Domain theory and quantum mechanics

Bob Coecke and Keye Martin
Oxford University Computing Laboratory
Wolfson Building, Parks Road
Oxford OX1 3QD
Bob.Coecke@comlab.ox.ac.uk, kmartin@comlab.ox.ac.uk

Abstract

We introduce a partial order on classical and quantum states which reveals that these sets are actually domains: Directed complete partially ordered sets with an intrinsic notion of approximation. The operational significance of the orders involved conclusively establishes that physical information has a natural domain theoretic structure.

In the same way that the order on a domain provides a rigorous qualitative definition of information, a special type of mapping on a domain called a measurement provides a formal account of the intuitive notion 'information content.' Not only is physical information domain theoretic, but so too is physical entropy: Shannon entropy is a measurement on the domain of classical states; von Neumann entropy is a measurement on the domain of quantum states.

These results yield a foundation for problem solving in computer science, quantum information and physics.

1. Introduction

One of the great lessons of the differential and integral calculus is that we can conquer the infinite, and in particular, the continuous, by means of the discrete. An infinite sum may be understood as a limit of finite sums, the area beneath a curve as the limit of areas of approximating rectangles, the line tangent to a curve at a point is the limit of the secant lines joining points nearby.

The philosophy espoused is unambiguous: The *ideal* can be realized as a *limit* of the *partial;* the abstract, as a limit of the concrete; the continuous, a limit of the discrete, and so on. And this powerful ideology, as it arises in the context of recursive functionals, is part of what the axioms of domain theory are intended to capture. But even in Scott's prelude to the subject, it is difficult to keep the imagination from wandering beyond computation [18]:

> "Maybe it would be better to talk about *information;* thus, $x \sqsubseteq y$ means that x and y want to approximate the same entity, but y gives more information about it. This means we have to allow *incomplete* entities, like x, containing only *partial* information."

In its purest interpretation, domain theory is a branch of mathematics which offers an exclusively qualitative account of information: A proposal for how we might find information structured in a universe where all things arise as a limit of the partial.

We prove that the density operator formulation of quantum mechanics [20] is an instance of domain theory: Its partial elements are the mixed states, its total or idealized elements are the pure states. To do so, we first order classical states recursively in terms of Bayesian state update, which corresponds to the process by which an observer looks for an object and updates his knowledge accordingly. This order on classical states, called the *Bayesian order,* combined with the predictions made by quantum mechanics about the measuring of certain observables, then enables us to naturally derive the *spectral order* on quantum states.

We then consider a few applications of these domains. The first is to physics, where we derive the logics of Birkhoff and von Neumann [2], $\mathcal{P}\{1, \ldots, n\}$ and $\mathbb{L}^n$, in a purely order theoretic manner. This is an exciting result given that $\mathbb{L}^n$ is the core of the axiomatic approach to quantum mechanics [12, 17]. The second is to domain theory, where we provide additional evidence that the domain theoretic notion *measurement* [13] can provide science with a formal mathematical definition of 'information content': Shannon entropy is a measurement on the domain of classical states, von Neumann entropy is a measurement on the domain of quantum states. Next we turn to quantum entanglement. Using an idea from the measurement formalism in domain theory, we show how more abstract measures of entanglement – "qualitative measures of entanglement" – can be used to precisely clarify the sense in which one state is more entangled than another. Last, we use the Bayesian order to calculate the complexity of Grover's algorithm [10], and to identify crucial qualitative properties it has that one must know about before implementing it experimentally.

All proofs of all theorems in sections 1–6 can be found in the research report [5]; section 7 is from [15].

2. Classical states

Definition 2.1 Let $n \geq 2$. The *classical states* are

$$\Delta^n := \left\{ x \in [0,1]^n : \sum_{i=1}^{n} x_i = 1 \right\}.$$

A classical state $x \in \Delta^n$ is *pure* when $x_i = 1$ for some $i \in \{1, \ldots, n\}$; we denote such a state by e_i.

Pure states $\{e_i\}_i$ are the actual states a system can be in, while general mixed states x and y are epistemic entities. If we know x and by some means determine that outcome i is not possible, our knowledge improves to

$$p_i(x) = \frac{1}{1 - x_i}(x_1, \ldots, \hat{x_i}, \ldots, x_{n+1}) \in \Delta^n,$$

where $p_i(x)$ is obtained by first removing x_i from x and then renormalizing. The partial mappings which result,

$$p_i : \Delta^{n+1} \rightharpoonup \Delta^n$$

with $\mathrm{dom}(p_i) = \Delta^{n+1} \setminus \{e_i\}$, are called the *Bayesian projections* and lead one directly to the following relation on classical states.

Definition 2.2 For $x, y \in \Delta^{n+1}$,

$$x \sqsubseteq y \equiv (\forall i)(x, y \in \mathrm{dom}(p_i) \Rightarrow p_i(x) \sqsubseteq p_i(y)). \quad (1)$$

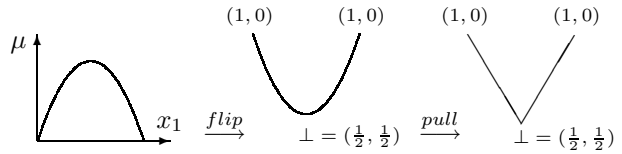
For $x, y \in \Delta^2$,

$$x \sqsubseteq y \equiv (y_1 \leq x_1 \leq 1/2) \text{ or } (1/2 \leq x_1 \leq y_1)\,. \quad (2)$$

The relation $\sqsubseteq$ on Δ^n is called the *Bayesian order.*

To motivate (1), if $x \sqsubseteq y$, then observer x knows less than observer y. If something transpires which enables each observer to rule out exactly e_i as a possible state of the system, then the first now knows $p_i(x)$, while the second knows $p_i(y)$. But since each observer's knowledge has increased by the same amount, the first must *still* know less than the second: $p_i(x) \sqsubseteq p_i(y)$.

The order on two states (2) is derived from the graph of Shannon entropy μ on Δ^2 (left) as follows:



The pictures above yield a canonical order on Δ^2:

Theorem 2.3 *There is a unique partial order on Δ^2 which has $\perp := (1/2, 1/2)$ and satisfies the mixing law*

$$x \sqsubseteq y \text{ and } p \in [0,1] \;\Rightarrow\; x \sqsubseteq (1-p)x + py \sqsubseteq y\,.$$

It is the Bayesian order on classical two states.

The set of *maximal elements* in a poset D is written $\max(D)$. The *least element* in a poset is denoted $\perp$, when it exists. A *domain* is a *dcpo* (directed complete partial order) with a notion of approximation. Here we are intentionally vague about 'approximation' since we will not explicitly make use of the idea until later; the details of this, as well as a more in depth derivation of the order, are in [5].

Theorem 2.4 *$(\Delta^n, \sqsubseteq)$ is a domain with maximal elements*

$$\max(\Delta^n) = \{e_i : 1 \leq i \leq n\}$$

and least element $\perp := (1/n, \ldots, 1/n)$.

The Bayesian order can also be described in a more direct manner, the *symmetric characterization*. Let $S(n)$ denote the group of permutations on $\{1, \ldots, n\}$ and $\Lambda^n := \{x \in \Delta^n : (\forall i < n)\, x_i \geq x_{i+1}\}$ denote the collection of *monotone* classical states.

Theorem 2.5 *For $x, y \in \Delta^n$, we have $x \sqsubseteq y$ iff there is a permutation $\sigma \in S(n)$ such that $x \cdot \sigma, y \cdot \sigma \in \Lambda^n$ and*

$$(x \cdot \sigma)_i (y \cdot \sigma)_{i+1} \leq (x \cdot \sigma)_{i+1} (y \cdot \sigma)_i$$

for all i with $1 \leq i < n$.

Thus, the Bayesian order is order isomorphic to $n!$ many copies of Λ^n identified along their common boundaries. This fact, together with the pictures of $\uparrow x$ at representative states x in Figure 1, will give the reader a good feel for the geometric nature of the Bayesian order.

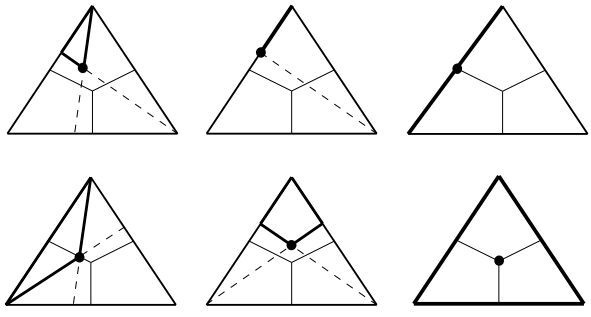


Figure 1. Pictures of $\uparrow x$ for $x \in \Delta^3$.

3. Quantum states

Let $\mathcal{H}^n$ denote an n-dimensional complex Hilbert space with specified inner product $\langle\cdot|\cdot\rangle$.

Definition 3.1 A *quantum state* is a density operator $\rho : \mathcal{H}^n \to \mathcal{H}^n$, i.e., a self-adjoint, positive, linear operator with $\mathrm{tr}(\rho) = 1$. The quantum states on $\mathcal{H}^n$ are denoted Ω^n.

Definition 3.2 A quantum state ρ on $\mathcal{H}^n$ is *pure* if

$$\mathrm{spec}(\rho) \subseteq \{0, 1\}.$$

The set of pure states is denoted Σ^n. They are in bijective correspondence with the one dimensional subspaces of $\mathcal{H}^n$.

Classical states are distributions on the set of pure states $\max(\Delta^n)$. By Gleason's theorem [8], an analogous result holds for quantum states: Density operators encode distributions on the set of pure states Σ^n.

Definition 3.3 A *quantum observable* is a self-adjoint linear operator $e : \mathcal{H}^n \to \mathcal{H}^n$.

An observable of a physical system is anything about it that we can measure. For example, *energy* is an observable. Observables in quantum mechanics are represented mathematically by self-adjoint operators. Time, on the other hand, *does not* have a representation as an operator in quantum mechanics; whether or not it is an observable is a topic well beyond this paper.

Now, if we have the operator e representing the energy observable of a system (for instance), then its set of eigenvalues $\mathrm{spec}(e)$, called the *spectrum* of e, consists of the actual energy values a system may assume. If our knowledge about the state of the system is represented by density operator ρ, then quantum mechanics predicts the probability that a measurement of observable e yields the value $\lambda \in \mathrm{spec}(e)$. It is

$$\mathrm{pr}(\rho \to e_\lambda) := \mathrm{tr}(p_e^\lambda \cdot \rho),$$

where p_e^λ is the projection corresponding to eigenvalue λ and e_λ is its associated eigenspace in the *spectral representation* of e.

Definition 3.4 Let e be an observable on $\mathcal{H}^n$ with $\mathrm{spec}(e) = \{1, \ldots, n\}$. For a quantum state ρ on Ω^n,

$$\mathrm{spec}(\rho|e) := (\mathrm{pr}(\rho \to e_1), \ldots, \mathrm{pr}(\rho \to e_n)) \in \Delta^n.$$

For the rest of the paper, we assume that all observables e have $\mathrm{spec}(e) = \{1, \ldots, n\}$. For our purposes it is enough to assume $|\mathrm{spec}(e)| = n$; the set $\{1, \ldots, n\}$ is chosen for the sake of aesthetics. Intuitively, then, e is an experiment on a system which yields one of n different outcomes; if our a priori knowledge about the state of the system is ρ, then our knowledge about what the result of experiment e *will be* is $\mathrm{spec}(\rho|e)$. Thus, $\mathrm{spec}(\rho|e)$ determines our ability to *predict* the result of the experiment e.

So what does it mean to say that we have more information about the system when we have $\sigma \in \Omega^n$ than when we have $\rho \in \Omega^n$? It could mean that there is an experiment e which (a) serves as a physical realization of the knowledge each state imparts to us, and (b) that we have a better chance of predicting the result of e from state σ than we do from state ρ. Formally, (a) means that $\mathrm{spec}(\rho) = \mathrm{Im}(\mathrm{spec}(\rho|e))$ and $\mathrm{spec}(\sigma) = \mathrm{Im}(\mathrm{spec}(\sigma|e))$, which is equivalent to requiring $[\rho, e] = 0$ and $[\sigma, e] = 0$, where $[a, b] = ab - ba$ is the commutator of operators.

Definition 3.5 Let $n \geq 2$. For quantum states $\rho, \sigma \in \Omega^n$, we have $\rho \sqsubseteq \sigma$ iff there is an observable $e : \mathcal{H}^n \to \mathcal{H}^n$ such that $[\rho, e] = [\sigma, e] = 0$ and $\mathrm{spec}(\rho|e) \sqsubseteq \mathrm{spec}(\sigma|e)$ in Δ^n.

This is called the *spectral order* on quantum states.

Theorem 3.6 *$(\Omega^n, \sqsubseteq)$ is a domain with maximal elements*

$$\max(\Omega^n) = \Sigma^n$$

and least element $\perp = I/n$, where I is the identity matrix.

There is one case where the spectral order can be described in an elementary manner.

Example 3.7 As is well-known, the 2×2 density operators can be represented as points on the unit ball in $\mathbb{R}^3$:

$$\Omega^2 \simeq \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}.$$

For example, the origin $(0, 0, 0)$ corresponds to the completely mixed state $I/2$, while the points on the surface of the sphere describe the pure states. The order on Ω^2 then amounts to the following: $x \sqsubseteq y$ iff the line from the origin $\perp$ to y passes through x.

Like the Bayesian order on Δ^n, the spectral order on Ω^n can also be characterized in terms of symmetries and projections. In its symmetric formulation, *unitary operators* on $\mathcal{H}^n$ take the place of permutations on $\{1, \ldots, n\}$, while the projective formulation of $(\Omega^n, \sqsubseteq)$ shows that each classical projection $p_i : \Delta^{n+1} \rightharpoonup \Delta^n$ is actually the restriction of a special quantum projection $\Omega^{n+1} \rightharpoonup \Omega^k$ with $k = n$.

4. The logics of Birkhoff and von Neumann

The logics of Birkhoff and von Neumann [2, 6] consist of the propositions one can make about a physical system. Each proposition takes the form "The value of observable e is contained in $E \subseteq \mathrm{spec}(e)$." For classical systems, the

logic is $\mathcal{P}\{1,\ldots,n\}$, while for quantum systems it is $\mathbb{L}^n$, the lattice of (closed) subspaces of $\mathcal{H}^n$. In each case, implication of propositions is captured by inclusion, and a fundamental distinction between classical and quantum – that there are pairs of quantum observables whose exact values cannot be simultaneously measured at a single moment in time – finds lattice theoretic expression: $\mathcal{P}\{1,\ldots,n\}$ is distributive; $\mathbb{L}^n$ is not.

We now establish the relevance of the domains Δ^n and Ω^n to theoretical physics: The classical and quantum logics can be *derived* from the Bayesian and spectral orders using the *same* order theoretic technique.

Definition 4.1 An element x of a dcpo D is *irreducible* when

$$\bigwedge(\uparrow x \cap \max(D)) = x$$

The set of irreducible elements in D is written $\mathrm{Ir}(D)$.

The order dual of a poset $(D, \sqsubseteq_D)$ is written D^*; its order is $x \sqsubseteq y \Leftrightarrow y \sqsubseteq_D x$.

Theorem 4.2 *For $n \geq 2$, the classical lattices arise as*

$$\mathrm{Ir}(\Delta^n)^* \simeq \mathcal{P}\{1,\ldots,n\} \setminus \{\emptyset\},$$

and the quantum lattices arise as

$$\mathrm{Ir}(\Omega^n)^* \simeq \mathbb{L}^n \setminus \{0\}.$$

It is worth pointing out that these logics consist exactly of the states traced out by the motion of a searching process on each of the respective domains. To illustrate, let $p_i^+ : \Delta^n \to \Delta^n$ for $1 \leq i \leq n$ denote the result of first applying the Bayesian projection p_i to a state, and then reinserting a zero in place of the element removed. Now, beginning with $\perp \in \Delta^n$, apply one of the p_i^+. This projects away a single outcome from $\perp$, leaving us with a new state. For the new state obtained, project away another single outcome; after $n-1$ iterations, this process terminates with a pure state e_i, and all the intermediate states comprise a path from $\perp$ to e_i. Now imagine all the possible paths from $\perp$ to a pure state which arise in this manner. This set of states is exactly $\mathrm{Ir}(\Delta^n)$. (See Figure 2).

The logic $\mathbb{L}^n$ is the canonical order theoretic structure corresponding to quantum mechanics in terms of only pure states. We are tempted to claim therefore that $(\Omega^n, \sqsubseteq)$ has a special place in physics: As a canonical order theoretic structure corresponding to quantum mechanics in terms of density operators. And if this idea proves to be correct, it means that $(\Omega^n, \sqsubseteq)$ offers a more complete picture of physical reality than does $\mathbb{L}^n$, due to the fact that the density operator formulation offers a more complete picture than simply working with pure states.

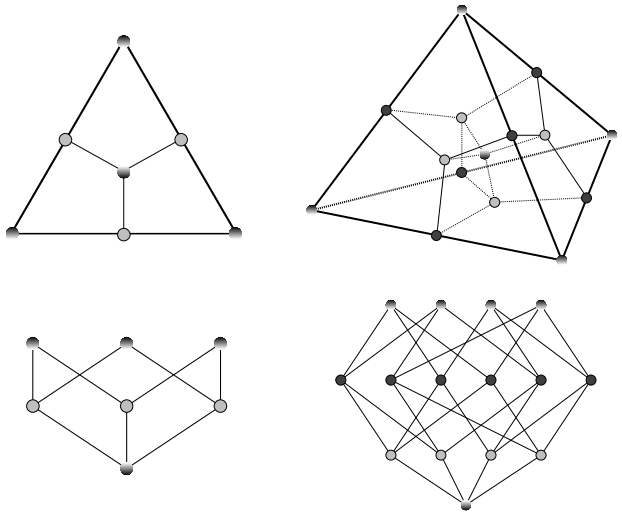


Figure 2. The irreducibles of Δ^3 and Δ^4 with their corresponding Hasse diagrams.

5. Entropy

A few of the ideas that the study of measurement [13] has led to include an informatic derivative, new fixed point theorems, the derivation of distance from content, techniques for treating continuous and discrete processes and data in a unified manner, a 'first order' view of recursion based on solving renee equations $\varphi = \delta + \varphi \circ r$ uniquely which establishes surprising connections between order and computability, and various approaches to complexity.

The original idea was that if a domain gave a formal account of 'information,' then a measurement on a domain should give a formal account of 'information content.' There is a stark difference between the view of information content taken in the study of measurement, and utterances of this phrase made elsewhere; it is this: Information content is a structural relationship between two classes of objects which, generally speaking, arises when one class may be viewed as a simplification of the other. The process by which a member of one class is simplified and thereby 'reduced' to an element of the the other is what we mean by 'the measurement process' in domain theory [14].

One of the classes may well be a subset of real numbers, but the 'structural relationship' underlying content should not be forgotten. Later we will use exactly this principle as the basis for a new approach to the study of entanglement. But right now, let us get to the point of this section: The formal notion of information content studied in measurement is broad enough in scope to capture Shannon's idea from information theory, as well as von Neumann's conception of entropy from quantum mechanics.

Definition 5.1 A Scott continuous map $\mu : D \to E$ between dcpo's is said to *measure the content of* $x \in D$ if

$$x \in U \Rightarrow (\exists \varepsilon \in \sigma_E)\, x \in \mu_\varepsilon(x) \subseteq U,$$

whenever $U \in \sigma_D$ is Scott open and

$$\mu_\varepsilon(x) := \mu^{-1}(\varepsilon) \cap \downarrow x$$

are the elements ε close to x in content. The map μ *measures* X if it measures the content of each $x \in X$.

Definition 5.2 A *measurement* is a Scott continuous map $\mu : D \to E$ between dcpo's that measures $\ker \mu := \{x \in D : \mu x \in \max(E)\}$.

The case $E = [0, \infty)^*$ is especially important. Then μ is a measurement iff for all $x \in D$ with $\mu x = 0$,

$$x \in U \Rightarrow (\exists \varepsilon > 0)\, x \in \mu_\varepsilon(x) \subseteq U,$$

whenever $U \subseteq D$ is Scott open. The elements ε close to $x \in \ker \mu$ are then given by

$$\mu_\varepsilon(x) := \{y \in D : y \sqsubseteq x \;\&\; |\mu x - \mu y| < \varepsilon\},$$

where for a *number* $\varepsilon > 0$ and $x \in \ker \mu$, we write $\mu_\varepsilon(x)$ for $\mu_{[0,\varepsilon)}(x)$. In this case, μx measures the *uncertainty* in x. Thus, an object with measure zero ought to have no uncertainty, which means it should be maximal.

Lemma 5.3 *If μ is a measurement, then* $\ker \mu \subseteq \max(D)$.

The converse is not true, and there are many important cases (like powerdomains [16]), where the applicability of measurement is greatly heightened by the fact that $\ker \mu$ need not consist of *all* maximal elements. However, in this paper, we are only interested in the case $\ker \mu = \max(D)$, so from here on we *assume* that this is part of the definition of measurement.

Theorem 5.4 *Shannon entropy*

$$\mu x = -\sum_{i=1}^{n} x_i \log x_i$$

is a measurement of type $\Delta^n \to [0, \infty)^*$.

A more subtle example of a measurement on classical states is the retraction $r : \Delta^n \to \Lambda^n$ which rearranges the probabilities in a classical state into descending order. We will apply it in our study of entanglement later on.

Theorem 5.5 *von Neumann entropy*

$$\sigma\rho = -\mathrm{tr}(\rho \log \rho)$$

is a measurement of type $\Omega^n \to [0, \infty)^*$.

Another natural measurement on Ω^n is the map $q : \Omega^n \to \Lambda^n$ which assigns to a quantum state its spectrum rearranged into descending order. It can be thought of as an important link between classical and quantum information theory.

By combining the quantitative and qualitative aspects of information, we obtain a highly effective method for solving a wide range of problems in the sciences. As an example, consider the problem of *rigorously* proving the statement "there is more information in the quantum than in the classical."

The first step is to think carefully about why we say that the classical is contained in the quantum; one reason is that for any observable e, we have an isomorphism

$$\Omega^n | e = \{\rho \in \Omega^n : [\rho, e] = 0\} \simeq \Delta^n$$

between the spectral and Bayesian orders. That is, each classical state can be assigned to a quantum state in such a way that *information is conserved*:

conservation of information
=
(qualitative conservation) + (quantitative conservation)
=
(order embedding) + (preservation of entropy).

This realization, that both the qualitative *and* the quantitative characteristics of information are preserved in passing from the classical to the quantum, solves the problem.

Theorem 5.6 *Let $n \geq 2$. Then*

(i) *There is an order embedding $\phi : \Delta^n \to \Omega^n$ with $\sigma \circ \phi = \mu$.*

(ii) *For any $m \geq 2$, there is no order embedding $\phi : \Omega^n \to \Delta^m$ with $\mu \circ \phi = \sigma$.*

Part (ii) is true for any pair of measurements μ and σ. The proof is fun: If (ii) is false, then ϕ restricts to an injection of $\max(\Omega^n)$ into $\max(\Delta^n)$, using $\ker \mu \subseteq \max(\Delta^n)$ and $\ker \sigma = \max(\Omega^n)$. But no such injection can actually exist: $\max(\Omega^n)$ is infinite, $\max(\Delta^n)$ is not.

6 Semantics of entanglement

Let (ψ_i) be a base of $\mathcal{H}^n$. Each pure state $s \in \Sigma^n$, being a one dimensional subspace of $\mathcal{H}^n$, can be written as a (normalized) vector $\sum_i c_i \psi_i \in s$. If we want to work with *qubits*, the case $n = 2$, we fix a *computational basis* $\{|0\rangle, |1\rangle\}$; in general, for *qunits* ($n > 2$), we fix a basis $\{|0\rangle, \ldots, |n-1\rangle\}$. Recall that a compound quantum system is described in the tensor product $\mathcal{H}^n \otimes \ldots \otimes \mathcal{H}^m$ of the

Hilbert spaces that describe the subsystems. Thus, a pure state of a compound quantum system has the form

$$\Psi = \sum_{i\ldots j} c_{i\ldots j}\psi_i \otimes \ldots \otimes \psi_j$$

since $(\psi_i \otimes \ldots \otimes \psi_j)$ is a base of $\mathcal{H}^n \otimes \ldots \otimes \mathcal{H}^m$ (by definition). For two qubits, $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ is a natural base, where we have abbreviated $|i\rangle \otimes |i\rangle \equiv |i\rangle|j\rangle \equiv |ij\rangle$. The tensor product allows one to capture a kind of intrinsic interaction between subsystems called *quantum entanglement*. In particular: *Quantum entanglement is the essential feature in quantum communication schemes and quantum cryptographic protocols that distinguishes them from their classical counterparts.* Concrete examples can be found in [3]. Below we illustrate by means of a series of examples how the results of this paper can be applied to the study of entanglement (from these examples it will be clear that the technique is generally applicable.) More details can be found in [5].

Example 6.1 *Measures of entanglement of bipartite quantum systems.* According to Schmidt's biorthogonal decomposition theorem, any bipartite state

$$\Psi = \sum_{ij} c_{ij}\psi_i \otimes \psi_j \in \mathcal{H}^n \otimes \mathcal{H}^n$$

can be rewritten as

$$\Psi = \sum_{i} c_i\psi_i^S \otimes \phi_i^S$$

with (ψ_i^S) and (ϕ_i^S) orthonormal bases and the (c_i) positive real coefficients (which as a set are uniquely defined). In particular we have $\sum_i c_i^2 = 1$ due to normalization of Ψ, so every $\Psi \in \mathcal{H}^n \otimes \mathcal{H}^n$ defines a unique classical state $c := (c_i^2)$. We can then *qualitatively measure entanglement* using the dcpo Λ^n as

$$\mathsf{Ent} : \mathcal{H}^n \otimes \mathcal{H}^n \to \Lambda^n : \Psi \mapsto r(c).$$

Every measurement $\mu : \Lambda^n \to [0,1]^*$ then gives rise to a *quantitative measure of entanglement*

$$\mu \cdot \mathsf{Ent} : \mathcal{H}^n \otimes \mathcal{H}^n \to [0,1]^* .$$

For μ Shannon entropy we find the usual quantitative measure of entanglement for bipartite quantum systems.

As an example, the state

$$\mathsf{S} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathcal{H}^3 \otimes \mathcal{H}^3.$$

has essentially a qubit nature, that is, we can express the state by only using a subbase of $\mathcal{H}^3$ that contains two vectors. For $1/3 < q < 1$ the states

$$\mathsf{T}_q := q(|00\rangle) + \frac{1-q}{2}(|11\rangle + |22\rangle) \in \mathcal{H}^3 \otimes \mathcal{H}^3$$

exhibit genuine qutrit entanglement. We obtain

$$\mathsf{Ent}(\mathsf{S}) = r\left(\frac{1}{2}, \frac{1}{2}, 0\right) \;\&\; \mathsf{Ent}(\mathsf{T}_q) = r\left(q, \frac{1-q}{2}, \frac{1-q}{2}\right).$$

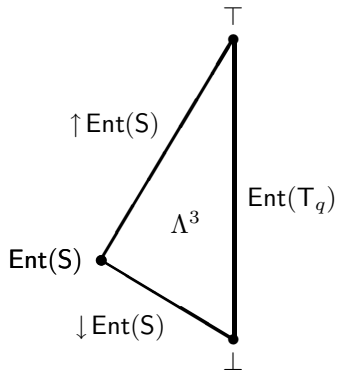
The states $\Psi \in \mathcal{H}^3 \otimes \mathcal{H}^3$ for which we have $\mathsf{Ent}(\mathsf{S}) \sqsubseteq \mathsf{Ent}(\Psi)$ are those such that $\mathsf{Ent}(\Psi) = r\,(q, 1-q, 0)$ for $0 \leq q \leq 1/2$, that is, convex combinations of S and the minimally entangled state in $\mathcal{H}^3 \otimes \mathcal{H}^3$ (the pure tensor $|00\rangle$) which provides a top

$$\top := \mathsf{Ent}(|00\rangle)\,.$$

The states $\Psi \in \mathcal{H}^3 \otimes \mathcal{H}^3$ for which we have $\mathsf{Ent}(\Psi) \sqsubseteq \mathsf{Ent}(\mathsf{S})$ are convex combinations of S and the maximally entangled state in $\mathcal{H}^3 \otimes \mathcal{H}^3$, which provides a bottom

$$\bot := \mathsf{Ent}\left(\frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)\right).$$

Graphically we have



We can further refine our qualitative representation of entanglement for bipartite states using the order on quantum states. The quantitative valuation $\mu \cdot \mathsf{Ent}$ with μ Shannon entropy can also be defined as the von Neumann entropy of one of the quantum states $\rho_1(\Psi)$ or $\rho_2(\Psi)$ for $\Psi \in \mathcal{H}^n \otimes \mathcal{H}^n$ that arise as partial traces, that is, the states of the subsystems. In fact, it is exactly the entanglement that causes a lack of knowledge about the actual pure state of the subsystems.

Example 6.2 *Qualitative entanglement of multipartite quantum systems.* In Example 6.1 we measured entanglement of bipartite quantum systems using unicity of the coefficients in the Schmidt biorthogonal decomposition. However, there does not exist a similar construction for arbitrary multipartite sytems. In particular, until now, there was not even a satisfactory notion of maximal entanglement e.g. see [19]. When considering three partite qubit states for the Greenberger-Horn-Zeilinger state [9]

$$\mathsf{GHZ} := \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$$

and the W-state [7]

$$\mathsf{W} := \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle)$$

there are conflicting arguments about which one is maximally entangled. The general favorite, however, is GHZ, especially in view of its maximal violation of certain types of inequalities that are characteristic for entanglement. The solution lies in the specification of context with respect to which one measures entanglement. Define

$$\begin{aligned}\mathsf{Ent}^{\Omega} : \mathcal{H}^n \otimes \ldots \otimes \mathcal{H}^n &\to \Omega^n \times \ldots \times \Omega^n \\ : \Psi &\mapsto \big(\rho_1(\Psi), \ldots, \rho_m(\Psi)\big)\end{aligned}$$

where $\rho_i(\Psi)$ arises by tracing over all systems except the ith. We can do this for example by considering the Schmidt decomposition for $\mathcal{H}^n \otimes (\mathcal{H}^n \otimes \ldots \otimes \mathcal{H}^n)$ where the single Hilbert space encodes the ith system.

We then obtain for the above examples that

$$\mathsf{Ent}^{\Omega}(\mathsf{GHZ}) = \left(\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}\right)$$

since

$$\mathsf{GHZ} = \frac{1}{\sqrt{2}}(|0\rangle|00\rangle + |1\rangle|11\rangle)$$

with respect to the 1st component and

$$\mathsf{Ent}^{\Omega}(\mathsf{W}) = \left(\begin{pmatrix} 2/3 & 0 \\ 0 & 1/3 \end{pmatrix}, \begin{pmatrix} 2/3 & 0 \\ 0 & 1/3 \end{pmatrix}, \begin{pmatrix} 2/3 & 0 \\ 0 & 1/3 \end{pmatrix}\right)$$

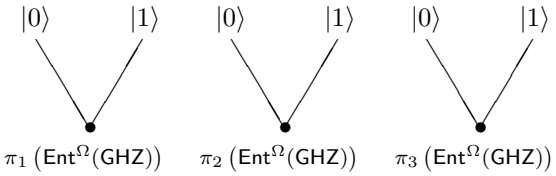
since eg.

$$\mathsf{W} := \frac{\sqrt{2}}{\sqrt{3}}|0\rangle\big(\frac{1}{\sqrt{2}}\left(|10\rangle + |01\rangle\right)\big) + \frac{1}{\sqrt{3}}|1\rangle|00\rangle$$

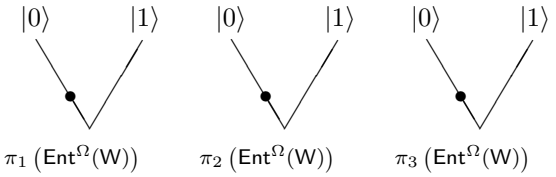
so it follows that

$$\mathsf{Ent}^{\Omega}(\mathsf{GHZ}) \sqsubset \mathsf{Ent}^{\Omega}(\mathsf{W})\,.$$

Writing only the part of Ω^2 containing the relevant pure states $|0\rangle$ and $|1\rangle$, i.e., a copy of Δ^2, our picture of $\mathsf{Ent}^{\Omega}(\mathsf{GHZ})$ is



while for $\mathsf{Ent}^{\Omega}(\mathsf{W})$ we have



where π_1, π_2 and π_3 represent the components of Ent^{Ω}.

Thus, with respect to Ent^{Ω} and the spectral order on Ω^n, GHZ is indeed the maximally entangled state.

These examples are worth giving some thought to. They use domains as a clarifying device. If we say, "this state is more entangled than that state," it has to be with respect to a qualitative measure of entanglement, call it Ent, and with respect to an order on the codomain of Ent. The pragmatic effect of this is that it allows for coherent arguing about which of two states is more entangled. The trouble with using numbers to measure entanglement is that numbers always compare, so one always gets an answer to the question 'which state is more entangled?' And that's only good if you have a genuine understanding of the process used to generate the number. The *semantics of entanglement* is thus motivated.

7 Grover's algorithm

The ideas and results of this section were first introduced in [15]. Grover's algorithm [10] for searching is the only known quantum algorithm whose complexity is *provably better* than its classical counterpart. We will now use the Bayesian order to analyze this algorithm. Here are some crucial things the approach yields:

(a) The complexity of the algorithm,

(b) A qualitative property the algorithm possesses called *antimonotonicity*. Without knowledge of this aspect, an experimental implementation would almost certainly fail (for reasons that will be clear later).

(c) An explanation of the algorithm as being an attempt to calculate a classical proposition.

Grover's algorithm searches a list L of length n (a power of two) for an element k known to occur in L precisely m times with $n > m \geq 1$. The register begins in the pure state

$$|\psi\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}|i\rangle$$

and after j iterations of the Grover operator G

$$G^j|\psi\rangle = \frac{\sin(2j\theta+\theta)}{\sqrt{m}}\sum_{L(i)=k}|i\rangle + \frac{\cos(2j\theta+\theta)}{\sqrt{n-m}}\sum_{L(i)\neq k}|i\rangle$$

where $\sin^2\theta = m/n$. The probability that a measurement yields i after j iterations is

$$\sin^2(2j\theta+\theta)/m \text{ if } L(i) = k$$

and

$$\cos^2(2j\theta+\theta)/(n-m) \text{ if } L(i) \neq k.$$

To get the answer, we measure the state of the register in the basis $\{|i\rangle : 1 \leq i \leq n\}$; if we perform this measurement after j iterations of G, when the state of the register is $G^j|\psi\rangle$, our knowledge about the result is represented by the classical state

$$\begin{aligned} x(j) &= \Big(\frac{\sin^2(2j\theta+\theta)}{m},\ldots,\frac{\sin^2(2j\theta+\theta)}{m}\ , \\ &\quad \frac{\cos^2(2j\theta+\theta)}{n-m},\ldots,\frac{\cos^2(2j\theta+\theta)}{n-m}\Big) \end{aligned}$$

The crucial step now is to *imagine* t iterations,

$$\begin{aligned} x(t) &= \Big(\frac{\sin^2(2t\theta+\theta)}{m},\ldots,\frac{\sin^2(2t\theta+\theta)}{m}\ , \\ &\quad \frac{\cos^2(2t\theta+\theta)}{n-m},\ldots,\frac{\cos^2(2t\theta+\theta)}{n-m}\Big) \end{aligned}$$

which defines a monotone state for $t \in \mathrm{dom}(x) = [a,b]$, $a = 0$ and $b = \pi/2\theta - 1$. The image of $x : [a,b] \to \Lambda^n$ is a chain in the Bayesian order, which is simplest to see by noting that it has the form

$$x = (f,\ldots,f,g,\ldots,g)$$

so that $f(s)g(t) \leq f(t)g(s) \Rightarrow x(s) \sqsubseteq x(t)$; otherwise, $x(t) \sqsubseteq x(s)$. We can now determine the exact nature of the motion represented by x with the following observation: If $x : [a,b] \to D$ is a curve on a domain D whose image is a chain and whose time derivative

$$\dot{x}_v(t) := \frac{d(v \circ x)}{dt}(t) = \lim_{s \to t} \frac{vx(s) - vx(t)}{s - t}$$

exists with respect to a *variable* v (i.e., a strictly monotone measurement $v : D \to [0,\infty)^*$), then

(i) The curve x has an absolute maximum on $[a,b]$: There is $t^* \in [a,b]$ such that

$$x(t^*) = \bigsqcup_{t \in [a,b]} x(t),$$

and

(ii) Either $t^* = a$, $t^* = b$ or $\dot{x}_v(t^*) = 0$.

Part of the power of this simple approach is that we are free to choose any variable we like. To illustrate, a tempting choice might be entropy $v = \mu$, but then solving $\dot{x}_v = 0$ means solving the equation

$$-m\dot{f}(1+\log f) - (n-m)\dot{g}(1+\log g) = 0$$

and we also have to determine the points where $\dot{x}_v$ is undefined, the set $\{t : g(t) = 0\}$. However, if we use the variable

$$v = 1 - \sqrt{x^+},$$

then we only have to solve a single elementary equation

$$\cos(2t\theta + \theta) = 0$$

for t, allowing us to conclude that the maximum must occur at $t = a$, $t = b$, or at points in

$$\{t : \dot{x}_v(t) = 0\} = \{b/2\}.$$

The absolute maximum of x is

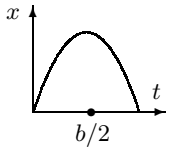
$$x(b/2) = (1/m,\ldots,1/m,0,\ldots,0)$$

because for the other points we find a minimum of

$$x(a) = x(b) = \bot.$$

The value of knowing the absolute maximum is that it allows us to calculate the complexity of the algorithm: It is $O(b/2)$, the amount of time required to move to a state from which the likelihood of obtaining a correct result by measurement is maximized. This gives $O(\sqrt{n/m})$ using $\theta \geq \sin\theta \geq \sqrt{m/n}$ and then $b/2 \leq (\pi/4)\sqrt{n/m} - 1/2$.

From $\dot{x}_v(t) \leq 0$ on $[a,b/2]$ and $\dot{x}_v(t) \geq 0$ on $[b/2,b]$, we can also graph x:



This is the 'antimonotonicity' of Grover's algorithm: If $j = b/2$ iterations will solve the problem accurately, $2j$ iterations will mostly unsolve it! This means that our usual way of reasoning about iterative procedures like numerical methods, as in "we must do at least j iterations," no longer applies. We must say "do exactly j iterations; no more, no less." As is now clear, precise estimates like these have to be obtained before going into a laboratory whenever possible.

Finally, we can view Grover's algorithm as an attempt to calculate as closely as possible the classical proposition

$$x(b/2) = (1/m,\ldots,1/m,0,\ldots,0) \in \mathrm{Ir}(\Delta^n).$$

It does so by generating *approximations*

$$x(t) \ll x(b/2)$$

for all $t \neq b/2$, where $\ll \subseteq \sqsubseteq$ is approximation in the sense of exact domains [5], a generalization of the usual notion that is equivalent to the way-below relation on continuous domains [1].

We believe this analysis suggests a special connection between Grover's algorithm and the Bayesian order. First, the sets $\downarrow x \cup \{x\}$ – the ones Grover's algorithm moves along – are chains in the Bayesian order *provided* $x \in \mathrm{Ir}(\Delta^n)$ is a proposition – and chains are just what we need in order to do calculus! Then, these very same sets – geometrically, the lines that join $\bot$ to a proposition – are all one needs to recover the entire Bayesian order as the result of a systematic procedure given in [4].

8 Closing remarks

It is worth pointing out that these results improve on some in the existing physics literature [11]. This is explained in [15] for those interested.

References

[1] S. Abramsky and A. Jung. Domain theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science, vol. III*. Oxford University Press, 1994.

[2] G. Birkhoff and J. von Neumann. The logic of quantum mechanics. *Annals of Mathematics*, 37:823–843, 1936.

[3] D. Bouwmeester, A. Ekert, and A. Zeilinger, editors. *The physics of quantum information*. Springer-Verlag, Berlin, 2001.

[4] B. Coecke. Entropic geomety from logic. In *Mathematical Foundations of Programming Semantics 19*, volume 83 of *Electronic Notes in Theoretical Computer Science*. Elsevier Science, 2003.

[5] B. Coecke and K. Martin. A partial order on classical and quantum states. Technical Report PRG-RR-02-07, Oxford University, 2002. http://web.comlab.ox.ac.uk/oucl/publications/tr/rr-02-07.html.

[6] B. Coecke, D. J. Moore, and A. Wilce, editors. *Current research in operational quantum logic: Algebras, categories, languages*. Kluwer Academic Publishers, Dordrecht, 2000.

[7] W. Dür, G. Vidal, and J. Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62:062314, 2000.

[8] A. M. Gleason. Measures on the closed subspaces of a hilbert space. *Journal of Mathematics and Mechanics*, 6:885–893, 1957.

[9] M. A. Greenberger, M. A. Horn, A. Shimony, and A. Zeilinger. Bell's theorem without inequalities. *American Journal of Physics*, 58:1131, 1990.

[10] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Physical Review Letters*, 78:325, 1997.

[11] J. I. Latorre and M. A. Martín-Delgado. The majorization arrow in quantum algorithm design. *Physical Review A*, 66:022305, 2002.

[12] G. M. Mackey. *The mathematical foundations of quantum mechanics*. W. A. Benjamin, New York, 1963.

[13] K. Martin. *A foundation for computation*. PhD thesis, Tulane University, 2000.

[14] K. Martin. The measurement process in domain theory. In *The 27th International Colloquium on Automata, Languages and Programming (ICALP 2000)*, volume 1853 of *Lecture Notes in Computer Science*, Berlin, 2000. Springer-Verlag.

[15] K. Martin. Epistemic motion in quantum searching. Technical Report PRG-RR-03-06, Oxford University, 2003. http://web.comlab.ox.ac.uk/oucl/publications/tr/rr-03-06.html.

[16] K. Martin, M. Mislove, and J. Worrell. Measuring the probabilistic powerdomain. In *The 29th International Colloquium on Automata, Languages and Programming (ICALP 2002)*, volume 2380 of *Lecture Notes in Computer Science*, Berlin, 2002. Springer-Verlag.

[17] C. Piron. *Foundations of quantum physics*. W. A. Benjamin, New-York, 1976.

[18] D. S. Scott. Outline of a mathematical theory of computation. Technical Report PRG-2, Oxford University, November 1970.

[19] P. van Loock and S. L. Braunstein. Multipartite entanglement. In A. K. Pati and S. L. Braunstein, editors, *Quantum information theory with continuous variables*. Kluwer Academic Publishers, 2002. arXiv:quant-ph/0205068.

[20] J. von Neumann. *Mathematische grundlagen der quantenmechanik*. Springer-Verlag, Berlin, 1932. Translation, *Mathematical foundations of quantum mechanics*, Princeton University Press, 1955.