Discrete Quantum Causal Dynamics

Richard F. Blute*
Ivan T. Ivanov*
Department of Mathematics
and Statistics
University of Ottawa
Ottawa, Ontario, Canada
rblute,iti@mathstat.uottawa.ca

Prakash Panangaden*
School of Computer Science
McGill University
Montréal, Québec, Canada
prakash@cs.mcgill.ca

February 4, 2003

Abstract

We give a mathematical framework to describe the evolution of an open quantum systems subjected to finitely many interactions with classical apparatuses. The systems in question may be composed of distinct, spatially separated subsystems which evolve independently, but may also interact. This evolution, driven both by unitary operators and measurements, is coded in a mathematical structure in such a way that the crucial properties of causality, covariance and entanglement are faithfully represented. We show how our framework may be expressed using the language of (poly)categories and functors. Remarkably an important physical consequence - covariance, i.e. independence of the slicing used - follows directly from functoriality.

We establish links between the physical picture we propose and linear logic. Specifically we show that the refined logical connectives of linear logic can be used to describe the entanglements of subsystems in a precise way. Furthermore, we show that there is a correspondence between the evolution of a given system and deductions in a certain formal logical system based on the rules of linear logic.

This framework generalizes both causal posets and the histories approach to quantum mechanics.

1 Introduction

We propose a uniform scheme for describing a quantum system, interacting with a network of classical objects. The system in question may be composed of distinct spatially separated subsystems which evolve independently, but may also interact with each other at various points as well as with the classical objects. When analyzing physical laboratory experiments on quantum systems, we frequently abstract away from the concrete experimental setup and from the particular details of the machinery involved. What we usually keep is the description of the quantum system - and its spatially separated subsystems - in terms of wave functions or density matrices and unitary operators as well as the changes of the quantum system induced by the interactions with classical devices. Crucial properties of the evolution such as the causal ordering, covariance of the description for different observers and quantum entanglement between distinct subsystems should be completely reflected in any such description.

*Research supported by NSERC (Canada).

The basis of our representation is the graph of events and causal links between them. An event could be one of the following: a unitary evolution of some subsystem, an interaction of a subsystem with a classical device (a measurement) or perhaps just the coming together or splitting apart of several spatially separated subsystems. Events will be depicted as vertices of a directed graph. The edges of the graph will represent a physical flow between the different events. The vertices of the graph are then naturally labelled with operators representing the corresponding processes.

Of course, the processes of unitary evolution and measurement take a certain amount of time; but we are only interested in the causal relations between such events and this allows us to consider them as point-like vertices on the graph. Thus we are thinking of the duration between events as being longer than the duration of an event so that no causal information is lost when we represent interactions as events.

The structure described thus far reflects the kinematical properties of the quantum system. To describe the dynamics we need a composition of the operators assigned to the vertices of the graph. This composition is most conveniently described in terms of a composition in a specific mathematical structure, namely a polycategory generated by the graph. The whole description could then be concisely summarized by noticing that we have a functor from this polycategory to the polycategory of Hilbert spaces. This functor captures the dynamics of the system.

Causal relations are made explicit and we prove that no influences breaking causality arise in our scheme. The possible entanglement between spatially separated subsystems - represented by distinct edges of the graph - is also accounted for. Thus, our framework allows one to represent locality of interaction - i.e. causal influences do not propagate outside the causal "cone" - while allowing the expression of nonlocal correlations which occur when one has quantum entanglement. The tension between causal evolution and quantum entanglement is resolved.

The categorical framework that we use is intimately connected with linear logic. Linear logic was originally introduced [Gir87] as a logic intended for a finer analysis of the way resources are consumed during the course of a proof. This logic has had a significant impact on the theory of computation as well as such far-flung areas as linguistics and pure mathematics. In the present paper, the connectives of (multiplicative) linear logic will be used to express the existence or nonexistence of nonlocal correlations. What we will introduce is a deductive system based on the graph-theoretic structure of the system that precisely picks out the spatial slices of physical interest. Thus evolution of the system corresponds to logical deductions within this deductive system. For an expository introduction to linear logic, see the review by Girard [Gir95] or the brief exposition in Appendix B.

1.1 Relation to other work

Next we outline the relations of our proposal to some recent approaches to quantum mechanics and quantum gravity.

1.1.1 Consistent and decoherent histories

The *consistent histories* approach to quantum mechanics due to Griffiths and Omnès [Gri96, Omn94] was formulated with the aim of shedding new light on the conceptual difficulties of the theory. A closely related proposal with different motivation is the *decoherent histories* approach to quantum cosmology of Gell-Mann and Hartle [GMH93]. The basic ingredient in both approaches is the notion of a *history* of the quantum system described by a sequence of projection operators in the Hilbert space of the system, for a succession of times. The goal of quantum mechanics is to determine the probability of an event or a sequence of events, thus one might hope to assign probabilities to the histories of the quantum system. In order for the probabilities to be additive in the usual sense, the histories have to be mutually noninterfering. Sets of histories obeying this condition are selected with the use of a special bilinear form on histories - the decoherence functional.

A particular history is mathematically represented as a linearly ordered sequence of projection operators in the Hilbert space of the quantum mechanical system. But the linear causal ordering of the events in a history is too restrictive in many experimental situations, in particular when analyzing spatially separated entangled quantum systems. This issue is even more pressing for quantum cosmology considerations. An application of the histories approach to quantum field theory on a curved space-time [Ble91] must assume the existence of a globally hyperbolic manifold, and thus via the associated foliation, a linear ordering of the histories of the quantum field.

Our proposal for describing the evolution of an open quantum system can be considered as describing a single history in a set of histories. The important point is that events are no longer linearly ordered by temporal order but, rather, partially ordered with respect to the causal order. This allows one to capture the notion of causal evolution in a manifestly covariant fashion. The consistency/decoherence condition for histories has an immediate generalization for histories described by more general graphs as proposed here.

1.1.2 Causal sets

Causal sets form the basis of an approach to quantum gravity mainly advocated by R. Sorkin and collaborators [BLMS87, Sor91], where the basic idea is to take the notion of causality as the primitive. In classical relativity, the structure of the space-time manifold together with a metric of Lorentzian signature determines the causality relation. An important observation is that the causal structure is conformally invariant, i.e. determined by only the conformal equivalence class of the metric and hence more primitive than the metric. Various proposals for quantum gravity - for example, the twistor program [PM72] - have taken as their point of departure the idea that the causal structure is more fundamental than the metric structure.

In the causal sets approach, one takes the point of view that, at the smallest length scales, spacetime is inherently discrete and that the causal structure, the "light cones", are fundamental. This leads naturally to the idea of a partially ordered set (poset for short) where the elements are events and two events are related by causality. The main interest is in approximating continuous spacetimes with such structures and defining processes that would generate these structures, with a view to an eventual theory of quantum gravity. Though the aims are rather different the issues connected with causality are closely related.

Causal sets are further motivated by the idea that a discrete structure would avoid the singularities that plague physics (both classical and quantum). The assumption that space-time should be a continuous manifold is one of the ingredients that leads to the problematic singularities of quantum field theory and general relativity. In the causal sets approach, space-time is a discrete structure, thus possibly avoiding these singularities, the idea being that at the Planck scale, continuous geometry gives way to discrete geometry.

One way to think of this is that one approximates a manifold as one "sprinkles" more and more points into the causal set in a uniform fashion. More formally, one would want to obtain a manifold as the categorical limit of a diagram of posets and embeddings [Mac98]. Applications and extensions of these ideas can be found in papers such as [Mar00, MS97, Rap00], although this list is by no means exhaustive. In our approach we are not thinking about generating the spacetime through such limiting processes, but the idea of a causal set is implicit in our work. For us, a finite causal set is the kinematical framework on which we describe evolution and information flow.

1.1.3 Quantum causal histories

The notion of *quantum causal history* was introduced by Markopoulou in [Mar00]. One begins with a poset (causal set) and assigns Hilbert spaces to the vertices and evolution operators to sets of edges. The assignment must satisfy properties analogous to functoriality. However, within this framework, one is quickly led to violations of causality - as the author herself notes - essentially because the slices used are "too global." She mentions the possibility of working with a dual view. In fact, in our work, we take such a dualized view as our starting point. In other words we assign operators representing evolution or measurement to vertices and Hilbert spaces to the edges, in a way satisfying (poly)functoriality. However, if we only work locally we get a causal theory but lose the possibility of capturing nonlocal correlations.

1.2 The Importance of Categories

A category can be seen as a generalization of a poset in the following sense. A poset merely records that an element x is less than y but a category keeps track of the different ways in which x might be less than y. For example, in logic one might consider formulas (denoted by Greek letters like ϕ, ψ etc.) and the relationship of provability between them. Thus one would write $\phi \vdash \psi$ to mean that starting from the assumption ϕ one can prove ψ. This gives rise to a transitive and reflexive relation; if one considers equivalence classes of formulas (two formulas being equivalent if each can be used to prove the other) we get a poset. However, if we are interested in distinguishing distinct proofs we need to keep track of the different ways in which ϕ can be used to prove ψ. Thus formulas as objects and proofs as morphisms can be organized into a category.

In a poset when one writes $x \leq y$ then, depending on the context, one is stating something like the following:

- x is less than y;
- x precedes y;
- x implies y.

or any of several other possibilities. In a causal set, we have in mind that x causally precedes y.

In the present work, we are particularly interested in modelling the idea that information can flow from one event to another in a number of different ways, *along different paths or channels.* We would like to keep track of all these various independent paths. The structure of a poset is inadequate for achieving this, as we would like to say that $x \leq y$ in several different ways. This naturally suggests that we pass from posets to more general graphs and eventually to categories.

Many recent experiments feature spatially distributed quantum systems. When entangled quantum subsystems come back together in the same spacetime region, the description of the resulting system is causally influenced by all events in the paths of the subsystems. In particular a past event could influence the future events in several distinct ways through different paths. Our scheme is well adapted for analyzing experiments featuring spatially separated quantum entangled entities and could be used in the field of quantum information processing to analyze information flow situations.

1.3 Contents of the present paper

Section 2 presents the basic ideas of our scheme via an example. Section 3 discusses the basic physical ideas involved. In the first subsection we review the notions of measurements and interventions. In the next subsection we give the dynamical prescription in a special case and in the final subsections we give the general prescription and prove covariance. In section 4 we review basic facts about polycategories and their construction. We also describe the polycategory of Hilbert spaces and intervention operators we will be using. In section 5 we give a logical presentation of polycategories and establish the connection between our structures and linear logic. A functorial version of our dynamical prescription is then presented. We end with a discussion on further applications of our scheme. We include an appendix on quantum mechanics, from the point of view of density matrices, for the benefit of readers who are not familiar with this topic. It is our hope that this paper will interest members of several different communities within mathematics, logic and physics.

2 Causal information flow via examples

Consider a quantum system evolving in space-time while being subjected to interactions with classical observers at a number of points. The causal and spatio-temporal relations in the system will be represented by a directed acyclic graph (hereafter called a *dag*). The vertices of the graph - which will be drawn as boxes - represent the events in the evolution of the system. An event could be a measurement by a classical observer, a local unitary evolution or just a splitting of a subsystem into several spatially separated subsystems, which however could still share an entangled common state. The propagation of the different subsystems will be indicated by the edges of the graph.

There are a number of causal relations between edges and vertices. A vertex v_1 is said to *immediately precede* v_2 if there is a (directed) edge from v_1 to v_2. We write $v_1 \leq v_2$ for the reflexive transitive closure of immediate precedence; thus $v \leq v$ always holds and $v_1 \leq v_2$ means that there is a *directed path* from v_1 to v_2 (possibly of length zero). When $v_1 \leq v_2$ we sometimes say v_1 is "to the past of" v_2 and dually "v_2 is to the future of v_1." When we draw a poset we typically leave out the self-loops and only draw the minimal number of edges needed to infer all the others; the so-called "Hasse diagram" of the poset. We note that our graphs will have initial and final "half-edges", i.e. edges with only one endpoint. Physically we have some quantum states incoming (or "prepared") followed by some interactions and some outgoing state.

The relation between vertices induces a causal relation between edges. We say that an edge e_1 is to the past of another edge e_2 if the terminal vertex of e_1, say v_1 and the initial vertex of e_2, say v_2, satisfy $v_1 \leq v_2$. Note that we could have $v_1 = v_2$. An initial edge is not to the future of any edge, nor is a final edge to the past of any other edge. If two edges are not causally related, we say that they are "spacelike separated" or acausal. Note that two spacelike separated edges could share a common terminal vertex or a common initial vertex, (but since we have a graph, not both). A *space-like slice* is defined as a set of pairwise acausal edges. Henceforth, whenever we say "slice" we will always mean "spacelike slice." Note that the initial (or final) edges form a spacelike slice. We call this the *initial (final) slice*.

For example for the graph of Figure 1 the set of edges $\{e_c, e_d, e_e\}$ form a space-like slice. Another example is the set $\{e_f, e_d, e_e\}$. The edges e_a and e_b form the initial slice. The edges e_a, e_b, e_f and e_g are half-edges, with e_a and e_b initial, and e_f and e_g final.

Associated with any edge e_i is an observer who has access to a subsystem of the complete quantum system. Thus the edges represent local information. Each edge e_i is assigned a density

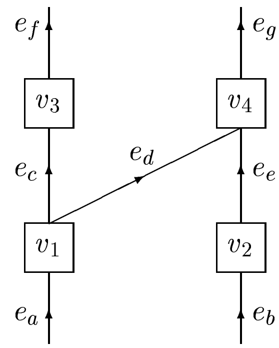


Figure 1:

matrix ρ_i in a Hilbert space $\mathcal{H}_i$[1]. The density matrix ρ_i describes the knowledge about the quantum system available to the local observer at the edge e_i. More generally density matrices will be associated to space-like slices. For a space-like slice consisting of edges $\{e_{i_1}, \ldots e_{i_p}\}$, the assigned density matrix will be denoted $\rho_{i_1,\ldots i_p}$. This density matrix describes the subsystem of the whole quantum system for that space-like slice. Every space-like slice has also a Hilbert space which is the tensor product of the Hilbert spaces of the edges forming the slice. However the density matrix associated with the slice is not in general a tensor product of the density matrices on the edges. If it were, we could not capture non-local quantum correlations.

The graph of Figure 1, represents a quantum system Q which starts evolving from a state in which Q consists of two spatially separated subsystems Q_a and Q_b described by density matrices ρ_a and ρ_b respectively, in Hilbert spaces $\mathcal{H}_a$ and $\mathcal{H}_b$. The initial edges e_a and e_b form the initial slice in this simple system. We will follow the convention that if the initial slice consists of several edges, the initial state of the whole system is a tensor product state, i.e. the subsystems are not entangled. For the above example, $\psi_{init} = \psi_a \otimes \psi_b$ and $\rho_a = |\psi_a\rangle\langle\psi_a|$ and $\rho_b = |\psi_b\rangle\langle\psi_b|$. Entangled subsystems on distinct edges will always have at least one event in the common past. Thus we always explicitly represent the interaction which caused the entanglement.

Each vertex v_i of the graph is labelled with an operator T_i which describes the process taking place at the corresponding event. The operator T_i at a given event v_i takes density matrices on the tensor product of Hilbert spaces living on the incoming edges at v_i to density matrices on the tensor product Hilbert space of outgoing edges. The process at a vertex could be an *intervention*[2] corresponding to a positive operator-valued measure (POVM) [NC00, Per95] or a unitary transformation. Or instead of an external or unitary action there could be several quantum subsystems that come together and then split apart, possibly in a different way. We will consider this last case as a particular instance of a unitary evolution with identity evolution operator. As a simple example, in the case of an event corresponding to unitary evolution by a unitary operator U, we have the usual expression:

$$\rho^{in} \mapsto \rho^{out} = U\rho^{in}U^\dagger \tag{1}$$

The general expression for an operator associated to an event will be discussed fully in the next section, see equation (3).

[1] Throughout the paper, we assume that the graph and the dimensions of all Hilbert spaces are finite.

[2] Interventions are generalized measurements where a quantum subsystem could be discarded [Per00a]. This will be discussed more fully below.

Here we will discuss some of the conditions such a dynamical scheme has to satisfy in order to reflect causality and other physical properties of the quantum system. Causality is the condition that the density matrix on a given edge should not depend on the actions performed at vertices which are acausal to this edge or are in its future. For example, referring back to Figure 1, we would like any quantum evolution rule to say that the density matrix at e_g is unaffected by the intervention at v_3 or the density matrix at e_f is unaffected by the intervention at v_2. A general unitary evolution between the states of two space-like slices is easily shown to violate this condition. Therefore we need to incorporate some sort of locality condition into the evolution scheme.

It is not hard to formulate such an evolution scheme. For example, one could work with the dual picture and have evolution occur along edges with density matrices at the vertices. It is not hard to formulate rules which would enforce causality properly in such a framework. Unfortunately this rules out quantum correlations across spatially separated subsystems. Thus, the evolution scheme cannot be too local because entangled subsystems of the quantum system could fly apart and later come together at a vertex.

Consider the system shown in Figure 2. The quantum system represented in this graph is as

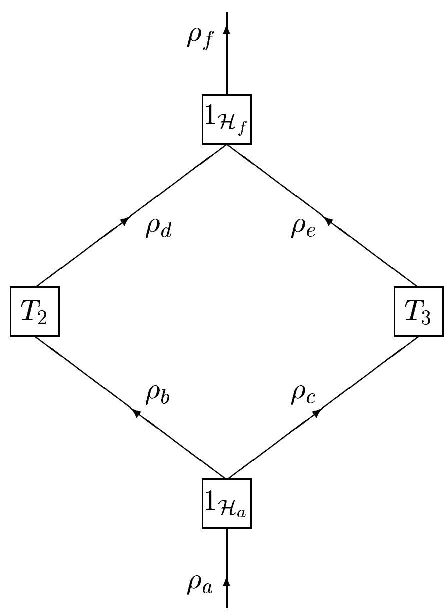


Figure 2:

follows. The system is prepared in a state ψ_a as indicated by the density matrix $\rho_a = |\psi_a\rangle\langle\psi_a|$ on the incoming edge. At the vertex v_1 the system splits into two spatially separated subsystems on the edges e_b and e_c which, in general, are still described by a global entangled state. The local transformations T_2 and T_3 will, in general, preserve the entanglement and the global state will be still entangled on the space-like slice $\{e_d, e_e\}$. The two subsystems come together at the vertex v_4. The two local density matrices ρ_d and ρ_e are not sufficient to reconstruct the entangled state of the system described by ρ_f. The off-diagonal terms of ρ_f are not reflected in the local density matrices, ρ_d and ρ_e. We need to include information about the history of the state on the space-like slice $\{e_d, e_e\}$ in order to reconstruct the global state. One possibility is to work with global space-like slices, and show that the scheme is generally covariant in the sense of being slice-independent. In our functorial approach, certain preferred (not necessarily global) spacelike slices account for all entanglement.

The rules for constructing and labeling the graphs given so far reflect the kinematics of the quantum system. Specifying the dynamics amounts to a prescription for how to obtain the density matrices on every edge from the density matrix on the initial slice and the operators at the vertices of the graph. This prescription will be given below in section 3.

3 Dynamics on Graphs

3.1 Measurements and Interventions

We begin with some standard material on density matrices and positive operator-valued measures (POVMs) [NC00, Pre], before introducing Peres' notion of *intervention operator* [Per00a].

Density matrices are used for describing quantum subsystems which are part of larger quantum systems. In particular a local observer who has access only to a subsystem Q_1 of a quantum system Q will associate a density matrix to his subsystem. Let $\mathcal{H}$ be the Hilbert space of state vectors of Q.

If the overall system Q is in a state described by a wave function $|\psi\rangle \in \mathcal{H}$, then its density Since Q can be decomposed into subsystems, its Hilbert space is a tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ of the Hilbert space $\mathcal{H}_1$ of the subsystem Q_1 and the Hilbert space $\mathcal{H}_2$ describing the remaining degrees of freedom. The density matrix of the subsystem Q_1 is then given by a partial trace with respect the Hilbert space $\mathcal{H}_2$: $\rho_1 = Tr^{\mathcal{H}_2}\rho$. If $\mathcal{H}$ is any Hilbert space, then the space of all density matrices will be denoted $\mathsf{DM}(\mathcal{H})$. For more details on density matrices, see Appendix A.

The *measurement* of a property of a quantum system involves interaction with a classical apparatus. When a classical apparatus measures an observable of a quantum subsystem sitting inside a larger system the appropriate mathematical formalism for such generalized measurement is that of *positive operator-valued measure* or POVM. Let the possible outcomes of the measurement be labelled by the letter $\mu \in \{1 \ldots N\}$. The measurement involves interaction between the apparatus and the quantum system, described by a unitary operator. The classical apparatus has a preferred basis of states indexed by μ. After the measurement, the apparatus appears in one these preferred states. Since we are only interested in describing our quantum subsystem Q_1, we trace out all the remaining degrees of freedom. Effectively to every outcome μ is associated an operator F_μ. The density matrix of Q_1 after the measurement with outcome μ is given by

$$\rho'_\mu = \frac{1}{p_\mu} F_\mu \rho F_\mu^\dagger \tag{2}$$

where ρ is the density matrix before the measurement and p_μ is a numerical factor normalizing the resulting density matrix to unit trace. Consider the family of positive operators $E_\mu = F_\mu^\dagger F_\mu$. For a generalized measurement these have to satisfy the condition $\sum_\mu E_\mu = I$. The probability p_μ for obtaining a measurement result labelled by μ is then given by: $p_\mu = Tr E_\mu \rho$. This justifies the name POVM.

Even more general measurement processes could be considered if the observer discards part of the quantum system during the process of measurement. The appropriate mathematical formalism for describing these generalized measurements is that of *intervention operators* [Per00a]. In the process of measurement, the density matrix changes according to:

$$\rho'_\mu \;=\; \frac{1}{p_\mu} \sum_m A_{\mu m} \; \rho \; A_{\mu m}^\dagger \tag{3}$$

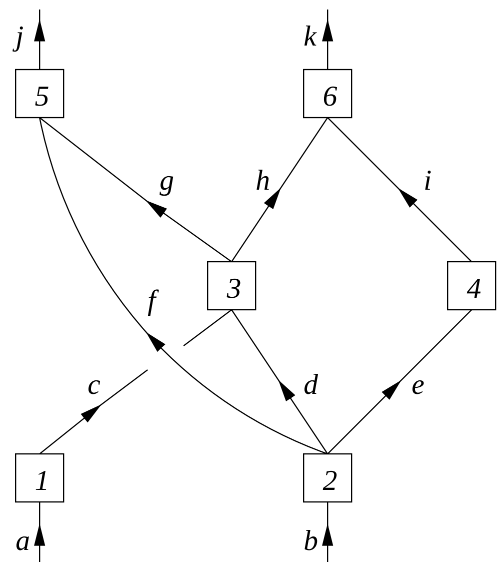


Figure 3:

The families of maps $A_{\mu m}$ now act in general from one Hilbert space to another, i.e for fixed μ and m they correspond to rectangular matrices.

The label μ again distinguishes the set of possible outcomes and the letter m labels the degrees of freedom discarded during this generalized measurement. Since the maps $A_{\mu m}$ come from measurements realized by unitary operator on some larger Hilbert space they again satisfy a completeness condition: $\sum_{\mu m} A^{\dagger}_{\mu m} A_{\mu m} = I$, where I is the identity operator in the appropriate Hilbert space. Notice that if the labels μ and m are absent in (3) the equation describes unitary evolution. Since the events we consider are generalized measurements or unitary evolutions, equation (3) is the appropriate mathematical representation of those processes in full generality. Such maps (3) on density matrices will be called *intervention operators*.

3.2 The dynamical prescription

We are now ready to start discussing the dynamics of a quantum system represented by a dag G. Dynamics will be described by supposing that we are given a density matrix on the initial spacelike slice, and then giving a prescription for calculating the density matrices of future spacelike slices. In essence, we are propagating the initial data throughout the system.

To each vertex $i \in G$ will be assigned an operator T_i, and to each edge e_j will be assigned a Hilbert space $\mathcal{H}_j$. We note that all incoming (or outgoing) edges of a given vertex are pairwise acausal and thus form a spacelike slice. Thus there will be a density matrix ρ_i^{in} associated to the slice of the incoming edges. Then one obtains the density matrix for the slice of the outgoing edges by:

$$\rho_i^{in} = T_i(\rho_i^{out}).$$

Notice that more generally, for two acausal vertices, the sets of incoming or outgoing edges are pairwise acausal. Thus, the associated intervention operators will act on different Hilbert spaces and hence commute.

We begin with an illustrative example. Consider the dag of Figure 3. Given the state on the initial slice, the operators at the events propagate the state to the future. In the example of Figure 3 we have: $\rho_c = T_1(\rho_a)$, $\rho_{fde} = T_2(\rho_b)$. However the next intervention operator T_3 must act on the so far undefined density matrix ρ_{cd}. T_3 takes density matrices on $\mathcal{H}_c \otimes \mathcal{H}_d$ to those on $\mathcal{H}_g \otimes \mathcal{H}_h$. By extending T_3 with the appropriate identity operators, we can view it as a map from $\mathsf{DM}(\mathcal{H}_c \otimes \mathcal{H}_d \otimes \mathcal{H}_e \otimes \mathcal{H}_f)$ to $\mathsf{DM}(\mathcal{H}_e \otimes \mathcal{H}_f \otimes \mathcal{H}_g \otimes \mathcal{H}_h)$. Then we can define the density matrix on another space-like slice, namely $\rho_{fghe} = T_3(\rho_c \otimes \rho_{fde})$. Similarly $\rho_{fdi} = T_4(\rho_{fde})$ and so on. Starting from density matrices on the initial edges and using the intervention operators associated with the vertices - extended with identities as needed - we obtain density matrices on specific space-like slices.

The above inductive process for propagating density matrices can be applied to any system described by a dag. However, the procedure only gives the density matrices for certain spacelike slices within the dag. For example, this procedure does not yet yield a matrix for the slice de. To calculate such density matrices, we will also have to make use of the trace operator. Before extending the procedure to such slices, we first consider those for which the above process is sufficient. We call these slices *locative*.

Definition 3.1 Let G be a dag, and L a slice of G. Consider the set of all vertices V which are to the past of some edge in L. Let I be the set of initial edges in the past of L. Consider all paths of maximal length beginning at an element of I and only going through vertices of V. Then L is *locative* if all such paths end with an edge in L.

In our example, the locative slices are the following:

$$a, b, ab, c, cb, def, adef, cdef, efgh, adfi, cdfi, fghe, fghi, fgk, hej, hij, jk$$

while, for example, de is not locative. Note that the fact that maximal slices are always locative follows immediately from the definition of locative.

We now describe the general rule for calculating the density matrices on locative slices. Associated with each locative slice L is the set I of initial edges in the past of L. We choose a family of slices that begins with I and ends with L in the following way. Consider the set of vertices V between the edges in I and the edges in L. Because L is locative we know that propagating slices forwards through the vertices in V will reproduce L. Let $M \subset V$ be such that the vertices in M are minimal in V with respect to causal ordering. We choose arbitrarily any vertex u in M, remove the incoming edges of u and add the outgoing edges of u to the set I obtaining a new set of edges I_1. It is clear that I_1 is spacelike and locative. Proceeding inductively in this fashion we obtain a sequence of slices $I = I_0, I_1, I_2, \ldots, I_n = L$, where n is the cardinality of V. Of course, this family of slices is far from unique.

The dynamics is obtained as follows. Recall that the states on initial edges are assumed not to be entangled with each other so that one can obtain the density matrix on any set of initial edges, in particular I, by a tensor product. Let ρ_0 be the density matrix on I. We look at the vertex u that was used to go from I to I_1 and apply the intervention operator T assigned to this vertex - possibly augmented with identity operators as in the example above. Proceeding inductively along the family of slices, we obtain the density matrix ρ_n on L.

The important point now is that ρ_n does not depend on the choice of slicing used in going from I to L. This can be argued as follows. Suppose we have a locative slice S and two vertices u and v which are both causally minimal above S and acausal with respect to each other. Then we have four slices to consider, S, S_u, S_v and S_{uv} where by S_u we mean the slice obtained from S by removing the incoming edges of u and adding the outgoing edges of u to S and similarly for the others. It is clear, in this case, that the intervention operators assigned to u and to v commute and the density matrix computed on S_{uv} is independent of whether we evolved along the sequence $S \longrightarrow S_u \longrightarrow S_{uv}$ or $S \longrightarrow S_v \longrightarrow S_{uv}$. Now when we constructed our slices at each stage we had the choice between different minimal vertices to add to the slice. But such vertices are clearly pairwise acausal and hence, by the previous argument applied inductively, the evolution prescription is independent of all possible choices. We summarize this argument in the following proposition

Proposition 3.2 (Covariance) *The density matrix on a locative slice is independent of the family of slices used to compute it.*

So far we have defined density matrices on locative slices only. To define density matrices on general spacelike slices we will need to consider partial tracing operations.

3.3 General Slices

Recall that when one has subsystems Q_1 and Q_2 of a quantum system Q, the Hilbert space for Q may be decomposed as $\mathcal{H}_1 \otimes \mathcal{H}_2$ where $\mathcal{H}_i$ represents Q_i. The density matrix for Q_1 is obtained by tracing over $\mathcal{H}_2$. To obtain the density matrix of a general spacelike slice S we proceeed as follows. First note that there is a natural partial ordering[3] on slices (recall that we always mean "spacelike slices" whenever we say "slices") given as follows

$$S_1 \sqsubseteq S_2 \text{ if } \forall x \in S_1 \exists y \in S_2 . x \leq y \text{ and } \forall y \in S_2 \exists x \in S_1 . x \leq y$$

where $x \leq y$ means that x causally precedes y. The collection of slices forms a finite poset. We find the *minimal* (in the order $\sqsubseteq$) locative slice M that contains S and trace over the Hilbert spaces on edges in $M \setminus S$. Such a locative slice M always exists because maximal spacelike slices are always locative.

Lemma 3.3 *Given a slice S, there is a* unique *minimal slice which contains S.*

Proof . We claim that the locative slice obtained by evolving through those vertices to the causal past of S is the unique minimal locative slice. More precisely, let $V = \{v_1, \dots, v_p\}$ be the set of vertices to the past of S. Let $I = \{e_1, \dots, e_q\}$ be the set of initial edges to the past of S. Constructing a sequence of slicesby incrementally incorporating the vertices of V in a manner similar to what we did in the previous subsection, we get a locative slice M containing S.

First, note that M is locative by definition and also it contains S. Second, note that any slice containing S (in particular S itself) must be to the future of all these vertices. Thus M is indeed minimal since it lies to the future of the fewest vertices. ∎

Definition 3.4 We shall refer to M as the *least locative slice* of the edge S.

Starting with the density matrices on the edges of I and applying the operators associated with the vertices of V, we obtain the density matrix on the locative slice M. Suppose that S consists of edges $\{e_1, \dots, e_i\}$. Let M - the minimal locative slice through S - consist of edges $\{e_1, \dots, e_i, e'_1, \dots, e'_j\}$. The density matrix $\rho_{1,\dots,i,1',\dots,j'}$ on M is an element of the space $End(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_i \otimes \mathcal{H}'_1 \otimes \cdots \otimes \mathcal{H}'_j)$. Let $Tr^{1' \dots j'}$ be the partial trace operation

$$End(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_i \otimes \mathcal{H}'_1 \otimes \cdots \otimes \mathcal{H}'_j) \rightarrow End(\mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_i)$$

which traces over the primed Hilbert spaces.

[3]This ordering is well-known in computation theory where it is called the Egli-Milner order[Plo76].

Definition 3.5 (Density matrix associated with a slice) The density matrix ρ at the slice S is defined to be:

$$\rho_S = Tr^{1'...j'}(\rho_{1,...,i,1',...,j'}). \tag{4}$$

Of course if M and S coincide then no tracing is done.

Proposition 3.6 *The prescription for computing density matrices is causal in the sense that the only vertices to the past of S can affect the density matrix on S.*

Proof . This is immediate from the construction and uniqueness of the minimal locative slice. The explicit construction of the minimal locative slice shows that only intervention operators to the past of S can affect the density matrix on S. ∎

For ease of presentation the rest of the discussion will focus on the case where S is a single edge. In general, an edge e_i is contained in many locative slices and we could just as well have defined ρ_i by tracing over the complementary degrees of freedom in any of these locative slices. However we have to compute the density matrices differently in order to ensure causality. To clarify the discussion consider the quantum system represented by the graph on Figure 4.

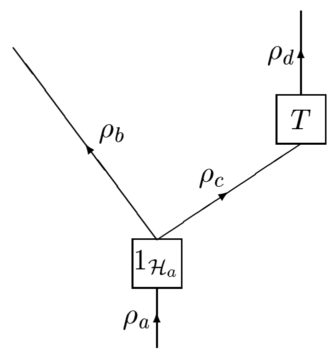


Figure 4:

Let the initial ρ_a be the density matrix of a maximally entangled state of two spin 1/2 subsystems: $\rho_a = |\psi_a\rangle\langle\psi_a|$, where $\psi_a = 1/\sqrt{2}\ (\psi_1^\uparrow \otimes \psi_2^\uparrow + \psi_1^\downarrow \otimes \psi_2^\downarrow)$. At the first vertex the two subsystems separate with no classical intervention. Therefore $\rho_{bc} = \rho_a$. The slice $\{e_b, e_c\}$ is the least locative slice for the edge e_b and we can compute the density matrix associated to this edge: $\rho_b = Tr^c\rho_{bc} = 1/2\ (|\psi_1^\uparrow\rangle\langle\psi_1^\uparrow| + |\psi_1^\downarrow\rangle\langle\psi_1^\downarrow|)$. Next, let the intervention at the second vertex be a measurement on the corresponding subsystem with the result that the spin was found to be in the state $\psi_2^\uparrow$. The intervention operator is the projection operator on this state of the second subsystem: $T(\rho) = 2\ P_2^\uparrow \rho P_2^\uparrow$. We obtain: $\rho_{bd} = T(\rho_{bc}) = (|\psi_1^\uparrow\rangle \otimes\ |\psi_2^\uparrow\rangle)(\langle\psi_1^\uparrow|\ \otimes \langle\psi_2^\uparrow|)$. If now we attempt to trace ρ_{bd} over the subsystem associated with the edge e_d, we will obtain an incorrect result for ρ_b, namely $|\psi_1^\uparrow\rangle\langle\psi_1^\uparrow|$. The resolution is well known. Since a classical observer located on the edge e_b is not aware of the result of the intervention at the second vertex, for him the density matrix ρ_{bd} has evolved from ρ_{bc} by an operator $\tilde{T}$ which includes all possible outcomes of the measurement: $\tilde{\rho}_{bd} = \tilde{T}(\rho_{bc}) = \sum_{s=\uparrow,\downarrow} P_2^s \rho_{bc} P_2^s$. Tracing out the d-subsystem in the expression for $\tilde{\rho}_{bd}$, we obtain the correct expression for ρ_b, namely $\rho_b = 1/2\ (|\psi_1^\uparrow\rangle\langle\psi_1^\uparrow| + |\psi_1^\downarrow\rangle\langle\psi_1^\downarrow|)$.

Now we give another general prescription for computing the density matrix on an edge e_i from an arbitrary locative slice L containing this edge. We first compute a density matrix $\tilde{\rho}_L$ for the slice L. This prescription has to deal with the possibility that some of the intervention operators used are from vertices that are not to the past of the edge e_i. As we saw in the example above, the density matrix at e_i cannot reflect the knowledge of the outcome of interactions at vertices that are not to the causal past of e_i.

This density matrix is computed from the initial data by applying intervention operators for the events in the past of L as before. But now, we will consider two types of events in the past of L: those that are to the past of e_i and those that are not. For the events that are to the past of the edge e_i, we use our regular intervention operators without a summation over the set of possible outcomes: $\rho \mapsto 1/p_\mu \sum_m A_{\mu m} \rho A^\dagger_{\mu m}$. We do not sum over the outcomes in this case precisely because the outcome is in fact known at e_i. For the events that are to the past of the slice L but not to the past of the edge e_i, we use operators that sum over all possible outcomes: $\rho \mapsto \sum_{\mu m} A_{\mu m} \rho A^\dagger_{\mu m}$. This time, of course, the summation is there because the outcome cannot be known at e_i since these events are not to the past of e_i.

After we have obtained $\tilde{\rho}_L$, we trace out those subsystems associated with edges in L except for e_i to obtain the density matrix $\tilde{\rho}_i$. This is the density matrix associated with our preferred edge e_i, as computed from the slice L. The independence of the result on the choice of L is expressed in the following proposition:

Proposition 3.7 *Let e_i be an edge in the dag G. The density matrix ρ_i associated with the edge e_i does not depend on the choice of locative slice used to compute it; provided that we use the second prescription to compute it.*

Proof. We have already demonstrated that to any edge e_i, there is a unique least locative slice M_i containing e_i. Let ρ_i be the density matrix for the edge e_i as computed from the least locative slice and let $\tilde{\rho}_i$ be the density matrix for the same edge but computed from an arbitrary locative slice, say L, containing e_i. We will prove the proposition by showing that $\rho_i = \tilde{\rho}_i$.

First note that M_i being less than L implies that there is a set V of events between M_i and L. The plan is to remove the effect of these events and show that, at each stage, the density matrix is unaffected. We begin by picking a maximal event, say k, with the intervention operator T_k. Since k is maximal and hence acausal with all other maximal elements of V, as well as with all the maximal elements to the past of e_i, the intervention operator at k commutes with all the intervention operators at the vertices just mentioned. Thus, we can choose the intervention operator T_k to be the outermost, i.e. the density matrix ρ_L obtained by propagating to L can be written as

$$\rho_L = T_k(\rho')$$

where ρ' is the density matrix on the (locative) slice obtained by removing the edges to the future of k from L and adding the edges to the past of k. Using the explicit general form for an intervention operator,

$$\rho_L = \sum_{\mu,m} A^{(k)}_{\mu,m} \rho' A^{\dagger(k)}_{\mu,m}.$$

In order to obtain the density matrix $\tilde{\rho}_i$, we trace over all Hilbert spaces associated with edges in L except e_i. In particular, we trace over the outgoing edges associated with k. Now we can use the cyclic property of trace and rewrite the expression for $\tilde{\rho}_i$ as,

$$\tilde{\rho}_i = Tr(\sum_{\mu,m} A^{\dagger(k)}_{\mu,m} A^{(k)}_{\mu,m} \rho').$$

Now we use the identity

$$\sum_{\mu m} A^{\dagger}_{\mu m} A_{\mu m} = I$$

to get

$$\tilde{\rho}_i = Tr(\rho').$$

We have eliminated the effect of the intervention operator at k. Proceeding inductively we can peel off the intervention operators associated with the rest of the vertices in V, thus

$$\tilde{\rho}_i = \rho_i.$$

■

A similar argument for the case of a simple system represented by the dag in Figure 2 is contained in [Per00b].

4 Polycategories

We now wish to give a more axiomatic treatment of the above construction. This will require the use of several concepts from category theory and logic, which we now present.

We begin by introducing the algebraic or categorical concepts necessary for our formulation of the dynamics of quantum information flow. While it might seem that these structures are excessively abstract, this level of abstraction has several advantages. First, it provides a great deal of generality. Our definition can be applied in many contexts, in particular it may be applied in situations other than the sorts of information flow considered here. Second, a crucial property - namely covariance - now becomes straightforward consequences of functoriality.

4.1 Posets, directed graphs and categories

For comparison, we recall briefly that a poset is a set P together with a binary relation on P (i.e. a subset of $P \times P$) denoted $\leq$ that satisfies the properties of antisymmetry, transitivity and reflexivity. It is a natural generalization of this idea to consider *directed graphs*. A directed graph is simply a set D, the set of *vertices* or *nodes*, together with a binary relation R on D. No properties of R are required in the definition of directed graph. In particular there is no implicit transitivity assumed. A directed graph has a natural geometric visualization. One considers the nodes as points in the plane, and if x and y are nodes with $\langle x, y \rangle \in R$, we draw an arrow from x to y.

As already remarked, the nodes of our directed graph will be events, and arrows will represent propagation from one event to another. To avoid temporal loops, we will add the single requirement that our directed graphs be *acyclic*, i.e. there does not exist a sequence of edges $x_1, x_2, \ldots, x_n$ such that for all $i \in \{1, 2, \ldots, n-1\}$, we have $\langle x_i, x_{i+1} \rangle \in R$, and $x_1 = x_n$. This of course corresponds to there being no directed cycles in the geometric representation. Hereafter, a directed acyclic graph will be called a *dag*. Note that every poset, considered as a directed graph, is acyclic. This is a consequence of transitivity and antisymmetry. But dags are a genuine generalization of posets.

This difference will become more apparent when we consider the space of *paths*. In a poset all the paths are already included (even if they are not explicitly drawn in the visualization of the poset). When we consider paths through a dag we may have multiple paths between the same two vertices. These multiple paths represent different ways that information flowed from one point to another, thus, we must regard them as distinct. Therefore - unlike the case with posets - we do not just want to regard the resulting structure as a binary relation, rather, we want to view it as a category.

It is natural to associate to any dag D, indeed to any directed graph, a category. We first briefly remind the reader of the basic definitions. See [Mac98] for a more extensive introduction.

Definition 4.1 A **category** C consists of two collections, the collection of *objects* and the collection of *morphisms*. Each morphism is assigned a domain and codomain, both being objects of C. Typically we write $f\colon A \to B$ to mean f is a morphism with domain A and codomain B. To every object A, we have a special morphism, the identity $id\colon A \to A$. There is also a composition law which takes morphisms $f\colon A \to B$ and $g\colon B \to C$ and returns a morphism $gf\colon A \to C$. All this data must satisfy several evident equations, as described for example in [Mac98]. We also remind the reader that a **functor** is a morphism of categories, i.e. a functor, denoted $F\colon C \to D$ consists of a function taking objects $c \in C$ to objects $F(c) \in D$ and taking morphisms $f\colon c \to d$ in C to morphisms $F(f)\colon F(c) \to F(d)$. A functor must preserve identities and composition.

To each dag D, we associate a category $\mathsf{C}(D)$. This is the category *freely generated* by the dag. See for example [Mac98] Chapter 2, for a detailed description. The objects of our free category will be the vertices of D. If x and y are vertices, a morphism from x to y is a directed path in our dag. Identities are paths of length 0, and composition is given by concatenation of paths. The verification of the axioms for a category is straightforward.

One of the key points of our work is that we are proposing passing from posets to categories. As we have remarked before, categories are more general than posets, indeed posets correspond to a degenerate class of categories in which there is at most one morphism between any two objects. The richer structure of categories allows us to retain more information about the system. Intuitively, the use of categories allows us not merely to note that x causally precedes y, but to keep track of the different ways that x may evolve into y. To make this more precise, we need a slightly different construction on dags, which will yield polycategories as opposed to categories.

4.2 Polycategories

Roughly speaking, the distinction between categories and polycategories is the following: A category allows one to have morphisms which go from single objects to single objects. A polycategory allows one to have morphisms from lists of objects to lists of objects. A typical morphism in a polycategory (hereafter called a polymorphism) would be denoted:

$$f\colon A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_m$$

There are a number of contexts in which such a generalization would be useful. Before giving the formal definition, we discuss two such contexts. The first arises in algebra. Consider Hilbert spaces, vector spaces or any class of modules in which one can form a tensor product. Then we can define a polycategory as follows. Our objects will be such spaces, and a morphism of the above form will be a linear function:

$$f\colon A_1 \otimes A_2 \otimes \ldots \otimes A_n \longrightarrow B_1 \otimes B_2 \otimes \ldots \otimes B_m$$

Thus polycategories have proven to be quite useful in the analysis of (ordinary) categories in which one can form tensor products of objects. Indeed this was the original motivation for their definition. See [Lam69, Sza75]. Categories in which one has a reasonable notion of tensor product are called *monoidal*, and have recently figured prominently in several areas of mathematical physics, most notably topological quantum field theory [Ati90, BD95].

The second well-known application of polycategories is to logic. Typically logicians are interested in the analysis of *sequents*, written:

$$A_1, A_2, \ldots, A_n \vdash B_1, B_2, \ldots, B_m$$

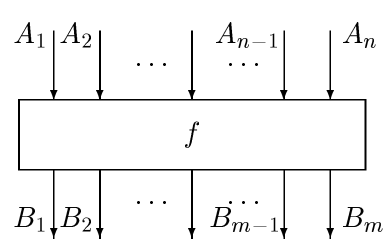
Now $A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m$ represent formulas in some logical system. We say that the above sequent holds if and only if the conjunction of $A_1, A_2, \ldots, A_n$ logically entails the disjunction of $B_1, B_2, \ldots, B_m$. There is a well-established correspondence between the sort of logical entailments considered here and categorical structures. See for example [LS86].

But notice the difference between this and our first example. When talking about vector spaces, the "commas" on the left and right were both interpreted as the tensor product. However in the logic example, we have two different interpretations. Commas on the left are treated as conjunction, while commas on the right are treated as disjunction. Thus for a proper categorical interpretation of polycategories, one needs categories with two monoidal structures which interact in an appropriate fashion. Such categories are called *linearly* or *weakly distributive*, a notion due to Cockett and Seely [CS97, BCST96]. Linearly distributive categories are the appropriate framework for considering a specific logical system known as *linear logic*, introduced by Girard [Gir87, Gir89]. For a brief exposition of linear logic, see Appendix B. As we will see, the refined logical connectives of linear logic will be used to express the entanglements of our system.

There is a very geometric or graphical calculus for representing morphisms in polycategories, which was introduced by Joyal and Street in [JS91], and given a logical interpretation in [BCST96]. A polymorphism of the form:

$$f\colon A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_m$$

is represented as follows:



Thus the polymorphism is represented as a box, with the incoming and outgoing arrows labelled by objects. Composition in polycategories then can be represented pictorially in a very natural fashion. Before giving a general discussion of composition in a polycategory, we illustrate this graphical representation. Suppose we are given two polymorphisms of the following form:

$$\begin{aligned} f\colon A_1, A_2, \ldots, A_n &\longrightarrow B_1, B_2, \ldots, B_m, C \\ g\colon C, D_1, D_2, \ldots, D_k &\longrightarrow E_1, E_2, \ldots, E_j \end{aligned}$$

Note the single object C common to the codomain of f and the domain of g. Then under the definition of polycategory, we can compose these to get a morphism of form:

$$g \circ_C f\colon A_1, A_2, \ldots, A_n, D_1, D_2, \ldots, D_k \longrightarrow B_1, B_2, \ldots, B_m, E_1, E_2, \ldots, E_j$$

The object C which "disappears" after composition is called the *cut object*, a terminology derived from logic. Note that we subscript the composition by the object being cut. This composition would be represented by the diagram on Figure 5:

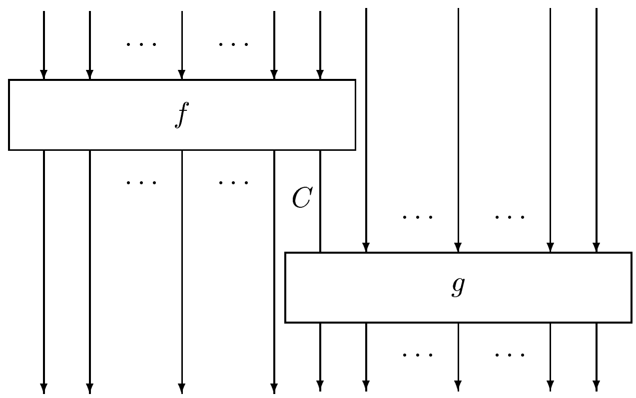


Figure 5:

We only label the segment corresponding to the cut object, for ease of reading. Thus composition in a polycategory is represented by the concatenation of the graphs of f and g, followed by joining the incoming and outgoing edges corresponding to the cut object. There are several other possibilities for applications of the composition rule. In some cases, the graphical representation requires our arrows to cross. This corresponds to having a *symmetric* polycategory. This is very much related to having a symmetric tensor or tensors, i.e. ones with the property that $A \otimes B \cong B \otimes A$. We will always assume our polycategories are symmetric.

We now give a more formal definition of polycategory. We refer the reader to [CS97, Sza75] for further details.

Definition 4.2 A **polycategory** C consists of the following data:

- A set of objects, denoted $|\mathsf{C}|$.
- If $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_m$ are finite sequences of objects, then we have a set of morphisms of the form $f\colon A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_m$. We note that technically one must consider these sequences of objects as being defined only up to permutation.
- For every object A, we have an identity morphism $id_A\colon A \to A$.

The composition law was already described pictorially. The data of course are subject to a number of axioms, of which most important for us is the one which requires associativity of composition. The notion of *polyfunctor* between polycategories is also straightforward to formulate. One first has a function F taking objects to objects, and then given a morphism $f\colon A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_m$, one assigns to it a morphism

$$F(f)\colon F(A_1), F(A_2), \ldots, F(A_n) \longrightarrow F(B_1), F(B_2), \ldots, F(B_m). \tag{5}$$

Again, a number of axioms must be satisfied, in particular the polyfunctor must commute with the composition of polymorphisms.

As suggested by the above, there is a relationship between polycategories and monoidal categories. It is summarized in the following lemma, which can be found for example in [CS97]:

Proposition 4.3 *Let* C *be a monoidal category. Then one can associate to* C *a polycategory (which will typically be denoted by* $P(\mathsf{C})$ *as follows:*

- *The objects of* $P(\mathsf{C})$ *will be the same as those of* C*.*
- *A polymorphism of the form* $f\colon A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_m$ *is a morphism* $f\colon A_1 \otimes A_2 \otimes \ldots \otimes A_n \longrightarrow B_1 \otimes B_2 \otimes \ldots \otimes B_m$*.*
- *Composition is induced by the composition in* C *in the following way. Suppose that we have two polymorphisms in* $P(\mathsf{C})$ *as follows:*

$$\begin{array}{c} f\colon A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_m, C \\ g\colon C, D_1, D_2, \ldots, D_k \longrightarrow E_1, E_2, \ldots, E_j \end{array}$$

Then since we are in a monoidal category, we have morphisms

$$\begin{array}{c} f\colon A_1 \otimes A_2 \otimes \ldots \otimes A_n \longrightarrow B_1 \otimes B_2 \otimes \ldots \otimes B_m \otimes C \\ g\colon C \otimes D_1 \otimes D_2 \otimes \ldots \otimes D_k \longrightarrow E_1 \otimes E_2 \otimes \ldots \otimes E_j \end{array}$$

The composite in $P(\mathsf{C})$ *is then given by:*

$$g \circ_C f = (id_{B_1 \otimes B_2 \otimes \ldots \otimes B_m} \otimes g) \circ (f \otimes id_{D_1 \otimes D_2 \ldots \otimes D_k}) \tag{6}$$

We note that the concepts of polycategory and monoidal category are not equivalent. To obtain an equivalence, one needs to replace monoidal categories with the more general notion of linearly distributive category mentioned above.

Now we will demonstrate that a dag generates a polycategory. In this construction, the nodes of the dag will be assigned morphisms and the edges will be assigned objects.

We consider the dag example of Figure 6. We have changed labels to be more appropriate for the present discussion.

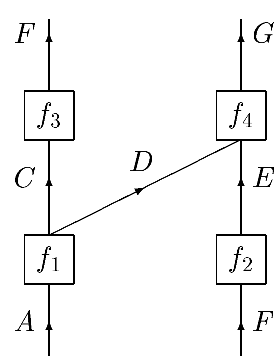


Figure 6:

The idea behind the construction is that the nodes of the dag (the boxes in our picture) will correspond to polymorphisms. For example, in the above picture, the box f_1 determines a polymorphism:

$$f_1\colon A \longrightarrow C, D$$

Similarly, f_4 determines a polymorphism $f_4\colon D, E \to G$. Thus we see that one has a polymorphism corresponding to each node. The domain of that polymorphism will be the labels of the incoming arrows, and the codomain is determined by the labels of the outgoing arrows. These are the basic morphisms of the polycategory. As in the previous construction, one must adjoin morphisms corresponding to the allowable compositions. For example, in the above case, we can compose the morphisms f_4 and f_1 along the cut object D to obtain a new polymorphism $f_4 \circ_D f_1\colon A, E \to C, G$. One must also add identities and must force these composites to satisfy the appropriate equations. This construction yields the *polycategory freely generated by the dag.* More generally, we would have the following definition.

Definition 4.4 We suppose that we are given a finite dag G. The *free polycategory generated by* G, denoted $P(G)$, is defined as follows. If a given vertex v has incoming edges $A_1, A_2, \ldots, A_n$ and outgoing edges $B_1, B_2, \ldots, B_m$ then the polycategory will have a polymorphism of the form $f_v\colon A_1, A_2, \ldots, A_n \to B_1, B_2, \ldots, B_m$. In general by induction, if $P(G)$ has polymorphisms of the form:

$$\begin{aligned} f\colon A_1, A_2, \ldots, A_n &\longrightarrow B_1, B_2, \ldots, B_m, C \\ g\colon C, D_1, D_2, \ldots, D_k &\longrightarrow E_1, E_2, \ldots, E_j \end{aligned}$$

then we require the existence of a composite $g \circ_C f$ as a new polymorphism. We assume the existence of an identity morphism for each edge of G. Finally we impose on this data the necessary equations implied by the definition of polycategory.

4.3 Categories of interventions

Next we describe an appropriate for our construction polycategory of intervention operators; there are several reasonable choices, this being the most straightforward. We start with the well known fact that the category Hilb of Hilbert spaces and bounded linear operators is a monoidal category. Hence by the construction of lemma 4.3, we obtain a polycategory. However this is not the category we will ultimately use. We will introduce a category Conj. Intuitively, the objects are Hilbert space endomorphisms and morphisms are conjugations. A more formal definition is as follows. Objects are finite-dimensional Hilbert spaces. A morphism from $\mathcal{H}_1$ to $\mathcal{H}_2$ is a finite family of maps $\{A_i\}_{i\in I}$ of linear morphisms $A_i\colon \mathcal{H}_1 \to \mathcal{H}_2$. Composition is then described as follows. If we have the following pair of maps:

$$\mathcal{H}_1 \xrightarrow{\{A_i\}_{i\in I}} \mathcal{H}_2 \xrightarrow{\{B_j\}_{j\in J}} \mathcal{H}_3$$

then the composite is:

$$\mathcal{H}_1 \xrightarrow{\{B_j \circ A_i\}_{\langle i,j\rangle \in I\times J}} \mathcal{H}_3$$

A morphism in Conj can be seen as taking endomorphisms of $\mathcal{H}_1$ to endomorphisms of $\mathcal{H}_2$ by the formula $\mathcal{O} \mapsto \sum_m A_m \mathcal{O} A_m^\dagger$. The monoidal structure on Hilb lifts to a monoidal structure on the category Conj. The tensor product operator is the usual tensor product of operators on Hilbert spaces, on maps we take all possible pairings. We next restrict the class of morphisms by considering only those families such that the corresponding conjugation is trace preserving. We call the resulting category Dio. This category also inherits a monoidal structure. As discussed in Lemma 4.3 any monoidal category canonically gives rise to a polycategory associated to it. We will denote by $\mathcal{P}(\mathsf{Dio})$ the polycategory associated with Dio.

5 The logic of polycategories

While definition 4.4 gives the free polycategory generated by a dag G, it will prove to be useful to have a more constructive description. Proof-theoretic techniques have proven to be useful in describing free polycategories. In our case, the logical structures necessary are quite simple, and so we digress briefly to put definition 4.4 in logical terms. Recall that one of the common interpretations of a polymorphism is as a logical sequent[4] of the form:

$$A_1, A_2, \ldots, A_n \vdash B_1, B_2, \ldots, B_m$$

Our system will have only one inference rule, called the *Cut rule*, which states:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma', A \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}$$

This should be interpreted as saying that if one has derived the two sequents above the line, then one can infer the sequent below the line. Proofs in the system always begin with *axioms*. Axioms are of the form $A_1, A_2, \ldots, A_n \vdash B_1, B_2, \ldots, B_m$, where $A_1, A_2, \ldots, A_n$ are the incoming edges of some vertex in our dag, and $B_1, B_2, \ldots, B_m$ will be the outgoing edges. There will be one such axiom for each vertex in our dag. For example, consider Figure 3. Then we will have the following axioms:

$$a \overset{1}{\vdash} c \quad b \overset{2}{\vdash} d,e,f \quad c,d \overset{3}{\vdash} g,h \quad e \overset{4}{\vdash} i \quad f,g \overset{5}{\vdash} j \quad h,i \overset{6}{\vdash} k$$

where we have labelled each entailment symbol with the name of the corresponding vertex. The following is an example of a deduction in this system of the sequent $a, b \vdash f, g, h, i$.

$$\dfrac{\dfrac{b \vdash d,e,f \qquad \dfrac{a \vdash c \quad c,d \vdash g,h}{a,d \vdash g,h}}{a,b \vdash e,f,g,h} \qquad e \vdash i}{a,b \vdash f,g,h,i}$$

This deduction corresponds to the fact that in the free polycategory generated by this dag, one has a morphism $a, b \to f, g, h, i$. In fact, it is easy to see that there is a precise correspondence between deductions in this logical system and nonidentity morphisms in the free polycategory.

As a first attempt at capturing quantum evolution on a dag G axiomatically, one might consider taking a polyfunctor from $P(G)$ to $P(\mathsf{Hilb})$, where Hilb is the usual category of finite-dimensional Hilbert spaces with its usual tensor product. Note that such a polyfunctor must necessarily take a sequence of, say, incoming edges $A_1, A_2, \ldots, A_n$ to $\mathcal{H}_1 \otimes \mathcal{H}_2 \ldots \otimes \mathcal{H}_n$ where $\mathcal{H}_i$ corresponds to A_i. Then one would (tentatively) define a set Δ of edges to be *valid* if there is a deduction in

[4] We note that for purposes of this paper sequents should always be considered "up to permutation", i.e. one may rearrange the order of premises and conclusions as one sees fit.

the logic generated by G of $\Gamma \vdash \Delta$ where Γ is a set of initial edges. Equivalently there must be a morphism $\Gamma \to \Delta$ in $P(G)$. Then the polyfunctor would take this to a morphism of Hilbert spaces $T\colon \mathcal{H}_\Gamma \to \mathcal{H}_\Delta$. The initial density matrices would always be assumed to be given, and one would just apply T to the appropriate initial density matrices to obtain the density matrix associated to Δ. The locative slices are the ones on which density matrices can be obtained without the trace operation and we are looking to equate the notions of locative and valid for slices. This approach would be genuinely axiomatic, and would evidently be applicable to other situations by simply using a category other than Hilbert spaces as the target of the polyfunctor. Furthermore we would suggest that using logic as the means of calculating the matrices gives the approach a very canonical flavor.

However, with this notion of validity, we would fail to capture all locative slices, and thus our tentative notion of validity will have to be modified. For example, consider the dag underlying the system of Figure 2 shown in Figure 7.

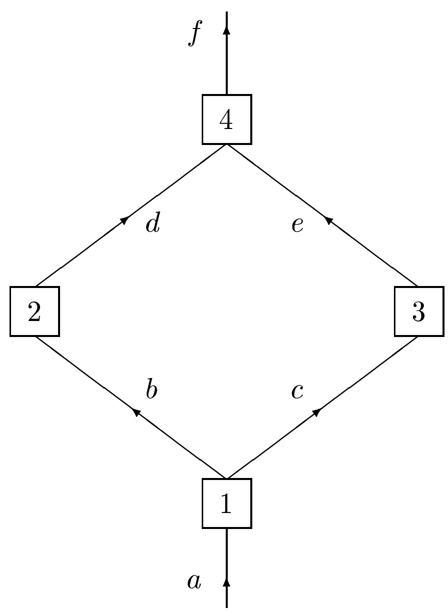


Figure 7:

Corresponding to this dag, we get the following basic morphisms (axioms):

$$a \vdash b, c \quad b \vdash d \quad c \vdash e \quad d, e \vdash f.$$

Evidently, the set $\{f\}$ is a locative slice, and yet the sequent $a \vdash f$ is not derivable. The sequent $a \vdash d, e$ is derivable, and one would like to cut it against $d, e \vdash f$, but one is only allowed to cut a single formula. Such "multicuts" are expressly forbidden, as they lead to undesirable logical properties [Blu93].

Physically, the reason for this problem is that the sequent $d, e \vdash f$ does not encode the information that the two states at d and e are correlated. It is precisely the fact that they are correlated that implies that one would need to use a multicut. To avoid this problem, one must introduce some notation, specifically a syntax for specifying such correlations. We will use the *logical connectives* of the multiplicative fragment of *linear logic* [Gir87, Gir95] to this end. The multiplicative disjunction of linear logic, denoted $\wp$ and called the *par* connective, will express such nonlocal correlations. In our example, we will write the sequent corresponding to vertex 4 as $d \wp e \vdash f$ to express the fact that the subsystems associated with these two edges are possibly entangled through interactions in their common past.

Note that whenever two (or more) subsystems emerge from an interaction, they are correlated. In linear logic, this is reflected by the following rule called the (right) *Par rule*:

$$\frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \wp B}$$

Thus we can always introduce the symbol for correlation in the right hand side of the sequent.

Notice that we can cut along a compound formula without violating any logical rules. So in the present setting, we would have the following deduction:

$$\dfrac{\dfrac{\dfrac{\dfrac{a \vdash b, c \quad b \vdash d}{a \vdash c, d} \qquad c \vdash e}{a \vdash d, e}}{a \vdash d \wp e} \qquad d \wp e \vdash f}{a \vdash f}$$

All the cuts in this deduction are legitimate; instead of a multicut we are cutting along a compound formula in the last step. So the first step in modifying our general prescription is to extend our polycategory logic, which originally contained only the cut rule, to include the connective rules of linear logic. These are described in Appendix B.

The above logical rule determines how one introduces a par connective on the righthand side of a sequent. For the lefthand side, one introduces pars in the axioms by the following general prescription. Given a vertex in a multigraph, we suppose that it has incoming edges $a_1, a_2, \ldots, a_n$ and outgoing edges $b_1, b_2, \ldots, b_m$. In the previous formulation, this vertex would have been labelled with the axiom $\Gamma = a_1, a_2, \ldots, a_n \vdash b_1, b_2, \ldots, b_m$. We will now introduce several pars ($\wp$) on the lefthand side to indicate entanglements of the sort described above. Begin by defining a relation $\sim$ by saying $a_i \sim a_j$ if there is an initial edge c and directed paths from c to a_i and from c to a_j. This is not an equivalence relation, but one takes the equivalence relation generated by the relation $\sim$. Call this new relation $\cong$. This equivalence relation, like all equivalence relations, partitions the set Γ into a set of equivalence classes. One then "pars" together the elements of each equivalence class, and this determines the structure of the lefthand side of our axiom. For example, consider vertices 5 and 6 in Figure 3. Vertex 5 would be labelled by $f \wp g \vdash j$ and vertex 6 would be labelled by $h \wp i \vdash k$. On the other hand, vertex 3 would be labelled by $c, d \vdash g, h$.

Just as the par connective indicates the existence of past correlations, we use the more familiar tensor symbol $\otimes$, which is also a connective of linear logic, to indicate the lack of nonlocal correlation. This connective also has a logical rule:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma' \vdash \Delta', B}{\Gamma, \Gamma' \vdash \Delta, \Delta', A \otimes B}$$

But we note that unlike in ordinary logic, this rule can only be applied in situations that are physically meaningful. We will say that two deductions π and π' are *spacelike separated* if all the the vertices of π and π' are pairwise spacelike separated. In the above formula, we require that the deductions of $\Gamma \vdash \Delta, A$ and $\Gamma' \vdash \Delta', B$ are spacelike separated. This restriction of application of inference rules is similar to the restrictions of *ludics* [Gir01]. From a categorical standpoint, the restrictions imply that the connectives are only partial functors, but this is only a minor issue.

Summarizing, to every dag G we associate its "logic", namely the edges are considered as formulas and vertices are axioms. We have the usual linear logical connective rules, including the cut rule which in our setting is interpreted physically as propagation. The par connective denotes correlation, and the tensor lack of correlation. Note that every deduction in our system will conclude with a sequent of the form $\Gamma \vdash \Delta$, where Γ is a set of initial edges.

Now one would like to modify the definition of validity to say that a set of edges Δ is *valid* if in our extended polycategory logic, one can derive a sequent $\Gamma \vdash \hat{\Delta}$ such that the list of edges appearing in $\hat{\Delta}$ was precisely Δ, and Γ is a set of initial edges. However this is still not sufficient as an axiomatic approach to capturing all locative slices. We note the example in Figure 8.

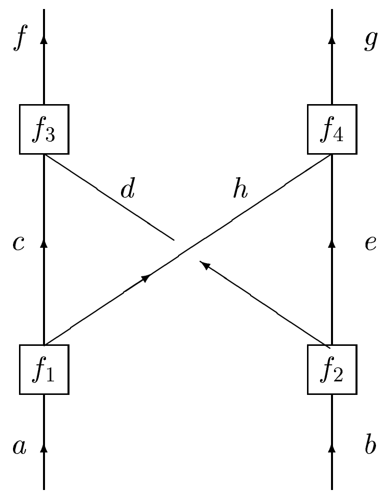


Figure 8:

Evidently the slice $\{f, g\}$ is locative, but we claim that it cannot be derived even in our extended logic. To this directed graph, we would associate the following axioms:

$$a \vdash c, h \quad b \vdash d, e \quad c, d \vdash f \quad h, e \vdash g$$

Note that there are no correlations between c and d or between h and e. Thus no $\invamp$-combinations can be introduced. Now if one attempts to derive $a, b \vdash f, g$, we proceed as follows:

$$\dfrac{\dfrac{a \vdash c, h \quad b \vdash d, e}{a, b \vdash c \otimes d, h, e} \quad \dfrac{c, d \vdash f}{c \otimes d \vdash f}}{a, b \vdash h, e, f}$$

At this point, we are unable to proceed. Had we attempted the symmetric approach tensoring h and e together, we would have encountered the same problem.

The problem is that our logical system is still missing one crucial aspect, and that is that correlations develop dynamically as the system evolves, or equivalently as the deduction proceeds. Thus our axioms must change dynamically as well. We give the following definition.

Definition 5.1 Suppose we have a deduction π of the sequent $\Gamma \vdash \Delta$ in the graph logic associated to the dag G, and that T is a vertex in G to the future or acausal to the edges of the set Δ with a

and b among the incoming edges of T. Then a and b are *correlated* with respect to π if there exist outgoing edges c and d of the proof π and directed paths from c to a and from d to b.

So the point here is that when performing a deduction, one does not assign an axiom to a given vertex until it is necessary to use that axiom in the proof. Then one assigns that axiom using this new notion of correlation and the equivalence relation defined above. This prescription reflects the physical reality that entanglement of local quantum subsystems could develop as a result of a distant interaction between some other subsystems of the same quantum system. We are finally able to give the following crucial definition:

Definition 5.2 A set Δ of edges in a dag G is said to be *valid* if there is a deduction in the logic associated to G of $\Gamma \vdash \hat{\Delta}$ where $\hat{\Delta}$ is a sequence of formulas whose underlying set of edges is precisely Δ and where Γ is a set of initial edges, in fact the set of initial edges to the past of Δ.

We are also ready to state the result relating the logical deduction and the dynamics of Section 3.2 in a graph.

Theorem 5.3 *A set of edges is valid if and only if it is locative. More specifically, if there is a deduction of* $\Gamma \vdash \hat{\Delta}$ *as described above, then* Δ *is necessarily locative. Conversely, given any locative slice, one can find such a deduction.*

Proof R. ecall that a locative slice L is obtained from the set of initial edges in its past by an inductive procedure. At each step, we choose arbitrarily a minimal vertex u in the past of L, remove the incoming edges of u and add the outgoing edges. This step corresponds to the application of a cut rule, and the method we have used of assigning the par connective to the lefthand side of an axiom ensures that it is always a legal cut. The tensor rule is necessary in order to combine spacelike separated subsystems in order to prepare for the application of the cut rule. ∎

Thus we have successfully given an axiomatic logic-based approach to describing evolution. In summary, to find the density matrix associated to a locative slice Δ, one finds a set of linear logic formulas whose underlying set of atoms is Δ and a deduction of $\Gamma \vdash \hat{\Delta}$ where Γ is as above. This deduction is interpreted as a morphism in the corresponding polycategory, and the polyfunctor to $\mathcal{P}(\mathsf{Dio})$ is applied to obtain a morphism in the category Dio. One then plugs in the given initial data to obtain the density matrix corresponding to that slice. Given a nonlocative slice, one simply finds a locative slice containing it, repeats the above procedure and then traces out the extraneous edges.

6 Conclusions

We have presented an axiomatic system for the analysis of quantum evolution. The dynamics is local as to preserve causality, but at the same time entanglement of separated quantum systems is faithfully represented. One could apply these ideas to other situations by using a category other than the category of intervention operators as the target of the functor. An appropriate categorical structure for the target is the notion of a *traced monoidal category* [AJV96] or the notion of a *traced ideal* [ABP99]. See also [BCS00]. One particular situation which might be analyzed in this framework is the notion of classical probabilistic information. The paper [ABP99] contains a category of *probabilistic relations* which might be of particular interest in this setting.

Our work also suggests a natural extension of the notion of *consistent* or *decoherent histories* [GMH93, Gri96]. Restricting the intervention operators at the vertices of our graph G to be projection operators we can consider G to denote a particular history within a set of histories. This relaxes the usual linear ordering of events considered in the literature thus far. An exposition of histories on graphs is under preparation.

Acknowledgements

The authors would like to thank NSERC for its financial support. We would also like to thank Phil Scott and Jean-Yves Girard for inviting us to present this work at the joint SMF-AMS conference in Lyon. The paper [Mar00], which led to our initial consideration of these ideas, was pointed out to us by Ioannis Raptis. We would like to especially thank Rafael Sorkin for a lengthy discussion on causal sets and related topics and for pointing out that an earlier proposal would not allow entanglement. Finally, the second author would like to thank the University of Ottawa Department of Mathematics and Statistics for its support.

References

[ABP99] S. Abramsky, R. Blute, and P. Panangaden. Nuclear and trace ideals in tensor-$*$ categories. *Journal of Pure and Applied Algebra*, 143(1-3):3–47, 1999.

[AJV96] Ross Street Andre Joyal and Dominic Verity. Traced monoidal categories. *Math. Proc. Camb. Phil. Soc.*, 119:447–468, 1996.

[Ati90] M. Atiyah. *The geometry and physics of knots.* Lezioni Lincee, 1990.

[Bar79] M. Barr. *$*$-Autonomous Categories.* Number 752 in Lecture Notes in Mathematics. Apringer-Verlag, 1979.

[BCS00] R. Blute, J.R.B. Cockett, and R.A.G. Seely. Feedback for linearly distributive categories: traces and fixpoints. *Journal of Pure and Applied Algebra*, 154:27–69, 2000.

[BCST96] R. Blute, J.R.B. Cockett, R.A.G. Seely, and T. Trimble. Natural deduction and coherence for weakly distributive categories. *Journal of Pure and Applied Algebra*, 113:229–296, 1996.

[BD95] J. Baez and J. Dolan. Higher-dimensional algebra and topological quantum field theory. *J. Math. Phys*, 36:6073–6105, 1995.

[Ble91] M. Blencowe. The consistent histories interpretation of quantum fields in curved spacetime. *Annals of Physics*, 211:87–111, 1991.

[BLMS87] L. Bombelli, J. Lee, D. Meyer, and R. Sorkin. Spacetime as a causal set. *Physical Review Letters*, 59:521–524, 1987.

[Blu93] R. Blute. Linear logic, coherence and dinaturality. *Theoretical Computer Science*, 115:3–41, 1993.

[CS97] J.R.B. Cockett and R.A.G. Seely. Weakly distributive categories. *Journal of Pure and Applied Algebra*, 114:133–172, 1997.

[Gir87] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[Gir89] J.-Y. Girard. Geometry of interaction i: Interpretation of system **F**. In *Logic Colloquium '88*. North Holland, 1989.

[Gir95] J.-Y. Girard. Linear logic: its syntax and semantics. In J.-Y. Girard, Y. Lafont, and L. Regnier, editors, *Advances in Linear Logic*, number 222 in LMS Lecture Note Series, pages 1–42. Cambridge University Press, 1995.

[Gir01] J.-Y. Girard. Locus solum. *Mathematical Structures in Computer Science*, 2001.

[GMH93] M. Gell-Mann and J.B. Hartle. Classical equations for quantum systems. *Physical Review D*, 47:3345–3382, 1993.

[Gri96] R.B. Griffiths. Consistent histories and quantum reasoning. *Physical Review A*, 54:2759–2774, 1996.

[JS91] A. Joyal and R. Street. The geometry of tensor calculus. *Advances in Mathematics*, 88:55–112, 1991.

[Lam69] J. Lambek. Deductive systems and categories II. In *Lecture Notes in Mathematics 87*. Springer-Verlag, 1969.

[LS86] J. Lambek and P. Scott. *Introduction to Higher-order Categorical Logic*. Cambridge Univ. Press, 1986.

[Mac98] S. Mac Lane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer-Verlag, 1998. Second Edition.

[Mar00] F. Markopoulou. Quantum causal histories. *Classical and Quantum Gravity*, 17:2059–2077, 2000.

[MS97] F. Markopoulou and L. Smolin. Causal evolution of spin networks. *Nuclear Physics B*, 508:409–430, 1997.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Omn94] R. Omnès. *The Interpretation of Quantum Mechanics*. Princeton Univ. Press, 1994.

[Per95] A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic, 1995.

[Per00a] A. Peres. Classical interventions in quantum systems: I. The measuring process. *Phys. Rev. A*, 61:022116, 2000.

[Per00b] A. Peres. Classical interventions in quantum systems: II. Relativistic invariance. *Phys. Rev. A*, 61:022117, 2000.

[Plo76] G. D. Plotkin. A powerdomain construction. *SIAM Journal of Computing*, 5(3):452–487, 1976.

[PM72] R. Penrose and M.A.H. MacCallum. Twistor theory: An approach to the quantization of fields and spacetime. *Phys. Repts.*, 6C:241–315, 1972.

[Pre] J. Preskill. Quantum information and computation. Lecture Notes, California Institute of Technology.

[Rap00] I. Raptis. Finitary spacetime sheaves. *Internat. J. Theoret. Phys.*, 39:1703–1720, 2000.

[Sor91] R. Sorkin. Spacetime and causal sets. In J. D'Olivo et. al., editor, *Relativity and Gravitation: Classical and Quantum*. World Scientific, 1991.

[Sza75] M. Szabo. Polycategories. *Communications in Algebra*, 3:663–689, 1975.

A The Density Matrix Formalism

This section can safely be skipped by physicists. We assume that all Hilbert spaces are finite dimensional for simplicity.

The postulates of quantum mechanics assert that the state of a system is described by a ray in some Hilbert space, $\mathcal{H}$. Following Dirac we denote a state by the "ket" $|\psi\rangle$ and the corresponding element of the dual space by the "bra" $\langle\psi|$. Typically the rays are represented by *normalized* vectors; $\langle\psi|\psi\rangle = 1$.

An *observable* is described by a self-adjoint operator A. The result of measuring the physical quantity associated with A one of the eigenvalues of A. Let λ_i be the eigenvalues of A and let $|\lambda_i\rangle$ be the associated eigenvectors. A state $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_i c_i|\lambda_i\rangle.$$

If a measurement of A is made on the system in the state ψ the result λ_i will be obtained with probability $|c_i|^2$.

In order to capture partial information about quantum systems one uses *density matrices*. Before we describe density matrices we review some linear algebra in the bra-ket notation. A projection operator is a linear operator $P : \mathcal{H} \to \mathcal{H}$ with $P^2 = P$. Given a ket $|\psi\rangle$ the notation $|\psi\rangle\langle\psi|$ denotes the projection operator onto the one dimensional subspace spanned by $|\psi\rangle$. To verify this note that

$$(|\psi\rangle\langle\psi|)(|\psi\rangle = |\psi\rangle\langle\psi|\psi\rangle = |\psi\rangle$$

and

$$(|\psi\rangle\langle\psi|)(|\phi\rangle) = |\psi\rangle\langle\psi|\phi\rangle = \langle\psi|\phi\rangle|\psi\rangle.$$

If $|\psi_i\rangle$ is an orthonormal basis for $\mathcal{H}$ the identity matrix is written $\sum_i |\psi_i\rangle\langle\psi_i|$. If Q is a linear operator with eigenvalues q_i and eigenvectors $|q_i\rangle$, which form an orthonormal basis for $\mathcal{H}$, we can represent Q as $\sum_i q_i|q_i\rangle\langle q_i|$. To see this, let $|\psi\rangle = \sum_i c_i|q_i\rangle$ then

$$Q|\psi\rangle = \sum_i c_i Q|q_i\rangle = \sum_i c_i q_i|q_i\rangle$$

now using our representation for Q we calculate

$$Q|\psi\rangle = \sum_i q_i|q_i\rangle\langle q_i|(|\psi\rangle) = \sum_{i,j} c_j q_i|q_i\rangle\langle q_i|q_j\rangle = \sum_i c_i q_i|q_i\rangle.$$

This is a toy version of the spectral theorem.

A state (i.e. a ray in $\mathcal{H}$) is called a *pure* state. The fact that $\mathcal{H}$ is a Hilbert space means that states can be added to obtain superpositions. This is part of the well-known indeterminacy of quantum mechanics. If a and b are distinct eigenvalues of some observable A with corresponding eigenvectors $|a\rangle$ and $|b\rangle$ it is perfectly possible to prepare a state of the form $\frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$. A measurement of A on such a state will yield either a or b each with probability $\frac{1}{2}$. However, it is also possible that a mixture is prepared. That is to say instead of a quantum superposition a classical stochastic mixture is prepared. In order to describe these we will use density matrices.

For a system in a pure state $|\psi\rangle$, the density matrix is just the projection operator $|\psi\rangle\langle\psi|$. If we have an observable Q with eigenvalues q_i - assumed nondegenerate for simplicity - then we can expand $|\psi\rangle$ in terms of the eigenvectors by

$$|\psi\rangle = \sum_i c_i |q_i\rangle.$$

Now the probability of observing q_i when measuring Q in the state $|\psi\rangle$ is $|\langle q_i|\psi\rangle|^2$. Recalling that the identity is given by $I = \sum_j |q_j\rangle\langle q_j|$ we get that

$$Prob(q_i, |\psi\rangle) = \sum_j \langle q_i|\psi\rangle\langle\psi|q_j\rangle\langle q_j|q_i\rangle$$

which after rearranging and using the definition of trace of an operator yields

$$Tr((|q_i\rangle\langle q_i|)(|\psi\rangle\langle\psi|)).$$

If as is typical we write ρ_ψ for the density matrix and P_i for the projection operator onto the subspace spanned by the eigenvector $|q_i\rangle$ we get

$$Prob(q_i, |\psi\rangle) = Tr(P_i\rho).$$

It is an easy calculation to show that the expectation value for Q in the state $|\psi\rangle$ is $Tr(Q\rho)$.

What if the state is not known completely? Suppose that we only know that a system is one of several possible states $|\psi_1\rangle, \ldots, |\psi_k\rangle$ with probabilities $p_1, \ldots, p_k$ respectively. We define the density matrix for such a state to be

$$\rho = \sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i|.$$

The same formulas for the probability of observing a value q_i , i.e. $Tr(P_i\rho)$ and for the expectation value of Q, i.e. $Tr(Q\rho)$ apply. One can check directly that a density matrix has the following two properties.

Proposition A.1 *An operator ρ on $\mathcal{H}$ is a **density matrix** if and only if*

- *ρ has trace 1 and*
- *ρ is a positive operator.*

Furthermore, if ρ is a density operator, $Tr(\rho^2) \leq 1$ with equality if and only if ρ is a pure state (i.e. a projection operator).

Suppose that we have a density matrix ρ describing a pure state of an $n+m$ dimensional system. Now suppose that an observer can only see the first n dimensions. The density matrix ξ describing what he can see is contained by taking the partial trace over the m dimensions that the observer cannot see. Doing this gives, in general, a nonpure state. Similarly a complementary observer who sees only the m dimensions would construct her density matrix σ by taking the appropriate partial trace. Taking these traces loses information; in fact, one cannot reconstruct ρ even from both xi and σ. Certainly the tensor product of xi and σ does not give back ρ. This is due to the loss of the cross-correlation information that was encoded in ρ but is not represented in either ξ or σ.

The axioms of quantum mechanics are easily stated in the language of density matrices. For example, if evolution from time t_1 to time t_2 is described by the unitary transformation U and ρ is the density matrix for time t_1, then the evolved density matrix ρ' for time t_2 is given by the formula $\rho' = U\rho U^\dagger$. Similarly, one can describe measurements represented by projective operators in terms of density matrices [NC00, Pre].

B Linear logic

This section can safely be skipped by logicians.

Linear logic [Gir87] is a logic introduced by Girard in 1987 to allow a finer analysis of how "resources" are consumed in the course of a deduction. As already remarked in the text, the primary objects of study in logic and especially proof theory are *sequents*, and the constructors of sequents, the *inference rules*. Several examples have already been given such as the *cut rule*:

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma', A \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\ CUT$$

So typically an inference rule is a prescription for creating a more complex sequent from one or possibly several simpler ones. Two typical inference rules are the rules of *contraction* and *weakening*. These are as follows:

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\ CONT$$

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta}\ WEAK$$

There are similar rules for the righthand side as well. These have long been standard in most logics, and indeed have a strong intuitive meaning. For example, contraction says that it is unnecessary to make the same assumption twice. However, in Girard's reexamination of the sequent calculus, he proposed an interpretation in which the formulas to the left of a sequent are resources to be consumed in the course of producing the output, i.e. the conclusions. From this perspective, the rules of contraction and weakening are quite dubious. The first step towards defining linear logic then is to recover these rules from the system. The result is a remarkably rich structure, the most notable aspect of which is that the usual connectives of logic, conjunction and disjunction, each split into two connectives. These connectives are naturally split into two classes, the *multiplicative* and the *additive* connectives. It is only the multiplicative connectives that will concern us here. Here are the rules for these connectives:

$$\frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \parr B}\ Right\text{-}\parr$$

$$\frac{\Gamma, A \vdash \Delta \quad \Gamma', B \vdash \Delta'}{\Gamma, \Gamma' A \wp B \vdash \Delta, \Delta'} \; Left{-}\wp$$

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma' \vdash \Delta', B}{\Gamma, \Gamma' \vdash \Delta, \Delta', A \otimes B} \; Right - \otimes$$

$$\frac{\Gamma, A, B \vdash \Delta}{\Gamma, A \otimes B \vdash \Delta} \; Left - \otimes$$

Categorically, the structure of linear logic has striking properties as well. As is traditional in categorical logic, one can form a category whose objects are formulas, and morphisms are proofs. This construction is described for example in [Lam69, LS86]. When one applies this construction to (multiplicative) linear logic (**MLL**), one obtains a special class of symmetric monoidal closed categories called $*$-*autonomous*. These were defined by Barr in [Bar79].

Subsequently it was demonstrated that the correspondence between proofs in **MLL** and morphisms in the free $*$-autonomous category is quite sharp. See [Blu93, BCST96]. This correspondence between morphisms and proofs is best expressed using *proof nets*, a graph-theoretic system for representing **MLL** proofs [Gir87]. Proof nets had already been seen to be a remarkable deductive system, exhibiting properties of great importance in the analysis of computation, especially concurrent computation. The precise connection between proof nets and free $*$-autonomous categories provides further evidence of their great utility.