# Strenghtened Lindblad inequality: Applications in nonequilibrium thermodynamics and quantum information theory

A. E. Allahverdyan and D. B. Saakian
*Yerevan Physics Institute, Alikhanian Brothers Street 2, Yerevan 375036, Armenia*
(Received 28 August 1997; revised manuscript received 10 March 1998)

A strengthened Lindblad inequality has been proved. We have applied this result to prove a generalized $H$ theorem in nonequilibrium thermodynamics. Information processing also can be considered as some thermodynamic process. From this point of view we have proved a strengthened data processing inequality in quantum information theory. [S1063-651X(98)05307-0]

PACS number(s): 05.30.−d, 03.65.Bz

There are close connections between statistical thermodynamics and information theory [1,2]. Physical ideas have played an important role as sources of information theory [1]. On the other hand, the concept of information is crucial for understanding some important physical problems such as Maxwell's ''demon'' [2] or the general problem of quantum correlation between two subsystems [3].

In this paper we concentrate on the two connected problems. These are the $H$-theorem problem in nonequilibrium quantum statistical thermodynamics and the problem of quantum data processing in quantum information theory. The concepts of entropy (or other entropy-like measures) and $H$ theorem are particularly important in *quantum* statistical physics because a correct definition is only possible in the framework of quantum mechanics. In classical theory entropy can only be introduced in a somewhat limited and artificial manner [4,5]. Suppose that a quantum system is described by a density matrix $\rho(t)$ at the moment $t$. In the general case evolution of the nonequilibrium system in the *Markovian* regime is described by some general quantum evolution operator

$$\hat{K}(t',t)\rho(t)=\rho(t'). \tag{1}$$

In the most general case $\hat{K}$ must be linear, completely positive, and trace preserving [6,7], and has standard representation

$$\hat{K}\rho=\sum_{\mu} A_{\mu}^{\dagger}\rho A_{\mu}, \quad \sum_{\mu} A_{\mu}A_{\mu}^{\dagger}=\hat{1}, \tag{2}$$

which was introduced in [7]. This representation is equivalent to the so-called unitary representation where the nonunitary evolution of the system is regarded as a part of the unitary evolution of some larger system. Equation (2) contains unitary transformations, nonselective measurements, partial traces, etc. For a non-Markovian case $\hat{K}$ also depends on the ''history'' from some initial time $t_0$ to $t'$ [8].

If the evolution of the system is in the *stationary Markovian* regime then

$$\hat{K}(t',t)=\hat{K}(t'-t). \tag{3}$$

Stationary Markovian regime is a reasonable conjecture if the system is not far from equilibrium [9,8], or it can be described by a non-Hermitian time-independent Hamiltonian [8] or by quantum Langevin equations [10].

One of the most important quantities which can be defined for statistical systems is entropy [9,8,1,2,4,5]. This quantity was introduced in quantum statistical physics by von Neumann

$$S(\rho)=-\mathrm{tr}\rho\ln\rho. \tag{4}$$

The concept of entropy has at least three main ingredients.

In the first, entropy of a macroscopic statistical system can only increase if the system tends to equilibrium. Therefore entropy is maximal at this state. This is the well known $H$ theorem. We want to stress that increasing of Eq. (4) with time can be violated if the evolution of the system is not stationary or Markovian [9,8] or the system is open or mesoscopic [11]. For example, the entropy of von Neumann can exhibit exactly periodic behavior for some open systems [12]. This means that Eq. (4) is not the relevant statistical function for such systems. Entropy is an additive function, and also invariant of a unitary transformation.

In the second, entropy can be regarded as a measure of the lack of information about a system. Therefore $S(\rho)$ should increase after a coarse-graining procedure [1,8,5]

$$S\left(\sum_i p_i\rho_i\right)\geqslant\sum_i p_iS(\rho_i), \quad \sum_i p_i=1, p_i\geqslant 0 \tag{5}$$

where the information about coordinate $i$ is lost ($i$ can also be continuous).

In the third, entropy can be considered as a measure of the amount of chaos, or, to what extent the density matrix $\rho$ can be considered as ''mixed.'' Indeed, the non-negative $S(\rho)$ is zero for a pure density matrix, and is maximal for homogeneous $\rho$. Entropy also is an additive quantity. Equation (4) can also be viewed as one of the basic statements of equilibrium statistical physics [9,8,4]. For example, after several assumptions the most important relation in thermostatics: $TdS=dE+pdV$, can be derived from Eq. (4) (where all symbols have their ordinary meaning).

Now the following questions arise. Is it possible to define an entropylike function for a mesoscopic statistical system or for an open system? Is it possible to save in this definition the main aspects of usual entropy? A large number of papers and books is devoted to these questions [9,8,13]. The answer is ''yes'' at least in the case when the evolution of a system is stationary Markovian, and has a well defined stationary distribution. The concrete form of this distribution is not important. Before the definition we need some mathematics.

Quantum relative entropy between two density matrices $\rho_1$, $\rho_2$ is defined as

$$S(\rho_1||\rho_2)=\mathrm{tr}(\rho_1\ln\rho_1-\rho_1\ln\rho_2). \quad (6)$$

This positive quantity was introduced by Umegaki [14] and characterizes the degree of ''closeness'' of density matrices $\rho_1$, $\rho_2$. The properties of quantum relative information were reviewed by Ohya [15]. Here only two basic properties are mentioned

$$S(\rho_1||\rho_2)\geqslant S(\hat{K}\rho_1||\hat{K}\rho_2), \quad (7)$$

$$S(\lambda\rho_1+(1-\lambda)\rho_2||\lambda\sigma_1+(1-\lambda)\sigma_2)$$
$$\leqslant\lambda S(\rho_1||\sigma_1)+(1-\lambda)S(\rho_2||\sigma_2), \quad (8)$$

where $0\leqslant\lambda\leqslant 1$. The first inequality was proved by Lindblad [16].

Now for a system with stationary distribution $\rho_{\mathrm{st}}$, and Markovian stationary evolution operator $\hat{K}$ the following function is defined:

$$-S(\rho(t)||\rho_{\mathrm{st}}). \quad (9)$$

This function is additive, and also increases after coarse-graining procedure as we see from Eq. (8). Further, Eq. (7), which can be written as

$$-S(\hat{K}\rho(t)||\rho_{\mathrm{st}})\geqslant -S(\rho(t)||\rho_{\mathrm{st}}), \quad (10)$$

is the $H$ theorem for Eq. (9).

The definition (9) is closely related to the functions which are used in usual equilibrium statistical physics. A very large closed statistical system can be described by microcanonical distribution where $\rho_{\mathrm{st}}$ can be represented as a unit matrix (up to some unessential factors). In this case Eq. (9) reduces to Eq. (4) (at least in the case of finite dimensional Hilbert space), and from Eq. (10) we have the usual $H$ theorem. Further, it is well known that for a closed macroscopic system canonical and microcanonical distributions are equivalent (except some special cases like second-order phase transitions). But in some sense canonical distribution has a larger area of application because it can describe some mesoscopic or quasiopen systems [8,4]. If we take $\rho_{\mathrm{st}}=\exp(-\beta H)/Z$ in Eq. (9) (where $\beta$ is inverse temperature, and $H$ is Hamiltonian) then

$$S(\rho(t)||\rho_{\mathrm{st}})-\ln Z=\mathrm{tr}(\rho\ln\rho)+\beta\,\mathrm{tr}(H\rho)=\beta F, \quad (11)$$

where $F$ is the usual free energy. Therefore for the case of canonical distribution we have a slightly different form of $H$ theorem: the free energy can only decrease if the system tends to equilibrium [9,8,4,13].

Is the physically relevant generalization of von Neumann entropy defined uniquely? This important question was investigated in [17]. The author showed that Eqs. (7) and (8) with some other mathematical conditions are sufficient for the determination of Eq. (6).

The conclusion is the following: Eq. (9) *is a correct generalization of von Neumann entropy to the more general case, and the generalized H theorem can be proved with the assumptions about the evolution of the system only.*

Now the following question arises. Can we generalize Eq. (7) without any restrictions? If the answer is yes, then we can prove with this result a more general relation. Let us assume in formula (7) that

$$\hat{K}=c\hat{C}_1+(1-c)\hat{C}_2, \quad (12)$$

where $\hat{C}_1$ is defined by Kraussian representation $A_\mu=|\mu\rangle\langle 0|$, $\langle\mu|\acute{\mu}\rangle=\delta_{\mu\mu'}$, $\langle 0|0\rangle=1$, $0\leqslant c\leqslant 1$. In other words, for any operator $\rho$, $\hat{C}_1\rho=|0\rangle\langle 0|$. Now from Eqs. (7) and (8) we get

$$S(\hat{K}\rho||\hat{K}\sigma)=S(c\hat{C}_1\rho+(1-c)\hat{C}_2\rho||c\hat{C}_1\sigma+(1-c)\hat{C}_2\sigma)$$
$$\leqslant cS(\hat{C}_1\rho||\hat{C}_1\sigma)+(1-c)S(\hat{C}_2\rho||\hat{C}_2\sigma)$$
$$\leqslant(1-c)S(\rho||\sigma). \quad (13)$$

We see that if $\hat{K}$ is represented in the form (12) the ordinary Lindblad inequality can be strengthened.

Now we need some general results from the theory of linear operators [18]. Let two Hermitian operators $A$ and $B$ have the spectra $a_1\leqslant\cdots\leqslant a_n$, $b_1\leqslant\cdots\leqslant b_n$. For the spectrum $c_1\leqslant\cdots\leqslant c_n$ of the operator $C=A+B$ we have

$$a_1+b_k\leqslant c_k\leqslant b_k+a_n, \quad b_1+a_k\leqslant c_k\leqslant a_k+b_n, \quad (14)$$

where $k=1,\ldots,n$. If

$$\rho'=\hat{K}\rho=c\hat{C}_1\rho+(1-c)\hat{C}_2\rho=c|0\rangle\langle 0|+(1-c)\sigma, \quad (15)$$

and $\rho'_1\leqslant\cdots\leqslant\rho'_n$, $\sigma_1\leqslant\cdots\leqslant\sigma_n$ are the spectra of $\rho'$, $\sigma$ then we have

$$\rho'_1-c\leqslant\sigma_1(1-c)\leqslant\min(\rho'_1,\rho'_n-c),$$
$$\max(\rho'_1,\rho'_k-c)\leqslant\sigma_k(1-c)\leqslant\rho'_k, \quad (16)$$

where $k=2,\ldots,n$. We define $c(\hat{K},\rho)$ as the minimal eigenvalue of $\rho'$ and $c(\hat{K})=\min_\rho c(\hat{K},\rho)$ where minimization is taken by all density matrices for the fixed Hilbert space. With the well known results of operator theory [18] we can write

$$c(\hat{K})=\min_{\rho}\min_{\langle\psi|\psi\rangle=1}\langle\psi|\hat{K}\rho|\psi\rangle, \quad (17)$$

where the second minimization is taken by all normal vectors in the Hilbert space. For any density matrix $\rho$ we get to the formula (12) where $c$ is defined in Eq. (17) and $\hat{C}_2$ is some general evolution operator. Now from Eqs. (12), (13), and (17) we get the strengthened Lindblad inequality

$$(1-c)S(\rho_1||\rho_2)\geqslant S(\hat{K}\rho_1||\hat{K}\rho_2). \quad (18)$$

Equations (17) and (18) are our general results. Of course there are many evolution operators $\hat{K}$ with $c(\hat{K})=0$ but later we shall show that our results can be nontrivial because for some simple but physically important case $c(\hat{K})$ is nonzero. From Eqs. (17) and (18) we immediately get to the strengthened $H$ theorem which gives us some information about the speed of relative entropy decrease. An analog of Eq. (18) exists also in classical information theory [19]. Equation (18) can also be regarded as a bound for entropy production. This quantity is very important in nonequilibrium statistical mechanics [9,8].

We now discuss application of this result to quantum data processing.

Quantum information theory is a new field with potential applications for the conceptual foundation of quantum mechanics. It appears to be the basis for a proper understanding of the emerging fields of quantum computation, communication, and cryptography (see [6] for references). Quantum information theory is concerned with quantum bits (qubits) rather than bits. Qubits can exist in superposition or entanglement states with other qubits, a notion completely inaccessible for classical mechanics. More generally, quantum information theory contains two distinct types of problem. The first type describes transmission of classical information through a quantum channel (the channel can be noisy or noiseless). In such a scheme bits are encoded as some quantum states and only these states or their tensor products are transmitted. In the second case arbitrary superposition of these states or entanglement states is transmitted. In the first case the problems can be solved by methods of classical information theory, but in the second case new physical representations are needed.

Mutual information is the most important ingredient of information theory. In classical theory this quantity was introduced by Shannon [19]. The mutual information between two ensembles of random variables $X$, $Y$ (for example, these ensembles can be input and output for a noisy channel),

$$I(X,Y)=H(Y)-H(Y/X), \tag{19}$$

is the decrease of the entropy of $X$ due to the knowledge about $Y$, and conversely with interchanging $X$ and $Y$. Here $H(Y)$ and $H(Y/X)$ are Shannon entropy and mutual entropy [19]. Mutual information in the quantum case must take into account the specific character of the quantum information as it is described above. The first reasonable definition of this quantity was introduced by Lloyd [20], and independently by Schumacher and Nielsen [21]. Suppose a quantum system with density matrix

$$\rho=\sum_i p_i|\psi_i\rangle\langle\psi_i|, \quad \sum_i p_i=1. \tag{20}$$

We only assume that $\langle\psi_i|\psi_i\rangle=1$ and the states may be nonorthogonal. The noisy quantum channel can be described by some general quantum evolution operator $\hat{K}$.

As follows from the definition of quantum information transmission, a possible distortion of entanglement of $\rho$ must be taken into account. In other words, a definition of mutual quantum information must contain the possible distortion of the relative phases of the quantum ensemble $\{|\psi_i\rangle\}$. Mutual quantum information is defined as [20,21]

$$I(\rho;\hat{K})=S(\hat{K}\rho)-S(\hat{1}^R\otimes\hat{K}(|\psi^R\rangle\langle\psi^R|)), \tag{21}$$

$$\hat{1}^R\otimes\hat{K}(|\psi^R\rangle\langle\psi^R|)=\sum_{i,j}\sqrt{p_ip_j}|\phi_i^R\rangle\langle\phi_j^R|\otimes\hat{K}(|\psi_i\rangle\langle\psi_j|). \tag{22}$$

Where $S(\rho)$ is the entropy of von Neumann and $\psi^R$ is a purification of $\rho$,

$$|\psi^R\rangle=\sum_i\sqrt{p_i}|\psi_i\rangle\otimes|\phi_i^R\rangle, \quad \langle\phi_j^R|\phi_i^R\rangle=\delta_{ij}, \tag{23}$$

$$\mathrm{tr}_R|\psi^R\rangle\langle\psi^R|=\rho. \tag{24}$$

Here $\{|\phi_i^R\rangle\}$ is some orthonormal set. The definition is independent of the concrete choice of this set [6]. Mutual quantum information is the decrease of entropy after the action of $\hat{K}$ due to the possible distortion of entanglement state. This quantity is not symmetric with respect to the interchanging of input and output and can be positive, negative, or zero in contrast with Shannon mutual information in classical theory.

It has been shown that Eq. (21) can be the upper bound of the capacity of a quantum channel [22]. Using this value the authors [22] have proved the converse coding theorem for a quantum source with respect to the so-called entanglement fidelity [6]. This fidelity is absolutely adequate for quantum data transmission or compression.

In Ref. [21] the authors prove a data processing inequality

$$I(\rho;\hat{K}_1)\geqslant I(\rho;\hat{K}_2\hat{K}_1). \tag{25}$$

The quantum information cannot increase after action of $\hat{K}$. In [22] we found an alternative derivation of this result which is simpler than the derivation of [21]. In the present paper we show that this inequality can be strengthened. The data processing inequality is a very important property of mutual information. This is an effective tool for proving general results and the first step toward identification of a physical quantity as mutual information.

Now we briefly recall the derivation of the data processing inequality. The formalism of relative quantum entropy is very useful in this context.

We have

$$S(\hat{1}^R\otimes\hat{K}(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R\otimes\hat{K}(\rho^R\otimes\rho)) \\ =-S(\hat{1}^R\otimes\hat{K}(|\psi^R\rangle\langle\psi^R|))+S(\rho^R)+S(\hat{K}\rho). \tag{26}$$

Here

$$\rho^R=\sum_{i,j}\sqrt{p_ip_j}|\phi_i^R\rangle\langle\phi_j^R|\langle\psi_i|\psi_j\rangle. \tag{27}$$

Now from the Lindblad inequality we have

$$S(\hat{1}^R\otimes\hat{K}(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R\otimes\hat{K}(\rho^R\otimes\rho)) \\ \geqslant S(\hat{1}^R\otimes\hat{K}_1\hat{K}_2(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R\otimes\hat{K}_1\hat{K}_2(\rho^R\otimes\rho)). \tag{28}$$

From this formula we have Eq. (25).

Now we can prove the strengthened data processing inequality. Let $\hat{K}_2$ in (28) be represented in the form (12). From Eqs. (7) and (12) we get

$$S(\hat{1}^R\otimes\hat{K}_2\hat{K}_1(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R\otimes\hat{K}_2\hat{K}_1(\rho^R\otimes\rho))$$
$$\leqslant -(1-c)S(\hat{1}^R\otimes\hat{C}_2\hat{K}_1(|\psi^R\rangle\langle\psi^R|))$$
$$+S(\rho^R)+(1-c)S(\hat{C}_2\hat{K}\rho). \tag{29}$$

And we have

$$(1-c(\hat{K}_2))I(\rho;\hat{K}_1)\geqslant I(\rho;\hat{K}_2\hat{K}_1). \tag{30}$$

Now we consider the simplest example of a noisy quantum channel: A two-dimensional, two Pauli channel [23] with the following Krauss representation:

$$A_1=\sqrt{x}\hat{1},\quad A_2=\sqrt{(1-x)/2}\sigma_1,$$
$$A_3=-i\sqrt{(1-x)/2}\sigma_2,\quad 0\leqslant x\leqslant 1 \tag{31}$$

where $\hat{1}$, $\sigma_1$, $\sigma_2$ are the unit matrix and the first and the second Pauli matrices. Equation (31) also has physical meaning as an evolution operator for a two-dimensional open system.

Any density matrix in two-dimensional Hilbert space can be represented in the Bloch form

$$\rho=(1+\vec{a}\vec{\sigma})/2, \tag{32}$$

where $\vec{a}$ is a real vector with $|\vec{a}|\leqslant 1$. Now we have

$$\hat{K}_{TP}[(1+\vec{a}\vec{\sigma})/2]=(1+\vec{b}\vec{\sigma})/2, \tag{33}$$

where $\vec{b}=(a_1x,a_2x,a_3(2x-1))$. After simple calculations we get

$$c(\hat{K}_{TP})=(1-|2x-1|)/2. \tag{34}$$

We conclude by reiterating the main results: the Lindblad inequality can be generalized. We have presented results not only about increasing of entropy and decreasing of mutual quantum information but also about the speed of these processes.


[1] R. L. Stratanovich, *Information Theory* (Nauka, Moscow, 1975).

[2] R. P. Poplavsky, *Thermodynamics of Information Processes* (Nauka, Moscow, 1981).

[3] S. M. Barnett and S. J. D. Phoenix, Phys. Rev. A **44**, 535 (1991).

[4] L. D. Landau and E. M. Lifshits, *Statistical Physics* (Nauka, Moscow, 1976).

[5] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).

[6] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).

[7] K. Kraus, Ann. Phys. (N.Y.) **64**, 311 (1971).

[8] R. L. Stratanovich, *Nonequilibriun, Nonlinear Thermodynamics* (Nauka, Moscow, 1985).

[9] F. Shlogl, Phys. Rep. **62**, 268 (1980).

[10] M. Lax, Phys. Rev. **145**, 110 (1966).

[11] For a comprehensive review and further references about nonequilibrium steady states in open systems, see B. Schmittmann and R. K. P. Zia, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic Press, New York, 1996).

[12] S. J. D. Phoenix and P. L. Knight, Ann. Phys. (N.Y.) **186**, 381 (1988).

[13] The literature on this particular point is too broad to cite every work. A recent book of M. C. Mackey, *Time's Arrow: The Origins of Thermodynamic Behavior* (Springer-Verlag, Berlin, 1992) provides a good review. See also M. C. Mackey, Rev. Mod. Phys. **61**, 981 (1989); Yu. L. Klimontovich, Usp. Fiz. Nauk **164**, 811 (1994); Yu. L. Klimontovich, *Statistical Physics* (Nauka, Moscow, 1982).

[14] H. Umegaki, Kodai Math. Sem. Rep. **14**, 59 (1962).

[15] M. Ohya, Rep. Math. **27**, 19 (1989).

[16] G. Lindblad, Commun. Math. Phys. **40**, 147 (1975).

[17] M. J. Donald, Commun. Math. Phys. **105**, 13 (1986).

[18] F. Gantmacher, *The Theory of Matrices* (Nauka, Moscow, 1983).

[19] I. Csiszar and J. Korner, *Information Theory* (Akademiai Kiado, Budapest, 1981).

[20] S. Lloyd, Phys. Rev. A **55**, R1613 (1997); also Report No. quant-ph/9604015.

[21] B. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996).

[22] A. E. Allahverdyan and D. B. Saakian, Report No. quant-ph/9702023.

[23] C. H. Bennett, C. A. Fuchs, and J. A. Smolin, Report No. quant-ph/9611006.