(A)] \underset{(\leqslant)}{\geqslant} \sum (\langle\phi_i|B\phi_i\rangle - \alpha_i)f'(\alpha_i) ,$$

which is just $\mathrm{Tr}(B-A)f'(A)$. [Since f is convex (concave), there exist at least both one-sided derivatives f'_r, f'_l, and the inequality is true for both of them.]

Now we return to the Gibbs state. Suppose that $\mathrm{Tr}\rho H$ is $\leqslant E$, and that $\mathrm{Tr}\sigma_\beta H = E$. Then

$$\mathrm{Tr}\rho \ln\sigma_\beta = -\beta\, \mathrm{Tr}\rho H - \ln \mathrm{Tr} e^{-\beta H} ,$$

$$\mathrm{Tr}\sigma_\beta \ln\sigma_\beta = -\beta\, \mathrm{Tr}\sigma_\beta H - \ln \mathrm{Tr} e^{-\beta H} ,$$

and, by assumption,

$$-\mathrm{Tr}\sigma_\beta \ln\sigma_\beta \geqslant -\mathrm{Tr}\rho \ln\sigma_\beta ,$$

hence $S(\rho) = -\mathrm{Tr}\rho\ln\rho \leqslant -\mathrm{Tr}\rho\ln\sigma_\beta$ (inequality 1.41) $\leqslant -\mathrm{Tr}\sigma_\beta\ln\sigma_\beta = S(\sigma_\beta)$.

In the classical case the same inequality for the entropy is obtained in a similar way, or directly by using Lagrangian multipliers.

One can look at our last inequalities in another way. Define the free energy

$$F(\rho,\beta,H) \equiv \mathrm{Tr}\rho H - \frac{1}{\beta}S(\rho) = \frac{1}{\beta}S(\sigma_\beta|\rho) + F , \qquad (1.45)$$
$$F \equiv -\frac{1}{\beta}\ln \mathrm{Tr} e^{-\beta H} .$$

Then, for $\rho=\sigma_\beta$, $F(\rho,\beta,H)$ is minimal (namely,$=F$). One easily verifies the standard relation

$$E = \mathrm{Tr}\sigma_\beta H = TS + F \qquad (1.46)$$

with $T = 1/\beta$.

If H is replaced by several operators, say, $A, B, C \ldots$, then the same argument as before shows: given numbers $\langle A\rangle, \langle B\rangle, \langle C\rangle, \ldots$ then the entropy of any density matrix ρ with $\mathrm{Tr}\rho A = \langle A\rangle$. etc ..., is $\leqslant S(\sigma)$, provided that a density matrix σ of the form $e^{\alpha A+\beta B+\gamma C+\cdots}/\mathrm{Tr}e^{\alpha A+\cdots}$ exists with $\mathrm{Tr}\sigma A = \langle A\rangle$ etc. (This need not be the case if the operators $A, B, C, \ldots$ do not commute.) A well-known example in Fock space is the grand-canonical density matrix $\rho_{gc} = e^{-\beta p V}e^{\alpha N-\beta H}$, which has maximal entropy among all density matrices with given expectation value of the energy and the particle number. (Of course, here α and β do not have the same meaning as above.) (See also Bayer and Ochs, Ochs and Bayer, 1973.)

At this place, the problem of the Third law should be mentioned: Does $S\to 0$ as $\beta\to\infty$? This cannot be true generally (take, for instance, a finite-dimensional Hilbert space, and let $H=0$); on the other hand, it is to be expected for realistic systems. But there does not exist a satisfactory solution to this problem yet.

6. Some properties of Gibbs states

a. $S(\sigma_\beta)$ is decreasing in β (or increasing in the temperature)

This is a consequence of the following

Lemma 1 (Wehrl, 1974): Let f be concave (convex), with $f(0)=0$. Then

$$S\left(\frac{f(\rho)}{\mathrm{Tr}f(\rho)}\right) \underset{(\leqslant)}{\geqslant} S(\rho) .$$

This lemma is itself a consequence of the next one, as we already have seen in our discussion of properties of the solutions of the master equation.

Lemma 2: The mappings $\rho\to f(\rho)/\mathrm{Tr}f(\rho)$ (for concave f) and $f(\rho)/\mathrm{Tr}f(\rho)\to\rho$ (for convex f), all with $f(0)=0$, are mixing-enhancing in the sense that, for the eigenvalues $p_1\geqslant p_2\geqslant\cdots$ of ρ, or $p'_1\geqslant p'_2\geqslant\ldots$ of $f(\rho)/\mathrm{Tr}f(\rho)$, the following relations are true:

$$p'_1\leqslant p_1,\quad p'_1+p'_2\leqslant p_1+p_2,\ldots$$

$$p'_1+p'_2+\ldots+p'_n\leqslant p_1+p_2+\ldots+p_n ,\ldots$$

in the first case, and $\leqslant$ being replaced by $\geqslant$ in the second case.

Proof. We consider the first case only. For concave fuctions f, with $f(0)=0$, if $x\leqslant y$, then $yf(x) \geqslant xf(y)$. Therefore $p_n f(p_m)\leqslant p_m f(p_n)$ for $m\leqslant n$, and

$$f(p_1)+\ldots+f(p_m)\sum_{r=1}^{\infty}p_r \leqslant (p_1+\ldots+p_m)\sum_{r=1}^{\infty}f(p_r) ,$$

i.e.,

$$p'_1 = \frac{f(p_1)}{\sum f(p_r)}\leqslant p_1,\quad p'_1+p'_2 = \frac{f(p_1)+f(p_2)}{\sum f(p_r)}\leqslant p_1+p_2,\ \text{etc.}$$

(Remember that $\sum p_r = 1$.)

b. Kubo-Martin-Schwinger (KMS) boundary condition

Define the thermal expectation value of an observable A by

$$\langle A\rangle_\beta \equiv \mathrm{Tr}\sigma_\beta A . \qquad (1.47)$$

Consider $\langle A_t B\rangle_\beta = \mathrm{Tr}\sigma_\beta e^{iHt}Ae^{-iHt}B = Z^{-1}\mathrm{Tr}e^{(-\beta+it)H} \times Ae^{-iHt}B = Z^{-1}\mathrm{Tr}Be^{(-\beta+it)H}Ae^{-iHt} = Z^{-1}\mathrm{Tr}e^{-\beta H}Be^{i(t+i\beta)H} \times Ae^{-i(t+i\beta)} = \langle BA_{t+i\beta}\rangle_\beta$. Hence, defining $F(z) = Z^{-1}\mathrm{Tr}e^{-\beta H}Be^{izH}Ae^{-izH}$, $F(t) = \langle BA_t\rangle_\beta$, $F(t+i\beta) = \langle BA_{t+i\beta}\rangle_\beta = \langle A_t B\rangle_\beta$.

As long as we are dealing with finite-dimensional Hilbert spaces, all operations we have performed are legitimate. On infinite-dimensional Hilbert spaces one has to worry about existence and analyticity questions but eventually arrives at the following:

KMS condition: There exists a function $F(z)$ that is analytic at least in the open strip $0<\mathrm{Im}z<\beta$ and continuous in the closed strip $0 \leqslant \mathrm{Im}z \leqslant \beta$ such that, for real t, $F(t) = \langle BA_t \rangle_\beta$, $F(t + i\beta) = \langle A_t B \rangle_\beta$.

The KMS condition is of greater importance since it extends to infinite systems (where σ_β does no longer exist). It turns out that it entails far-reaching consequences for the structure of infinite systems (Kubo, 1957; Martin and Schwinger, 1959; cf. the last section. The role of the KMS-condition in infinite systems was realized by Haag, Hugenholtz, and Winnink, 1967.)

Finally it should be mentioned that there is an interesting inequality between the quantum-mechanical and the classical partition function, somewhat similar to inequality (1.5) but much more powerful. If the Hamiltonian is of the form $H = -\sum_{i=1}^{N} p_i^2/2 + V(x)$ then the quantum-mechanical partition function is

$$Z = \mathrm{Tr}\, e^{-\beta H} = e^{-\beta F},$$

whereas the classical partition function is

$$Z^{\mathrm{cl}} = \int \frac{d^{3N}p\, d^{3N}q}{(2\pi)^{3N}} \exp\left[-\beta \sum \frac{p_i^2}{2} + V(q)\right] = e^{-\beta F^{\mathrm{cl}}} \tag{1.48}$$

(we put $m = \hbar = 1$ and, for simplicity, suppose Boltzmann statistics; F^{cl} = classical free energy). Define the convolution

$$V_\omega(q) = \left(\frac{\omega}{\pi}\right)^{3N/2} \int d^{3N}q'\, V(q') \exp\left(-\omega \sum (q_i - q_i')^2\right) + \frac{3N}{2}\omega, \tag{1.49}$$

and $Z_\omega^{\mathrm{cl}}, F_\omega^{\mathrm{cl}}$ as $Z^{\mathrm{cl}}, F^{\mathrm{cl}}$ above, but with $V(q)$ being replaced by $V_\omega(q)$. Then, for all ω,

$$Z_\omega^{\mathrm{cl}} \leqslant Z^{\mathrm{cl}},$$

or

$$F^{\mathrm{cl}} \leqslant F \leqslant F_\omega^{\mathrm{cl}}. \tag{1.50}$$

The upper bound for Z relies on the Golden–Thompson inequality (Golden, 1965; Thompson, 1965): $\mathrm{Tr}\, e^{A+B} \leqslant \mathrm{Tr}\, e^A e^B$. Inserting for A = kinetic energy, $B = V(x)$, one arrives at

$$Z \leqslant \mathrm{Tr} \exp\left(-\beta \sum \frac{p_i^2}{2}\right) \exp(-\beta V).$$

To $\exp(-\beta \sum p_i^2/2)$ there corresponds an integral kernel K with

$$K(x, x) = (2\pi\beta)^{-3N/2} = \int \frac{d^{3N}p}{(2\pi)^{3N}} \exp\left(-\beta \sum \frac{p_i^2}{2}\right),$$

which immediately yields the desideratum. For the lower bound we will consider the case of one degree of freedom only. The general case is obtained in a straightforward manner. Using coherent states $|z\rangle$ (see the first section)

$$Z = \int \frac{dz}{2\pi} \langle z | e^{-\beta H} | z \rangle \geqslant \int \frac{dz}{2\pi} e^{-\beta \langle z | H | z \rangle}$$

and by an easy explicit calculation one obtains $\langle z | H | z \rangle = p^2/2 + \frac{1}{2} + V_1(q)$. That 1 can be replaced by ω follows from our remarks concerning the definition of coherent states in Sec. A. A similar inequality holds for spin systems (Lieb, 1973a).

C. The classical ensembles

The most important kinds of density matrices to be considered in statistical mechanics are the classical ensembles (see again Ruelle, 1969).

The microcanonical ensemble is the density matrix in the Hilbert space $\mathsf{H}_N^{\pm}(V)$ defined by

$$\rho_{\mathrm{mc}} = e^{-S_{\mathrm{mc}}} \chi_{[E-\epsilon, E]}(H) \tag{1.51}$$

($\chi_{[E-\epsilon, E]}$ = characteristic function of the interval $[E - \epsilon, E]$, ϵ = "thickness" of the energy shell). Here S_{mc}, the microcanonical entropy, is

$$S_{\mathrm{mc}} = \ln \mathrm{Tr}\, \chi_{[E-\epsilon, E]}(H)$$

= logarithm of the number of energy levels between $E - \epsilon$ and E.

The number ϵ is undetermined to a certain extent. In classical mechanics it can be chosen equal to zero (cf. Sec. B; it certainly is not necessary here to display the corresponding classical probability distribution $\rho_{\mathrm{mc}}^{\mathrm{cl}}$); on the contrary, in quantum mechanics it is very convenient to choose $\epsilon = E$, thus

$$\rho_{\mathrm{mc}} = e^{-S_{\mathrm{mc}}}\, \mathrm{Tr}\, \Theta(E - H). \tag{1.52}$$

(We suppose $H \geqslant 0$.) It turns out that in fact under certain conditions it does not matter how big ϵ is chosen, as we will indicate below.

The canonical ensemble in $\mathsf{H}_N^{\pm}(V)$ is given by

$$\rho_c \equiv \sigma_\beta, \tag{1.53}$$

σ_β being the Gibbs state at inverse temperature β. The entropy is

$$S_c = \beta(E - F)$$

[Eq. (1.46)].

The grand-canonical ensemble is defined in Fock space $\mathsf{H}_N(V)$ by

$$\rho_{gc} \equiv e^{-\beta p V} e^{\alpha N - \beta H} \tag{1.54}$$

($\alpha = \beta\mu$, μ = chemical potential, p = pressure.)

There are also other kinds of ensembles that are sometimes of use in physics (for instance, the "pressure ensemble": Lewis and Siegert, 1956. If there are more parameters, like electric and magnetic fields, one clearly also has to consider more complicated ensembles.) However, I do not want to go into details since these things are covered in all textbooks on statistical mechanics.

I would rather like to concentrate on only a few aspects that are of some importance for the rest of this paper.

1. The thermodynamic limit

This is the question whether, if a sequence of volumes V tends to infinity, the limits

$$\frac{1}{|V|} S_{\mathrm{mc}}, \quad \text{or} \quad \frac{1}{|V|} F,$$

or of p as defined by Eq. (3.4), exist, provided that $N/|V|, E/|V|$ (or α, β) are kept fixed. (We write $|V|$ for the measure of the region V if there is a risk of confusion.) The existence of such a limit only justifies the

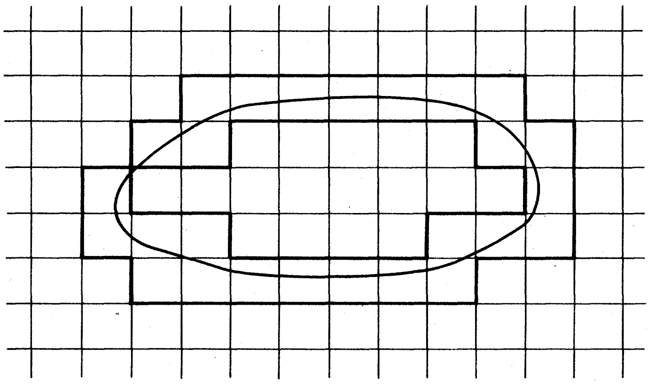

FIG. 4. Definition of n_V^- and n_V^+.

usual distinction between "extensive" quantities, like entropy, free energy, etc., and "intensive" ones like temperature, pressure, etc.

At first one has to make clear what is meant by "V tends to infinity." The least restrictive notion in this direction is "tends to infinity in the sense of van Hove" (van Hove, 1949, 1950):

Let $V(a)$, with $a \in \mathbf{R}^d$ or Z^d, be a parallelepiped $\{x \in \mathbf{R}^d\colon\ 0 \leqslant x_i \leqslant a_i\}$. Consider all translates of $V(a)$ of the form $\{x\colon\ x_i = n_i a_i + \xi_i, n_i = \text{integer}, 0 \leqslant \xi_i \leqslant a_i\}$. Let n_V^- or n_V^+, respectively, be the number of all translates of $V(a)$ that are contained in a given volume V, or have a nonempty intersection with it, respectively (see Fig. 4). A sequence of volumes is said to tend to infinity in the sense of van Hove, if, for any a,

$$n_V^- \to \infty\,, \quad n_V^+/n_V^- \to 1\,.$$

More restrictive is the tendency to infinity in the sense of Fisher. Define V_h to be the set of all points that have a distance less than h to the boundary of V. Let $D(V)$ be the diameter of V. Then $V \to \infty$ in the sense of Fisher if there is a function $\pi(\alpha)$, with $\pi(\alpha) \to 0$ for $\alpha \to 0$, such that, for sufficiently small α and all V,

$$\frac{|V_{\alpha D(\Lambda)}|}{|V|} \leqslant \pi(\alpha)$$

and, in addition, $|V| \to \infty$.

One often considers less sophisticated kinds of limits $V \to \infty$ such as sequences of parallelepipeds $V(a)$, with all $a_i \to \infty$, or sequences of balls with radii going to infinity. However, the disadvantage of these limits is that one never can be sure that the limit is shape independent and, for instance, for a sequence of cubes it could be different from that for a sequence of balls.

Now the result concerning the thermodynamic limit is that it exists for N-particle Hamiltonians of the form

$$H_N = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i<j} \Phi(x_j - x_i) \tag{1.55}$$

provided that there is a lower bound for the Hamiltonian

$$H_N \geqslant -N \cdot C\,,$$

C being a fixed constant, and if, for $|x|$ being sufficiently large,

$$\Phi(x) \leqslant A\, |x|^{-\lambda}\,, \tag{1.56}$$

$\lambda > d$. (One can also formulate a similar theorem for Hamiltonians including many-body interactions.) The proof of this theorem is quite involved and shall be omitted here. (See Ruelle, 1969. The idea of using "corridors" in order to prove the existence of the thermodynamic limit is due to Yang and Lee, 1952, and van Hove. Rigorous proofs are due to Fisher, 1964, and Griffiths, 1965.)

Unfortunately, however, these considerations do not really apply to physics, because, after all, the interaction between particles is not "tempered" in the sense of Eq. (1.56) but, with great accuracy, goes as $1/x$: there is one contribution from gravitation and an electrostatic part. Also, one usually has to consider several species of particles (electrons, nuclei, ...). If one neglects electrostatic forces (i.e., if one considers neutron stars or something similar), then a thermodynamic limit in the usual sense no longer exists. If one takes the free energy as a function of particle number, inverse temperature, and volume, the limit

$$\lim_{N \to \infty} N^{-7/3} F(N, N^{-4/3}\beta, N^{-2}V)$$

exists instead of

$$\lim_{N \to \infty} N^{-1} F(N, \beta, V)\,,$$

i.e., the scaling behavior of the thermodynamical quantities is entirely different from the usual one (Hertel and Thirring, 1971; Hertel, Narnhofer, and Thirring, 1972). This shows that it is by no means self-evident, for instance, that entropy is an extensive quantity.

If gravitation can be neglected, but there are only electric forces, then for *neutral* systems (and provided that at least one species of particles are fermions) it can be shown that a lower bound $H_N \geqslant -N \cdot C$ is true (Dyson and Lenard, 1967, 1968; Dyson, 1967. A much better method for obtaining a lower bound was presented by Lieb and Thirring, 1975) and a thermodynamic limit in the usual sense exists (Lebowitz and Lieb, 1969; Lieb and Lebowitz, 1972). This, undoubtedly, is one of the deepest results of mathematical physics: *stability of matter*. (For all that concerns stability of matter cf. the review article by Lieb, 1976.)

2. Equivalence of ensembles

The next problem is that of equivalence of ensembles: Are the thermodynamic quantities, computed with the various ensembles, asymptotically equal in the thermodynamic limit? This, of course, need not always be true (for instance, in the regions of phase transitions). On the other hand, it is "normally" to be expected. Let me illustrate this question by two examples only.

a. Equivalence of the microcanonical ensembles

Let

$$s(\eta) = \lim \frac{1}{|V|} \ln \mathrm{Tr}\,\Theta(\eta |V| - H) \tag{1.57}$$

(η = energy density; the underlying Hilbert space is that of N particles, with $N/|V|$ kept fixed). It turns out that s is a concave, nondecreasing function of η. As long as it is *strictly increasing*, all the other ensembles defined by (1.51) yield the same entropy density too (see Ruelle, 1969).

b. Comparison between microcanonical and canonical ensemble

Assuming differentiability of s, take $\beta = \partial s/\partial\eta$, and define $f = \epsilon - \beta^{-1}s$. Then the limit of the density of the free energy of the canonical ensemble exists and equals that f. (For more details, once more Ruelle's book should be consulted.)

What one can learn from all this is that for large systems, provided that the thermodynamic limit exists, the precise structure of the density matrix is not so important. To come back to our last example, the difference of the entropies in the microcanonical and the canonical ensemble is of order $\ln N$, which is big but on the other hand is negligible in the thermodynamic limit. Therefore our starting point of Sec. B, the maximum entropy principle, must be formulated as follows: given some intensive quantities (such as temperature, density, or energy density, etc.), the entropy density of the corresponding equilibrium state is maximal.

D. Historical remarks

I would like to conclude this introductory chapter by some remarks concerning history. There are many reviews, historical surveys as well as reprint volumes, containing the decisive papers in this field (for instance, by Brush, 1965, 1966; Klein, 1972; Koenig, 1959; Roller, 1950), that certainly describe the historical development of the subject much better than I could do. I only want to make a very few comments that concern "entropy" itself, without pretending that they are exhaustive or take into account all important steps in the past.

Thermodynamics in the modern sense has its origin in the work of Mayer (1842) and Joule (1845), to whom major credit is to be given in the recognition of the first law of thermodynamics (conservation of energy, which in times past was called "force"), and of Clausius (1850) and Thomson (Lord Kelvin) (1852, 1857), who, based on previous work of Carnot (1824), formulated the principle of dissipation of energy and the second law of thermodynamics. That this principle leads to the heat death was worked out by von Helmholtz (1854). The notion of entropy finally was introduced by Clausius in 1865.

At about the same time the kinetic theory was put forward by Maxwell (1860) and Clausius. An important step towards the understanding of irreversibility was Maxwell's Demon (Maxwell in a letter to Tail, 1867) which illustrated the statistical nature of it.

The main contributions, however, in this direction, are due to L. Boltzmann: the Boltzmann equation and the H theorem (1872), the relation between entropy and probability theory (1877), etc. etc.

Boltzmann's ideas caused many controversies and there were many objections, such as the reversibility paradox (Loschmidt, 1876) and the recurrence paradox (Zermelo, 1896; based on work of Poincaré: 1890), etc. and as we have seen, there are still open problems with them, at least if one desires full mathematical rigor.

The next step towards the modern concept of entropy was taken by J. W. Gibbs (1902), who adopted the ensemble point of view and gave a definition of entropy as the *average index of probability-in-phase* (this probability-in-phase is just the classical probability distribution of Sec. A: the "index" is $\ln 1/\rho$, hence the average is $\int \rho \ln 1/\rho$).

How the paradox that entropy, after all, should remain constant, could be resolved, was pointed out by the Ehrenfests (1911), who recognized the role of coarse-graining.

Finally the quantum-mechanical expression for the entropy was given by von Neumann in 1927.

As far as more recent developments are concerned, I have tried to give the relevant literature in the text. There have been significant contributions concerning classical entropy and classical statistical mechanics, and there has been a strong impetus in creating such fields as ergodic theory and theory of dynamical systems. In contrast, the properties of quantum-mechanical entropy have not been investigated in detail for a very long time, and it certainly is to the credit of E. Leib to put forward their study in the last few years.

II. PROPERTIES OF ENTROPY

A. Simple properties

In this chapter we come to the very object of this review, namely, to describe the various general properties of entropy. Let me start with a few extremely simple ones. Some of them we already have discussed in Sec. I as, for instance, the following:

Entropy is defined for every density matrix, it is always ≥ 0, possibly $= \infty$. For the pure states, and only for them, $S = 0$.

Again this shows a weak point in a purely classical approach because in the classical case the "pure states" certainly are density distributions that are concentrated at one point (i.e., δ functions), and their entropy is $= -\infty$. This does not fit into the interpretation of entropy as "lack of information."

One easily verifies that the range of $S(\rho)$ is the whole extended real half-line $[0, \infty]$, i.e., to every number c: $0 \leq c \leq \infty$ there exists a density matrix ρ such that $S(\rho) = c$.

The range of the generalized Boltzmann–Gibbs–Shannon (cf. Sec. I.A) entropy is $[u, v]$ or (u, v), $u \equiv \inf \ln\mu(X)$ [X = non-null measurable subset of Ω, v: $\ln\mu(\Omega)$, depending on whether the infimum is attained or not (Ochs, 1976)]. For the classical Boltzmann–Gibbs entropy, in particular, it is all of R.

An important property of entropy is invariance. Since $S(\rho) = -\sum p_\alpha \ln p_\alpha$, p_α being the eigenvalues of ρ, S only de-

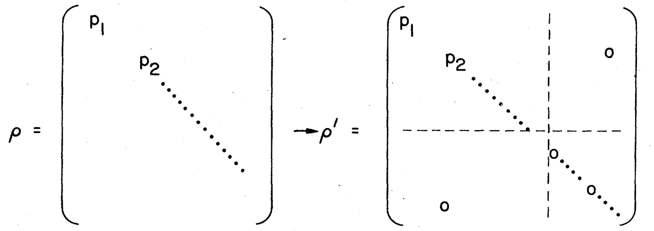

FIG. 5. Construction of $\rho' = \rho \oplus 0$.

pends on the (strictly) positive part of the spectrum of ρ. Any mapping $\rho \to \rho'$ that leaves the positive spectrum unchanged also leaves the entropy unchanged. Examples for such mappings are the following:

1. $\rho' = U^*\rho U$ (U = unitary). (cf. Sec. I.B.) By the way, for the coarse-grained entropy the above invariance property is not true, hence, it may change with time. If, in particular, U is a permutation matrix with respect to the eigenvectors of ρ, then the invariance property is also called symmetry. In other words: S is a symmetric function of the p_α's.

2. Let $\mathrm{H}' = \mathrm{H} \oplus \mathrm{H}''$, $\rho' = \rho \oplus 0$. In that case the invariance property is called expansibility. In graphical language one adds zeros to ρ as seen in Fig. 5.

Another simple property is insensitivity. S is a continuous function of *finitely* many eigenvalues p_α, provided that the rest of them are kept fixed. (This, however, does not mean that S is a continuous function of ρ; cf. Sec. D.)

Let us give an example of a function of a density matrix for which the insensitivity does not hold: The quantum-mechanical version of the Hartley entropy $S_0(\rho)$ is defined by $S_0(\rho)$ = logarithm of the number of eigenvalues that are $\neq 0$. If, for instance, $p_1 = \cdots = p_n = 1/n, p_{n+1} = p_{n+2} = \cdots = 0$, then $S_0(\rho) = S(\rho)$. If now $p_3 = p_4 = \cdots = 0$, then $S(\rho) = -p_1 \ln p_1 - p_2 \ln p_2$ is continuous in p_1 and p_2, whereas $S_0(\rho) = \ln 2$ if $0 < p_1, p_2 < 1$, otherwise $=0$, hence $S_0(\rho)$ is discontinuous in p_1 and p_2. However, it should be remarked that, apart from insensitivity, many properties of the Hartley entropy are shared in common with the "right" entropy, such as additivity, subadditivity, and the above invariance property.

If ρ is of finite rank, i.e., if only finitely many eigenvalues are >0, then $S(\rho) < \infty$. Now let ρ be an arbitrary density matrix. The canonical approximations are obtained as follows:

$$\rho^{(N)} \equiv \sum_{k=1}^{N} p_k |k\rangle\langle k| \Big/ \sum_{k=1}^{N} p_k .$$

Since

$$S(\rho^{(N)}) = -\sum_{k=1}^{N} \frac{p_k}{\sum_{k=1}^{N} p_k} \ln \frac{p_k}{\sum_{k=1}^{N} p_k}$$

$$= -\sum_{k=1}^{N} (p_k \ln p_k) \Big/ \sum_{k=1}^{N} p_k + \ln \sum_{k=1}^{N} p_k ,$$

this tends to $S(\rho)$ because $-\sum_{k=1}^{N} p_k \ln p_k \to S(\rho)$ and $\sum_{k=1}^{N} p_k \to 1$. This fact may be of use in generalizing theorems that are first established for finite-dimensional matrices only, to the general case.

In the next section we now want to turn to less trivial properties of entropy.

B. Concavity

Concavity states that for the density matrix $\rho = \lambda_1\rho_1 + \lambda_2\rho_2 (\rho_1, \rho_2$ density matrices $\lambda_1, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1)$

$$S(\rho) \geq \lambda_1 S(\rho_1) + \lambda_2 S(\rho_2) \tag{2.1}$$

(see Lieb, 1975, or Ruelle, 1969). [There is an equality for $\lambda_1, \lambda_2 > 0$ only if $\rho_1 = \rho_2$ or $S(\rho_1)$, or $S(\rho_2)$, respectively is equal to ∞.]

The proof is very simple. Let $\rho = \sum p_k |k\rangle\langle k|$. Then, because of the concavity of the function $s(x) = -x \ln x$, $S(\rho) = -\sum p_k \ln p_k = \sum s(\langle k|\rho|k\rangle) \geq \lambda_1 \sum s(\langle k|\rho_1|k\rangle) + \lambda_2 \sum s(\langle k|\rho_2|k\rangle) \geq \lambda_1 \sum \langle k|s(\rho_1)|k\rangle + \lambda_2 \sum \langle k|s(\rho_2)|k\rangle = \lambda_1 S(\rho_1) + \lambda_2 S(\rho_2)$.

Of course, we did not use any special property of $s(x)$ besides concavity; thus, for any concave function f, the mapping $\rho \to \mathrm{Tr} f(\rho)$ is concave.

Why is concavity considered to be important? Entropy is a measure of lack of information. Hence if two ensembles are fitted together (what in mathematical language is described by the convex combination $\lambda_1\rho_1 + \lambda_2\rho_2$), one loses the information that tells from which ensemble a special sample stems, and therefore entropy increases (Wigner and Yanase, 1963).

Let us illustrate things by a simple example. Let ρ_1, ρ_2 be one-dimensional projections, i.e., pure states. In a case in which $\lambda_1, \lambda_2 > 0$, $\rho = \lambda_1\rho_1 + \lambda_2\rho_2$ is a mixed state (unless $\rho_1 = \rho_2$). Therefore, $S(\rho) > \lambda_1 S(\rho_1) + \lambda_2 S(\rho_2) = 0$. By the way, in that case it is no longer possible to reconstruct ρ_1, ρ_2 from ρ.

Mixing of pure states (the forming of convex combinations of them) yields a mixed state. More generally, it seems to be plausible to argue: if we are given two mixed states that are unitarily equivalent ($\rho_2 = U^*\rho_1 U$), then mixing of them yields a new state that is more strongly mixed than the two original ones, and $S(\rho) \geq \lambda_1 S(\rho_1) + \lambda_2 S(\rho_2) = S(\rho_1)$. We will make this consideration precise in the next section.

It should be anticipated that concavity is a consequence of *subadditivity* (Sec. F).

Clearly concavity generalizes as follows: let $\rho_1, \rho_2, \ldots, \rho_n$ be density matrices, $\lambda_1, \lambda_2, \ldots, \lambda_n$ numbers ≥ 0, with $\sum \lambda_i = 1$. Then

$$S\Big(\sum \lambda_i \rho_i\Big) \geq \sum \lambda_i S(\rho_i) . \tag{2.2}$$

Let us now take a fixed density matrix ρ and let ρ_t be its time evolution $\rho_t = e^{-iHt} \rho\, e^{iHt}$. The time-averaged density matrix $(1/T)\int_0^T \rho_t\, dt$ shall be denoted by $\bar{\rho}_T$. Then $S(\bar{\rho}_T) \geq S(\rho)$ by a straightforward modification of our previous proof. If $\bar{\rho}_\infty = \lim_{T\to\infty} \bar{\rho}_T$ exists, then, for all T, $S(\bar{\rho}_T) \leq S(\bar{\rho}_\infty)$.

In the next section we will see that the mapping $\rho \to \bar{\rho}_T$ is even mixing-enhancing.

Since the entropy of a time-averaged density matrix always is bigger than the entropy of a density matrix at a fixed time, one should not be surprised that many physical measurements yield a value for the entropy that is bigger than the correct one, because they last rather long compared with the "relaxation time." One should bear this in mind in discussions about the second law of thermodynamics.

There is an upper bound for the convex combination $\rho = \sum \lambda_i \rho_i$ given by

$$S(\rho) \leqslant \sum \lambda_i S(\rho_i) - \sum \lambda_i \ln\lambda_i \, . \tag{2.3}$$

(see Lanford and Robinson, 1968). Let us first assume (2.3) in the special case that all ρ_i are one-dimensional projections, $\rho_i = P_i$:

$$S(\rho) \leqslant -\sum \lambda_i \ln\lambda_i \, . \tag{2.4}$$

Inequality (2.3) in the general case follows from the special case (2.4) because if one decomposes the ρ_i into one-dimensional spectral projections,

$$\rho_i = \sum p_k^{(i)} Q_k^{(i)} \, ,$$

then

$$\begin{aligned} S(\rho) = S\Big(\sum_{ik} \lambda_i p_k^{(i)} Q_k^{(i)}\Big) &\leqslant -\sum_{ik} \lambda_i p_k^{(i)} (\ln\lambda_i + \ln p_k^{(i)}) \\ &= -\sum \lambda_i \ln\lambda_i - \sum \lambda_i \sum p_k^{(i)} \ln p_k^{(i)} \\ &= -\sum \lambda_i \ln\lambda_i + \sum \lambda_i S(\rho_i) \, . \end{aligned}$$

Now let us turn to the inequality (2.4). Let $p_1 \geqslant p_2 \geqslant \cdots$ be eigenvalues of ρ, arranged in decreasing order, and let also the numbers $\lambda_1 \geqslant \lambda_2 \geqslant \cdots$ be arranged in decreasing order. We will show that

$$\begin{gathered} p_1 \geqslant \lambda_1, \; p_1 + p_2 \geqslant \lambda_1 + \lambda_2, \ldots \\ p_1 + p_2 + \cdots + p_n \geqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n, \ldots \end{gathered} \tag{2.5}$$

As already discussed in Sec. I.B, this implies that $-\sum p_i \ln p_i \leqslant -\sum \lambda_i \ln\lambda_i$. In order to prove the inequalities (6.5) we will use Ky Fan's inequality (Ky Fan, 1949): for any set of pairwise orthogonal, normed vectors ϕ_i,

$$p_i + \cdots + p_n \geqslant \sum_{i=1}^{n} \langle \phi_i | \rho | \phi_i \rangle \, . \tag{2.6}$$

One can prove this as follows: let $\psi_1, \ldots, \psi_n$ be another set of pairwise orthogonal, normed vectors, all of them contained in the subspace spanned by $\phi_1, \ldots, \phi_n$. Then it is immediately seen that $\sum \langle \phi_i | \rho | \phi_i \rangle = \sum \langle \psi_i | \rho | \psi_i \rangle$. Now choose $\psi_n \perp |1\rangle, |2\rangle, \ldots, |n-1\rangle$ (remember that $|k\rangle$ = eigenvector of ρ belonging to the eigenvalue p_k). This is always possible since the dimension of the subspace spanned by $\phi_1 \cdots \phi_n$ is $> n-1$. Continue by choosing $\psi_{n-1} \perp |1\rangle, |2\rangle \ldots, |n-2\rangle, \psi_n$; $\psi_{n-2} \perp |1\rangle, \ldots, |n-3\rangle, \psi_n, \psi_{n-1}; \ldots$; $\psi_1 \perp \psi_2, \ldots, \psi_n$. Since $\langle \psi_i | \rho | \psi_i \rangle \leqslant p_i$, we arrive at inequality (2.6).

Define $H^{(n)}$ as the subspace generated by $P_1 H, \ldots, P_n H$, and denote by d its dimension. Let $\phi_1, \ldots, \phi_d$ be an orthonormal basis for $H^{(n)}$. Because of Ky Fan's inequality,

$$\begin{aligned} p_1 + \cdots + p_n \geqslant p_1 + \cdots + p_d &\geqslant \sum_{i=1}^{d} \langle \phi_i | \rho | \phi_i \rangle \\ &= \sum_{i=1}^{d} \sum_{j=1}^{\infty} \lambda_j \langle \phi_i | P_j | \phi_i \rangle \\ &\geqslant \sum_{j=1}^{n} \sum_{i=1}^{d} \lambda_j \langle \phi_i | P_j | \phi_i \rangle = \sum_{j=1}^{n} \lambda_j \, . \end{aligned}$$

Remark: Whereas inequalities (2.1) and (2.2) extend to the continuous case, the continuous analogs of (2.3) and (2.4) are false in general. For instance, consider the projection onto coherent states, $|z\rangle\langle z|$, which are pure states, with entropy = 0, and let $f(z)$ be a non-negative function with $\int dz f(z)/\pi = 1$. Let $\rho = \int (dz f(z)/\pi) |z\rangle\langle z|$. Then

$$S(\rho) \geqslant -\int \frac{dz}{\pi} f(z) \ln f(z)$$

rather than $\leqslant$ (Wehrl, 1977).

The term $-\sum \lambda_i \ln\lambda_i$ occurring on the right-hand side of inequality (2.4) is called mixing entropy. It is most important if the ranges of the ρ_i are pairwise orthogonal; in that case, there is equality.

It should be remarked that this fact allows an axiomatic characterization of entropy: Let Φ be a mapping of the set of density matrices into the extended real half-line, which fulfills the invariance and continuity properties of Sec. A. Also let $H = H_1 \oplus \cdots \oplus H_n$, ρ_i = density matrices in H_i, $\rho = \lambda_1 \rho_1 \oplus \cdots \oplus \lambda_n \rho_n$. If $\Phi(\rho)$ always satisfies $\Phi(\rho) = \sum \lambda_i \Phi(\rho_i) + \Phi(\Lambda)$, Λ being a diagonal matrix in the Hilbert space C^n with entries $\lambda_1, \ldots, \lambda_n$, then $\Phi(\rho)$ is a constant multiple of $S(\rho)$. [This is a quantum-mechanical version of the characterization theorem of Faddeev and Khintchine (see Renyi, 1966). The above form of the theorem was written down by Thirring, 1975.] Here is a sketch of the proof: Because of our remarks of the last section we can suppose that all H_i are finite-dimensional. Let ρ be an $n \times n$-density matrix. Because of invariance, $\Phi(\rho)$ is a symmetric function of its eigenvalues only, i.e., $\Phi(\rho) = I_n(p_1, \ldots, p_n)$. Now the mixing property implies that $I_2(1,0) = I_1(1) + 0 \cdot I_1(1) + I_2(1,0)$, i.e., $I_1(1) = 0$. (Take H_1, H_2 to be one-dimensional.) Furthermore we have (using $p' = p_1 + p_2$) $I_n(p_1, \ldots, p_2) = p' I_2(p_1/p', p_2/p') + (1-p') \cdot I_{n-2}(p_3/(1-p'), \ldots, p_n/(1-p')) + I_2(p', 1-p')$ and $I_{n-1}(p', p_3, \ldots, p_n) = p' I_1(1) + I_2(p', 1-p') + (1-p') I_{n-1}(p_3/(1-p'), \ldots, p_n/(1-p'))$, hence $I_n(p_1, \ldots, p_n) = I_{n-1}(p_1 + p_2, p_3, \ldots, p_n) + (p_1 + p_2) I_2(p_1/(p_1+p_2), p_2/(p_1+p_2))$. See Fig. 6; the left-hand side refers to the first equality,

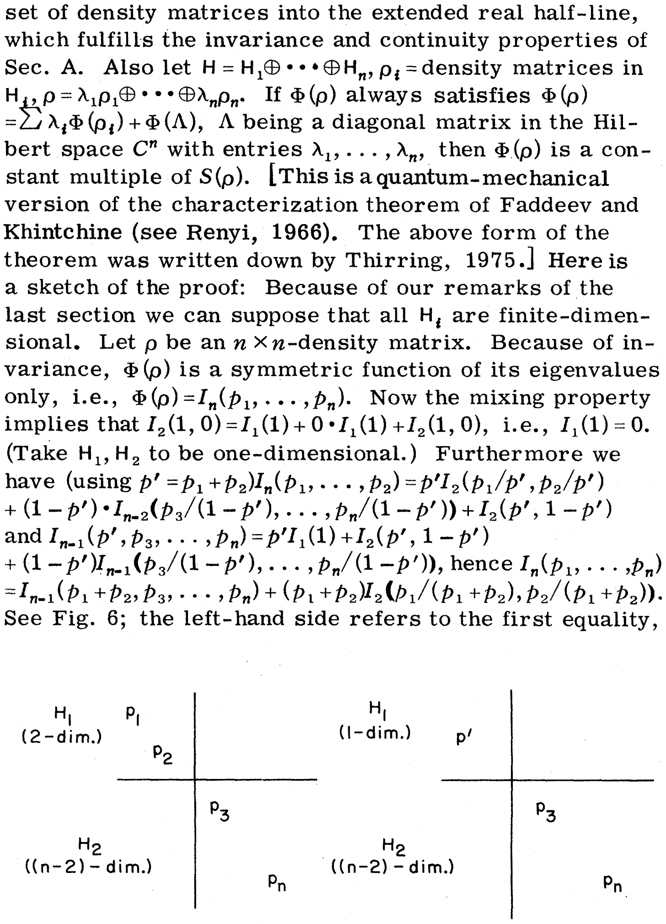

FIG. 6. Illustration of the preceding identity.

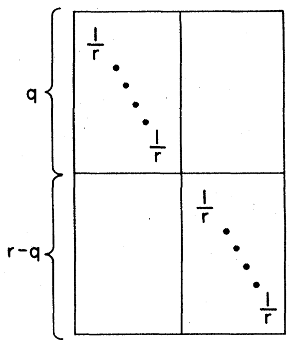

FIG. 7. How to obtain I_2 from α_n.

the right-hand side to the second one. We want to show that the "information function" $I_2(p, 1-p) = c(-p \ln p - (1-p)\ln(1-p))$. Once this is established, one proceeds by induction: $I_3(p_1, p_2, p_3) = I_2(p_1+p_2, p_3) + (p_1+p_2) I_2(p_1/(p_1+p_2), p_2/(p_1+p_2)) = c(-p_1 \ln p_1 - p_2 \ln p_2 - p_3 \ln p_3)$ and so on. Now let $\alpha_n = I_n(1/n, \dots, 1/n)$. One easily verifies that $\alpha_{nm} = \alpha_n + \alpha_m$. It remains to rule out all other solutions of this equation except $\alpha_n = c \ln n$. There is a theorem of Erdös which states that this is the case if $\lim_{n\to\infty}(\alpha_n - \alpha_{n-1}) = 0$. In order to prove the latter relation, we note that, similarly as before,

$$\alpha_n = I_2(1/n, 1-1/n) + (1-1/n)\alpha_{n-1}.$$

Let $\beta_n = \alpha_n - \alpha_{n-1}$, $\gamma_n = I_2(1/n, 1-1/n)$. Here $\gamma_n \to 0$ as $n \to \infty$ by continuity (insensitivity). On the other hand,

$$\gamma_n = \beta_n + \frac{\beta_2 + \beta_3 + \cdots + \beta_{n-1}}{n},$$

so by Mercer's theorem also $\beta_n \to 0$. Turning back to the information function, it suffices, of course, to show that $I_2(p, 1-p) = c[-p \ln p - (1-p)\ln(1-p)]$ for *rational* p. Let $p = q/r$; q, r being integers. Take $\dim \mathsf{H}_1 = q$, $\dim \mathsf{H}_2 = r - q$, and let $\rho = (1/b)1$. Then

$$\alpha_r = \frac{q}{r}\alpha_q + \frac{r-q}{r}\alpha_{r-q} + I_2\left(\frac{q}{r}, 1-\frac{q}{r}\right),$$

and inserting $\alpha_n = c \ln n$,

$$I_2\left(\frac{q}{r}, 1-\frac{q}{r}\right) = c\left[-\frac{q}{r}\ln\frac{q}{r} - \left(1-\frac{q}{r}\right)\ln\left(1-\frac{q}{r}\right)\right]$$

which completes the proof (see Fig. 7).

C. Uhlmann's theory

On several occasions we have already met the notion of *mixing-enhancement*. It states that for the eigenvalues of two density matrices ρ and ρ', arranged in decreasing order, the inequalities

$$\begin{aligned} &p_1(\rho) \le p_1(\rho'), \quad p_1(\rho) + p_2(\rho) \le p_1(\rho') + p_2(\rho'), \dots \\ &p_1(\rho) + \cdots + p_r(\rho) \le p_1(\rho') + \cdots + p_r(\rho'), \dots \end{aligned} \tag{2.7}$$

hold. One then says that ρ is *more mixed* then ρ' (or *more chaotic*), or that ρ' is *purer* than ρ, and writes $\rho \succ \rho'$, or $\rho' \prec \rho$, respectively. Properties of this relation were first investigated by Uhlmann (1971, 1972, 1973); therefore we propose the name "Uhlmann's theory" for this field. [Uhlmann's original papers referred to the finite-dimensional case only; the generalization to the infinite-dimensional case as well as quite a few of the theorems presented below are due to Wehrl (1974).]

Of course, this notion also makes sense in the classical discrete case. There, "density matrix" is replaced by "discrete probability distribution," and "biggest eigenvalue" by "biggest probability," etc. (See, for instance, Ruch, 1975.)

Let us recall where mixing-enhancement has appeared so far:

(1) In our discussion of the master equation.

(2) The mapping $\rho \to f(\rho)/\mathrm{Tr} f(\rho)$, f being concave and non-negative, is mixing-enhancing (lemma 2 of Sec. I.B).

(3) "Deleting off-diagonal matrix elements" implies mixing-enhancement because of Ky Fan's inequality: Let ρ be any density matrix, ϕ_i be an orthonormal basis, and define ρ^d by

$$\langle \phi_i | \rho^d | \phi_k \rangle = \delta_{ik} \langle \phi_i | \rho | \phi_i \rangle.$$

Then the eigenvalues of ρ^d are the diagonal elements $\langle \phi_i | \rho | \phi_i \rangle$ and, by Ky Fan's inequality, the sum of, say, n of them—in particular of the n biggest ones—is $\le p_1(\rho) + \cdots + p_n(\rho)$.

(4) If $\rho = \sum \lambda_i P_i$, (the P_i being one-dimensional projections, $\lambda_i \ge 0$, $\sum \lambda_i = 1$), then $\rho \prec \rho'$ where $\rho' = \sum \lambda_i Q_i$, the Q_i being any family of pairwise orthogonal one-dimensional projections.

Other examples of Uhlmann's relation will be given below. This relation plays a role quite frequently, one reason being its connection with monotone increase of entropy (see Sec. I.B).

Here $\rho \succ \rho'$ implies that $S(\rho) \ge S(\rho')$, but even more, namely that $\mathrm{Tr} f(\rho) \ge \mathrm{Tr} f(\rho')$ for *every* concave function f (we have seen this in Sec. I.B already). This yields, by the way, a characterization of "$\rho \succ \rho'$": $\rho \succ \rho'$ if, and only if, for every concave function f, $\mathrm{Tr} f(\rho) \ge \mathrm{Tr} f(\rho')$.

The proof is obtained by considering functions of the form $f(x) = x$ if $0 \le x \le c$, $f(x) = c$ if $x \ge c$. Suppose that $\rho \succ \rho'$ is not true. Then there exists a smallest integer n such that $p_1(\rho) + \cdots + p_n(\rho) > p_1(\rho') + \cdots + p_n(\rho')$. Now choose $c = p_n(\rho')$. Then $\mathrm{Tr} f(\rho') = n p_n(\rho') + \sum_{i=n+1}^{\infty} p_n(\rho') > n p_n(\rho') + \sum_{i=n+1}^{\infty} p_n(\rho) \ge \mathrm{Tr} f(\rho)$ since $p_n(\rho') < p_n(\rho)$.

Similarly, $\rho \succ \rho'$ if, and only if, for every *convex* function f, $\mathrm{Tr} f(\rho) \le \mathrm{Tr} f(\rho')$. (Unfortunately it is not true that $\mathrm{Tr}\rho^p \le \mathrm{Tr}\rho'^p$ for all p: $1 \le p < \infty$ implies that $\rho \succ \rho'$.)

Now let us consider another example of mixing-enhancement:

(5) Let $U_1, \dots, U_n, \dots$ be unitary operators, $\lambda_i \ge 0$, $\sum \lambda_i = 1$. Then the convex combination $\rho = \lambda_1 U_1^* \rho' U_1 + \cdots + \lambda_n U_n^* \rho' U_n + \cdots$ is more mixed than ρ'.

This is again a consequence of Ky Fan's inequality. Let ϕ_k be the eigenvectors belonging to the $p_k(\rho)$. Then $p_1(\rho) + \cdots + p_n(\rho) = \sum_{k=1}^{n} \sum_i \lambda_i \langle \phi_k | U_i^* \rho' U_i \phi_k \rangle = \sum\sum \lambda_i \langle U_i \phi_k | \rho' U_i \phi_k \rangle \le \sum\sum \lambda_i p_k(\rho') = \sum_{k=1}^{n} p_k(\rho')$.

This explains the word "mixing-enhancement." Remember our remark of Sec. B: ρ' and $U^* \rho' U$, U being an arbitrary unitary transformation, have to be considered as equally mixed; although they do not contain the *same* information, they certainly contain the *same* *amount* of information in any sensible interpretation. (Cf. also Sec. G.) Now ρ is obtained from the $U^*\rho' U$ by a mixing procedure, hence there is a loss of information. (Note that the constituents of ρ cannot be reconstructed.)

By the way, in example 5 it is not necessary that the U_i be unitary; they may be only isometric (i.e., $U_i^* U_i = 1$).

Clearly the relation $\vdash$ is transitive and reflexive, i.e., a preorder. Thus it generates an equivalence relation: $\rho \sim \rho'$ if, and only if, $\rho \vdash \rho'$ and $\rho' \vdash \rho$, hence if ρ and ρ' have the same positive spectrum. This equivalence relation may be regarded as the most general concept of invariance (Sec. A): from the entropy, or more generally from the information-theoretical point of view, density matrices with the same positive spectra are equally good.

Uhlmann's theorem states that, in essence, mixing-enhancement is always produced by the mechanism described in example 5: $\rho \vdash \rho'$ if, and only if, ρ is in the (weak) closure of the convex hull of $\{U^*\rho' U\colon\ U$ unitary$\}$.

We will only sketch the proof. It consists of four steps:

(1) The set A of all operators $A \geqslant 0$ such that $p_1(A) \leqslant p_1(\rho')$, $p_1(A) + p_2(A) \leqslant p_1(\rho') + p_2(\rho'), \ldots$ is by virtue of Ky Fan's inequality convex and weakly closed, hence weakly compact.

(2) Its extremal elements are exactly those A for which $p_i(A) = p_i(\rho')$ for all i or $p_i(A) = p_i(\rho')$ for $i \leqslant n$, $p_i(A) = 0$ otherwise.

(3) Apply the theorem of Krein and Milman: A = closed convex hull of the extremal A.

(4) All extremal A are in the weak closure of $\{U^*\rho' U\}$.

In the finite-dimensional case, there is another way of proving the theorem invoking Birkhoff's theorem mentioned in Sec. I.B. Namely, for two sequences of numbers $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$, or $\beta_1 \geqslant \beta_2 \geqslant \cdots \geqslant \beta_n$ ($\alpha_i \geqslant 0$, $\beta_i \geqslant 0$, $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i = 1$) the relations $\alpha_1 \leqslant \beta_1$, $\alpha_1 + \alpha_2 \leqslant \beta_1 + \beta_2$ etc. are true if and only if there is a doubly-stochastic matrix T such that $\alpha_i = \sum_j T_{ij}\beta_j$ (see Hardy, Littlewood, and Polya, 1934). Then a straightforward application of Birkhoff's theorem yields Uhlmann's theorem.

Let us now make a short remark on the order structure of density matrices (this expression is due to Thirring, 1975. The lattice structure of density matrices was recognized by Wehrl, 1974):

For any two density matrices ρ_1, ρ_2 there exist (up to equivalence) a "purest" density matrix $\vdash \rho_1, \rho_2$ and a "most mixed" one $\dashv \rho_1, \rho_2$. Thus the equivalence classes of density matrices form a lattice. Its "purest" element clearly is the equivalence class of the pure states. A most mixed element does not exist in infinite-dimensional Hilbert space, only in the finite-dimensional case, namely, $\rho = 1/\dim \mathsf{H}$.

Next, let us generalize example 3: let P_i be a family of pairwise orthogonal projections (not necessarily one-dimensional) with $\sum P_i = 1$. Then $\sum P_i \rho P_i \vdash \rho$.

This means intuitively that deleting off-diagonal matrix elements reduces the information and increases the degree of mixture. The proof is easily obtained by means of Ky Fan's inequality.

The coarse-grained density matrix $\rho_{\mathrm{cg}} = \sum \lambda_i P_i$ ($\lambda_i \operatorname{Tr} P_i = \operatorname{Tr} \rho P_i$, cf. Sec. I.B is $\vdash \sum P_i \rho P_i$, hence $\vdash \rho$; therefore not only $S(\rho_{\mathrm{cg}}) \geqslant S(\rho)$, but also $\operatorname{Tr} f(\rho_{\mathrm{cg}}) \geqslant \operatorname{Tr} f(\rho)$ for any concave function f. (Remember Fig. 3: both mappings are mixing-enhancing.)

It is worth mentioning that Uhlmann's theory has been generalized to arbitrary von Neumann algebras by Wehrl (1975), Alberti (thesis, 1973), and Uhlmann himself. It turns out that this theory provides a powerful tool in the investigation of the structure of von Neumann algebras and, in a certain sense, is the "dual" of the von Neumann–Murray dimension theory.

D. Continuity properties

In infinite-dimensional Hilbert spaces, entropy, as a function of density matrices, is discontinuous in the usual topologies. There are only a few restricted continuity properties. The problems that arise in this connection may be divided into two groups:

(1) Those which are of more mathematical interest and which we will not treat in great detail here.

(2) Technical considerations that are of use in extending theorems that can be proven for finite-dimensional matrices, to the general case. (Cf. the end of Sec. A. For a typical example, see Sec. III.A.)

From section A we already know insensitivity. Other restricted continuity properties are:

Lower Semicontinuity. (This fact seems to have been well known for a long time, but was written down only by Naudts, 1969; Wehrl, 1976. For other proofs, cf. Secs. III.B and IV.B). Let ρ_n, ρ be density matrices, such that $\operatorname{Tr}|\rho - \rho_n| \to 0$. Then $S(\rho) \leqslant \liminf S(\rho_n)$.

Ky Fan's inequality tells us that, for the eigenvalues of ρ_n, or ρ, respectively, arranged in decreasing order,

$$p_1(\rho_n) + \cdots + p_k(\rho_n) \leqslant \operatorname{Tr}|\rho_n - \rho| + p_1(\rho) + \cdots + p_k(\rho)$$

since $\rho_n = (\rho_n - \rho) + \rho$. Vice versa,

$$p_1(\rho) + \cdots + p_n(\rho) \leqslant \operatorname{Tr}|\rho_n - \rho| + p_1(\rho_1) + \cdots + p_k(\rho_n)\,,$$

hence $|(p_1(\rho) - p_1(\rho_n)) + \cdots + (p_k(\rho) - p_k(\rho_n)| \to 0$ and eventually $p_k(\rho_n) \to p_k(\rho)$. Thus

$$-\sum_{i=1}^{k} p_i(\rho) \ln p_i(\rho) = \lim\Big[-\sum_{i=1}^{k} p_i(\rho_n) \ln p_i(\rho_n)\Big]$$

and

$$S(\rho) = \sup_k \sum_{i=1}^{k} (-p_i(\rho) \ln p_i(\rho))$$

$$\leqslant \liminf \sup_k \sum_{i=1}^{k} (-p_i(\rho_n) \ln p_i(\rho_n))\,.$$

Remark: It is also true that, if $\rho_n \xrightarrow{\text{weakly}} \rho$, $S(\rho) \leqslant \liminf S(\rho_n)$, *provided that ρ is a density matrix.* In this case, also $\operatorname{Tr}|\rho_n - \rho| \to 0$. If ρ is not a density matrix, it can happen that $S(\rho) > \liminf S(\rho_n)$(Wehrl, 1976; see also Davies, 1972, and dell 'Antonio, 1967).

Unboundedness in Every Neighborhood. Let ρ be a density matrix and $\epsilon > 0$ be an arbitrary number. Then there always exists another density matrix ρ' with $\mathrm{Tr}|\rho - \rho'| < \epsilon$ and $S(\rho') = \infty$. (Clearly this implies that $S(\rho)$ is discontinuous.)

For this one only has to change, beginning from a sufficiently large index l, the eigenvalues of ρ in such a manner that $p_1' = p_1, \ldots, p_l' = p_l$, and

$$p_k' \sim \frac{1}{k\,(\ln k)^2} \quad \text{for } k > l .$$

Entropy is Almost Always Infinite. Due to lower semicontinuity, the sets $\{\rho\colon S(\rho) \leqslant n\}$ are closed; also they are nowhere dense because of our above statement, hence $\{\rho\colon S(\rho) < \infty\} = \cup\{\rho\colon S(\rho) \leqslant n\}$ is a *set of first category* (i.e., the topological analog of a set of measure 0).

The set of density matrices with finite entropy also has an interesting algebraic property: its finite linear combinations are a two-sided ideal in the set of all bounded operators (G. and G. Lassner, 1977).

The proof is obtained by using the following criterion (see Dixmier, 1957): A set of positive operators, say J^+, is the positive part of a two-sided ideal J, if and only if the following conditions are fulfilled:

(i) If $A \in J^+$, then $U^*AU \in J^+$ for all unitaries U.
(ii) If $A \in J^+$ and $0 \leqslant B \leqslant A$, then $B \in J^+$.
(iii) If $A, B \in J^+$, then also $A + B \in J^+$.

(Note that any element of an ideal is a finite linear combination of elements of its positive part.) Now if $S(A) < \infty$, then $S(U^*AU) = S(A) < \infty$. Furthermore, let $B \leqslant A$. Denote by $\beta_1 \geqslant \beta_2 \geqslant \ldots$, or $\alpha_1 \geqslant \alpha_2 \geqslant \ldots$, the eigenvalues of B, or A, respectively. We have $\beta_i \leqslant \alpha_i$, hence for all except possibly a finite number of indices, $s(\beta_i) \leqslant s(\alpha_i)$, thus $S(B) < \infty$. Concerning condition (iii), due to the fact that $S(\lambda A) = \lambda S(A) - \lambda \ln\lambda$ for $\lambda \geqslant 0$, we can assume that $\mathrm{Tr}(A + B) = 1$. (One easily verifies that $S(A) < \infty$ implies that $\mathrm{Tr}A < \infty$.) Define $A' = A/\mathrm{Tr}A$, $B' = B/\mathrm{Tr}B$. By assumption, $S(A') < \infty$ and $S(B') < \infty$. Then, $S(A + B) = S((\mathrm{Tr}A)A' + (\mathrm{Tr}B)B') \leqslant (\mathrm{Tr}A)S(A') + (\mathrm{Tr}B)S(B') - \mathrm{Tr}A\,(\ln \mathrm{Tr}A) - \mathrm{Tr}B\,(\ln \mathrm{Tr}B)$ [inequality (2.3)], and, consequently, is $<\infty$.

From these results one may doubt that entropy is a sensible concept in infinite-dimensional Hilbert spaces. But fortunately these theorems do not really affect physics. Let H be a reasonable Hamiltonian such that $\mathrm{Tr}e^{-\beta H} < \infty$, and let $\sigma_\beta = e^{-\beta H}/\mathrm{Tr}e^{-\beta H}$ be the Gibbs state. Suppose that the energy $\mathrm{Tr}\sigma_\beta H$ is finite. Whenever, for some density matrix ρ, $\mathrm{Tr}\rho H \leqslant \mathrm{Tr}\sigma_\beta H$, then $S(\rho) \leqslant S(\sigma_\beta) < \infty$. Hence a density matrix ρ with $S(\rho) = \infty$ would also have infinite energy. The assertion that $S(\rho)$ is mostly infinite therefore is as good as the assertion that $\mathrm{Tr}\rho H$ is mostly infinite. However, this is a trivial fact since H is an unbounded operator, but it has no physical significance at all.

Also the discontinuity of $S(\rho)$ is not as bad as it may look at first. We will show in Sec. III.B that the relative entropy $S(\sigma|\rho) = \mathrm{Tr}\rho(\ln\rho - \ln\sigma)$, and, consequently, the free energy $F(\rho, \beta, H) = \mathrm{Tr}\rho H - B^{-1}S(\rho)$ (cf. Sec. I.B) is lower semicontinuous in ρ. Therefore if not only $\mathrm{Tr}|\rho_n - \rho| \to 0$ but also $\mathrm{Tr}\rho_n H \to \mathrm{Tr}\rho H$, i.e., if the energy expectation values converge, then $S(\rho) \leqslant \lim\inf S(\rho_n)$, $-S(\rho) \leqslant \lim\inf[-S(\rho_n)] = -\lim\sup S(\rho_n)$, hence $S(\rho) = \lim S(\rho_n)$.

One can even dispense with the requirement that $\mathrm{Tr}\rho_n H \to \mathrm{Tr}\rho H$ if $\mathrm{Tr}e^{-\beta H} < \infty$ for *all* β: $0 < \beta < \infty$. Then $S(\rho)$ in continuous on the sets $\{\rho\colon \mathrm{Tr}\rho H \leqslant C < \infty\}$, even if $\mathrm{Tr}\rho_n H \to \mathrm{Tr}\rho H$. Namely, $\mathrm{Tr}\rho(\beta H) - S(\rho) \leqslant \lim\inf[\mathrm{Tr}\rho_n(\beta H) - S(\rho_n)]$; hence $-S(\rho) \leqslant \lim\inf[-S(\rho_n)] + \lim\sup|\mathrm{Tr}\rho_n(\beta H)| + \mathrm{Tr}\rho(\beta H)$ for all $\beta > 0$. Since the sum of the last two terms is $\leqslant 2\beta C$, we have as above $-S(\rho) \leqslant \lim\inf[-S(\rho_n)]$ and consequently $S(\rho) = \lim S(\rho_n)$.

Now let us turn to the second group of theorems concerning continuity properties, those that are of use for practical computations. Let me mention two of them (Simon, 1973):

1. Let A_n, A be compact operators, $\geqslant 0$ (not necessarily being density matrices), $w - \lim A_n = A$. Suppose that for the eigenvalues, arranged in decreasing order, $p_k(A_n) \leqslant p_k(A)$ for all k. Then $S(A_n) \to S(A)$.

2. Dominated convergence theorem for entropy. Again let A_n, A as above, and $w - \lim A_n = A$. If there is a compact operator B such that $A_n \leqslant B$ and $S(B) < \infty$, then $S(A_n) \to S(A)$.

E. Additivity

Additivity states the following: let H_1, H_2 be two Hilbert spaces, and ρ_1, ρ_2 be two density matrices. Then the entropy of the density matrix $\rho_1 \otimes \rho_2$ in $\mathrm{H}_1 \otimes \mathrm{H}_2$ is

$$S(\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2) . \tag{2.8}$$

Of course, this generalizes to

$$S(\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n) = S(\rho_1) + S(\rho_2) + \cdots + S(\rho_n) . \tag{2.9}$$

The proof is very simple: let ϕ_k or ψ_j be the eigenvectors of ρ_1, or ρ_2, respectively, belonging to the eigenvalues p_k, or q_j, respectively. Then the $\phi_k \otimes \psi_k$ are the eigenvectors of $\rho_1 \otimes \rho_2$ and the corresponding eigenvalues are $p_k q_j$.

$$\begin{aligned} S(\rho_1 \otimes \rho_2) &= -\sum_{kj} p_k q_j \ln(p_k q_j) \\ &= -\sum p_k \ln p_k - \sum q_j \ln q_j = S(\rho_1) + S(\rho_2) . \end{aligned}$$

From the information-theoretical point of view, this property is quite clear: if we are given two independent systems, described by H_1, ρ_1 or H_2, ρ_2, respectively, then the information about the total system, described by $\mathrm{H}_1 \otimes \mathrm{H}_2$, $\rho_1 \otimes \rho_2$, equals the sum of the information about its constituents. As concerns physics, additivity must not be confused with the scaling behavior of entropy. It is often thought to be of *apodictical truth* that the entropy of a piece of matter composed of two parts equals the sum of the entropies of these parts. However, *it is not*, as we have seen in Sec. I.C. It is approximately true for "normal" matter, and exactly true only if there are no correlations at all between the parts. If there are considerable correlations between them, the entropy of the total system may be *much smaller* (even 0).

In the classical case additivity reads as follows: let Ω_1, Ω_2 be two phase spaces (with elements w_1, w_2), and let $\rho_1(w_1), \rho_2(w_2)$ be two probability distributions. For $\rho(w_1, w_2) \equiv \rho_1(w_1)\rho_2(w_2)$, $S(\rho) = S(\rho_1) + S(\rho_2)$.

In the earlier sections the special role of entropy did not appear; rather, the traces of any concave function was taken to be more or less as good as entropy. However, the property of additivity distinguishes entropy among all functionals of the form $\rho \rightarrow \mathrm{Tr} f(\rho)$, where f is a measurable function: if $\mathrm{Tr} f(\rho_1 \otimes \rho_2) = \mathrm{Tr} f(\rho_1) + \mathrm{Tr} f(\rho_2)$, then $f(\rho) = \mathrm{const}\, \rho \ln \rho$.

Due to the assumption $\sum f(p_k q_j) = \sum f(p_k) + \sum f(q_j) = \sum [q_j f(p_k) + p_k f(q_j)]$ for all sequences p_k, q_j, hence $f(p_k q_j) = q_j f(p_k) + p_k f(q_j)$. For $g(x) \equiv f(x)/x$, $g(p_k q_j) = g(p_k) + g(q_j)$, i.e., $g(x) = \mathrm{const} \ln x, f(x) = \mathrm{const}\, x \ln x$.

There are, of course, other additive functionals of ρ, but they are not of the form $\rho \rightarrow \mathrm{Tr} f(\rho)$.

An example is provided by the so-called α entropies. [In the classical case they were introduced by Renyi. See Renyi, 1966. For the quantum-mechanical case, see Wehrl, 1976; Thirring, 1975. The case $\alpha = 0$ was (classically) invented by Hartley, 1928. The case $\alpha = 2$ has been considered occasionally in the past, for instance by Fano, 1957, and Prigogine, 1972.] $S_\alpha(\rho) \equiv 1/(1-\alpha) \ln \mathrm{Tr} \rho^\alpha$ for $\alpha > 0, \neq 1, \neq \infty$; $S_0(\rho) \equiv$ Hartley entropy (see Sec. A), $S_1(\rho) \equiv S(\rho), S_\infty(\rho) \equiv -\ln \|\rho\|$. (One verifies easily that, for fixed ρ, the mapping $\alpha \rightarrow S_\alpha(\rho)$ is continuous for $1 \leq \alpha \leq \infty$.) In infinite-dimensional Hilbert space, these α entropies, for $\alpha > 1$, are not concave, however (and also not subadditive; this property only holds for S_0 and S. Cf. next section). On the contrary, they are compatible with Uhlmann's relation: $\rho \vdash \rho' \Rightarrow S_\alpha(\rho) \geq S_\alpha(\rho')$.

F. Subadditivity

Like additivity, subadditivity refers to a system composed of two subsystems; however, they are no longer supposed to be independent. That means that, in the Hilbert space $\mathsf{H} = \mathsf{H}_1 \otimes \mathsf{H}_2$, we consider now density matrices of a more general kind than $\rho_1 \otimes \rho_2$.

What information about, say, the first subsystem is contained in ρ? It is given by the partial trace (reduced density matrix) $\rho_1 \equiv \mathrm{Tr}_{\mathsf{H}_2} \rho$. ρ_1 is a density matrix in H_1 with matrix elements $\langle \phi | \rho_1 | \psi \rangle$ $(\phi, \psi \in \mathsf{H}_1)$ defined as follows: let $e_1, e_2, \ldots$ be an orthonormal basis in H_2. Then

$$\langle \phi | \rho_1 \psi \rangle = \sum \langle \phi \otimes e_i | \rho | \psi \otimes e_i \rangle \ . \tag{2.10}$$

It can be shown that the right side of (2.10) is independent of the e_i basis. (ρ_1 is, so to say, some average over the second system.) In an analogous way, $\rho_2 \equiv \mathrm{Tr}_{\mathsf{H}_1} \rho$ is constructed. Of course, for density matrices of the form $\rho = \rho_1 \otimes \rho_2$, $\mathrm{Tr}_{\mathsf{H}_2} \rho = \rho_1$, and $\mathrm{Tr}_{\mathsf{H}_1} \rho = \rho_2$. [The notion of partial trace is a special case of a so-called "conditional expectation," (Umegaki, 1954; cf. Takesaki, 1972; Guichardet, 1974).]

If $T \in \mathsf{B}(\mathsf{H})$ is an operator of the form $T_1 \otimes 1$ $(T_1 \in \mathsf{B}(\mathsf{H}_1))$, then $\mathrm{Tr} \rho T = \mathrm{Tr} \rho_1 T_1$. (This property may equally be used as a *definition* of ρ_1.) Since the state of a system is determined by the knowledge of all expectation values of observables, one can say that in that sense ρ_1 contains all the information about the first system, and ρ_2 all the information about the second one.

In the classical case one constructs the reduced distributions $\rho_1(w_2)$, or $\rho_2(w_2)$, from a distribution $\rho(w_1, w_2)$, by integrating over the other variable:

$$\rho_1(w_1) = \int dw_2\, \rho(w_1, w_2) \tag{2.11}$$

and vice versa.

Now *subadditivity* states that

$$S(\rho) \leq S(\rho_1) + S(\rho_2) = S(\rho_1 \otimes \rho_2) \ . \tag{2.12}$$

This appears plausible since, when forming ρ_1 and ρ_2, one loses the information about the correlations. (Also, one cannot reconstruct ρ from ρ_1 and ρ_2.) However, it is *false* that $\rho_1 \otimes \rho_2 \vdash \rho$. (Lieb, private communication. This follows also from the fact that the α entropies for $\alpha \neq 0, 1$ are not subadditive; see below.)

A proof is obtained from the inequality for the relative entropy (1.41) $S(\rho_1 \otimes \rho_2 | \rho) \geq 0$. Now $S(\rho_1 \otimes \rho_2 | \rho) = \mathrm{Tr} \rho [\ln \rho - \ln(\rho_1 \otimes \rho_2)] = \mathrm{Tr} \rho (\ln \rho - \ln \rho_1 \otimes 1 - 1 \otimes \ln \rho_2) = \mathrm{Tr} \rho \ln \rho - \mathrm{Tr} \rho_1 \ln \rho_1 - \mathrm{Tr} \rho_2 \ln \rho_2$. In the classical case the proof is quite analogous, even for the generalized Boltzmann–Gibbs–Shannon entropy.

Subadditivity is a stronger property than concavity. In fact, our next example shows that concavity is a consequence of subadditivity.

In order to verify this, take two Hilbert spaces H_1 and H_2, and let $e_1, e_2, \ldots$ be an orthonormal basis in H_2. Then $\mathsf{H} = \mathsf{H}_1 \otimes \mathsf{H}_2 = \oplus \mathsf{H}^{(i)}$, $\mathsf{H}^{(i)} \equiv \mathsf{H}_1 \otimes e_i$. Now let ρ be the density matrix $\oplus \lambda_i \rho_i$ (ρ_i = density matrix in $\mathsf{H}_1, \lambda_i \geq 0$, $\sum \lambda_i = 1$). $\mathrm{Tr}_{\mathsf{H}_2} \rho = \sum \lambda_i \rho_i$. $\mathrm{Tr}_{\mathsf{H}_1} \rho$ is a matrix in H_2 that is diagonal with respect to the basis e_i, and its entries are the λ_i. Then $S(\rho) = \sum \lambda_i S(\rho_i) - \sum \lambda_i \ln \lambda_i$ by subadditivity is $\leq S(\rho_1) + S(\rho_2) = S(\sum \lambda_i \rho_i) - \sum \lambda_i \ln \lambda_i$, hence $S(\sum \lambda_i \rho_i) \geq \sum \lambda_i S(\rho_i)$. One can argue in the opposite direction: subadditivity is implied by the properties (a) that deleting off-diagonal matrix elements increases entropy (example 3 of Sec. C), (b) the mixing property $S(\sum \lambda_i \rho_i) = \lambda_i \sum S(\rho_i) - \sum \lambda_i \ln \lambda_i$, if the ranges of the ρ_i are orthogonal, and (c) concavity (Wehrl, 1976).

We already have indicated in the preceding section that the Hartley entropy is subadditive too. There are no other functionals than linear combinations of S and S_0 that are invariant (in the sense of Sec. A), additive, and subadditive (except for some trivial possibilities like "entropy" = 0 if ρ has finite rank, otherwise = ∞) (characterization theorem of Aczel, Forte, and Ng, 1974. The quantum-mechanical version is due to Ochs, 1975). If, in addition, insensitivity is required, we are left with the constant multiples of S only.

The proof is rather involved, therefore we can give here only an outline. As in the proof of the characterization theorem of Khintchine and Faddeev (Sec. B), it suffices to consider the finite-dimensional case only. Then $\Phi(\rho)$ is a symmetric function of the eigenvalues of ρ; $\Phi(\rho) = I_n(p_1, \ldots, p_n)$. Because of expansibility (Sec. A),

$$I_n(p_1, \ldots, p_n) = I_{n+1}(p_1, \ldots, p_n, 0) \ .$$

Additivity implies that for n numbers $p_1, \ldots, p_n$ with $\sum p_i = 1$ and m numbers $q_1, \ldots, q_m$ with $\sum q_i = 1$

$$I_{nm}(p_1 q_1, \ldots, p_k q_j, \ldots, p_n q_m) = I_n(p_1, \ldots, p_n) + I_m(q_1, \ldots, q_m) ,$$

whereas subadditivity states that

$$I_{nm}(r_{11}, \ldots, r_{kj}, \ldots, r_{nm}) \leqslant I_n(p_1, \ldots, p_n) + I_m(q_1, \ldots, q_m) .$$

In the last relation, the r_{kj} are a double sequence of non-negative numbers with $\sum_{k,j} r_{kj} = 1$, and $p_i = \sum_k r_{ik}$, $q_j = \sum_k r_{kj}$. As in Sec. B, the first part of the proof consists in showing that, for $0 < p < 1$, the information function $I_2(p, 1-p) = aS(p, 1-p) + A_2$; with $S(p, 1-p) = -p \ln p - (1-p)\ln(1-p)$ and a and A_2 being constants. For this, we need three lemmas:

(1) Let $f(p) = I_2(p, 1-p)$. We note that $f(p) = f(1-p)$. From the next lemma it follows that f is nondecreasing in $[0, 1/2]$ and nonincreasing in $[1/2, 1]$, and that it is concave.

(2) Symmetry, additivity, and subadditivity can be used to obtain the inequality

$$\begin{aligned} I_2(1-q, q) - I_2((1-p)(1-q) + p(1-r), (1-p)q + pr) &\leqslant I_m(p(1-q), pq, p_3, \ldots, p_m) \\ &\quad - I_m(p(1-r), pr, p_3, \ldots, p_m) \\ &\leqslant I_2(p(1-q) + (1-p)(1-r), pq + (1-p)r) - I_2(1-r, r) . \end{aligned}$$

Therefore

$$I_2(1-q, q) - I_2((1-p)(1-q) + p(1-r), (1-p)q + pr) \leqslant I_2(p(1-q) + (1-p)(1-r), pq + (1-p)r) - I_2(1-r, r) .$$

Inserting $r = 1-q$, one arrives at

$$f(q) \leqslant f(p(1-q) + (1-p)q) .$$

Given q, the set $\{p(1-q) + (1-p)q : 0 \leqslant p \leqslant 1\}$ is just $[q, 1-q]$, hence $f(p)$ is nondecreasing in $[0, 1/2]$ and nonincreasing in $[1/2, 1]$. Inserting $p = 1/2$, we obtain

$$f(\tfrac{1}{2} q + \tfrac{1}{2} r) \geqslant \tfrac{1}{2} f(q) + \tfrac{1}{2} f(r) ,$$

hence f is weakly concave, and, because of monotonicity, concave.

(3) The above lemma is also the essential ingredient in proving the following recursive relation: For $0 < q < 1$, the difference

$$I_m((1-q)p, pq, p_3, \ldots, p_m) - p I_2(1-q, q)$$

is independent of q,

$$I_m(\cdots) - p I_2(\cdots) = J_{m-1}(p, p_3, \ldots, p_m) .$$

Now

$$I_3(p_1, p_2, p_3) = (p_1 + p_2) I_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) + J_2(p_1 + p_2, p_3)$$

Let $g(x) \equiv J_2(1-x, x)$. Because of the symmetry of I_3 one obtains the functional equation

$$\begin{aligned} (p_1 + p_2) f\left(\frac{p_2}{p_1 + p_2}\right) + g(p_3) &= I_3(p_1, p_2, p_3) \\ &= I_3(p_1, p_3, p_2) \\ &= (p_1 + p_3) f\left(\frac{p_3}{p_1 + p_3}\right) + g(p_2) , \end{aligned}$$

or, with $p_2 = y$, $p_3 = x$,

$$(1-x) f\left(\frac{y}{1-x}\right) + g(x) = (1-y) f\left(\frac{x}{1-y}\right) + g(y) .$$

Differentiating with respect to x and y yields

$$\frac{y}{(1-x)^2} f''\left(\frac{y}{1-x}\right) = \frac{x}{(1-y)^2} f''\left(\frac{x}{1-y}\right) .$$

(One can prove that all derivatives exist.) For $s = y/1-x$, $t = x/1-y$, this becomes

$$s(1-s) f''(s) = t(1-t) f''(t) .$$

The left-hand side depending on s only, and the right-hand side depending on t only, both sides must be a constant, and by integration one arrives at

$$f(t) = -a[t(\ln t - 1) + (1-t)(\ln(1-t) - 1)] + bt + C .$$

Because of the symmetry, $b = 0$, and thus we have obtained the expression for $f(p)$ asserted above, with $A_2 = C + a$. How do we get from I_2 to I_m? Suppose all $p_i > 0$ and consider the expression

$$\psi_m(p_1 + p_2, p_3, \ldots, p_m) \equiv J_{m-1}(p_1 + p_2, p_3, \ldots, p_m) + (p_1 + p_2) A_2 - aS(p_1 + p_2, p_3, \ldots, p_m) .$$

It is symmetric in $p_3, \ldots, p_m$, and

$$\psi_m(p_1 + p_2, p_3, p_4, \ldots, p_m) = \psi_m(p_1 + p_3, p_2, p_4, \ldots, p_m) .$$

From this one concludes easily that it is a constant, A_m. Therefore

$$\begin{aligned} I_m(p_1, \ldots, p_m) &= (p_1 + p_2) aS\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) \\ &\quad + aS(p_1 + p_2, p_3, \ldots, p_m) + A_m \\ &= aS(p_1, p_2, \ldots, p_m) + A_m . \end{aligned}$$

It is immediately seen that $A_{mn} = A_m + A_n$. Inequality (2) and our last result yield

$$\liminf (A_{m+1} - A_m) \geqslant 0 .$$

This relation is similar to the one we have obtained in Sec. B. There we used a theorem of Erdös to conclude that $A_m = b \ln m$. In our case this is guaranteed by a stronger theorem of Katai, 1967. The proof is now completed by a simple application of expansibility:

$$I_m(p_1, \ldots, p_m) = aS(p_1, \ldots, p_m) + A_k ,$$

where k = number of p's that are $\neq 0$. A final remark applies: there are various other characterization theorems (see Aczel and Daroczy, 1974). However, from the physical point of view, the two theorems of this chapter seem to be the only "natural" ones.

Before discussing some applications of subadditivity, let me make some remarks on the question of monotonicity. Neither quantum-mechanically nor in the classical continuous case is it true that $S(\rho_1) \leqslant S(\rho)$! In the classical case, one could, for instance, have $S(\rho_2) < 0$, then $S(\rho)$, for $\rho(w_1, w_2) = \rho_1(w_1)\rho_2(w_2)$, is $< S(\rho_1)$. (One can easily give other examples not involving negative values of entropy.) In quantum mechanics, ρ may be pure, but ρ_1 may not be, hence $S(\rho_1) > 0 = S(\rho)$.

Take, as an example, $\mathsf{H}_1 = \mathsf{H}_2 = C^2$, ϕ_1, ϕ_2 (or ψ_1, ψ_2, respectively) = orthonormal basis in H_1 (or H_2, respectively). Let ρ = projection onto

$$\frac{1}{\sqrt{2}}(\phi_1 \otimes \psi_1 + \phi_2 \otimes \psi_2) \cdot \rho_1 = \frac{1}{2}(|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|),$$

$$S(\rho_1) = \ln 2 .$$

This is a strange phenomenon, even though its formal origin is quite simple. In Sec. III.A we will discuss why it is not observed in "real matter."

In connection with the heat death, it has been noted that one could imagine that the universe is in a pure state, with entropy = 0, but nevertheless the entropy of sufficiently small subsystems (earth, galaxies, ...) increases (Lieb, 1975). After all, such a possibility is not excluded. However, this field certainly is very speculative and I do not want to proceed further in this direction.

Monotonicity *is* valid for the classical discrete case and (in the opposite direction) for the relative entropy (even in the quantum-mechanical case): $S(\sigma_1 | \rho_1) \leq S(\sigma | \rho)$. (For a proof, see Sec. III.B.) It is also valid for the "right" classical continuous entropy (not the conventional one), i.e., for ρ^{cl} as defined in Sec. I.A (Wehrl, 1977).

Let us come back to our example above (ρ pure, ρ_1 not). Two remarks are appropriate:

(1) If ρ is pure, then $S(\rho_1) = S(\rho_2)$; moreover, the positive spectra of ρ_1 and ρ_2 coincide.

Let ϕ_i, ψ_k be orthonormal bases in H_1, or H_2, respectively. Let χ be the vector $\sum c_{ik} \phi_i \otimes \psi_k$. $\rho \equiv |\chi\rangle\langle\chi| = \sum c_{ik} c_{jl}^* |\phi_i\rangle\langle\phi_j| \otimes |\psi_k\rangle\langle\psi_l|$. Then $\rho_1 = \sum c_{ik} c_{jl}^* |\phi_i\rangle\langle\phi_j| \delta_{kl} = \sum c_{ik} c_{jk}^* |\phi_i\rangle\langle\phi_j|$. Let C be an infinite matrix with entries c_{ik}. The eigenvalues of ρ_1 equal those of CC^*. Similarly, one finds that the eigenvalues of ρ_2 equal those of C^*C. But it is well known that CC^* and C^*C have the same positive spectrum.

(2) Given ρ_1, one always can find a Hilbert space H_2 and a pure density matrix in $\mathsf{H}_1 \otimes \mathsf{H}_2$ such that $\rho_1 = \mathrm{Tr}\mathsf{H}_2 \rho$.

Let $\rho_1 = \sum p_i |\phi_i\rangle\langle\phi_i|$. Take for H_2 a Hilbert space with the same dimension as H_1, and with an orthonormal basis $\psi_1, \psi_2, \ldots$

$$\rho_1 = \mathrm{Tr}_{\mathsf{H}_2} \rho \quad \text{if } \rho = |\chi\rangle\langle\chi|, \chi \equiv \sum \sqrt{p_i}\, \phi_i \otimes \psi_i .$$

From remarks 1 and 2 one can derive the triangle inequality (Araki and Lieb, 1970) which gives a partial compensation for the failure of monotonicity:

$$|S(\rho_1) - S(\rho_2)| \leq S(\rho) \leq S(\rho_1) + S(\rho_2) . \tag{2.13}$$

(Of course, the right-hand side is merely subadditivity.)

We want to prove the inequality $S(\rho_1) \leq S(\rho) + S(\rho_2)$; interchanging remarks 1 and 2 yields the rest. ρ is a density matrix in $\mathsf{H}_1 \otimes \mathsf{H}_2$. Due to remark 2 there exists a Hilbert space H_3 and a pure density matrix σ in $\mathsf{H}_1 \otimes \mathsf{H}_2 \otimes \mathsf{H}_3$ such that $\rho = \mathrm{Tr}_{\mathsf{H}_3} \sigma$. Let $\sigma_3 = \mathrm{Tr}_{\mathsf{H}_1 \otimes \mathsf{H}_2} \rho$. $S(\sigma_3) = S(\rho)$ because of remark 1. $\rho_1 = \mathrm{Tr}_{\mathsf{H}_2 \otimes \mathsf{H}_3} \sigma$. $S(\rho_1) = S(\sigma_{23})$, $\sigma_{23} \equiv \mathrm{Tr}_{\mathsf{H}_1} \sigma$. By subadditivity, $S(\rho_1) = S(\sigma_{23}) \leq S(\rho_2) + S(\sigma_3) = S(\rho_2) + S(\rho)$.

It should be noted that the triangle inequality is *false* in the classical continuous case, because the analog of 2 above does not hold.

Now let us describe some applications of subadditivity.

1. Existence of mean entropy for translationally invariant systems

Let $V \subset \mathrm{R}^d$ (or Z^d) be a bounded region. We attach to it the Fock space $\mathsf{H}(V)$ (Sec. I.A; as concerns statistics, we do not care for ±; our results are independent of whether there are Fermi or Bose statistics). Remember that $\mathsf{H}(V \cup V') = \mathsf{H}(V) \otimes \mathsf{H}(V')$ if $V \cap V' = \phi$ (in the sense of Sec. I.A).

A state on $\mathsf{H}(V)$ is described by a density matrix ρ_V. It is plausible to require that all these density matrices are *consistent* in the sense that $\rho_V = \mathrm{Tr}_{\mathsf{H}_{V'}} \rho_{V \cup V'}$ if $V \cap V' = \phi$. Note: ρ_V is not necessarily a Gibbs state.

The entropy of a subvolume is defined by

$$S(V) \equiv S(\rho_V) . \tag{2.14}$$

Subadditivity implies that $S(V \cup V') \leq S(V) + S(V')$ if $V \cap V' = \phi$. The problem we want to study is the following: Does there exist a limit $S(V)/|V|$ ($|V|$ being the volume of V, or the number of lattice points in V, respectively), as $V \to \infty$ in some suitable sense (for instance, in the sense of van Hove; see Sec. I.C), provided that the system is translationally invariant, i.e., $S(V + a) = S(V)$ for all $a \in \mathrm{R}^d$, or Z^d, respectively.

To begin with let us consider the case of a one-dimensional lattice system. Let V be an interval of length l: $V = \{k, k+1, \ldots, k+l-1\}$. Because of translational invariance, $S(V)$ is a function of l only: $S(V) = F(l)$, $l \equiv |V|$. By subadditivity of the entropy, F is also a subadditive function of l:

$$F(l_1 + l_2) \leq F(l_1) + F(l_2) . \tag{2.15}$$

Using a classical theorem of analysis (e.g., Polya and Szegö, 1970) one concludes that the limit $F(l)/l$ exists:

$$\lim_{l \to \infty} \frac{F(l)}{l} = \lim \frac{S(V)}{|V|} = s, \quad \text{with } s \equiv \inf \frac{F(l)}{l} . \tag{2.16}$$

The same argument would work in the continuous case (Z being replaced by R) too, provided that one would have some bound on $F(l)$. However, subadditivity is not sufficient to provide such a bound.

As an example, consider for V intervals $[a, b]$ and define "S"$([a, b]) = 0$ if $b - a$ is rational, $= \infty$ if $b - a$ is irrational.

We will see in Sec. III.A that such a bound is in fact, for quantum systems, provided by strong subadditivity, which is a sharpening of the subadditivity property discussed in this section. Namely, strong subadditivity yields, for $l' < l_0$, the inequality

$$F(l') + F(2l_0 - l') \leq 2F(l_0) , \tag{2.17}$$

hence in the quantum case, $F(l') \leq 2F(l_0)$, because the quantum-mechanical entropy is always ≥ 0.

But inequality (2.17), which also holds in the classical case (see Sec. III.A), allows us to prove the existence of the mean entropy even for classical continuous systems, where it is not true that $F(l') \leq 2F(l_0)$. We have, with $s=\inf F(l)/l$,

$$F(l')\leqslant 2F(l_0)-(2l_0-l')s\,.$$

Choose l_0 such that $F(l_0)\leqslant l_0(s+\epsilon)$. Any l can be written as $l=nl_0+l'$ (n = integer, $l'<l_0$). Therefore $F(l)\leqslant nF(l_0)+F(l')$ (by subadditivity) and

$$s\leqslant\frac{F(l)}{l}\leqslant\frac{(2+n)F(l_0)-(2l_0-l')s}{nl_0+l'}\,,$$

and, since for $l\to\infty$ also $n\to\infty$,

$$\lim_{l\to\infty}\sup\frac{F(l)}{l}\leqslant\frac{F(l_0)}{l_0}\leqslant s+\epsilon\,,$$

thus $\lim F(l)/l=s$.

2. Monotonicity

If S were monotonic there would be no need to refer to strong subadditivity in order to establish the existence of $\lim F(l)/l$. But monotonicity is generally true only in the classical discrete case. It will turn out, however, that it is also true in our case of translationally invariant systems; again, this result relies on strong subadditivity.

Nevertheless there are some instances where some sort of monotonicity can be proved even without any knowledge of strong subadditivity. We will discuss two of them:

(a) *Quantum lattice systems*. (This, of course, also comprises the case of classical lattice systems, cf. Sec. I.A.) Let us drop the assumption $d=1$. Remember that $\mathsf{H}(V)=\bigotimes_{x\in V}\mathsf{H}_x$, H_x being Hilbert spaces of fixed finite dimension, say κ. One readily verifies that

$$S(V)\leqslant|V|\ln\kappa\,.$$

Let $V\subset V'$. By subadditivity, $S(V')\leqslant S(V)+S(V'')$, $V''=V'\backslash V$), hence

$$S(V')-S(V)\leqslant S(V'')\leqslant|V''|\ln\kappa=(|V'|-|V|)\ln\kappa\,.$$

In the classical case, also $S(V)\leqslant S(V')$. In the quantum-mechanical case, let

$$\tilde S(V)=S(V)-|V|\ln\kappa\,.$$

Then

$$\tilde S(V)\leqslant 0\,,$$
$$\tilde S(V')\leqslant\tilde S(V)\quad\text{if } V'\supset V \tag{2.18}$$

and

$$\tilde S(V')\leqslant\tilde S(V)+\tilde S(V'\backslash V)\,.$$

(b) *Configurational entropy of classical statistical mechanics* (Robinson and Ruelle, 1967). In classical statistical mechanics it is often sufficient to consider only the probability distribution in configuration space instead of the complete distribution in phase space. Taking into account the possibility of a variable number of particles moving in the (bounded) volume V, one thus arrives at a family of symmetric distributions

$$\rho_V^{(N)}(q_1,\ldots,q_N),\quad N=0,1,\ldots \qquad (\rho_V^{(0)}\equiv 1)\,. \tag{2.19}$$

Since in a classical theory there is no need of introducing the "correct" normalization condition (1.3) one chooses it as follows:

$$\sum_{N=0}^{\infty}\frac{e^{-|V|}}{N!}\int_{V^N}dq_1\cdots dq_N\rho_V^{(N)}=1 \tag{2.20}$$

and defines as the configurational entropy

$$S_{\text{conf}}(V)=-\sum_{N=0}^{\infty}\frac{e^{-|V|}}{N!}\int_{V^N}dq_1\cdots dq_N\rho_V^{(N)}\ln\rho_V^{(N)}\,. \tag{2.21}$$

Note that this kind of entropy is defined in some sort of "classical Fock space." It is very different from the grand-canonical entropy of Sec. I for two reasons: (i) the normalization condition, and (ii) the kinetic energy is omitted.

One can show, extending our previous arguments, that the analog of (2.18) holds:

$$S_{\text{conf}}(V)\leqslant 0\,,$$
$$S_{\text{conf}}(V')\leqslant S_{\text{conf}}(V)\quad\text{if } V'\supset V\,, \tag{2.22}$$

and also that subadditivity (and even strong subadditivity) holds.

Namely,

$$S_{\text{conf}}(V)=\sum_{N=0}^{\infty}\frac{e^{-|V|}}{N!}\int_{V^N}dq_1\cdots dq_N\,\rho_V^{(N)}(\ln 1-\ln\rho_V^{(N)})$$
$$\leqslant\sum_{N=0}^{\infty}\frac{e^{-|V|}}{N!}\int_{V^N}dq_1\cdots dq_N(1-\rho_V^{(N)})\ \text{(by Klein's inequality; cf. Secs. I.A and I.B)}=1-1=0\,.$$

The second inequality follows from the first one and subadditivity: For $V\subset V'$,

$$S_{\text{conf}}(V')\leqslant S_{\text{conf}}(V)+S_{\text{conf}}(V'\setminus V)\leqslant S_{\text{conf}}(V)\,.$$

Turning to subadditivity, we first have to define $S_{\text{conf}}(V)$ and $\rho_V^{(N)}$ in terms of $\rho_{V'}^{(N)}$:

$$\rho_V^{(N)}(x_1,\ldots,x_N)\equiv\sum_{M=0}^{\infty}\frac{e^{-|V''|}}{M!}\int_{V''^M}dy_1\cdots dy_M\,\rho_{V'}^{(N+M)}(x_1,\ldots,x_N,y_1,\ldots,y_M)$$

etc. ($V''=V'\setminus V, x_i\in V$, $y_k\in V''$). Similarly, $S_{\text{conf}}(V'')$ and $\rho_{V''}^{(N)}$ are defined. We have

$$S_{\rm conf}(V')=-\sum_{N=0}^{\infty}\frac{e^{-|V'|}}{N!}\int_{V'^N}dq_1\cdots dq_N\,\rho_{V'}^{(N)}\ln\rho_{V'}^{(N)}$$

$$=-\sum_{N,M=0}^{\infty}\frac{e^{-|V'|}}{N!M!}\int_{V^N}dx_1\cdots dx_N\int_{V''^M}dy_1\cdots dy_M\,\rho_{V'}^{(N+M)}(x_1,\ldots,x_N,y_1,\ldots,y_M)\cdot\ln\rho_{V'}^{(N+M)}(\cdots)\ ,$$

since every point $q\in V'$ must either belong to V or to V'', and

$$\int_{V'^N}dq_1\cdots dq_N=\sum_{K=0}^{N}\binom{N}{K}\int_{V^K}dx_1\cdots dx_K\int_{V''^{N-K}}dy_1\cdots dy_{N-K}\ .$$

Now,

$$S_{\rm conf}(V)=-\sum_N\frac{e^{-|V|}}{N!}\int_{V^N}dx_1\cdots dx_N\,\rho_V^{(N)}(x_1,\ldots,x_N)\ln\rho_V^{(N)}$$

$$=-\sum_{N,M}\frac{e^{-|V|}}{N!}\frac{e^{-|V''|}}{M!}\int_{V^N}dx_1\cdots dx_N\int_{V''^M}dy_1\cdots dy_M$$

$$\times\,\rho_{V'}^{(N+M)}(x_1,\ldots,x_N,y_1,\ldots,y_M)\ln\rho_V^{(N)}(x_1,\ldots,x_N)\ .$$

Note that $|V|+|V''|=|V'|$. A similar formula holds for $S_{\rm conf}(V'')$, and therefore

$$S_{\rm conf}(V)+S_{\rm conf}(V'')-S_{\rm conf}(V')$$

$$=\sum_{N,M}\frac{e^{-|V'|}}{N!M!}\int_{V^N}dx_1\cdots dx_N\int_{V''^M}dy_1\cdots dy_N\,\rho_{V'}^{(N+M)}(x_1,\ldots,x_N,y_1,\ldots,y_M)$$

$$[\ln\rho_{V'}^{(N+M)}(x_1,\ldots,y_M)-\ln\rho_V^{(N)}(x_1,\ldots,x_N)-\ln\rho_{V''}^{(M)}(y_1,\ldots,y_M)]$$

$$\geqslant\sum_{N,M}\frac{e^{-|V'|}}{N!\,M!}\int_{V^N}dx_1\cdots dx_N\int_{V''^M}dy_1\cdots dy_N(\rho_{V'}^{(N+M)}(x_1,\ldots,y_M)-\rho_N^{(N)}(x_1,\ldots,x_N)\rho_{V''}^{(M)}(y_1,\ldots,y_M))$$

(again by Klein's inequality) $\geqslant 1-1=0$.

Let us return to our example and indicate how monotonicity can be used in order to establish the existence of $\lim\tilde{S}(V)/|V|$ or $\lim S_{\rm conf}(V)/|V|$. (Of course, for quantum lattice systems the existence of $\lim\tilde{S}(V)/|V|$ also implies the existence of $\lim S(V)/|V|$.) We shall consider the above case (a) only (for the configurational entropy things work in quite the same manner) since for our following arguments we need only relations (2.18), or (2.22), respectively.

Choose ϵ, l_0, l, and n as before in example 1, replacing S, however, by $\tilde{S}$ in the definition of $F(l)$, thus defining $s=\inf F(l)/l$, etc. Inequality (2.17) is then replaced by

$$0\leqslant\frac{F(l)}{l}\leqslant\frac{F(nl_0)}{l}\leqslant\frac{nl_0}{l}\,\frac{F(l_0)}{l_0}\ . \tag{2.23}$$

Consequently,

$$\lim\frac{F(l)}{l}=\inf\frac{F(l)}{l}$$

exists.

3. Dimensions > 1

In the case of dimension >1, subadditivity is definitely too weak to establish the existence of the mean entropy, even in lattice systems. Also monotonicity is not sufficient.

To illustrate this let us consider $\tilde{S}$ or the configurational entropy. For simplicity we will use the same letter $\tilde{S}$ for both S itself and the configurational entropy. Let $a=(a_1,\ldots,a_d)$ be a vector in Z^d or R^d, and let $V(a)$ be the box $\{x\in Z^d \text{ or } \mathrm{R}^d\colon\ 0\leqslant x_i\leqslant a_i\}$. Let us also define

$$s\equiv\inf\frac{\tilde{S}(V(a))}{|V(a)|}\ .$$

Suppose now that a sequence of volumes V tends to infinity in the sense of van Hove (cf. Sec. I.C). We choose a_0 such that $\tilde{S}(V(a_0))\leqslant|V(a_0)|(\tilde{s}+\epsilon)$. Define n_V^-, or n_V^+, respectively., as in Sec. I.C. By assumption, $n_V^-/n_V^+\rightarrow 1$. Monotonicity and subadditivity imply in the same way as before that

$$\frac{\tilde{S}(V)}{|V|}\leqslant\frac{n_V^-}{n_V^+}(\tilde{s}+\epsilon)\ . \tag{2.24}$$

It remains to show an inequality of the kind

$$\frac{S(V)}{|V|}\geqslant\frac{n_V^+}{n_V^-}\,\tilde{s} \tag{2.25}$$

which would be a consequence of the inequality

$$S(\Gamma_V^+)\geqslant n_V^-\,\tilde{s}\,|V(a_0)|\ , \tag{2.26}$$

where Γ_V^+ denotes the union of the n_V^+ translates of $V(a_0)$ that cover V. However, the latter inequality can only be obtained by invoking strong subadditivity (cf. Sec. III.A).

4. The Kolmogorov-Sinai invariant

The construction of the Kolmogorov-Sinai invariant of classical ergodic theory also makes use of subadditivity (see Sec. IV.A).

Let us once more come back to the mean entropy. One easily can show that the mean entropy, should it exist, is affine on the set of translationally invariant states. Again let us consider case (a) only. Since, by Eq. (2.3) for $\rho_V = \lambda\rho_{1,V} + (1-\lambda)\rho_{2,V}$,

$$\begin{aligned}\lambda S(\rho_{1,V}) + (1-\lambda) S(\rho_{2,V}) \leq S(\rho_V) &\leq \lambda S(\rho_{1,V}) + (1-\lambda) S(\rho_{2,V}) \\ &\quad - \lambda \ln\lambda - (1-\lambda)\ln(1-\lambda) \\ &\leq \lambda S(\rho_{1,V}) + (1-\lambda) S(\rho_{2,V}) + \ln 2 \ ,\end{aligned}$$

in the limit $|V| \to \infty$

$$\lim \frac{S(\rho_V)}{|V|} = \lambda \lim \frac{S(\rho_{1,V})}{|V|} + (1-\lambda) \lim \frac{S(\rho_{2,V})}{|V|} \ . \qquad (2.27)$$

G. Entropy and information theory

In principle we should now treat strong subadditivity. However, since it is

(1) closely related to the concepts of relative entropy and skew entropy, and

(2) requires quite a lot of nontrivial mathematical preparations, we prefer to devote an extra part of this review to these problems and to close this second part with a rough account of the connection between entropy and information.

The principle that entropy is a measure of our ignorance about a given physical system was recognized very early (see, for example, Weaver, Appendix to Shannon and Weaver, 1949; v. Smoluchowski, 1914); Boltzmann was also aware of it.

On the other hand, the *mathematical* theory of information (Shannon and Weaver, 1943) originally was intended as a theory of communication. The simplest problem it deals with is the following: take any message (for instance consisting of words or of digits). One can represent it as a sequence of binary digits and thus, if the length of the "word" is n, one needs n digits to characterize it. The set E_n of all words of length n contains 2^n elements, therefore the amount of information needed to characterize one element of its is $\log_2$ of (the number of elements of E_n) $= \log_2 N$, with $N = 2^n$. Elaborating on this a little bit, one arrives at the result that the amount of information which is needed to characterize an element of any set of power N (not necessarily of the form $N = 2^n$) is $\log_2 N$. Now let E be a union $E_1 \cup \cdots \cup E_K$ of pairwise disjoint sets, N_i = number of elements of E_i. Let $p_i = N_i/N$, $N = \sum N_i$. If one knows that an element of E belongs to some E_i, one needs $\log_2 N_i$ additional information in order to determine it completely. Hence the average amount of information needed to determine an element, provided that one already knows to which E_i it belongs, is $\sum (N_i/N) \log_2 N_i = \sum p_i \log_2 N p_i = \sum p_i \log_2 p_i + \log_2 N$. Now we just have seen that $\log_2 N$ is the information that is needed if one does not know to which E_i a given element belongs, hence the corresponding average lack of information is

$$-\sum p_i \log_2 p_i \ . \qquad (2.28)$$

This is usually called Shannon's formula, although it was discovered by Wiener independently.

From here there is a short way to physics: if the set E is intereted as a set of N measurements, and the p_i are the probabilities of finding the system in the pure state $|i\rangle$, then, expect for an irrelevant factor ln2, Shannon's expression equals the definition of entropy.

This enables one to apply quite a few mathematical results from information theory to entropy, and we already have done this on several occasions. Examples are, for instance, the characterization theorems of Secs. B and F.

However, one has to bear in mind that information theory does not contain any quantum mechanics, so that it can be applied directly to the classical discrete case only, i.e., if there is no noncommutativity involved at all. If one wants to apply it to the general quantum-mechanical case, one usually has to worry about problems arising from noncommutativity, so that not every result of information theory has a quantum-mechanical "translation."

As concerns noncommutativity, it seems quite natural to ask for quantities that measure the amount of noncommutativity of two operators rather than the amount of information contained in *one* density matrix. We will do this in Secs. III.B. and C., as we just have said, information theory does not cover that subject.

Turning back to Shannon's formula, it has to be added that one can conceive of many other measures of the amount of information contained in a probability distribution or a density matrix. These measures usually have only little importance, as was already pointed out in the introduction. In the previous section we occasionally were concerned with the quantum analogs of Renyi's α entropies. Other quantities one could think of were, for instance,

$$-\ln f^{-1}(\mathrm{Tr}\, \rho f(\rho)) \ ,$$

f being an increasing convex or concave function, or

$$f^{-1}[\mathrm{Tr}\, \rho f(-\ln \rho)]$$

(Renyi, 1965; Aczel and Daroczy, 1963) or

$$\frac{1}{1-\alpha} (\mathrm{Tr}\, \rho^{\alpha} - 1) \ ,$$

etc. (Daroczy, 1970).

What one can learn from considering these "entropies" is that mixing-enhancement means loss of information in the worst possible way because not only does entropy increase but also all the other measures of lack of information increase.

By means of information theory it is possible to rephrase the maximum entropy principle in other terms: suppose that for some system you know only a few, macroscopic quantities, and you have no further knowledge of it. Then the system is expected to be in the state with maximal entropy, because if it were in a state with lower entropy it would also contain more information than previously specified (Jaynes' principle, Jaynes, 1957). However, as we have already discussed in Sec. I.B, one has to be careful with such arguments because they only make plausible, but do not actually prove, the maximum entropy principle.

On the other hand it is amusing to note that in practical applications of information theory, such as in technology, biology, etc., the second law of thermodynamics has been adopted and there called the negentropy principle (see Brillouin, 1962). Thus we find a mutual interaction between physics and information theory rather than a perfect understanding of statistical mechanics on the grounds of information theory.

III. STRONG SUBADDITIVITY AND LIEB'S THEOREM

A. Strong subadditivity

In Sec. II.F it turned out that mere subadditivity often is too weak a property, and that strong subadditivity is needed. By this, the following is meant: given three Hilbert spaces $\mathsf{H}_1, \mathsf{H}_2, \mathsf{H}_3$, let ρ be a density matrix in $\mathsf{H}_1 \otimes \mathsf{H}_2 \otimes \mathsf{H}_3$. Define the partial traces $\rho_1 = \mathrm{Tr}_{\mathsf{H}_2 \otimes \mathsf{H}_3} \rho$, $\rho_{12} = \mathrm{Tr}_{\mathsf{H}_3} \rho$, etc. (In order to have a less cumbersome notation, from now on instead of ρ we will write ρ_{123}, instead of $\mathrm{Tr}_{\mathsf{H}_2 \otimes \mathsf{H}_3}$ we will write Tr_{23}, instead of $\mathrm{Tr}_{\mathsf{H}_3}$ we will write Tr_3, etc.) Then

$$S(\rho_{123}) + S(\rho_2) \leq S(\rho_{12}) + S(\rho_{23}) \ . \tag{3.1}$$

If H_2 is one-dimensional, this reduces to normal subadditivity. The same inequality holds in the classical case, there being given three "phase spaces" $\Omega_1, \Omega_2, \Omega_3$; ρ_{123} is a probability distribution in $\Omega_1 \times \Omega_2 \times \Omega_3$, $\rho_1(w_1) = \int dw_2 dw_3\, \rho_{123}(w_1, w_2, w_3)$, $\rho_{12}(w_1, w_2) = \int dw_3\, \rho_{123}(w_1, w_2, w_3)$, etc.

In the classical case, the proof of inequality (3.1) is very simple: one only has to use the inequaltiy

$$\int dw_1 dw_2 dw_3 \rho_{123}(\ln\rho_{123} - \ln\sigma) \geq 0 \ ,$$

valid for every probability distribution σ, and to take $\sigma = \rho_{12}\rho_{23}/\rho_2$. Then

$$\int dw_1 dw_2 dw_3 \rho_{123}(\ln\rho_{123} + \ln\rho_2 - \ln\rho_{12} - \ln\rho_{23}) \geq 0 \ ,$$

which is just the assertion.

On the other hand, in quantum mechanics the proof is extremely difficult. Therefore, before turning to it, let us consider how strong subadditivity can be used in physical problems.

Let us first put inequality (3.1) into another form: consider two volumes V, V' (not necessarily disjoint) and the associated Fock spaces $\mathsf{H}(V), \mathsf{H}(V'), \mathsf{H}(V \cap V'), \mathsf{H}(V \cup V')$ (or the Hilbert spaces for lattice systems as indicated in Sec. I.A). Then, with the notation of Sec. II.F,

$$S(V \cap V') + S(V \cup V') \leq S(V) + S(V') \ . \tag{3.2}$$

Now we are in a position to state some applications.

(1) In Sec. II.F we were concerned with the problem of the existence of $\lim S(V)/|V|$ for one-dimensional, translationally invariant, continuous systems and saw that some bound on the function $F(l)$ would suffice for this purpose. Strong subadditivity provides this bound in the quantum-mechanical case: let $l' < l$. Then,

$$F(l') \leq 2F(l) \ . \tag{3.3}$$

This is true because any interval of length l' can be represented as the intersection of two intervals of length l; thus by strong subadditivity

$$F(l') \leq F(l') + F(2l - l') \leq F(l) + F(l) \ ,$$

On the other hand, as we have seen in Sec. II.F, the second inequality of the latter relation is also sufficient to prove the existence of the mean entropy in the classical continuous case (see Sec. II.F).

The following remark, due to E. Lieb, applies: let $x = 2l - l'$, $y = l'$ in the last formula. Then $2F(x+y)/2 \geq F(x) + F(y)$, i.e., F is weakly concave. To show that F is concave, i.e., $F(\lambda x + (1-\lambda)y) \geq \lambda F(x) + (1-\lambda)F(y)$, it is sufficient to have F bounded above in *any* interval. Conversely, if F is concave, this implies strong subadditivity.

(2) This problem is closely related to the problem of *monotonicity* of the quantum-mechanical entropy, i.e., of proving that $F(l') \leq F(l)$ (cf., our remarks of Sec. II.F). If there is no translational invariance, we already have seen that this need not be true. However, if the system *is* translationally invariant, then one can use strong subadditivity to show that

$$F(l) - F(l') \geq F(l+m) - F(l'+m)$$

for every $m \geq 0$, in particular for $m = n(l - l')$, n being an integer. Consequently,

$$F(l) - F(l') \geq \frac{1}{N} \sum_{n=1}^{N} [F(l + n(l-l')) - F(l' + n(l-l'))]$$

$$= \frac{1}{N} [F(l + N(l-l')) - F(l)] \ .$$

If $N \to \infty$, the right-hand side is $\geq \inf F(l'')/l''$, which we already know to be $\lim F(l'')/l''$, and which, in quantum mechanics, is ≥ 0.

3. Now let us consider translationally invariant systems in dimensions >1. If we are given a lattice system and consider a sequence of boxes whose lengths tend to infinity, then again, as in example 1 of Sec. II.F, $\lim S(V)/|V|$ exists.

V being a parallelepiped $\{x\colon\ 0 \leq x_i \leq a_i\}$, $S(V)$ is a function F of $a_1, \ldots, a_d$ and $|V|$ is $a_1 \cdots a_d$. By subadditivity alone, F is a subadditive function of every variable a_i separately. A straightforward modification of the theorem used before shows that

$$\lim_{a_1, \ldots, a_d \to \infty} \frac{F(a_1, \ldots, a_d)}{a_1 \cdots a_d} = \inf \frac{F(a_1, \ldots, a_d)}{a_1 \cdots a_d} \ .$$

On the other hand, if we are given a continuous system, again we have to make use of strong subadditivity. The following argument is due to Araki and Lieb (1970). Choose the box $V(a_0)$ with edges $a_1^{(0)}, \ldots, a_d^{(0)}$ so that

$$\frac{S[V(a_0)]}{|V(a_0)|} \leq (s + \epsilon) \ ,$$

$$s \equiv \inf \frac{S(V(a))}{|V(a)|} = \inf \frac{F(a_1, \ldots, a_d)}{a_1 \cdots a_d} \ .$$

Now, as in Sec. II.F, for large boxes $V(a)$ with lengths $a_1, \ldots, a_n$ there are integers $n_1, \ldots, n_d$ such that $a_i = n_i a_i^{(0)} + b_i$, $0 \leq b_i \leq a_i^{(0)}$. Then

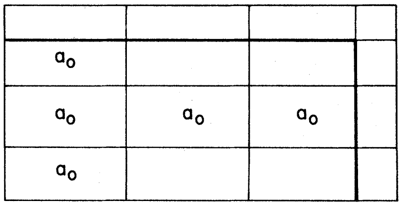

FIG. 8. Construction of the mean entropy for boxes.

$$s \leq \frac{S[V(a)]}{|V(a)|}$$

and $S[V(a)] \leq S[V(na_0)]+$contributions of smaller boxes: see Fig. 8. [$V(Na_0)$ is the box with lengths $n_1a_1^{(0)}, \ldots, n_da_d^{(0)}$]

$$\frac{S[V(na_0)]}{|V(na_0)|} \leq (s+\epsilon) \text{ by subadditivity.}$$

As concerns the smaller boxes, in quantum mechanics, due to strong subadditivity, the entropy of any one of them is $\leq 2S[V(a_0)]$, because they can be represented as the intersection of two translates of $V(a_0)$. The number of these smaller boxes being of the order of the surface of $V(a)$ only, one concludes as in Sec. II.F that

$$\lim \frac{S[V(a)]}{|V(a)|} = s\,.$$

(4) One might think it easily possible to generalize this proof to volumes that are not boxes but arbitrary, provided that they tend to infinity in the sense of von Hove. Unfortunately strong subadditivity does not give a bound for the entropy of the "surface terms" of a covering of V by translates of $V(a_0)$, i.e., those translates of $V(a_0)$ that have a nonempty intersection with V but are not entirely contained in V. Namely, in general these volumes cannot be represented as an intersection of two translates of a and therefore strong subadditivity does not help. (As E. Lieb has remarked, the previous argument *does* work for states that are also rotationally invariant, if, in three dimensions, the surface is composed of finitely many flat polygonal pieces. In this case the shaded part of the volume, according to Fig. 9, can be decomposed into tetrahedrons, and every tetrahedron can be represented as an intersection of four boxes. Of course, a similar statement holds in arbitrary dimensions $d>1$.)

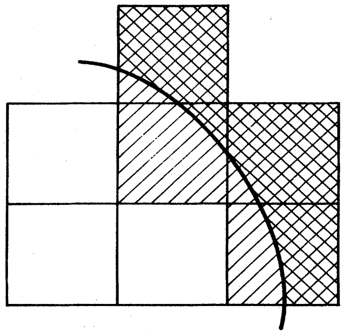

FIG. 9. "Surface terms" of a covering of V by translates of $V(a_0)$.

One rather needs some monotonicity, which, for instance, is the case for the *configurational entropy,* as we have discussed already in example 2 of Sec. II.F. This provides a bound for $S(V)$ in one direction, but remember that we were left with the necessity of proving a bound in the other direction, namely that

$$S(\Gamma_V^+) \geq n_V^+ s|V(a_0)|\,.$$

[One checks easily that the configurational entropy is strongly subadditive too, in fact, even the generalized Boltzmann–Gibbs–Shannon entropy is strongly subadditive (Ochs, 1976).] This inequality, which is due to Robinson and Ruelle, is obtained by a more elaborate version of the method we used in example 2, and is based on a combinatorial argument that is a little bit tricky. Note that in the case of boxes there was no need of proving such an estimate because *by definition,* always $S[V(a)]/|V(a)| \geq s$; however, since Γ_V^+ need not itself be a box, *a priori* it is not impossible that $S(\Gamma_V^+) < s|\Gamma_V^+|$, and this is the point where strong subadditivity comes in to exclude this possibility.

Now, after having indicated some applications, which, as I hope, show why strong subadditivity is an important property and what it is good for, I should like to make a few remarks about history.

Strong subadditivity was known for many years in information theory, but not generally called that. For statistical mechanics, at least, it was Robinson and Ruelle (1967) who coined the word and who first realized that is was important. They proved it in the classical case. Then Lanford and Robinson (1968) conjectured it in quantum mechanics.

For five years this conjecture was an open problem. There were several attempts either to prove or to disprove it but only two partial results:

(a) a proof by Baumann and Jost (1969) for 2×2 matrices and (b) a weak version of strong subadditivity by Araki and Lieb (1970), which was powerful enough to establish the existence of the mean entropy for translationally invariant states of continuous quantum systems.

Finally, the "Lanford–Robinson conjecture" was proven by Lieb and Ruskai (1973), using the results of Lieb concerning the so-called "Wigner–Yanase–Dyson conjecture" (see Sec. C).

Let us turn to the proof of strong subadditivity for quantum-mechanical entropy. The crucial quantity to be considered is the conditional entropy (Lieb, 1975)

$$S(2|1) \equiv S(\rho_{12}) - S(\rho_1)\,.$$

For simplicity we will from now on write S_{12} for $S(\rho_{12})$, S_1 for $S(\rho_1)$, etc.

We will later prove that the conditional entropy is concave in ρ_{12} (Lieb and Ruskai, 1973). This is true both quantum-mechanically and classically, but we will consider the quantum case only. Lieb (1975) uses the expression "relative entropy." For finite-dimensional Hilbert spaces, it differs from our relative entropy, see next section, by a term $\ln \dim \mathrm{H}_2$.

The concavity of $S_{12}-S_1$ implies the following inequality: for a density matrix ρ_{123} in $\mathrm{H}_1 \otimes \mathrm{H}_2 \otimes \mathrm{H}_3$,

$$S_1+S_2\leqslant S_{13}+S_{23} \tag{3.5}$$

[$S_1=-\mathrm{Tr}\rho_1\ln\rho_1$ (Lieb and Ruskai, 1973), etc.] This statement, somewhat similar to monotonicity, is obtained by considering

$$\Delta\equiv(S_{13}-S_1)+(S_{23}-S_2)\,.$$

The mapping $\rho_{123}\to\rho_{13}$ being linear, $S_{13}-S_1$ is concave in ρ_{123}, and, similarly, the same is true for $S_{23}-S_2$. Hence Δ is concave. For pure states, $\Delta=0$ since $S_{13}=S_2, S_{23}=S_1$ (remark 1 of Sec. II.F). By concavity, for mixed states Δ must be $\geqslant 0$. [It should be remarked that inequality (3.5) is false in the classical continuous case.]

Let us now proceed by choosing a fourth Hilbert space H_4 such that, according to remark 2 of Sec. II.F, there is a pure ρ_{1234} in $(\mathsf{H}_1\otimes\mathsf{H}_2\otimes\mathsf{H}_3)\otimes\mathsf{H}_4$ such that $\rho_{123}=\mathrm{Tr}_4\,\rho_{1234}$. Then

$$S_{123}+S_2-S_{12}-S_{23}=S_4+S_2-S_{12}-S_{14}\leqslant 0$$

by (3.5), which establishes strong subadditivity.

One might think that there are other inequalities of the type of the above ones, for instance between $S_{123}+S_1+S_2+S_3$ and $S_{12}+S_{13}+S_{23}$, but this is not the case. Also $S_{12}-S_1-S_2$ is neither concave nor convex. For a further discussion of which inequalities are true and which are not, see Lieb (1975).

The above is the original proof of Lieb and Ruskai of strong subadditivity. There is another way, due to Uhlmann, of proving strong subadditivity from the concavity of $S_{12}-S_1$. Let all Hilbert spaces under consideration be finite-dimensional. Now, as above, $S_{123}-S_{23}$ is concave. Denote by dU_3 the Haar measure of the group of unitary operators in H_3. Then

$$\int dU_3U_3\rho_{123}U_3^{-1}=\frac{1}{d_3}\,\rho_{12}\otimes 1$$

($d_3=$dimension of H_3; U_3 is identified with $1\otimes U_3$). Thus

$$\int dU_3(S_{123}-S_{23})(U_3\rho_{123}U_3^{-1})\leqslant(S_{123}-S_{23})\left(\frac{1}{d_3}\,\rho_{12}\otimes 1\right),$$

or

$$S_{123}-S_{23}\leqslant(S_{12}-\ln d_3)-(S_2-\ln d_3)\,.$$

So what we have to do is to prove the concavity of $S_{12}-S_1$. We will do this for the finite-dimensional case; the general case follows from an application of our results of Sec. II.D. To make things more transparent, let us also abuse language and write ρ_1 instead of $\rho_1\otimes 1$ in $\mathsf{H}_1\otimes\mathsf{H}_2$.

The essential ingredient of the proof is the following.

Lemma (Lieb, 1973b). For finite-dimensional matrices, the mapping $A\to\mathrm{Tr}\exp(K+\ln A)$ (for $A>0, K$ self-adjoint) is concave.

Now let $\rho_{12}=\lambda\rho_{12}'+(1-\lambda)\rho_{12}''(0\leqslant\lambda\leqslant 1)$. Define

$$\Delta=\mathrm{Tr}\rho_{12}(\ln\rho_{12}-\ln\rho_1)-\lambda\,\mathrm{Tr}\rho_{12}'(\ln\rho_{12}'-\ln\rho_1')$$
$$-(1-\lambda)\mathrm{Tr}\rho_{12}''(\ln\rho_{12}''-\ln\rho_1')\,,$$
$$\Delta'=\mathrm{Tr}\rho_{12}'(\ln\rho_{12}-\ln\rho_1-\ln\rho_{12}'+\ln\rho_1')\,,$$

and Δ'' similarly. $\Delta=\lambda\Delta'+(1-\lambda)\Delta''$. We want to show that $\Delta\leqslant 0$, or $e^{\Delta}\leqslant 1$. Because of the convexity of the exponential function,

$$e^{\Delta}\leqslant\lambda e^{\Delta'}+(1-\lambda)e^{\Delta''}\,.$$

On the other hand, the Peierls–Bogoliubov inequality (Sec. I.B)

$$\mathrm{Tr}e^{A+B}\leqslant\mathrm{Tr}e^{A+\langle B\rangle}\,,$$
$$\langle B\rangle\equiv\mathrm{Tr}Be^{A}/\mathrm{Tr}e^{A}\,,$$

with $A=\ln\rho_{12}'$, $B=(\ln\rho_{12}-\ln\rho_1-\ln\rho_{12}'+\ln\rho_1')$ yields $e^{\Delta'}\leqslant\mathrm{Tr}e^{K+\ln\rho_1'}$, with $K\equiv\ln\rho_{12}-\ln\rho_1$, and in the same way $e^{\Delta''}\leqslant\mathrm{Tr}e^{K+\ln\rho_1''}$. By Lieb's lemma,

$$\lambda\,\mathrm{Tr}e^{K+\ln\rho_1'}+(1-\lambda)\,\mathrm{Tr}e^{K+\ln\rho_1''}\leqslant\mathrm{Tr}e^{K+(\lambda\ln\rho_1'+(1-\lambda)\rho_1'')}$$
$$\leqslant\mathrm{Tr}e^{K+\ln\rho_1}=\mathrm{Tr}e^{\ln\rho_{12}}=1\,.$$

So we have got a proof of the concavity of the conditional entropy by assuming the validity of the lemma. Unfortunately, the proof of the latter is not easy at all (see Sec. C).

B. Relative entropy

We have met the concept of relative entropy, which in general form is due to Umegaki (1962) and Lindblad (1973), on several occasions already, the first being in Secs. I.A (as a special case of the generalized Boltzmann–Gibbs–Shannon entropy) and I.B. [in our discussion of the free energy $F(\rho,\beta,H)$].

Remember that it was defined as $S(\sigma|\rho)\equiv\mathrm{Tr}\rho(\ln\rho-\ln\sigma)$. We have proven that $S(\sigma|\rho)\geqslant 0$ for all density matrices σ,ρ; by the way, going through our proof of Klein's inequality, one sees that $S(\sigma|\rho)=0$ if and only if $\sigma=\rho$;

The second important property is joint convexity for density matrices $\rho_1,\rho_2,\sigma_1,\sigma_2$ and λ: $0\leqslant\lambda\leqslant 1$,

$$S(\sigma|\rho)\leqslant\lambda S(\sigma_1|\rho_1)+(1-\lambda)S(\sigma_2|\rho_2)\,, \tag{3.6}$$

where $\sigma\equiv\lambda\sigma_1+(1-\lambda)\sigma_2, \rho\equiv\lambda\rho_1+(1-\lambda)\rho_2$.

Joint convexity arises from Lieb's concavity theorem, which we will discuss in the next section. The latter states that $\mathrm{Tr}KA^tK^*B^{1-t}$, for $0\leqslant t\leqslant 1$ and $A\geqslant 0$, $B\geqslant 0$, and any K is jointly concave in A and B. Hence setting $K=1$, taking the derivative for $t=0$, one finds that

$$\frac{d}{dt}\,\mathrm{Tr}(A^tB^{1-t})|_{t=0}=\mathrm{Tr}B(\ln A-\ln B) \tag{3.7}$$

is concave, or $S(\sigma|\rho)$ is convex.

As a consequence, the conditional entropy $S_{12}-S_1$ is concave: suppose all Hilbert spaces to be finite-dimensional. Then

$$S_{12}-S_1=-S\!\left(\rho_{12}|\rho_1\,\otimes\,\frac{1}{d_2}\right)+\ln d_2$$

($d_2=$dimension of H_2), observing that $\ln(\rho_1\otimes 1/d_2)=\ln\rho_1\otimes 1-1\otimes(\ln d_2)$. (The transition to the infinite-dimensional case follows the methods indicated in Sec. II.D.)

There is a representation of $S(\sigma|\rho)$ in which the argument of the trace does not contain a product of two noncommuting operators:

$$S(\sigma|\rho)=\sup_{\lambda}S_{\lambda}(\sigma|\rho)\,,$$

$$S_{\lambda}(\sigma|\rho)\equiv(1/\lambda)[S(\lambda\sigma+(1-\lambda)\rho)-\lambda\,S(\sigma)-(1-\lambda)S(\rho)] \tag{3.8}$$

$(0<\lambda<1)$. First note that $d/d\lambda(\lambda S_\lambda)(\lambda=0)=S(\sigma|\rho)$. Second, $\lambda\rightarrow\lambda S_\lambda(\sigma|\rho)$ is concave by Eq. (2.1); hence the differences $(\lambda S_\lambda-0.S_0)/\lambda$, where clearly $0.S_0$ is understood to be $=0$, are decreasing (Lindblad, 1974).

Our aim is to show the lower semicontinuity of $S(\sigma|\rho)$. Let us, as a preparation, give another proof of lower semicontinuity of the usual entropy:

The function $s(x)=-x\ln x$ being continuous on the compact interval $[0,1]$, one finds easily that $\mathrm{Tr}|\rho_n-\rho|\rightarrow 0$ implies that $\|s(\rho_n)-s(\rho)\|\rightarrow 0$ since $\|A\|=\sup\langle\phi|A|\phi\rangle/\langle\phi|\phi\rangle$. Thus, for every finite-dimensional projection P, $\mathrm{Tr}P[s(\rho_n)-s(\rho)]\rightarrow 0$ because of the standard inequality $\mathrm{Tr}PA<\mathrm{Tr}P\|A\|$ (see, for instance, Dixmier, 1957). On the other hand,

$$\mathrm{Tr}PA\leqslant\|P\|\,\mathrm{Tr}A\,,$$

and

$$\mathrm{Tr}A=\sup_P\mathrm{Tr}PA\,.$$

Therefore

$$S(\rho)=\sup_P\mathrm{Tr}Ps(\rho)\leqslant\liminf_n\,[\sup_P\mathrm{Tr}Ps(\rho_n)]$$
$$=\liminf_n S(\rho_n)\,.$$

Now,

$$S(\sigma|\rho)=\sup_{P,\lambda}\mathrm{Tr}P[s(\lambda\sigma+(1-\lambda)\rho)-\lambda s(\sigma)-(1-\lambda)s(\rho)]$$
$$\leqslant\liminf_n\{\sup_{P,\lambda}\mathrm{Tr}P[s(\lambda\sigma_n+(1-\lambda)\rho_n)+\lambda s(\sigma_n)$$
$$-(1-\lambda)s(\rho_n)]\}$$
$$=\liminf_n S(\sigma_n|\rho_n);$$

i.e., the relative entropy is, like the usual entropy, lower semicontinuous. We have already used this fact in Sec. II.D.

There are other theorems similar to those at the end of Sec. II.D. For instance, define, for general $A,B>0$, not necessarily being density matrices,

$$S(A|B)\equiv\mathrm{Tr}[A\,(\ln A-\ln B)+(B-A)]\,. \tag{3.9}$$

Then, if $P_n\uparrow 1$ (P_n = finite-dimensional projections), $S(A_n|B_n)\rightarrow S(A|B)$.

We may use convexity to prove monotonicity. Equation (3.6) generalizes to $S(\sigma|\rho)\leqslant\sum\lambda_i S(\sigma_i|\rho_i)$ (λ_i, etc, being defined as usual), in finite dimensions. Also, noting that

$$S(U^*\sigma U|U^*\rho U)=S(\sigma|\rho) \tag{3.10}$$

for unitary operators U, we have for density matrices σ,ρ in a tensor product $\mathrm{H}_1\otimes\mathrm{H}_2$ of finite-dimensional Hilbert spaces, using a representation similar to that used in the last section,

$$\sigma_1\otimes\frac{1}{d_2}=\int dU_2(1\otimes U_2^*)\sigma(1\otimes U_2)\,,$$
$$\rho_1\otimes\frac{1}{d_2}=\int dU_2(1\otimes U_2^*)\rho(1\otimes U_2)\,, \tag{3.11}$$

and

$$S(\sigma_1|\rho_1)=S(\sigma_1\otimes 1|\rho_1\otimes 1)$$
$$=S\left(\int\cdots\sigma\cdots\Big|\int\cdots\rho\cdots\right)$$
$$\leqslant\int dU_2 S[(1\otimes U_2^*)\sigma(1\otimes U_2)|\,(1\otimes U_2^*)\rho(1\otimes U_2)] \tag{3.12}$$

(cf. Sec. II.F. In particular, $\tilde{S}=s-|V|\ln\kappa$ for lattice systems is just $-S(\sigma|\rho)$, with $\sigma=1/\kappa^{|V|}$.)

Equation (3.12) is a special case of a theorem of Lindblad (1975), which states that $S(\Phi\sigma|\Phi\rho)\leqslant S(\sigma|\rho)$ for every completely positive, trace-preserving mapping Φ, which maps $\mathrm{B}(\mathrm{H})$, the bounded operators in the Hilbert space H, into the $\mathrm{B}(\mathrm{H}_1)$, the bounded operators in another space H_1.

The notion of complete positivity was introduced by Stinespring (1955). It means the following: let $\tilde{A}$ be an $n\times n$ matrix with entries $A_{ik}\in\mathrm{B}(\mathrm{H})$, and let $\Phi_n\tilde{A}$ be the $n\times n$ matrix with entries $\Phi(A_{ik})$. If Φ_n is positive for all n, then Φ is called completely positive. Important special cases are: (a) the partial traces, and (b) doubly stochastic mappings of finite-dimensional spaces (cf. Sec. I.B). In the latter case, $S(\rho)=-S(\sigma|\rho)=\ln r$, with $\sigma=1/r$ (r = dimension of the space). Thus we recover our result of Sec. I.B: $S(M\rho)=-S(\sigma|M\rho)+\ln r$ $=-S(M\sigma|M\rho)+\ln r\geqslant-S(\sigma|\rho)+\ln r=S(\rho)$. For a discussion of the physical meaning of complete positivity see, for example, Lindblad (1977) or Kraus (1970).

In the classical case Lieb's theorem is not needed for a proof but one can argue directly:

$$S(\sigma|\rho)-S(\sigma_1|\rho_1)=\int dw_1dw_2\rho(w_1,w_2)\ln\frac{\rho}{\sigma}$$
$$-\int dw_1dw_2\rho\ln\frac{\rho_1(w_1)}{\sigma_1(w_1)}$$
$$=\int dw_1dw_2\rho\ln\frac{\rho/\sigma}{\rho_1/\sigma_1}$$
$$\geqslant\int dw_1dw_2\rho\left[1-\frac{\rho_1}{\sigma_1}\,\frac{\sigma}{\rho}\right]=1-1+0$$

due to the inequality $\ln x\geqslant 1-1/x$.

Taking three Hilbert spaces, application of Eq. (13.7) yields $S(\rho_{123}|\rho_1\otimes\rho_{23})=S(\rho_1)+S(\rho_{23})-S(\rho_{123})\geqslant S(\rho_{12}|\rho_1\otimes\rho_2)$ $=S(\rho_1)+S(\rho_2)-S(\rho_{12})$, hence $S(\rho_{123})+S(\rho_2)\leqslant S(\rho_{12})+S(\rho_{23})$, i.e., strong subadditivity. Hence the latter is a special case of the monotonicity of the relative entropy, which, in turn, follows from the convexity.

As an application which, of course, we could already have mentioned in Sec. II.F, let us consider a variational principle for lattice systems. Let ρ_V be an arbitrary family of consistent density matrices as in Sec. II.F. Define $S(V)$ as usual, and $E(V)=\mathrm{Tr}\rho_V H_V$, H_V being the Hamiltonian of the volume $|V|$. Also define P_V $=(1/|V|)\ln\mathrm{Tr}e^{-\beta H_V}$. The limits, for $V\rightarrow\infty$ in some appropriate sense (cf. Sec. II.F), or $S(V)/|V|$, $E(V)/|V|$, and P_V are known to exist (Araki, 1975; Ruelle, 1969). The usual techniques establish, by combining the methods of Sec. II.F and our arguments about the Gibbs state which, as one might remember, were just expressing the positivity of $S(\sigma_\beta|\rho)$ (Sec. I.B), the variational principle

$$p \geqslant s - \beta\eta \,, \tag{3.13}$$

with $p \equiv \lim P_V$, $s \equiv \lim S(V)/|V|$, $\eta = \lim E(V)/|V|$. The left side of (3.13) is the "true" pressure (it does not depend on ρ_V). The right side is a function of ρ_V.

It is possible to think, at least in the classical discrete case, of generalizations of relative entropy that are analogous to the generalizations of entropy to $\mathrm{Tr} f(\rho)$, f being a concave function. Namely one could consider expressions like

$$F(\sigma,\rho) = \sum p_i f\left(\frac{q_i}{p_i}\right),$$

f being a concave function, and p_i, q_i being the values of the probability distributions ρ and σ. Based on this observation, it has turned out to be possible to modify Uhlmann's theory and to define a relation $\underset{\rho}{\succ}$ roughly by

$$\sigma_1 \underset{\rho}{\succ} \sigma_2 \leftrightarrow \sum p_i f\left(\frac{q_i^{(1)}}{p_i}\right) \geqslant \sum p_i f\left(\frac{q_i^{(2)}}{p_i}\right).$$

However, these ideas have not been completely worked out yet (Uhlmann, 1977; Ruch and Mead, 1976).

C. Skew entropy and the Wigner-Yanase-Dyson conjecture

In 1963, Wigner and Yanase proposed a measure for the noncommutativity between a density matrix ρ and a fixed observable K, which they called "skew information":

$$I(\rho,K) = -\tfrac{1}{2}\mathrm{Tr}[\rho^{1/2},K]^2 \,. \tag{3.14}$$

The "skew entropy" is its negative:

$$S(\rho,K) = \tfrac{1}{2}\mathrm{Tr}[\rho^{1/2},K]^2 \,. \tag{3.15}$$

They were able to prove a fundamental property that is valid for ordinary entropy, for skew entropy, namely, concavity in ρ. [Of course, it cannot be expected that invariance holds, except for the trivial statement that, for U = unitary, $S(U^*\rho U, U^*KU) = S(\rho,K)$.] (Actually, they did not suppose K to be bounded, so one usually has to worry about whether $[\rho^{1/2},K]$, etc., makes sense, but we will neglect these problems and henceforth hitherto suppose K to be bounded.) Of course, $S(\rho,K) \leqslant 0$ for all ρ, K, and it is exactly $=0$ if, and only if, ρ and K commute.

Later on, Dyson generalized the Wigner–Yanase entropy to

$$S_p(\rho,K) = \tfrac{1}{2}\mathrm{Tr}([\rho^p,K][\rho^{1-p},K]) \tag{3.16}$$

($0<p<1$), (the degenerate case $p=0$ becomes

$$\tfrac{1}{2}\mathrm{Tr}\{[\rho,K][\ln\rho,K]\}$$

and conjectured that it was concave too. Since

$$S_p(\rho,K) = -\mathrm{Tr}\rho K^2 + \mathrm{Tr}\rho^{1-p}K\rho^pK \,, \tag{3.17}$$

which follows from $\mathrm{Tr}(AB\cdots C) = \mathrm{Tr}(CAB\cdots)$, the assertion that S_p is concave in ρ is equivalent to the assertion that $\mathrm{Tr}\rho^{1-p}K\rho^pK$ is concave in ρ. [$-\mathrm{Tr}\rho K^2$ is linear, hence concave.]

No one had any idea that the Wigner–Yanase–Dyson conjecture was related to strong subadditivity until Lieb (1973) realized this connection. In fact, the concavity in ρ of $\mathrm{Tr}\rho^PK^*\rho^{1-p}K$ constitutes the key to the quantum-mechanical strong subadditivity problem. Certainly, without Lieb's work only a few experts would have known of the Wigner–Yanase–Dyson conjecture. The Wigner–Yanase–Dyson conjecture was proven by Lieb (1973). (A proof of the conjecture for 2×2 matrices was given by Baumann, 1971.) Lieb proved even more, namely that $\mathrm{Tr}[\rho^P,K^*][\rho^{1-p},K]$ is also concave for K not necessarily being self-adjoint. This statement follows from the fact that, for $0<p<1$,

$$\mathrm{Tr}\rho^PK^*\rho^{1-p}K$$

is concave in ρ. But Lieb even succeeded in proving that, for $A, B \geqslant 0$, the mapping

$$(\mathrm{Tr}A^pK^*B^qK)^{1/r} \,,$$

where $r \geqslant 1/(p+q)$ is jointly concave in A and B (for $A, B \geqslant 0$) and, for $r \leqslant 2$, convex in K. This is Lieb's concavity theorem. (We already have used the special case $K=1$ to derive the convexity of the relative entropy and the concavity of the conditional entropy from it.) It has to be mentioned that shortly after the appearance of Lieb's proof, Epstein (1973), motivated by Lieb's work, found another proof. Epstein's method is based on the theory of Herglotz functions and is very powerful, because it also provides quite a few other examples of concave maps.

The following elementary proof is a variant, due to Simon (1977), of Uhlmann's proof (1977). The original Lieb proof is shorter, but uses complex interpolation and so is less direct.

Lemma. Let, for $i=1,2, R_i, S_i, T_i \geqslant 0$, $[R_1,R_2]=[S_1,S_2]=[T_1,T_2]=0$. If $R_1 \geqslant S_1+T_1$ and $R_2 \geqslant S_2+T_2$, then $R_1^{1/2}R_2^{1/2} \geqslant S_1^{1/2}S_2^{1/2} + T_1^{1/2}T_2^{1/2} (\equiv Q)$.

Proof: $|\langle\phi|Q|\psi\rangle| \leqslant \|S_1^{1/2}\phi\|\ \|S_2^{1/2}\psi\| + \|T_1^{1/2}\phi\|\ \|T_2^{1/2}\psi\| \leqslant \langle\phi|(S_1+T_1)|\phi\rangle^{1/2}\langle\psi|(S_2+T_2)|\psi\rangle^{1/2}$ because of Schwarz's inequality $s_1s_2+t_1t_2 \leqslant (s_1^2+t_1^2)^{1/2}(s_2^2+t_2^2)^{1/2}$. ($s_1 \equiv \|S_1^{1/2}\phi\|$ etc.) Hence $\|R_1^{-1/2}QR_2^{-1/2}\| \leqslant 1$, by taking $\phi = R_1^{-1/2}\chi$, $\psi = R_2^{-1/2}\chi$. Consequently

$$\begin{aligned}\|R_1^{-1/4}R_2^{-1/4}QR_1^{-1/4}R_2^{-1/4}\| &= \text{largest eigenvalue of } (\cdots)\\ &= \text{(where "spr" denotes}\\ &\quad\ \ \text{the spectral radius)}\\ &= \mathrm{spr}\,(R_1^{-1/2}QR_2^{-1/2})\\ &\leqslant \|R_1^{-1/2}QR_2^{-1/2}\| \leqslant 1.\end{aligned}$$

(We have used that $\mathrm{spr}\,(AB) = \mathrm{spr}(BA)$, $\mathrm{spr}(A) \leqslant \|A\|$, and the commutativity of R_1 and R_2.)

Consider now the space of Hilbert-Schmidt operators, i.e., the operators with $\mathrm{Tr}A^*A<\infty$. It is well known that they become a Hilbert space with the scalar product $(A|B) = \mathrm{Tr}A^*B$. Define R_1, R_2, etc., by

$$(X|R_1Y) = \mathrm{Tr}X^*(\lambda A_1 + (1-\lambda)A_2)Y \,,$$

$$(X|R_2Y) = \mathrm{Tr}X^*Y(\lambda B_1 + (1-\lambda)B_2) \,,$$

$$(X|S_1Y) = \lambda\,\mathrm{Tr}X^*A_1Y \,,$$

etc. Then the lemma tells us that

$$
\begin{aligned}
(X|R_1^{1/2}R_2^{1/2}X) &= \mathrm{Tr}X^*(\lambda A_1+(1-\lambda)A_2)^{1/2}X(\lambda B_1+(1-\lambda)B_2)^{1/2}\\
&\geq \mathrm{Tr}X^*(\lambda A_1)^{1/2}X(\lambda B_1)^{1/2}\\
&\quad + \mathrm{Tr}X^*((1-\lambda)A_2)^{1/2}X((1-\lambda)B_1)^{1/2}\\
&= \lambda\,\mathrm{Tr}X^*A_1^{1/2}X\,B_1^{1/2}+(1-\lambda)\,\mathrm{Tr}X^*A_2^{1/2}X\,B_2^{1/2}\,,
\end{aligned}
$$

i.e., the lemma proves Lieb's theorem for $p=1/2$.

Our method shows even more. Define, more generally,

$$
\begin{aligned}
&R_1:\ (X|R_1Y)=\mathrm{Tr}X^*A^pYB^{1-p}\,,\\
&R_2:\ (X|R_2Y)=\mathrm{Tr}X^*A^qYB^{1-q}\,,\\
&S_1:\ (X|S_1Y)=\lambda\,\mathrm{Tr}X^*A_1^pY\,B_1^{1-p}\,,\\
&S_2:\ (X|S_2Y)=\lambda\,\mathrm{Tr}X^*A_1^qY\,B_1^{1-q}\,,\\
&T_1:\ (X|T_1Y)=(1-\lambda)\,\mathrm{Tr}X^*A_2^pY\,B_2^{1-p}\,,\\
&T_2:\ (X|T_2Y)=(1-\lambda)\,\mathrm{Tr}X^*A_2^qY\,B_2^{1-q}\,,\\
&(A=\lambda A_1+(1-\lambda)A_2,\ B=\lambda\,B_1+(1-\lambda)B_2)\,.
\end{aligned}
$$

The validity of Lieb's concavity theorem at p, or q, states that $R_1\geq S_1+T_2$, $R_2\geq S_2+T_2$, hence by the lemma $R_1^{1/2}R_2^{1/2}\geq S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2}$, i.e., its validity at $(p+q)/2$. The rest follows from a simple induction and continuity argument. (Since $(X|R_1X)\geq 0$, a standard Schwarz inequality-type argument shows convexity in X.)

We have seen in the last sections that Lieb's theorem $\Rightarrow$convexity of the relative entropy and concavity of the conditional entropy$\Rightarrow$strong subadditivity. Remember, however, that our first proof of concavity of the conditional entropy was based on a lemma which looked rather innocent, namely on the concavity of the function $A\to\mathrm{Tr}e^{K+\ln A}$. We did not give a proof there and also will not do this now since it is surprisingly complicated. I want to indicate only that the concavity of $\mathrm{Tr}e^{K+\ln A}$ can be obtained from Lieb's concavity theorem through a sequence of lemmas. (See Lieb's *Advances in Mathematics* paper, 1973b. Epstein's proof of Lieb's theorem also gives, among other results, a *direct* proof of the concavity of $\mathrm{Tr}e^{K+\ln A}$.) So, in conclusion, Lieb's theorem is in fact the essential tool in all the considerations of this section.

Let us come back to skew entropy. In consideration of another most important property of ordinary entropy, Wigner and Yanase proposed the following generalization of subadditivity:

$$S_p(\rho_{12},L)\leq S_p(\rho_1,K_1)+S_p(\rho_2,K_2) \tag{3.18}$$

with $L\equiv K_1\otimes 1+1\otimes K_2$. This assertion can be rewritten as

$$
\begin{aligned}
\mathrm{Tr}\rho_1^pK_1\rho_1^{1-p}K_1+\mathrm{Tr}\rho_2^pK_2\rho_2^{1-p}K_2 &\geq -2\,\mathrm{Tr}\rho_{12}[K_1\otimes K_2]\\
&\quad+\mathrm{Tr}\rho_{12}^p\,L\rho_{12}^{1-p}L\,.
\end{aligned}\tag{3.19}
$$

This is true if $\rho_{12}=\rho_1\otimes\rho_2$, or if K_1, or K_2, respectively, $=0$ (Lieb, 1973). Also is can be proven for $p=1/2$ provided that ρ_{12} is pure. It is an open question whether it is generally true. Anyway, after all, skew entropy is a sort of "relative α entropy" (cf. Sec. II.E), and for α entropies subadditivity does not hold, although they also have the property that $S_\alpha(\rho_1\otimes\rho_2)=S_\alpha(\rho_1)+S_\alpha(\rho_2)$ (in fact, both sides are equal), and for $\rho_{12}=\text{pure}$, $S_\alpha(\rho_{12})\leq S_\alpha(\rho_1)+S_\alpha(\rho_2)$.

IV. RELATED CONCEPTS

A. Dynamical entropies

Let me start this last part with a description of the celebrated Kolmogorov-Sinai invariant of classical dynamical systems (Kolmogorov, 1959; Sinai, 1961, 1965).

In classical mechanics, one is given the phase space Ω and a time evolution, which is a one-parameter group of mappings Φ_t: $\Omega\to\Omega$. By Liouville's theorem, these mappings are measure-preserving; in addition, they are diffeomorphisms.

The Kolmogorov–Sinai (KS) invariant (or KS entropy) is constructed as follows: take a partition $\{\Omega_i\}$ of Ω, i.e., let $\Omega=\cup\Omega_i$, the Ω_i being measurable subsets of Ω, $\Omega_i\cap\Omega_k\neq\phi$ for $i\neq k$. (We do not concern ourselves with sets of measure 0 with respect to the Liouville measure introduced in Sec. I.B.) The entropy of the partition $\omega=\{\Omega_i\}$ is defined as

$$S(\omega)\equiv-\sum W(\Omega_i)\ln W(\Omega_i) \tag{4.1}$$

(remember our notation of Sec. I.B).

Now consider two partitions $\omega_1=\{\Omega_i^{(1)}\}$, $\omega_2=\{\Omega_j^{(2)}\}$. They generate a partition $\omega_1\vee\omega_2$ which consists of all intersections $\Omega_i^{(1)}\cap\Omega_j^{(2)}$. Setting $p_i=W(\Omega_i^{(1)})$, $q_j=W(\Omega_j^{(2)})$, $r_{ij}=W(\Omega_i^{(1)}\cap\Omega_j^{(2)})$, $S(\omega_1\vee\omega_2)=-\sum r_{ij}\ln r_{ij}$, hence, by subadditivity,

$$S(\omega_1\vee\omega_2)\leq S(\omega_1)+S(\omega_2)\,.$$

Let Φ be one of the mappings Φ_t, for fixed t (i.e., some sort of "discrete" time evolution). Since $W(\Phi A)=W(A)$ for any measurable subset $A\subset\Omega$, one has

$$S(\Phi\omega)=S(\omega) \tag{4.3}$$

for every partition (with $\Phi\omega\equiv\{\Phi\Omega_i\}$). Therefore our arguments of Sec. II.F show the existence of

$$\lim_{n\to\infty}\frac{1}{n}\,(S(\omega\vee\Phi\omega\cdots\vee\Phi^{n-1}\omega))\equiv s(\omega,\Phi)\,. \tag{4.4}$$

There is another way of looking at this limit. Define the conditional entropy $S(\omega_1,\omega_2)$ of two partitions as

$$
\begin{aligned}
S(\omega_1,\omega_2)=\sum q_is\left(\frac{r_{ij}}{q_j}\right) &= -\sum r_{ij}\ln r_{ij}+\sum q_j\ln q_j\\
&=S(\omega_1\vee\omega_2)-S(\omega_2)\,,
\end{aligned}\tag{4.5}
$$

i.e., as the classical analog of the conditional entropy we frequently were concerned with in Sec. III. (The quantity

$$\frac{r_{ij}}{q_j}=\frac{W(\Omega_i^{(1)}\cap\Omega_j^{(2)})}{W(\Omega_j^{(2)})}$$

is called "conditional expectation" in probability theory; this explains the word "conditional entropy.") Note that $S(\omega_1,\omega_2)\geq 0$ since entropy is monotonic in the classical discrete case.

Therefore the difference Δ_n equals

$$\begin{aligned}\Delta_n &\equiv S(\omega \vee \cdots \vee \Phi^n \omega) - S(\omega \vee \cdots \vee \Phi^{n-1}\omega)\\ &= S(\Phi^n \omega, \omega \vee \cdots \vee \Phi^{n-1}\omega)\,. \end{aligned} \tag{4.6}$$

Due to strong subadditivity,

$$S(\alpha, \beta) \geq S(\alpha, \beta \vee \gamma) \tag{4.7}$$

for any partitions α, β, γ [because of Eq. (15.5)], hence $\Delta_n \leq \Delta_{n-1}$; consequently, $\lim \Delta_n$ exists, and since

$$S(\omega \vee \cdots \vee \Phi^{n-1}\omega) = S(\omega) + \Delta_1 + \cdots + \Delta_{n-1}\,,$$

$$\begin{aligned}\lim \frac{1}{n}\,(S(\omega \vee \cdots \vee \Phi^{n-1}\omega)) &= \lim S(\Phi^{n-1}\omega, \omega \vee \cdots \vee \Phi^{n-2}\omega)\\ &= \lim \Delta_n = s(\omega, \Phi)\,. \end{aligned} \tag{4.8}$$

The entropy of Φ (KS invariant) is defined as

$$s(\Phi) \equiv \sup s(\omega, \Phi)\,, \tag{4.9}$$

the sup being taken over all finite partitions ω.

It should be noted that, in contradistinction to usual notions of entropy, this kind of entropy is *not* a function of a state but rather a function of the *dynamics* of the system.

The Kolmogorov-Sinai invariant has the following important properties:

(1) It is an *invariant* of the dynamical system in the following sense: the system is described by Ω, the Liouville measure μ, and the measure-preserving one-to-one mapping Φ: $\Omega \to \Omega$. Suppose there is another triple Ω', μ', Φ' with the same properties, and an isomorphism f: $\Omega \to \Omega'$, etc., such that the diagram

$$\begin{tikzcd} \Omega \arrow[r, "\Phi"] \arrow[d, shift right, "f"'] & \Omega \arrow[d, shift right, "f"'] \\ \Omega' \arrow[u, shift right, "f^{-1}"'] \arrow[r, "\Phi'"] & \Omega' \arrow[u, shift right, "f^{-1}"'] \end{tikzcd}$$

is commutative. Then $s(\Phi') = s(\Phi)$.

(2) *Kolmogorov's theorem.* The partition α is called a generator if the σ algebra generated by the sets $\Phi^m(A)(m = 0, \pm 1, \pm 2, \ldots, A \in \alpha)$ is all measurable subsets of Ω. Then

$$s(\Phi) = s(\alpha, \Phi) \tag{4.10}$$

(Kolmogorov, 1953, 1959).

Before stating the next important property, one remark should be made. Namely, all our considerations above apply to abstract dynamical systems too, where Ω need not be phase space or even any smooth manifold, but can be any set. Also Φ need not have anything to do with time evolution but can be any automorphism, for instance, space translation, or any symmetry operation. There is the following theorem relevant to *classical* dynamical systems.

Kouchnirenko's theorem. The KS entropy of finite classical dynamical systems is finite. (For abstract systems it may be infinite.) (Kouchnirenko, 1965, 1967.)

The construction of the KS entropy is very similar to the one of the mean entropy in Sec. II.F. It can be shown that for classical lattice systems one can find a transformation such that they become an abstract dynamical system (clearly, Φ then means "space translation"; Robinson and Ruelle, 1967). Then, if s (the mean entropy) is $<\infty$, KS and mean entropy coincide. This means that the KS entropy, in essence, is a mean entropy. (By the way, it is possible to generalize KS entropy by replacing the group of discrete time translation by more general groups, for instance, Z^d. Many of the important results then, after obvious modifications, remain valid.)

There is a serious problem with the KS entropy because it refers to a discrete time evolution. In the more realistic case of a continuous one-parameter group Φ_t of time evolutions the construction presented above does not work for two reasons: (a) it is not obvious by what quantity $S(\omega \vee \Phi\omega \vee \cdots \vee \Phi^{n-1}\omega)$ has to be replaced. In particular, there may arise measurability questions because in the continuous case uncountable unions and intersections of the sets $\Phi_t \Omega_i$ are involved which need not be measurable. (b) If we adopt the view that the KS entropy is a mean entropy, then, certainly, in the continuous case strong subadditivity enters in a very essential way. Thus, in any case, the construction of an analog of the KS entropy in the continuous case must be much more sophisticated.

As concerns quantum mechanics, one could think of imitating the original method of Kolmogorov and Sinai according to our "translation table" of Sec. I.B. However, this does not work in general. The difficulty lies in the possible noncommutativity. In quantum mechanics, clearly a partition ω has to be defined as a set of pairwise orthogonal projections P_i, with $\sum P_i = 1$. However, if we are given two partitions $\omega_1 = \{P_i^{(1)}\}$ and $\omega_2 = \{P_j^{(2)}\}$, then it is unclear how to define $\omega_1 \vee \omega_2$, since the products $P_i^{(1)}P_j^{(2)}$ in general will not be projections; they will not even be Hermitian. Also the dimension of the algebra generated by ω_1 and ω_2 can be exceedingly large, so that in any case subadditivity arguments cannot be used.

There is partial success in constructing a KS entropy for quantum-mechanical K systems (Emch, 1976). They are analogs of the classical K systems (Kolmogorov, 1953), which are systems with a mixing property that is much stronger than the mixing property we have used in Sec. I.B. Unfortunately, this is rather lengthy to describe and demands a good knowledge of the theory of von Neumann algebras, so I must refer the reader to the original papers. There is also a construction for Bernoulli shifts on the hyperfinite II_1 factor by Connes and Størmer (1975).

Recently, Lindblad (1977) succeeded in giving a definition of a quantum analog of the *KS* entropy which is not based on a noncommutative generalization of partitions but is rather analogous to the definition of the mean entropy for quantum lattice systems.

Besides its interpretation as a mean entropy, *KS* entropy can also be taken as a measure of the strength of mixing of Φ. Remember that

$$s(\omega, \Phi) = \lim_{n\to\infty} [S(\omega \vee \cdots \vee \Phi^n \omega) - S(\omega \vee \cdots \vee \Phi^{n-1}\omega)]\,.$$

Let $n = 1$; the following argument can easily be transferred to the general case. If $S(\omega \vee \Phi\omega) = S(\omega)$, then $\Phi\omega = \omega$, i.e., Φ leaves the sets Ω_i unchanged. If, on the other hand, the difference $S(\omega \vee \Phi\omega) - S(\omega)$ is big, this

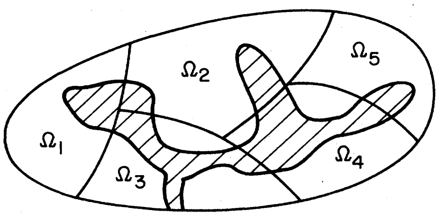

FIG. 10. Interpretation of the Kolmogorov–Sinai-invariant.

means that the intersection of every set $\Phi\Omega_i$ with the original Ω_j must be quite significant. (See Fig. 10. The shaded area is $\Phi\Omega_1$; we have not represented the other sets $\Phi\Omega_2 \cdots \Phi\Omega_5$ for reasons of clearness.) Thus big KS entropy means that the sets of any partition ω get rapidly distributed over the whole phase space, and that the system exhibits strong mixing properties.

Similar to *KS* entropy are Kouchnirenko's A entropies: let A be a sequence of integers $a_1 < a_2 < a_3 < \cdots$. Then

$$s_A(\omega, \Phi) = \limsup_n \frac{1}{n}[S(\Phi^{a_1}\omega \vee \cdots \vee \Phi^{a_n}\omega)]$$

and

$$s_A(\Phi) = \sup_\omega s_A(\omega, \Phi) .$$

They also are invariants of dynamical systems(cf. Arnold and Avez, 1969).

B. Various other concepts

On several occasions we already have met entropylike concepts that were of a certain use, either directly in physics, as, for instance, the coarse-grained entropy, or in order to show that certain properties of the "right" entropy were not as obvious as one might think at first.

Let me write down a short list of these concepts, as far as we were concerned with them, or as they seem to have a certain relevance for physics.

(1) Coarse-grained entropy (see Sec. I.B).

(2) α entropies (see Sec. II.E). One property of α entropies should be added: for $\alpha > 1$, they are continuous, i.e., $\mathrm{Tr}|\rho_n - \rho| \to 0$ implies $S_\alpha(\rho_n) \to S(\alpha)$. For fixed ρ, the mappings $\alpha \to S_\alpha(\rho)$ are convex and decreasing; since $S(\rho) = \sup_{\alpha>1} S_\alpha(\rho)$, this provides a third proof of lower semicontinuity of entropy.

(3) Daroczy and other entropies (see Sec. I.G).

(4) Measures of noncommutativity (see Sec. III.C).

(5) Ingarden–Urbanik (IU) entropy (Ingarden and Urbanik, 1962; Ingarden, 1965, 1973). This concept in fact appeared very early, namely in the papers of the Ehrenfests (1911), Pauli (1928), and von Neumann (1929), but was intensively studied in the 1960s. It arises in connection with considerations about the measurement process. Let $\omega = \{P_i\}$ be a partition of one-dimensional projections, i.e., commensurable "counters" in physical language. Then a measurement yields the numbers $p_i = \mathrm{Tr}\rho P_i$, and the amount of information obtained by this measurement clearly is

$$S_{\mathrm{IU}}(\rho, \omega) \equiv -\sum p_i \ln p_i . \qquad (4.11)$$

[i.e., $S_{\mathrm{IU}}(\rho, \omega)$ is just the $S(\omega)$ of the last section, but with ω now being a partition of one-dimensional projections rather than a finite partition.] The effect of a measurement may be described by transforming the original density matrix ρ into

$$\rho_\omega \equiv \sum P_i \rho P_i . \qquad (4.12)$$

We know that $\rho_\omega \succ \rho$, hence $S(\rho_\omega) \geq S(\rho)$. Performing another measurement corresponding to another partition ω', one obtains

$$(\rho_\omega)_{\omega'} \succ \rho_\omega ,$$

i.e., again a loss of information. For details see Wehrl, 1977, and Staszewski, 1977. It also arises in other situations. Let ω be the set of spectral projections of a Hamiltonian H (and let us assume that there are no degeneracies). Then, using the notation after Eq. (2.2),

$$S_{\mathrm{IU}}(\rho, \omega) = S(\bar{\rho}_\infty) .$$

IU entropy has (of course, besides invariance) many properties in common with classical discrete entropy, for instance concavity, additivity, and subadditivity (the latter ones in some appropriate sense). There are also continuous analogs of it (Grabowski, 1977). Since $S(\rho) = \inf_\omega S_{\mathrm{IU}}(\rho, \omega)$, and $= S_{\mathrm{IU}}(\rho, \omega)$ if and only if ω consists of the spectral projections of ρ, the quantity $S_{\mathrm{IU}}(\rho, \omega) - S(\rho)$ may be considered as a measure of noncommutativity between ρ and the partition ω.

Some concepts measuring the amount of information have been described. The list is not exhaustive and it is left to everyone to invent new such quantities. However, it will be very hard to establish their physical meaning.

C. Systems with infinitely many degrees of freedom

Many theorems of statistical mechanics refer to the infinite case, i.e., systems with infinitely many particles moving in an infinite volume. We have seen that only in this case phenomena such as quantum-mechanical ergodicity, etc. can be expected to hold in a rigorous manner.

1. Description of infinite systems

In Sec. II.F we obtained a description of infinitely extended systems by attaching to every bounded region V a Hilbert space H_V and a density matrix ρ_V; thus we supposed the family of density matrices to be compatible. Remember that in the continuous case H_V was the Fock space $\oplus \mathsf{H}_N^{\pm}(V)$, with $\mathsf{H}_N^{\pm}(V)$ being the space of symmetric (or, antisymmetric, respectively) square-integrable functions $\psi(x_1, \ldots, x_N)$, where the arguments x_i were restricted to V.

It seems to be quite natural to describe an infinitely extended system in d dimensions simply by replacing V by $\mathbf{R}^d$. The Hilbert space then would be

$$\mathsf{H}^{\pm}(\mathbf{R}^d) = C \oplus L^2(\mathbf{R}^d) \oplus [L^2(\mathbf{R}^d) \underset{s(a)}{\otimes} L^2(\mathbf{R}^d)] \oplus \cdots . \qquad (4.13)$$

This construction makes perfect sense. The unfortunate thing, however, is that, in general, there is no density matrix in this space describing the state. To be more precise, for any bounded region V, as in Eq. (1.9),

$$\mathsf{H}^{\pm}(V)\otimes\mathsf{H}^{\pm}(\mathrm{R}^{d}\backslash V)=\mathsf{H}^{\pm}(\mathrm{R}^{d})\,, \tag{4.14}$$

but there is no density matrix ρ in $\mathsf{H}^{\pm}(\mathrm{R}^{d})$ such that

$$\mathrm{Tr}_{\mathsf{H}^{\pm}(\mathrm{R}^{d}\backslash V)}\,\rho=\rho_{V}\,.$$

If such a density matrix existed, then, for instance, the particle density $n=\lim N(V)/|V|$ would be $=0$. This means that the Hilbert space (4.13) cannot be the right one for the description of the system.

The algebraic approach (see Ruelle, 1969; Eckmann and Guenin, 1969; Emch, 1972) now essentially proposes the following procedure:

Since it is at first unclear what the right Hilbert space of the system is, one should not worry too much about it. One should rather concentrate on the operators representing the observables of the system. Let $\mathsf{A}(V)$ be the algebra for all operators on $\mathsf{H}(V)$, for V bounded. If $V'\supset V$, then every operator $T\in\mathsf{A}(V)$ can be identified with the operator $T'=T\otimes 1$ on $\mathsf{H}(V')=\mathsf{H}(V)\otimes\mathsf{H}(V'/V)$, hence $\mathsf{A}(V)\subset\mathsf{A}(V')$ (isotony). Define

$$\mathsf{A}\equiv\cup\,\mathsf{A}(V)\,. \tag{4.15}$$

This is again an algebra since sums and products make sense. Also there is a norm defined on it, namely, if $T\in\mathsf{A}(V)$, just the usual operator norm. Hence A is a normed algebra, and, if we take its norm completion (which by abuse of language we also will denote by A), it becomes a C^* algebra.

Every family of density matrices ρ_V defines a state on this algebra: let $T\in\mathsf{A}(V)$. Define

$$\omega(T)=\mathrm{Tr}\rho_V T\,. \tag{4.16}$$

We know that, for $V'\supset V$,

$$\mathrm{Tr}\rho_{V'}(T\otimes 1)=\mathrm{Tr}\rho_V T=\omega(T)\,,$$

$T\otimes 1$ being an operator on $\mathsf{H}(V)\otimes\mathsf{H}(V'\backslash V)=\mathsf{H}(V')$. Hence, this definition of ω extends to every element of all of A and makes sense. Note that $\omega(1)=1$, and, for $T\geqslant 0$, $\omega(T)\geqslant 0$. Clearly, ω also is linear, i.e., in the language of mathematics, it is a positive, normed, linear functional on A.

There is now a canonical way of constructing a Hilbert space for the system: the so-called Gelfand–Naimark–Segal (GNS) construction (Segal, 1951; cf. also Dixmier, 1964). It tells us that (up to isomorphisms) there is exactly one Hilbert space H_ω, a representation π_ω of A [i.e., a homomorphism π_ω: $\mathsf{A}\to\mathsf{B}(\mathsf{H}_\omega)$], and a unique cyclic vector $\Omega_\omega\in\mathsf{H}_\omega$ such that

$$\omega(A)=\langle\Omega_\omega|\pi_\omega(A)|\Omega_\omega\rangle$$

for all $A\in\mathsf{A}$. By "cyclic vector" is meant that the set $\pi_\omega(A)|\Omega_\omega\rangle$ is dense in H_ω.

Of course, as a rule, this Hilbert space will be entirely different from Fock space. The mathematical reason for this is that, in the case of infinitely many degrees of freedom, there are infinitely many inequivalent representations of the canonical commutation (or anticommutation) relations (cf. Emch, 1972). It is important to note that the Hilbert space H_ω depends explicitly on ω. For a Gibbs state this means that it depends on the temperature.

We have not yet said anything about time evolution. Regarding that, from the algebraic point of view, the time evolution in $\mathsf{B}[\mathsf{H}(V)]$, the bounded linear operators on $\mathsf{H}(V)$,

$$T\to e^{iHt}T\,e^{-iHt}\,,$$

is nothing else than an automorphism of the algebra $\mathsf{A}(V)=\mathsf{B}[\mathsf{H}(V)]$; it is natural to consider the time evolution in A also as an automorphism of A (or, better, as a one-parameter group of automorphisms τ_t: $\mathsf{A}\to\mathsf{A}$). In general there will not be a Hamiltonian $H\in\mathsf{A}$ such that

$$\tau_t A=e^{iHt}A\,e^{-iHt}\,.$$

However, in the *GNS* construction performed with a time-invariant state ω, i.e., an ω such that $\omega(A)=\omega(\tau_t A)$ for all t, there exists a H_ω such that

$$\pi_\omega(\tau_t A)=e^{iH_\omega t}\pi_\omega(A)\,e^{-iH_\omega t}\,. \tag{4.17}$$

In that case, Ω_ω is invariant:

$$e^{iH_\omega t}\Omega_\omega=\Omega_\omega\,. \tag{4.18}$$

[Note that H_ω in general neither belongs to $\pi_\omega(\mathsf{A})$ nor can be constructed from it by some limiting procedure.]

Admittedly the above scheme looks a little bit complicated but on the other hand it is very powerful because it not only gives a description of infinite continuous quantum systems but also covers the cases of finite continuous quantum systems, finite and infinite quantum lattice systems, and all sorts of classical systems. [Once more, Ruelle's book (1969) should be consulted for these questions.]

2. Mixing

Utilizing the algebraic approach we are now in the position to deal with the problems of ergodicity and mixing (cf. Sec. I.B) in quantum mechanics. We have discussed in Sec. I.B the fact that mixing means

$$\omega(P_tQ)\to\omega(P)\omega(Q)\quad\text{as }t\to\pm\infty$$

[Eq. (2.23a)]. Whereas this turned out to be impossible for finite quantum systems, it is very well possible in infinite systems that there is an invariant state ω such that, for any two elements $A,B\in\mathsf{A}$,

$$\lim_{t\to\pm\infty}\ \omega(\tau_t A\cdot B)=\ \omega(A)\omega(B)\ . \tag{4.19}$$

(The concept of ergodicity for invariant states on C^* algebras was introdued by Segal, 1951.)

In the GNS representation this reads

$$e^{iH_\omega t}\ \xrightarrow{\text{weakly}}\ |\Omega_\omega\rangle\langle\Omega_\omega| \tag{4.20}$$

for $t\to\pm\infty$, by virtue of Eq. (4.18). H_ω does not have any other eigenvector than Ω_ω, because $H_\omega\psi=\lambda\psi$, ψ orthogonal to Ω_ω, implies

$$\|\psi\|^2=\langle\psi|\psi\rangle=|\langle\psi|e^{iH_\omega t}\psi\rangle|\to|\langle\psi|\Omega_\omega\rangle|^2=0\ .$$

Hence H_ω has a continuous spectrum on the orthogonal complement of Ω_ω.

It is usually supposed that the time evolution is asymptotically Abelian (Doplicher, Kastler, and Robinson, 1966; Ruelle, 1966; Doplicher, Kastler, Kadison, and Robinson, 1967), i.e., that the commutator $[\tau_t A, B]$ vanishes in some appropriate sense as $t \to \pm\infty$. (There are different notions of this property which we do not want to discuss in detail here. So let us for simplicity suppose that $\|[\tau_t A, B]\| \to 0$ as $t \to \infty$.) This is certainly true for free systems where the commutator goes as $t^{-3/2}$. For systems with repulsive forces only, one can expect even stronger commutation properties, and for attractive forces, asymptotic Abelianness will presumbaly hold as long as the attraction is not too strong.

If asymptotic Abelianness is true, then $\lim \omega(C(\tau_t A)B) = \lim \omega(\tau_t A \cdot BC) = \omega(A)\omega(BC)$, hence in the GNS construction $\pi_\omega(\tau_t A) \to \omega(A)$ times the unit operator in H_ω. Now let ω' be a state that is normal with respect to ω, by which we mean that there exists a density matrix ρ' in H_ω such that $\omega'(A) = \mathrm{Tr}\rho'\pi_\omega(A)$. Then $\omega'(\tau_t A) = \mathrm{Tr}\rho'\pi_\omega(\tau_t A) \to \omega(A)$; i.e., states not too far from a mixing state converge towards the latter. This is a rigorous result concerning approach to equilibrium (Sec. I.B).

Now take a bounded subvolume V. For any $A \in \mathsf{A}(V)$, $\omega_t'(A) \equiv \omega'(\tau_t A) \to \omega(A)$. Let ρ_V', ρ_V be density matrices on $\mathsf{H}(V)$ defined by

$$\omega'(A) = \mathrm{Tr}\rho_V' A,$$

$$\omega(A) = \mathrm{Tr}\rho_V A\ ,$$

and let $\rho_V'(t)$ be the time evolution of ρ_V', defined in an obvious manner. Then

$$\mathrm{Tr}\,\rho_V'(t)A \to \mathrm{Tr}\rho_V A,$$

and, consequently, $\mathrm{Tr}\,|\rho_V'(t) - \rho_V| \to 0$ (see Davies, 1972; Wehrl, 1976). As we have seen in Sec. II.D, this does not necessarily imply that $S(\rho_V'(t)) \to S(\rho_V)$, but under some weak additional assumptions (which in general can be expected to be fulfilled) this will in fact be true.

It usually is not possible to define the entropy of the state ω of the whole system; any sensible definition would give $S = \infty$. However, in addition to the mean entropy (Secs. II.F and III.A), one can define the relative entropy of two states by

$$S(\omega' \,|\, \omega) \equiv \lim S(\rho_V' \,|\, \rho_V)\ . \tag{4.21}$$

This concept turns out to be very useful for infinite systems too; however, due to mathematical complications (one has to know about Tomita–Takesaki theory), we have to refer the reader to the literature (Araki, 1975).

3. KMS states

In general, for an infinite system there exists no operator H belonging to A or which can be constructed as a limit of elements of A such that the time evolution is given by $A \to e^{iHt}Ae^{-iHt}$. Therefore one also cannot use Eq. (1.39) to describe Gibbs states. But one can use the KMS condition (Sec. I.B) in order to obtain an analog of them: A state ω is called a KMS state at inverse temperature β if there is a function $F(z)$ with the analyticity properties stated after Eq. (1.47), namely

$$\omega(B\tau_t A) = F(t), \quad \omega(\tau_t AB) = F(t + i\beta)\ . \tag{4.22}$$

KMS states have attracted great interest in recent years both from the physical and the mathematical side, and there is a rich literature about them (the study of KMS states was initiated by Haag, Hugenholtz, and Winnink, 1967). Let me just mention a few results.

(1) KMS states are automatically time invariant.

(2) To a given state ω there is exactly one group of time automorphisms τ_t such that ω is KMS for them. (However, there may be more than one KMS state for a given time evolution.)

(3) KMS states can be decomposed into extremal ones, i.e., those that cannot be written as a genuine convex combination of two other KMS states. These extremal KMS states are factorial, i.e., $\pi_\omega(\mathsf{A})''$ is a factor. [$\pi_\omega(\mathsf{A})'$ = set of all operators on H_ω that commute with all of $\pi_\omega(\mathsf{A})$, $\pi_\omega(\mathsf{A})''$ = all operators that commute with $\pi_\omega(\mathsf{A})'$. $\pi_\omega(\mathsf{A})'$ is called the commutant of $\pi_\omega(\mathsf{A})$, $\pi_\omega(\mathsf{A})''$ the bicommutant. "Factor" means that the center $\pi_\omega(\mathsf{A})' \cap \pi_\omega(\mathsf{A})''$ consists of the multiples of the identity only.]

(4) Factorial states (whether they are KMS or not) are always mixing.

(5) $\pi_\omega(\mathsf{A})'$ and $\pi_\omega(\mathsf{A})''$ are anti-isomorphic. There is a deep theory studying this symmetry: the so-called Tomita–Takesaki theory, which is one of the most fruitful recent concepts in the field of operator algebras (Takesaki, 1970).

4. Stability

We already have mentioned stability properties of equilibrium states: small perturbations of the dynamics do not lead to global changes of the state. Let me sketch one result in this direction (Haag, Kastler, and Trych-Pohlmeyer, 1974; Haag and Trych-Pohlmeyer, 1977; for another approach, cf. Araki and Sewell, 1977).

A small, local perturbation of the dynamics may be described by changing τ_t to $\tau_t^{\lambda h}$, where $\tau_t^{\lambda h}$ is defined via its infinitesimal generator (which is in mathematical language a derivation of the algebra A) as

$$i\frac{\partial}{\partial t}\tau_t^{\lambda h} = i\frac{\partial}{\partial t}\tau_t + \lambda h\ ,$$

with $h \in \mathsf{A}$. (One can also write down $\tau_t^{\lambda h}$ directly as an infinite series involving time-ordered integrals of multi-commutators.) Let ω (or $\omega^{\lambda h}$, respectively) where $\omega^{\lambda h}$ is defined in a similar way to $\tau_t^{\lambda h}$, be a time-invariant state of the unperturbed, or perturbed, system, respectively, and suppose that for every h

$$\|\omega^{\lambda h} - \omega\| \to 0 \quad \text{as } \lambda \to 0\ . \tag{4.23}$$

If $\int \|[\tau_t A, B]\|dt < \infty$, $\int \|[\tau_t^{\lambda h} A, B]\|dt < \infty$ (for small λ), then, for factorial states ω, they turn out to be KMS for some β. (The β comes in as some "modulus of stability.") On the other hand, every factorial KMS state has the stability property (4.23).

Let me close with a few words about a very general concept of entropy that refers to von Neumann algebras, i.e., weakly closed*-algebras of operators containing the identity. [Examples are $\pi(\mathsf{A})'$, $\pi(\mathsf{A})''$; in fact von Neumann algebras are exactly those operator algebras N for which $\mathsf{N} = \mathsf{N}''$.]

5. Segal entropy (Segal, 1960)

We have remarked at the end of Sec. I.A that this is in some sense the most general concept of entropy. It is defined as follows: let $\mathsf{N} \subset \mathsf{B}(\mathsf{H})$ be a von Neumann algebra. Let Φ be a faithful normal semifinite trace on N, i.e., a mapping of the positive part of N into $[0,\infty]$ such that $\Phi(R) \neq 0$ if $R \neq 0$, $\Phi(\lambda R) = \lambda\Phi(R)(\lambda \geqslant 0)$, $\Phi(R+S) = \Phi(R) + \Phi(S)$, $\Phi(U^*RU) = \Phi(R)$ for $U =$unitary, $\in \mathsf{N}$; furthermore if $R_n \uparrow R$, then $\Phi(R_n) \uparrow \Phi(R)$; and finally that, to every R there exists $S \neq 0, \leqslant R$, with $\Phi(S) < \infty$.

(The usual trace $\mathrm{Tr}\,\cdot$ fulfills all requirements; however, there are algebras such that $\mathrm{Tr}\, T = \infty$ for every positive $T \neq 0$).

If ψ is a normal state that can be written as $\psi(\cdot) = \Phi(\rho\,\cdot)$, for some ρ (the set of all those ψ is dense), then let $\rho = \int_0^\infty \lambda dE(\lambda)$ be its spectral decomposition. [All $E(\lambda)$ belong to N.] The Segal entropy is defined as

$$S(\psi \,|\, \Phi) = -\int_0^\infty \lambda \ln\lambda\, d\Phi(E(\lambda)) \ . \tag{4.24}$$

The most important special cases are:

(1) Let (Ω, μ) be a measure space. Take $\mathsf{H} = L^2(\Omega, \mu)$, $\mathsf{N} = L^\infty(\Omega, \mu)$ (which is a von Neumann algebra). The mapping $\Phi(f) = \int f d\mu$ for $f \in L^\infty$ is a trace in the above sense. Let $\psi(f) = \int g f d\mu$. g is the Radon–Nikodym derivative of the measure ν defined by

$$\nu(M) = \int_M g f d\mu \ ,$$

and hence

$$S(\psi \,|\, \Phi) = S(\nu \,|\, \mu) \ ,$$

i.e., the generalized BGS entropy (Sec. I.A).

(2) Let $\mathsf{N} = \mathsf{B}(\mathsf{H})$, $\Phi =$the usual trace Tr, ψ be the state given by $\psi(\cdot) = \mathrm{Tr}\rho\,\cdot$. Then $S(\psi \,|\, \Phi) =$the usual quantum-mechanical entropy $S(\rho)$.

Many of the properties of classical and quantum-mechanical entropy have generalizations to Segal entropy, and there are also characterization theorems similar to those of Sec. II.B and F (Ochs and Spohn, 1976).

ACKNOWLEDGMENTS

I am most grateful to Professors S. B. Treiman and E. H. Lieb for suggesting this paper and for their encouragement. Moreover, I have obtained invaluable aid from Professor Lieb during the preparation of this work. I also acknowledge useful comments by H. Spohn. I have learned many facts about entropy verbally from my teacher, Professor W. Thirring. Several theorems were communicated to me by Professor A. Uhlmann. Finally I want to thank my colleagues Heide Narnhofer, A. Pflug, and G. Siegl, as well as Franzi Wagner.

REFERENCES

Aczel, J., and Z. Daroczy, 1963, C. R. Acad. Sci. Paris **257**, 1581.

Aczel, J., 1974, *On Measures of Information and Their Characterizations* (Academic, New York).

Aczel, J., B. Forte, and C. T. Ng, 1974, Adv. Appl. Prob. **6**, 131.

Alberti, P., 1973, thesis, Leipzig.

Alberti, P., 1977, to be published in Wiss. Z. Karl-Marx-Univ., Leipzig.

dell'Antonio, G. F., 1967, Commun. Pure Appl. Math. **20**, 413.

Araki, H., 1975, preprint RIMS 190, Kyoto.

Araki, H., and E. Lieb, 1970, Commun. Math. Phys **18**, 160.

Araki, H., and G. L. Sewell, 1977, Commun. Math. Phys. **52**, 103.

Arnold, V., and A. Avez, 1969, *Ergodic Problems of Classical Mechanics* (Benjamin, New York).

Baumann, F., and R. Jost, 1969, in *The Problems of Theoretical Physics: Essays Devoted to N. N. Bogoliubov* (Nauka, Moscow).

Baumann, F., 1971, Helv. Phys. Acta **44**, 95.

Bayer, W., and W. Ochs, 1973, Z. Naturforsch. A **28a**, 693.

Berezin, F., 1972a, Izv. Akad. Nauk SSSR Ser. Mat. **36**, 5.

Berezin, F., 1972b, Mat. Sb. Nov. Ser. **86** (128), 4 (12); **88** (130); 2 (6).

Billingsley, P., 1965, *Ergodic Theory and Information* (Wiley, New York).

Birkhoff, G. D., Proc. Natl. Acad. Sci. USA **17**, 656.

Blau, J. M., 1959, Prog. Theor. Phys. **22**, 745.

Boltzmann, L., 1872, Wiener Ber. **66**, 275.

Boltzmann, L., 1877a, Wiener Ber. **75**, 67.

Boltzmann, L., 1877b, Wiener Ber. **76**, 373.

Boltzmann, L., 1896, *Vorlesungen über Gastheorie* (Barth, Leipzig).

Brillouin, L., 1962, *Science and Information Theory* (Academic, New York).

Brush, S., 1965, 1966, *Kinetic Theory I, II* (Pergamon, Oxford).

Burbury, S., 1890, Philos. Mag. **30**, 301.

Carnot, S., 1824, *Reflexions sur la puissance matrice du feu* (Bachelier, Paris).

Clausius, R., 1850, *Über die bewegende Kraft der Wärme* (Ostwalds Klassiker, Leipzig, 1898).

Clausius, R., 1865, Ann. Phys. **125**, 353.

Cohen, E. G. D., and W. Thirring, eds., 1972, *The Boltzmann Equation* (Springer, Vienna).

Connes, A., and E. Størmer, 1975, Acta Math. **134**, 289.

Daroczy, Z., 1970, Inf. Control **16**, 36.

Davies, E., 1972, Commun. Math. Phys. **27**, 309.

Davies, E., 1976, *Quantum Theory of Open Systems* (Academic, London).

Dixmier, J., 1957 (1969), *Les algêbres d' opêrateurs dans l'espace Hilbertien* (Gauthier-Villars, Paris).

Dixmier, J., 1964, *Les C*algêbres et leurs reprêsentations* (Gauthiers-Villars, Paris).

Doplicher, S., D. Kastler, R. Kadison, and D. Robinson, 1967, Commun. Math. Phys. **6**, 101.

Doplicher, S., D. Kastler, and D. Robinson, 1966, Commun. Math. Phys. **3**, 1.

Dyson, F. J., 1967, J. Math. Phys. **8**, 1538.

Dyson, F. J., and A. Lenard, 1967, J. Math. Phys. **8**, 423.

Eckmann, J. P., and M. Guenin, 1969, *Mêthodes algebriques dans la mêcanique statistique* (Springer, Berlin).

Ehrenfest, P. and T., 1911, Encyclopädie der math. Wiss. **4**, Article 32.

Einstein, A., 1914, Verh. Dtsch. Phys. Ges. **12**, 820.

Emch, G., 1972, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley, New York).

Emch, G., 1976a, Acta Phys. Austriaca, Suppl XV, 79.

Emch, G., 1976b, Commun. Math. Phys. **49**, 191.

Epstein, H., 1973, Commun. Math. Phys. **31**, 317.

Erdös, P., 1946, Ann. Math. **47**, 1.

Fano, U., 1957, Rev. Mod. Phys. **29**, 74.

Feller, W., 1950, *An Introduction to Probability Theory and Its Applications* (Wiley-Chapman, New York).

Fierz, M., 1955, Helv. Phys. Acta **28**, 705.

Fisher, M. E., 1964, Arch. Ration. Mech. Anal. **17**, 377.

Fréchet, M., 1938, *Méthode des fonctions arbitraires, théorie des évenements en chaîne dans le cas d'un nombre fini d'états possibles* (Dunod, Paris).

Gibbs, J. W., 1902, *Elementary Principles in Statistical Mechanics* (Yale University, New Haven, Conn.).

Glasman, I. M., and Y. I. Gubich, 1969, *Finite-dimensional Linear Analysis,* in Russian (Gostekhisdat, Moscow).

Golden, S., 1965, Phys. Rev. B **137**, 1127.

Gorini, V., A. Kossakowski, and E. C. G. Sudarshan, 1976, J. Math. Phys. **17**, 821.

Gorini, V., A. Frigerio, A. Kossakowski, E. C. G. Sudarshan, and M. Verri, 1976, to be published in Rep. Math. Phys.

Grabowski, M., 1977, to be published in Rep. Math. Phys.

Grad, H., 1958, in *Handbuch der Physik, Vol. XII: Thermodynamik der Gase,* edited by S. Flügge (Springer, Berlin).

Griffiths, R. B., 1965, J. Math. Phys. **6**, 1447.

Guichardet, A., 1974, *Systèmes dynamiques non-commutatifs* (Gauthier-Villars, Paris).

Haag, R., N. M. Hugenholtz, and M. Winnink, 1967, Commun. Math. Phys. **5**, 215.

Haag, R., D. Kastler, and E. Trych-Pohlmeyer, 1974, Commun. Math. Phys. **38**, 143.

Haag, R., and E. Trych-Pohlmeyer, 1977, to be published in Commun. Math. Phys.

Halmos, P. R., 1956, *Lectures in Ergodic Theory* (Chelsea, New York).

Hardy, G. H., J. E. Littlewood, and G. Polya, 1934 (1967), *Inequalities* (Cambridge University, Cambridge, England).

von Helmholtz, H., 1854, reprinted in *Popular Scientific Lectures,* 1962 (Dover, New York).

Hertel, P., H. Narnhofer, and W. Thirring, 1972, Commun. Math. Phys. **28**, 159.

Hertel, P., and W. Thirring, 1971, Commun. Math. Phys. **24**, 22.

Hopf, E., 1932, Proc. Natl. Acad. Sci. USA **18**, 204.

van Hove, L., 1949, Physica **15**, 951.

van Hove, L., 1950, Physica **16**, 137.

van Hove, L., 1962, in *Fundamental Problems in Statistical Mechanics,* edited by E. G. D. Cohen (North-Holland, Amsterdam).

Huang, K., 1963, *Statistical Mechanics* (Wiley, New York).

Ingarden, R. S., 1965, Fortschr. Phys. **13**, 755.

Ingarden, R. S., 1973, Acta Phys. Pol. **43**, 3.

Ingarden, R. S., and K. Urbanik, 1962, Acta Phys. Pol. **21**, 281.

Jacobs, K., 1960, *Neuere Methoden und Ergebnisse der Ergodentheorie* (Springer, Berlin).

Jacobs, K., 1963, *Lectures on Ergodic Theory* (Aarhus University, Aarhus).

Jaynes, E. T., 1957, Phys. Rev. **106**, 620.

Joule, J. P., 1845, Phil. Mag. **27**, 205.

van Kampen, N. G., 1962, in *Fundamental Problems in Statistical Mechanics,* edited by E. G. D. Cohen (North-Holland, Amsterdam).

Katai, S., 1967, Ann. Univ. Sci. Budapest, Eötvös Sect. Math. **12**, 81.

Klauder, J. R., and E. C. G. Sudarshan, 1968, *Fundamentals of Quantum Optics* (Benjamin, New York).

Klein, M., 1972, in *The Boltzmann Equation,* edited by E. G. D. Cohen and W. Thirring (Springer, Vienna).

Koenig, F. O., 1959, in *Men and Moments in the History of Science,* edited by H. M. Evans (University of Washington, Seattle).

Kolmogorov, A. N., 1953, Dokl. Akad. Nauk SSSR **119**, 861.

Kolmogorov, A. N., 1959, Dokl. Akad. Nauk SSSR **124**, 754.

Koopman, B. O., 1931, Proc. Natl. Acad. Sci. USA **17**, 315.

Kossakowski, A., 1972, Rep. Math. Phys. **3**, 247.

Köthe, G., 1960, *Topologische lineare Räume I* (Springer, Berlin).

Kouchnirenko, A. G., 1965, Dokl. Akad. Nauk SSSR **161**, 37.

Kouchnirenko, A. G., 1967, Funkz. Analys i ego Prilojenija, Moscow, **1**, 103.

Kraus, K., 1970, Ann. Phys. **64**, 311.

Kubo, R., 1957, J. Phys. Soc. Japan **12**, 570.

Ky Fan, 1959, Proc. Natl. Acad. Sci. USA **35**, 652.

Landau, L. D., and E. M. Lifschitz, 1966, *Statistische Physik* (Akademie, Berlin).

Lanford, O. E., 1975, in *Dynamical Systems, Theory and Applications,* edited by J. Moser (Springer, Berlin).

Lanford, O. E., and D. Robinson, 1968, J. Math. Phys. **9**, 1120.

Lassner, G. and G., 1978, to be published in Rep. Math. Phys.

Lebowitz, J., and E. Lieb, 1969, Phys. Rev. Lett. **22**, 631.

Lebowitz, J., and H. Spohn, 1977, preprint, Yeshiva University.

Lenard, A., and F. J. Dyson, 1968, J. Math. Phys. **9**, 698.

Lewis, M. B., and A. Siegert, 1956, Phys. Rev. **101**, 1227.

Lieb, E., 1973a, Commun. Math. Phys. **31**, 327.

Lieb, E., 1973b, Adv. Math. **11**, 267.

Lieb, E., 1975, Bull. Am. Math. Soc. **81**, 1.

Lieb, E., 1976, Rev. Mod. Phys. **48**, 553.

Lieb, E., and J. Lebowitz, 1972, Adv. Math. **9**, 316.

Lieb, E., and M. B. Ruskai, 1973a, Phys. Rev. Lett. **30**, 434.

Lieb, E., and M. B. Ruskai, 1973b, J. Math. Phys. **14**, 1938.

Lieb, E., and W. Thirring, 1975, Phys. Rev. Lett. **35**, 687.

Lindblad, G., 1973, Commun. Math. Phys. **33**, 305.

Lindblad, G., 1974, Commun. Math. Phys. **39**, 111.

Lindblad, G., 1975, Commun. Math. Phys. **40**, 147.

Lindblad, G., 1976a, Commun. Math. Phys. **48**, 119.

Lindblad, G., 1976b, Lett. Math. Phys. **1**, 219.

Lindblad, G., 1977, Preprint, Stockholm.

Loschmidt, J., 1876, Wiener Ber. **73**, 128.

Maison, D., 1971, Commun. Math. Phys. **22**, 166.

Martin, P. C., and J. Schwinger, 1959, Phys. Rev. **115**, 1342.

Maxwell, J. C., 1860, Philos. Mag. **19**, 19; **20**, 21.

Maxwell, J. C., 1871, *Theory of Heat* (Longmans, London).

Mayer, J. E., 1961, J. Chem. Phys. **34**, 1207.

Mayer, R., 1842, *Die Mechanik der Wärme* (Ostwalds Klassiker, Leipzig, 1911).

Mead, A., 1977, J. Chem. Phys. **66**, 459.

von Mises, R., 1931, *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik* (Deuticke, Leipzig).

Naudts, J., 1969, thesis, Brussels, unpublished.

von Neumann, J., 1927, Gött. Nachr. 273.

von Neumann, J., 1929, Z. Phys. **57**, 30.

von Neumann, J., 1932, Proc. Natl. Acad. Sci. USA **18**, 710.

Ochs, W., 1975, Rep. Math. Phys. **8**, 109.

Ochs, W., 1976, Rep. Math. Phys. **9**, 135.

Ochs, W., and W. Bayer, 1973, Z. Naturforsch. A **28a**, 1571.

Ochs, W., and H. Spohn, 1976, preprint, München.

Pauli, W., 1928, in *Festschrift zum 60. Geburtstage A. Sommerfelds,* edited by P. Debye (Hirzel, Leipzig).

Pauli, W., and M. Fierz, 1937, Z. Phys. **106**, 572.

Peierls, R., 1936, Proc. Camb. Philos. Soc. **32**, 477.

Planck, M., 1906, *Vorlesungen uber die Theorie der Wärmestrahlung* (Barth, Leipzig).

Poincaré, H., 1890, Acta Math. **13**, 1.

Polya, G., and G. Szegö, 1970, *Aufgaben und Lehrsätze der Analysis* (Springer, Heidelberg).

Prigogine, I., 1972, in *The Boltzmann Equation,* edited by E. G. D. Cohen and W. Thirring (Springer, Vienna).

Reed, M., and B. Simon, 1972, *Methods of Modern Mathematical Physics* (Academic, New York).

Reif, F., 1965, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York).

Renyi, A., 1965, Rev. Int. Statist. Inst. **33**, 1.

Renyi, A., 1966, *Wahrscheinlichkeitsrechnung* (Deutscher Verlag der Wissenschaften, Berlin).

Robertson, A. and W., 1964, *Topological Vector Spaces* (Cambridge University, Cambridge, England).

Robinson, D. W., 1971, *The Thermodynamical Pressure in*

Quantum Statistical Mechanics (Springer, Berlin).

Robinson, D. W., and D. Ruelle, 1967, Commun. Math. Phys. **5**, 288.

Roller, D., 1950, *The Early Development of the Concepts of Temperature and Heat* (Harvard University, Cambridge, Mass.).

Ruch, E., 1975, Theor. Chim. Acta **38**, 167.

Ruch, E., and A. Mead, 1976, Theor. Chim. Acta **41**, 95.

Ruelle, D., 1966, Commun. Math. Phys. **3**, 133.

Ruelle, D., 1969, *Statistical Mechanics* (Benjamin, New York).

Schlögl, F., 1976, Z. Phys. B **25**, 411.

Schrödinger, E., 1927, Naturwissenschaften **14**, 644.

Segal, I. E., 1951, Duke Math. J. **18**, 221.

Segal, I. E., 1960, J. Math. Mech. **9**, 623.

Shannon, C., and W. Weaver, 1949, *The Mathematical Theory of Communication* (University of Illinois, Urbana).

Simon, B., 1973, appendix to Lieb and Ruskai, J. Math. Phys. 1973.

Simon, B., private communication.

Sinai, Ya. G., 1961, Dokl. Akad. Nauk SSSR **25**, 899.

Sinai, Ya. G., 1965, Usp. Math. Nauk **20**, 232.

von Smoluchowski, M., 1914, *Vorträge über die kinetische Theorie der Materie und Elektrizität* (Teubner, Leipzig).

Spohn, H., private communication.

Spohn, H., 1977, preprint.

Staszewski, P., 1977, to be published in Rep. Math. Phys.

Stinespring, W. F., 1955, Proc. Am. Math. Soc. **6**, 211.

Szilard, L., 1925, Z. Phys. **32**, 777.

Takesaki, M., 1970, *Tomita's Theory of Modular Hilbert Algebras and Its Applications* (Springer, Berlin).

Takesaki, M., 1972, J. Funct. Anal. **9**, 306.

Thirring, W., 1975, *Vorlesungen über mathematische Physik* (University of Vienna).

Thompson, C., 1965, J. Math. Phys. **6**, 1812.

Thomson, W. (Lord Kelvin), 1852, Philos. Mag. **4**, 304.

Thomson, W. (Lord Kelvin), 1857, Proc. R. Soc. Edinb. **3**, 139.

Uhlmann, A., 1971, Wiss. Z. Karl-Marx-Univ. Leipzig **20**, 633.

Uhlmann, A., 1972, Wiss. Z. Karl-Marx-Univ. Leipzig **21**, 427.

Uhlmann, A., 1973, Wiss. Z. Karl-Marx-Univ. Leipzig **22**, 139.

Uhlmann, A., 1975, Rep. Math. Phys. **7**, 449.

Uhlmann, A., 1976, Commun. Math. Phys. **54**, 21.

Uhlmann, A., 1977, preprint, Leipzig.

Umegaki, H., 1954, Tôhoku Math. J. **6**, 177.

Umegaki, H., 1962, Kodai Math. Sem. Rep. **14**, 59.

Wehrl, A., 1973, Acta Phys. Austriaca **37**, 361.

Wehrl, A., 1974, Rep. Math. Phys. **6**, 15.

Wehrl, A., 1975, Acta Phys. Austriaca **41**, 197.

Wehrl, A., 1976a, to be published in Rep. Math. Phys.

Wehrl, A., 1976b, Lectures at the CIME Summer School in Bressanone.

Wehrl, A., 1976c, Rep. Math. Phys. **10**, 159.

Wehrl, A., 1977, to be published in Rep. Math. Phys.

Wigner, E. P., and M. M. Yanase, 1963, Proc. Natl. Acad. Sci. USA **49**, 910.

Yang, C. N., and T. D. Lee, 1952, Phys. Rev. **87**, 404.

Zermelo, E., 1896, Ann. Phys. **57**, 485; **59**, 793.