International Journal of Theoretical Physics, Vol. 21, Nos. 3/4, 1982

Physical Limitations of Rate, Depth, and Minimum Energy in Information Processing

L. B. Levitin

Syracuse University, Syracuse, New York

Received May 6, 1981

The effect of the quantum nature of matter on the maximum information-processing potentialities is considered. It is shown that the degeneracy of the energy levels of a physical information-processing system results in the fact that a universal limit of information-processing rates does not exist, though for any specific system this rate is indeed bounded. A physical interpretation is then proposed for an elementary act of information-processing and the concept of information-processing depth is introduced. The example of a system of quantum oscillators is used to show that the maximal information-processing depth is bounded, only a very small fraction of the possible system states being used. The effect of thermal noise on information processing is briefly discussed.

1. INTRODUCTION

An important question in cybernetics (we use this word in the strict sense of Norbert Wiener) is whether there exist certain limiting relationships, following from the fundamental physical nature of information-processing processes, which in principle limit the potentialities of any natural or artificial cybernetical system. A quantitative consideration of this question should proceed from the fact that any processing on information, whatever its logical or semantic aspects are (logical or arithmetical, computation, analog simulation, random search, etc.), is in its physical essence nothing other than a transfer of information from one physical system to another in the process of their interaction. Thus, information processing may be described as transmission of information over communication channels inside a cybernetical system, using the Shannon measure of information.

There is a profound relationship between the concepts of information theory and statistical physics (Brillouin, 1960; Lebedev and Levitin, 1966).

0020-7748/82/0400-0299$03.00/0

The "entropy defect principle" (Lebedev and Levitin, 1966) makes it possible to interpret the Shannon information as a measure of the deviation of the system from a state of thermodynamic equilibrium and to express it in terms of the physical entropy of the systems constituting the communication channel. Thus, no information can exist by itself: it is "attached" to a certain ensemble of states of some physical system. Speaking somewhat loosely, one might say that thought (understood in this context as an information-processing process) is not only not a nonmaterial, "imponderable" entity, as supposed by some philosophers of the past, but may even be described quantitatively (though not semantically) in physical terms.

A consequence of this is that the amount of information which can be processed by any real system restricted in time, space, and energy cannot be arbitrarily large. However, this does not yet imply that there are universal limits, expressible in terms of world constants and valid for all concrete systems irrespective of their macroparameters and microscopic structure.

Bremermann (1962, 1967a, 1967b) concluded that there exists such a universal limit of the information rate per unit of mass, owing to the quantum nature of matter, equal to

$$c^2/h = 1.35 \times 10^{47} \text{ bit/g sec}$$

where c is the velocity of light and h is Planck's constant. More precisely, Bremermann asserts that the information-processing rate in any system cannot exceed E/h (bits/sec), where E is the total energy of the system, or, according to the equivalence of energy and mass, mc^2/h (bits/sec), where m is the mass of the system.

In other words, there is an upper bound for the ratio of the amount of processed information to the product of the energy used in its processing and the processing time:

$$\frac{I}{Et} \leqslant \frac{1}{h} \text{ bit/erg sec} \tag{1}$$

Inequality (1) is indeed very remarkable and it has been widely accepted. However, it is valid only under certain conditions. As will be shown below, the arguments utilized to derive (1) implicitly ignore some essential factors: the degeneracy of energy levels of the physical systems that transmit the information, the possibility of repeated use of energy, and parallel transmission of information. These factors imply that the quantity I/Et has no upper limit common to all physical systems; in the absence of noise, it may assume arbitrarily large values, though for any specific system it is of course bounded.

2. EFFECT OF DEGENERACY OF ENERGY STATES

Let us assume that the state of an information-processing system varies in time in a completely deterministic manner. In the language of physics, this means that the initial state and Hamiltonian of the system are known exactly, and that there are no interactions of a statistical nature (such as interaction with a thermostat). In that case there is no loss of information due to random thermal motion and the information-processing potentialities of the system are limited only by the quantum properties of matter and field.

What are the premises under which inequality (1) is valid? To see what these are, we present (with some specifications) the derivation of (1) first given by Bremermann (1962).

Let the total energy of the system be E and the transmission time of the signal t. If the energy state of the signal is nondegenerate (i.e., there is only one quantum state with given energy) then the number of distinguishable signals (i.e., almost-orthogonal quantum states) of length t is $Et/h+1=E/\Delta E+1$, where $\Delta E=h/t$ is the quantum mechanical uncertainty of the signal energy. Hence the amount of information transmissible in unit time is (in natural units)

$$\frac{I}{t}=\frac{1}{t}\ln\left(\frac{E}{\Delta E}+1\right)=\frac{E}{h}\frac{\Delta E}{E}\ln\left(\frac{E}{\Delta E}+1\right)\leqslant\frac{E}{h}\ln 2 \text{ nat/sec} \tag{2}$$

(since $E/\Delta E\geqslant 1$, we have $E/\Delta E+1=x\geqslant 2$, and

$$\frac{\Delta E}{E}\ln\left(\frac{E}{\Delta E}+1\right)=\frac{1}{x-1}\ln x\leqslant \ln 2)$$

This is precisely Bremermann's limit.

The situation changes in an essential way if the energy levels are degenerate, i.e., there exist several orthogonal states with the same energy. Such states are uniquely distinguishable, not by their energies but by measurement of other variables ("quantum numbers"). Let the degeneracy of all the states except the ground one be K. Then the transmission rate is

$$\frac{I}{t}=\frac{E}{\hbar}\cdot\frac{\Delta E}{E}\cdot\ln\left(K\frac{E}{\Delta E}+1\right)$$

Take $\Delta E=E$ (this means that we are using only two different values of signal energy). Then

$$\frac{I}{t}=\frac{E}{h}\ln(K+1)\text{ nat/sec} \tag{3}$$

small:

$$\bar{n} = \frac{Ec^2}{\Omega S t h \nu^3 \Delta\nu} \ll 1 \tag{7}$$

and if the following three criteria hold, the system should be "narrow band":

$$\alpha = \nu/\Delta\nu \gg 1 \tag{8}$$

"three dimensional":

$$\beta = \frac{\Omega S \nu^2}{c^2} \gg 1 \tag{9}$$

and "quasiclassical":

$$\gamma = t\,\Delta\nu \gg 1 \tag{10}$$

It is easy to see that the above criteria (7)–(10) may be observed while allowing I/Et to assume arbitrarily large values (e.g., by choosing β arbitrarily large for fixed α, γ, E, and t). Note that in the case considered the energy E was used once during the time t.

The calculation for an ideal corpuscular channel (Levitin, 1981b), using particles with a rest mass m instead of photons, gives a similar result:

$$\frac{I}{Et} = \frac{1}{\varepsilon t} \ln \frac{2e\Omega S m t \varepsilon^2 \Delta\varepsilon}{h^3 E} \tag{11}$$

where ε is the average energy of one particle and $\Delta\varepsilon$ the interval of particle energies used. The conditions parallel to criteria (7)–(10) are

$$\bar{n} = \frac{h^3 E}{2\Omega S m t \varepsilon^2 \Delta\varepsilon} \ll 1, \qquad \alpha = \varepsilon/\Delta\varepsilon \gg 1$$

$$\beta = \frac{2\Omega S m \varepsilon}{h^3} \gg 1, \qquad \gamma = \frac{t\,\Delta\varepsilon}{h} \gg 1 \tag{12}$$

It is obvious that in this case too I/Et may be arbitrarily large.

3. MAXIMUM DEPTH OF INFORMATION PROCESSING

We now consider another aspect in which the quantum structure of matter may limit information-processing potentialities. It is well known in cybernetics that a finite automaton is capable of processing information only to a finite depth (for example, it can deduce only finitely many theorems from a given axiomatic system), since the number of different states of a finite automaton is finite. However, the questions of whether there exist limits to "microminiaturization" and what are the potentialities of an "ideal" finite automaton belong to the realm of physics.

From the standpoint of classical physics, the number of different states in any finite phase volume is infinite. According to quantum mechanics, the number of different orthogonal states ("cells") in a phase volume Γ is (in the quasiclassical approximation, i.e., for large phase volumes):

$$N = \frac{\Gamma}{(2\pi\hbar)^n} \tag{13}$$

where n is the number of degrees of freedom of the system. It seems reasonable to define an elementary act of "information processing" to be a transition of the system to a state orthogonal to the preceding state, since only orthogonal states can be unambiguously distinguished by measurement. Thus, the information-processing depth, i.e., the number of different ways in which a system can transform the initial information (specified by the choice of the initial state or, in other words, the number of different possible forms in which the initial information can be represented in the system, is equal to the number of different orthogonal states through which the system passes from the given initial state in the course of a Poincaré cycle (i.e., until it returns to a state close to the initial one). The important point is that systems with a large number of degrees of freedom may have many single-valued integrals of motion. Therefore, if we consider information processing in a closed system, every phase trajectory will contain only a small part of the set of all orthogonal states localized near a given energy surface. Consequently, the information-processing depth is not equal to the number N of different states given by equation (9), but is much smaller.

Note that states which are orthogonal at a certain time remain orthogonal throughout further evolution of the system. Consequently, if the information is specified by the choice of some state from a set of orthogonal states, then, in principle, no indeterminacy will arise in the course of its further processing. However, generally speaking, the combination of measurements (complete set of variables) permitting unambiguous discrimination of these states depends on the instant of time and on the evolution law of the system, which in turn is determined by its Hamiltonian. In other words, to carry out the necessary measurement at any instant of time one must calculate the time dependence of the complete set of variables whose eigenstates are just our orthogonal states. But this means that one must perform a calculation equivalent to the very processing of information being done by the system! The paradox is resolved if the motion of the system is sufficiently near to being classical, viz., at definite intervals of time the system comes into states that belong to the initial set of orthogonal states (i.e., to the initial partition of the phase space into cells). In the general case, this condition is satisfied for sufficiently small time intervals, during which the "dispersion of the wave packet" may be ignored (nevertheless, the number of different orthogonal states gone through within this time may be very large). However, there are cases in which motion of this type may go on for an unlimited time. An example of such a situation will now be analyzed.

Consider a system which can be brought by a canonical transformation of variables to the form of a system of noninteracting oscillators of the same frequency ω. It is convenient to take the dimension of the phase variables equal to the square root of the energy. The Hamiltonian of the system is then

$$\hat{H} = \sum_{i=1}^{n} \tfrac{1}{2}(\hat{p}_i^2 + \hat{q}_i^2) \tag{14}$$

where $\hat{p}_i$ and $\hat{q}_i$ are operators of, respectively, momentum and coordinate of the ith oscillator, and n is the number of oscillators. To each quantum state of the oscillator system there corresponds a cell of volume $(2\pi\hbar\omega)^n$ in $2n$-dimensional phase state. The energy surface corresponding to energy E is a $2n$-dimensional sphere of radius $(2E)^{1/2}$.

We now assume that all the oscillators are in states of an "oscillating wave packet," corresponding most closely to the classical motion of an oscillator with definite amplitude and phase (Schrödinger, 1926). [Nowadays, they are usually called "coherent" states (Senitzky, 1962; Glauber, 1963).] In such states the shape of the wave packet, the average energy, and phase of the oscillators are preserved. These states are not exactly orthogonal; in the quasiclassical case, however, if the average energy of each oscillator and the number of oscillators are sufficiently large ($E/n\hbar\omega \gg 1$, $n \gg 1$), one can choose a set of coherent states such that the states are asymptotically orthogonal. These states are concentrated in $2n$-dimensional spherical cells of radius $r = (n\hbar\omega)^{1/2}$, centered at points lying on the sphere of energy E. The number of such orthogonal states when $n \gg 1$ and $E/n\hbar\omega \gg 1$ is asymptotically equal to

$$N \approx \left(\frac{2E}{n\hbar\omega}\right)^n \tag{15}$$

and the information associated with the choice of one of the states is

$$I = \ln N \approx n \ln \frac{2E}{n\hbar\omega} \tag{16}$$

In this situation, the cells may be placed in such a way that the centers of closest neighbors are spaced apart by a distance $r\sqrt{2} = (2n\hbar\omega)^{1/2}$. (Thus neighboring cells intersect, despite the orthogonality of the corresponding states.) The phase trajectory of the center of one of these cells is a circle of radius $(2E)^{1/2}$.

Thus, the number of different orthogonal cells "threaded" on the phase trajectory of the center of one cell is

$$D = \frac{2\pi(2E)^{1/2}}{r\sqrt{2}} = 2\pi\left(\frac{E}{n\hbar\omega}\right)^{1/2} \tag{17}$$

This is the very number of different orthogonal states through which the system passes during a Poincaré cycle (which is clearly equal to the oscillator period $\tau = 2\pi/\omega$).

Thus, the number of states D through which the system passes from some given initial state is very small compared with the total number N of asymptotically orthogonal states with the given average energy:

$$D \approx \pi\sqrt{2}\,N^{1/2n} \tag{18}$$

Metaphorically speaking, a system of this type is capable, starting from different "premises," of developing a great number of different "theories," though each of these will be relatively short.

Note that the rate of information processing per unit energy is in this case bounded from above by a quantity which exceeds Bremermann's limit by a factor approximately 9.4. Indeed, the average length of one processing act is

$$\Delta t = \frac{\tau}{D}$$

Hence, in view of (16) and (17), we have

$$\frac{I}{E\Delta t} = \frac{\sqrt{2}}{\hbar}\left(\frac{n\hbar\omega}{2E}\right)^{1/2}\ln\frac{2E}{n\hbar\omega} \leqslant \frac{\sqrt{2}}{\hbar}\cdot\frac{2}{3}\ln 3 = \frac{6.508}{h} \tag{19}$$

[according to equation (2), Bremermann's limit is $\ln 2/h = 0.693/h$]. The existence of a limit is due to the fact that all oscillators take part simultaneously and "synchronously" in processing the information while their energy states are nondegenerate (i.e., the number of degrees of freedom of each oscillator is unity).

The nature of these results is not qualitatively altered if the oscillators have different frequencies. The average of each oscillator is an integral of motion, and there are a great number of phase trajectories corresponding to different distributions of energy among the oscillators such that states belonging to different trajectories are orthogonal (at least, asymptotically). Thus, the number of different disjoint subsets of states is not less than the number of corresponding choices of energy values for the oscillators, the total E being given. The situation is exactly the same in any physical system possessing integrals of motion which are single-valued functions of the state of the system (Landau and Lifshitz, 1958).

4. EFFECT OF THERMAL NOISE

The most serious limitation on the possibilities of information processing is the statistical nature of physical processes. In reality we are usually in a position to control only a few, predominately macroscopic degrees of freedom of the system. The remaining degrees of freedom, of which there are a tremendous number, interact chaotically with these "useful" degrees of freedom, causing dissipation of signal energy and appearance of thermal noise. Unfortunately, there are at present few general results concerning the theoretical limitations imposed on information processing by relaxation processes. Brillouin (1960) showed through a number of examples that the minimal energy per natural unit of information is kT (where k is Boltzmann's constant and T the absolute temperature of the system). However, this assertion has actually been proved rigorously only for ideal physical channels with additive statistically independent noise (Levitin, 1982). It is shown in that paper that the minimal amount of energy necessary to transmit a unit of information over a channel of this type is

$$E_{\min} = \frac{E_0}{I_{\max}} \geqslant kT_1 \tag{20}$$

where E_0 is the signal energy, $I_{\max}$ the maximal amount of information, and T_1 the noise temperature. Equality holds asymptotically when the signal is weak, i.e., when

$$\frac{E_0}{E_l} \ll 1$$

where E_1 is the noise energy. In particular, for the one-dimensional photon channel (Lebedev and Levitin, 1966),

$$E_{\min} = kT_1 + \frac{3}{\pi}\hbar R$$

where R is the information transmission rate. It is to be expected that (20) will also be valid under considerably more general conditions.

Intuitively, it is clear that the information-processing potentialities of a closed physical system depend on the nonequilibrium of the system. Subsystems which are in complete thermodynamic equilibrium with each other are clearly incapable of exchanging information. For instance, a television camera placed in a cavity filled with black-body radiation can "see" something therein only if there is a temperature difference between cavity and camera. On the other hand, one should expect that more intensive information processing will increase energy dissipation and speed up relaxation of the system.

ACKNOWLEDGMENTS

The author is deeply grateful to Professor H. Bremermann of the University of California at Berkeley for many fruitful discussions, for making available copies of his papers, and rendering moral support.

REFERENCES

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 of the 5th Berkeley Symposium on Mathematical Statistics and Information, University of California Press, Berkeley, Vol. 4, pp. 15–22.

Brillouin, L. (1960). *Science and Information Theory*, 2nd ed. Academic Press, New York.

Glauber, R. J. (1963). "The quantum theory of optical coherence," *Physical Review*, **130**, 2529–2539.

Landau, L. D., and Liftshitz, E. M. (1953). *Mechanics*. Fizmatgiz, Moscow (in Russian).

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 an ideal photon channel," *Problems of Information Transmission*, **1**, No. 3.

Levitin L. B. (1981). "A thermodynamic characterization of ideal physical information channels," *Journal of Information and Optimization Sciences*, **2**, No. 3.

Levitin, L. B. (1982). "Ideal corpuscular information channels," (to be published).

Schrödinger, E. (1926). "Der stetige übergang von der Mikro zur Makromechanik," *Naturwissenschaft*, **14**, 664–676.

Senitzky, I. R. (1962). "Incoherence, quantum fluctuations and noise," *Physical Review*, **128**, 2864–2870.