Nuclear and Trace Ideals in Tensored $*$-Categories

Samson Abramsky*
Department of Computer Science
University of Edinburgh
Edinburgh, Scotland

Richard Blute†
Department of Mathematics and Statistics
University of Ottawa
Ottawa, Ontario, Canada

Prakash Panangaden‡
Department of Computer Science
McGill University
Montreal, Quebec, Canada

June 8, 1998

Presented to Mike Barr on the occasion of his 60th birthday.

Abstract

We generalize the notion of nuclear maps from functional analysis by defining nuclear ideals in tensored $*$-categories. The motivation for this study came from attempts to generalize the structure of the category of relations to handle what might be called "probabilistic relations". The compact closed structure associated with the category of relations does not generalize directly, instead one obtains nuclear ideals.

Most tensored $*$-categories have a large class of morphisms which behave as if they were part of a compact closed category, i.e. they allow one to transfer variables between the domain and the codomain. We introduce the notion of *nuclear ideals* to analyze these classes of morphisms. In compact closed tensored $*$-categories, all morphisms are nuclear, and in the tensored $*$-category of Hilbert spaces, the nuclear morphisms are the *Hilbert-Schmidt maps.*

We also introduce two new examples of tensored $*$-categories, in which integration plays the role of composition. In the first, morphisms are a special class of *distributions*, which we call *tame distributions*. We also introduce a category of *probabilistic relations*.

Finally, we extend the recent work of Joyal, Street and Verity on *traced monoidal categories* to this setting by introducing the notion of a *trace ideal.* We establish a close correspondence between nuclear ideals and trace ideals in a tensored $*$-category, suggested by the correspondence between Hilbert-Schmidt operators and trace operators on a Hilbert space.

---

*Research supported in part by EPSRC
†Research supported in part by NSERC.
‡Research supported in part by NSERC. Much of this work was done while on leave at BRICS, Aarhus University.

1 Introduction

This paper develops a new categorical structure, called a *nuclear ideal*, which comes from two independent, seemingly unrelated, developments. These are Grothendieck's concept of nuclearity in functional analysis, see for example [56], and the usual notion of binary relations. The original motivation for this investigation was the need to generalize ordinary binary relations to probabilistic relations with an eye towards certain applications in computer science. However a satisfactory notion of what this generalization should be comes from the concept of nuclearity in functional analysis. This paper presents the new concept and gives several nontrivial examples of nuclear ideals.

Relations form a basic and ubiquitous mathematical structure. There has been much activity in formulating what relations are "abstractly", so that one can generalize the concept to new situations. Typical examples of such formulations are the concept of cartesian bicategories [19] and allegories [27]. One of the key aspects of the category **Rel** is the fact that one has "transfer of variables" i.e. one can use the closed structure and the involution to move variables from "input" to "output". Intuitively speaking, this reflects the idea that the source and target of a binary relation are a matter of convention and a binary relation is an inherently symmetric object. In many situations that otherwise resemble relations, one finds that the closed structure does not exist and hence one loses the ability to transfer variables. A typical analogue of binary relations are the "probabilistic" binary relations, described at length later in the paper. Even in the absence of detailed definitions it ought to be clear that one cannot (indeed *should not*) rearrange the inputs and outputs of a probabilistic relation because there may be dependencies present among different inputs. What remains then in lieu of closed structure? We claim that it is precisely the nuclear ideals of the present paper.

In these settings, there appears to be a tension between having identities and having compact closed structure. If one looks only at the nuclear ideal, one has a compact closed "category" without identities. On the other hand, the ambient category lacks closed structure. Others have observed that there are "categories without identities", and given a wide range of examples and applications [4, 54]. However the interplay between the ideal and the ambient category is the point of the present work, not just the lack of identities.

Another motivation for this work comes from considering Hilbert spaces. The tensored $*$-category of Hilbert spaces and bounded linear maps (hereafter denoted **Hilb**) shares much of the same structure as **Rel**. One of the goals of this paper is to measure the extent of this correspondence. Like the category of relations, **Hilb** has a tensor product and a tensor-preserving involution, which is the identity on objects. In the case of **Hilb**, it is given by the adjoint operation. However, the category of Hilbert spaces lacks the closed structure of **Rel**. The structure of **Hilb** has been axiomatized as the notion of a *tensored $*$-category* [29, 23]. (In fact, it is a *tensored C^*-category*, but we will not consider its normed structure here.)

In this paper, we argue that a tensored $*$-category should be thought of as a category of (generalized) relations. The category of relations is compact closed, and this property is frequently taken to be fundamental in axiomatizing relational categories [1, 19]. However, the categories of relations which we consider are not compact closed, but rather contain a large class of morphisms, in fact an ideal, which has the basic structure of a compact closed category. To axiomatize this notion, we introduce the new notions of *nuclear ideal* and *nuclear morphism.* This idea is based on the definition of a nuclear morphism between Banach spaces, due to Grothendieck [33], which was subsequently axiomatized by Higgs and Rowe [37]. The concept of nuclearity in analysis can be viewed as describing when one can think of linear maps as matrices. Of course, in the finite-dimensional case one can always do this and it will be the case that all maps between finite-dimensional vector spaces are nuclear. The Higgs-Rowe theory applies only to autonomous (symmetric monoidal closed) categories, while our definition applies to the somewhat different setting of tensored $*$-categories. In the case of a compact closed $*$-category, all morphisms are nuclear, while in **Hilb** with its usual tensored $*$-structure, the nuclear morphisms are precisely the Hilbert-Schmidt maps [42]. Note that since we are only considering **Hilb** with the L_2 tensored $*$-structure, the notion of nuclear map we obtain is different from Grothendieck's notion arising from the category of Banach spaces (with, of course, the L_1 tensor product).

A further goal of this paper is to introduce two new examples of tensored $*$-categories, in which integration plays the role of composition. The first such category is a category of *generalized functions* or *distributions* [6, 56]. Since a discrete relation on $X \times Y$ can be viewed as a function $f : X \times Y \to \{0, 1\}$, it seems reasonable to model a "smeared out" relation as a continuous function $f : U \times V \to \mathsf{R}$, where U and V are open subsets of Euclidean space. However, the identity for such a category would be the *Dirac Delta* which is not a function, but a distribution. We choose a particular class of distributions, the *tame distributions*, which are sufficiently functional to allow composition. We then present a nuclear ideal for this category. It will consist of the tame distributions with functional kernel.

To build a category of *probabilistic relations*, one would like a category where the objects are probability spaces, and a morphism is a measure on the product space. The structure we eventually arrive at is the notion of *conditional probability distribution*, described in section 9. Categories of conditional probability distributions have previously been studied by Giry [32] and Wendt [57, 58]. Our formulation differs from theirs in that in our category, objects are equipped with measures and morphisms are measures on the product space satisfying an absolute continuity property. To each morphism, we are then able to associate a pair of conditional probability distributions. Again, in this case the nuclear ideal will consist of measures having a functional kernel.

We also extend the recent work of Joyal, Street and Verity on *traced monoidal categories* [40] to the present setting by introducing the notion of a *trace ideal*. For a given symmetric monoidal category, it is not generally the case that arbitrary endomorphisms can be assigned a trace. However, one can often find ideals on which a trace can be defined satisfying equations analogous to those of Joyal, Street and Verity. Our abstract definition is suggested by the usual trace construction in the category of Hilbert spaces, where there is a well-established relationship between maps in the trace class and Hilbert-Schmidt maps. In this case, we obtain the usual notion of trace of a bounded linear operator in the trace class.

Acknowledgments- The authors would like to thank Mayer Alvo, John Baez, Michael Barr, Robin Cockett, Thomas Ehrhard, Jean-Yves Girard, Martin Hyland, Vojkan Jaksic, William Lawvere, Robert Seely and Michael Wendt for helpful discussions. We also received a number of extremely helpful comments from the anonymous referees.

2 Categorical Preliminaries

We assume the reader is familiar with the notion of a *symmetric monoidal*[1] *category.* A suitable reference is [43]. We now review some of the different closed structures such a category could have.

Definition 2.1 A symmetric monoidal category is *closed* or *autonomous* if, for all objects A and B, there is an object $A \multimap B$ and an adjointness relation:

$$Hom(A \otimes B, C) \cong Hom(B, A \multimap C)$$

The unit and counit of this adjunction are the familiar morphisms:

$$ev\colon A \otimes (A \multimap B) \to B \qquad coev\colon A \to B \multimap (A \otimes B)$$

Examples of autonomous categories include the category of vector spaces and the category of relations. We obtain a pair of autonomous categories by considering Banach spaces. We can either consider $\mathbf{Ban}_\infty$, the category of Banach spaces and bounded linear maps, or we can consider the category $\mathbf{Ban}_1$, of Banach spaces and maps of norm less than or equal to 1. In either case, the internal Hom is the Banach space of all bounded linear maps, and the tensor product is the completed projective tensor product [56].

Definition 2.2 A *compact closed category* is a symmetric monoidal category such that for each object A there exists a dual object A^*, and canonical morphisms:

$$\nu\colon I \to A \otimes A^*$$
$$\psi\colon A^* \otimes A \to I$$

such that the usual adjunction equations hold:

$$\begin{tikzcd}
 & & I \otimes A \arrow[drr, "\nu \otimes id"] & & \\
A \arrow[urr, "\cong"] \arrow[dr, "id"'] & & & & A \otimes A^* \otimes A \arrow[dl, "id \otimes \psi"] \\
 & A & & A \otimes I \arrow[ll, "\cong"] &
\end{tikzcd}$$

together with the dual diagram for A^*. In the case of a strict monoidal category, these equations reduce to the usual adjunction triangles. It is easy to see that a compact closed category is indeed closed and that $A \multimap B \cong A^* \otimes B$.

[1] We observe that the notation $c\colon A \otimes B \to B \otimes A$ is used for the symmetry, and I is used for the tensor unit.

Compact categories could also be defined as $*$-autonomous categories [13, 15] with the additional canonical isomorphism $A^* \otimes B^* \cong (A \otimes B)^*$. $*$-Autonomous categories provide the basic framework for the model theory of the multiplicative fragment of linear logic [30].

We briefly describe the prototypical example, the *category of relations*.

Definition 2.3 The *category of relations*, **Rel**, has sets as objects, a morphism from X to Y will be a relation on $X \times Y$, with the usual relational composition.

In what follows, X, Y, Z will denote sets, and x, y, z will denote elements. A binary relation on $X \times Y$ will be denoted $x\mathcal{R}y$. The identity relation will be denoted $\mathcal{ID}$, and is defined as $x\mathcal{ID}x$, for all $x \in X$. Given a relation $\mathcal{R}: X \to Y$, we let $\overline{\mathcal{R}}: Y \to X$ denote the converse relation.

We verify that **Rel** is compact. The tensor product $\otimes$ is given by taking the products of sets, and on morphisms, we have:

$$\mathcal{R}: X \to Y \qquad \mathcal{S}: X' \to Y'$$

$$(x, x')\mathcal{R} \otimes \mathcal{S}(y, y') \text{ if and only if } x\mathcal{R}y \text{ and } x'\mathcal{S}y'$$

The unit for the tensor is given by any one point set. We define the functor $(\ )^*: \mathbf{Rel} \to \mathbf{Rel}$ by:

$$X^* = X \qquad \mathcal{R}^* = \overline{\mathcal{R}}$$

The relation $\nu: I \to X \otimes X^*$ is given by $*\nu(x, x)$ for all $x \in X$ and similarly for ψ.

3 The Tensored $*$-Category of Hilbert Spaces

Our notation for this section will be as follows. We will use brackets of the form $\langle -, - \rangle$ to denote the inner product, which will be linear in the first variable. The associated norm will be denoted $|| - ||$. If α is an element of the base field, then $\overline{\alpha}$ will denote its conjugate. If H is a Hilbert space, then $\overline{H}$ will denote the conjugate space. An orthonormal basis will be denoted $\{e_i\}_{i \in I}$. A suitable reference for basic Hilbert space theory is [42].

Let **Hilb** denote the category of Hilbert spaces and bounded linear maps, where "bounded" always means bounded in the norm associated to the inner product. We now discuss the structure of this category which is relevant to this paper. The first structure we need is the *adjoint function* [42].

Definition 3.1 Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, and $f: \mathcal{H} \to \mathcal{K}$ a bounded linear map. Then the *adjoint* of f, denoted f^*, is defined to be the unique bounded linear map $f^*: \mathcal{K} \to \mathcal{H}$ such that, for all $a \in \mathcal{H}, \quad b \in \mathcal{K}$, we have:

$$\langle a, f^*(b) \rangle = \langle f(a), b \rangle$$

Lemma 3.2 *The adjoint construction satisfies the following properties:*

- $(id_{\mathcal{H}})^* = id_{\mathcal{H}}$

- $(fg)^* = g^* f^*$
- $f^{**} = f$
- $(f \otimes g)^* = f^* \otimes g^*$ *(The tensor product will be discussed below.)*

These conditions tell us that the adjoint operation provides a contravariant, tensor-preserving, involutive functor on **Hilb** which is the identity on objects. Given such a functor, it is clear that the category **Hilb** is much closer in its categorical structure to the category of relations than to the category of Banach spaces.

3.1 Hilbert-Schmidt Maps

We now discuss a crucial class of bounded linear maps, called the *Hilbert-Schmidt maps.* The material in this section can be found in [42].

Definition 3.3 If $f\colon \mathcal{H} \to \mathcal{K}$ is a bounded linear map, we call f a *Hilbert-Schmidt map* if the sum

$$\sum_{i \in I} ||f(e_i)||^2$$

is finite for an orthonormal basis $\{e_i\}_{i \in I}$. The sum is independent of basis chosen.

It is straightforward to see that:

Lemma 3.4 *If $f\colon \mathcal{H} \to \mathcal{K}$ is a Hilbert-Schmidt map and $g\colon \mathcal{H}_1 \to \mathcal{H}$, $g'\colon \mathcal{K} \to \mathcal{K}_1$ are arbitrary bounded linear maps, then $g'f$ and fg are Hilbert-Schmidt.*

Thus the Hilbert-Schmidt operators on a space form a 2-sided ideal in the set of all bounded linear operators. A proof of the following theorem may be found in [42].

Theorem 3.5 *Let* $\mathbf{HSO}(\mathcal{H}, \mathcal{K})$ *denote the set of Hilbert-Schmidt maps from $\mathcal{H}$ to $\mathcal{K}$. Then* $\mathbf{HSO}(\mathcal{H}, \mathcal{K})$ *is a Hilbert space with:*

$$\langle f, g \rangle = \sum_{i \in I, j \in J} \langle f(e_i), e'_j \rangle \langle e'_j, g(e_i) \rangle$$

Here, $\{e_i\}_{i \in I}$ is an orthonormal basis for $\mathcal{H}$ and $\{e'_j\}_{j \in J}$ is an orthonormal basis for $\mathcal{K}$.

3.2 The Tensor Product

It is standard to construct the tensor product of Hilbert spaces $\mathcal{H} \otimes \mathcal{K}$ as the completion of the algebraic tensor product with respect to the inner product:

$$\langle x_1 \otimes y_1, x_2 \otimes y_2 \rangle = \langle x_1, x_2 \rangle \langle y_1, y_2 \rangle$$

One then completes with respect to the L_2 norm to obtain a Hilbert space. (Note that it is also possible to give an equivalent presentation that emphasizes the universal mapping property of the tensor. This involves the notion of a *weak Hilbert-Schmidt mapping.* This is explained in [42], page 132.)

Remark 3.6 *We wish to emphasize that, in this paper, we will only be considering the L_2 tensor product. Furthermore, the category* **Hilb** *will always be the category of Hilbert spaces with bounded linear maps, equipped with this tensored $*$-structure.*

Thus, our notion of nuclearity will not coincide with the notion obtained by viewing Hilbert spaces as Banach spaces and applying Grothendieck's definition, which of course uses the L_1 tensor.

For us, the most important property of the Hilbert tensor product is its relation to Hilbert-Schmidt maps. This is given by the following theorem [42], p.142:

Theorem 3.7 *We define a linear mapping $U : \overline{\mathcal{H}} \otimes \mathcal{K} \to \mathbf{HSO}(\mathcal{H}, \mathcal{K})$ by $U(x \otimes y)(u) = \langle x, u \rangle y$, where $x \otimes y \in \mathcal{H} \otimes \mathcal{K}$. Then U is a unitary transformation of $\overline{\mathcal{H}} \otimes \mathcal{K}$ onto $\mathbf{HSO}(\mathcal{H}, \mathcal{K})$. In particular, we note that the morphism U is a linear bijection.*

4 Tensored $*$-categories

The category **Hilb**, of Hilbert spaces and bounded linear maps, shares many of the properties of a compact closed category, except for the closed structure. **Hilb** is in fact an example of a *tensored $*$-category.* We now develop this theory.

Definition 4.1 A category $\mathcal{C}$ is a *$*$-category* if it is equipped with a functor $(-)^* : \mathcal{C}^{op} \to \mathcal{C}$, which is strictly involutive and the identity on objects. (Note that the strict involution may be replaced with a coherent involution, but we will not require that level of generality.) A $*$-category is *tensored* if it is symmetric monoidal, $(f \otimes g)^* = f^* \otimes g^*$, and there is a covariant *conjugate functor*, $\overline{(\ )} : \mathcal{C} \to \mathcal{C}$, which commutes with the $*$-functor and has natural isomorphisms:

- $\overline{\overline{A}} \cong A$ (We will generally take this to be an equality.)
- $\overline{A \otimes B} \cong \overline{A} \otimes \overline{B}$ (We will generally take this to be an equality.)
- $\overline{I} \cong I$.

satisfying the usual monoidal equations, and the following equation. Suppose that $f : I \to I$.

$$\begin{tikzcd} I \arrow[r, "f^*"] \arrow[d, "\cong"'] & I \\ \overline{I} \arrow[r, "\overline{f}"'] & \overline{I} \arrow[u, "\cong"'] \end{tikzcd}$$

In all of our examples except for those involving complex Hilbert spaces, conjugation will simply be taken to be the identity. In this case, the previous diagram implies that if $f : I \to I$, then $f^* = f$.

The notion of a tensored $*$-category is the first step towards defining a tensored *C^*-category*, or multiobject C^*-algebra [29, 23]. This theory has been developed quite extensively in the previously cited references. Among the results established is a representation theorem stating that such categories have faithful structure-preserving embeddings in **Hilb**. This should be thought of as a multiobject version of the Gelfand-Naimark-Segal theorem.

Examples of tensored $*$-categories

- **Rel**
- **Hilb**
- $\mathbf{Hilb}_{fd}$, the category of finite-dimensional Hilbert spaces.
- $\mathbf{URep}(G)$, the category of unitary representations of a compact group G.
- $\mathbf{URep}_{fd}(G)$, the category of finite-dimensional unitary representations of a compact group G.

Further examples can be found in [29, 23]. Note that examples 2 and 4 are tensored $*$-categories which are not closed. We will present other examples of tensored $*$-categories which are not closed.

Even though tensored $*$-categories are not compact closed, they share much of the same structure. One of the goals of this paper is to introduce a structure for measuring the extent to which such a category is closed.

5 Nuclearity

One of the characteristic features of compact closed categories is the ability to distribute the dual functor across the tensor product. This is represented by the equation $(A \otimes B)^\circ \cong A^\circ \otimes B^\circ$. ($A^\circ$ denotes the dual object. We temporarily adopt this notation to avoid confusion with the $*$-functor we will be discussing later. In the context of tensored $*$-categories, one should keep in mind the equation $A^\circ = \overline{A^*}$.) This allows one to arbitrarily repartition the morphism or "interface" in the terminology of interaction categories [1]. The categories we will encounter typically allow such repartitioning for some maps, but do not meet all the requirements of being a compact closed category.

We now introduce the related notion of *nuclearity* in a symmetric monoidal closed category, due to Rowe [50], and subsequently studied by Rowe and Higgs [37]. The idea is suggested by Grothendieck's work on topological tensor products and nuclear spaces [33]. Grothendieck defined a continuous linear map $f\colon A \to B$ between Banach spaces to be *nuclear* if it can be written as $f(a) = \Sigma f_i(a) b_i$ where $\Sigma f_i \otimes b_i$ is an element of the completed projective tensor product $A^\circ \otimes B$. We begin by noting that in any symmetric monoidal closed category, there is a morphism of the form:

$$\varphi\colon B \otimes A^\circ \to A \multimap B$$

Here, $A^\circ = A \multimap I$, where I is the unit for the tensor.
This is constructed as the transpose of the composite:

$$B \otimes A^\circ \otimes A \xrightarrow{id \otimes \psi} B \otimes I \xrightarrow{\cong} B$$

Definition 5.1 Let $\mathcal{C}$ be a symmetric monoidal closed category. Let φ denote the canonical morphism $\varphi: B \otimes A^\circ \to A \multimap B$. If $f: A \to B$ in $\mathcal{C}$, then let $n(f): I \to A \multimap B$ be the name of f. We say that f is *nuclear* if there exists $p(f): I \to B \otimes A^\circ$ such that the following diagram commutes:

$$\begin{tikzcd} I \arrow[rr, "p(f)"] \arrow[dr, "n(f)"'] & & B \otimes A^\circ \arrow[dl, "\varphi"] \\ & A \multimap B & \end{tikzcd}$$

We will refer to $p(f)$ as a *pseudoname* for f. (We should point out that there are some cases in which a pseudoname is not unique.) We say that an object of $\mathcal{C}$ is *nuclear* if its identity map is nuclear.

Lemma 5.2 *Suppose that $f: A \to B$ and $g: C \to D$ are nuclear, then so are:*

- $f^\circ: B^\circ \to A^\circ$
- $f'f: A \to E$ *for any morphism* $f': B \to E$
- $fh: F \to B$ *for any morphism* $h: F \to A$
- $f \otimes g: A \otimes C \to B \otimes D$

All of the above can be obtained by straightforward diagram chasing. For example, in the third item, one can choose $p(fh) = p(f); (id \otimes h^\circ)$. It is not in general the case that if f, g are nuclear, then so is $f \multimap g$. However, if $\mathcal{C}$ is $*$-autonomous with unit as dualizing object, then $f \multimap g$ will also be nuclear, [37] p. 70.

In a compact closed category, the map φ is an isomorphism, and thus every map is nuclear. Furthermore, we can see the following:

Theorem 5.3 ([37], Thm. 2.5) *For an arbitrary object A in $\mathcal{C}$, a symmetric monoidal closed category, the following are equivalent.*

- *A is nuclear*
- *The morphism $\varphi: A \otimes A^\circ \to A \multimap A$ is an isomorphism.*
- *The morphism $\varphi: B \otimes A^\circ \to A \multimap B$ is an isomorphism, for arbitrary objects B.*

Theorem 5.4 *For any symmetric monoidal closed category, the full subcategory of nuclear objects is compact-closed.*

Proof. Suppose that A is a nuclear object. Then, choosing a pseudoname for the identity gives a morphism of the form $I \to A \otimes A^\circ$. It only remains to show that the adjunction triangles commute. We will consider one of the two adjunction triangles.

$$\begin{tikzcd}
A \cong I \otimes A \arrow[r, "p(id) \otimes id"] \arrow[d, "n(id) \otimes id"'] & A \otimes A^\circ \otimes A \arrow[d, "id \otimes ev"] \arrow[dl, "\varphi \otimes id"'] \\
(A \multimap A) \otimes A \arrow[r, "ev"'] & A \otimes I \cong A
\end{tikzcd}$$

It is standard that the lower leg of the diagram is the identity. The upper leg of the diagram corresponds to the adjunction triangle. The upper triangle in the above square is the definition of pseudoname. The lower triangle is a straightforward exercise.

∎

In $\mathbf{Ban}_\infty$ or $\mathbf{Ban}_1$, we recover Grothendieck's original definition of nuclearity. The nuclear objects are the finite-dimensional Banach spaces. In the category of vector spaces, a morphism is nuclear if and only if its image is finite-dimensional. Again, a vector space is nuclear if and only if it is finite-dimensional. In [37], the authors explore the notion of nuclearity in the category of complete join semilattices **CJSL**. It is well known that this is a symmetric monoidal closed category, in fact $*$-autonomous [41]. The authors completely characterize nuclearity in this case (This result is closely related to Raney's notion of a *tight* morphism [48].):

Theorem 5.5 *(Higgs, Rowe) A morphism $f: A \to B$ in* **CJSL** *is nuclear if and only if there exists $g: B \to A$ such that for all $a \in A, f(a) = sup\{b | a \not\leq g(b)\}$. An object is nuclear if and only if it is completely distributive.*

Remark 5.6 *Following recent work of Joyal, Street and Verity [40] on traced monoidal categories, one can now observe that, in a symmetric monoidal closed category, it is possible to define a trace on the nuclear morphisms as follows, under the assumption that pseudonames are unique. If $f: A \to A$ is nuclear, then $tr(f): I \to I$ is given by:*

$$tr(f) = p(f); ev: I \to A \otimes A^\circ \to I$$

where $ev: A \otimes A^\circ \to I$ is the usual evaluation map. Then given $h: A \to B$ a nuclear map, and $g: B \to A$ arbitrary, one can verify the usual trace equation $tr(gh) = tr(hg)$. This is seen by the following diagram:

$$\begin{tikzcd}
 & A \otimes A^\circ \arrow[dr, "ev"] & \\
I \arrow[ur, "p(gh)"] \arrow[r, "p(h)"] \arrow[dr, "p(hg)"'] & B \otimes A^\circ \arrow[u, "g \otimes id"'] \arrow[d, "id \otimes g^\circ"] & I \\
 & B \otimes B^\circ \arrow[ur, "ev"'] &
\end{tikzcd}$$

The righthand diamond is the usual (di)naturality of evaluation. The two triangles on the left are the equations for $p(gh)$ and $p(hg)$.

While this theory is satisfactory when considering symmetric monoidal closed categories, there are nonclosed categories which exhibit similar structure. For example, the category of Hilbert spaces is not closed, but the class of Hilbert-Schmidt maps seem to have something like a nuclearity property. We will soon exhibit other such categories. One of the goals of this paper is to extend the above notions to a larger class of categories, specifically to $*$-categories. We now introduce a new notion, that of a *nuclear ideal.*

Definition 5.7 Let $\mathcal{C}$ be a tensored $*$-category. A *nuclear ideal* for $\mathcal{C}$ consists of the following structure:

- For all objects $A,\ B \in \mathcal{C}$, a subset $\mathcal{N}(A,B) \subseteq Hom(A,B)$. We will refer to the union of these subsets as $\mathcal{N}(\mathcal{C})$ or $\mathcal{N}$. We will refer to the elements of $\mathcal{N}$ as *nuclear maps.* The class $\mathcal{N}$ must be closed under composition with arbitrary $\mathcal{C}$-morphisms, closed under $\otimes$, closed under $(\,)^*$, and the conjugate functor.

- A bijection $\theta : \mathcal{N}(A,B) \to Hom(I, \overline{A} \otimes B)$. If $f : A \to B$ is a nuclear morphism, note that we can use the bijection θ and the $*$-functor to construct morphisms of the form:

  1. $\theta(f) : I \to \overline{A} \otimes B$
  2. $\theta(f)^* : \overline{A} \otimes B \to I$
  3. $\theta(f^*) : I \to \overline{B} \otimes A$
  4. $\theta(f^*)^* : \overline{B} \otimes A \to I$

  We shall frequently refer to these morphisms as *transposes* of f. It will always be clear from the context which transpose is being considered. The bijection θ must also satisfy the following properties:

  1. **Preservation of tensored $*$-structure** The bijection θ must preserve all of the tensored $*$-structure. In other words,

     (a) If $f : A \to B$ and $g : C \to D$ are nuclear, then $\theta(f \otimes g) = \theta(f) \otimes \theta(g)$. More precisely, the map $\theta(f \otimes g) : I \to \overline{A \otimes C} \otimes B \otimes D$ is given by the composite:

     $$I \cong I \otimes I \to \overline{A} \otimes B \otimes \overline{C} \otimes D \cong \overline{A} \otimes \overline{C} \otimes B \otimes D = \overline{A \otimes C} \otimes B \otimes D$$

     Furthermore, the transposes of a map of the form $f : I \to A$ are given by composition with the evident isomorphism.

     (b) $\theta(\overline{f}) = \theta(f^*) = \overline{\theta(f)}$. Again, more precisely, we would say:

     $$\theta(\overline{f}) = c \circ \theta(f^*) = \overline{\theta(f)} \circ \iota$$

     where c is the symmetry and ι is the isomorphism $\iota : I \to \overline{I}$.

2. **Naturality** For any $f\colon A \to C$ and $g\colon B \to D$, the following diagram commutes:

$$\begin{tikzcd} \mathcal{N}(A,B) \arrow[r, "\theta"] \arrow[d, "\mathcal{N}(f^*,g)"'] & Hom(I, \overline{A} \otimes B) \arrow[d, "Hom(I,\overline{f} \otimes g)"] \\ \mathcal{N}(C,D) \arrow[r, "\theta"] & Hom(I, \overline{C} \otimes D) \end{tikzcd}$$

Note that since the class of nuclear morphisms is closed under composition with arbitrary $\mathcal{C}$-morphisms, the function $\mathcal{N}(f^*, g)$ is well defined.

3. **Compactness** Let $f\colon A \to B$ and $g\colon B \to C$ be nuclear.
Then the following should commute.

$$\begin{tikzcd} A \arrow[r, "\cong"] \arrow[d, "gf"] & I \otimes A \arrow[r, "\theta(g) \otimes id_A"] & \overline{B} \otimes C \otimes A \arrow[d, "c"'] \\ C & C \otimes I \arrow[l, "\cong"] & C \otimes \overline{B} \otimes A \arrow[l, "id_C \otimes \theta(f^*)^*"] \end{tikzcd}$$

This completes the definition of nuclear ideal. In the case where A is a nuclear object and $f = g = id_A$, then this last equation reduces to the usual adjunction equation for a compact closed category. We will see that it is also related to the "yanking" axiom of [40].

Given a category $\mathcal{C}$ and a nuclear ideal $\mathcal{N}$, we say that an object A of $\mathcal{C}$ is $\mathcal{N}$-nuclear if we have that $\mathcal{N}(A,-) = Hom(A,-)$. Note that by the ideal property, this is equivalent to saying that the identity map for A is nuclear. Typically, this notion of nuclear object is capturing the "finite-dimensional" subcategory. It should not be thought of as describing Grothendieck's much richer theory of nuclear spaces.

Note that we are not claiming that the transposition map is in any way unique; different choices of θ could conceivably give different nuclear ideal structures. The usual uniqueness arguments, see for example [43] pp. 80-82, do not apply here in that we may not transpose the identity map. Thus it is possible that several distinct nuclear structures may exist on a given category. We are still pursuing this question. However, we know of no such examples. In the examples presented in this paper, the choice of the transpose is obvious and canonical, given the structures under consideration.

One of the consequences of the above definition is the "sliding" equation of Joyal, Street and Verity [40]:

Lemma 5.8 *Suppose $f\colon A \to B$ and $g\colon B \to A$ are nuclear. Then the following diagram commutes for any nuclear ideal:*

$$\begin{tikzcd}
 & \overline{A} \otimes B \arrow[dr, "\theta(g^*)^*"] & \\
I \arrow[ur, "\theta(f)"] \arrow[dr, "\theta(g)"'] & & I \\
 & \overline{B} \otimes A \arrow[ur, "\theta(f^*)^*"'] &
\end{tikzcd}$$

This equation is a straightforward consequence of the axioms. We will see in Section 8 that it corresponds to the familiar trace equation $tr(fg) = tr(gf)$.

Theorem 5.9 *Let $(\mathcal{C}, \mathcal{N})$ be a nuclear ideal for which all objects are nuclear, then $\mathcal{C}$ is a compact-closed category.*

Proof. If A is an object of $\mathcal{C}$, then the transpose of the identity will be a morphism of the form $I \to A \otimes \overline{A}$. The commutativity of the adjunction triangles follows from the compactness requirement of the definition.

∎

Theorem 5.10 *The set of Hilbert-Schmidt maps forms a nuclear ideal for* **Hilb**.

Proof. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, and let $\mathcal{N}(\mathcal{H}, \mathcal{K})$ be the set of all Hilbert-Schmidt maps from $\mathcal{H}$ to $\mathcal{K}$. It is evident that $Hom(I, \overline{\mathcal{H}} \otimes \mathcal{K}) \cong \overline{\mathcal{H}} \otimes \mathcal{K}$. So the morphism U, defined in 3.7, will act as a transpose operator. We saw in section 3 that this map was a linear bijection. It only remains to check the equations. These are a straightforward consequence of linearity and properties of the adjointness operator.

∎

The nuclear objects in this case are precisely the finite-dimensional Hilbert spaces. Thus we recover the familiar compact closed subcategory. The same program can be carried out for categories of representations such as **URep**(G).

5.1 Partial Injective Functions

Define a category **PInj** as follows. Its objects will be sets, and morphisms will be partial injective functions, that is to say partial functions which are monomorphic when restricted to the domain. These partial functions were used by Danos in his modeling of the geometry of interaction [21].

If $f: X \to Y$ is a morphism, let $Dom(f)$ be its domain, i.e. $Dom(f) = \{x \in X | f(x) \text{ is defined}\}$. This category has an evident $*$-structure, and if we choose the cartesian product of sets as a tensor, then we evidently have a tensored $*$-category. We now demonstrate that this category has an evident nuclear ideal. Define:

$$\mathcal{N}(X, Y) = \{f: X \to Y | Dom(f) \text{ has cardinality 0 or 1}\}$$

Then one can see that we have an obvious bijection between $Hom(I, X \otimes Y)$ and $\mathcal{N}(X, Y)$.

Theorem 5.11 *The above construction defines a nuclear ideal for* **PInj**.

5.2 Crossed M-Sets

The following is based on Freyd and Yetter's notion of a *crossed G-Set*, which they use in their work on braided compact closed categories [28]. In this paper, we will only consider a commutative monoid, which gives a symmetric monoidal category. We hope to explore the nonsymmetric and braided versions of this construction in future work, as well as the connections to topological quantum field theory [10].

Definition 5.12 Let M be a commutative monoid with identity e. Define a *crossed M-set* to be a (left) M-set X, together with a function $|\ | : X \to M$ such that $|mx| = |x|$. (This formula is more complicated in the nonabelian case. With a nonabelian group, we would require that $|gx| = g^{-1}|x|g$.)

Now define a category **XRel** as follows. Objects are crossed M-sets, and maps are relations $R : X \to Y$ such that:

- $xRy \ \Rightarrow \ mxRmy$
- $xRy \Rightarrow |x| = |y|$

Freyd and Yetter construct a category where the objects are functions satisfying precisely these requirements. They use a nonabelian group and the braiding is the symmetry adjusted appropriately by the action of G. They then use this category to develop knot invariants [28]. In subsequent work, Yetter uses crossed G-sets to construct topological quantum field theories [59]. See also [47].

If X and Y are crossed M-sets, define $X \otimes Y$ as cartesian product with componentwise action, and $|(x, y)| = |x||y|$. The unit is the one element set $I = \{*\}$. Define $|*| = e$.

Theorem 5.13 **XRel** *is a tensored $*$-category.*

Note that **XRel** is not compact. The counit of the adjunction would be required to satisfy $*R(x, x)$ for all $x \in X$, but this would hold if and only if $|x|^2 = e$. This will be our definition of nuclear object.

Now for all X, Y, define $\mathcal{N}(X, Y) \subseteq Hom(X, Y)$ by:

$$R : X \to Y \text{ is nuclear if and only if } xRy \Rightarrow |x|^2 = |y|^2 = e$$

Theorem 5.14 *This defines a nuclear ideal for XRel.*

6 Distributions as Relations

In this section, we introduce a generalized category of relations based on the idea of *distributions*. The guiding intuition is that composition should be determined by an integral of the form:

$$\varphi(x, y); \psi(y, z) = \int \varphi(x, y)\psi(y, z)dy.$$

The viewpoint here is that the notion of integration generalizes the existential quantification that appears in the definition of relational composition. We will refer to this formula as the "convolution formula." We now introduce a framework in which this makes sense. A naive approach is to view $\varphi(x, y)$ and $\psi(y, z)$ as real-valued functions. However, for such a "category" to have identities would require an equation of the form:

$$\int \varphi(x, y)\delta(y, y')dy = \varphi(x, y')$$

and similarly for left composition. The "function" playing this role is in fact the Dirac δ which is not a function but a *generalized function* or *distribution* in the sense of Schwartz [52, 56, 6]. Unfortunately multiplication of distributions is not always well-defined. Formulas like the one above are sensible only for certain limited kinds of distributions. In the rest of this section, we review basic facts about distributions and then develop a theory of what we call "tame" distributions for which the above integral formula makes sense.

Tame distributions are mentioned in the extant literature (see, for example, Dieudonné's "Treatise on Analysis", volume 7, chapter 23, sections 9 and 10 [22]), but are not given a name.

6.1 Basics of Distributions

Let Ω denote a nonempty open subset of R^n. Let $\mathcal{E}(\Omega)$ denote the set of C^∞ (smooth) functions on Ω and $\mathcal{D}(\Omega)$ denote the smooth (complex-valued) functions of compact support on Ω. We will refer to the elements of $\mathcal{D}(\Omega)$ as *test functions*. In what follows, we use Greek letters such as ϕ, ψ, η as test functions. $\mathcal{D}(\Omega)$ is given the structure of a topological vector space as follows. This structure is described for example in [6, 56].

We begin by considering a compact subset $K \subseteq \Omega$, and letting $\mathcal{D}(\Omega; K)$ be the set of continuous functionals on Ω with support contained in K. Then we define a family of seminorms on $\mathcal{D}(\Omega; K)$ by the following formula, where $\partial^{x_1^i \ldots x_n^j}$ denotes the partial derivative with respect to the listed variables:

$$|\varphi|_m = sup\{|\partial^{x_1^i \ldots x_n^j}\varphi(x)| : x \in K \text{ and } i + \ldots + j \leq m\}$$

We then give $\mathcal{D}(\Omega; K)$ the least topology such that each of these seminorms is continuous. The existence of such a topology is proved on page 12 of [6]. With this topology, $\mathcal{D}(\Omega; K)$ is a *Fréchet space*, i.e. it is locally convex, metrizable and complete.

Now observe that

$$\mathcal{D}(\Omega) = \bigcup\{\mathcal{D}(\Omega; K) |\ K \subseteq \Omega \text{ and } K \text{ is compact}\}$$

We then give $\mathcal{D}(\Omega)$ the finest locally convex topology such that the inclusions $\mathcal{D}(\Omega; K) \subseteq \mathcal{D}(\Omega)$ are continuous for every compact K. This is known as the *inductive limit* of the topologies on $\mathcal{D}(\Omega; K)$.

Theorem 6.1 *(p.25 [6]) The topology that $\mathcal{D}(\Omega; K)$ inherits as a subspace of $\mathcal{D}(\Omega)$ is the same as its original topology for every compact K. A linear functional on $\mathcal{D}(\Omega)$ is continuous if and only if the restriction to $\mathcal{D}(\Omega; K)$ is continuous for every compact K.*

With this topology, $\mathcal{D}(\Omega)$ is not metrizable. However it is an *LF space (locally Fréchet)* in the sense of [56] p.126. As such, it is locally convex, Hausdorff and complete.

Then we define a *distribution* on Ω to be a continuous, linear (complex-valued) functional on $\mathcal{D}(\Omega)$. Let $\mathcal{D}'(\Omega)$ denote the set of all distributions on Ω. Let $\mathcal{D}'(\Omega)$ be given the weak topology, p. 45 [6] or p.197 [56]. This is equivalent to the topology of pointwise convergence, and $\mathcal{D}'(\Omega)$ is locally convex, Hausdorff and complete. We will also have need of the following extension theorem [56] p.39.

Theorem 6.2 *Let E, F be two Hausdorff topological vector spaces, with A a dense subset of E and f a continuous linear mapping of A into F. If F is complete, then there is a unique continuous linear mapping $\overline{f}$ from E into F which extends f.*

We now describe some examples.

1. Let $\mathsf{L}_{loc}(\Omega)$ denote the space of locally integrable functions. Suppose that $f \in \mathsf{L}_{loc}(\Omega)$. Define a distribution T_f by:

$$T_f(\varphi) = \int_\Omega f(x)\varphi(x)dx$$

   Note that two locally integrable functions determine the same distribution if and only if they are equal almost everywhere [56]. A distribution of this form is called *regular*, and the function f is called the *kernel of the distribution.* A distribution which does not arise in this way is called *singular*. Regular distributions are fundamental examples, in fact there are a number of strong results regarding the approximation of distributions by regular distributions [56]. This justifies thinking of distributions as *generalized functions*.

2. As a special case of the previous example, we observe that every test function is itself locally integrable, and so induces a regular distribution. Thus we have a canonical inclusion

$$\iota \colon \mathcal{D}(X) \hookrightarrow \mathcal{D}'(X)$$

   given as follows:

$$\phi(x) \mapsto [\psi(x) \in \mathcal{D}(X) \mapsto \int \phi(x)\psi(x)dx]$$

   There are similar inclusions for the set of locally integrable functions or smooth functions.

3. For any point $x \in \Omega$, let $\delta_x(\varphi) = \varphi(x)$. If $0 \in \Omega$, we denote δ_0 simply as δ and refer to it as the (one-variable) *Dirac delta.* One can show that this distribution is singular, see for example [6, 56].

4. If $\Omega \subseteq \mathsf{R}$, we may also "differentiate" the previous distribution *via* the formula:

$$\delta'_x(\varphi) = -\varphi'(x)$$

This distribution is also singular. More generally, if $\Omega \subseteq \mathsf{R}^n$ and $T \in \mathcal{D}'(\Omega)$, we have the formulas:

$$\frac{\partial}{\partial x_i}(T)(\varphi) = -T(\frac{\partial}{\partial x_i}(\varphi))$$

$$\partial^{x_1^i \dots x_n^j}(T)(\varphi) = (-1)^{i+\dots+j} T(\partial^{x_1^i \dots x_n^j}(\varphi))$$

These formulas allow one to "differentiate" nondifferentiable functions, and are one of the many advantages of distributions. See, for example, [6], Chapter 2.3.

5. When considering $\Omega \times \Omega$, we have the *trace distribution*, [39] Example 5.2.2, given by:

$$\varphi \in \mathcal{D}(\Omega \times \Omega) \mapsto \int_\Omega \varphi(x, x)$$

6.2 The Schwartz kernel theorem

One is often interested in distributions on product spaces, especially in the theory of differential equations and their associated Green's functions. In this situation the analogy between distributions and "infinite-dimensional matrices" is quite striking. The theory of kernel distributions can be seen as a formalization of this analogy. In the analysis literature, the notion of "kernel distribution" is studied at length, see for example the massive treatise of Dieudonné [22] or the book by Treves [56]. When considering a space of test functions of the form $\mathcal{D}(X \times Y)$, there is a canonical subspace of fundamental importance. Consider the tensor product $\mathcal{D}(X) \otimes \mathcal{D}(Y)$. A typical element of this space is of the form

$$\sum_{i=1}^{n} \varphi_i \otimes \psi_i \text{ where } \varphi_i \in \mathcal{D}(X) \text{ and } \psi_i \in \mathcal{D}(Y)$$

There is a canonical inclusion of $\mathcal{D}(X) \otimes \mathcal{D}(Y)$ into $\mathcal{D}(X \times Y)$ given by:

$$\varphi \otimes \psi \mapsto [(x, y) \mapsto \varphi(x)\psi(y)]$$

The result we will have use for is:

Proposition 6.3 *The space $\mathcal{D}(X) \otimes \mathcal{D}(Y)$ is sequentially dense in $\mathcal{D}(X \times Y)$.*

Now we have a chance of defining functions on $\mathcal{D}(X \times Y)$ as the unique continuous extension of functions defined on $\mathcal{D}(X) \otimes \mathcal{D}(Y)$ using Theorem 6.2.

One of the fundamental results in the theory of distributions is the Schwartz kernel theorem, which gives conditions under which maps from $\mathcal{D}(X)$ to $\mathcal{D}'(Y)$ can be realized as distributions on $X \times Y$. We need the following notations to state the theorem. If f is a distribution on $X \times Y$ and $\phi \in \mathcal{D}(X)$ then $f_*(\phi)$ will be the function from $\mathcal{D}(Y)$ to the base field given by $\psi \in \mathcal{D}(Y) \mapsto f(\phi \otimes \psi)$ and $f^*(\psi)$ is given by the evident "transpose" formula. We have not yet said that $f_*(\phi)$ and $f^*(\psi)$ are distributions; that is part of the content of the kernel theorem.

The Schwartz kernel theorem states:

Theorem 6.4 *Let X and Y be two open subsets of R^n and R^m.*

1. *Let f be a distribution on $X \times Y$. For all functions $\phi \in \mathcal{D}(X)$ the linear map $f_*(\phi)$ is a distribution on Y. Furthermore, the map $\phi \mapsto f_*(\phi)$ from $\mathcal{D}(X)$ to $\mathcal{D}'(Y)$ is continuous, when $\mathcal{D}'(Y)$ is given the weak topology.*
2. *Let f_* be a continuous linear map from $\mathcal{D}(X)$ to $\mathcal{D}'(Y)$. Then there exists a unique distribution on $X \times Y$ such that for $\phi \in \mathcal{D}(X)$ and $\psi \in \mathcal{D}(Y)$ the following holds:*

$$f(\phi \otimes \psi) = f_*(\phi)(\psi)$$

Evidently, by symmetry, the same result applies for f^*. In light of the kernel theorem, we may now state the following definition.

Definition 6.5 *Suppose that f is a distribution on $X \times Y$, then we obtain the following continuous maps (supposing that $\phi \in \mathcal{D}(X), \psi \in \mathcal{D}(Y)$ are arbitrary):*

1. *$f_*: \mathcal{D}(X) \to \mathcal{D}'(Y)$ is given by $f_*(\phi)(\psi) = f(\phi \otimes \psi)$*
2. *$f^*: \mathcal{D}(Y) \to \mathcal{D}'(X)$ is given by $f^*(\psi)(\phi) = f(\phi \otimes \psi)$*

6.3 Tame Distributions

To pass from the "discrete" category of ordinary relations to a category of "continuously varying" relations, we should replace the usual notion of morphism in **Rel**, a function $X \times Y \to 2$, with an integrable function $X \times Y \to \mathsf{R}$, where X and Y are now open subsets of some Euclidean space. We have already seen, however, that functions do not suffice. One must pass to a class of generalized functions or distributions. While distributions satisfy many properties of functions, they cannot be multiplied and hence the composition formula that we had proposed does not make sense. Thus our goal is to introduce a class of distributions which are sufficiently "functional" as to allow us to compose them using the integral formula discussed above.

We will use a notion defined by Dieudonné in [22]. It will provide the first step towards defining a composable class of distributions. Note that $\mathcal{E}(X)$ is the space of all smooth complex-valued functions on X (not necessarily of compact support). Unfortunately, Dieudonné uses the term *regular* which conflicts with the terminology above. We therefore use the term *Dieudonné-regular*.

Definition 6.6 We say that a distribution $f \in \mathcal{D}'(X \times Y)$ is *Dieudonné-regular* if

1. For all functions $\phi \in \mathcal{D}(X)$, $f_*(\phi)$ is in $\mathcal{E}(Y)$, that is to say there exists $\hat{\phi} \in \mathcal{E}(Y)$ such that the distribution $f_*(\phi) \in \mathcal{D}'(Y)$ is defined by:

$$f_*(\phi)(\psi) = \int_Y \hat{\phi}(y)\psi(y)$$

2. Similarly, for all functions $\psi \in \mathcal{D}(Y)$, $f^*(\psi)$ is in $\mathcal{E}(X)$.

An equivalent statement is that the function $f_* : \mathcal{D}(X) \to \mathcal{D}'(Y)$ specified by the kernel theorem factors through the inclusion $\mathcal{E}(Y) \hookrightarrow \mathcal{D}'(Y)$, and similarly for f^*.

We would like to define our composition as follows. Given distributions $f \in \mathcal{D}(X \times Y), g \in \mathcal{D}(Y \times Z)$ which are Dieudonné-regular, we try to define a distribution $f; g \in \mathcal{D}(X \times Z)$ using the following formula (with $\phi \in \mathcal{D}(X), \gamma \in \mathcal{D}(Z)$).

$$f; g(\phi \otimes \gamma) = \int_Y \hat{\phi} \ \ \hat{\gamma}$$

Here $\hat{\phi}$ is the element of $\mathcal{E}(Y)$ associated to the distribution $f^*(\phi)$, and $\hat{\gamma}$ is the element of $\mathcal{E}(Y)$ associated to the distribution $f_*(\gamma)$.

However, the above integral may well be infinite. Thus we must add an additional assumption which assures the finiteness of this integral. One possibility is to require not only that the two kernels be smooth, but that they have compact support.[2] Thus, we have the following:

Definition 6.7 A *tame distribution* on $X \times Y$ is a distribution f on $X \times Y$ such that each of f^* and f_* factor continuously through the appropriate ι, where ι is the inclusion of the space of test functions into the space of distributions. Explicitly, there exist continuous linear maps

$$f_L \colon \mathcal{D}(X) \to \mathcal{D}(Y)$$

$$f_R \colon \mathcal{D}(Y) \to \mathcal{D}(X)$$

such that for every $\phi \in \mathcal{D}(X)$ and $\psi \in \mathcal{D}(Y)$, we have:

$$f_*(\phi)(\psi) = f^*(\psi)(\phi) = f(\phi \otimes \psi) = \int f_L(\phi)\psi dy = \int \phi f_R(\psi) dx$$

Note that we are not saying that f_L and f_R have functional kernels and certainly not that f has a functional kernel. But rather that f^* and its adjoint f_* map test functions to distributions with test functions as kernels. In some sense, tame distributions are allowed to be mildly singular, in that composing with a test function "tames" the singularity.

Dieudonné, in [22], page 77, examines the question of when the operators f^* and f_* map test functions to test functions, and he derives the following theorem.

Theorem 6.8 *Let f be a Dieudonné-regular distribution on $X \times Y$. The following are equivalent:*

1. *The operator f_* extends to a continuous linear map from the Fréchet space $\mathcal{E}(X)$ to the Fréchet space $\mathcal{E}(Y)$.*
2. *The operator f^* maps $\mathcal{D}(Y)$ to $\mathcal{D}(X)$.*
3. *The operator f^* maps $\mathcal{E}'(Y)$ to $\mathcal{E}'(X)$, where $\mathcal{E}'(Y)$ is the space of distributions of compact support (see [6] for the definition of support of a distribution).*

[2] In fact, one could use a more general class of functions, such as the square integrable functions, but we prefer the symmetry of the present definition.

6.4 Examples

- Let X be an open subset of R^n. The trace distribution on $X \times X$ is given by $Tr(\eta) = \int \eta(x,x)dx$ where $\eta(x,x') \in \mathcal{D}(X \times X)$. From this definition it follows that $Tr_*(\phi)(\psi) = Tr^*(\psi)(\phi) = Tr(\phi \otimes \psi) = \int \phi(x)\psi(x)dx$. Thus we clearly have $Tr_L(\phi) = Tr_R(\phi) = \phi$, which shows that δ is tame. This tame distribution will act as the identity in our category.

- Suppose that T is a regular distribution on $X \times Y$ with a test function $\beta(x,y)$ as its kernel, that is to say:

$$T(\alpha(x,y)) = \int_{X \times Y} \beta(x,y)\alpha(x,y)$$

Then T is tame with its associated functions being given by:

$$T_L(\phi) = \int_X \beta(x,y)\phi(x)$$

$$T_R(\psi) = \int_Y \beta(x,y)\psi(y)$$

We write $\mathcal{T}(X,Y)$ for the tame distributions on $X \times Y$.

6.5 Composing tame distributions

Given tame distributions we can define the following operation which will serve as composition. Suppose that $f \in \mathcal{T}(X,Y), g \in \mathcal{T}(Y,Z)$. We define $f;g \in \mathcal{T}(X,Z)$ as follows. Given that f is tame, we have a continuous function $f_L : \mathcal{D}(X) \to \mathcal{D}(Y)$. Applying the first part of the Schwartz kernel theorem to g, we obtain a morphism $g_* : \mathcal{D}(Y) \to \mathcal{D}'(Z)$. Composition gives a continuous map $\mathcal{D}(X) \to \mathcal{D}'(Z)$. By the second part of the kernel theorem, we obtain a distribution on $X \times Z$.

Alternatively, we could use the extension theorem, Theorem 6.2. Let $\phi \in \mathcal{D}(X), \psi \in \mathcal{D}(Z)$. We set

$$(f;g)(\phi \otimes \psi) = \int f_L(\phi) g_R(\psi) dy.$$

This, of course, only defines $f;g$ on $\mathcal{D}(X) \otimes \mathcal{D}(Z)$ rather than on $\mathcal{D}(X \times Z)$. We then use the fact that the tensor product is a dense subspace to extend composition to all of $\mathcal{D}(X \times Z)$. One observes that $f;g$ is tame as can be seen by an elementary calculation, noting $(f;g)_L = f_L; g_L$ and $(f;g)_R = g_R; f_R$ and the tameness of f and g.

6.6 The category DRel

Definition 6.9 The category **DRel** has as objects open subsets on R^n, and, as morphisms, tame distributions. Composition is as described above.

Theorem 6.10 **DRel** *is a tensored $*$-category.*

Proof. Evidently we can verify properties of the composition $f;g$ by carrying out calculations on the distribution defined on $\mathcal{D}(X)\otimes\mathcal{D}(Z)$ and appealing to continuity and the density of $\mathcal{D}(X)\otimes\mathcal{D}(Z)$ in $\mathcal{D}(X\times Z)$. We have already noted above that $f;g$ is tame. A simple calculation shows that the trace distribution is the identity for composition.

To verify associativity we calculate as follows. Let $f\in\mathcal{D}(X\times Y)$, $g\in\mathcal{D}(Y\times Z)$ and $h\in\mathcal{D}(Z\times W)$ be tame distributions. Then we have:

$$
\begin{aligned}
((f;g);h)(\phi(x)\otimes\rho(w)) &= \int (f;g)_L(\phi)h_R(\rho)dz \\
&= \int g_L(f_L(\phi))h_R(\rho)dz \\
&= \int f_L(\phi)g_R(h_R(\rho))dy \\
&= (f;(g;h))(\phi\otimes\rho)
\end{aligned}
$$

Thus we have shown that **DRel** is a category. The tensor product is given as follows. Given objects X and Y we define $X\otimes Y$ as the cartesian product space $X\times Y$. Given morphisms in **DRel** $f\colon X\to Y$ and $g\colon X'\to Y'$ we can define $f\otimes g\colon X\otimes X'\to Y\otimes Y'$ as follows. We first define $f\otimes g$ as a distribution on $\mathcal{D}(X)\otimes\mathcal{D}(X')\otimes\mathcal{D}(Y)\otimes\mathcal{D}(Y')$ by the formula $(f\otimes g)(\phi(x)\otimes\phi'(x')\otimes\psi(y)\otimes\psi'(y'))=f(\phi\otimes\psi)g(\phi'\otimes\psi')$. It is routine to verify that this is tame. We extend $f\otimes g$ to all of $\mathcal{D}(X\times X'\times Y\times Y')$ as above. The one-point space, written $I=\{*\}$, is the unit for the tensor (with measure $\mu(\{*\})=1$).

Finally the $*$-structure is the identity on objects. On morphisms, the only thing that changes is the role of f_L and f_R. The conjugate functor is taken to be the identity.

∎

Remark 6.11 *As an example, we will describe $Hom(I,X)$, where X is an arbitrary object. Clearly, $\mathcal{D}(I)$ is isomorphic to the base field. We must have two functions:*

$$f_L\colon\mathcal{D}(I)\to\mathcal{D}(X)$$

$$f_R\colon\mathcal{D}(X)\to\mathcal{D}(I)$$

such that, for all $\psi\in\mathcal{D}(X)$:

$$\int_X f_L(1)\psi = \int_I 1 f_R(\psi)$$

But evidently $\int_I 1f_R(\psi)=f_R(\psi)$. So the function f_R is uniquely determined by the function f_L. Hence we may conclude that $Hom(I,X)$ is in bijective correspondence to test functions on X.

We now display a nuclear ideal for **DRel**. We remarked that not all tame distributions can be viewed as integral operators with functions as kernels. In particular the identity morphisms do not have this property. However, we will see that tame distributions with functional kernels form a nuclear ideal.

Definition 6.12 Given objects Y and Z of **DRel** we define the set of *nuclear morphisms*, written $\mathcal{N}(Y,Z)$, as the collection of tame distributions $g: Y \to Z$ such that $\exists \beta(y,z) \in \mathcal{D}(Y \times Z)$ with the property that for every $\phi(y,z) \in \mathcal{D}(Y \times Z)$:

$$g(\phi) = \int \beta(y,z)\phi(y,z)dydz$$

Note that the test function $\beta(y,z) \in \mathcal{D}(Y \times Z)$ associated to to the tame distribution g is unique. Thus, the set $\mathcal{N}(Y,Z)$ is in bijective correspondence to $\mathcal{D}(Y \times Z)$.

Theorem 6.13 *The sets $\mathcal{N}(Y,Z)$ form a nuclear ideal for* **DRel**.

Proof. As already remarked, if $g \in \mathcal{N}(Y,Z)$ and if β is its kernel, then:

$$\forall \psi \in \mathcal{D}(Y),\ g_L(\psi) = \int \beta(y,z)\psi(y)dy$$

To verify that we have an ideal, we have to show that for any $f \in \mathcal{T}(X,Y)$ the composite $f;g$ is nuclear and symmetrically for composition on the other side of g. In order to verify this we need to find a kernel for $f;g$. We claim that this kernel is $\alpha(x,z) =_{df} f_R(\beta(y,z))$ where we interpret this formula as follows. For each fixed $z \in Z$ $\beta(y,z)$ is a smooth function of compact support in Y; f_R acts on this function to produce a function of compact support in X. The function $\alpha(x,z)$ evidently has compact support, and its smoothness is a consequence of the continuity of f_R. It suffices to prove this for functions β of the form $\beta(y,z) = a(y)b(z)$ where $a \in \mathcal{D}(Y)$ and $b \in \mathcal{D}(Z)$. This follows from Proposition 6.3 which implies that arbitrary β can be written:

$$\beta(y,z) = \lim_{n\to\infty} \sum_{i=1}^{m_n} a_{i,n}(y)b_{i,n}(z)$$

The general result then follows from the linearity and continuity of f_R.
Now observe that for a fixed z:

$$f_R(\beta(y,z)) = f_R(a(y)b(z)) = f_R(a)(x)b(z)$$

Now we calculate as follows, again letting $\beta(y,z) = a(y)b(z)$ and relying on linearity and continuity for the general result:

$$\begin{aligned}
(f;g)(\phi(x) \otimes \psi(z)) &= \int_Y f_L(\phi)(y)g_R(\psi)(y)dy \\
&= \int_Y f_L(\phi)(y)[\int_Z \beta(y,z)\psi(z)dz]dy \\
&= \int_Z [\int_Y f_L(\phi)(y)\beta(y,z)dy]\psi(z)dz \\
&= \int_Z [\int_Y f_L(\phi)(y)a(y)b(z)dy]\psi(z)dz \\
&= \int_Z [\int_Y f_L(\phi)(y)a(y)dy]b(z)\psi(z)dz
\end{aligned}$$

$$
\begin{aligned}
&= \int_Z [\int_X \phi(x) f_R(a)(x)dx] b(z)\psi(z)dz \\
&= \int_X \phi(x)[\int_Z [f_R(a)(x)b(z)]\psi(z)dz]dx \\
&= \int_X \int_Z \phi(x)\alpha(x,z)\psi(z)dxdz.
\end{aligned}
$$

It follows that $f;g$ is an integral operator with α as its kernel. The verification for composition on the other side of g is very similar.

To complete the proof, we need to show that $Hom(I, X \otimes Y) \cong \mathcal{N}(X,Y)$. This isomorphism is described in Remark 6.11. It remains to verify the equations. Naturality requires an argument similar to the previous calculation. Compactness is quite straightforward.

∎

7 The Category PRel

In this section, we define a category of *probabilistic relations*, and describe a nuclear ideal for it. We will see that we indeed get most of the important properties of the category of relations, i.e. we have a tensored $*$-category with a nuclear ideal. Thus one may think of this category as representing relations "smeared out probabilistically". Once again, as in **DRel** we have a situation where the identity maps are too singular to be in the nuclear ideal. The nuclear ideal can be thought of as functions but the ambient category has to be described in terms of measures.

7.1 Basic Definitions of Measure Theory

We assume the reader is familiar with the basic concepts of measure theory. We recall the basic definitions for completeness. A reader who remembers these definitions can skip to the start of the next section without loss of continuity.

Definition 7.1 A *σ-field* Σ on a set X is a collection of subsets of X which

1. includes the whole space X,
2. is closed under complementation, and
3. is closed under finite and countable unions.

A *measurable space* is a set together with a σ-field. A *measurable function* from a measurable space (X, Σ_X) to (Y, Σ_Y) is a function from X to Y such that for all $B \in \Sigma_Y$ we have $f^{-1}(B) \in \Sigma_X$.

Given a measurable space (X, Σ_X), we call the members of Σ_X *measurable sets*. If B is a measurable set then the characteristic function of B is denoted χ_B and is clearly measurable.

Definition 7.2 A *measure* μ on a measurable space (X, Σ_X) is a function $\mu : \Sigma_X \longrightarrow [0, \infty]$ such that

1. $\mu(\emptyset) = 0$

2. if $\{A_i | i \in I\}$ is a pairwise-disjoint family of measurable sets, with I countable, then

$$\mu(\cup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i).$$

If we have a measure taking values in $[0, 1]$ we call it a *sub-probability measure* and if the measure ("mass") of the whole space is 1 we say that it is a *probability measure.* A σ-field equipped with a measure is called a *measure space* and equipped with a probability measure it is called a *probability space.*

Sets of measure zero play an important role. The phrase *almost everywhere* is frequently used to assert that a certain property holds everywhere except on a set of measure zero. If there is confusion about which measure is intended we might say, for example, P-almost everywhere.

The set of real numbers and the closed unit interval $[0, 1]$ play a central role in the subsequent discussion. As measurable spaces, each has two σ-fields which are often used, the *Borel* σ-field and the *Lebesgue* σ-field. Any collection of subsets of a set X generates a σ-field, namely the least σ-field containing all the sets of the given collection. If we take the open sets of any topological space and generate a σ-field we get the Borel σ-field. In particular we get the Borel σ-field on the reals. This σ-field on the reals can be given a measure in such a way that the measure of an interval is its length. The resulting measure space has the property that there are subsets of sets of measure zero that are not measurable. There is a canonical "completion" procedure which yields an extended σ-field and measure, such that any previously measurable set has the same measure and all subsets of sets of measure 0 are measurable (and have measure 0). When applied to the Borel subsets of the reals with the Lebesgue measure one gets Lebesgue measurable sets (with the Lebesgue measure). *In our discussion we always mean Borel measurable whenever we talk about a measurable subset of the reals.*

In some older books [34, 51], a measurable function from the reals to the reals is defined to be a function where the inverse image of an open set has to be a Lebesgue measurable set rather than a Borel measurable set. This has the unfortunate effect that the composite of two measurable functions need not be measurable. A suitable reference for the above discussion is [44], but any good book on probability theory such as Ash [7], Billingsley [14], or Dudley [25] covers this material.

7.2 A category of stochastic kernels

Probability theory has been examined in the past from a categorical perspective. For example, Giry [32] has given the following construction, based on hints in unpublished notes of Lawvere. Wendt has examined this construction extensively [57, 58].

Let **Meas** denote the category of measurable spaces and measurable functions. We will now describe a triple T on the category **Meas**. In what follows, when we talk about measurable functions into $[0, 1]$, we always mean the Borel σ-field on $[0, 1]$, denoted $\mathcal{B}$. If (X, Σ) is an object of **Meas**, then we define $T(X, \Sigma)$ to be the set of probability measures on (X, Σ) equipped with the least σ-algebra making the evaluations

$$e_B : T(X) \to [0, 1] \text{ defined by } e_B(P) = P(B)$$

measurable, where B ranges over the measurable sets of X. T acts on maps by the formula:

$$T(f)(P)(B') = P(f^{-1}(B'))$$

where $f: X \to Y$ and $B' \in \Sigma_Y$.

The unit for the triple $\eta: id \to T$ is defined by the formula:

$$\eta_X(x)(B) = \chi_B(x)$$

where $x \in X$ and χ_B is the characteristic function of B.

The multiplication $\mu: T^2 \to T$ is defined as follows. If $P' \in T^2(X)$, then P' defines a measure on $T(X)$, and we use it to form the following integral:

$$\mu_X(P')(B) = \int_{T(X)} e_B dP'$$

With these definitions, one can then prove [32]:

Theorem 7.3 (T, η, μ) *form a triple on* **Meas**.

To understand the structure of the Kleisli category, we require the following definition.

Definition 7.4 If (X, Σ) and (X', Σ') are measurable spaces, then a *stochastic kernel* on $X \times X'$ is a function

$$\rho: X \times \Sigma' \to [0, 1]$$

that is measurable in its first argument, for each fixed measurable set and a probability measure in its second argument for each point in X.

Stochastic kernels are closely related to regular conditional probability distributions [7, 25].

If τ is a stochastic kernel on $X \times Y$ and ρ is a stochastic kernel on $Y \times Z$, then we can compose ρ and τ to obtain a stochastic kernel $\tau \circ \rho: X \times \Sigma_Z \to [0, 1]$, using the following formula:

$$\tau \circ \rho(x, C): \int_Y \rho(-, C) d\tau(x, -) \text{ for all } x \in X, C \in \Sigma_Z$$

Note that in the above formula $\rho(-, C)$ is acting as the measurable function, and $\tau(x, -)$ as the measure. The associativity of this composition follows easily from the monotone convergence theorem.

So we obtain a category **Stoch**, whose objects are measurable spaces, and whose morphisms are stochastic kernels. The identity for this category is given by the δ-formula:

$$\delta(x, A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

One can now derive [32]:

Theorem 7.5 *The Kleisli category for the triple* T *is equivalent to* **Stoch**.

Given a morphism $f: X \to TY$ in the Kleisli category, one obtains a stochastic kernel *via* the formula:

$$F: X \times \Sigma_Y \to [0, 1] \text{ is defined by } F(x, B') = f(x)(B')$$

7.3 Probabilistic Relations

While the category **Stoch** allows valuable insights into probability theory - for example, the Chapman-Kolmogorov equation is simply functoriality [32] - it lacks some of the structure one requires of a category of relations; notably the ability to take the converse. To pass to a category which is more relational in nature, we will use measures on the product space. Unfortunately one cannot compose measures in any simple way. Given measures on the product space, there is no obvious sense in which one can integrate them to compose as in the category **Stoch**. The idea is to rely on a basic theorem which says that given such product measures, *on suitable spaces*, one can construct a pair of stochastic kernels – which, together with the marginal distributions, determine the original measure on the product space – and then compose them in the manner described for **Stoch**.

We now give the details of the construction. First suppose that we have a pair of measurable spaces (X, Σ_X) and (Y, Σ_Y), a probability measure P_X on (X, Σ_X), and a stochastic kernel $h(x, B)\colon X \times \Sigma_Y \to [0,1]$. Then we have a *unique* measure P on the product such that for all $A \in \Sigma_X$:

$$P(A \times B) = \int_A h(x, B) dP_X(x).$$

Thus if we have a pair of stochastic kernels $h : X \times \Sigma_Y \longrightarrow [0,1]$ and $k : Y \times \Sigma_X \longrightarrow [0,1]$ and probability distributions P_X on (X, Σ_X) and P_Y on (Y, Σ_Y) – satisfying an evident compatibility condition – we can reconstruct a unique probability measure on the product space.

Conversely, given a measure P on the product $X \times Y$ we can construct a measure on each of the factor spaces by setting $P_X(A) := P(A \times Y)$ and $P_Y(B) := P(X \times B)$. These are called the *marginals*. Knowing one of the marginals and the appropriate stochastic kernel is equivalent to knowing the product measure. Clearly the pair of stochastic kernels does not uniquely determine the product measure; it does not even determine the marginals. We now need to show how to go from the product measure to the stochastic kernels.

The situation we have is: a pair of measure spaces (X, Σ_X, μ_X) and (Y, Σ_Y, μ_Y) and a measure, say α, on the product space equipped with the product σ-field, $\Sigma_X \otimes \Sigma_Y$. We want to construct a stochastic kernel $h : X \times \Sigma_Y \longrightarrow [0,1]$. The product space is a product in the category **Meas** and is equipped with the usual projections π_1 and π_2 to X and Y respectively. We want to construct $h : X \times \Sigma_Y \longrightarrow [0,1]$ as in the diagram

$$\begin{tikzcd}
 & X \times Y \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\
X \arrow[rr, "h"'] & & Y
\end{tikzcd}$$

such that

$$\int_A h(x, B)\mu_X = \alpha(A \times B).$$

where h is the morphism (of the category **Stoch**) that we are trying to construct and π_1, π_2, the projections, are morphisms of the category **Meas**. However, this construction requires some assumption on the spaces involved.

More precisely, we require that the spaces are *Polish spaces*[3]. Recall that a Polish space is the topological space underlying a complete separable metric space. This assumption is quite common in probability theory and allows the construction of *regular conditional probability distributions* [7, 14, 25]. We will not invoke these general concepts here.

We state a slightly more general theorem from which the construction of h in the preceding paragraph follows immediately.

Theorem 7.6 *Suppose that (U, Σ_U, P) is a probability space, V is a Polish space with the Borel σ-field, written Σ_V, and (W, Σ_W) is a measurable space. Suppose that f is a measurable function from U to V and that g is a measurable function from U to W. Then there exists a* **Stoch** *morphism, i.e. a stochastic kernel, $Q : W \longrightarrow V$ as shown in the diagram*

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "g"'] & V \\ W \arrow[ur, "Q"] & \end{tikzcd}$$

such that for all $A \in \Sigma_W, B \in \Sigma_V$:

$$\int_{g^{-1}(A)} Q(g(u), B)\, dP(u) = P(g^{-1}(A) \cap f^{-1}(B)).$$

This Q is unique in the sense that if Q' is another stochastic kernel satisfying the same equation then for P-almost all $u \in U$ $Q(u, \cdot)$ and $Q'(U, \cdot)$ are identical.

Roughly speaking, this says that Q composed with g agrees with f at least when evaluated on the measures P. In probability texts this theorem is stated in terms of existence of regular conditional probability distributions *relative to a sub σ-field.* We have essentially the same situation since the set of inverse images under g of the W-measurable sets forms a sub-σ-field of Σ_U. With this identification, theorem 7.6 is equivalent to theorem 10.2.2 of [25].

We are now ready for the corollary of chief interest.

Corollary 7.7 *Given Polish spaces X and Y with their Borel σ-fields and a probability measure α on the product space, there is a stochastic kernel $Q_1(x, B)$ (i.e. a* **Stoch** *morphism from X to Y), where $B \in \Sigma_Y$ and a stochastic kernel $Q_2(y, A)$ (i.e. a* **Stoch** *morphism from Y to X), where $A \in \Sigma_X$, such that*

$$\int_A Q_1(x, B)\, d\alpha_X = \alpha(A \times B) = \int_B Q_2(y, A)\, d\alpha_Y.$$

Proof. We use the theorem 7.6 with $X \times Y$ as U, X as W and Y as V and the projection maps as f and g. Now we immediately get Q_1. To see that the equation is satisfied we check as follows:

$$\alpha(\pi_1^{-1}(A) \cap \pi_2^{-1}(B)) = \alpha(A \times B).$$

[3] We could have more general spaces, for example analytic spaces [38].

On the other hand the left hand side of the equation asserted in theorem 7.6 is, in this case,

$$\int_{A\times Y} Q_1(\pi_1(\langle x,y\rangle),B)d\alpha.$$

This can be rewritten as

$$\int_A Q_1(x,B)d\alpha\circ\pi_1^{-1} = \int_A Q_1(x,B)d\alpha_X$$

which is the desired result. One gets the result for Q_2 similarly. ∎

Here are two simple example applications of corollary 7.7. For the first we take the product measure α to be $\mu\otimes\nu$. In this case the stochastic kernel $h: X\times\Sigma_Y\to[0,1]$ is $h(x,B)=\nu(B)$, i.e. it is independent of x. If we take the product $X\times X$ with the measure Δ defined by $\Delta(A\times B)=\mu(A\cap B)$, we get the usual Dirac delta $\delta(x,A)$.

Finally, to define morphisms in our category, we proceed as follows. Given two measures, μ and ν, on a measurable space we say ν is *absolutely continuous* with respect μ, written $\nu << \mu$, if for any measurable set A, $\mu(A)=0$ implies that $\nu(A)=0$. We now assume that the marginal α_X is absolutely continuous with respect to μ. By applying the Radon-Nikodym theorem [14], we obtain a measurable function $h(x): X\to\mathsf{R}$ such that

$$\alpha_X(A) = \int_A h(x)d\mu(x)$$

From which it follows that:

$$\int_A Q(x,B)d\alpha_X(x) = \int_A Q(x,B)h(x)d\mu(x)$$

We refer to the function $F(x,B)=Q(x,B)h(x)$ as the *stochastic kernel associated to* α.

Definition 7.8 We define a category **PRel** as follows. The objects of **PRel** are triples (X,Σ,μ), where X is a Polish space, Σ the associated σ field and μ is a probability measure on (X,Σ). A morphism $\alpha: (X,\Sigma,\mu)\to(X',\Sigma',\mu')$ is a probability measure on $\Sigma\otimes\Sigma'$ whose marginals are absolutely continuous with respect to μ and μ'.

To compose morphisms $\alpha: X\to Y$ and $\beta: Y\to Z$, we calculate their associated stochastic kernels $F(x,B)$ and $G(y,C)$ and compose as in the above Kleisli category to obtain a stochastic kernel $H(x,C)$. We then obtain a measure on $X\times Z$ *via* the formula:

$$\gamma(A\times C) = \int_A H(x,C)d\mu(x)$$

Theorem 7.9 **PRel** *is a category.*

Proof. The only thing remaining to consider is the identity. If (X,Σ,μ) is an object, its identity is given by $\Delta(A\times A')=\mu(A\cap A')$, with the associated conditional distribution given by the Dirac δ. ∎

Theorem 7.10 **PRel** *is a tensored* $*$*-category.*

Proof. The $*$-structure of **PRel** is evident, and the tensor product on objects is given by the product in the category **Meas**, that is, one takes the product of the 2 sets, the tensor of the σ-algebras, and the product measure. The necessary equations are all straightforward to verify. ∎

It is worth understanding the nature of isomorphisms in **PRel** in order to get a better sense of the role of the measures on the **PRel** objects. We consider first objects with the same underlying Polish space and hence σ-field. We will show that two such objects are isomorphic exactly when they define the same ideal of sets of measure zero.

Proposition 7.11 *Consider two* **PRel** *objects* X_1 *and* X_2 *where* $X_1 = (X, \Sigma, \mu)$ *and* $X_2 = (X, \Sigma, \nu)$*. They are isomorphic in* **PRel** *if and only if* $\mu << \nu$ *and* $\nu << \mu$*.*

Proof. Suppose first that $\mu << \nu$ and $\nu << \mu$. We define an isomorphism $H : X_1 \longrightarrow X_2$ and $K : X_2 \longrightarrow X_1$ as follows[4]. We set $H(A \times B) = \mu(A \cap B)$ and $K(A \times B) = \nu(A \cap B)$. The marginals are

$$H_1 = H_2 = \mu \text{ and } K_1 = K_2 = \nu.$$

By the absolute continuity assumptions these are **PRel** morphisms. The associated stochastic kernels are just the Dirac delta distributions and the composite of these distributions are again Dirac delta distributions. As we have observed before the Dirac delta distribution is the stochastic kernel associated with the identity morphism. Thus H and K form an isomorphism.

Conversely, suppose that we have an isomorphism $H : X_1 \longrightarrow X_2$ and $K : X_2 \longrightarrow X_1$. Suppose that $\mu(A) = 0$ for some $A \in \Sigma$. Let h' be the stochastic kernel from X_2 to X_1 associated with H, then we have

$$\int_{X_2} h'(x, A)d\nu(x) = H(A \times X) = H_1(A) = 0$$

where the last equality follows from $H_1 << \mu$ as required for H to be a **PRel** morphism. We are writing integrals over X_2 and X_1 rather than over X in order to avoid confusion; of course X_1 and X_2 are both X as sets. Since h' is always nonnegative we have that it is ν-almost everywhere 0. Let k be the stochastic kernel from X_1 to X_2 associated to K. Since H and K form an isomorphism, we have

$$\int_{X_2} h'(x', A)k(x, dx') = \delta(x, A).$$

Integrating both sides of this equation over X_1 using ν, we get

$$\int_{X_1} [\int_{X_2} h'(x', A)k(x, dx')]d\nu(x) = \int_{X_1} \delta(x, A)d\nu(x) = \nu(A)$$

It can easily be shown, using the monotone convergence theorem, that we can rewrite the left hand side as

$$\int_{X_2} h'(x', A)[\int_{X_1} k(x, dx')d\nu(x)]$$

where the integral in square brackets defines the measure used for the outer integration. This measure is absolutely continuous with respect to ν since it is defined by k. Since the integrand $h'(x', A)$ is ν-almost everywhere 0, the whole integral is 0. Thus $\nu(A) = 0$ and $\nu << \mu$. Similarly $\mu << \nu$. ∎

[4]As usual we define measures on product spaces by specifying them on the semi-ring of "rectangles" and then relying on the standard extension theorems [14] to obtain the unique extension to the whole space.

Observation 7.12 Similarly, given two Polish spaces and a Borel isomorphism between them, one can show that the two objects are isomorphic if and only if the Borel isomorphism preserves and reflects sets of measure zero.

In view of proposition 7.11 and observation 7.12 the following important theorem of classical measure theory (see, for example, theorem 13.1.1 in [25]) almost completes the analysis of isomorphisms in **PRel**.

Theorem 7.13 *If X and Y are Polish spaces, then X and Y are Borel isomorphic if and only if X and Y have the same cardinality. Moreover this cardinality must be either finite, countable or the cardinality of the continuum.*

Now we can state the main theorem about isomorphisms in **PRel**.

Theorem 7.14 *Let X and Y be two objects in* **PRel**. *Then X and Y are isomorphic if and only if there is a Borel isomorphism between them and that isomorphism preserves and reflects sets of measure 0.*

Proof. In view of theorem 7.13, it remains to show that isomorphic objects in **PRel** always have the same cardinality. First note that for finite or countable objects in **PRel** the stochastic kernels are just stochastic matrices. Thus an elementary rank argument suffices.

In the case that one of the objects has an uncountable underlying set we argue as follows. It is easy to see that in an uncountable set, with any σ-field and with any probability measure, say P, there can be at most countably many points, x, with $P(\{x\}) \neq 0$.

Now suppose that (X, Σ_X, μ), with X a countable set, and (Y, Σ_Y, ν), with Y uncountable, are **PRel** objects. Suppose, for the moment, that $\mu(\{x\})$ is nonzero for every $x \in X$. Now suppose that we have an isomorphism $H : X \longrightarrow Y$ with inverse $K : Y \longrightarrow X$. Thus we have stochastic kernels as follows:

$$h^+, k^- : X \times \Sigma_Y \longrightarrow [0,1] \text{ and } h^-, k^+ : Y \times \Sigma_X \longrightarrow [0,1].$$

Since these are isomorphisms, we obtain the equation

$$\int_X k^-(x, B) h^-(y, dx) = \delta(y, B).$$

Since X is countable, this reduces to

$$\sum_{x \in X} k^-(x, B) h^-(y, \{x\}) = \delta(y, B).$$

Let $B = \{y\}$, where $\{y\}$ is a set with ν-measure zero. Now observe that k^- must satisfy

$$\int_X k^-(x, \{y\}) d\mu = K(X \times \{y\}) = K_Y(\{y\}) = 0$$

where the last equality is a consequence of the absolute continuity requirement. But

$$\int_X k^-(x, \{y\}) d\mu = \sum_{x \in X} k^-(x, \{y\}) \mu(\{x\}).$$

By assumption, for every $x \in X$, we have that $\mu(\{x\}) \neq 0$. Thus, for every $x \in X$, it is the case that $k^-(x, \{y\}) = 0$. So we conclude:

$$\sum_{x \in X} k^-(x, \{y\}) h^-(y, \{x\}) = 0 \neq \delta(y, \{y\}).$$

This is a contradiction.

Finally, recall that the stochastic kernels are uniquely defined only *almost* everywhere. In particular, for a countable probability space, the set of all points of measure zero itself has measure zero. Thus, at points where $\mu(\{x\}) = 0$, we can define $k^-(x, B)$ to be 0, and the above argument still applies. ∎

7.4 A nuclear ideal for PRel

To determine a nuclear ideal for **PRel**, we must consider the set $Hom(I, X \otimes Y)$. By definition, this consists of measures α which are absolutely continuous with respect to the product measure $\mu \times \mu'$. By Radon-Nikodym, we can construct a measurable function $f: X \times Y \to \mathsf{R}$ such that for all $C \in \Sigma_X \otimes \Sigma_Y$:

$$\int_C f(x, y)\ d_{\mu \times \mu'}(x, y) = \alpha(C)$$

As usual two measures are equal if and only if their associated functions agree almost everywhere.

Thus, we will define $\mathcal{N}(X, Y)$ to be the set of all measures on $X \times Y$ for which there exists a measurable function f such that the previous formula holds. It is immediate that the marginals associated to such a measure are absolutely continuous with respect to μ and μ', respectively. While f itself is only unique almost everywhere, the measure with which f is associated is easily viewed - in a canonical way - both as a member of $Hom(X, Y)$ and as a member of $Hom(I, X \times Y)$. Thus every element of the set $Hom(I, X \otimes Y)$ is associated with a measure that has a functional kernel which is in turn one of the members of the set $\mathcal{N}(X, Y)$.

To see that we have a 2-sided ideal, suppose that $\alpha \in \mathcal{N}(X, Y)$. Hence we have a function $f: X \times Y \to \mathsf{R}$ satisfying the above equation. Suppose $\beta \in Hom(Y, Z)$. let $G_2: \Sigma_Y \times Z \to \mathsf{R}$ be the associated stochastic kernel. Then we define a function $h: X \times Z \to \mathsf{R}$ by the formula:

$$h(x, z) = \int_Y f(x, y) G_2(-, z)$$

As usual, we are viewing $f(x, y)$ as a measurable function of y for the fixed x, and $G_2(-, z)$ as a measure on Y for the fixed z. The construction for right composition is essentially identical. One can readily verify that the functions so constructed are indeed functional kernels for the composite measures.

Finally, we observe that in the case when both α and β are nuclear, then there exist functions $f(x, y)$ and $g(y, z)$ which act as functional kernels. The functional kernel of the composite is given by:

$$\int_Y f(x, y) g(y, z) d\mu(y)$$

We conclude:

Theorem 7.15 *The above construction determines a nuclear ideal for* **PRel**.

The verification of the requirements for a nuclear ideal are routine. The calculations involve computing transposes and can be done just the same way as proving associativity of composition in **Stoch**. We call this nuclear ideal **MRel**. One can generalize the setting to *analytic spaces* [25] which are continuous (or measurable) images of $^\infty$ in Polish spaces.

8 Trace Ideals

In [40], Joyal, Street and Verity develop an abstract theory of *trace operators* in a monoidal category. A *trace* is a function of the form:

$$tr_A: Hom(A, A) \to Hom(I, I)$$

satisfying appropriate equations. (In fact, the authors introduce a more general parametrized trace which we discuss below.) The authors demonstrate that in a symmetric (in fact, braided) compact closed category, one obtains a trace via the formula (using the notation of section 2 and using c to represent the symmetry):

$$(h: A \to A) \mapsto (\nu; h \otimes id; c; \psi: I \to I)$$

For example, in the compact closed category of finite-dimensional Hilbert spaces, one obtains the usual notion of trace of an endomorphism. This notion of trace also underlies such ideas as feedback in a computation and braid closure [40, 36].

When one passes from the category of finite-dimensional Hilbert spaces to the category of arbitrary Hilbert spaces, one finds endomorphisms which do not have a trace, for example the identity on an infinite-dimensional space. However, each endomorphism monoid contains an ideal of endomorphisms which do have a trace. This ideal is called the *trace class* and these trace maps are closely related to Hilbert-Schmidt morphisms. After reviewing this relationship, we describe a general theory of *trace ideals* for symmetric monoidal categories. We then show that if a tensored $*$-category has a nuclear ideal satisfying certain additional structure, then one can recover a trace ideal, as in the compact closed case.

8.1 Hilbert Spaces

Appropriate references for this material are [49, 55].

Definition 8.1 An operator $B \in \mathcal{L}(H)$, the space of bounded linear operators on $\mathcal{H}$, is called *positive* if $\langle Bx, x\rangle \geq 0$, for all $x \in \mathcal{H}$. In this case, we write $B \geq 0$ and $B \geq A$ if $A - B \geq 0$.

Note for example that AA^* and A^*A are always positive.

Theorem 8.2 *([49] page 196) Suppose $A \geq 0$. Then there exists a unique $B \geq 0$ such that $B^2 = A$.*

Definition 8.3 The unique operator B of the previous theorem is denoted $\sqrt{A}$. Let $A \in \mathcal{L}(H)$. Define $|A| = \sqrt{A^*A}$.

Theorem 8.4 *Let $\mathcal{H}$ be separable and $\{e_i\}$ an orthonormal basis. If A is a positive operator, we define $tr(A) = \sum\langle Ae_n, e_n\rangle$. This is independent of orthonormal basis. It has the following properties:*

- $tr(A+B) = tr(A) + tr(B)$
- $tr(\lambda A) = \lambda tr(A)$, *for all* $\lambda \geq 0$
- *If* $0 \leq A \leq B$, *then* $tr(A) \leq tr(B)$.

Definition 8.5 An operator A is called *trace class* if $tr(|A|) < \infty$. The family of all trace class operators is denoted by $\mathcal{I}(\mathcal{H})$ or just $\mathcal{I}$.

Theorem 8.6 *$\mathcal{I}$ has the following properties:*

- *$\mathcal{I}$ is a vector space.*
- *It is a 2-sided ideal in the monoid $Hom(\mathcal{H},\mathcal{H})$.*
- *If $A \in \mathcal{I}$, then $A^* \in \mathcal{I}$*

These last two conditions say that we have a $*$-*ideal.* We now extend the notion of trace to arbitrary endomorphisms in the trace ideal.

Theorem 8.7 *([49], p.211) If $A \in \mathcal{I}$ and $\{e_i\}$ is an orthonormal basis, then $\sum_{n=1}^{\infty}\langle Ae_n, e_n\rangle$ converges absolutely and is independent of the basis. (We call this map the* **trace of** *A, $tr(A)$.)*

Using the notion of trace class, it is possible to give an equivalent formulation of the notion of Hilbert-Schmidt map:

Proposition 8.8 *([49], p.211) A mapping $f: \mathcal{H} \to \mathcal{K}$ is Hilbert-Schmidt if and only if $f^*f \in \mathcal{I}(\mathcal{H})$.*

The converse of this observation is also true:

Proposition 8.9 *([49], p.211) If h is a bounded linear operator on $\mathcal{H}$, then $h \in \mathcal{I}$ if and only if there exist Hilbert-Schmidt operators f and g on $\mathcal{H}$ such that $h = fg$.*

Remark 8.10 *Let $\mathcal{H}$ be a Hilbert space, and suppose we consider H as a Banach space. Then $\mathcal{H}$ is an object in the category* $\mathbf{Ban}_\infty$*, where we consider* $\mathbf{Ban}_\infty$ *with its usual L_1 tensor product. Thus we can apply Grothendieck's original definition of nuclear morphism, and we see that we recover precisely the trace class maps.*

8.2 Trace Ideals

The previous discussion suggests the following abstract definition. We suppose for the remainder that $\mathcal{C}$ is a symmetric monoidal category.

Definition 8.11 A *trace ideal* in $\mathcal{C}$ is a choice of subsets

$$\mathcal{I}(U) \subseteq Hom(U,U) \text{ for each object } U \text{ in } \mathcal{C}$$

and a function

$$tr_U: \mathcal{I}(U) \to Hom(I,I) \text{ for each } U \text{ in } \mathcal{C}$$

such that

1. $\mathcal{I}(U)$ is a 2-sided ideal in the monoid $Hom(U,U)$.
2. (*Dinaturality* or *Sliding*) Suppose that $f: U \to V$ and $g: V \to U$ are such that $gf \in \mathcal{I}(U)$. Then $fg \in \mathcal{I}(V)$, and $tr_U(gf) = tr_V(fg)$.
3. (*Vanishing*) If $f \in \mathcal{I}(U)$, then $f \otimes id_I \in \mathcal{I}(U \otimes I)$ and $tr_{U\otimes I}(f \otimes id_I) = tr_U(f)$. Furthermore, we require that $\mathcal{I}(I) = Hom(I,I)$. If $f : I \to I$, then $tr_I(f) = f$.
4. (*Tensor Axiom*) If $f \in \mathcal{I}(U)$ and $g \in \mathcal{I}(V)$, then $f \otimes g \in \mathcal{I}(U \otimes V)$ and $tr_{U\otimes V}(f \otimes g) = tr_U(f)tr_V(g)$.
5. Furthermore, if the category has a tensored $*$-structure, then we require that trace maps are closed under tensored $*$-structure, and the trace operators respect this structure, i.e.
   - If $f \in \mathcal{I}(U)$, then so is f^*, and $tr_U(f^*) = tr_U(f)^*$.
   - If $f \in \mathcal{I}(U)$, then $\overline{f} \in \mathcal{I}(\overline{U})$ and $tr_{\overline{U}}(\overline{f}) = tr_U(f)^*$.

An alternative approach to partial traces is presented in [18], which considers traces on a *linearly distributive category*. The trace operator works on a certain subcategory, the *core*, which has the same sort of "type degeneracy" as a compact closed category.

We would like to extend the relationship between compact closed categories and traced monoidal categories to a relationship between nuclear ideals and trace ideals. Keeping in mind the correspondence between Hilbert-Schmidt maps and the trace class, we define:

Definition 8.12 Suppose that $\mathcal{C}$ is a tensored $*$-category equipped with a nuclear ideal. Suppose also that A is an object in $\mathcal{C}$. We define the *trace class* of A to be:

$$\mathcal{I}(A) = \{h: A \to A| \textit{ There exists an object } B, \textit{ and morphisms } f: A \to B, g: B \to A$$
$$\textit{with } f,g \textit{ nuclear and } h = gf\}$$

More generally, given two objects $A, B \in \mathcal{C}$, one can define:

$$\mathcal{I}(A,B) = \{h: A \to B| \textit{ There exists an object } C, \textit{ and morphisms } f: A \to C, g: C \to B$$
$$\textit{with } f,g \textit{ nuclear and } h = gf\}$$

Lemma 8.13 *$\mathcal{I}(A)$ is a 2-sided ideal in the monoid $Hom(A, A)$. $\mathcal{I}(A, B)$ is a 2-sided ideal in $\mathcal{C}$.*

While one can define the notion of trace class for arbitrary morphisms in $\mathcal{C}$ as above, note that the actual trace function only acts on $\mathcal{I}(A) = \mathcal{I}(A, A)$. In other words, the trace function acts only on the diagonal of the functor $\mathcal{I}(-,-)$. This is analogous to the notion of *dinatural transformation*, which is the appropriate notion of naturality for multivariate functors. These are families of morphisms between the two given functors, *instantiated along the diagonals*, satisfying an appropriate commutative hexagon [26, 24, 12, 15]. Hence the alternate name "dinaturality" for the sliding axiom.

If $h \in \mathcal{I}(A)$, we would like to define a morphism $tr_A(h): I \to I$ (or just $tr(h)$ if there is no confusion) by the formula (where $\hat{g}, \hat{f}$ denote the evident transposes):

$$tr(h) = \hat{g}\hat{f}: I \to \overline{A} \otimes B \to I$$

However, there is no guarantee that if h is also equal to $f'g'$ that we will obtain the same trace. Therefore we make the following definition:

Definition 8.14 A nuclear ideal is *traced* if it satisfies the following uniqueness property:

- If $f: A \to B, g: B \to A, f': A \to C, g': C \to A$ are nuclear and $gf = g'f'$, then $\hat{g}\hat{f} = \hat{g'}\hat{f'}: I \to I$.

Theorem 8.15 *The above construction assigns a trace ideal to each traced nuclear ideal.*

The proof of this theorem is simply a matter of checking the necessary diagrams. For example, lemma 5.8 gives the sliding axiom. One can also check that:

Theorem 8.16 *The canonical nuclear ideal in* **Hilb** *is traced.*

8.3 Traces in DRel

We now examine the trace construction in our category of distributions.

Theorem 8.17 *The canonical nuclear ideal in* **DRel** *is traced.*

Proof. Suppose that $f: X \to Y, g: Y \to X, f': X \to Z, g': Z \to X$ are nuclear and $gf = g'f'$. Since f is nuclear, we have a morphism $\hat{f}: I \to X \otimes Y$, which has associated to it $\hat{f}_L: \mathcal{D}(I) \to \mathcal{D}(X \times Y)$. As already remarked, $\mathcal{D}(I)$ is isomorphic to the base field, hence the map $\hat{f}_L$ simply picks out an element of $\mathcal{D}(X \times Y)$, which we denote by β_f. Similarly for f', g, g'.

To verify the uniqueness property, recall that if $\phi \in \mathcal{D}(X)$, then

$$f_L(\phi) = \int_X \beta_f(x, y)\phi(x)$$

Since $gf = g'f'$, we have that for $\phi_1, \phi_2 \in \mathcal{D}(X)$:

$$\int_Y f_L(\phi_1) g_R(\phi_2) = \int_Z f'_L(\phi_1) g'_R(\phi_2)$$

After rearranging the order of integration one can conclude:

$$\int_X \int_X \int_Y \beta_f(x,y)\beta_g(y,x')\phi_1(x)\phi_2(x') = \int_X \int_X \int_Z \beta_{f'}(x,z)\beta_{g'}(z,x')\phi_1(x)\phi_2(x')$$

The left-hand side corresponds to the distribution on $X \times X$ with kernel $\int_Y \beta_f(x,y)\beta_g(y,x')$, and the right-hand side has kernel $\int_Z \beta_{f'}(x,z)\beta_{g'}(z,x')$. We know that two integrable functions induce the same distribution if and only if they are equal almost everywhere, but since these are smooth functions on $X \times X$, we conclude:

$$\int_Y \beta_f(x,y)\beta_g(y,x') = \int_Z \beta_{f'}(x,z)\beta_{g'}(z,x')$$

Thus we have:

$$\int_X \int_Y \beta_f(x,y)\beta_g(y,x) = \int_X \int_Z \beta_{f'}(x,z)\beta_{g'}(z,x)$$

And we conclude $tr(gf) = tr(g'f')$.

■

Actually, there is a more succinct description of the trace operator in **DRel**. Since $h = gf$ is nuclear, it has a kernel, $\alpha(x,x')$. Recall from theorem 6.13 that the formula for α is given by:

$$\alpha(x,x') = f_R(\beta_g(y,x')) = \int_Y \beta_f(x,y)\beta_g(y,x')$$

Hence we may conclude that:

$$tr_A(h) = \int_X \alpha(x,x)$$

We leave the details of the following to the reader. The result is quite similar to the case of **DRel**.

Theorem 8.18 *The canonical nuclear ideal in* **PRel** *is traced.*

8.4 The parametric trace operator

In [40], the authors actually have a parametrized trace operator. This means that there is a function of the form:

$$tr_U : Hom(A \otimes U, B \otimes U) \to Hom(A,B)$$

which reduces to the usual trace when $A = B = I$. There is an evident generalization to the ideal setting:

Definition 8.19 We suppose again that $\mathcal{C}$ is a symmetric monoidal category. A *(parametric) trace ideal* in $\mathcal{C}$ is a choice of a family of subsets, for each object U of $\mathcal{C}$, of the form:

$$\mathcal{I}^U_{A,B} \subseteq Hom(A \otimes U, B \otimes U) \text{ for all } A, B \text{ in } \mathcal{C}$$

and functions

$$tr^U_{A,B} : \mathcal{I}^U_{A,B} \to Hom(A,B)$$

such that the families are ideals in the sense that:

- If $f \in \mathcal{I}^U_{A,B}$ and $h: U \to U$ is arbitrary, then $(id \otimes h) \circ f$ and $f \circ (id \otimes h)$ are in $\mathcal{I}^U_{A,B}$.
- If $f \in \mathcal{I}^U_{A,B}$ and $g: B \to C, h: D \to A$ are arbitrary, then $(g \otimes id_U) \circ f \circ (h \otimes id_U) \in \mathcal{I}^U_{D,C}$.

These are subject to the ideal-theoretic versions of the Joyal-Street-Verity axioms. In particular, (dropping sub- and superscripts if there is no chance of confusion)

- (Vanishing)
  1. $\mathcal{I}^I_{A,B} = Hom(A \otimes I, B \otimes I)$, and the trace is calculated in the evident way.
  2. Suppose $g: A \otimes U \otimes V \to B \otimes U \otimes V$. Then $g \in \mathcal{I}^{U \otimes V}_{A,B}$ if and only if $g \in \mathcal{I}^v_{A \otimes U, B \otimes U}$ and $tr^V_{A \otimes U, B \otimes U}(g) \in \mathcal{I}^U_{A,B}$. Furthermore,

$$tr^{U \otimes V}_{A,B}(g) = tr^U_{A,B}(tr^V_{A \otimes U, B \otimes U}(g))$$

- (Superposing) Suppose $f \in \mathcal{I}^U_{A,B}$ and $g: C \to D$ is arbitrary. Then $g \otimes f \in \mathcal{I}^U_{C \otimes A, D \otimes B}$, and $tr(g \otimes f) = g \otimes tr(f)$.
- (Yanking) Suppose $f: A \to U$ and $g: U \to B$. If $c_{U,B} \circ (f \otimes g) \in \mathcal{I}^U_{A,B}$, then

$$tr^U_{A,B}(c_{U,B} \circ (f \otimes g)) = gf: A \to B$$

- (Sliding) Suppose $f: A \otimes U \to B \otimes V$ and $u: V \to U$. Then $(id \otimes u) \circ f \in \mathcal{I}^U_{A,B}$ if and only if $f \circ (id \otimes u) \in \mathcal{I}^V_{A,B}$, and the two traces are equal.
- (Tightening) Suppose $f \in \mathcal{I}^U_{A,B}$ and $g: B \to C, h: D \to A$ are arbitrary. Then

$$tr((g \otimes id_U) \circ f \circ (h \otimes id_U)) = g \circ tr(f) \circ h$$

- Furthermore, if $\mathcal{C}$ is a tensored $*$-category, then the trace must preserve this structure in an evident sense.

Some discussion of our version of the Yanking axiom is in order. The Joyal-Street-Verity version of this axiom is essentially the requirement that the trace of a symmetry morphism is the identity. However, in our framework, one cannot make this requirement since the symmetry map will generally not be in the trace class. In the forthcoming thesis of Haghverdi [35], it is observed that the following requirement is equivalent to the Joyal-Street-Verity version:

Generalized Yanking Rule:

Suppose $f: A \to U$ and $g: U \to B$. Then,

$$tr^U_{A,B}(c_{U,B} \circ (f \otimes g)) = gf: A \to B$$

8.5 U-nuclear ideals

As before, we would like to construct trace ideals from nuclear ideals. An analogous construction can be carried out using the notion of a *U-nuclear ideal.* We now outline this idea, but leave most of the details to the reader. The generalization amounts to introducing the notion of a *U-nuclear morphism.* We will say that a morphism $f: A \otimes U \to B$ is U-nuclear, if it has a transpose $\hat{f}: A \to \overline{U} \otimes B$. More specifically, for each object U, we introduce a family of morphisms $\mathcal{N}_U(A \otimes U, B) \subseteq Hom(A \otimes U, B)$. These families should be closed under all of the operations and furthermore an ideal in the sense that if

$$f \in \mathcal{N}_U(A \otimes U, B)$$

and $h: V \to U$ is arbitrary, then

$$((id \otimes h); f) \in \mathcal{N}_V(A \otimes V, B)$$

Similarly for the variables A and B.

Also there should be a natural bijection of the form:

$$\mathcal{N}_V(A \otimes V, B) \cong \mathcal{N}_{\overline{V}}(B \otimes \overline{V}, A)$$

satisfying appropriate equations. For example, the *compactness* requirement becomes:

- (Compactness) Suppose $f: A \to C \otimes B$ and $g: B \otimes D \to E$. Then we have:

$$\begin{tikzcd}
A \otimes D \arrow[r, "f \otimes id_D"] \arrow[d, "id_A \otimes \hat{g}"'] & C \otimes B \otimes D \arrow[d, "id_C \otimes g"] \\
A \otimes \overline{B} \otimes E \arrow[r, "\hat{f} \otimes id_E"'] & C \otimes E
\end{tikzcd}$$

If a tensored $*$-category is equipped with such structure, we will refer to it as a *parametrized nuclear ideal.*

Given such a construction, one defines the *U-trace class* $\mathcal{I}_U(A \otimes U, B \otimes U) \subseteq Hom(A \otimes U, B \otimes U)$ by saying that:

$$h \in \mathcal{I}_U(A \otimes U, B \otimes U)$$

if and only if there exist

$$f \in \mathcal{N}_U(A \otimes U, C), g \in \mathcal{N}_U(B \otimes U, C) \qquad \text{such that } h = g^* f$$

One then constructs the U-trace of h via the formula:

$$tr^U_{A,B}(h): A \to C \otimes \overline{U} \to B$$

where the components are the evident transposes of f and h. Again, one must add conditions to ensure that the trace satisfies appropriate equations. In particular, we note that with the above axioms, we can only obtain the following weaker version of the yanking axiom:

Lemma 8.20 *Suppose that $\mathcal{C}$ is a tensored $*$-category equipped with a parametrized nuclear ideal. If $f: X \to U$ and $g: U \to Y$ are nuclear, then $c \circ (f \otimes g): X \otimes U \to Y \otimes U$ is in the U-trace class, and*

$$tr^U_{X,Y}(c \circ (f \otimes g)) = gf$$

This is a consequence of the compactness requirement of section 5.

8.6 Traces in PInj

We now discuss the traced structure of **PInj**. First it is evident that unlike in **Hilb**, we have that $\mathcal{I}(A) = \mathcal{N}(A, A)$ for all objects A. If $f: A \to A$ is a trace map, then we have the following formula:

$$tr(f) = \begin{cases} id & \text{if } |Dom(f)| = 1, \text{ and } f \text{ is the identity when restricted to its domain.} \\ \emptyset & \text{otherwise} \end{cases}$$

The parametrized trace also has a very simple description. We will say that a morphism $f: X \otimes U \to Y$ is U-nuclear if it satisfies:

$$\forall x \in X \text{ if } (x, u) \in Dom(f) \text{ and } (x, u') \in Dom(f), \text{ then } u = u'$$

Given this definition, there is an evident bijection $\mathcal{N}(X \otimes U, Y) \cong \mathcal{N}(Y \otimes U, X)$.
The class $\mathcal{I}(X \otimes U, Y \otimes U)$ is described by having the above requirement for both the domain and codomain. Then we can say that if $f \in Tr(X \otimes U, Y \otimes U)$, $(x, u) \in Dom(f)$ and $f(x, u) = (y, u')$, then:

$$tr(f)(x) \begin{cases} \text{undefined} & \text{if } u \neq u' \\ y & \text{if } u = u' \end{cases}$$

9 Conclusions

Our investigations began with an attempt to define probabilistic relations in analogy with ordinary relations. Unexpectedly, ideas from functional analysis [33] were essential. The key idea, expressed in our abstract definition of nuclear ideals, is that certain morphisms can be thought of as behaving like "matrices".

Our work naturally follows on from the development of Higgs and Rowe [37], the fundamental difference being that we have no closed structure. Crudely speaking, Higgs and Rowe generalize Banach space theory while we generalize Hilbert space theory.

A key application of our work is that we can now work with structures that are not categories but which are nuclear ideals inside some tensored $*$-category. For example, the nuclear ideal **MRel**, described in Section 7, is of interest but is not a category. (As an example of its possible applications, we note that **MRel** has partially additive structure [45, 35].) However, **MRel** is indeed a nuclear ideal in **PRel**.

An important open question is the computational significance of trace ideals. It is already well-established that a trace structure can be used to model feedback in denotational semantics [40, 36]. But what can be said when one only has these operations on an ideal? The geometry of interaction program, due to Girard [31], can be used to obtain a compact closed category from a traced monoidal category [2, 3, 40]. It seems possible that a similar construction applied to a category with a traced ideal will give a nuclear ideal.

Another area of application of the theory of compact closed categories is *topological quantum field theory* [8, 9], which evolved, in part, from Segal's work on *conformal field theory* [53]. In topological quantum field theory, one considers a compact closed category of *cobordisms* in which composition is defined by gluing along boundaries. Then a TQFT is given by a compact closed functor to the compact closed category of finite-dimensional Hilbert spaces. In Segal's formulation of conformal field theory, one works with arbitrary Hilbert spaces and a similar "category" of Riemann surfaces with boundary. This structure is essentially a compact closed category, except that it fails to be a category in that it lacks identity morphisms. Thus it seems reasonable to suspect that it is a nuclear ideal in some larger ambient tensored $*$-category. One of our goals in future work will be to find such a category. A conformal field theory would then be a *nuclear functor* to the tensored $*$-category **Hilb**.

A related issue is the extension of our work to higher-dimensional categories. The theory of *n-Hilbert spaces* [11], a higher-dimensional analogue of Hilbert space, has become quite important in TQFT [10]. Baez has developed the theory of 2-Hilbert spaces with this in mind, and extended some of the work of Doplicher and Roberts to this setting [23].

Finally, the category **DRel** suggests several further topics of investigation. One possible extension of **DRel** is to the theory of *noncommutative distributions* [5]. Roughly speaking, these are distributions which take values in a Lie group. They are useful in the representation theory of gauge groups. Finally, we hope to take advantage of the fact that distributions form a *$\mathcal{D}$-module*, that is to say they provide representations of the Weyl algebra [20]. It would be interesting to attempt to extend the work of [16, 17], where full completeness theorems are obtained by considering representations of the additive group of integers and a noncocommutative Hopf algebra.

References

[1] S. Abramsky, S. Gay, and R. Nagarajan. Interaction categories and foundations of typed concurrent programming. In M. Broy, editor, *Deductive Program Design: Proceedings of the 1994 Marktoberdorf International Summer School*, NATO ASI Series F. Springer-Verlag, 1994. Also available as theory/papers/Abramsky/marktoberdorf.ps.gz via anonymous ftp to theory.doc.ic.ac.uk.

[2] S. Abramsky, R. Jagadeesan. New foundations for the geometry of interaction. *Information and Computation*, 111(1):53–119, May 1994.

[3] S. Abramsky. Retracing some paths in process algebra. In Montanari and Sassone, editors, *Proceedings of CONCUR 96*, number 1119 in Lecture Notes In Computer Science, pages 1–17. Springer-Verlag, 1996.

[4] P. Ageron. Effective taxonomies and crossed taxonomies. *Cahiers de Top. et Geom. Diff.*, 37:82–90, 1996.

[5] S. Albeverio, R. Hoegh-Krohn, J. Marion, D. Testard, and B. Torrésani, *Noncommutative Distributions.* Dekker Pure and Applied Mathematics, 1993.

[6] M. A. Al-Gwaiz. *Theory of Distributions.* Dekker Pure and Applied Mathematics, 1992.

[7] R. B. Ash. *Real Analysis and Probability.* Academic Press, 1972.

[8] M. Atiyah. Topological quantum field theories. *Publ. Math. Inst. Hautes Etudes Sci. Paris*, 68:175–186, 1989.

[9] M. Atiyah. *The Geometry and Physics of Knots.* Cambridge University Press, 1990.

[10] J. Baez, J. Dolan. Higher-dimensional algebra and topological quantum field theory. *Journal of Mathematical Physics,* 36:6073-6105, 1995.

[11] J. Baez. Higher-dimensional algebra II:2-Hilbert spaces. preprint, 1996.

[12] E. Bainbridge, P. Freyd, A. Scedrov, P. Scott. Functorial Polymorphism *Theoretical Computer Science,* 70:35-64, 1990.

[13] M. Barr. $*$*-autonomous categories.* Number 752 in Lecture Notes in Mathematics. Springer-Verlag, 1980.

[14] P. Billingsley. *Probability and Measure.* Wiley-Interscience, 1995.

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

[16] R. Blute, P. Scott. Linear Läuchli semantics. *Annals of Pure and Applied Logic*, 77:101-142, 1996.

[17] R. Blute, P. Scott. The shuffle Hopf algebra and noncommutative full completeness. to appear in *Journal of Symbolic Logic,* 1998.

[18] R. Blute, J.R.B. Cockett, and R.A.G. Seely. Feedback for linearly distributive categories: traces and fixed points. In preparation, 1998.

[19] A. Carboni and R. F. C. Walters. Cartesian bicategories i. *Journal of Pure and Applied Algebra*, 49:11–32, 1987.

[20] S. Coutinho. *A primer of algebraic* $\mathcal{D}$*-modules.* London Mathematical Society Student Texts, Cambridge University Press, 1995.

[21] V. Danos, *Logique linéaire: Une représentation algébrique du calcul,* preprint

[22] J. Dieudonne. *Treatise on Analysis - VII.* Number 10 in Pure and Applied Mathematics. Academic Press, 1988.

[23] S. Doplicher and J. Roberts. A new duality theory for compact groups. *Inventiones Mathematicae*, 98:157–218, 1989.

[24] E. Dubuc, R. Street. Dinatural Transformations *Springer Lecture Notes in Mathematics Volume 137*, Springer-Verlag, 1970

[25] R. M. Dudley. *Real Analysis and Probability.* Wadsworth and Brookes/Cole, 1989.

[26] S. Eilenberg, G.M. Kelly. A generalization of the functorial calculus *Journal of Algebra*, 3:366-375, 1966.

[27] P. J. Freyd, A. Scedrov. *Categories, Allegories.* North-Holland, 1990.

[28] P. Freyd, D. Yetter. Braided compact closed categories with applications to low dimensional topology. *Advances in Mathematics*, 77:156-182, 1989.

[29] P. Ghez, R. Lima, and J. Roberts. w^*-categories. *Pacific Journal of Mathematics*, 120:79–109, 1985.

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

[31] J.Y. Girard. Geometry of interaction I: interpretation of system F. *Proceedings of the ASL Meeting, Padova, 1988.*

[32] M. Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Proceedings of a Conference on Categorical Aspects of Topology and Analysis*, number 915 in Lecture Notes In Mathematics, pages 68–85. Springer-Verlag, 1980.

[33] A. Grothendieck. *Products Tensoriels Topologiques et Espaces Nucleaires.* AMS Memoirs. American Mathematical Society, 1955.

[34] P. Halmos. *Measure Theory.* Graduate Texts in Mathematics 18. Springer-Verlag, 1974. Originally published in 1950.

[35] E. Haghverdi. Thesis, In preparation, 1998.

[36] M. Hasegawa. Recursion from cyclic sharing: traced monoidal categories and models of cyclic lambda calculi. *Springer Lecture Notes in Computer Science 1210*, p.196-213, 1997.

[37] D. A. Higgs, K. Rowe. Nuclearity in the category of complete semilattices. *Journal of Pure and Applied Algebra*, 57:67–78, 1989.

[38] J. Hoffman-Jörgenson. *Probability With a View Towards Applications - 2 volumes.* Chapman and Hall, 1994.

[39] L. Hörmander. *The Analysis of Linear Partial Differential Operators I.* Grundleheren der mathematischen Wissenscaften 256, Springer-Verlag 1990.

[40] A. Joyal, R. Street and D. Verity. Traced monoidal categories. *Mathematical Proceedings of the Cambridge Philosophical Society,* 119:425–446, 1996.

[41] A. Joyal, M. Tierney. *An Extension of the Galois Theory of Grothendieck.* Memoirs of the AMS. American Mathematical Society, 1984.

[42] R. Kadison, J. Ringrose. *Fundamentals of the Theory of Operator Algebras.* Academic Press, 1983.

[43] S. Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate texts in Mathematics.* Springer-Verlag, New York, 1971.

[44] P. Malliavin. *Integration and Probability*. Graduate Texts in Mathematics 157. Springer-Verlag, 1995. French edition appeared in 1993.

[45] E. Manes, M. Arbib. *Algebraic Approaches to Program Semantics.* Springer-Verlag, 1986.

[46] J. Pachl. Disintegration and Compact Measures. *Math. Scand.*, 43:157-168, 1978

[47] T. Porter. Interpretations of Yetter's notion of G-coloring: simplicial fibre bundles and non-abelian cohomology. preprint, 1995.

[48] G. Raney. Tight Galois connections and complete distributivity. *Transactions of the American Mathematical Society,* 97:418–426, (1960)

[49] M. Reed, B. Simon. *Functional Analysis, Methods of Mathematical Physics, Volume I.* Academic Press 1972

[50] K. A. Rowe. Nuclearity. *Canad. Math. Bull.*, 31(2):227–235, 1988.

[51] W. Rudin. *Real and Complex Analysis.* McGraw-Hill, 1966.

[52] L. Schwartz. *Théorie des Distributions.* Hermann, 1957.

[53] G. Segal. The definition of conformal field theory. In K. Bleuler and M. Werner, editors, *Differential Geometric Methods in Theoretical Physics*, pages 165–171. Kluwer Academic Publishers, 1988.

[54] P. Selinger. First order axioms for concurrency. In *Proceedings of CONCUR 97*, number 1243 in Lecture Notes In Computer Science, 1997.

[55] B. Simon, *Trace Ideals and Their Applications* Cambridge University Press, 1979

[56] F. Treves. *Topological Vector Spaces, Distributions and Kernels.* Pure and Applied Mathematics 25. Academic Press, 1967.

[57] M. Wendt. *On Measurably Indexed Families of Hilbert Spaces.* PhD thesis, Dalhousie University, 1993.

[58] M. Wendt. The category of disintegrations. *Cahiers de Topologie et Geometrie Differentielle Categoriques*, 35:291–308, 1994.

[59] D. Yetter. Topological quantum field theories associated to finite groups and crossed G-sets. *Jourbal of Knot Theory and its Ramifications* 2:113-123, 1993.