

# Entropy, Information and Quantum Measurements

Göran Lindblad

Department of Theoretical Physics, Royal Institute of Technology, Stockholm, Sweden

Received March 20, 1973

**Abstract.** The conditional entropy between two states of a quantum system is shown to be nonincreasing when a complete measurement is performed on the system. The information between two quantum systems is defined and is shown to be bounded above by the logarithmic correlation. This inequality is then applied to the measurement process. The entropy changes in the observed system and the measuring apparatus are compared with the information gain in the measurement.

## 1. Introduction

The object of this paper is to discuss the measurement process in quantum theory using the concepts of information and entropy. The starting point is a model of the measurement process introduced by von Neumann [1] which treats both the system to be observed ( $S$ ) and the measuring instrument ( $M$ ) as finite quantum systems. The aim is to obtain relations between the information obtained in the measurement and the entropy changes in  $S$  and  $M$ .

In § 2 the fundamental quantities of information, entropy and conditional entropy are introduced. The entropy used here is the statistical (or Gibbs') entropy, which cannot, in general, be identified with the thermodynamic entropy. The further development is based on an inequality for the conditional entropy which says that the conditional entropy is nonincreasing when a complete measurement is performed on the system (§ 3). The inequality is believed to hold also under a general measurement, and this conjecture can be shown to be equivalent to the conjecture that the entropy for a quasilocal quantum system is strongly subadditive [2]. The information between two subsystems of a quantum systems in a given state is defined, and from the inequality above it is proved to be bounded above by the logarithmic correlation (§ 4). This result is applied to the measurement process. If the system  $S + M$  is closed, then the time evolution is a unitary transformation and the total entropy is conserved. Under the assumption that any eigenstate of the observable to be measured is conserved under the interaction of  $S$  and  $M$ , it is proved that the entropy of  $M$  increases by at least as much as the information gained during the measurement

(§ 6). Furthermore we consider the entropy change in  $S$  during the interaction and the reduction in the entropy of  $S$  that can be achieved by using the information contained in  $M$ . We then have a quantum-mechanical analogue of Szilard's discussion of Maxwell's demon where he argues for the identification of entropy with lack of information [3–6], but the entropy concept considered there is the thermodynamic entropy and consequently he has to postulate the nondecrease of entropy. Szilard's argument has the weakness that it does not keep within the bounds of thermodynamics, and therefore it has been criticized for using procedures not allowed in that theory [7]. In order to beat the second law of thermodynamics the demon tries to use the fluctuations inherent in the picture of equilibrium in statistical mechanics. The argument goes roughly like this: local fluctuations always take place, but due to their random character it is impossible to exploit them in order to obtain macroscopic violations of the second law. If, however, our information on the microstate could be increased, such a process would become feasible unless the entropy of the measuring instrument increased by a corresponding amount. There are several difficulties in this argument. First, fluctuations are difficult to include in classical thermodynamics so one must use statistical mechanics. Secondly, Szilard has to identify the thermodynamic and statistical entropies for a simple system (a one-molecule gas in a box). This identification cannot be general as we have already stated, especially it does not hold for the measuring instrument as will be discussed further in § 8. Finally, in classical statistical mechanics even the statistical entropy is not uniquely defined. In fact the microstate of a classical system may be defined with arbitrary precision, hence an infinite amount of information may be obtained about it, and it is then necessary to consider states with negative entropy. For an ideal gas it is necessary to introduce ad hoc minimal phase space cells of volume  $\hbar^3$ , which define the maximal information that can be obtained about the system. This points to the conclusion that quantum theory is the best framework for a discussion of the relation between statistical entropy and information.

Attempts have been made before to adapt the concept of information to quantum theory in general [8, 9] and the measurement process in particular [10], but as far as we know the fundamental inequalities in § 3 and 4 are new.

## 2. Formalism

The set of observables  $\mathcal{O}$  will be the bounded self-adjoint operators with discrete spectra in a Hilbert space  $\mathcal{H}$ . The set of normal states  $\mathcal{S}$  are then the density operators i.e. positive operators of trace 1. Let  $A$

be an observable with spectrum  $\{a_k\}$  and eigenspaces defined by projectors  $P_k$ . When  $A$  is measured on the system in the state  $W$  by a classical apparatus, the value  $a_k$  is obtained with probability  $p_k = \text{Tr } P_k W$  and after obtaining  $a_k$  the state of the system is  $W'_k = p_k^{-1} P_k W P_k$ .

Averaging over all possible outcomes of the experiment gives the state

$$W' = \sum p_k W'_k = \sum P_k W P_k,$$

i.e. a mixture of states in each of which  $A$  has a definite value. Note that  $W'$  gives the same expectations as  $W$  for all observables commuting with  $A$ , and we can therefore call this a minimum-disturbance measurement [11]. We call the measurement (and the observable  $A$ ) *complete* if the  $P_k$  are one-dimensional.

The map  $T_A : W \rightarrow W'$  extended to  $\mathcal{B}(\mathcal{H})$  is a special case of an expectation in the operator algebra  $\mathcal{B}(\mathcal{H})$  i.e. a linear map from  $\mathcal{B}(\mathcal{H})$  into a von Neumann subalgebra satisfying [12, 13]  $TI = I$  where  $I$  is the identity operator

$$T(A \cdot TB) = TA \cdot TB \quad \text{for all } A, B \in \mathcal{B}(\mathcal{H}).$$

It then follows that  $T$  is idempotent and positive. In this special case we have furthermore that  $T$  is tracepreserving ( $\text{Tr } TA = \text{Tr } A$ ), hence it maps states into states.

The entropy of a state  $W$  is defined as

$$\begin{aligned} S(W) &= -\text{Tr } W \log W && \text{if } W \log W \text{ is of trace class} \\ &= +\infty && \text{otherwise.} \end{aligned}$$

The entropy functional satisfies the following inequalities. Let  $W_i \in \mathcal{S}$ ,  $\lambda_i > 0$ ,  $\sum \lambda_i = 1$  and  $s\{\lambda_i\} = -\sum \lambda_i \log \lambda_i$ . Then

$$S(\sum \lambda_i W_i) \geq \sum \lambda_i S(W_i) \tag{2.1}$$

(concavity) with equality iff all  $W_i$  are equal,

$$S(\sum \lambda_i W_i) \leq \sum \lambda_i S(W_i) + s\{\lambda_i\} \tag{2.2}$$

(subadditivity) with equality iff  $W_i W_j = 0$  for  $i \neq j$ ,

$$S(W) \leq S(T_A W) \leq S(W) + s\{p_k\} \tag{2.3}$$

where the first equality holds iff  $T_A W = W$  and the second iff  $S(W'_k) = S(W)$  for all  $k$ . For proofs see [14–16]. The conditional entropy between two states is given by

$$S(W_1 | W_2) = \text{Tr}(W_1 \log W_1 - W_1 \log W_2).$$

This formal expression will be interpreted in the following way: if  $A, B$  are positive trace class operators with complete orthonormal sets of eigenvectors  $\{|a_i\rangle\}$  and  $\{|b_j\rangle\}$  define

$$\begin{aligned} S(A|B) &= \sum_{i,j} |\langle a_i|b_j\rangle|^2 (a_i \log a_i - a_i \log b_j + b_j - a_i) \\ &= \Sigma \langle a_i|(A \log A - A \log B + B - A)|a_i\rangle \\ &= \Sigma \langle b_j|(A \log A - A \log B + B - A)|b_j\rangle. \end{aligned}$$

From the convexity of  $x \log x$  follows that the terms of the first sum are nonnegative [14]. The sum is then welldefined (eventually infinite) and the order of summation is irrelevant. Furthermore the definition is independent of the choice of eigenvectors when  $A, B$  have degenerate spectra [17]. Consequently the two other expressions are obtained by summation over  $i$  or  $j$ .

In order that  $S(W_1|W_2) < \infty$  it is necessary that  $P_{W_1} \leq P_{W_2}$  where  $P_W = \text{supp } W$  is the support projection of  $W$  (we will write this as  $W_1 < W_2$ ). From the definition follows that

$$S(W_1|W_2) \geq 0 \tag{2.4}$$

with equality iff  $W_1 = W_2$  [15]. If  $\lambda W_1 \leq W_2$  for some  $\lambda \in (0, 1)$  then  $S(W_1|W_2) \leq -\log \lambda$  from the operator monotony of  $\log x$  [18]. If  $W = \Sigma \lambda_i W_i$  then

$$S(W) = \Sigma \lambda_i S(W_i) + \Sigma \lambda_i S(W_i|W)$$

which gives (2.1) and (2.2). If  $T$  is a trace-preserving expectation then  $W < TW$  [8] and

$$S(TW) = S(W) + S(W|TW)$$

i.e.  $T$  is entropy-increasing.

The uncertainty of  $A \in \mathcal{O}$  in the state  $W$  is defined as the entropy of the probability distribution  $\{p_k\}$  obtained in a measurement of  $A$ :

$$H(A, W) = s\{p_k\}.$$

Let  $A, B \in \mathcal{O}$ . We can then define the average uncertainty in  $B$  after  $A$  has been measured

$$H(B|A, W) = \Sigma p_k H(B, W_k).$$

If  $[A, B] = 0$  then

$$H(B|A, W) = H(A \cup B, W) - H(A, W),$$

where  $A \cup B$  is the observable obtained by measuring  $A$  and  $B$  simultaneously. The information between commuting  $A$  and  $B$  is defined by (note the symmetry in  $A$  and  $B$ )

$$\begin{aligned} I(B|A, W) &= H(B, W) - H(B|A, W) \\ &= H(A, W) + H(B, W) - H(A \cup B, W). \end{aligned}$$

(The first expression works also for noncommuting  $A$  and  $B$  but the information is then not necessarily positive.) In case some of the terms in (2.5) are infinite the information is defined as follows: let  $\{p_i\}, \{q_j\}, \{r_{ij}\}$  be the probability distributions of  $A, B,$  and  $A \cup B,$  respectively, then

$$I(B|A, W) = \sum_{i,j} (r_{ij} \log(r_{ij}/p_i q_j) + p_i q_j - r_{ij}).$$

The following inequalities hold

$$0 \leq H(A, W) \leq S(T_A W)$$

where the second equality holds if  $A$  is complete,

$$\begin{aligned} \sum \lambda_i H(A, W_i) &\leq H(A, \sum \lambda_i W_i) \leq \sum \lambda_i H(A, W_i) + s\{\lambda_i\} \\ H(A, W) &\leq H(A \cup B, W) \leq H(A, W) + H(B, W) \\ H(A|B \cup C, W) &\leq H(A|B, W) \\ 0 &\leq I(B|A, W) \leq H(A, W) \\ I(A|B, W) &\leq I(A|B \cup C, W) \end{aligned} \tag{2.5}$$

where  $A, B, C$  is a commuting set of observables. The proofs follow from the concavity and subadditivity properties of the function  $s\{ \}$  as in classical information theory [19].

### 3. A Fundamental Inequality

**Theorem 1.** *Let  $W_1, W_2 \in \mathcal{S}, W_1 < W_2$  and let the expectation  $T: W \rightarrow W'$  correspond to a complete measurement. Then*

$$S(W'_1|W'_2) \leq S(W_1|W_2). \tag{3.1}$$

The proof proceeds via a number of lemmas.

*Definition.* An operator function  $f(x)$  is operator convex in an interval  $I$  if the operator inequality

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda) f(B)$$

holds for  $\lambda \in (0, 1)$  and arbitrary self-adjoint  $A, B \in \mathcal{B}(\mathcal{H})$  with spectra in  $I$  [17].

From Theorem 3.3 and corollaries of [18] follows that  $f(x)$  is operator convex in  $I = (x_0 - R, x_0 + R)$  iff it can be represented there as a Stieltjes integral

$$f(x) = f(x_0) + (x - x_0) f'(x_0) + (x - x_0)^2 \int_{-1/R}^{+1/R} (1 - t(x - x_0))^{-1} d\psi(t)$$

where  $\psi(t)$  is a bounded nondecreasing function, continuous on the left.

**Lemma 1.**  $x^\alpha, \alpha \in [1, 2]$  and  $-\log x$  are operator convex in  $(0, \infty)$ .

*Proof.*  $x^\alpha$  satisfies the Stieltjes integral formula above with  $x_0 = 1, R = 1$  and

$$\begin{aligned} d\psi &= -\pi^{-1} \sin \alpha \pi (1+t)^\alpha (-t)^{1-\alpha} dt & \text{for } t \in (-1, 0) \\ &= 0 & \text{for } t \in (0, 1) \end{aligned}$$

hence it is operator convex in  $(0, 2)$ . Dilatation  $x \rightarrow \lambda x$  gives operator convexity in  $(0, \infty)$ .

Derivation at  $\alpha = 0$  gives for  $f(x) = -\log x$

$$d\psi = -t dt \quad \text{for } t \in (-1, 0), = 0 \quad \text{for } t \in (0, 1)$$

and the same conclusion follows.

Let  $A_0, A_1$  be positive trace class operators,  $A_1 < A_0, \{|k\rangle\}$  an arbitrary complete orthonormal set in  $\mathcal{H}$ . Put  $A_\lambda = \lambda A_1 + (1 - \lambda) A_0, P_k = |k\rangle \langle k|, A' = \sum P_k A P_k$  and  $p_{k,\alpha}(A) = (\langle k | A^\alpha | k \rangle)^{1/\alpha}$ .

**Lemma 2.**  $p_{k,\alpha}(A_\lambda) \leq \lambda p_{k,\alpha}(A_1) + (1 - \lambda) p_{k,\alpha}(A_0)$  for  $\alpha \in [1, 2]$ .

*Proof.*  $p_{k,\alpha}$  convex  $\Leftrightarrow \{A; p_{k,\alpha}(A) \leq 1, A \geq 0\}$  convex  $\Leftrightarrow \{A, p_{k,\alpha}^\alpha(A) \leq 1, A \geq 0\}$  convex  $\Leftrightarrow \langle k | A^\alpha | k \rangle$  convex.

But the convexity of  $\langle k | A^\alpha | k \rangle$  follows from Lemma 1. Note that equality holds for  $\alpha = 1$ .

**Lemma 3.**  $\lambda S(A'_1 | A'_1) + (1 - \lambda) S(A'_0 | A'_0) \leq \lambda S(A_1 | A_\lambda) + (1 - \lambda) S(A_0 | A_\lambda)$ .

*Proof.*  $\frac{d}{d\alpha} p_{k,\alpha}(A)|_{\alpha=1} = \langle k | A \log A | k \rangle - \langle k | A | k \rangle \log \langle k | A | k \rangle = \langle k | A \log A - A' \log A' | k \rangle$ .

From Lemma 2 we obtain by derivation at  $\alpha = 1$ ;

$$\langle k | B_\lambda(A'_0, A'_1) | k \rangle \leq \langle k | B_\lambda(A_0, A_1) | k \rangle$$

where

$$B_\lambda(A_0, A_1) = \lambda A_1 \log A_1 + (1 - \lambda) A_0 \log A_0 - A_\lambda \log A_\lambda.$$

Due to the operator convexity of  $x \log x$  [20]  $B_\lambda(A_0, A_1) \geq 0$ , hence the trace of  $B_\lambda$  is well defined and independent of the basis chosen [21] i.e.

$$\text{Tr } B_\lambda(A'_0, A'_1) \leq \text{Tr } B_\lambda(A_0, A_1)$$

by summation over  $k$ . But

$$\begin{aligned} B_\lambda(A_0, A_1) &= \lambda(A_1 \log A_1 - A_1 \log A_\lambda + A_\lambda - A_1) \\ &\quad + (1 - \lambda)(A_0 \log A_0 - A_0 \log A_\lambda + A_\lambda - A_0). \end{aligned}$$

Calculating the trace in an eigenbasis of  $A_\lambda$  we obtain

$$\text{Tr } B_\lambda(A_0, A_1) = \lambda S(A_1 | A_\lambda) + (1 - \lambda) S(A_0 | A_\lambda) \equiv f(\lambda) \leq s(\lambda, 1 - \lambda).$$

**Lemma 4.**  $\lim_{\lambda \rightarrow 0} \lambda^{-1} f(\lambda) = S(A_1 | A_0)$ .

*Proof.* First we note that

$$\lim_{\lambda \rightarrow 0} S(A_1 | A_\lambda) = S(A_1 | A_0).$$

This follows from the definition of  $S(A_1 | A_\lambda)$  using an eigenbasis of  $A_1$ . In fact every finite partial sum

$$g_n(\lambda) = \sum_{i=1}^n (a_i \log a_i - a_i \langle a_i | \log A_\lambda | a_i \rangle + \langle a_i | A_\lambda | a_i \rangle - a_i)$$

is continuous in  $\lambda = 0$ :  $g_n(0) = \lim_{\lambda \rightarrow 0} g_n(\lambda)$  as  $A_1 < A_0$ .  $g_n$  is convex in  $\lambda$  by the operator concavity of  $\log x$  and  $\{g_n(\lambda)\}$  form a monotonely increasing sequence,  $g_n(\lambda) \rightarrow g(\lambda) = S(A_1 | A_\lambda)$ , hence by a simple argument  $\lim_{\lambda \rightarrow 0} g(\lambda) = g(0)$ . In the same way  $S(A_0 | A_\lambda) \rightarrow S(A_0 | A_0) = 0$  and  $f(\lambda) \rightarrow f(0) = 0$ .

From the convexity of  $x \log x$  follows that  $f(\lambda)$  is concave, hence  $\lambda^{-1} f(\lambda)$  is monotonely increasing when  $\lambda \rightarrow 0$  and  $\lim_{\lambda \rightarrow 0} \lambda^{-1} f(\lambda)$  is uniquely defined. Then  $\lim(\lambda^{-1} - 1) S(A_0 | A_\lambda) \geq 0$  also exists and obviously

$$\lim \lambda^{-1} f(\lambda) \geq S(A_1 | A_0).$$

If  $S(A_1 | A_0) = \infty$  the lemma holds, if not then we can write

$$f(\lambda) = \lambda S(A_1 | A_0) - S(A_\lambda | A_0)$$

using an eigenbasis of  $A_0$  in calculating the trace. Hence

$$\lim \lambda^{-1} f(\lambda) = S(A_1 | A_0) - \lim \lambda^{-1} S(A_\lambda | A_0) \leq S(A_1 | A_0),$$

consequently  $\lim \lambda^{-1} f(\lambda) = S(A_1 | A_0)$ .

*Proof of Theorem 1.* Dividing by  $\lambda$  in Lemma 3 and taking the limit  $\lambda \rightarrow 0$  gives by Lemma 4  $S(A'_1 | A'_0) \leq S(A_1 | A_0)$  which proves the theorem. In analogy with the classical case [8, 22] it is natural to make the following hypothesis.

*Conjecture.* If  $W_1, W_2 \in \mathcal{S}$ ,  $W_1 < W_2$ , and if  $T: W \rightarrow \Sigma P_k W P_k$  corresponds to an arbitrary measurement then

$$S(TW_1 | TW_2) \leq S(W_1 | W_2). \quad (3.2)$$

It turns out that this inequality is equivalent to the property of strong subadditivity for the entropy of a quasilocal system. This property has been conjectured to be valid for quantum systems [2, 17]<sup>1</sup>. A proof of the equivalence will be published elsewhere.

*Remark.* From the proofs of Lemmas 3 and 4 follows that we can define an operator conditional entropy.

*Definition.*  $S_{\text{op}}(A_1 | A_0) = \lim_{\lambda \rightarrow 0} \lambda^{-1} B_\lambda(A_0, A_1)$  where the limit exists if  $S(A_1 | A_0) < \infty$ .

From the operator convexity of  $x \log x$  follows that  $\lambda^{-1} B_\lambda(A_0, A_1)$  is monotonely increasing when  $\lambda \rightarrow 0$ . Furthermore

$$\lambda^{-1} B_\lambda(A_0, A_1) \leq \lambda^{-1} \text{Tr}(B_\lambda(A_0, A_1)) \cdot I \leq S(A_1 | A_0) \cdot I$$

hence the limit exists and  $\lambda^{-1} B_\lambda(A_0, A_1) \rightarrow S_{\text{op}}(A_1 | A_0)$  strongly (Lemma 1 of [23]).

#### 4. Correlation and Information

Consider a system made up of two subsystems such that  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ . From the state  $W$  of the whole system we define the partial states for the subsystems

$$W_1 = \text{Tr}_2 W, \quad W_2 = \text{Tr}_1 W.$$

The corresponding entropies satisfy the inequalities [17]

$$|S(W_1) - S(W_2)| \leq S(W) \leq S(W_1) + S(W_2).$$

The logarithmic correlation between the subsystems is defined as

$$C_{12}(W) = S(W | W_1 \otimes W_2).$$

From (2.4) follows that

$$C_{12}(W) = 0 \quad \text{iff} \quad W = W_1 \otimes W_2. \quad (4.1)$$

<sup>1</sup> After the completion of the manuscript a proof was announced by Lieb, E.H., Ruskai, M. B., in: Phys. Rev. Letters **30**, 434 (1973).



We note that  $W < W_1 \otimes W_2$ . In fact  $P_{W_1 \otimes W_2} = P_{W_1} \otimes P_{W_2}$ , hence

$$\begin{aligned} 0 &\leq \text{Tr}(I - P_{W_1} \otimes P_{W_2})W = -\text{Tr}(I - P_{W_1}) \otimes (I - P_{W_2})W \leq 0 \\ &\Rightarrow (I - P_{W_1 \otimes W_2})W = 0 \Rightarrow (I - P_{W_1 \otimes W_2})P_W = 0 \Rightarrow P_W \leq P_{W_1 \otimes W_2}. \end{aligned}$$

If  $S(W) < \infty$  we can write

$$C_{12}(W) = S(W_1 \otimes W_2) - S(W) = S(W_1) + S(W_2) - S(W).$$

It then follows that

$$\begin{aligned} 0 &\leq C_{12}(W) \leq 2 \min_i S(W_i) \\ S(W_1) = 0 &\Rightarrow S(W) = S(W_2), \quad C_{12}(W) = 0 \quad \text{and} \quad W = W_1 \otimes W_2. \end{aligned} \tag{4.2}$$

Define the information between the two subsystems in the state  $W$  as

$$I_{12}(W) = \sup_{A, B} I(A|B, W), \quad A \in \mathcal{O}(\mathcal{H}_1), \quad B \in \mathcal{O}(\mathcal{H}_2).$$

**Theorem 2.**  $I_{12}(W) \leq C_{12}(W).$  (4.3)

*Proof.* From (2.5) follows that it is sufficient to take the supremum over  $A, B$  complete in  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively. For such  $A, B$

$$I(A|B, W) = S(T_A T_B W | T_A W_1 \otimes T_B W_2) = S(T_A T_B W | T_A T_B W_1 \otimes T_B W_2).$$

As  $T_A T_B$  corresponds to a complete measurement in  $\mathcal{H}$ , it follows from Theorem 1 that

$$S(T_A T_B W | T_A T_B W_1 \otimes T_B W_2) \leq S(W | W_1 \otimes W_2) = C_{12}(W)$$

and the theorem is proved.

In general it is necessary to use observables not of tensor product form to detect the total correlation  $C_{12}(W)$ , which implies an interaction between the subsystems. If we consider the physically realizable observables for separated subsystems to be those of tensor product form, then we can call  $I_{12}$  the measurable correlation which is thus always less than  $C_{12}$ . In the opposite direction we have the following easy result.

**Proposition 1.**  $I_{12}(W) = 0 \Rightarrow C_{12}(W) = 0.$

*Proof.*  $I_{12}(W) = 0 \Rightarrow C_{12}(T_A T_B W) = 0$  for all complete  $A \in \mathcal{O}(\mathcal{H}_1), B \in \mathcal{O}(\mathcal{H}_2)$ . Hence by (4.1)  $T_A T_B W = T_A W_1 \otimes T_B W_2$ . If  $\{|k\rangle\}, \{|l\rangle\}$  are arbitrary complete orthonormal systems in  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively, then the equality reads  $\langle k, l | W | k, l \rangle = \langle k | W_1 | k \rangle \langle l | W_2 | l \rangle$ . By polariza-

tion we obtain  $\langle k, l | W | k', l' \rangle = \langle k | W_1 | k' \rangle \langle l | W_2 | l' \rangle$  and for arbitrary  $\psi \in \mathcal{H}$   $\langle \psi | W | \psi \rangle = \langle \psi | W_1 \otimes W_2 | \psi \rangle$ , hence  $W = W_1 \otimes W_2$  and  $C_{12}(W) = 0$ .

### 5. The Measurement Process

Consider the following model of the measurement process. The system  $S$  as well as the apparatus  $M$  are described as quantum systems with Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  respectively. Before the interaction the two systems are supposed to be uncorrelated i.e.  $W = W_1 \otimes W_2$ . As the total system  $S + M$  is taken to be closed, its time development during the interaction is given by a unitary operator in  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ :

$$W \rightarrow W' = U W_1 \otimes W_2 U^+.$$

We will call this interaction a *perfect measurement* of  $A = \sum a_k P_k \in \mathcal{O}(\mathcal{H}_1)$  if the following two conditions are satisfied:

$$(a) \quad W'_1 = \text{Tr}_2 W' = T_A W_1 = \sum P_k W_1 P_k$$

i.e. the reduced state of  $S$  is that after a measurement of  $A$  by a classical apparatus,

$$(b) \quad \inf_B H(A|B, W') = 0, \quad B \in \mathcal{O}(\mathcal{H}_2) \quad \text{or equivalently} \\ \sup_B I(A|B, W') = H(A, W'_1) = s\{p_k\}$$

i.e. complete information on  $A$  is obtainable by a measurement on  $M$ . Note that the preceding conditions in general only hold for some special choice of  $W_2$ . Apart from this we want to make no more a priori assumptions on  $M$ .

The transformation  $W_1 \rightarrow \sum P_k W_1 P_k$  can be viewed as a loss of the information contained in the nondiagonal terms  $P_k W_1 P_{k'}$ ,  $k \neq k'$ , in  $W_1 = \sum_{k, k'} P_k W_1 P_{k'}$ , or, equivalently, as an averaging over the relative phases between the subspaces  $P_k \mathcal{H}_1$ . In fact [13]

$$\sum P_k W_1 P_k = \mathcal{M}(V W_1 V^+)$$

where  $\mathcal{M}$  denotes the invariant mean over the abelian group of unitary operators in the von Neumann algebra generated by the projectors  $P_k$ . We now consider a more general process where the relative phases are not completely lost: The interaction is called an *imperfect measurement* of  $A$  if the following condition is satisfied

$$(c) \quad W'_1 = W_1 \text{ for all } W_1 \text{ such that } T_A W_1 = W_1, \text{ i.e. every eigenstate of } A \text{ is mapped onto itself.}$$

We will see below that (b) and (c) imply (a), so in the general case (c) we will have only partial information on  $A$  by measurements on  $M$  (obviously (a)  $\Rightarrow$  (c)).

We want to show that (c) implies that  $W'_2$  is independent of the relative phases present in  $W_1$ .

**Lemma 5.** *If (c) is satisfied, then  $W'_2 \equiv \text{Tr}_1(U W_1 \otimes W_2 U^+) = \text{Tr}(U T_A W_1 \otimes W_2 U^+) = \text{Tr}(U V W_1 V^+ \otimes W_2 U^+)$  for all unitary  $V$  generated by  $\{P_k\}$ .*

*Proof.* Put  $W'_{kk'} = U P_k W_1 P_{k'} \otimes W_2 U^+$ . Then, by (c)  $\text{Tr}_2 W'_{kk} = P_k W_1 P_k$ , hence  $\text{Tr}(P_{k'} \otimes I \cdot W'_{kk}) = 0$  for  $k \neq k'$  i.e.  $P_{k'} \otimes I W'_{kk} = 0$  for  $k \neq k'$ . But

$$W'_{kk'} W'_{k'k} = U P_k W_1 P_{k'} W_1 P_k \otimes W_2^2 U^+ \leq U P_k W_1 P_{k'} W_1 P_k \otimes W_2 U^+ \leq W'_{kk}$$

where we have used that  $W^2 \leq W$  for  $W \in S$  and that  $W P W \leq W^2$  for any projection  $P$ . Hence

$$P_{k''} \otimes I W'_{kk'} = 0 \quad \text{for } k \neq k''$$

and

$$\begin{aligned} T_A W' &= \sum P_k \otimes I W' P_k \otimes I = \sum W'_{kk} = U T_A W_1 \otimes W_2 U^+ \\ W'_2 &= \text{Tr}_1 W' = \text{Tr}_1 T_A W' = \text{Tr}_1 U T_A W_1 \otimes W_2 U^+ . \end{aligned}$$

If  $V$  is generated by  $\{P_k\}$  then  $T_A V W_1 V^+ = T_A W_1$  and the lemma follows.

*Remarks.* 1)  $T_A W'$  is the state obtained when  $A$  is measured by a classical apparatus  $M_1$  after the interaction of  $S$  and  $M$ . The partial state  $W'_2$  is unchanged by the interaction with  $M_1$ . The lemma says that it is impossible to get information on the relative phases in  $W_1$  by making a measurement on  $M$  and simultaneously making a measurement of  $A$  on  $S$ .

2) The lemma implies that in order to calculate the state of  $M$  after the interaction (and hence the entropy) we can replace  $W_1$  by  $T_A W_1$ .

**Lemma 6.** *Let  $W_1$  be an eigenstate of  $A$  with eigenvalue  $a_k$  i.e.  $P_k W_1 = W_1$ . Then (c) implies that  $W' = W_1 \otimes W'_{2k}$  where  $W'_{2k}$  depends only on  $k$ .*

*Proof.* In the proof of Lemma 5, replace  $\{P_k\}$  by any complete set of projectors  $\{P_{kr}\}$  in  $P_k \mathcal{H}$ ,  $k$  fixed, and let  $V$  be any unitary operator in  $P_k \mathcal{H}$ . As the set of pure states is transitive under this group of unitaries and  $W'_2 = \text{Tr}(U V W_1 V^+ \otimes W_2 U^+)$ ,  $W'_2$  is the same for all pure states and thus by superposition for all states in  $P_k \mathcal{H}$ . Consequently  $W'_2 = W'_{2k}$  is a function only of  $k$ . If  $W_1$  is a pure state we must have by (4.2)  $W' = W_1 \otimes W'_{2k}$  and by superposition this holds for all  $W_1$  belonging to the eigenvalue  $a_k$ .

**Corollary.**  $T_A W' = \sum P_k W_1 P_k \otimes W'_{2k}$  for arbitrary  $W_1$ .

**Lemma 7.** If the  $W'_{2k}$  are orthogonal then  $W'_1 = \sum P_k W_1 P_k$  i.e. the measurement satisfies condition (a).

*Proof.* If  $Q_k = \text{supp } W'_{2k}$  then  $Q_k \cdot Q_{k'} = 0$  for  $k \neq k'$ . From the proof of Lemma 5 and from Lemma 6

$$W'_{kk'} W'_{k'k} \leq W'_{kk} = P_k W_1 P_k \otimes W'_{2k},$$

hence  $I \otimes Q_{k'} \cdot W'_{kk'} = 0$  for  $k \neq k'$  and

$$W'_1 = \text{Tr}_2 W' = \text{Tr}_2 \sum I \otimes Q_k W' I \otimes Q_k = \text{Tr}_2 \sum W'_{kk} = \sum P_k W_1 P_k.$$

## 6. Entropy Increase in the Apparatus

Define the information content in  $M$  on the observable  $A$  in the final state  $W'$  for  $S + M$  as

$$I_2(A, W') = \sup_B I(A|B, W'), \quad B \in \mathcal{O}(\mathcal{H}_2).$$

As the probability distributions of  $A$  and  $B$  are invariant under  $W' \rightarrow T_A W'$  we have  $I(A|B, W') = I(A|B, T_A W')$ , hence

$$I_2(A, W') = I_2(A, T_A W') \leq I_{12}(T_A W').$$

Assume that the entropies appearing below are finite.

**Proposition 2.** If condition (c) is satisfied then

$$I_2(A, W') \leq S(W'_2) - S(W_2).$$

*Proof.* By the remarks after Lemma 5 we can replace the initial state  $W_1$  of  $S$  by  $T_A W_1$ . Then the final state of  $S$  is  $T_A W_1$  by condition (c) and the final state of  $S + M$  is  $T_A W'$  by Lemma 6. The unitary invariance of the entropy means that

$$S(T_A W_1) + S(W_2) = S(T_A W_1) + S(W'_2) - C_{12}(T_A W')$$

i.e.  $S(W'_2) - S(W_2) = C_{12}(T_A W')$ .

But from Theorem 2

$$I_2(A, W') \leq I_{12}(T_A W') \leq C_{12}(T_A W')$$

hence the statement is proved.

**Proposition 3.** If conditions (b) and (c) are satisfied then

$$S(W'_2) - S(W_2) = I_2(A, W') = s\{p_k\}$$

and furthermore condition (a) is satisfied.

*Proof.* Put  $T_A W_1 = \sum p_k W'_{1k}$ , then it follows from Lemma 6 that

$$\begin{aligned} S(W'_2) - S(W_2) &= C_{12}(T_A W') = S(\sum p_k W'_{1k}) + S(W'_2) - S(\sum p_k W'_{1k} \otimes W'_2) \\ &= S(W'_2) - \sum p_k S(W'_{2k}) = \sum p_k S(W'_{2k} | W'_2). \end{aligned}$$

For complete  $B \in \mathcal{O}(\mathcal{H}_2)$

$$\begin{aligned} I(A|B, W') &= H(B, W'_2) - \sum p_k H(B, W'_{2k}) = S(T_B W'_2) - \sum p_k S(T_B W'_{2k}) \\ &= \sum p_k S(T_B W'_{2k} | T_B W'_2). \end{aligned}$$

From Theorem 1 we obtain (this is an alternative proof of Proposition 2)

$$I(A|B, W') \leq S(W'_2) - S(W_2).$$

If (b) is satisfied then  $S(W'_2) - S(W_2) \geq s\{p_k\}$ . But according to (2.2)

$$S(W'_2) - S(W_2) = S(W'_2) - \sum p_k S(W'_{2k}) \leq s\{p_k\}$$

with equality iff  $\{W'_{2k}\}$  are mutually orthogonal. Hence the equality holds and from Lemma 7 follows that condition (a) is satisfied.

*Remark.* Alternatively one can use Lemma 6 and the unitary invariance of entropy to show that  $S(W'_{2k}) = S(W_2)$ . Hence  $S(W_2) = \sum p_k S(W'_{2k})$  and  $S(W'_2) - S(W_2) = \sum p_k S(W'_{2k} | W_2)$ .

It is natural to ask whether conversely condition (a) implies (b). This is not true in general. Let  $W_2 = \sum \lambda_r |\varphi_r\rangle \langle \varphi_r|$  where  $\{|\varphi_r\rangle\}$  are orthonormal. From Lemma 6 follows that  $W'_{2k} = \sum \lambda_r |\varphi_{kr}\rangle \langle \varphi_{kr}|$  where  $\langle \varphi_{kr} | \varphi_{k'r'} \rangle = 0$  for  $r \neq r'$  and all  $k, k'$ . A simple calculation gives

$$\begin{aligned} W' &= \sum \lambda_r P_k W_1 P_k \otimes |\varphi_{kr}\rangle \langle \varphi_{k'r'}|, \\ W'_1 &= \sum \lambda_r \langle \varphi_{k'r'} | \varphi_{kr} \rangle P_k W_1 P_{k'}. \end{aligned} \tag{6.1}$$

The condition  $W'_1 = T_A W_1$  for all  $W_1$  is equivalent to

$$\sum \lambda_r \langle \varphi_{k'r'} | \varphi_{kr} \rangle = \delta_{kk'}.$$

This relation can be fulfilled without necessarily

$$\text{Tr } W'_{2k} W'_{2k'} = \sum \lambda_r^2 |\langle \varphi_{k'r'} | \varphi_{kr} \rangle|^2 = 0 \quad \text{for } k \neq k'$$

i.e.  $\langle \varphi_{k'r'} | \varphi_{kr} \rangle = 0$  for  $k \neq k'$  and all  $r$ .

This is easily seen e.g. when the number of different eigenvalues of  $A$  is finite. Consequently (b) need not be satisfied.

## 7. Entropy Changes in the Object System

We have already stated in (2.3) that for a perfect measurement the entropy of  $S$  is nondecreasing

$$S(W_1) \leq S(T_A W_1). \tag{7.1}$$

Furthermore, using the information contained in  $M$  about the state  $S$  we can reduce the average entropy of  $S$  to  $\sum p_k S(W'_{1k})$ . This is done by choosing for  $B$  an operator in  $\mathcal{O}(\mathcal{H}_2)$  with the eigenprojectors  $Q_k$  of Lemma 7. Then from Lemmas 6 and 7

$$T_B W' = \sum I \otimes Q_k W' I \otimes Q_k = \sum p_k W'_{1k} \otimes W'_{2k},$$

hence the state of  $S$  after measuring  $B$  and obtaining an eigenvalue  $b_k$  (with probability  $p_k$ ) is  $W'_{1k}$  and the average entropy of  $S$  is  $\sum p_k S(W'_{1k})$ . The average entropy decrease in  $S$  due to the measurement of  $B$  is then  $S(W_1) - \sum p_k S(W'_{1k})$ , which is in fact positive by (2.3). From (2.3) it also follows that

$$S(W_1) - \sum p_k S(W'_{1k}) \leq s\{p_k\} = I_2(A). \tag{7.2}$$

Equality holds only if  $W_1 = W'_1$ , hence the average entropy decrease in  $S$  is in general strictly less than the entropy increase in  $M$ , in spite of the fact that the process is reversible. The origin of this is of course the relative phases present in  $W_1$  and  $W'$  which are lost in the partial states  $W'_1$  and  $W'_2$ . In other words,  $W'$  contains an extra correlation

$$C_{12}(W') - C_{12}(T_A W') = S(T_A W') - S(W') = S(T_A W_1) - S(W_1)$$

that does not contribute to the information when measuring pairs of observables  $A, B$  with the given choice of  $A$ . This extra correlation is of course the cause of the Einstein-Podolsky-Rosen “paradox” and is thus a typical quantum effect.

The entropy increase property (7.1) is also easily proved for imperfect measurements.

**Proposition 4.** *If property (c) holds then  $S(W'_1) \geq S(W_1)$ .*

*Proof* (Compare [24]). From (6.1) follows

$$\begin{aligned} \text{Tr } W'_1 &= \text{Tr } W_1, \\ \text{Tr } W'^2_1 &= \sum_{k,k'} \left| \sum_r \lambda_r \langle \varphi_{k'r} | \varphi_{kr} \rangle \right|^2 \text{Tr } P_k W_1 P_{k'} W_1 \\ &\leq \sum_{k,k'} \text{Tr } P_k W_1 P_{k'} W_1 = \text{Tr } W^2_1. \end{aligned}$$

Hence the map  $T: W_1 \rightarrow W'_1$  satisfies  $\|T W_1\|_1 = \|W_1\|_1$  and  $\|T W_1\|_2 \leq \|W_1\|_2$ , where  $\|W\|_p = (\text{Tr } W^p)^{1/p}$ .

From an interpolation theorem due to Kunze [25] follows that

$$\|T W_1\|_p \leq \|W_1\|_p \quad \text{for } p \in [1, 2].$$

Hence  $\text{Tr}(T W_1)^p \leq \text{Tr } W^p_1$  for  $p \in [1, 2]$  with equality for  $p = 1$ . Differentiation at  $p = 1$  gives  $S(T W_1) \geq S(W_1)$ .

The analogue of (7.2) for imperfect measurements is probably true, but a proof seems to depend on the conjecture (3.2). From conservation of total entropy and the proof of Proposition 2 we have

$$S(W'_1) - S(W_1) = C_{12}(W') + S(W_2) - S(W'_2) = C_{12}(W') - C_{12}(T_A W').$$

Let  $B$  be complete in  $\mathcal{O}(\mathcal{H}_2)$  with eigenprojectors  $Q_i$  and put

$$q_i W'_{1i} = \text{Tr}_2(I \otimes Q_i W').$$

Then simple calculations give

$$\begin{aligned} S(W'_1) - \sum q_i S(W'_{1i}) &= C_{12}(T_B W'), \\ I(A|B, W') &= C_{12}(T_A T_B W'). \end{aligned}$$

The average entropy of  $S$  when  $B$  has been measured on  $M$  is  $\sum q_i S(W'_{1i})$  and the average entropy decrease in  $S$  achieved through this measurement is

$$S(W_1) - \sum q_i S(W'_{1i}) = C_{12}(T_A W') + C_{12}(T_B W') - C_{12}(W').$$

A simple rearrangement gives that the r.h.s. is equal to

$$I(A|B, W') - S(W'|T_A W') + S(T_B W'|T_B T_A W').$$

If the conjecture (3.2) were proved we could conclude that  $S(T_B W'|T_B T_A W') \leq S(W'|T_A W')$  and hence that

$$S(W_1) - \sum q_i S(W'_{1i}) \leq I(A|B, W').$$

## 8. Discussion

a) The following objection has been made to the scheme given in § 5 [26, 27]. One can prove that the condition (c) can be satisfied only if  $A$  commutes with all additive conserved observables of  $S + M$ . Thus if  $S$  is a spin- $\frac{1}{2}$  system,  $A = \sigma_z$ , and if the total angular momentum of  $S + M$  commutes with  $U$ , then condition (c) cannot hold. By making  $M$  large in the sense that the expectation of  $J^2$  is large, it is nevertheless possible to realize the condition in an approximate manner [28].

This objection, however, does not make the results above irrelevant for the following reason. If we consider a Stern-Gerlach experiment which measures  $\sigma_z$ , then we are free to consider the static magnetic field as an external potential, i.e. as an additional term in the Hamiltonian of  $S + M$ . Hence it is not included in the apparatus  $M$  which then consists only of the counters detecting the particles. In this way the conservation

law is broken, but the time development of  $S + M$  is still unitary, which is all that is needed in § 5.

b) We note that the conditions in § 5 imply an asymmetry between  $S$  and  $M$  and a special form of the final state  $W'$ . On the other hand the definition of information is symmetric, so one may be tempted to obtain a symmetric relation between the entropies of  $S$  and  $M$ . Consider therefore any interaction of  $S$  and  $M$ , and any  $B \in \mathcal{O}(\mathcal{H}_2)$ . Then

$$\begin{aligned} S(W_1) - S(W'_1) &= S(W'_2) - S(W_2) - C_{12}(W'), \\ S(W_1) - \sum q_l S(W'_{1l}) &\leq C_{12}(T_B W') \end{aligned}$$

where  $q_l, W'_{1l}$  are defined as in § 7. Hence

$$\begin{aligned} S(W_1) - \sum q_l S(W'_{1l}) &\leq S(W_1) - S(W'_1) + C_{12}(T_B W') \\ &= C_{12}(T_B W') - C_{12}(W') + S(W'_2) - S(W_2). \end{aligned}$$

If the conjecture (3.2) is correct, then we can conclude that

$$S(W_1) - \sum q_l S(W'_{1l}) \leq S(W'_2) - S(W_2)$$

as well as the symmetric relation for any  $A \in \mathcal{O}(\mathcal{H}_1)$ :

$$S(W_2) - \sum p_k S(W'_{2k}) \leq S(W'_1) - S(W_1).$$

c) From the results of the preceding paragraphs we may conclude that the relation between statistical entropy and information for quantum systems is satisfactory and would be even better if the conjecture (3.2) could be proved. The obvious question, however, is whether these results have anything to do with the second law of thermodynamics. Obviously this question applies equally to the classical discussions of Maxwell's demon mentioned in the introduction.

The definition of entropy used above and the description of the measurement process do not differentiate between microscopic and macroscopic quantities. But the thermodynamic entropy  $S_{th}$  should be defined as a function of a few parameters of a macroscopic nature like temperature, density etc. Hence if a state  $W$  is a superposition of states that correspond to different values of some macroscopic variable (like the state of a macroscopic instrument  $M$  after a measurement), an observation that determines the macrostate (reading the instrument  $M$ ) cannot decrease the thermodynamic entropy. As the average entropy of  $M$  after reading it is  $\sum p_k S(W'_{2k}) = S(W_2)$ , the entropy increase  $S(W'_2) - S(W_2)$  cannot be given a thermodynamic interpretation. Furthermore, if we measure a macroscopic quantity on  $S$  an increase in  $S_{th}$  for  $M$  is not necessary in order to satisfy the entropy nondecrease property for  $S + M$ . Consequently any relation between information and  $S_{th}$  cannot be



general, but requires some way of defining the macroscopic variables of the system. In this connection the point of view that the amplification process from the microscopic to the macroscopic level must necessarily be an irreversible process may be relevant [29]. Then the initial state of  $M$  is at most a quasiequilibrium state, which then develops in some ergodic manner towards an equilibrium final state. In fact if  $M$  is initially in an equilibrium state, the measurement process could be considered as a microscopic fluctuation (the state of  $S$ ) growing spontaneously to macroscopic dimensions (the final state of  $M$ ) which would be contrary to the entropy law. It seems natural to conjecture that this amplification of information will always imply an increase of  $S_{t,\mu}$  for  $M$  not less than the information gain.

*Acknowledgements.* I want to thank Professor M. Flato for stimulating my interest in the quantum measurement problem and Professor B. Nagel for many fruitful discussions.

### References

1. Neumann, J. von: Mathematical foundations of quantum mechanics. Princeton University Press 1965
2. Lanford, O. E., Robinson, D. W.: J. Math. Phys. **9**, 1120—1125 (1968)
3. Szilard, L.: Z. Phys. **53**, 840—856 (1929)
4. Brillouin, L.: Science and information theory. New York: Academic Press 1962
5. Penrose, O.: Foundations of statistical mechanics. Oxford: Pergamon Press 1970
6. Jaynes, E. T.: Phys. Rev. **106**, 620—630 (1957) and **108**, 171—190 (1957)
7. Jauch, J. M., Baron, J. G.: Helv. Phys. Acta **45**, 220—232 (1972)
8. Umegaki, H.: Kodai Math. Sem. Rep. **14**, 59—85 (1962)
9. Ingarden, R. S., Kossakowski, A.: Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom et Phys. **16**, 61—65 (1968)
10. Majernik, V.: Acta Phys. Austriaca **31**, 271—299 (1970) and **34**, 83—92 (1972)
11. Lüders, G.: Ann. Physik, **8**, 322 (1951)
12. Umegaki, H.: Tôhoku Math. J. **6**, 177—181 (1954)
13. Arveson, W. B.: Am. J. Math. **89**, 578—672 (1967)
14. Ruelle, D.: Statistical mechanics. New York: Benjamin 1969
15. Falk, H.: Am. J. Phys. **38**, 858—869 (1970)
16. Lindblad, G.: Commun. math. Phys. **28**, 245—249 (1972)
17. Araki, H., Lieb, E. H.: Commun. math. Phys. **18**, 160—170 (1970)
18. Bendat, J., Sherman, S.: Trans. Am. Math. Soc. **79**, 58—71 (1955)
19. Khinchin, A. I.: Mathematical foundations of information theory. New York: Dover 1957
20. Nakamura, M., Umegaki, H.: Proc. Japan. Acad. **37**, 149—154 (1961)
21. Schatten, R.: Norm ideals of completely continuous operators. Berlin Göttingen Heidelberg: Springer 1960
22. Kullback, S.: Information theory and statistics. New York: Wiley 1959
23. Topping, D. M.: Lectures on von Neumann algebras. London: Van Nostrand Reinhold 1971
24. Segal, I. E.: J. Math. Mech. **9**, 623—629 (1960)
25. Kunze, R. A.: Trans. Amer. Math. Soc. **89**, 519—540 (1958)

26. Wigner, E. P.: Z. Phys. **133**, 101 (1952)
27. Araki, H., Yanase, M. M.: Phys. Rev. **120**, 622—626 (1960)
28. Yanase, M. M.: Phys. Rev. **123**, 666—668 (1961)
29. Daneri, A., Loinger, A., Prospero, G. M.: Nucl. Phys. **33**, 297—319 (1962)

G. Lindblad  
Department of Theoretical Physics  
Royal Institute of Technology  
S-10044 Stockholm 70, Sweden