Illustrating the concept of quantum information

R. Jozsa

Over the past decade, quantum information theory has developed into a vigorous field of research despite the fact that quantum information, as a precise concept, is undefined. Indeed, the very idea of viewing quantum states as carriers of some kind of information (albeit unknowable in classical terms) leads naturally to interesting questions that might otherwise never have been asked, and corresponding new insights. We discuss some illustrative examples, including a strengthening of the well-known no-cloning theorem leading to a property of permanence for quantum information, and considerations arising from information compression that reflect on fundamental issues.

Introduction

Perhaps the most intriguing product of quantum information theory is the concept of quantum information itself. In the early 1990s Charles Bennett was one of the first workers to recognize and promote this new concept, establishing the foundations of a new subject. Taken as a primary ingredient, quantum information cannot be defined, but the viewpoint it fosters is richly suggestive, leading to new interesting questions and modes of interpretation for some quantum processes. In this paper we explore a few examples.

A quantum state $|\psi\rangle$ may be viewed as a carrier of information in two fundamentally different ways. First, $|\psi\rangle$ may be regarded as carrying the *classical* information of the state identity. As an example, a sender may prepare one of the two (non-orthogonal) states $|\psi_0\rangle$ and $|\psi_1\rangle$ to encode the bit values 0 and 1, respectively. Then the receiver's task is to regain the value of i from $|\psi_i\rangle$. If $p_{j|i}$ denotes the probability that he generated the output j when the state was $|\psi_i\rangle$ and q_i is the probability that $|\psi_i\rangle$ was sent, then he may, for example, apply a procedure that minimizes the error probability $p_{2|1}q_1 + p_{1|2}q_2$. In this way, the available information in the quantum state is similar to the result of classical communication through a noisy channel; it is well known that if $\langle\psi_1|\psi_2\rangle \neq 0$, the minimum error probability cannot be zero; i.e., the state $|\psi_i\rangle$ cannot be perfectly identified by any physical process.

In a second way, $|\psi\rangle$ may be viewed as the carrier of "quantum information" which, although we leave it undefined in more fundamental terms, we intuitively think of as "the state itself." Quantum information is a new concept with no classical analog, and it is important to distinguish it from the state identity. For example, given a physical realization of one of the two states $|\psi_i\rangle$ above, quantum theory considerably restricts (in a richly structured way) the allowable manipulations that we can perform, in contrast to what is possible if we are given the identity of i. Indeed, "being given the quantum state $|\psi_i\rangle$" is very different from being given any kind of classical information, and by an analogy of terminology we apply the phrase *quantum information* to describe what we have received. In more formal terms, we would aim to formulate and interpret quantum physics in a way that has a concept of information as a primary fundamental ingredient. Primary fundamental concepts are *ipso facto* undefined (as a definition amounts to a characterization in yet more fundamental terms) and they acquire meaning only afterward, from the structure of the theory they support.

As a first example, consider the process of quantum teleportation (cf. [1] for details): Alice (A) succeeds in transferring a qubit state to Bob (B) (distantly separated in space) by sending only two classical bits of information.

Note: This paper is based on a talk presented in May 2003 at a symposium at the IBM Thomas J. Watson Research Center in Yorktown Heights, New York, honoring Charles Bennett on the occasion of his sixtieth birthday.

0018-8646/04/$5.00 © 2004 IBM

A and B also need to share an entangled Einstein–Podolsky–Rosen (EPR) pair which is destroyed in the process. We would like to think of teleportation as the transmission of quantum information from A to B. If we accept the intuitively appealing tenet that a transfer of information from sender to receiver must always be mediated by a channel connecting the two participants, teleportation appears paradoxical: If only two classical bits were sent, how did the full quantum information pass from A to B? Looking at the standard space–time diagram of the teleportation process (cf. Figure 1(a) of [1]), we see that there is indeed a second (V-shaped) path connecting A to B, which is defined by the two world lines of the distributed EPR particle pair. This leads to an intriguing interpretation (first proposed by Bennett soon after the discovery of teleportation): In addition to the two bits, the remaining quantum information must have been propagated *backward in time* from A to the EPR source and thence forward in time to B.[1] Indeed, if we insist that information transmission requires a physical channel, there appears to be no other possible interpretation of the teleportation process! It is remarkable that this interpretation is entirely consistent: The principles of quantum measurement theory imply that the information sent backward in time is random and independent of the teleported state, so long as the two classical bits remain unknown. Hence, the well-known classical causal paradoxes of backward-in-time information propagation are neatly circumvented. This analysis, inspired by our informational point of view, also reveals a new significance for entanglement in quantum theory (beyond the traditional issues of nonlocal correlations of measurement outcomes): Entanglement can be viewed as providing a channel for the transmission of quantum information.

In the following sections we discuss two further issues in which an informational point of view leads to interesting considerations. First we revisit the quantum no-cloning theorem [3] and prove a new stronger form of this result. Together with the Pati–Braunstein no-deleting principle [4], this leads to a property of "permanence" for quantum information. Second, we discuss the concept of information compression. In classical information theory, this provides one of the clearest approaches to the concept of information. By mimicking this theory in a quantum context, we obtain some surprising relationships between the concept of information and the geometry of Hilbert space (i.e., the basis of the conventional formulation of quantum theory).

The idea that a concept of information should be regarded as a fundamental ingredient in physical theory has also been proposed by Horodecki et al. [5] from a different point of view. In that work a notion of quantum information is based on the presence of entanglement in (multipartite) quantum systems, leading to an information conservation principle (corresponding to the fact that in a closed system, entanglement cannot be changed by local operations). In [6] the no-cloning and no-deleting principles are discussed from this point of view. However, this notion of quantum information and its conservation (involving the extra ingredients of locality and entanglement) appear not to be evidently related to the property of permanence that we discuss below.

A stronger no-cloning theorem

It is well known that (non-orthogonal) pure quantum states cannot be cloned [3]; i.e., if $\{|\psi_i\rangle\}$ is a set of pure states containing at least one non-orthogonal pair, no physical operation can achieve the transformation $|\psi_i\rangle \rightarrow |\psi_i\rangle|\psi_i\rangle$. Although the impossibility of cloning in quantum theory can be attributed to the fact that such a process is nonunitary or nonlinear, from an informational point of view we can intuitively understand it by saying that two copies of a quantum state embody strictly more "information" than is available in just one copy, so cloning must be impossible. Extending this particular line of thought, it is then natural to go on to ask: What additional (quantum) information is needed to supplement one copy $|\psi_i\rangle$ in order to be able to produce two copies $|\psi_i\rangle|\psi_i\rangle$? For classical information, no supplementary information at all is needed, and one might guess that as the set $\{|\psi_i\rangle\}$ becomes "more classical," the necessary supplementary information should decrease in some suitable way. However, we prove below that this is not the case: We show (for mutually non-orthogonal states) that the supplementary information must always be as large as it can possibly be; i.e., the second copy $|\psi_i\rangle$ can always be generated from the supplementary information alone, independently of the first (given) copy. Thus, in effect, cloning of $|\psi_i\rangle$ is possible only if the second copy is fully provided as an additional input.

We now give a precise formulation of the main result in this section. By a physical operation we mean a trace-preserving, completely positive map. Note that this definition excludes the collapse of wavefunction in a quantum measurement, as a valid physical process. (This will be relevant to our later discussion of the no-deleting principle.) By an abuse of notation for pure states, we write $|\psi\rangle\langle\psi| \otimes \rho$ as just $|\psi\rangle \otimes \rho$ and sometimes also omit the tensor product symbol, writing $|\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|$ as $|\psi\rangle|\psi\rangle$.

Theorem 1

Let $\{|\psi_i\rangle\}$ be any finite set of pure states containing no orthogonal pairs of states. Let $\{\rho_i\}$ be any other set of

[1] A similar interpretation involving propagation backward in time was proposed earlier for the Bennett–Wiesner dense coding protocol in [2] and attributed there to B. Schumacher, Kenyon College, Gambier, OH. Further developments of this idea by Schumacher are unpublished.

(generally mixed) states indexed by the same labels. Then there is a physical operation

$$|\psi_i\rangle \otimes \rho_i \rightarrow |\psi_i\rangle|\psi_i\rangle$$

if and only if there is a physical operation

$$\rho_i \rightarrow |\psi_i\rangle;$$

i.e., the full information of the clone must already be provided in the ancillary state ρ_i alone.

Remark[2]

If the set $\{|\psi_i\rangle\}$ contains some orthogonal pairs, the unassisted clonability of orthogonal states spoils the simplicity of the statement of Theorem 1. As an example, consider

$$|\psi_1\rangle = |0\rangle, \qquad |\alpha_1\rangle = |a\rangle,$$
$$|\psi_2\rangle = |1\rangle, \qquad |\alpha_2\rangle = |a\rangle,$$
$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \qquad |\alpha_3\rangle = |b\rangle,$$

where $|a\rangle$ and $|b\rangle$ are orthogonal. Then clearly $|\psi_i\rangle|\alpha_i\rangle \rightarrow |\psi_i\rangle|\psi_i\rangle$ is possible (because $\{|\psi_i\rangle|\alpha_i\rangle\}$ is an orthonormal set), but $|\alpha_i\rangle \rightarrow |\psi_i\rangle$ is not possible (since $|\alpha_1\rangle = |\alpha_2\rangle$ but $|\psi_1\rangle \neq |\psi_2\rangle$). Indeed, the $|\alpha_i\rangle$ states here provide reliable distinguishability of i values exactly when this is not already provided by the $|\psi_i\rangle$ themselves.

To prove the theorem we use the following lemma, which is proved as Lemma 1 of [7].

Lemma 1

Let $\{|\alpha_i\rangle\}$ and $\{|\beta_i\rangle\}$ be two sets of pure states (indexed by the same labels). Then the two sets have equal matrices of inner products (i.e., $\langle\alpha_i|\alpha_j\rangle = \langle\beta_i|\beta_j\rangle$ for all i and j) if and only if the sets are unitarily equivalent (i.e., there exists a unitary operation U on the direct sum of the state spaces of the two sets with $U|\alpha_i\rangle = |\beta_i\rangle$ for all i).

Proof of Theorem 1

Suppose that there is a physical operation $\rho_i \rightarrow |\psi_i\rangle$. Then clearly $|\psi_i\rangle \otimes \rho_i \rightarrow |\psi_i\rangle|\psi_i\rangle$ is allowed.

Conversely, suppose that there is a physical operation

$$|\psi_i\rangle \otimes \rho_i \rightarrow |\psi_i\rangle|\psi_i\rangle. \tag{1}$$

Consider first the case that ρ_i are pure states, which we write as $|\alpha_i\rangle$. The physical operation [Equation (1)] may be expressed as a unitary operation if we include an environment space, initially in a fixed state denoted $|A\rangle$. For clarity we include an extra register, initially in a fixed state $|0\rangle$, that is to receive the clone of $|\psi_i\rangle$. Then Equation (1) may be written as a *unitary* transformation,

$$|\psi_i\rangle|0\rangle|\alpha_i\rangle|A\rangle \rightarrow |\psi_i\rangle|\psi_i\rangle|C_i\rangle,$$

where $|C_i\rangle$ (generally depending on i) is the output state of the two registers that initially contained $|\alpha_i\rangle|A\rangle$. Hence, by unitarity, the two sets $\{|\psi_i\rangle|\alpha_i\rangle\}$ and $\{|\psi_i\rangle|\psi_i\rangle|C_i\rangle\}$ have equal matrices of inner products, and then so do the sets $\{|\alpha_i\rangle\}$ and $\{|\psi_i\rangle|C_i\rangle\}$ (by a simple cancellation of $\langle\psi_i|\psi_j\rangle$ from the two initial matrices). Thus, by Lemma 1 these two sets are unitarily equivalent; hence, $|\psi_i\rangle$ can be generated from $|\alpha_i\rangle$ alone (by applying the unitary equivalence and discarding the $|C_i\rangle$ register).

If ρ_i are mixed, we express them as probabilistic mixtures of pure states,

$$\rho_i = \sum_k p_k^{(i)} |\alpha_k^{(i)}\rangle\langle\alpha_k^{(i)}|$$

(where all $p_k^{(i)}$ are nonzero). Then a physical operation achieves

$$|\psi_i\rangle \otimes \rho_i \rightarrow |\psi_i\rangle|\psi_i\rangle \qquad \text{for all } i$$

if and only if it achieves

$$|\psi_i\rangle \otimes |\alpha_k^{(i)}\rangle \rightarrow |\psi_i\rangle|\psi_i\rangle \qquad \text{for all } i \text{ and } k. \tag{2}$$

By the pure state analysis above, a physical operation achieving Equation (2) exists if and only if there is a physical operation achieving

$$|\alpha_k^{(i)}\rangle \rightarrow |\psi_i\rangle \qquad \text{for all } i \text{ and } k.$$

We then get $\rho_i \rightarrow |\psi_i\rangle$ as well. *QED*.

In the particular case of cloning assisted by *classical* information (i.e., the states ρ_i are required to be mutually commuting), we deduce that this supplementary data must contain the full identity of the states as classical information. Indeed, if the ρ_i are classical, they can be copied any number of times; thus, if we can make one clone of $|\psi_i\rangle$ from ρ_i, we can make arbitrarily many clones and hence determine the identity of $|\psi_i\rangle$; i.e., the classical information of the label i must be contained in the supplementary classical information.

The proof of Theorem 1 is easily adapted to prove the following generalization: Let $\{|\psi_i\rangle\}$ be any finite set of pure states containing no orthogonal pairs of states. Let $|\psi_i\rangle^{\otimes n}$ denote the state of n copies of $|\psi_i\rangle$. Then there is a physical operation

$$|\psi_i\rangle^{\otimes n} \otimes \rho_i \rightarrow |\psi_i\rangle^{\otimes(n+1)}$$

if and only if there is a physical operation

$$\rho_i \rightarrow |\psi_i\rangle.$$

Curiously, the increasing information contained in n copies of $|\psi_i\rangle$ (as n increases) can never be used to assist in the creation of even a single extra copy.

[2] Thanks to H. R. Thomann (Consultant, Grosswiesenstrasse 80, CH-8051 Zurich) and A. Winter (Computer Science Department, University of Bristol UK) for pointing out an error in an earlier version of Theorem 1.

No deleting

Our techniques may also be used to give a simple proof of the Pati–Braunstein no-deleting principle [4] for sets $\{|\psi_i\rangle\}$ that contain no orthogonal pairs. The issue here is the following. Suppose we have two copies $|\psi_i\rangle|\psi_i\rangle$ of a state and we wish to delete one copy by a physical operation:

$$|\psi_i\rangle|\psi_i\rangle \rightarrow |\psi_i\rangle|0\rangle, \tag{3}$$

where $|0\rangle$ is any fixed state of the second register. As before, any such physical operation may be expressed as a unitary operation if we include an environment space, initially, say, in a fixed state $|A\rangle$. Then Equation (3) is equivalent to the unitary transformation

$$|\psi_i\rangle|\psi_i\rangle|A\rangle \rightarrow |\psi_i\rangle|0\rangle|A_i\rangle, \tag{4}$$

where the final state $|A_i\rangle$ of the environment may depend on $|\psi_i\rangle$ in general. One way of achieving this would be to simply swap (a constant) part of the environment into the second register, but then the second copy of $|\psi_i\rangle$ would remain in existence (albeit in the environment now). The no-deleting principle states that the second copy of $|\psi_i\rangle$ can *never* be "deleted" in the sense that $|\psi_i\rangle$ can *always* be resurrected from $|A_i\rangle$. Note, however, that if wavefunction collapse is also allowed as a valid physical process, deletion is possible. (We perform a complete measurement on $|\psi_i\rangle$ and rotate the seen post-measurement state to $|0\rangle$ by a unitary transformation that depends on the measurement outcome.)

To see the no-deleting principle with our methods, note that the unitarity of Equation (4) implies that the sets $\{|\psi_i\rangle|\psi_i\rangle|A\rangle\}$ and $\{|\psi_i\rangle|0\rangle|A_i\rangle\}$ have equal matrices of inner products, and then, as before, so do the sets $\{|\psi_i\rangle\}$ and $\{|A_i\rangle\}$. Thus Lemma 1 states that these sets are unitarily equivalent, which is just the no-deleting principle.

Permanence of quantum information

Deleting and cloning have a common feature: In cloning we saw that the existence of the first copy $|\psi_i\rangle$ provided no assistance in constructing the second copy from the supplementary information. Similarly for deletion, the existence of the first copy provides no assistance in deleting the second copy—in effect, the only way to delete the second copy is to transform it out into the environment [i.e., $|0\rangle|A_i\rangle$ in Equation (4) is a unitary transform of $|\psi_i\rangle|A\rangle$ alone], again making no use of the first copy. Considering no-cloning and no-deleting together (and excluding wavefunction collapse as a valid physical process), we see that quantum information (of non-orthogonal states) has a quality of "permanence": Creation of copies can be achieved only by importing the information from some other part of the world where it had *already existed*; destruction (deletion of a copy) can be achieved only by exporting the information to some other part of the world where it must *continue to exist*. This property is different from the *preservation* of information by any reversible dynamics. For example, consider the classical reversible C-NOT operation mapping $|b_1\rangle|b_2\rangle$ to $|b_1\rangle|b_1 \oplus b_2\rangle$ (where b_1, b_2 are bit values and $\oplus$ is addition modulo 2). This operation can imprint copies of a bit b into a standard state via $|b\rangle|0\rangle \rightarrow |b\rangle|b\rangle$ and also delete copies via $|b\rangle|b\rangle \rightarrow |b\rangle|0\rangle$. In both cases, however, the first copy is used in an essential way in the process and the information content of one copy is the same as that of two copies. In contrast, in the quantum (non-orthogonal) case, copying and deleting can only occur independently of the first copy, and then reversibility of dynamics implies that the information of the second copy must have already existed separately in the environment (for cloning) or continue to exist separately in the environment (for deletion). But in any reasonable intuitive sense, $|\psi_i\rangle|\psi_i\rangle$ does not have double the information content of $|\psi_i\rangle$ (and similarly, $|\psi_i\rangle^{\otimes n}$ cannot have n times the information content, since n is unbounded). One might interpret this as an overlap of information content of the two copies; then Theorem 1 implies that this common part cannot be duplicated from within a single copy and merely extended to give the second copy.

Information compression and Hilbert space geometry

So far our discussion of quantum information has been qualitative. However, it would be interesting to develop a quantitative theory of this concept, being able to say that one quantum system has more quantum information than another; and we would like to have corresponding dynamical laws for the manipulation of quantum information. In classical information theory there exists a well-defined quantitative notion of information. As a first attempt, we consider importing it into a quantum context.

In Shannon's classical information theory, we begin with a classical information source which is defined by a prior probability distribution $\{p_i\}$ of signals s_i. The information content is then quantified by the Shannon entropy $H(p_i) = -\Sigma_i\, p_i \log_2 p_i$ bits. This definition has a compelling physical interpretation: It characterizes the minimal resources (H bits per signal) that are necessary and sufficient to faithfully represent the source (in a suitable asymptotic sense [8] that we need not elaborate here). To approach the concept of quantum information, a natural avenue is to mimic this very successful classical theory in a quantum context. Thus we introduce a quantum source, characterized by a prior probability distribution $\{p_i\}$ of quantum signals $|\alpha_i\rangle$ and define its quantum information content to be S, the least number of qubits that are necessary and sufficient to faithfully represent the source (in an asymptotic sense that naturally

generalizes the classical case). Following Schumacher, one may prove [9] that S is then the von Neumann entropy $S(\rho)$ of the source density matrix $\rho = \Sigma_i\, p_i|\alpha_i\rangle\langle\alpha_i|$, establishing a fundamental role for von Neumann entropy in quantum information theory.

This notion of quantum information, while interesting, is perhaps not entirely satisfactory in that it still involves an essentially classical ingredient, viz., the prior classical probability distribution. However, in this context it should be pointed out that that there is an unexpected and remarkable harmony in such classical mixing of quantum information: If $\mathscr{E}_1 = \{|\alpha_i\rangle;\, p_i\}$ and $\mathscr{E}_2 = \{|\beta_j\rangle;\, q_j\}$ are two quantum sources with the same density matrix $\Sigma_i\, p_i|\alpha_i\rangle\langle\alpha_i| = \Sigma_j\, q_j|\beta_j\rangle\langle\beta_j|$, then $\mathscr{E}_1$ and $\mathscr{E}_2$ are entirely indistinguishable by any physical process. The quantum information of the $|\alpha_i\rangle$s probabilistically mixed by p_i is exactly the same as the quantum information of the $|\beta_j\rangle$s mixed by q_j; no trace of the component states remains in the mixture! This indistinguishability can be seen to be related to various other consistency requirements of a physical theory such as the no-superluminal-signaling principle [10].

A second difficulty with the proposal of identifying the quantum information content of a source with its von Neumann entropy S is the fact that very different sources can have the same entropy, yet some look "more quantum" than others! This was realized soon after the appearance of Schumacher's compression theorem and discussed by Bennett and other participants during the first ELSAG–Bailey Quantum Computation Workshop at the Institute for Scientific Interchange in Torino, Italy, in July 1993 (subsequently leading to [11]). It was suggested that a quantum source might be decomposable into a classical and a quantum part, with the von Neumann entropy quantifying both parts together. Then we would seek to separate out a maximal classical part and quantify the quantum part alone [11]. To illustrate the situation, consider a source which emits one of two orthogonal states $|\psi_0\rangle$ and $|\psi_1\rangle$ with equal prior probabilities of $\frac{1}{2}$. Since the states can be reliably identified by a measurement, this source can be represented entirely in classical terms with $S = 1$ classical bit per signal. Suppose now that the states are not quite orthogonal, e.g., $|\langle\psi_0|\psi_1\rangle| = 10^{-9}$. The von Neumann entropy is still very close to 1, and we ask: Is this source "almost classical"? That is, can we extract approximately one classical bit of information, leaving behind a very small amount of quantum information (e.g., almost parallel states) in such a way that the signals can still be faithfully reconstructed from the classical and quantum information parts? This question was settled only recently [12], in the negative: Let $\mathscr{E}$ be a quantum source whose signal states do not lie in a family of orthogonal subspaces. Then $\mathscr{E}$ can be faithfully compressed to α qubits per signal plus any number of classical bits per signal if and only if α is at least as large as the von Neumann entropy S; i.e., it is generically impossible to separate a source nontrivially into a classical and a quantum part, and the classical representation of exactly orthogonal states is therefore a singular feature of infinite precision.

Thus, we need to look at more subtle properties of quantum compression to distinguish sources with equal entropy. Following a further suggestion of Bennett, we can study features of so-called *visible* quantum compression. In this scenario the source is described by giving the *classical* information of the identity of the emitted signal state (rather than just the quantum state itself, as quantum information). Our task again is to faithfully represent the signal states with minimal resources. Since we now have more prior information about the signals (viz., the full classical information of their identities), we have more possibilities available for constructing such a minimal (compressed) representation. As an example, consider a source of four signal states, with equal prior probabilities of $\frac{1}{4}$ and having von Neumann entropy 1. In visible compression, we can represent this source with two classical bits per signal and no qubits (since there are four equiprobable possibilities) or with one qubit per signal and no classical bits (by creating the signal states and performing Schumacher's quantum compression on them). Between these two extremes, there is a tradeoff curve $q(c)$: If we have c classical bits per signal (with $0 \le c \le 2$), then $q(c)$ is the least number of qubits per signal that is sufficient to represent the source [so the above gives $q(0) = 1$ and $q(2) = 0$]. Thus, instead of trying to extract classical information from a quantum source, we *start* by giving a fully classical description of the source and consider the tradeoff involved in coding the source back into quantum terms. An extensive study of this tradeoff curve is given in [13], and it is found that the shape of the curve does indeed distinguish different sources with the same entropy.

Returning now to information compression and the insight it may provide into the notion of information, it is interesting to ask why compression is possible at all. Evidently some kind of redundancy in the raw signals is being eliminated. For a classical source, it is well known that nontrivial compression is possible if and only if the prior probability distribution is not uniform. For example, consider the case of two signals with unequal probabilities. In that case we already have some prior knowledge of the signal before it is received, in the sense that we can guess the signal (choosing the more probable one) and be correct more often than not. In this sense, part of the signal (if sent in full) is redundant.

The quantum situation is considerably more subtle. For any quantum source $\{|\alpha_i\rangle;\, p_i\}$ we have $S(\rho) \le H(p_i)$ with equality if and only if the signals are all mutually

orthogonal, suggesting that there is a quantum redundancy associated specifically with non-orthogonality. For example consider two qubit signals $|\psi_1\rangle$ and $|\psi_2\rangle$ at 45° with equal prior probabilities of $\frac{1}{2}$. Then $H(p_i)$ is 1 bit but $S(\rho)$ is 0.601 qubits. Moreover, $S(\rho)$ decreases monotonically from 1 to 0 as the overlap $|\langle\psi_1|\psi_2\rangle|^2$ is increased from 0 to 1.

The interpretation of non-orthogonality is one of the enigmas of quantum theory. Conventionally the overlap $|\langle\psi_1|\psi_2\rangle|^2$ provides a measure of the nondistinguishability of $|\psi_1\rangle$ and $|\psi_2\rangle$, and this is reflected in the properties of the von Neumann entropy above, viz., an increasing redundancy of quantum information with increasing overlap. Thus, if $\mathcal{E}_1 = \{|\alpha_i\rangle; 1/n\}$ and $\mathcal{E}_2 = \{|\beta_i\rangle; 1/n\}$ are two quantum sources with n states each (having all prior probabilities equal, for simplicity) and with larger overlaps $|\langle\beta_i|\beta_j\rangle|^2 > |\langle\alpha_i|\alpha_j\rangle|^2$ for each pair in $\mathcal{E}_2$ compared to the corresponding pair in $\mathcal{E}_1$, we would expect a decrease of information content in passing from $\mathcal{E}_1$ to $\mathcal{E}_2$. While this is true for $n = 2$, it can be shown to fail generically [7] for $n = 3$ and higher; i.e., it is generically possible to increase the von Neumann entropy of a source while increasing the overlaps of every pair of signal states! Evidently the quantum information content depends on more subtle structural properties of the geometry of the signals in the Hilbert space, beyond just the pairwise overlaps.

The relationship of quantum information to the geometry of the Hilbert space is largely unstudied, but for $n = 3$ we can say a little more [7]. The von Neumann entropy S of three equiprobable states $|\psi_1\rangle$, $|\psi_2\rangle$, $|\psi_3\rangle$ is a function of four independent real parameters: the three overlaps $a_{12} = |\langle\psi_1|\psi_2\rangle|^2$, $a_{23} = |\langle\psi_2|\psi_3\rangle|^2$, $a_{31} = |\langle\psi_3|\psi_1\rangle|^2$, and the phase ξ of the triple product $\Upsilon = \langle\psi_1|\psi_2\rangle\langle\psi_2|\psi_3\rangle\langle\psi_3|\psi_1\rangle$ (noting that the squared modulus of Υ is the dependent quantity $a_{12}a_{23}a_{31}$). Then, keeping a_{12}, a_{23}, and a_{31} fixed, we can vary ξ, and it can be shown [7] that S is actually a monotonically decreasing function of $\cos\xi$ [or $\mathrm{Re}(\Upsilon)$]. Despite this clean relationship, we still lack an intuitive understanding of why increasing the phase ξ allows increased compressibility.

Concluding remarks

In this paper we have promoted a viewpoint that attempts to place a notion of information at a primary fundamental level in the formulation of quantum physics. In the spirit of Landauer's slogan "Information is physical!" we would declare "Physics is informational!" Physical theories have traditionally been formulated in conceptual and mathematical terms that are, at root, *geometrical*. As such, they have an intuitive accessibility which has facilitated many developments (for example, the powerful guiding principles of symmetry and coordinate invariance in the construction of Lagrangians and field equations). Similarly, the concept of information has an intuitive basis, although not geometrical (and evidently having a complicated relation to the geometry of state space). Hence, it offers a potentially new perspective on quantum physics with its own guiding principles. For example, we might adopt the principle that any prospective physical theory should not allow the efficient solution of an NP-complete problem (where NP denotes the complexity class of nondeterministic polynomial time algorithms). This principle greatly restricts the form of the theory, yet, remarkably, it appears to hold in the established formalisms of classical and quantum physics, which developed from entirely different perspectives. Although an informational and geometrical formulation of a given physical theory would be mathematically equivalent, both points of view are valuable for further developments: Natural generalizations that these respective formalisms suggest would be quite different and no longer equivalent as theories.

Acknowledgments

The author's work is supported by the U.K. Engineering and Physical Sciences Research Council.

References

1. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters, "Teleporting an Unknown Quantum State Via Dual Classical and EPR Channels," *Phys. Rev. Lett.* **70,** 1895–1899 (1993).
2. C. H. Bennett and S. Wiesner, "Communication Via One- and Two-Particle Operators on Einstein–Podolsky–Rosen States," *Phys. Rev. Lett.* **69,** 2881–2884 (1992).
3. W. Wootters and W. Zurek, "A Single Quantum Cannot Be Cloned," *Nature* **299,** 802–803 (1982).
4. A. Pati and S. Braunstein, "Impossibility of Deleting an Unknown Quantum State," *Nature* **404,** 164–165 (2000).
5. M. Horodecki and R. Horodecki, "Are There Basic Laws of Quantum Information Processing?," *Phys. Lett. A* **244,** 473 (1998); R. Horodecki, M. Horodecki, and P. Horodecki, "On Balance of Information in Bipartite Quantum Communication Systems: Entanglement-Energy Analogy," *Phys. Rev. A* **63,** 022310 (2001).
6. M. Horodecki, R. Horodecki, A. Sen De, and U. Sen, "No-Deleting and No-Cloning Principles as Consequences of Conservation of Quantum Information"; see *http://arxiv.org/quant-ph/0306044/*.
7. R. Jozsa and J. Schlienz, "Distinguishability of States and Von Neumann Entropy," *Phys. Rev. A* **62,** 012301-1–01203-11 (1999).
8. T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.
9. B. Schumacher, "Quantum Coding," *Phys. Rev. A* **51,** 2738–2747 (1995); R. Jozsa and B. Schumacher, "A New Proof of the Quantum Noiseless Coding Theorem," *J. Mod. Opt.* **41,** 2343–2349 (1994).
10. L. Hughston, R. Jozsa, and W. Wootters, "A Complete Classification of Quantum Ensembles Having a Given Density Matrix," *Phys. Lett. A* **183,** 14–18 (1993); N. Gisin, "Stochastic Quantum Dynamics and Relativity," *Helv. Phys. Acta* **62,** 363–371 (1989).

11. C. H. Bennett, G. Brassard, R. Jozsa, D. Mayers, A. Peres, B. Schumacher, and W. Wootters, "Reduction of Quantum Entropy by Reversible Extraction of Classical Information," *J. Mod. Opt.* **41,** 2307–2314 (1994).
12. H. Barnum, P. Hayden, R. Jozsa, and A. Winter, "On the Reversible Extraction of Classical Information from a Quantum Source," *Proc. Roy. Soc. (Lond.) A* **457,** 2019–2039 (2001); M. Koashi and N. Imoto, "Teleportation Cost and Hybrid Compression of Quantum Signals"; see *http://arxiv.org/quant-ph/0104001/*.
13. P. Hayden, R. Jozsa, and A. Winter, "Trading Quantum for Classical Resources in Quantum Data Compression," *J. Math. Phys.* **43,** 4404–4444 (2002).

Received May 20, 2003; accepted for publication June 24, 2003

Richard Jozsa *Department of Computer Science, University of Bristol, Woodland Road, Bristol BS8 1UB, U.K. (R.Jozsa@bristol.ac.uk).* Professor Jozsa received a B.Sc.(Hons) from Monash University, Australia, in 1975 and a D.Phil. from Oxford University in 1981, both in mathematics. Throughout the 1980s he held postdoctoral positions at Oxford, McGill, Sydney, and other universities in Australia. In 1996 he was a Royal Society Leverhulme Senior Research Fellow and subsequently became Professor of Mathematical Physics at the University of Plymouth UK. Currently Professor Jozsa is an EPSRC Senior Research Fellow and Professor of Computer Science at the University of Bristol UK. His current research work is in the area of quantum algorithms and complexity, and quantum information theory.