The Two Extremes of Information in Quantum Mechanics

William K. Wootters

Department of Physics, Williams College
Williamstown, MA 01267

In the problem of estimating the amount of information one can extract from a quantum system, the entropy S of the system is known to play a special role: it is an upper bound on the accessible information. This paper calls attention to another property of quantum systems, called Q, which appears to play an analogous role in defining a lower *bound on the accessible information. Some parallels between S and Q are noted.*

Suppose one is presented with a physical system whose quantum state one does not know. The simplest example is a single photon that could be in any of a number of different polarization states. Such a system might be a signal in a quantum communication scheme or the output of a quantum computer, or it might come from a natural source. In any case, let us assume that one's goal is to extract as much information as possible about the system's state. The first question I want to address is this: How much information can one extract? The quantity of interest here is the difference between the amount of information [1] one lacks about the system's state prior to the measurement and the expectation value of the amount of information one will lack at the end of the measurement. The maximum possible value of this difference has rightly been called the "accessible information." [2]

The problem of extracting information from a quantum system is not trivial, because there are always infinitely many possible measurements that one can make, and these measurements are mutually exclusive. Moreover, performing any one measurement typically destroys information about the outcomes of other measurements that could have been performed. One therefore wants to choose one's measurement carefully. No general algorithm is known for finding an optimal measurement, nor is any formula known for computing the amount of accessible information. [3]

There is, however, one central theorem on this subject, due to Levitin [4] and Kholevo [5], based on what is known as the *density matrix* of the system. In the sort of problem we are considering, what is given initially is a set of possible quantum states (e.g., polarization states) along with their *a priori* probabilities. Let us call such a description a "state distribution." For any given state distribution, there is a corresponding density matrix, which is a much more compact description of the same system. [6] For example, if for a given photon there are 100 possible pure polarization states having non-zero probability, then it takes 299 independent real numbers to specify the state distribution (two numbers to specify each state and another number to specify its probability; but the probabilities must add to unity; hence 299), whereas the density matrix is a simple 2x2 matrix. Clearly the density matrix is a less detailed description, but for many purposes it is the only description one needs. In particular, in order to compute the probabilities of the outcomes of any future measurement on the system, one does not need to know the details of the state distribution but only the density matrix. For this reason the density matrix is widely used in applications of quantum mechanics.

The Levitin-Kholevo theorem states that the accessible information is never greater than the entropy $S[\rho]$ of the system's density matrix ρ. This entropy is defined as

$$S[\rho] = -\sum_i p_i \log p_i,$$

where the p_i's are the eigenvalues of ρ. These eigenvalues have a simple physical meaning. They are the probabilities of the outcomes of that measurement whose outcome one can predict the best. In other words, $S[\rho]$ measures the minimum amount of information one lacks about the outcome of any measurement. For thermodynamic systems, $S[\rho]$ is normally interpreted as the thermodynamic entropy. [7] Thus the Levitin-Kholevo theorem provides an important link between thermodynamic entropy and accessible information.

Two comments on the Levitin-Kholevo theorem will be helpful. First, the upper bound given in the theorem depends only on the density matrix, while the actual accessible information depends on details of the state distribution that are not contained in the density matrix. A typical density matrix is consistent with many different state distributions, and some of these distributions permit one to extract more information than others. For example, a photon that is known to have either vertical or horizontal polarization, with equal probability, has the same density matrix as a photon about which one knows only that it is in *some* pure state. Yet in the former case one can extract 1 bit of information from the photon, while in the latter case one can extract only 0.28 bits. The second comment is this: the Levitin-Kholevo bound is the *best* possible upper bound that depends only on the density matrix. This is because for any given density matrix ρ, one can find a state distribution consistent with ρ for which the accessible information equals $S[\rho]$.

I now come to the new part of this paper. Just as there is a best *upper* bound depending only on the density matrix, there must also be a best *lower* bound depending only on the density matrix. Only by having both bounds can one know the full range of accessible information consistent with a given density matrix. I wish to propose a formula for this best lower bound. The formula has not been proved correct, but the evidence suggests strongly that it is. The conjectured lower bound is $Q[\rho]$, defined by

$$Q[\rho] = -\sum_i \left[\left(p_i \log p_i\right) \prod_{j\neq i} \frac{p_i}{p_i - p_j} \right].$$

Here the p_i's are again the eigenvalues of the density matrix. (If two or more of the p_i's are equal, then Q is defined as the limit of the above expression.) Specifically, the conjecture is this: For any distribution of pure states having density matrix ρ, the accessible information I satisfies

$$Q[\rho] \leq I \leq S[\rho];$$

moreover for any density matrix ρ, there exists a distribution of pure states consistent with ρ for which $I = Q[\rho]$. One can show that the conjecture is true for a wide variety of cases, and this fact is what leads me to think that the conjecture is probably true in general.

The above inequality depends on the assumption that the actual state of the system is a pure state, the density matrix arising only from one's ignorance of the state. It is also possible to consider cases where the set of possible states includes mixed states as well as pure states. For example, perhaps one knows that a given photon is in one of the following three polarization states: vertical, right-circular, or unpolarized. The last of these is a mixed state, so the above inequality cannot be used directly to estimate the amount of information one could gain in this situation.

When mixed states are allowed, the Kholevo-Levitin theorem takes the modified form [4,5]

$$I \leq S[\rho] - \sum_j P_j S[\rho_j],$$

where ρ_j is the density matrix describing the j^{th} state of the initial state distribution, and P_j is its *a priori* probability. What the modification says is this: if some of the states one is trying to distinguish are mixed states, then their presence in the ensemble tends to reduce the maximum amount of information one can hope to extract.

Similarly, the conjectured *lower* bound must also be modified when mixed states are allowed. Remarkably, the evidence so far suggests that the appropriate modification is of exactly the same form as the above modification of the upper bound. That is, it seems that the correct lower bound may be given by

$$I \geq Q[\rho] - \sum_j P_j Q[\rho_j].$$

The similarity between the two bounds, if the latter is indeed correct, is somewhat surprising because Q, unlike S, is not additive when one combines independent probability distributions. And yet the above equation manifests a certain kind of additivity. My main reason for presenting these modified bounds is to emphasize the apparent similarity of the roles played by S and Q.

Supposing that the conjecture is correct, what consequences does it have and what questions does it raise? The first thing to note is that Q never exceeds a particular finite value--it happens to be 0.60995 bits--whereas S can get arbitrarily large. This means that the accessible information is not specified very precisely by the density matrix alone, being allowed to range over what is typically a rather large interval. The largeness of the interval makes it desirable to find an alternative way of estimating the accessible information, based on something other than the density matrix. Work is currently being done on that problem, but that is not the issue I want to emphasize here. The density matrix remains an important concept in these matters, and my main point is that S and Q seem to be the two characteristics of the density matrix that are relevant for the purpose of estimating the accessible information.

Another interesting parallel between S and Q is worth noting. S can be interpreted as the loss of information about the outcome of the most predictable measurement, owing to the fact that one does not have a single pure-state description of the quantum system. Similarly, Q can be interpreted as the loss of information about the outcome of a *typical* measurement owing to the same lack of knowledge. [8,9] (The estimation of information about measurement outcomes is the context in which Q has appeared in the literature previously.) Thus both of these quantities, S and Q, are of importance not only in the problem of inferring past states but also in the problem of predicting future outcomes.

One might ask whether the parallel between S and Q extends into the thermodynamic realm. If S is indeed the thermodynamic entropy, does Q likewise play a role, hitherto unknown, in thermodynamics? If it does it must be a very subtle role, if only because the magnitude of Q is always small compared to macroscopic quantities. But the question has not been studied in any detail and is certainly worth investigating.

Finally, it is interesting to ask whether the above formula for Q has any meaning outside of quantum mechanics. The formula has been derived specifically within the quantum context, and it would be different if the world operated according to a different set of laws (e.g., a quantum mechanics based on the real numbers instead of the complex numbers). If Q has a significance outside of quantum mechanics, then this alternative interpretation of Q would almost certainly shed light on the logical origin of quantum mechanics itself.

1. C. E. Shannon, *Bell Syst. Tech. J.* **27** (1948) 379, 623.
2. B. Schumacher, "Information from Quantum Measurements," in *Complexity, Entropy, and the Physics of Information*, ed. by W. H. Zurek (Addison-Wesley, 1990).
3. L. B. Levitin, "Information Theory for Quantum Systems," in *Information, Complexity, and Control in Quantum Physics*, ed. by A. Blaquière, S. Diner, and G. Lochak (Springer, Vienna, 1987).
4. L. B. Levitin, "On the quantum nature of information," Proceedings of the Fourth All-Union Conference on Information and Coding Theory, Sec. II (Tashkent, 1969), pp. 111-5.
5. A. S. Kholevo, *Probl. Inform. Transmission* **9** (1973), 177 (translated from *Problemy Peredachi Informatsii*).
6. B. H. Bransden and C. J. Joachain, *Introduction to Quantum Mechanics* (Longman, Essex 1989), pp. 601-9.
7. A. Peres, "Thermodynamic Constraints on Quantum Axioms," in *Complexity, Entropy, and the Physics of Information*, ed. by W. H. Zurek (Addison-Wesley, 1990).
8. W. K. Wootters, *Found. Phys.* **20** (1990), 1365.
9. K. R. W. Jones, *J. Phys. A: Math. Gen.* **24** (1991), 1237-44.