Physical Information Theory Part I. Quasiclassical Systems

Lev B. Levitin

Boston University, Boston, MA

Abstract

A short review of the development of the physical information theory is presented. The entropy defect principle is formulated for quasiclassical systems. The concept of ideal physical information channels is iintroduced. It is shown that information properties of such a channel with independent additive noise can be ontained from the thermodynamic description of the physical system that transmits information. Brillouin's conjecture of minimum enerrgy per unit of information is proved for this type of channels.

1. Introduction

The origin of the physical information theory ascends to L. Boltzmann and L. Szilard [1] , who attributed an information meaning to the thermodynamic notion of entropy long before the quantitative measure of information was rigorously introduced by C.E. Shannon [2] . However, the information theory leading off with fundamental Shannon's work developed at first as a pure mathematical branch of science. An impression arose that the laws of transmission and processing of information were not physical and the concepts of the information theory could not be defined on the base of physical concepts. The erroneousness of such views was noted as far back as 1950 by D. Gabor [3] , who pointed out that "the communication theory should be considered as a branch of physics." Later in the remarkable book of L. Brillouin [4] a profound relationship between physical entropy and information was suggested in the form of the "negentropy principle of information". Though the term "negentropy" was not rigorously defined, the principle implied that any information is represented by a certain state of a physical system and associated with its deviation from the thermodynamic equilibrium. The "entropy defect principle" [5,6] implements this basic idea in a mathematically consistent and rigorous form. According to this principle, any information is represented by an ensemble of states of a physical system and associated with the average deviation from the thermodynamic equilibrium caused by the choice of one of those states. Thus the information properties of real systems can be described in purely physical terms and a way is open to develop a consistent physical information theory.

In the first applications of information theory the communication systems were described on macroscopic level and their physical properties, in so far as they were taken into account, were considered from the point of view of classical physics.

It should, however, be borne in mind that the physical foundations of information theory must be essentially based on quantum theory. Indeed, in classical statistical physics the entropy dimension is that of logarithm of action, and entropy includes, as was noted by Planck [7] , an infinite additive constant, so that only the difference of two values of entropy has physical meaning, but not the absolute value of entropy. Such a definition of entropy corresponds to the so called "differential entropy" (the term coined by Kolmogorov [8]) in information theory. But differential entropy depends on the choice of variables and does not allow calculating (without additional assumptions) the amount of information in the continous message (formally the amount of information becomes infinite, which is a nonsense from a physical point of view). In the conventional theory the properties of continous message ensembles are usually described in terms of ε-entropy. Such a description, however, is purely phenomenological and is valid only in the range where quantum effects make no appearance. According to quantum theory, every real physical system has a well-defined finite ("absolute" by Planck) entropy which leads to a natural absolute measure of information associated with statistical ensembles of quantum macrostates.

On the other hand, the vigorous development of quantum electronics, which brought into use the infrared and optical range of electromagnetic waves and allowed us to build systems with extremely low thermal noise, calls for consideration of the quantum nature of information transmission processes.

0-8186-3420-0/93 $03.00

The physical information theory developed historically along the following three lines:

1. Investigation of the interrelations between the basic concepts of statistical physics and information theory.
2. Investigation of the physical nature of information transmission and processing.
3. Application of information-theoretical concepts and approaches to the problems of statistical physics.

Though numerous papers have been devoted to the first two areas (e.g., [9-19]), the only book on this subject is [20] (in Russian). The use of quantum carriers of information brings into consideration the fundamental properties of quantum measurements (e.g. [21-27]). Results on quantum detection and estimation theory have been systemized in the comprehensive monograph by C. Helstrom [28] . It was found that a consistent quantum-mechanical treatment of information transmission calls for generalization of basic information-theoretical concepts [6,29-32] . The effect of the quantum-mechanical irreversibility of the measurement on the information transfer has been revealed in the quantum-mechanical formulation of the entropy defect principle [6,29] . The controversial problem of information efficiency of direct and indirect quantum measurements has been solved in [33].

A new remarkable application of quantum physics to information theory is quantum cryptography [45-47] which employs a new principle of secure communicaion based on the properties of quantum measurements.

Investigations in the third of above-mentioned areas (e.g. [4,34-39]) promise to shed a new light on the foundations of statistical physics, in particular, regarding the origin and the meaning of the Second Law of Thermodynamics. It seems that there exists fundamental relationships between information, energy and work which manifest itself in such important and still controversial physical situations as Gibbs' paradox, Maxwell's demon, etc. Results obtained in [37] lead to a conjecture that in the case of irreversible processes the change of information rather than that of entropy determines the amount of work produced by a non-equilibrium system.

Information-theoretical concepts seem to be fruitfully applicable to some classical problems of statistical physics, such as the Ising model [40,41] .

In general, one can conclude that physical information theory is, today, a well-established area of science with a number of rigorous general results and important applications in communication engineering.

This part deals with quasiclassical systems only. The generalization for quantum systems is presented in Part II of this paper.

2. Entropy Defect Principle (Quasiclassical Formulation)

The entropy defect principle was first formulated for the quasiclassical case in [5], with the view to obtain a general expression for information (in Shannon's sense) in terms of physical entropy. The quasiclassical description of a physical system implies, as usual, that all the operators corresponding to physical observables commute, which leads to two important premises:

1. All different microstates ("complexions", by Planck) of the system are inequivocally distinguishable by the same (complete) measurement;
2. Any macrostate of the system is uniquely represented by a statistical ensemble of the microstates.

Consider and ensemble $X=\{x_i,p_i\}$ of signals xi with corresponding probabilities p_i (i is the subscript of a signal; for the sake of simplicity the set of signals is presumes to be denumerable) and let this ensemble act on a physical system. The signal x_i brings the system to the state s_i. As a result, we obtain an ensemble of macrostates $S=\{s_i,p_i\}$ which is in one-to-one correspondence with the ensemble of signals X. Each s_i is, generally, a certain macrostate of the system, i.e., a certain ensemble of various microstates defined by probabilities of microstates w_{ik} (k is the subscript of a microstate; the set of microstates is presumed to be denumerable). Macrostate s_i has physical entropy H_i defined by the following expression:

$$H_i=-\sum_k w_{ik} \ln w_{ik}$$

(Information and entropy are henceforth expressed in natural units--nats.)

We introduce now the *average* entropy of macrostates s_i

$$\overline{H}=\sum_i p_i H_i$$

Note that $\overline{H}$ does not have, in general, the meaning of the entropy of a particular macrostate but it is the conditional expected value of the entropy of a macrostate arising under the action of a determined (prescribed) signal.

On the other hand, it makes sense to speak about the mactostate s of the physical system, arising when the signal is chosen from the ensemble at random. The corresponding probabilities of microstates are

$$w_k=\sum_i p_i w_{ik}$$

and the entropy is

$$H=-\sum_k w_k \ln w_k$$

We call the quantity

$$I_0 = H - \overline{H} \quad (1)$$

"entropy defect".

The entropy defect shows how far (on the average) the state of the system, arising under the action of a determined signal, is from thermodynamic equilibrium (when entropy is a maximum and probabilities of microstates are given by Gibbs' distribution) in comparison with the state arising under the action of signal chosen completely at random. In short, the entropy defect shows how far the *determinate choice* of the signal deflects the system from its thermal equilibrium state.

The entropy defect principle claims the following. The amount of information about the signal A obtained by a physical system is equal to its entropy defect

$$I = I_0 \quad (2)$$

Indeed, since distinctions between the microstates are the finest possible distinctions between the states of a physical system, the maximum information about the signal is obtainable by measuring the microstate, i.e. by performing a complete measurement that distinguishes between all the microstates. Let $M = \{m_k\}$ be the random variable representing the microstate of the system. The joint probability distrubution of two random variables - the signal X and the microstate M - is given by

$$\Pr\{X = x_i,\ M = m_k\} = \Pr\{S = s_i,\ M = m_k\} = p_i w_{ik} \quad (3)$$

Therefore,

$$I_0 = H(M) - H(M|X) = I(M;X), \quad (4)$$

where I(M;X) is Shannon's information about the signal X in the microstate M.

Though the equality (2) seems to be almost evident, it is far from being trivial. Indeed, its validity is based on two crucial assumption 1. and 2. about the nature of the physical world. These assumptions are not valid for a consistent quantum-mechanical description, where the equality (2) turns into inequality, as will be shown in Part II. The importance of the entropy defect principle is that it allows us to interpret information as a measure of the deviation of a physical system from the thermodynamic equilibrium state and to establish not only mathematical analogy but also identity of fundamental concepts of information theory and statistical physics.

3. Ideal Physical Information Channels

The entropy defect principle has been used to analyze properties of certain systems which are, physically speaking, the simplest information transmission systems - ideal physical channels. We define an ideal physical information channel to be a channel in which the transmitter uniquely specifies the microstate of the transmitted signal - the physical agent that carries the information, and the receiver uniquely determines the microstate of the received signal. Thus, the signal ensemble in an ideal physical channel is the ensemble of microstates of a certain physical system.

The following results have been obtained in [42].

Theorem 1. The capacity of a nonideal channel does not exceed the capacity of an ideal channel with the same transmission probabilities, obeying the same constraints on the ensemble of microstates of the transmitted signal.

Consider in more detail an ideal physical channel with statistically independent additive noise. In a channel of this type, the received signal is formed by addition of the transmitted signal and noise (the latter being of the same physical nature as the signal). The received signal and the noise may, therefore, be regarded as two different states of the same physical system. We shall assume that the microstate of this system is uniquely determined by the energies assigned to the various degrees of freedom (for example, to the various quantum states of the particles constituting the system). Additivity means here that the energy assigned to any degree of freedom of the received signal is the sum of the respective energies for the transmitted signal and the noise. (In particular, if one is using a description in terms of occupation numbers of quantum states of the particles constituting the system, the occupation numbers of the received signal are the sums of the respective occupation numbers of the transmitted siganl and the noise.) Statistical independence means that the distribution of noise microstates is independent of the transmitted signal.

An ideal physical channel with statistically independent additive noise is an idealized model of systems in which the information carriers are particles obeying Bose-Einstein statistics (such as photons or elementary acoustical excitations - phonons). If the information carriers are particles obeying Fermi-Dirac statistics, this description is apt only in the limiting case of Boltzmann statistics (i.e., small average occupation numbers).

Let the ensemble average of the energy of the transmitted signal be E_0, that of noise E_1. Then, by virtue of additivity, the average energy of the received signal is $E_2 = E_0 + E_1$. Assume, moreover, that the noise is thermal, i.e., the distribution of noise microstates corresponds to thermodynamic equilibrium of the physical system in question (i.e., of the agent transmitting the information) at temperature $T_1 = T(E_1)$, where T(E) is the temperature of the system as a function of its average energy. Under these conditions the entropy H_1 of the noise has the maximum value possible given the average energy E_1. We then have the following.

Theorem 2. The maximum amount of information that can be transmitted over an ideal physical channel with statistically independent additive noise is

$$I_{max} = \int_{E_1}^{E_2} \frac{dE}{kT} = \int_{T_1}^{T_2} \frac{1}{kT} \frac{dE(T)}{dT} dT \qquad (5)$$

and this maximum is achieved when the distribution of microstates of the received signal corresponds to thermodynamic equilibrium at the temperature T_2 determined by

$$\int_{T_1}^{T_2} \frac{dE(T)}{dT} dT = E_0 \qquad (6)$$

Here E(T) is the average energy of the system as a function of temperature; $E(T_2) = E_2$.

Corollary. The minimum amount of energy necessary to transmit one natural unit of inforamtion over a channel with statistically independent additive noise satisfies the inequality

$$E_{min} = \frac{E_0}{I} \geq kT_1 \qquad (7)$$

where T_1 is the temperature of the noise. Equality holds asymptotically in case of weak signal, i.e., when

$$\frac{T_2 - T_1}{T_2} << 1 \qquad (8)$$

(For the usual physical systems, in which E(T) is a monotonic and smooth function, condition (8) is equivalent to $E_0/E_1<<1$).

This corollary is a rigorous specialization of the well-known conjecture of Brillouin [4] that the minimum energy necessary to obtain one bit of information at temperature T_1 is $kT_1 \ln 2$.

In particular, the minimum energy per nat has been calculated in [5,10] for a broadband one-dimensional photon channel:

$$E_{min} = kT + \frac{3}{2\pi^2} hR \qquad (9)$$

Here h is Plank's constant, R is the rate of information transmission (nats/s). This elegant formula clearly shows both the effect of thermal noise and the limitation imposed by the quantum structure of the electromagnetic field: the faster one transmits information, the more energy required per each unit of information. Other types of ideal physical channels have been investigated in [5,9-11,43,44]. For corpuscular channels (with particles of non-zero rest mass as information carriers) the specific effects of degeneracy (Bose-Einstein condensation and Fermi saturation) which affect channel capacity have been demonstrated [44].

References:

1. Szilard, L.: Über die Entropieverminerung in einem thermodynamischen System bei Eingriff intelligenter Wesen, Z. Physik, B. 53, No. 5 (1929), 840-856 (in German).

2. Shannon, C.E. and Weaver, W.: The Mathematical Theory of Communication, Urbana Univ. Press., Chicago, Ill. 1949.

3. Gabor, D.: Communication theory and physics, Phil. Mag., 41 (1950), 1161-1187.

4. Brillouin, L.: Science and Information Theory, Acad. Press, New York 1956.

5. Lebedev, D.S. and Levitin, L.B.: Information transmission by electromagnetic field. Inform. Contr. 9, No. 1 (1966), 1-22.

6. Levitin, L.B.: On the quantum measure of the amount of information, in: Proc. of the IV National Conf. on Information Theory, Tashkent, 1969, 111-115.

7. Plank, M.: Wärmestrahlung, Berlin 1913 (English translation: Theory of Heat Radiation, Dover, New York 1959).

8. Kolmogorov, A.N.: On the Shannon theory of information in the case of continuous signals, IRE Trans. on Inform. Theory, IT-2 (1956), 102-108.

9. Gordon, J.P.: Quantum effects in communication systems. Proc. IRE 60, No. 9 (1962), 1898-1908.

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

11. Takahasi, H.: Information theory of quantum-mechanical channels, in: Advances in Communication Systems, v.1, Acad. Press, New York 1965, 227-310.

12. Stratonovich, R.L.: Information transmission rate in some quantum communication channels, Probl. Info. Transm., 2 (1966), 45-57.

13. Mityugov, V.V.: On quantum theory of information transmission, Probl. Info. Transm., 2 (1966), 48-58.

14. Ingarden, R.S.: Quantum information theory, Institut of Physics, N. Copernicus Univ., Torun, Poland, 1975.

15. Vainshtein, V.D. and Tvorogov, S.D.: Some problems of measurement of quantum observables and determination of joint entropy in quantum statistics, Comm. Math. Phys., 43 (1975), 273-278.

16. Drikker, A.S.: Homodine reception of a quantum electromagnetic signal, Probl. Info. Transm., 12 (1976), 57-68.

17. Poplavskii, R.P.: Thermodynamic models of information processes, Sov. Phys. Uspekhi, 115 (1975), 222-241.
18. Pierce, J.B., Posner, E.C. and Rodemich, E.C.: The capacity of the photon counting channel, IEEE Trans. on Info. Theory, IT-27 (1981), 61-67.
19. Kosloff, R.: Thermodynamic aspects of the quantum-mechanical measuring process, in: Advances in Chemical Physics, v. 46, eds. I. Prigogine, S.A. Rice, Wiley and Sons, 1981, 153-193.
20. Mitiugov, V.V.: Physical foundations of information theory, Sovietskoe Radio, Moscow, 1976.
21. Arthurs, E. and Kelly, J.L., Fr.: On the simultaneous measurement of a pair of conjugate observables, Bell System Tech. J., 44 (1965), 725-729.
22. Gordon, J.P. and Louisell, W.J.: Simultaneous measurement of non-commuting observables, in: Physics of Quantum Electronics (Eds. P. Kelley, M. Lax, B. Tannenwald), McGraw-Hill, New York 1966, 833-840.
23. Levitin, L.B., Mitiugov, V.V.: Reception of coherent signals by splitting of the beam, 1st Conf. on Problems of Information Transmission by Laser Radiation, Kiev, 1968.
24. Helstrom, C.W.: Detection theory and quantum mechanics, Inform. Contr. 10 (1967), 254-291.
25. Davis, E.B. and Lewis, J.T.: An operational approach to quantum probability, Comm. Math. Phys., 17 (1970), 239-260.
26. Benioff, P.A.: Decision procedures in quantum mechanics, J. Math. Phys., 13 (1972), 908-915.
27. Helstrom, C.W. and Kennedy, R.S.: Noncommuting observables in quantum detection and estimation theory, IEEE Trans. Inform. Theory, IT-20 (1974), 16-24.
28. Helstrom, C.W.: Quantum Detection and Estimation Theory, Academic Press, New York 1976.
29. Levitin, L.B.: Amount of information and the quantum-mechanical irreversibility of measurement, in: Proc. of the II Intern. Symp. on Inform. Theory, Yerevan, 1971, 144-147.
30. Holevo, A.S.: Informational aspects of quantum measurements, Probl. Info. Transm., 9 (1973), 31-42.
31. Holevo, A.S.: Certain estimates of information transmissible over a quantum communication channel, Probl. Info. Transm., 9 (1973), 3-11.
32. Davies, E.B.: Information and quantum measurement, IEEE Trans. on Inform. Theory, IT-24 (1979), 596-599.
33. Levitin, L.B.: Direct and indirect quantum measurements yield equal maximum information, 1981 IEEE Intern. Symp. on Information Theory, Santa Monica, CA., USA, 1981.
34. Jaynes E.T.: Information theory and statistical mechanics, Phys. Rev., Part I, 106 (1957), 620-630; Part II, 108 (1959), 171-190.
35. Katz, A.: Principles of Statistical Mechanics. The Information Theory Approach, W.H. Freeman, 1967.
36. Ingarder, R.S.: Information Theory and Thermodynamics. Part I, Torun, Poland, 1974. Part II, Torun, Poland, 1976.
37. Levitin, L.B.: Quantum amount of information and maximum work, in: Proc. of the 13th IUPAP Conf. on Statistical Physics (Eds. D. Cabib, D.G. Kuper and I. Riess), A. Hilger, Bristol, England, 1978.
38. Poplavskii, R.P.: Maxwell demon and correlations between information and entropy, Sov. Phys. Uspekhi, 128 (1979), 165-176.
39. Mityugov, V.V.: Entropy, information and work in quantum statistics, Probl. Control and Inform. Theroy, 2 (1973), 243-256.
40. Berger, T.: Communication theory via random fields, IEEE Intern. Symp. on Inform. Theory, St. Jovite, Quebec, Canada, 1983.
41. Levitin, L.B.: Information-theoretical approach to Ising problem, IEEE Intern. Symp. on Inform. Theory, Brighton, England, 1985.
42. Levitin, L.B.: A thermodynamic characterization of ideal physical information channels, Journal of Information and Optimization Sciences, 2 (1981), 259-266.
43. Levitin, L.B.: Information transmission in an ideal photon channel, Probl. Inform. Transm., 1 (1965), 55-62.
44. Levitin, L.B.: Ideal corpuscular information channels, IEEE Intern. Symp. on Inform. Theory, St. Jovite, Quebec, Canada, 1983.
45. Bennett, C.H., Brassard, G., Breidbart, S., and Weisner, S.: Quantum cryptography, or unforgeable subway tokens, Advances in Cryptology: Proceedings of Crypto '82, August 1982, Plenum, New York, 267-275.
46. Bennett, C.H., Bessette, F., Brassard, G., Salvail, L., and Smolin, J.: Experimental quantum cryptography, J. of Cryptology, 5 (1992), 3-28.
47 Bennett, C.H., Brassard, G., and Mermin, N.D.: Quantum cryptography without Bell's theorem, Phys. Rev. Letters, 68, No5 (1992), 557-559.