Physical Information Theory Part II: Quantum Systems

Lev B. Levitin

Boston University, Boston, MA

Abstract

The concept of information (by Shannon) is generalized for quantum statistical ensembles, and the consistent quantum-theoretical form of the entropy defect principle is obtained. The effect of the irreversiblility of quantum-mechanical measurements on information transmission is analyzed. For the ensemble of two pure quantum states, the optimal measurement procedure is found in the explicit form.

1. Introduction.

The quasiclassical form of the entropy defect principle [1] is valid when we can neglect the effect of noncommutativity of the density operators which describe quantum-mechanical ensembles. However, to obtain a consistent quantum-mechanical description of information transmission processes, we should generalize the concepts of entropy defect and Shannon's information on the basis of the quantum theory of measurements. This generalization has been first presented in [2].

In quantum mechanics a macrostate s of a physical system is described by a density operator (density matrix) $\hat{\rho}$ which is a self-adjoint nonnegative definite operator with $\mathrm{Tr}\,\hat{\rho}=1$ in a separable Hilbert space H. A microstate (pure state) of a physical system is described by a state vector (wave function) $\psi \in H$ of a length equal to 1, or, alternatively, by a density operator $\hat{\rho}$ of a special form, namely such that any matrix element $\rho_{nn'} = \psi_n \overline{\psi}_{n'}$, (the bar denotes complex conjugation). A complete set of observables is such that all the observables can take on simultaneously definite values, and a set of those values, determines uniquely the microstate of the physical system. A measurement whose outcome is a set of values of a complete set of observables is called a complete measurement. We consider here observables which take on a countable set of values. Then any complete direct quantum-mechanical measurement is associated with a certain countable orthonormal basis L in the space H, the probabilities of the m-th outcome being given by the diagonal element $\rho_{kk}(L)$ of the density matrix $\hat{\rho}$ in the basis L:

$$\rho_{kk}(L) = \mathrm{Tr}(\hat{\rho}\hat{P}_k) = \sum_n \sum_{n'} \rho_{nn'}(L)\delta_{kn'}\delta_{n'n} \qquad (1)$$

Here $\hat{P}_k$ is a projection operator, having in the basis L a form:

$$\|P_{k,nn'}(L)\| = \|\delta_{kn}\delta_{nn'}\| \qquad (2)$$

Obviously, $\hat{P}_k$ is a nonnegative definite operator, and the set of all $\hat{P}_k$ (k = 1,2,...) forms an orthogonal resolution of the identity $\hat{1}$ in the space H:

$$\hat{P}_k \geq \hat{0};\ \sum_k \hat{P}_k = \hat{1};\ \mathrm{Tr}\,\hat{P}_k\hat{P}_{k'} = 0\ (k \neq k') \qquad (3)$$

Note that each $\hat{P}_k$ is a density operator corresponding to a microstate m_k described by a state vector ψ_k which is one of the basis vectors: $\psi_k \in L$. Thus, the physical meaning of the m-th outcome of the mearsurement is that the system is found in the microstate m_k. A complete set of observables can include physical quantities which take on a continuum of values (e.g., coordinates or momenta of particles). In a separable Hilbert space, however, any measurement of continuous variables can be arbitrarily closely approximated by measurements with a countable set of outcomes. Therefore henceforth we restrict ourselves to the measurements of type (3).

The entropy of a quantum system in a macrostate s described by a density matrix $\hat{\rho}$ is defined as follows [3,4]

$$H = -\mathrm{Tr}\,\hat{\rho}\ln\hat{\rho} \qquad (4)$$

A quantum-mechanical measurement changes the state of a system in an irreversible way: the entropy of the ensemble of microstates obtained as a result of a measurement is, generally speaking, larger than the entropy of the initial state. Namely, by Klein's lemma [3,4]

0-8186-3420-0/93 $03.00

$$-\operatorname{Tr}\hat{\rho}\ln\hat{\rho} \le -\sum_k \rho_{kk}(L)\ln\rho_{kk}(L) \qquad (5)$$

where the equality holds if $\hat{\rho}$ is diagonal in basis L.

2. Entropy Defect Principle (Quantum Formulation).

Consider now an ensemble $S=\{s_i,p_i\}$ of macrostates of a physical system, each macrostate s_i occuring with probability p_i and being described by a density matrix $\hat{\rho}^{(i)}$. (The preparation of a certain macrostate s_i can be interpreted as "physical encoding" of a signal x_i taken from an ensemble of signals $X = \{x_i,p_i\}$). The a priori state of the system s chosen at random from the ensemble S is described by a density matrix $\hat{\rho}$:

$$\hat{\rho} = \sum_i p_i \hat{\rho}^{(i)} \qquad (6)$$

Definition 1. The entropy defect of a system described by the ensemble of macrostates S is the quantity

$$I_0 = -\operatorname{Tr}\hat{\rho}\ln\hat{\rho} + \sum_i p_i \operatorname{Tr}\hat{\rho}^{(i)}\ln\hat{\rho}^{(i)} \qquad (7)$$

Entropy defect characterizes the average decrease of the entropy of the system, when it becomes known, which macrostate (out of the ensemble S) has been chosen.

Theorem 1.

$$0 \le I_0 \le -\sum_i p_i \ln p_i \qquad (8)$$

The lefthand equality holds iff all $\hat{\rho}^{(i)}$ are identical, the right-hand equality holds iff all $\hat{\rho}^{(i)}$ are orthogonal (i.e. $\operatorname{Tr}\hat{\rho}^{(i)}\hat{\rho}^{(i')}=0,\ i \ne i'$).

Thus, entropy defect does not exceed the entropy of the signal

$$H(X) = -\sum_i p_i \ln p_i$$

As it has been shown in [1], in the quasiclassical case, the entropy defect is equal to the information about the macrostate of the system obtained by measuring its microstate. The situation is different and more complex in quantum theory. Information about the macrostate (or about the signal, since the random macrostate S and signal X are in one-to-one correspondence) depends on the choice of the complete set of observables to be measured.

Using Shannon's concept of information, we come to the following definition:

Definition 2. Information about the macrostate S of a physical system obtained in a complete measurement associated with a countable orthonormal basis L in the Hilbert space *H* is the quantity

$$I_L = -\sum_k \rho_{kk}(L)\ln\rho_{kk}(L) + \sum_i\sum_k p_i\,\rho_{kk}^{(i)}(L)\ln\hat{\rho}_{kk}^{(i)}(L) \qquad (9)$$

Here $\rho_{kk}(L)$ and $\rho_{kk}^{(i)}(L))$ are diagonal elements of $\hat{\rho}$ and $\hat{\rho}^{(i)}$, respectively, in basis L.

Since the quantity I_L depends on L, it determines not an absolute, but a conditional maximum of amount of data transmittable over a quantum channel under a fixed choice of L.

Definition 3. Information about the macrostate S of a physical system obtainable by direct measurements is the quantity

$$I = \sup_L I_L \qquad (10)$$

where the least upper bound is taken over all possible choices of L (or, in other words, over all possible orthogonal resolutions of identity in *H*).

The quantity I plays the same role in information transmission as information by Shannon in the classical theory. (Note that here we restrict ourselves to direct measurements only, i.e. to the usual quantum-mechanical measurements in the sense of von Neumann [4] performed in the Hilbert space *H* of the system. The problem of so called indirect (or generalized) measurements is considered in [5]). Indeed, the following theorem is valid:

Theorem 2. Let $X(\tau)=\{x_i(\tau),p_i(\tau)\}$ be a set of signals of duration τ which can be transmitted over a channel, the signal $x_i(\tau)$ being used with probability $p_i(\tau)$. Let $S(\tau)=\{s_i(\tau),p_i(\tau)\}$ be a set of macrostates of a physical system carrying the information, each macrostate $s_i(\tau)$ corresponding to the signal $x_i(\tau)$ and being described by a density matrix $\rho^{(i)}(\tau)$. Let $I(\tau)$ be defined according to (10) with $p_i(\tau)$ and $\rho^{(i)}(\tau)$ substituted for p_i and $\rho^{(i)}$, respectively. Then the capacity C of the channel is equal to

$$C = \lim_{\tau\to\infty} \frac{I(\tau)}{\tau} \qquad (11)$$

(It implies, of course, that the limit exists).

A crucial question is, what is the relation between information I and entropy devect I_0.

Theorem 3.

$$I \leq I_0 \tag{12}$$

and the equality holds iff all the matrices $\hat{\rho}^{(i)}$commute. In this case

$$I = I_{L_0} = I_0 \tag{13}$$

for a basis L_o in which all $\hat{\rho}^{(i)}$are diagonal.

Theorem 3 is the quantum-mechanical counterpart of the entropy defect principle (cf. [1]). It shows that the irreversibility of a quantum measurement results in inevitable loss of information (except for the case when all the density matrices are diagonal, and the quasiclassical approximation is applicable).

Note that I can be written in a form:

$$I_0 = -\sum_n \lambda_n \ln\lambda_n + \sum_i \sum_n p_i \lambda_n^{(i)} \ln\lambda_n^{(i)} \tag{14}$$

where $\lambda_n$and $\lambda_n^{(i)}$are eigenvalues of $\hat{\rho}$ and $\hat{\rho}^{(i)}$ respectively, i.e. the probabilities of orthogonal (i.e. perfectly distinguishable) microstates whose mixtures are the macrostates described by $\hat{\rho}$ and $\hat{\rho}^{(i)}$. (But, of course, when $\hat{\rho}^{(i)}$do not commute, $\lambda_n \neq \sum_i p_i \lambda_n^{(i)}$.)

Thus, I_0 can be interpreted as information in the macrostate about the microstate. On the other hand, I is the information in the micrtostate (specified by the measurement) about the macrostate. Since, in general, $I \neq I_0$ it means that the quantum measurement breaks the symmetry between the input and the output of a channel (in the classical theory I(X;Y) = I(Y;X), where X and Y are input and output variables, respectively).

Suppose now that the information carrier is a closed system with a Hamiltonian $\hat{H}$. Then the time evolution of the system is described by a unitary transformation

$$\hat{\rho}_t^{(i)} = \exp(-\frac{i\hat{H}t}{\hbar})\hat{\rho}_0^{(i)} \exp(\frac{i\hat{H}t}{\hbar}) \tag{15}$$

where $\hat{\rho}_0^{(i)}$ and $\hat{\rho}_t^{(i)}$ are density matrices corresponding to the signal x_i at the initial moment of time and at the moment t, respectively. Since both I and I_0 are unitary invariant, the following theorem is valid:

Theorem 4.

$$\frac{dI}{dt} = \frac{dI_0}{dt} = 0. \tag{16}$$

Thus, information and entropy defect are integrals of motion for a closed system. One should not forget, however, that the operators associated with the optimal measurement changes in time in the same way as the density matrices. Therefore, generally speaking, the optimal measurement becomes more and more complicated with the increase of time.

3. Ensemble of Two Pure Quantum States

In the general case it seems not possible to derive an explicit expression for I or to determine explicitly a basis L for which I_L=I (when such a basis exists). This is a difficult problem even in simple special cases. (Some related results have been obtained in [6]). In this section explicit expressions will be found for the amount of information I and for an optimal basis L in the simple case of two pure states. We shall also present some resluts for the case of two mixed states described by second-order matrices (such as spin polarization matrices).

Consider a quantum system which has a probability p of being in a state $\hat{\rho}^{(1)}$ and a probability 1-p of being in a state $\hat{\rho}^{(2)}$. Suppose that these are pure states, i.e., their density matrices $\hat{\rho}^{(1)}$, $\hat{\rho}^{(2)}$ are of the form

$$\rho_{nn'}^{(1)} = \psi_n^{(1)} \overline{\psi}_{n'}^{(1)} \; ; \; \rho_{nn'}^{(2)} = \psi_n^{(2)} \overline{\psi}_{n'}^{(2)} \tag{17}$$

where $\psi^{(1)}$, $\psi^{(2)}$ are wave functions (state vectors). The following theorem holds.

Theorem 5. The information $I_L$for an ensemble of two pure states, occurring with probabilities p_1=p, p_2=1-p, achieves its maximum I_L=I in a basis L= $\{\phi_1\}$ (i=1, 2, ...) such that two basis vectors (say ϕ_1 and ϕ_2) lie in the plane spanned by the state vectors $\Psi^{(1)}$, $\Psi^{(2)}$ and the following condition holds: $\hat{\sigma} = p\hat{\rho}^{(1)} - (1-p)\hat{\rho}^{(2)}$is a diagonal matrix.

This result means that in the optimal basis L only two components of the vectors $\Psi^{(1)}$ and $\Psi^{(2)}$are different from zero (say $\psi_1^{(i)}, \psi_2^{(i)}$) and moreover

$$p\psi_1^{(1)} \overline{\psi}_2^{(1)} = (1-p)\psi_1^{(2)} \psi_2^{(2)} \tag{18}$$

In terms of density matrices, this means that in the optimal basis

$$\rho_{\ell\ell'}^{(1)} = \rho_{\ell\ell'}^{(2)} = 0 \quad \text{for } \ell > 2,\ \ell' > 2$$

$$p\rho_{\ell\ell'}^{(1)} = (1-p)\rho_{\ell\ell'} \quad \text{for } \ell \neq \ell' \qquad 19$$

Since the phase of a wave function (a constant factor of the form $e^{i\alpha}$) can be chosen arbitrarily (it does not affect the form of the density matrix), condition (18) implies that the phases of states $\psi^{(1)}$, $\psi^{(2)}$ and the basis vectors $\phi 1$ and $\phi 2$ can be chosen in such a way that all the components of $\psi^{(1)}$, $\psi^{(2)}$ in this bais are real. Condition (18) may then be given a simple geometric meaning. Let (x_1, x_2) be the Cartesian coordinate system formed by the basis vectors ϕ_1, ϕ_2. Then the endpoints of the vectors $\sqrt{p\psi^{(1)}}$ and $\sqrt{1-p\psi^{(2)}}$ must lie on a hyperbola x_1x_2=const (see Fig. 1).

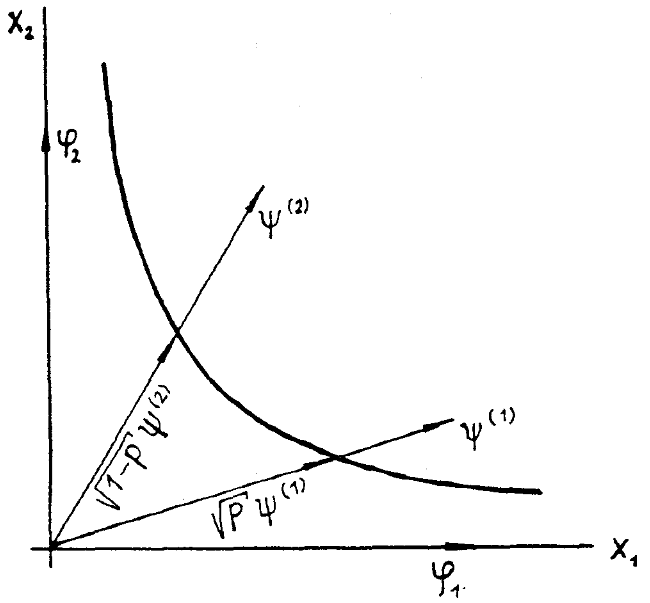

Fig. 1. Two nonorthogonal pure quantum states and optimal basis.

For two pure states, the values of I_0 and I depend only on the probability p and on the only joint invariant of $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$ - the trace r of their product:

$$Tr\hat{\rho}^{(1)}\hat{\rho}^{(2)} = r = |s|^2$$

where $s = <\psi^{(1)}, \psi^{(2)}>$ is the scalar product of the wave vectors $\psi^{(1)}$ and $\psi^{(2)}$. The calculations give the following results.

$$I_0 = -Tr\hat{\rho}\ln\hat{\rho} =$$
$$\ln 2 - 1/2[(1+\sqrt{1-4p(1-p)(1-r)})\ln(1+\sqrt{1-4p(1-p)(1-r)}) +$$
$$+(1-\sqrt{1-4p(1-p)(1-r)})\ln(1-\sqrt{1-4p(1-p)(1-r)})] \qquad (20)$$

$$I = 1/2\{p[(\sqrt{1-4p(1-p)r}+1-2(1-p)r)\ln(\sqrt{1-4p(1-p)r}+1-2(1-p)r)+$$
$$+(\sqrt{1-4p(1-p)r}-1+2(1-p)r)\ln(\sqrt{1-4p(1-p)r}-1+2(1-p)r)]+$$
$$+(1-p)[(\sqrt{1-4p(1-p)r}-1+2pr)\ln(\sqrt{1-4p(1-p)r}-1+2pr)+$$
$$+(\sqrt{1-4p(1-p)r}+1-2pr)\ln(\sqrt{1-4p(1-p)r}+1-2pr)]-$$
$$-(\sqrt{1-4p(1-p)r}+1-2p)\ln(\sqrt{1-4p(1-p)r}+1-2p)-$$
$$-(\sqrt{1-4p(1-p)r}-1+2p)\ln(\sqrt{1-4p(1-p)r}-1+2p)\}\bullet$$
$$\bullet(1-4p(1-p)r)^{-\frac{1}{2}} \qquad (21)$$

In the general case, the formulae for I_0 and I are very cubersome. We, shall compare them for the case p=1/2. It is convenient to introduce the angle α between the vectors $\psi^{(1)}$ and $\psi^{(2)}$; then $r=\cos^2\alpha$. In this case the vectors $\psi^{(1)}$ and $\psi^{(2)}$ are symmetric about the bisectrix of the angle between the basis vectors ϕ_1, ϕ_2:

$$I_0 = \ln 2 - \frac{1}{2}(1+\cos\alpha)\ln(1+\cos\alpha) +$$
$$+(1-\cos\alpha)\ln(1-\cos\alpha) \qquad (22)$$

$$I = \frac{1}{2}(1+\sin\alpha)\ln(1+\sin\alpha) +$$
$$+(1-\sin\alpha)\ln(1-\sin\alpha) \qquad (23)$$

The following interesting relationship is noteworthy:

$$I(\frac{\pi}{2} - \alpha) = \ln 2 - I_0(\alpha)$$

The difference I_0-I achieves its maximum at $\alpha = \pi/4$. In that case I_0+I=ln2 and I≈2/3 I_0 Plots of I_0 and I for the case p=1/2 are shown in Fig. 2.

In the general case of two mixed states described by second-order density matrices (for example, different spin states) the quantities I_0 and I depend on invariants of the matrices $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$: their determinants d_1 and d_2 and the trace r of their product. In the special case when $d_1=d_2=d$ and p=1/2 we obtain:

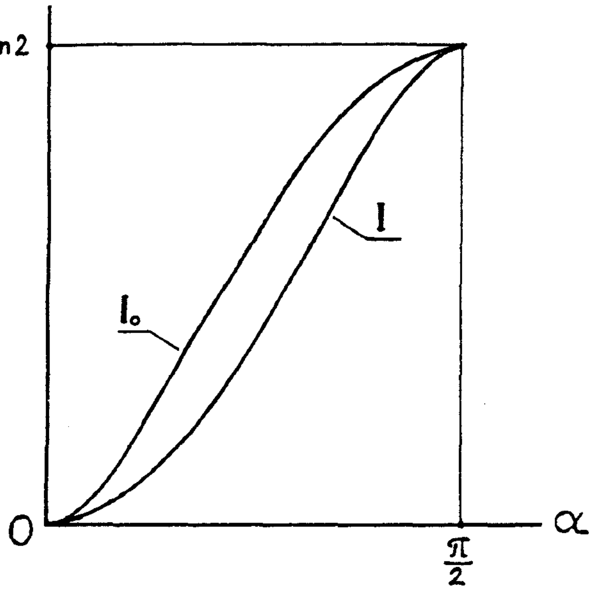

Fig. 2 Entropy defect and information for two pure quantum states, p=1/2. The point $(\frac{\ln 2}{2}, \frac{\pi}{4})$ is the center of symmetry of the graph.

$$I_0 = \frac{1}{2}[(1+\sqrt{1-4d})\ln(1+\sqrt{1-4d}) + (1-\sqrt{1-4d})\ln(1-\sqrt{1-4d}) - (1+\sqrt{r-2d})\ln(1+\sqrt{r-2d}) - (1-\sqrt{r-2d})\ln(1-\sqrt{r-2d})]. \quad (24)$$

$$I = \frac{1}{2}[(1+\sqrt{1-r-2d})\ln(1+\sqrt{1-r-2d}) + (1-\sqrt{1-r-2d})\ln(1-\sqrt{1-r-2d}) \quad (25)$$

For fixed r, the values of I_0 and I decrease with increasing d.

It is interesting that in the case of two pure states the condition for maximum information coincides with the condition for optimal detection that guarantees the minimum average error probability [7,8]. However, this is not true in the general case of two mixed states.

4. Conclusion.

Information theory for quantum systems is based on two imporatnt concepts: the entropy defect and the information obtainable in quantum measurements. The consistent quantum-mechanical formulation of the entropy defect principle (Secion 2) shows that these two quantities, being identical in the classical case, split apart from each other in the realm of quantum theory, which reflects the effect of the irreversibility caused by a quantum-mechanical measurement. This expresses the fundamental limitations imposed on information transmission and storage by quantum nature of information carriers even in the absence of any external noise.

The results of Section 3 provide the only known at present example of explicit expressions for information and the optimal basis for the case of states described by noncommuting density matrices. The optimal procedure is proved to be a direct measurement, which anticipates the general result presented in [5].

References.

1. Levitin, L.B.: Physical information theory. Part I. Quasiclassical systems, in this book.
2. Levitin, L.B.: On the quantum measure of information, Proc. of the IV National Conf. on Information Theory, Tashkent, 1969, 111-115.
3. Klein, O.: Zur quantenmechanischen Begründung des zweiten Hauptsatzes der Warmelehre, Z. Phys., 72 (1931), 767-775.
4. Neumann, J.: Mathematische Grundlagen der Quantenmechanik, Springer Verlag, Berlin, 1932 (English translation: Mathematical Foundations of Quantum Mechanics, Princeton Univ. Press, Princeton, NJ, USA, 1955).
5. Levitin, L.B.: Information in direct and indirect quantum measurements, in this book.
6. Davies, E.B.: Information and quantum measurement, IEEE Tran. on Inform. Theory, IT-24 (1979), 596-599.
7. Helstrom, C.W.: Detection theory and quantum mechanics, Inform. Contr., 10 (1967), 254-291.
8. Helstrom, C.W.: Quantum Detection and Estimation Theory, Academic Press, New York, 1976.