International Journal of Theoretical Physics, Vol. 37, No. 1, 1998

Energy Requirements in Quantum Communication

Lev B. Levitin[1]

Received July 4, 1997

It is shown that the minimum energy required to transmit a natural unit of information (nat) over a noisy channel is kT, where T is the effective noise temperature, which depends on the interaction between the information carrier and the environment. The result follows from the entropy defect principle and general laws of quantum dynamics without any specific assumptions.

1. INTRODUCTION

The problem of the minimum energy required to transmit a unit of information over a noisy channel has been extensively discussed for many years (Felker, 1952; Jaynes, 1957; Brillouin, 1956; Bremermann, 1962, 1967a, b; von Neumann, 1966; see also references in Landauer, 1994). Still, the problem remains controversial, as stressed in a recent paper by Landauer (1994), who argues against "a widespread presumption that it takes $kT \ln 2$ to send a bit from one place to another."

The problem is, indeed, far from being settled. Most results obtained are model dependent. General resuts seem to be hard to achieve, and the difference of opinions arises, as Landauer correctly points out, "from the difficulty of the problem, when compared to our ability."

This paper is an attempt to clarify the situation and to draw some general conclusions regarding the process of information transmission over a quantum mechanical communication channel.

First of all, it is clear that, in speaking about kT, we assume that there is *noise*, i.e., interaction between the physical system carrying information and another, uncontrolled physical system. (Of course, by "system" we mean not a physical "body," but a set of degrees of freedom, that is, certain dynamic variables.) Indeed, if we control completely the degrees of freedom that are

[1] College of Engineering, Boston University, Boston, Massachusetts 02215.

0020-7748/98/0100-048 7$15.00/0

used to carry the information, and there is no interference with other degrees of freedom that could change the state of the system in an unpredictable way, it means that the system is isolated and operates at absolute zero: $kT = 0$, and, hence, there is no problem with energy dissipation.

The fact that *in the absence of noise* the energy required for transmission of a unit of information can be made arbitrarily small "if we are willing to do it slowly" (Landauer, 1994) has been well known for a long time. In particular, this was demonstrated in Lebedev and Levitin (1963, 1966) and Levitin (1965, 1969a, 1983) for various types of boson and fermion channels. Furthermore, it was shown in Levitin (1970, 1982) that, in principle, we do not need to slow down the transmission rate to make energy arbitrarily small: the Bremermann limit (Bremermann, 1962, 1967a, b) $I/t \leq E/\hbar$ (often quoted in the 1960s) was proved not to be valid in general, though, of course, for each given physical system, the value of I/Et is limited (here I denotes information in nats; t, time; E, energy; and $\hbar$, Planck's constant).

The situation changes in principle if there is interaction between the controlled and uncontrolled degrees of freedom. As shown in Section 2, the result of the interaction can be described as the effect of thermal noise characterized by a certain temperature. This temperature depends, understandably, on the strength of the interaction and does not have to be equal, in general, to the temperature of the environment. The energy put into the controlled degrees of freedom must be sufficient to make the signals discernible against the background of noise. This does not mean that all this energy must be "lost." It seems plausible that a part of it can be "reused" or converted into mechanical work, at least in principle. On the other hand, another part of energy is indeed dissipated. Thus we have to answer two different questions:

1. What is the minimum energy per unit of information that is required for transmission under noisy conditions?
2. What is the minimum energy per unit of information that inevitably dissipates as a result of the interaction with uncontrollable degrees of freedom?

It is important to understand clearly what we mean by energy dissipation. This question is far from trivial. It touches upon fundamental problems related to the Second Law of Thermodynamics. Dissipation does not mean "loss" or "disappearance" of energy. After all, energy is conserved in an isolated system. What "dissipation" really means is the increase of *entropy* of the system, which reduces the ability of the system to carry information or produce mechanical work, so that a part of the energy of the system cannot no longer be used for those purposes.

2. GENERAL BOUND

The analysis given below is based on the entropy defect principle first established in Levitin (1969b; (see also Levitin 1993a).

Let X be the input random variable X taking on values x_i with probabilities p_i. Suppose that each x_i is encoded by a state s_i of a quantum system A described by a density matrix ρ_i. Let ρ be the *a priori* density matrix of the entire ensemble of states $S = \{s_i, p_i\}$:

$$\rho = \sum_i p_i \rho_i \tag{1}$$

Denote by D the *entropy defect* of the system:

$$D = H - \overline{H} \tag{2}$$

where

$$H = -\mathrm{Tr}\, \rho \ln \rho \tag{3}$$

and

$$\overline{H} = -\sum_i p_i \,\mathrm{Tr}\, \rho_i \ln \rho_i \tag{4}$$

Let I be the maximum information about the random variable X obtainable by measurements over the random system A. [Some authors call it "accessible information" (Fuchs, 1994; Schumacher *et al.*, 1996).] Then the following inequalities are valid (Levitin, 1969b, 1993a):

$$I \leq D \leq -\sum_i p_i \ln p_i \stackrel{\Delta}{=} H(X) \tag{5}$$

The left-hand equality holds iff all density matrices ρ_i commute, and the right-hand equality holds iff all ρ_i are orthogonal.

Now consider the case when the ensemble of signals $S = \{s_i, p_i\}$ is encoded initially by an ensemble of pure orthogonal states $\{s_i^0, p_i\}$ of system A, each s_i^0 being described by a density matrix ρ_i^0. This means that we have complete control over the initial state of the information carrier—the system A. In this case the conditional enthropy is $H_0 = 0$, and the initial information is given by

$$I_0 = D_0 = H_0 = H(X) \tag{6}$$

In other words, our "quantum encoding" preserves completely the information in the ensemble of signals [which is equal to $H(X)$].

Suppose now that starting with time $t = 0$, system A interacts with another quantum system B (the environment) which is beyond our control.

Let the initial state of B be described by the density matrix σ. Without loss of generality we assume that A and B together form an isolated system with Hamiltonian $\mathbf{H}$. Then the time evolution of their joint density matrices $\rho_i^{(A+B)}$ is given by a unitary transformation:

$$\rho_i^{(A+B)}(t) = \exp\left(-\frac{i\mathbf{H}t}{\hbar}\right) \rho_i^0 \otimes \sigma \exp\left(-\frac{i\mathbf{H}t}{\hbar}\right) \tag{7}$$

where $\rho_i^0 \otimes \sigma$ is the tensor product of initial density matrices of systems A and B in their joint tensor-product Hilbert space. The factorization of the joint density matrix at time $t = 0$ reflects the independence of the initial states of systems A and B, which is due to the absence of interaction prior to time $t = 0$. At time $t \neq 0$ the joint density matrices $\rho_i^{(A+B)}(t)$ cannot be, in general, factorized. The density matrices ρ_i of the system A at the given time t are obtained by taking the trace over the variables of system B

$$\rho_i = \mathrm{Tr}_B\, \rho_i^{(A+B)}(t) \tag{8}$$

The density matrices ρ_i are, in general, nonorthogonal. Moreover, they represent, in general, mixed, rather than pure, states. The exceptional cases when this is not so are equivalent to the absence of interaction, i.e., to the free time evolution of system A with an appropriate Hamiltonian $\mathbf{H}_A$. Therefore, in general, the conditional entropy is nonzero:

$$\overline{H} = -\sum_i p_i \,\mathrm{Tr}\rho_i \ln \rho_i \neq 0 \tag{9}$$

and the initial information is partially lost:

$$I \leq H - \overline{H} < H(X) \tag{10}$$

Here ρ and H are defined by (1) and (3), respectively.

The fact that the conditional entropy (which is the average entropy of the ensemble of states $\{s_i, p_i\}$, s_i being represented by ρ_i) has increased reflects the irreversibility of the process. The ensemble of states $S = \{s_i, p_i\}$ now carries less information about X than prior to the interaction with system B and cannot be "reused" to encode as much information as the initial ensemble $S_0 = \{s_i^{(0)}, p_i\}$.

Consider now the situation in terms of energy. The initial energy of the ensemble of states is

$$E_0 = \sum_i p_i E_i^{(0)} = \sum_i p_i \,\mathrm{Tr}\rho_i^{(0)}\mathbf{H}_A = \mathrm{Tr}\, \rho^{(0)}\mathbf{H}_A \tag{11}$$

The final energy is

$$E = \text{Tr}\, \rho \mathbf{H}_A \tag{12}$$

Denote by $E(T')$ the energy of the system in the state of thermal equilibrium at temperature T'. The increase of the average entropy $\overline{H}$ means that a part of energy has been transformed into heat in the system, or transferred to the system in the form of heat. How large is this part? It is equal to the energy $E(T)$ of the system in the state of thermal equilibrium at temperature T, where T is determined by the value of entropy increase

$$\overline{H} = H(T) = \int_0^T \frac{1}{kT'} \frac{dE(T')}{dT'} \, dT' \tag{13}$$

Thus, T has the meaning of the effective noise temperature. The noise temperature depends on the interaction between system A and the environment and is not equal, in general, to the temperature of the environment (if the latter is at thermal equilibrium).

Note that $E(T')$ is the minimum possible energy of the system in a state with given entropy $H(T')$. It follows that the final energy E of the system obeys the inequality

$$E \geq E(T_1) \tag{14}$$

where T_1 is determined as the temperature of the system in the thermal equilibrium state with entropy $H = -\text{Tr}\, \rho \ln \rho$,

$$H = H(T_1) = \int_0^{T_1} \frac{1}{kT'} \frac{dE(T')}{dT'} \, dT' \tag{15}$$

$E - E(T)$ is the excess of the total energy of the system over the thermal equilibrium level that makes the ensemble of states capable of carrying information. Here we assume that interaction with the environment does not include amplification of the signal. Indeed, any amplification process increases both the energy of the signal and the noise temperature, so that the information in the signal (as well as the signal-to-noise ratio) does not increase (Helstrom, 1976). Therefore, following the approach used in communication engineering, we relate both signal and noise to the "input of the amplifier." Then

$$E_0 \geq E - E(T) \tag{16}$$

Hence, for the initial energy of the system per unit of information transmitted, we obtain

$$\frac{E_0}{I} \geq \frac{E - E(T)}{H - \bar{H}} \geq \frac{E(T_1) - E(T)}{\displaystyle\int_T^{T_1} \frac{1}{kT'} \frac{dE(T')}{dT'}\, dT'} \tag{17}$$

$$= \frac{E(T_1) - E(T)}{\displaystyle\frac{1}{kT^*} \int_T^{T_1} \frac{dE(T')}{dT'}\, dT'} = kT^*$$

where

$$T \leq T^* \leq T_1 \tag{18}$$

Thus, the following theorem has been proved:

Theorem 1. The minimum energy ε per unit of information required to transmit information over a channel with effective noise temperature T satisfies the inequality

$$\mathscr{E} = \frac{E_0}{I} \geq kT \tag{19}$$

Inequality (19) establishes a general lower bound for the minimum energy per unit of information (nat). It follows from the laws of quantum dynamics without any specific assumptions. Note that expression (19) approaches equality iff $E_0/kT \to 0$ (the signal-to-noise ratio is small).

In the limiting case when $E = E(T)$ the information is completely lost: $H - \bar{H} = 0$, all density matrices ρ_i become identical, and the ensemble of states cannot be used for information transmission or for producing mechanical work. [More precisely, if all ρ_i are identical, the process is impossible whose sole result would be converting heat taken from a thermostat into mechanical work, without changing the total energy of our system and without any other changes in the universe. Such a process is always possible, in principle, if ρ_i are not indentical; cf. Levitin (1993b).]

It is more difficult to give a general answer to the second question stated in Section 1 concerning the dissipation of signal energy, because the information carrier, i.e., system A, is not isolated from the unobservable system B, the environment. Therefore, there is no way to identify clearly which part of heat production and heat transfer is due to system A or to system B. However, the change of the entropy of the environment as the result of the interaction with the signal is limited. In particular, if the initial state of the environment was pure, its final entropy is $H_B = \bar{H}$. Therefore, if the environment is very large compared to system A (i.e., has many more degrees of freedom), then the part of its energy corresponding to the change of entropy is very small compared to $E(T)$. Thus, there is a good reason to consider $E(T)$ as a measure of energy dissipation resulting from the interaction of information carrier with the environment. One can see that the ratio $E(T)/I$ can be larger or smaller than kT, depending on the concrete model. Let us emphasize once again that we do not consider here possible changes of entropy resulting from the fact of measurement (except for the inequality $I \leq D$).

Then, what is the origin of the "widespread presumption" that kT is the minimum energy to be dissipated per one nat of information? Von Neumann's (1966), reasoning as well as that of many other authors, is based on the assumption that all the energy dissipates in the process of measurement performed after each act of information transmission. This assumption is quite realistic and, indeed, takes place in many practical situations. Then, inequality (19) applies, where T means the temperature of the environment.

REFERENCES

Bremermann, H. J. (1962). Optimization through evolution and recombination, in *Conference on Self-Organizing Systems*, M. C. Yovits, G. T. Jacobi, and G. D. Goldstein, eds., Spartan Books, Washington, D.C.

Bremermann, H. J. (1967a). Quantitative aspects of goal seeking self-organizing systems, *Progress in Theoretical Biology,* **1**, 59–77.

Bremermann, H. J. (1967b). Quantum noise and information, In *Proceedings 5th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, Vol. 4, pp. 15–22.

Brillouin, L. (1956). *Science and Information Theory*, Academic Press, New York.

Felker, J. H. (1952). A link between information and energy, *Proceedings IRE,* **40**, 728–729.

Fuchs, C., and Caves, C. (1994). Ensemble-dependent bounds for accessible information in quantum mechanics, *Physical Review Letters,* **73**, 3047–3050.

Helstrom, C. W. (1976). *Quantum Detection and Estimation Theory*, Academic Press, New York.

Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical Review,* **106**, 620–630.

Jozsa, R., Robb, D., and Wooters, W. (1994). Lower bound for accessible information in quantum mechanics, *Physical Review, A,* **49**, 668–677.

Landauer, R. (1994). Zig-zag path to understanding, In *Proceedings Workshop on Physics and Computation PhysComp'94*, IEEE Computer Society Press, Los Alamitos, California, pp. 54–59.

Lebedev, D. S., and Levitin, L. B. (1963). The maximum amount of information transmissible by electromagnetic field, *Soviet Physics-Doklady,* **8**, 377–379.

Lebedev, D. S., and Levitin, L. B. (1966). Information transmission by electromagnetic field, *Information and Control,* **9**, 1–22.

Levitin, L. B. (1965). Information transmission in the ideal photon channel, *Problems of Information Transmission,* **1**, 55–62.

Levitin, L. B. (1969a). Physical limitations in information transmission channels, Ph.D. Thesis, Gorky University, Gorky, Russia.

Levitin, L. B. (1969b). On the quantum measure of information, in *Proceedings 4th Conference on Information Theory*, Tashkent, pp. 111–116 [English translation, *Annales de la Fondation Louis de Broglie,* **21**(3) (1996)].

Levitin, L. B. (1970). Does there exist a fundamental limit of information processing rate? in *Third International Symposium on Radio Electronics*, Varna, Bulgaria, Part 3, pp. 1–15.

Levitin, L. B. (1982). Physical limitations of rate depth, and minimum energy in information processing, *International Journal of Theoretical Physics,* **21**, 299–309.

Levitin, L. B. (1983). Ideal corpuscular information channels, in *IEEE International Symposium on Information Theory, St. Jovite, Quebec, Canada.*

Levitin, L. B. (1993a). Physical information theory. Part II: Quantum systems, In *Proceedings Workshop on Physics and Computation. PhysComp' 92*, IEEE Computer Society Press, Los Alamitos, California, pp. 215–219.

Levitin, L. B. (1993b). Gibbs paradox and equivalence relation between quantum information and work, in *Proceedings Workshop on Physics and Computation. PhysComp' 92*, IEEE Computer Society Press, Los Alamitos, California, pp. 223–226.

Schumacher, B., Westmoreland, M., and Wooters, W. K. (1996). Limitations on the amount of accessible information in a quantum channel, *Physical Review Letters,* **76**, 3452.

Von Neumann, J. (1966). *Theory of Self-Reproducing Automata*, Lecture 3, University of Illinois Press, Urbana, Illinois.