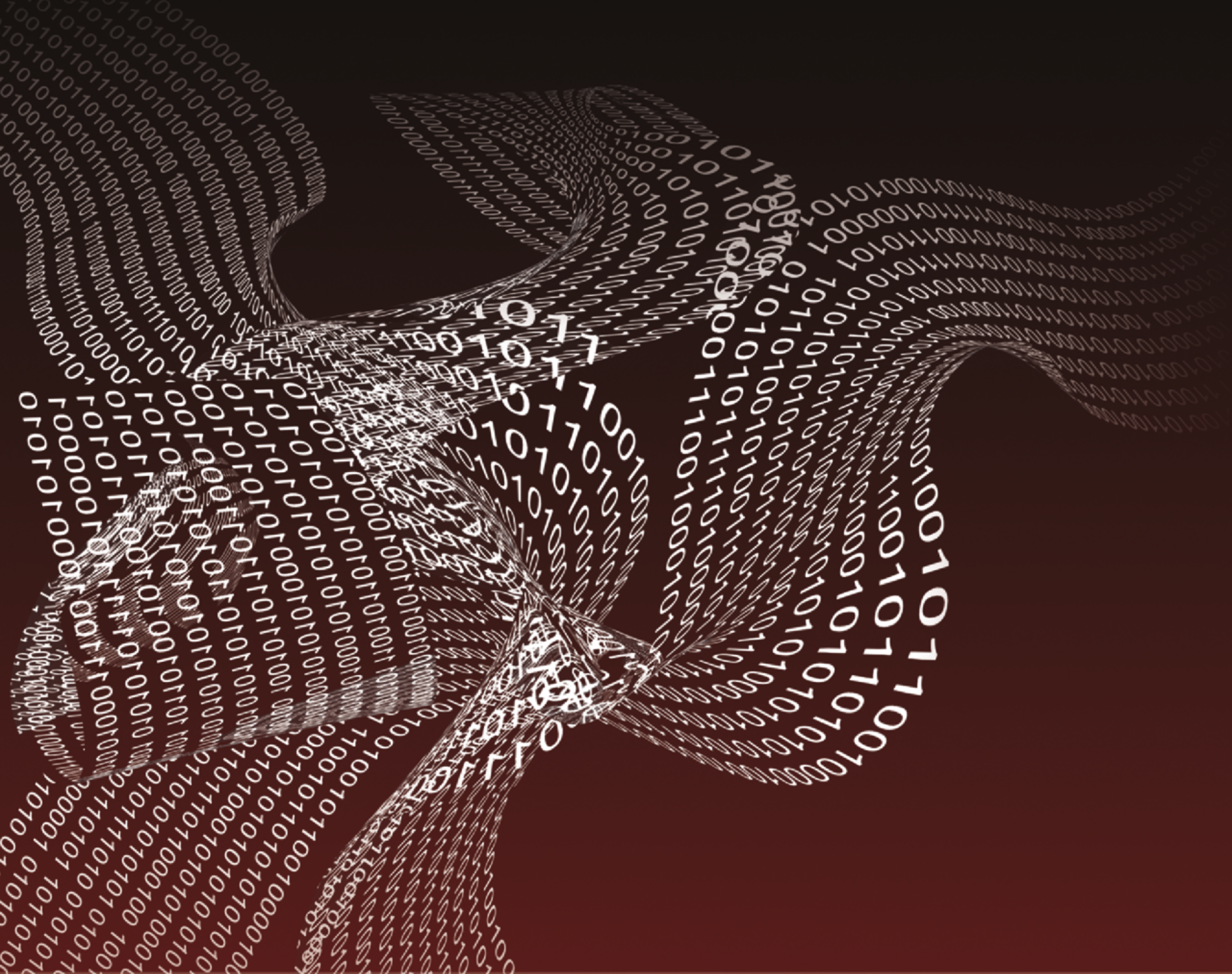
What is Category Theory?



Giandomenico Sica
editor

Polimetrica
International Scientific Publisher

What is Category Theory?

Giandomenico Sica
Editor

Polimetrica
International Scientific Publisher
Monza Italy

Editor:
Giandomenico Sica
Polimetrica
Corso Milano 26
20052 Monza (MI)
Italy
e-mail: g.sica@polimetrica.org

ISBN 88-7699-031-3

Cover design by Arch. Diego Recalcati

International Scientific Publisher
Corso Milano 26
20052 Monza - Milano - Italy
Website: http://www.polimetrica.com
Printed in Italy

Contents

7 *Preface*

Giandomenico Sica

9 Abstract and Variable Sets in Category Theory

John L. Bell

17 Categories for Knotted Curves, Surfaces and Quandles

J. Scott Carter

45 Introducing Categories to the Practicing Physicist

Bob Coecke

75 Some Implications of the Adoption of Category Theory for Philosophy

David Corfield

95 Sets, Categories and Structuralism

Costas A. Drossos

127 A Theory of Adjoint Functors – with some Thoughts about their Philosophical Significance

David Ellerman

185 Enriched Stratified Systems for the Foundations of Category Theory

Solomon Feferman

205 Category Theory, Pragmatism and Operations Universal in Mathematics

Ralf Krömer

221 What is Category Theory?

Jean-Pierre Marquis

257 Category Theory: an abstract setting for analogy and comparison

Ronald Brown – Tim Porter

275 On Doing Category Theory within Set Theoretic Foundations

Vidhyānāth K. Rao

Preface

"What is Category Theory?", third volume of the series "Advanced Studies in Mathematics and Logic", collects contributions written by some of the most representative working scientists in this field of research: John L. Bell, Ronald Brown, Scott Carter, Bob Coecke, David Corfield, Costas A. Drossos, David Ellerman, Solomon Feferman, Ralf Krömer, Jean-Pierre Marquis, Tim Porter, Vidhyānāth K. Rao.

When I thought to develop the plan of this publication, by choosing the invited authors and the related topics, I wanted to achieve the following goals:

- To stimulate the debates about the foundations of Category Theory (see the contributions written by Bell, Feferman, Marquis, Rao).
- To examine the possibility to use Category Theory as a general framework in which to unify different kinds of mathematical branches (see the contributions written by Carter, Drossos, Ellerman, Krömer).
- To analyse the idea to apply Category Theory in fields of scientific enquiry different from Mathematics (see the contributions written by Brown, Coecke, Corfield, Porter).

At the end of the work, it seems that Category Theory is living a moment of great expansion, with a growing number of possible applications in different kinds of scientific sectors; for this reason, the question "What is Category Theory?" has a particular importance in order to understand which is the specificity of this field of study.

Giandomenico Sica

Abstract and Variable Sets in Category Theory[1]

John L. Bell

In 1895 Cantor gave a definitive formulation of the concept of set (*menge*), to wit,

> *A collection to a whole of definite, well-differentiated objects of our intuition or thought.*

Let us call this notion a *concrete set*. More than a decade earlier Cantor had introduced the notion of cardinal number (*kardinalzahl*) by appeal to a process of abstraction:

> *Let* M *be a given set, thought of as a thing itself, and consisting of definite well-differentiated concrete things or abstract concepts which are called the elements of the set. If we abstract not only from the nature of the elements, but also from the order in which they are given, then there arises in us a definite general concept…which I call the power or the* cardinal number *belonging to M.*

As this quotation shows, one would be justified in calling *abstract sets* what Cantor called termed cardinal numbers[2]. An abstract set may be considered as what

---

[1] This paper has its origins in a review [2] of Lawvere and Rosebrugh's book [5].

[2] This usage of the term "abstract set" is due to F. W. Lawvere: see [4] and [5]. Lawvere's usage contrasts strikingly with that of Fraenkel, for example, who on p. 12 of [3] remarks:

arises from a concrete set when each element has been purged of all intrinsic qualities *aside from the quality which distinguishes that element from the rest.* An abstract set is then an image of pure discreteness, an embodiment of raw plurality; in short, it is an assemblage of featureless *but nevertheless distinct* "dots" or "motes"[3]. The sole intrinsic attribute of an abstract set is the number of its elements.

Concrete sets are typically obtained as *extensions of attributes*. Thus to be a member of a concrete set *C* is precisely to possess a certain attribute *A*, in short, to *be* an *A*. (It is for this reason that Peano used ∈, the first letter of Greek εστι, "is", to denote membership.) The identity of the set *C* is completely determined by the attribute *A*. As an embodiment of the relation between object and attribute, membership naturally plays a central role in concrete set theory; indeed the usual axiom systems for set theory such as Zermelo-Fraenkel and Gödel-Bernays take membership as their sole primitive relation. Concrete set theory may be seen as a theory of extensions of attributes.

By contrast, an abstract set cannot be regarded as the extension of an attribute, since the sole "attribute" possessed by the featureless dots—to which we shall still refer as *elements*—making up an abstract set is that of bare distinguishability from its fellows. Whatever abstract set theory is, it cannot be a theory of extensions of attributes. Indeed the object/attribute relation, and so *a fortiori* the membership relation between objects and sets cannot act as a primitive within the theory of abstract sets.

The key property of an abstract set being *discreteness*, we are led to derive the principles governing abstract sets from that fact. Now it is characteristic of discrete collections, and so also of abstract sets, that relations between them are reducible to relations between their constituting elements[4]. Construed in this way, relations between abstract sets provide a natural first basis on which to build a theory thereof[5]. And here categorical ideas can first be glimpsed, for relations can be *composed* in the evident way, so that abstract sets and relations between them form a *category*, the category **Rel.**

In fact **Rel** does not play a central role in the categorical approach to set theory, because relations have too much specific "structure" (they can, for example, be intersected and inverted). To obtain the definitive category associated with abstract sets, we replace arbitrary relations with *maps* between sets. Here a map from an abstract set *X* to an abstract set *Y* is a relation *f* between *X* and *Y* which correlates each element of *X* with a *unique* element of *Y*. In this situation we

---

Whenever one does not care about what the nature of the members of the set may be one speaks of an abstract set.

Fraenkel's "abstraction" is better described as "indifference".

[3] Perhaps also as "marks" or "strokes" in Hilbert's sense.

[4] This is to be contrasted with relations between *continua*. In the case of straight line segments, for example, the relation of being double the length is clearly not reducible to any relation between points or "elements". In the case of continua, and geometric objects generally, the relevant relations take the form of *mappings*.

[5] We conceive a relation *R* between two abstract sets *X* and *Y* is as correlating (some of) the elements of *X* with (some of) the elements of *Y*.

write $f: X \to Y$, and call X and Y the *domain*, and *codomain*, respectively, of f. Since the composite of two maps is clearly a map[6], abstract sets and mappings between them form a category **Set** known simply as the *category of abstract sets*.

While definitions in concrete set theory are presented in terms of membership and extensions of predicates, in the category of abstract sets definitions are necessarily formulated in terms of maps, and correlations of maps. This is the case in particular for the concept of membership itself. Thus in **Set** an element of a set X is defined to be a mapping $1 \to X$, where 1 is any set "consisting of a single dot", that is, satisfying the condition, for any set Y, that there is a unique mapping $Y \to 1$. In categorical terms, 1 is a *terminal element* of **Set.** In **Set** 1 has the important property of being a *separator* for maps in the sense that, for any maps f, g with common domain and codomain, if the composites of f and g with any element of their common codomain agree, then f and g are identical.

The "empty" set $\varnothing$ may be characterized as an *initial* object of **Set**, i.e., such that, for any set Y, there is a unique map $\varnothing \to Y$.

In **Set** the concept of set inclusion is replaced by that of monic (or one-to one) map, where a map $m: X \to Y$ is monic if, for any $f, g: A \to X$, $m f = m g \Rightarrow f = g$. A monic map to a set Y is also known as a *subobject* of Y.

Any two-element set 2 (characterized categorically as the sum of a pair of 1s; to be specific, we may choose 2 to be the set $\{\varnothing, 1\}$) plays the role of a *subobject classifier* or *truth-value object* in **Set.** This means that, for any set X, maps $X \to 2$ correspond naturally to subobjects of X. Maps $X \to 2$ correspond to *attributes* on X, with the members of 2 playing the role of *truth values*: $\varnothing$ "false" and 1 "true".

Along with $\in$, in concrete set theory the concept of identity or equality of *sets* —essentially defined in terms of $\in$—plays a seminal role. In abstract set theory, i.e. in **Set,** by contrast, it is the equality of maps which is crucial; it is, in fact, taken as a primitive notion. Equality for sets is, to all intents and purposes, replaced by the notion of *isomorphism,* that is, the existence of an invertible map between assemblages of dots. An abstract set is then defined "up to isomorphism"—the precise identity of the "dots" composing the set in question being irrelevant, the sole identifying feature is the "form" of the set.

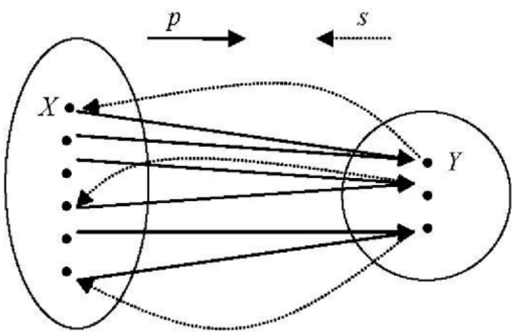
In abstract or categorical set theory sets are identified not as extensions of predicates but through the use of the omnipresent categorical concept of *adjunction*. Consider, for instance, the definition of exponentials. In concrete set theory the exponential B^A of two sets A, B is defined to be the set whose elements are all functions from A to B. In categorical set theory B^A is introduced in terms of an adjunction, that is, the postulation of an appropriately defined natural bijective correspondence, for each set X, between maps $X \to B^A$ and mappings $X \times A \to B$. (Here $X \times A$ is the Cartesian product of X and A, itself defined by means of a suitable adjunction expressing the fact that maps from an arbitrary set Y to $A \times X$ are in natural bijective correspondence with pairs of maps from Y to X and A. We note in passing that relations between X and A can be identified with subobjects of

[6] If $f: X \to Y$ and $g: Y \to Z$, we write $g f: X \to Z$ for the composite of g and f.

G.Sica (ed.) *What is Category Theory?*

$X \times A$.) Thus defined, B^A is then determined uniquely up to isomorphism, that is, as an assemblage of dots. We note that the exponential 2^X then corresponds to the *power set* of X.

The *axiom of choice* is a key principle in the theory of abstract sets. Stated in terms of maps, it takes the following form. Call a map $p: X \to Y$ *epic* if $f, g: Y \to A$, $fp = gp \Rightarrow f = g$: p is then epic if it is "onto" Y in the sense that each element of Y is the image under p of an element of X. A map $s: Y \to X$ is a *section* of p if the composite ps is the identity map on Y. Now the axiom of choice for abstract sets is the assertion that any epic map in **Set** has a section (and, indeed usually many). This principle is taken to be correct for abstract sets because of the totally arbitrary nature of the maps between them. Thus in the figure below the choice of a section s of the epic map p can be made on purely combinatorial grounds since no constraint whatsoever is placed on s (aside, of course, from the fact that it must be a section of p).



An abstract set X is said to be *infinite* if there exists an isomorphism between X and the set $X + 1$ obtained by adding one additional "dot" to X. It was the discovery of Dedekind in the 19th century that the existence of an infinite set in this sense is equivalent to that of the system of natural numbers. The *axiom of infinity*, which is also assumed to hold in abstract set theory, is the assertion that an infinite set exists.

The category **Set** is thus supposed to satisfy the following axioms:

1. There is a 'terminal' object 1 such that, for any object X, there is a unique arrow $X \to 1$
2. Any pair of objects A, B has a Cartesian product $A \times B$.
3. For any pair of objects A, B one can form the 'exponential' object B^A of all maps $A \to B$.

4. There is an object of truth values Ω such that for each object X there is a natural correspondence between subobjects (subsets) of X and arrows $X \to \Omega$. (In Set, as we have observed, one may take Ω to be the set 2 = {∅, 1}.)
5. 1 is not isomorphic to ∅.
6. The axiom of infinity.
7. The axiom of choice.
8. "Well-pointedness" axiom: 1 is a separator.

A category satisfying axioms **1. – 6.** (suitably formulated in the first-order language of categories) is called an elementary nondegenerate topos with an infinite object, or simply a *topos*[7]. The category of abstract sets is thus a topos satisfying the special additional conditions **7.** and **8.**

The objects of the category of abstract sets have been conceived as pluralities which, in addition to being discrete, are also *static* or *constant* in the sense that their elements undergo no change. There are a number of natural category-theoretic approaches to bringing *variation* into the picture. For example, we can introduce a simple form of *discrete* variation by considering as objects *bivariant sets*, that is, maps $F : X_0 \longrightarrow X_1$ between abstract sets. Here we think of X_0 as the "state" of the bivariant set F at stage 0, or "then", and X_1 as its "state" at stage 1, or "now". The bivariant set may be thought of having undergone, via the "transition" F, a change from what it was then (X_0) to what it is now (X_1). Any element x of X_0, that is, of F "then" becomes the element Fx of X_0 "now". Pursuing this metaphor, two elements "then" may become one "now" (if F is not monic), or a new element may arise "now", but because F is a map, no element "then" can split into two or more "now" or vanish altogether[8].

The appropriate maps between bivariant sets are pairs of maps between their respective states which are compatible with transitions. Thus a map from $F : X_0 \longrightarrow X_1$ to $G : Y_0 \longrightarrow Y_1$ is a pair of maps $h_0 : X_0 \longrightarrow Y_0$, $h_2 : X_1 \longrightarrow Y_1$ for which G E $h_1 = h_2$ E F. Bivariant sets and maps between them defined in this way form the category **Biv** of bivariant sets.

Now, like **Set, Biv** is a topos but the introduction of variation causes several new features to emerge. To begin with, the subobject classifier Ω in **Biv** is no longer a two -element constant set but the bivariant set i: $\Omega_0 \to \Omega_1 = 2$, where Ω_0 is the *three-element* set {∅, Φ, 1}, that is, 2 together with a new element Φ, and i sends ∅ to ∅ and both Φ and 1 to 1.

[7] See, e.g., [6] or [7].

[8] Note that had we employed relations rather than maps the latter two possibilities would have to be allowed for, complicating the situation considerably.

G.Sica (ed.) *What is Category Theory?*

And while axioms **5** and **6** continue to hold in **Biv,** axioms **7** and **8** fail[9]. In short, the axiom of choice and well-pointedness are incompatible with even the most rudimentary form of discrete variation.

Abstract sets can also be subjected to *continuous* variation. This can be done in the first instance by considering, in place of abstract sets, *bundles over topological spaces*. Here a *bundle* over a topological space X is a continuous map p from some topological space Y to X. If we think of the space Y as the union of all the "fibres" $A_x = p^{-1}(x)$ for $x \in X$, and A_x as the "value" at x of the abstract set A, then the bundle p itself may be conceived as the *abstract set A varying continuously over X*. A *map* $f: p \to p'$ between two bundles $p: Y \to X$ and $p': Y' \to X$ over X is a continuous map $f: Y \to Y'$ respecting the variation over X, that is, satisfying $p' \circ f = p$. Bundles over X and maps between them form a category **Bun**(X), the *category of bundles* over X.

While categories of bundles do represent the idea of continuous variation in a weak sense, they fail to satisfy the topos axioms **3.** and **4.** and so fall short of being suitable generalizations of the category of abstract sets to allow for such variation. To obtain these, we confine attention to special sorts of bundles known as *sheaves*. A bundle $p : Y \to X$ over X is called a *sheaf* over X when p is a *local homeomorphism* in the following sense: to each $a \in Y$ there is an open neighbourhood U of a such that pU is open in X and the restriction of p to U is a homeomorphism $U \to pU$. The domain space of a sheaf over X "locally resembles" X in the same sense as a differentiable manifold locally resembles Euclidean space. It can then be shown that the category **Shv**(X) of sheaves over X and maps between them (as bundles) is a topos the elements of whose truth-value object correspond bijectively with the open subsets of X[10]. *Categories of sheaves are appropriate generalizations of the category of abstract sets to allow for continuous variation*, and the term *continuously varying set* is taken to be synonymous with the term *sheaf*. In general, both the axiom of choice and the axiom of well-pointedness fail in sheaf categories[11], showing that both principles are incompatible with continuous variation.

If we take X to be a space consisting of a single point, a sheaf over X is a discrete space, so that the category of sheaves over X is essentially the category of abstract sets. In other words, an abstract set varying continuously over a one-point

[9] The axiom of choice fails in **Biv** since it is easy to construct an epic from the identity map on {0, 1} to the map {0, 1} → {0} with no section. That 1 is not a separator follows from the fact that ∅ → 1 has many different maps from it but no maps from 1 → 1 to it.

[10] See, e.g., [6].

[11] This is most easily seen by taking X to be the unit circle S^1 in the Euclidean plane, and considering the "double-covering" D of S^1 in 3-space. The obvious projection map $D \to S^1$ is a sheaf over S^1 with no elements in **Shv**(S^1), so the latter cannot be well-pointed. The same fact shows that the natural epic map in **Shv**(S^1) from $D \to S^1$ to the identity map $S^1 \to S^1$ (the terminal object of **Shv**(S^1) has no section, so that the axiom of choice fails in **Shv**(S^1).

space is just a (constant) abstract set. In this way arresting continuous variation leads back to constant discreteness[12].

There is an alternative description of sheaf categories which brings forth their character as categories of variable sets even more strikingly. For **Shv**(X) can also be described as the category of sets *varying* (in a suitable sense) *over open sets in* X[13]. This type of variation can be further generalized to produce categories of sets *varying over a given* (small) *category*. Each such category is again a topos[14]. Further refinements of this procedure yield so-called *smooth toposes*, categories of variable sets in which the form of variation is honed from mere continuity into smoothness[15]. Regarded as universes of discourse, in a smooth topos all maps between spaces are smooth, that is, arbitrarily many times differentiable. Even more remarkably, the objects in a smooth topos can be seen as being generated by the motions of an infinitesimal object—a "generic tangent vector"—as envisioned by the founders of differential geometry.

So, starting with the category of abstract sets, and subjecting its objects to increasingly strong forms of variation leads from discreteness to continuity to smoothness. The resulting unification of the continuous and the discrete is one of the most startling and far-reaching achievements of the categorical approach to mathematics.

Bibliography

[1] Bell, John L. [1988] *A Primer of Infinitesimal Analysis*. Cambridge University Press.

[2] ——————— Review of [5], Mathematical Reviews **MR 1965482.**

[3] Fraenkel, A. [1976] *Abstract Set Theory,* 4th Revised Edition. North-Holland.

[4] Lawvere, F. W. [1994] *Cohesive toposes and Cantor's "lauter Einsen"*. Philosophia Mathematica **2**, no. 1, 5-15.

[5] —— and Rosebrugh, R. [2003] *Sets for Mathematics*. Cambridge University Press.

[6] Mac Lane, S. and Moerdijk, I. [1994] *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. Springer.

[12] We note that had we chosen categories of bundles to represent continuous variation, then the corresponding arresting of variation would lead, not to the category of abstract sets—constant discreteness—but to the category of topological spaces—constant continuity. This is another reason for not choosing bundle categories as the correct generalization of the category of abstract sets to incorporate continuous variation.

[13] See, e.g. [6].

[14] See, e.g. [6].

[15] See [1], [7] and [8].

[7] McLarty, C. [1988] *Elementary Categories, Elementary Toposes*. Oxford University Press.

[8] Moerdijk, I. and Reyes, G. [1991] *Models for Smooth Infinitesimal Analysis*. Springer.

John L. Bell
Department of Philosophy
University of Western Ontario
LONDON, Ontario
Canada N6A 3K7

Categories for Knotted Curves, Surfaces and Quandles

J. Scott Carter

Abstract. This paper is an overview of categorical structures that are associated to embedded surfaces in 3-space, generic surfaces in 3-space, and knotted surfaces in 4-space. The paper is short on technical details, but replete with descriptions and pictures. The motivation is to give a sense in which topological phenomena can be algebratized, and how invariants can be constructed from such translations. Many steps have been completed in such a program. However, there is much that remains to be done.

1 Introduction

The discovery of the Jones polynomial lead to a broad research area that is called Quantum Topology since it combines ideas from quantum mechanics and elementary knot theory. Throughout this development, category theory has played an essential role. The diagrammatical nature of the Jones polynomial lead to the notion of a braided monoidal category [23]. Functors from the category of tangles to representations of quantum groups lead to generalizations of the Jones polynomial such as the colored Jones polynomial, and non-associative tangle theory was used [8] for finite type invariants via Kontsevich integrals. The representations at roots of unity were used to define quantum 3-manifold invariants [43] that are a mathematical definition of Witten's [47] invariants. Topological Quantum Field Theories (TQFT) as axiomatized by [2] are formed as functors from a cobordism category to a category of Hilbert spaces for example, and a variety of TQFT have been constructed (see for example, [25, 38, 46]). The literature on these developments is extensive; for example, [49] develops the theory from the point of view of the category of tangles, [38] investigates TQFTs on 3-manifolds, and [45] gives a particularly detailed and clear exposition of the Reshetikhin-Turaev [43] and Turaev-Viro [46] invariants (see also [14]). By using the Kauffman bracket definition of the Jones polynomial, Khovanov has been able to produce an homology theory in which the Jones polynomial can be interpreted as a graded Euler characteristic. His approach was motivated by categorifying the representations of $U_q(sl_2)$.

Since the discovery of the Jones polynomial, it has been questioned whether higher dimensional analogues exist. That is, "Can invariants of higher dimensional knottings be defined via diagrammatic methods?" As it was successful to investigate category theoretical structures of tangles and 3-manifolds and their functors in constructing generalizations of the invariants for knots and 3-manifolds, one of the approaches to higher dimensions was to investigate category theoretical structures for knots and manifolds in higher dimensions, in particular for knotted surfaces in 4-space [18, 5, 27, 37] and 4-manifolds [16, 19, 21, 20]. In these approaches, categorifications of algebraic systems and 2-categories have been investigated that effectively represent surfaces and 4-manifolds. Categorifications and 2-categories have been investigated for a long time from purely algebraic points of view and motivations as higher dimensional algebras (for example, [10, 33] and [6] through [3]).

These notions motivated my collaborators and me to define invariants via quandle cohomology theory [15]. The notion was defined as a modification of rack homology theory [30], and the quandle cocycle invariant was constructed [15] in a state-sum form. Some aspects of quandles and their knot invariants have been studied from algebraic [1, 26, 28] and category theoretical [3, 22] points of view. The purpose of this paper is to revisit category theoretical aspects of surfaces, quandles and their cohomology theories and invariants, from a point of view with these recent developments in mind. In Section 2, I revisit categories related to surfaces, and in Section 3, I review the definition of the quandle cocycle invariants.

2 Categories for Surfaces

2.1 Surfaces embedded in 3-space and non-associativity

The idea of a 2-category can be exemplified by embedded surfaces in 3-space, generic surfaces in 3-space, and knotted surfaces in 4-space. In this section, I develop these three 2-categories from the ground up. It is my hope that such examples will indicate clearly the ideas of objects, morphisms, 2-morphisms, relations among these, and the so-called 4-square relation. From my provincial point of view (certainly incorrect), the notion of a 2-category is invented to study surfaces. But more generally, the result of Fisher [27], and more rigorously Baez-Langford [5], indicates that the 2-category of knotted surfaces is a free 2-category of a very specific type — a free braided monoidal 2-category with duals on an

G. Sica (ed.) *What is Category Theory?*

unframed self-dual object. Thus it provides evidence for the tangle hypothesis, and invariants of knotted surfaces can be found as representations into another braided monoidal 2-category with the same structures.

Associators. We construct a 2-category $\mathcal{EMB}_0$ based upon embedded arcs and surfaces bounded by these. Consider the set of k-fold subsets, $S \subset 2^{\mathbb{N}}$ of the natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$. A typical object in $\mathcal{EMB}_0$ is a subset $\{i_1 < i_2 < \cdots < i_k\} \subset \mathbb{N}$. If $k = 0$, such a subset is the empty set. A groupoid is constructed as a collection of morphisms between such objects. Here *groupoid* means a set upon which an associative binary operation is partially defined. That is the composition ab may or may not be defined. Alternatively, a groupoid is a category in which the collection of objects forms a set.

For $j = 1, \ldots, k$, if $i_{j+1} \neq i_j + 1$ there is a morphism

$$|_{j-1} \nearrow |_{k-j-1} : \{i_1 < i_2 < \cdots < i_j < \cdots < i_k\}$$

$$\to \{i_1 < i_2 < \cdots < i_{j-1} < i_j + 1 < \cdots < i_k\}.$$

Under analogous circumstances there is a morphism back:

$$|_{j-1} \nwarrow |_{k-j-1} : \{i_1 < i_2 < \cdots < i_{j-1} < i_j + 1 < \cdots < i_k\}$$

$$\to \{i_1 < i_2 < \cdots < i_j < \cdots < i_k\}.$$

If $i_{j+1} = i_j + 1$, then there is a morphism,

$$|_{j-1} \cap_j |_{k-j-2} : \{i_1 < i_2 < \cdots i_j < i_j + 1 < \cdots < i_k\}$$

$$\to \{i_1 < i_2 < \cdots i_{j-1} < i_{j+2} < \cdots < i_k\}.$$

Similarly, when $i_{j+1} - i_j > 2$, there is a morphism:

$$|_{j-1} \cup_{i_j+1} |_{k-j-2} : \{i_1 < i_2 < \cdots i_j < i_{j+1} < \cdots < i_k\}$$

$$\to \{i_1 < i_2 < \cdots i_j < i_{j'} < i_{j'+1} < i_{j+1} < \cdots < i_k\}$$

where the symbol j' does not carry any meaning until the resulting set is reindexed.

There is a tensor product of two subsets A and B that is obtained by shifting B (if necessary) to a set B' that is completely to the right of A, and juxtaposing A and B. Similarly morphisms may be tensored. Thus the morphism $|_{j-1} \nearrow_j |_{k-j-1}$ is strictly speaking the tensor product of an identity morphism on a $(j-1)$-element subset, a northeastern morphism on one object, and another identity morphism.

At the level of objects the structure is hardly associative. Indeed, the initial shift of the object B can affect the grouping of the resulting object. The next paragraphs contain more details of the lack of associativities in the tensor product.

G. Sica (ed.) *What is Category Theory?*

There is a category whose objects are finite elements in $2^{\mathbb{N}}$ and whose morphisms are generated by $\cup$, $\cap$, $|$, $\nearrow$ and $\nwarrow$. The identity object is the empty set, the identity morphism on $\{i_1 < i_2 < \ldots < i_k\}$ is the tensor product of $|$; specifically $|_{i_1,i_2,\ldots,i_k}$. The geometric depiction of the five generating morphisms is given in Fig. 1.

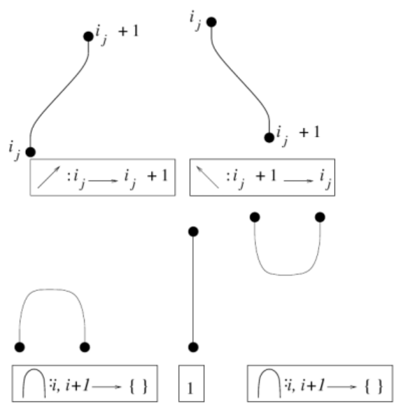


Figure 1: The generating 1-morphisms in the 2-category $\mathcal{EMB}_0$

The objects in this category determine partially associated products of indeterminants as follows: A given set $\{i_1 < i_2 < \ldots < i_k\}$ has its elements associated by their proximity to each other. Thus elements that are closer together are viewed as having been associated. For example, $\{1, 2, 4, 7\}$ is related to the associated product $((ab)c)d$ since the first two elements are closest and since 4 is closer to 2 than it is to 7. If a subset consist of elements in an arithmetic progression, then these element can be treated as a partially parenthesized product. For example, the set $\{1, 2, 3, 4, 5, 8, 10, 12\}$ is determines the partially parenthesized product $(x_1x_2x_3x_4x_5)(x_8x_{10}x_{12})$, and so in this case, we assume that the products, $x_1x_2x_3x_4x_5$ and $x_8x_{10}x_{12}$ are well-defined. Once an un-associated product is presented, then we treat it as a single entity.

More rigorously, to a given set $S{=}\{i_1 < i_2 < \ldots < i_k\}$ there is a corresponding monomial in non-associative variables $x_{i_1}, \ldots, x_{i_k}$. There is an association of these variable that is determined by the proximity of their subscripts. To determine the association compute the minimum among the differences $i_{j+1} - i_j$. If j achieves the minimum value, then parenthesize (x_jx_{j+1}). If the sequence $i_j, i_{j+1}, \ldots i_{j+\ell}$ is in arithmetic progression and this achieves the minimum difference, then group all of the corresponding variables together. By induction, an association scheme can be constructed. Here is one further example: $\{1, 3, 4, 5, 7, 10\}$ corresponds to $(x_1(x_3x_4x_5)x_7)x_{10}$. The products $(x_3x_4x_5)$ and $x_1(x_3x_4x_5)x_7$ are only partially associated since the corresponding subscripts are equally close.

The correspondence between a given subset and a partially parenthesized

G. Sica (ed.) *What is Category Theory?*

product is well-defined, but different subsets can determine the same product. It is possible to choose standard representatives of each partially parenthesized product by rescaling the subset.

We can think of a pair of subsets, $\{i_1 < i_2 < \ldots < i_k\}$ and $\{j_1 < \ldots < j_{k'}\}$ where $i_k < j_1$ as being a distinct entity from $\{i_1 < \ldots i_k < j_1 < \ldots < j_{k'}\}$. Geometrically, to make this distinction we can add an increment to each element of the latter set so that there is a parenthetical insertion forced between the two sets. To exemplify these ideas consider $\{1, 2, 4, 7\}$ and $\{1, 4, 6, 7\}$. The tensor product is the set $\{1, 2, 4, 7, 8, 11, 13, 14\}$. This corresponds to the parenthesized string $((x_1x_2)x_4)(x_7x_8)(x_{11}(x_{13}x_{14}))$. Whereas $\{1, 2, 4, 7, 11, 14, 16, 17\}$ corresponds to $(((x_1x_2)x_4)x_7)(x_{11}(x_{14}(x_{16}x_{17})))$. Thus the structure of the 1-category at this level is not monoidal.

If we want to shift $\{j, j+1, \ldots, j+k\}$ right one unit to the set $\{j+1, j+2, \ldots, j+k+1\}$, the topography of the morphism $\nearrow$ allows all of these points to be moved simultaneously. Thus $\nearrow_{\{1,\ldots,k\}}$ denotes the tensor product of $\nearrow_1, \ldots, \nearrow_k$. And right shifts to $\{j, j+1, \ldots, j+k\}$ can be achieved using the same scheme.

The process of moving elements from one subset to another by shifting them individually achieves a method of re-associating the objects. In this manner, the morphisms $\nearrow$ and $\nwarrow$ are associators. Thus the 1-category, $\mathcal{EMB}_0$, is premonoidal provided the following two transformations (and their variants) of the set $\{1, 2, 4, 7\} \to \{1, 4, 6, 7\}$ given as

$$\begin{aligned}
&|_1 \nearrow_3 |_{6,7}(|_1 \nearrow_2 |_{6,7}(|_{1,2} \nearrow_5 |_7(|_{1,2} \nearrow_4 |_7(\{1, 2, 4, 7\})))) \\
&= |_1 \nearrow_3 |_{6,7}(|_1 \nearrow_2 |_{6,7}(|_{1,2} \nearrow_5 |_7(\{1, 2, 5, 7\}))) \\
&= |_1 \nearrow_3 |_{6,7}(|_1 \nearrow_2 |_{6,7}(\{1, 2, 6, 7\})) \\
&= |_1 \nearrow_3 |_{6,7}(\{1, 3, 6, 7\}) \\
&= \{1, 4, 6, 7\}
\end{aligned}$$

and

$$\begin{aligned}
&|_{1,4} \nearrow_5 |_7(|_1 \nearrow_{3,4} |_7(|_1 \nearrow_2 |_{4,7}(\{1, 2, 4, 7\}))) \\
&= |_{1,4} \nearrow_5 |_7(|_1 \nearrow_{3,4} |_7(\{1, 3, 4, 7\})) \\
&= |_{1,4} \nearrow_5 |_7(\{1, 4, 5, 7\}) \\
&= \{1, 4, 6, 7\}
\end{aligned}$$

are equal. Note that at this stage, only objects and morphisms have been defined. If we want a 1-category, then we need to assert identities among 1-morphisms. So if we want a premonoidal 1-category with objects the finite elements of $2^{\mathbb{N}}$ and with morphisms $\cup$, $\cap$, $\nearrow$, $\nwarrow$, and $|$, then we would assert (among other things) that these two transformations are equal. The transformations are encoding the pentagon relation.

However, in the 2-category, we might assert the existence of a 2-morphism $\phi_{1,2,4,7}$ such that

G. Sica (ed.) *What is Category Theory?*

$$\phi_{1,2,4,7}(|_{1,4} \nearrow_5 |_7) \circ (|_1 \nearrow_{3,4} |_7) \circ (|_1 \nearrow_2 |_{4,7})$$
$$= (|_1 \nearrow_3 |_{6,7}) \circ (|_1 \nearrow_2 |_{6,7}) \circ (|_{1,2} \nearrow_5 |_7) \circ (|_{1,2} \nearrow_4 |_7)$$

The map $\phi_{1,2,4,7}$ is called the *pentagonator*. An inspection of the superficial diagram of the pentagonator 2 reveals that it can be decomposed in terms of a more simple operation. Each arrow $\nearrow$ (or $\nwarrow$) is depicted as a neighborhood of an inflection point in its diagrammatic representation. Distant inflection points commute. And thus we can assert the existence of a more basic invertible 2-morphism, R, that represents the commutation of distant inflection points and that satisfies a Yang-Baxter type relation. In Baez and Langford [5], this morphism is a tensorator between the pair of distant arrows.

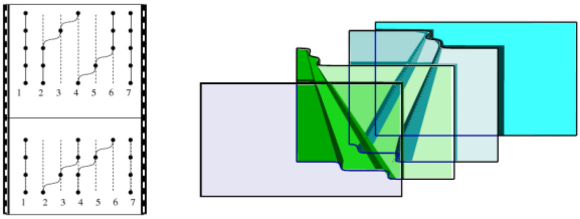


Figure 2: A surface that envelops the pentagonator

Observation. *The pentagonator can be written in terms of a tensorator.*
Proof.

$$R_{2,4}R_{3,5}R_{2,4}(|_{1,4} \nearrow_5 |_7) \circ (|_1 \nearrow_{3,4} |_7) \circ (|_1 \nearrow_2 |_{4,7})$$
$$= R_{2,4}R_{3,5}R_{2,4}(|_{1,4} \nearrow_5 |_7) \circ (|_{1,5} \nearrow_3 |_7) \circ (|_{1,2} \nearrow_4 |_7) \circ (|_1 \nearrow_2 |_{4,7})$$
$$= R_{2,4}R_{3,5}(|_{1,4} \nearrow_5 |_7) \circ (|_1 \nearrow_3 |_{5,7}) \circ (|_{1,4} \nearrow_2 |_7) \circ (|_{1,2} \nearrow_4 |_7)$$
$$= R_{2,4}(|_1 \nearrow_3 |_{6,7}) \circ (|_{1,3} \nearrow_4 |_7) \circ (|_1 \nearrow_2 |_{5,7}) \circ (|_{1,2} \nearrow_4 |_7)$$
$$= (|_1 \nearrow_3 |_{6,7}) \circ (|_1 \nearrow_2 |_{6,7}) \circ (|_{1,2} \nearrow_5 |_7) \circ (|_{1,2} \nearrow_4 |_7)$$

See also Fig. 3. This completes the proof.

The pentagonator is not, strictly speaking, restricted to acting on 1-morphisms of subset of size 4. A given partially associated product, $((AB)C)D$ with four constituents, $A = \{i_1 < i_2 < \ldots i_j\}$ and $(i_{k+1} - i_k)$ constant for $k = 1, \ldots j$ (similar descriptions for B, C, etc.) can be reassociated by similar inflection points (literally $\nearrow_B$ and so forth), and the resulting pentagonator is a "pleated multi-surface" in which the pleats commute past each other via corresponding Yang-Baxterators.

In addition to Yang-Baxterators between distant (and therefore commuting inflection points), we assert the existence of Yang-Baxterators between commuting critical points. (Baez and Langford treat all such 2-morphisms as tensorators).

G. Sica (ed.) *What is Category Theory?*

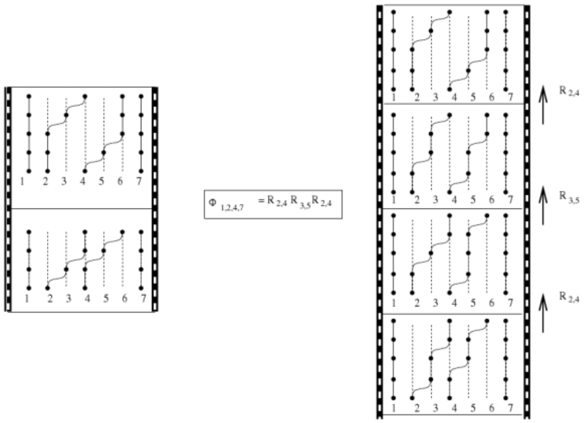


Figure 3: Factoring the pentagonator via Yang-Baxterators

Thus the compositions $(\cap_{1,2}) \circ (|_{1,2}\cap_{3,4})$ and $(\cap_{3,4}) \circ (\cap_{1,2}|_{3,4})$ are related by a 2-morphism (also called R) and other commutations of critical points or inflection points are given by similar 2-morphisms.

Further contrasts in the 1-category and the 2-category. At the level of a 1-category, one may wish to assert an identity among 1-morphisms such as $\cup_{1,2} \circ \cap_{1,2} = |_{1,2}$. The geometric reason is that a given set has a pair of successive elements annihilated by means of a $\cap$, and then the pair is recreated via a $\cup$. In a physical analogy, a pair of particles annihilate each other, emit a photon, and then the photon splits into a pair of identical particles. If the emission of the photon is so short lived that it can't be detected, then we would not want to include its life in a measurement.

Similarly, we may want to assert an identity between $(\cap_{1,2}|_3)\circ(|_1\cup_{2,3})$ and $|_1$. Such identities have obvious analogues that are dependent upon shifting indices, and the second identity has an analogue obtained by turning the diagram for $(\cap_{1,2}|_3) \circ (|_1\cup_{2,3})$ upside-down.

In a rigid monoidal category with duals, one asserts that $(\cap_{1,2}|_3) \circ (|_1\cup_{2,3}) = |$. (In the premonoidal case, the equality might be written as $(\cap_{1,2}|_3)\circ(|_1\cup_{2,3}) = \nearrow$). In this context, there is a self-dual object generator V. The maps $\cup$ and $\cap$ represent respectively represent a double map $\cup : 1 \to V \otimes V$, and an evaluation map $\cap : V \otimes V \to 1$. Here 1 represents the identity object (the empty set), and V represents a single point. In [42] associators are given to axiomatize a monoidal category with duals, but therein there are no diagrammatic representations given of associators are given.

In the current 2-category, we assert the existence of 2-morphisms: $S : \cup_{1,2} \circ \cap_{1,2} \Rightarrow |_{1,2}$ — a saddle point, $C : (\cap_{1,2}|_3) \circ (|_1\cup_{2,3}) \Rightarrow |_1$— a cusp (called a triangulator in [5]), $B : \emptyset \Rightarrow \cap \circ \cup$ — a birth of a simple closed curve, and

G. Sica (ed.) *What is Category Theory?*

$D : \cap \circ \cup \Rightarrow \emptyset$ — a death of a simple closed curve. These 2-morphisms are depicted via their diagrams on the top row of Fig. 4. We note here that the rest of the figure refers to the next subsection.

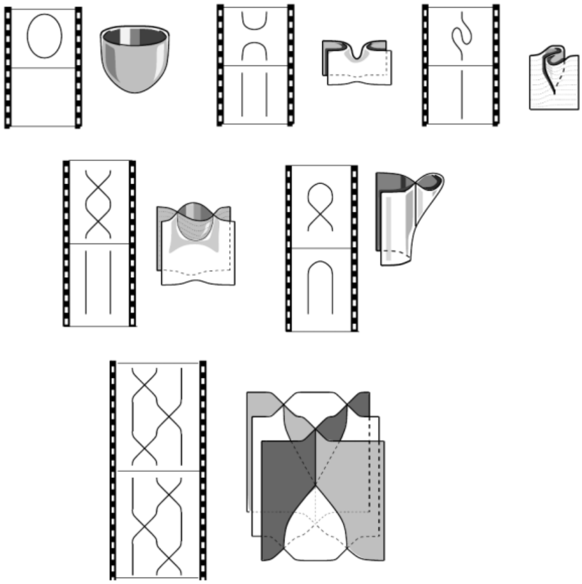


Figure 4: The 2-morphisms in the category of immersed surfaces

Furthermore, one asserts the existence of the two 2-morphisms that are depicted in Fig. 5. The first represents a 2-morphism bend $: \nwarrow \circ \nearrow :\Rightarrow |$, and the second represents a 2-morphisms curve : $\cap_{1,2} \circ \nearrow_{1,2} \Rightarrow \cap_{1,2}$. The former replaces the assumption that $\nearrow$ and $\nwarrow$ are inverse morphisms, and the latter replaces the assumption that the location of the cap is immaterial.

The generating objects and morphisms in the 2-category $\mathcal{EMB}_0$. The objects in the category are finite subsets of the positive integers. The generating 1-morphisms are $\nearrow$, $\nwarrow$, $|$, $\cup$, and $\cap$. The set of morphisms between two subsets $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_{k'}\}$ is empty if $k \equiv k' + 1 \pmod 2$, otherwise, it consists of the set of diagrams of disjoint arcs that inner-connect the points in the subsets. An arc that connects a point i_ℓ at the bottom to j'_ℓ at the top, has the opportunity to meander to the left or right, up or down before it makes the connection. We can decompose this meandering in terms of the generating morphisms by choosing suitable height functions. The generating 2-morphisms consist of births, B, and deaths, D, of simple closed curves ($\emptyset \Rightarrow \cap \circ \cup$ or $\cap \circ \cup \Rightarrow \emptyset$), saddles, S, ($\cup \circ \cap \Rightarrow ||$ and vice-versa), cusps, C ($(\cap \otimes |) \circ (| \otimes \cup) \Rightarrow |$ and vice-versa), tensorators, bends, curves and their inverses.

A topological/categorical question. In the current set up, suppose that two

G. Sica (ed.) *What is Category Theory?*

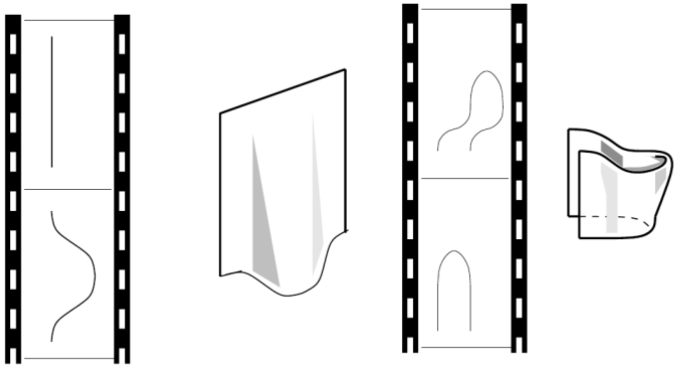


Figure 5: Extra 2-morphisms in the pre-monoidal setting

1-morphisms both with source $A = \{i_1, \ldots, i_k\}$ and target $B = \{j_1, \ldots, j_\ell\}$. are given. An uncategorical question to ask would be, "Are these 1-morphisms equal?" It is uncategorical, since we have been reminded that in higher category theory, that it is undesirable and unnatural to consider different things to be equal. Instead we can ask, "Is there a 2-morphism connecting them?" To ask this becomes a topological/geometric question. We have two pairs of fixed arcs meandering and inter-connecting among the specific subsets A and B. The topological question is if there is a surface with boundary that interpolates between the two arcs and if that surface can be decomposed in terms of the generating 2-morphisms that have been written down.

To address this question, suppose there is a surface, project it to a plane (called the retinal plane below), and track critical arcs and arcs of inflection points in the retinal plane. I expect that a Cerf theory type argument like that given in [17] will show that these 2-morphisms suffice to connect the two 1-morphisms. In general, it is not hard to construct a surface that interpolates between the pair of 1-morphisms. The first question is whether or not such a surface can be decomposed in terms of our generating 2-morphisms.

If so, then we ask if there is a 2-morphism that connects these two 1-morphisms that does not factor with any saddle points maxima or minima. Clearly, the number of arcs that start and end at either A, B, or travel between them has to be the same for each 1-morphism. The reason for excluding the critical points is that these are precisely the 2-morphisms that we do not want to consider to be invertible. Thus the categorical question is, "Are the two 1-morphisms equivalent in the sense that they are related by a sequence of invertible 2-morphisms?"

Composition of 2-morphisms. In $\mathcal{EMB}_0$ or any of our subsequent 2-categories, 1-morphisms are composed by juxtaposing vertically. In my conventions, the com-

G. Sica (ed.) *What is Category Theory?*

position $f \circ g$ is represented with f above and g below. If f and g were functions, then the domain of the function would be represented at the bottom of the diagram, g would be applied first, and then f would be applied. The advantage is that when reading from left to right we can draw from top to bottom. In the abstract tensor notation of Kauffman's work [34], this means that the matrices represented by the diagrams are applied to column vectors that appear on the right of an equation.

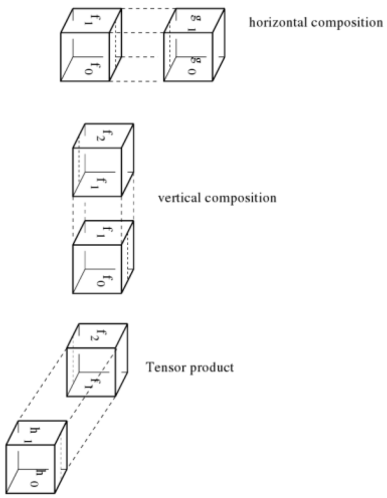


Figure 6: Compositions of 2-morphisms

In any 2-category, there are vertical and horizontal composition of 2-morphisms. The horizontal composition of two morphisms $F : f_0 \Rightarrow f_1$, and $G : g_0 \Rightarrow g_1$ is defined when the sources of the arrows g_0 and g_1 coincide with the targets of the arrows f_0 and f_1. In our movie description of the 2-morphisms, successive stills in the movie differ by a composition of 2-morphism. To compose two 2-morphisms horizontally, we create two movies representing the 2-morphisms. The target of the 1-morphisms f_0 and f_1 consists of, say, m points and these are the sources of the 1-morphisms g_0 and g_1. We can compose the two 2-morphisms by super-imposing the two changes in scenes that are represented by F and G. In other words, both F and G are represented by surfaces with boundary, and the right boundary of F is glued to the left boundary of G. The vertical composition of two 2-morphisms is represented by the succession of stills.

To be more specific, Suppose that the source 1-morphism $\hat{n}$ of f_0 is depicted on the left bottom edge of a box that contains the surface representing the 2-morphism F. The left to right axis on the 1-morphism diagram translates to the back to front axis (x-axis) on the box. The left face (xz-plane, $y = 0$) of the box contains the arcs $\hat{n} \times I$. The right face of the box contains the arcs, $\hat{m} \times I$, that

G. Sica (ed.) *What is Category Theory?*

consist of the targets of f_0 and f_1 multiplied by an interval factor. The bottom face (xy-plane $z = 0$) contains the diagram for f_0. The top face ($z = 1$) contains the diagram for f_1. The two sides of the box, $x = 0$ and $x = 1$, are empty. The description of the box holding a 2-morphism, G, is similar. Now the vertical composition of F and G is obtained by stacking G on top of F and rescaling the vertical dimension of the box. The horizontal composition is obtained by gluing the boxes together with F on the left and G to the right. Meanwhile, the tensor product of 2-morphisms is obtained by juxtaposing the boxes in the x direction. Figure 6 indicates the various compositions. Observe that the 4-square relation is a consequence of the fact that these four boxes can be stacked two on top of two, or two to the right of a stack of two.

A further simplification. To finish specifying the 2-category, we need to specify relations among the 2-morphisms. The situation is getting complicated, because heretofore we have assumed a pre-monoidal structure. Henceforth, an object $\hat{k}$ will be any set of numbers with k elements where $k = 0, 1, \ldots$. We depict such a set as $\hat{k} = \{1, \ldots, k\}$, and the case $k = 0$ corresponds to the empty set. However, the location of the points on the line is assumed to be immaterial. Thus we abandon the need to specify left-to-right ($\nearrow$) or right-to-left ($\nwarrow$) motion in the diagrams for the 1-morphisms. The tensor product $\hat{k} \otimes \hat{k}' = \widehat{k + k'}$ is strictly associative, and any arc that has no critical points is assumed to be the identity. Otherwise, arcs are assumed to have generic critical points (x^3 is excluded for example). The generating 1-morphisms are $\cup$, $\cap$, and $|$. The generating 2-morphisms are B, D, S, C, and tensorators— those that involve the commutation of distant critical points. In this 2-category $\mathcal{EMB}$, we will describe the relations among 2-morphisms. The pre-monoidal structure has been adjusted to be monoidal.

The cusp (triangulator) 2-morphism is invertible. There are two types of cusps: those that create a pair of critical points, and those that cancel a pair of critical points. Thus there are two possible vertical compositions of cusps. The terminology from singularity theory is *lips* or *beak-to-beak*. A saddle point can cancel with a local maximum or minimum. There is a *swallow-tail* cancelation of a pair of cusps in the presence of a Yang-Baxterator for the critical points. Finally, a cusp can be turned upside-down in a the presence of a saddle point. All of these relations can be easily depicted in terms of the fold lines and cusps that appear on the projection of embedded surfaces in 3-space. These relations are depicted in Figs. 7, 8, 11, and 9. Observe that the right had parabolic cylinder is unnecessary to the relationship — it is drawn only to anchor the idea.

In a rigid monoidal 1-category with duals, the maps $\cup$ and $\cap$ satisfy the relation $(\cap \otimes |) \circ (| \otimes \cup) = |$. In the corresponding 2-category we relax this condition to the existence of a 2-isomorphism $(\cap \otimes |) \circ (| \otimes \cup) \Rightarrow |$. This 2-morphism is depicted in the graphical notation as a cusp on the projection of a surface. The fact that it is a 2-isomorphism is depicted in Fig. 11; the figure indicates that the two movies representing the vertical composition of 2-morphisms represent the same surface. Or in categorical language there is a commuting polytope of the

G. Sica (ed.) *What is Category Theory?*

corresponding 2-morphism. Another relation that is imposed upon the duality 2-morphism by the geometry is the swallowtail relation that is depicted in Fig. 8. In general, a full list of relations among 2-morphisms consist of the following:
Relations among the 2-morphisms.

- commuting distant 2-morphisms;
- Yang-Baxter relations among nearby tensorators;
- the cancelation or introduction of a pair of similar such tensorators ($ab \Rightarrow ba \Rightarrow ab$)
- the cancelation or introduction of a saddle point and a birth or death (Fig. 7);
- lips cancelations or introduction of cusps (Fig. 11);
- beak-to-beak cancelation or introduction of cusps (Fig. 11);
- swallow-tail cancelation or introduction of cusps (Fig. 8);
- interchanging an upward cusp and a saddle point with a saddle point and a downward cusp (Fig. 9).

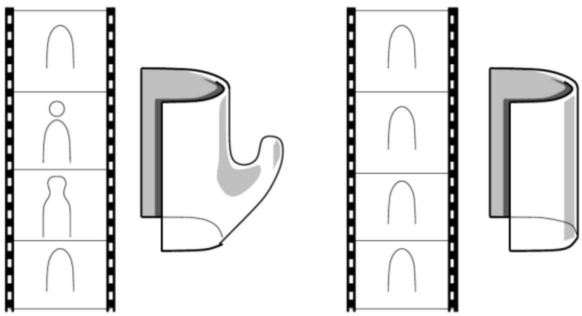


Figure 7: The cancelation of a maximum point and a saddle point

Now all of these relations can be written in purely categorical language. See [5] for example. Moreover each is quantified as a codimension 1 singularity between surface maps, and each such relation can be thought of kinematically. I encourage you to determine if the free monoidal 3-category with duals on one self dual object generator consists of the non-negative integers as objects with addition of integers as the monoidal structure, 1-morphisms generated by $\cup$, $\cap$ and $|$. Generating 2-morphisms would be given as cusps, saddles, births, and deaths. And generating 3-morphisms would be the singularities above, births and

G. Sica (ed.) *What is Category Theory?*

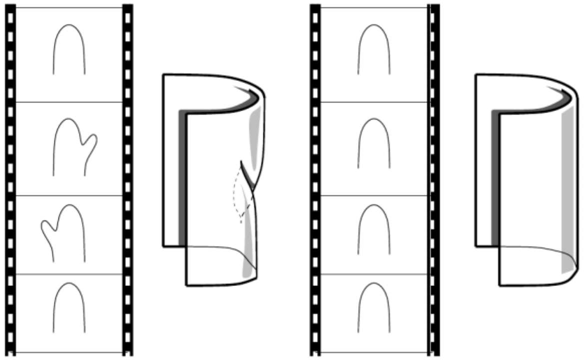


Figure 8: The swallow-tail relation among 2-morphisms

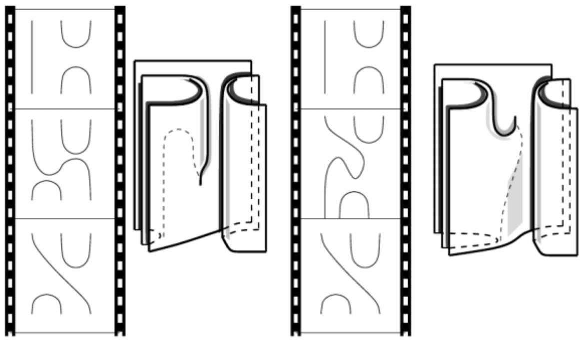


Figure 9: The horizontal cusp relation among 2-morphisms

deaths of spheres, and the attachments of 1-handles and 2-handles. Relations among the 3-morphisms would be generated by ambient isotopy of embedded 3-manifolds in 4-space. I expect that a complete list of relations can be given via the study of singularities between smooth 3-dimensional manifolds.

Conjecture. *The free monoidal* 3*-category with duals on one self dual (unframed) object generator is the* 3*-category of embedded* 3*-manifolds in* 4*-space.*

The 2-category of embedded surfaces. Given an embedding of a closed surface, F, in 3-space, it can be decomposed as a sequence of 2-morphisms as follows. A plane in 3-space that is disjoint from the surface F is chosen so that the projection of F to this plane has generic cusps and folds.This plane is called *the retinal plane.* General position considerations and Whitney's theorem guaranty that such a plane can be found; the set of planes form an open dense subset of the set of all planes. A height direction in the retinal plane is chosen so that the

G. Sica (ed.) *What is Category Theory?*

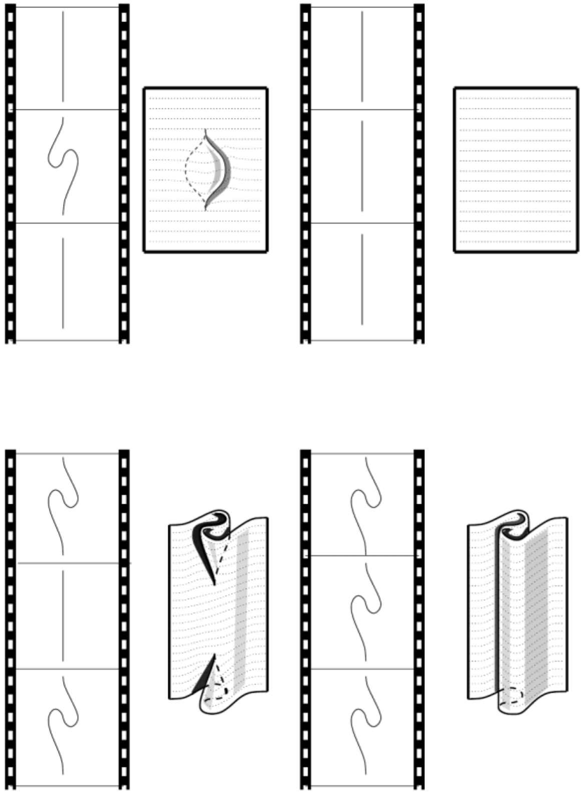


Figure 10: The lips and beak-to-beak relations among 2-morphisms

subsequent projection of the surface onto this line has non-degenerate critical points; that is, in local coordinates, the Hessian is non-singular. The projection of the surface onto the height axis can be given as the restriction of the projection of 3-space onto this axis. Choose a third direction perpendicular to the plane of projection: *a line of sight*, and a particular line parallel to this third direction. That is a line in 3-space whose direction vector is a scalar multiple of the vector defining the line of sight.

The intersection of the surface F with a general line parallel to the line of sight will consist of a finite collection of points such that the tangent to the surface at any such intersection maps injectively to the retinal plane. The distances between successive points on this line can be measured, and the set of these distances can be compared. In general, the intersections will not be integral multiples of a fixed length, but a subset of the positive integers can be established so that there is an order preserving map from the set of intersections of the given

G. Sica (ed.) *What is Category Theory?*

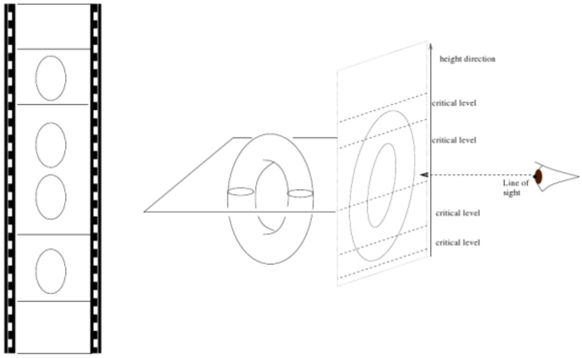


Figure 11: The retinal plane, height function, and resulting movie

line to a subset of integers that satisfies the following additional property: Points of intersection that are closer on the line of intersection correspond to similarly closer points on the subset of the integers. Thus not only is the order of intersection preserved but a new scale is chosen such that the relative distances are also preserved. The set of distances between points of F on the line of intersection can be partially ordered, and the subset of the integers is chosen so that if adjacent points of intersection x and y are closer together than the adjacent points z and w on the line of intersection, then the corresponding distances are closer ($|i_x - i_y| < |i_z - i_w|$ where $i_x, i_y, i_z, \&\ i_w$ are the corresponding integral points).

Any two lines parallel to the line of sight determine a plane, and the intersection of the given surface with this plane is a sequence of arcs. Critical points of the arcs, will project to points on the fold set in the retinal plane. If we consider one of the original lines as the bottom of an infinite rectangular strip, and the other line to be the top of such a strip, then the arcs undulate in this rectangular strip. This undulation is modeled via a sequence of generating 1-morphisms between the corresponding subsets of the positive integers.

As outlined in the previous two paragraphs, it is possible to start from the image of a given generic surface, measure carefully, and develop a sequence of 2-morphisms in the 2-category $\mathcal{EMB}_0$ that approximates the topography of the surface. To do so rigorously would require us to show that given surface can be suitably approximated. For example, for two lines of intersection at the same height level, results in two model subsets. The intersection of the surface with the thin strip that is bounded by these lines, consists of a sequence of arcs. These arcs have to be replaced by a sequence of 1-morphisms in $\mathcal{EMB}_0$, and the replacements have to be consistent from height to height. Specifically, folds and bulges have to be preserved. While I have not proven that there is a 2-isomorphism of 2-categories (a pair of 2-functors with a natural equivalences between their compositions and the identity functor), I feel confident that such

G. Sica (ed.) *What is Category Theory?*

a 2-equivalences can be constructed. My level of confidence is bolstered from my experience in drawing and shading surfaces. The drawings themselves are projections to the plane, and the technique that I use for depicting subtle details is exactly the idea of putting certain layers of the surface into a standard position, and isolating the singularities in specific locations.

Now in general, we can choose a sequence of parallel planes arranged as cosets of the height function. Literally the height function is defined on all of 3-space, and each plane is $\pi_t = h^{-1}(t)$ for some $t \in \mathbb{R}$. That is, for a given height, we choose a plane at that height, perpendicular to the height axis. For all but a finite number of heights, t, the intersection of $\pi_t = h^{-1}(t)$ with the surface, F, is generic and consists of a collection of closed loops. As such, it is modeled by the composition of 1-morphisms whose ultimate source and target are the empty sequence. A pair of nearby planes, between which are no critical or cuspal levels, will have intersection loops that differ only in position. On the other hand, two planes on either side of a critical point or cusp, will contain arcs with differing critical behavior, where now critical points are measured via a height function in this plane at a given vertical height.

The non-critical changes correspond to portions of the surface being closer or further from the retinal plane. It is these changes that we cease attempting to measure. The critical points corresponds to saddles, births, and deaths with respect to the height function, and the cusps correspond to changes in the fold set of the projection onto the retina.

Two embeddings of the surface that are ambiently isotopic can be related to one another via a sequence of relations among the 2-morphisms. In [18], we gave a graphical description of the ambient isotopy that converts a coffee cup into a doughnut. This isotopy involved the basic relations among folds that we have given here. It is amusing and satisfying to carefully watch the motions of a person while considering the changes in folds and cusps as the person moves. For example, the motion of the legs of a walking person as viewed from the side involves the commutation of distant fold lines. A knee bending and unbending can be approximated by a swallow-tail change. The junction of an arm and a shoulder involves a pair of cusps, the folds that are the profile of the biceps and triceps, and the fold of the arm pit. Line drawings, comics, and artist sketches exploit Whitney's theorem (or realize it) that states that the generic projection of one surface onto another is locally one-to-one and non-singular on all but a set of measure zero, and the measure zero set consists of fold lines that close or end in cusps.

2.2 Generic surfaces in 3-space

A closed surface in general position in 3-space has arcs of double points that end at branch points or triple points. In a neighborhood of a double point, the surface looks like the intersection of 2 coordinate planes in 3-space. In a neighborhood of a branch point the surface looks like the cone on a figure 8, and in a neighborhood

G. Sica (ed.) *What is Category Theory?*

of a triple point the surface resembles $\{(x, y, z) : xyz = 0\}$ — the set of three coordinate planes in 3-space.

We define the 2-category, $\mathcal{GEN}$, by introducing a new generating 1-morphism, X, that consist of the commutation of a pair of dots, and we introduce new 2-morphisms and relations among these. More specifically, we take objects to correspond to integers, $n = 0, 1, 2, \ldots$ as above. the object $\hat{n}$ is arranged as n dots along a line. Define an associative tensor product $\hat{n} \otimes \hat{m} = \widehat{n+m}$, and consider generating 1-morphisms $|, \cup, \cap$, and X. The diagrammatic depiction of X is given in Fig. 12.

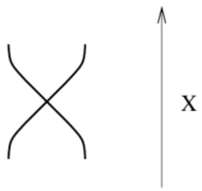


Figure 12: The generating 1-morphism X

The generating 2-morphisms consist of births and deaths of simple closed curves ($\emptyset \Rightarrow \cap \circ \cup$ or $\cap \circ \cup \Rightarrow \emptyset$), saddles ($\cup \circ \cap \Rightarrow ||$ and vice-versa), cusps ($(\cap \otimes |) \circ (| \otimes \cup) \Rightarrow |$ and vice-versa), branch points ($X \circ \cup \Rightarrow \cup$, vice-versa, $\cap \circ X \Rightarrow ||$ and vice-versa), cancelation of a pair of successive commutations ($X \circ X \Rightarrow ||$ and vice versa), the commutation of distant 2-morphisms, and a Yang-Baxter 2-morphism:

$$(X \otimes |) \circ (| \otimes X) \circ (X \otimes |) \Rightarrow (| \otimes X) \circ (X \otimes |) \circ (| \otimes X).$$

Representatives of each type of these 2-morphisms are given in Fig. 4. There is one additional 2-morphism to include. It is depicted separately in Fig. 13. (The reasons for separating this figure from Fig. 4 was forgetfulness rather than any mathematical reason. As I was finishing the manuscript, I realized I didn't have the source code for Fig. 4 at the place I was preparing the text). The 2-morphism can be described as $\psi : | \circ (\cap \otimes |) \circ (| \otimes X) \Rightarrow | \circ (| \otimes \cap) \circ (X \otimes |)$.

The 1-morphisms are akin to representatives of elements of the Brauer group, and the 2-morphisms are formed from relations therein, saddles, cusps, and births and deaths of simple closed curves.

A generic closed surface, F, in 3-space gives rise to a sequence of 2-morphisms from the empty 2-morphism and back as in the embedded surface case. The difference is the existence of the 1-morphisms X and the resulting 2-morphisms. These are branch points and triple points of the image of the surface.

From a categorical point of view, the 1-category that has objects $\{\hat{n}, n \in \{0, 1, 2, \ldots\}\}$ and morphisms generated by $|, \cup, \cap$, and X is a symmetric monoidal category. Upon imposing relations among the morphisms that correspond to some

G. Sica (ed.) *What is Category Theory?*

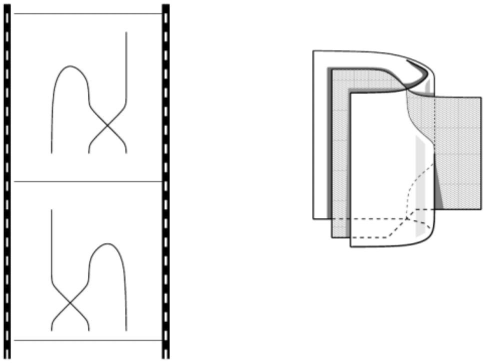


Figure 13: Moving a crossing over a fold

of the 2-morphisms (specifically the cusp relation gives rigidity and the branch point gives the category being pivotal), this symmetric category can be described algebraically. As an exercise, the reader should formulate a Freyd-Yetter type theorem that states that this category is the free symmetric, rigid, pivotal, monoidal category with duals on one self-dual object generator. The typical knot diagrammatic calculi can be used to perform computations in this category.

Homma-Nagase [31] and independently Roseman [44] demonstrated that there is a set of 7 local moves such that any two isotopic generic maps of closed surfaces in 3-dimensional manifolds can be transformed, one to the other, via a finite sequence of applications of these moves. In 3-space, any two orientable generic surfaces are isotopic if and only if the underlying surfaces are homeomorphic. Subsequently, Goryunov [24] gave a list of these moves as real pictures of codimension 1 singularities in the space of multi-germs of complex maps. The Roseman moves for generic surfaces have movie parameterizations. It is these movie parameterizations and movie moves that parameterize the interactions with the fold set (such as lips, beak-to-beak, swallowtails), and the interactions of the double point set, branch point set, and triple point set with the fold lines. See [18] for a full description. To understand these results in categorical language see [5].

At a category theory conference in the distant past, a mathematician raised strong objections to our use of diagrams to encode these categorical aspects. The defense of the diagrammatic point of view here is that the free symmetric monoidal 2-category on one self dual object generator is almost certainly the 2-category whose 1-morphisms are generated by $|$, X, $\cup$, and $\cap$, whose 2-morphisms are generated by the surfaces depicted in Fig. 4, and whose relations among 2-morphisms are given as projections of the movie moves. Here *projections* means that the classical knot crossings are projected to the pair of intersecting arcs depicted in Fig. 12. Therefore, a calculation in any symmetric monoidal 2-category,

G. Sica (ed.) *What is Category Theory?*

can be encoded as a surface manipulation. It has become standard to depict computations in braided monoidal categories in graphical notation. The advantage to me, is that the calculation can be followed by comparison of diagrams, and this comparison is easier for me, than the comparison of algebraic expressions. In the 2-category setting, we wind up manipulating surfaces. Most of the truly tedious algebraic work in my papers with Masahico and others is informed via diagrams. My colleague who objected to their use, obviously, had not seen the mathematics in ironing a shirt or shaping a metal plate via hammering.

The images of any two generic surfaces of the same genus are isotopic, and an isotopy can be constructed as an application of a finite sequence of moves taken from among the Roseman moves. If height function information is also preserved, then the moves are selected from among the projections of the movie moves given in [17] (see also the chart moves below). The categorical meaning of these results (modulo the proof that the category of generic surfaces is the free symmetric monoidal 2-category on one self-dual unframed object generator) is that invariants of generic surfaces defined categorically can, at best, detect genus.

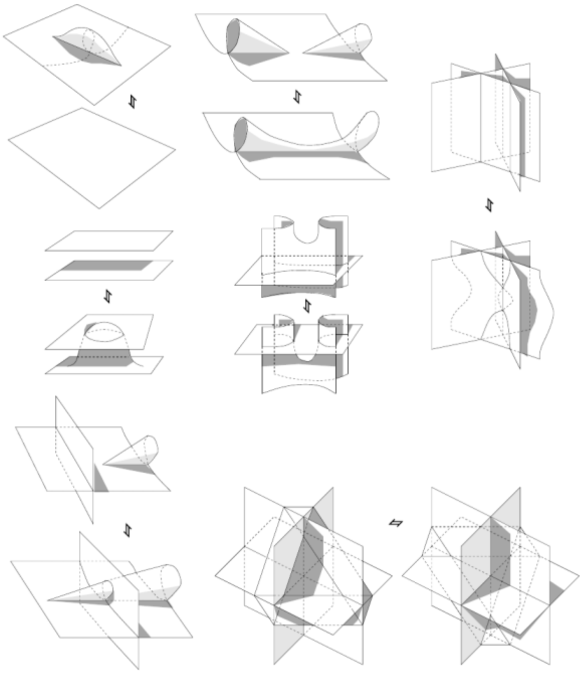


Figure 14: The Roseman moves

G. Sica (ed.) *What is Category Theory?*

2.3 The 2-category of knotted surfaces in 4-space

The current section is completed by replacing the 1-morphism X — which is a transposition of adjacent elements in the permutation group — with a pair of 1-morphisms X and $\bar{X}$ which are positive and negative braid generators in the braid group, respectively. Having introduced over and under crossing information in the 1-morphisms, the crossing information is extended to the 2-morphisms. Let me describe this further.

The notion of a classical knot diagram is a planar picture that depicts crossing information. The diagram represents an observer's view of the knot; arcs at crossings that are further from the observer are depicted as haven been broken in the diagram. If the 3 space in which the knot lives is very thin in the observation direction, then most of the knot is on the plane of observation. The under crossing arcs bend behind that plane. Metaphorically, imagine a collection of wires mounted on a wall. In order to avoid shorting out a circuit, when wires might cross on the wall, the under crossing arc is fed behind the wall through a pair of small holes.

I am belaboring this point with knot diagrams because I want to generalize it to knotted surfaces. Now suppose that a surface is embedded in a 4-space that consists of $\mathbb{R}^3 \times (-3\epsilon, 0]$. Most of the surface is embedded in the 3-space $\mathbb{R}^3 \times \{0\}$ in which you are reading this article. However, there are small sections of the surface, along would-be double arcs, triple points and near branch points at which the surface bends below $\mathbb{R}^3 \times \{0\}$, and protrudes in to the "vinn" — a term coined by Rudy Rucker (Up is to down, as right is to left, as fore is to aft, as vinn is to vout). Along arcs of double points the surface has the structure of an interval times a semi-circle. At branch points this structure tapers off, and at the lower sheet at a triple point, there is a pair of canals that accommodates the canal at the middle sheet. Figures for these surfaces are given in [18].

The 2 category of 2-tangles is an algebraic model for knotted surfaces in 4-space. Here is the complete description. As before the objects are finite sets of points along a line. The 1-morphisms consist of $|$, $\cup$, $\cap$, X and $\bar{X}$. The 2-morphisms consists of cusps $C : (\cap \otimes |) \circ (| \otimes \cup) \Rightarrow |$ (and variants, births and deaths of simple closed curves, $B : \emptyset \Rightarrow (\cap \circ \cup)$ and $D : (\cap \circ \cup) \Rightarrow \emptyset$ respectively, and saddles $S : \cup \circ \cap \Rightarrow ||$. Each of these 2-morphisms has an analogue in which source and target 1-morphisms are switched. Finally, the lifts of the 2-morphisms depicted in Figs. 4 and 13 are included. Here *lifts* means that the 1-morphism X in these figures is replaced by the classical knot crossing, either X or $\bar{X}$, so that the resulting broken surfaces have consistent crossing information along their double curves. More specifically, the resulting 2-morphisms are the classical Reidemeister moves (including the ψ move), births, deaths, saddles, and cusps.

The relations among 2-morphisms can be encoded in terms of their effects on their projections to the retinal plane. In [17] we gave a set of chart moves that encrypted all of the movie moves in the resulting theory; the figures also appear in [18]. Again all of these can be described in purely categorical means as

G. Sica (ed.) *What is Category Theory?*

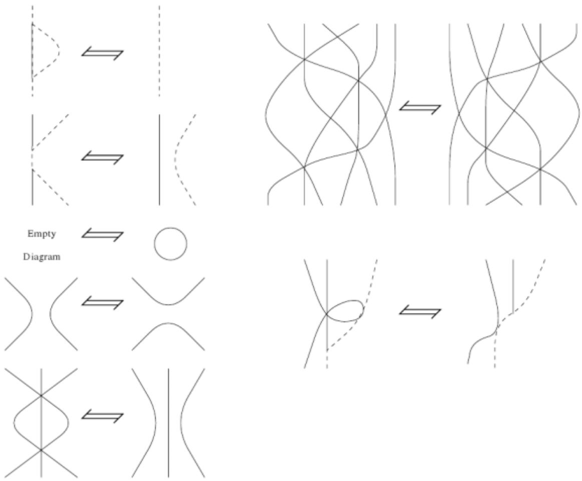


Figure 15: The chart moves page 1

in [5]. The following conventions were used in preparing these diagrams. A thin dashed line indicates a fold line in the retinal plane. A thick dashed line indicates either a fold or a crossing line. The crossings among thick dashed lines, then, are tensorators. A thin solid line represents a double arc from the projection of the knotted surface (originally in 4-space) to its diagram (in 3-space), and finally into the retinal plane. The three-fold intersection of solid lines represent triple points in the projection to 3-space (Reidemeister type III moves). The apparently tangential intersection between solid lines and dashed lines represents the 2-morphism ψ. The points at which arcs of double points appear to end at a fold line represent branch points (Reidemeister type I moves).

2.4 Why anyone else should care

An analogue of the Yang-Baxter equation is the Zamolodchikov equation (ZE) from statistical mechanics. In a braided monoidal 2-category, there is a solution to the ZE, and it has been shown [33], that a braided monoidal 2-category can be constructed from a solution to the ZE. Now our case of knotted surfaces has the added duality structure much of which I have hidden in the closet. Most of the duality has to do with the ability to include variants of the 2-morphisms obtained by reflecting them in various planes parallel to the faces of the boxes containing them. It has turned out that braided monoidal 2-categories with duals have been relatively difficult to find.

On the other hand some progress has been made. Most notably, is the Baez-Crans development [22, 3] in which categorifications of Lie Algebras give solutions

[9] Brieskorn, E. [1988] *Automorphic sets and braids and singularities,* Braids (Santa Cruz, CA, 1986), 45–115, Contemp. Math., 78, Amer. Math. Soc., Providence, RI.

[10] Brown R., Porter T., *Category Theory and Higher Dimensional Algebra: potential descriptive tools in neuroscience*, math.CT/0306223.

[11] Carter, J. S., Elhamdadi, M., Graña, M., Saito, M., [2005] *Cocycle knot invariants from quandle modules and generalized quandle homology*, Osaka J. Math. 42, no. 3, 499–541.

[12] Carter, J.S., Elhamdadi, M., Saito, M, [2004] *Homology theory for the set-theoretic Yang-Baxter equation and knot invariants from generalizations of quandles,* Fund. Math. **184**, 31–54.

[13] Carter, J.S , Elhamdadi, M., Saito, M., [2002] *Twisted Quandle homology theory and cocycle knot invariants* Algebraic and Geometric Topology, 95–135.

[14] Carter, J. S., Flath, D. E., Saito, M. [1995] "The classical and quantum $6j$-symbols," Mathematical Notes, 43. Princeton University Press, Princeton, NJ.

[15] Carter, J.S., Jelsovsky, D., Kamada, S., Langford, L., Saito, M., [2003] *Quandle cohomology and state-sum invariants of knotted curves and surfaces,* Trans. Amer. Math. Soc. **355**, no. 10, 3947–3989.

[16] Carter, J.S., Kauffman, L. H., and Saito, M., [1999] *Structures and diagrammatics of four dimensional topological lattice field theories*, Advances in Math. **146**, 39–100.

[17] Carter, J. S., Rieger, J. H., Saito, M., [1997] *A combinatorial description of knotted surfaces and their isotopies*, Adv. Math. **127**, no. 1, 1–51.

[18] Carter, J. S., Saito, M. [1998] "Knotted Surfaces and Their Diagrams," AMS Surveys and Mongraphs, **55** , Providence.

[19] Crane, L., Frenkel, I.B., [1994] *Four-dimensional topological quantum field theory, Hopf categories, and the canonical bases, Topology and physics,* J. Math. Phys. **35**, 5136–5154.

[20] Crane, L., Kauffman, L.H., Yetter, D., *Evaluating the Crane-Yetter invariant*, Quantum topology, 131–138, Ser. Knots Everything, 3, World Sci. Publishing, River Edge, NJ, 1993.

[21] Crane, L., Yetter, D., [1993] *A categorical construction of 4D topological quantum field theories*, Quantum topology, 120–130, Ser. Knots Everything, 3, World Sci. Publishing, River Edge, NJ.

[22] Crans, A., *Lie 2-Algebras*, math.QA/0409602.

[23] Freyd, Peter J., Yetter, D. N. [1989]*Braided compact closed categories with applications to low-dimensional topology*, Adv. Math. 77, no. 2, 156–182.

[24] Goryunov, V. V., [1997]*Local invariants of mappings of surfaces into three-space*, The Arnold-Gelfand mathematical seminars, 223–255, Birkhäuser Boston, Boston, MA.

G. Sica (ed.) *What is Category Theory?*

[25] Deligne, P. *et al.* [1999] "Quantum fields and strings: a course for mathematicians. Vol. 1, 2. Material from the Special Year on Quantum Field Theory held at the Institute for Advanced Study, Princeton, NJ, 1996–1997." AMS, Providence.

[26] Etingof, P., Graña, M. [2003]*On rack cohomology*, J. Pure Appl. Algebra 177, no. 1, 49–59.

[27] Fischer, J.E., Jr., [1994] *2-categories and 2-knots*, Duke Math. J. **75** , 493–526.

[28] Graña, M., [2002] *Quandle knot invariants are quantum knot invariants*, J. Knot Theory Ramifications **11**, 673–681.

[29] Fenn, R., Rourke, C., [1992] *Racks and links in codimension two,* Journal of Knot Theory and Its Ramifications Vol. 1 No. 4, 343–406.

[30] Fenn, R., Rourke, C., Sanderson, B., [1993]*An introduction to species and the rack space*, Topics in knot theory (Erzurum, 1992), 33–55, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., **399**, Kluwer Acad. Publ., Dordrecht.

[31] Homma,T., Nagase, T., [1985]*On elementary deformations of the maps of surfaces into 3-manifolds I*, Yokahama Mathematical Journal **33**, 103-119.

[32] Joyce, D., *A classifying invariant of knots, the knot quandle,* J. Pure Appl. Alg., 23, 37–65.

[33] Kapranov, M. M., Voevodsky, V. A., [1994] *2-categories and Zamolodchikov tetrahedra equations*, Algebraic groups and their generalizations: quantum and infinite-dimensional methods (University Park, PA, 1991), 177–259, Proc. Sympos. Pure Math., 56, Part 2, Amer. Math. Soc., Providence, RI.

[34] Kauffman, L.H., [1991]*Knots and Physics*, World Scientific, Series on knots and everything, vol. 1, Singapore.

[35] Kauffman, L.H., [1993] *Knots and topological quantum field theory,* Proceedings of the Conference on Quantum Topology (Manhattan, KS, 1993), 137–186.

[36] Kauffman, L.H., [1999] *Virtual Knot Theory*, European J. Combin. **20**, 663–690.

[37] Kharlamov, V. M., Turaev, V. G., [1996] *On the definition of the 2-category of 2-knots*, Mathematics in St. Petersburg, 205–221, Amer. Math. Soc. Transl. Ser. 2, 174, Amer. Math. Soc., Providence, RI.

[38] [2001] Kerler, T.; Lyubashenko, V.V., *Non-semisimple topological quantum field theories for 3-manifolds with corners,* Lecture Notes in Mathematics, 1765. Springer-Verlag, Berlin.

[39] Lopes, P., [2003]*Quandles at finite temperatures. I*, J. Knot Theory Ramifications **12**, no. 2, 159–186.

[40] Matveev, S., [1982] *Distributive groupoids in knot theory,* (Russian) Mat. Sb. (N.S.) 119(161), no. 1, 78–88, 160.

[41] Mochizuki, T., [2003] *Some calculations of cohomology groups of finite Alexander quandles*, J. Pure Appl. Algebra, 179, 287–330.

G. Sica (ed.) *What is Category Theory?*

[42] Ng, S.H.; Schauenburg, P. *Frobenius-Schur Indicators and Exponents of Spherical Categories*, math.QA/0601012.

[43] Reshetikhin, N., Turaev, V. G., [1991] *Invariants of* 3*-manifolds via link polynomials and quantum groups,* Invent. Math. **103**, 547–597.

[44] Roseman, D. [1998], *Reidemeister-type moves for surfaces in four-dimensional space,* Knot theory (Warsaw, 1995), 347–380, Banach Center Publ., 42, Polish Acad. Sci., Warsaw.

[45] Turaev, V. G. [1994] "Quantum invariants of knots and 3-manifolds," de Gruyter Studies in Mathematics, **18**Walter de Gruyter & Co., Berlin.

[46] Turaev, V. G., Viro, O. Ya., [1992] *State sum invariants of* 3*-manifolds and quantum* $6j$*-symbols,* Topology **31**, 865–902.

[47] Witten, E., [1989] *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. **121**, 351–399.

[48] Witten, E., [1988] *Topological quantum field theory,* Comm. Math. Phys. **117**, 353–386.

[49] Yetter, D.N., [1994] *State-sum invariants of* 3*-manifolds associated to Artinian semisimple tortile categories.,* Topology Appl. **58**, 47–80.

J. Scott Carter
Department of Mathematics and Statistics
University of South Alabama
MOBILE, AL 36688

G. Sica (ed.) *What is Category Theory?*

Introducing Categories to the Practicing Physicist

Bob Coecke

Abstract. It is our aim to convince the physicist, and more specific the quantum physicist and/or informatician, that *category theory* should become a part of their daily practice. The reason for this is not that category theory is a better way of doing mathematics, but that *monoidal categories* constitute the actual *algebra of practicing physics.* We will not provide rigorous definitions or anything resembling a coherent mathematical theory, but we will take the reader for a journey introducing concepts which are part of category theory in a manner that the physicist will recognize them.

1 Why?

Why would a physicist care about category theory, why would he want to know about it, why would he want to show off with it? There could be many reasons. For example, you might find John Baez's webside one of the coolest in the world. Or you might be fascinated by Chris Isham's and Lee Smolin's ideas on the use of topos theory in Quantum Gravity. Also the connections between knot theory, braided categories, and sophisticated mathematical physics such as quantum groups and topological quantum field theory might lure you. Or, if you are also into pure mathematics, you might just appreciate category theory due to its incredible unifying power of mathematical structures and constructions. But there is a far more on-the-nose reason which is never mentioned. Namely,

a category is the exact mathematical structure of practicing physics!

What do I mean here by a practicing physics? Consider a physical system of type A (e.g. a qubit, or two qubits, or an electron, or classical measurement data) and perform an operation f on it (e.g. perform a measurement on it) which results in a system possibly of a different type B (e.g. the system together with classical data which encodes the measurement outcome, or, just classical data in the case that the measurement destroyed the system). So typically we have

$$A \xrightarrow{\;f\;} B$$

where A is the initial type of the system, B is the resulting type, and f is the operation. One can perform an operation

$$B \xrightarrow{\;g\;} C$$

after f since the resulting type B of f is also the initial type of g, and we write $g \circ f$ for the consecutive application of these two operations. Clearly we have $(h \circ g) \circ f = h \circ (g \circ f)$ since putting the brackets merely adds the superficial data of conceiving two operations as one. If we further set

$$A \xrightarrow{\;1_A\;} A$$

for the operation 'doing nothing on a system of type A' we have

$$1_B \circ f = f \circ 1_A = f\,.$$

Hence we have a *category*! (a concept introduced by Samuel Eilenberg and Saunders Mac Lane in 1945 in [15]) When we also want to be able to conceive two systems A and B as one whole which we will denote by$A \otimes B$, and hence also need to consider the compound operations

$$A \otimes B \xrightarrow{\;f \otimes g\;} C \otimes D$$

inherited from the operations on the individual systems, then we pass from ordinary categories to a particular case of the 2-dimensional variant of categories called *monoidal categories*. (a concept introduced by Jean Benabou in 1963 in [8]) We will define these monoidal categories in Section 5.

2 What?

The (almost) formally precise definition of a category is the following:

G. Sica (ed.) *What is Category Theory?*

Definition. A *category* $\mathbf{C}$ consists of:

- *objects* $A, B, C, \ldots,$
- *morphisms* $f, g, h, \ldots \in \mathbf{C}(A, B)$ for each pair A, B ,
- *composition* of each $f \in \mathbf{C}(A, B)$ with each $g \in \mathbf{C}(B, C)$ resulting in $g \circ f \in \mathbf{C}(A, C)$ and this composition is such that

$$(h \circ g) \circ f = h \circ (g \circ f) ,$$

- *identity morphisms* $1_A \in \mathbf{C}(A, A)$ for all A which satisfy

$$f \circ 1_A = 1_B \circ f = f .$$

For the same *operational* reasons as discussed above (and which extend to the far more compelling case of monoidal categories as we shall see below), category theory could be expected to play an important role in other fields where operations/processes play a central role e.g. Programing (programs as morphisms) and Logic & Proof Theory (proofs as morphisms), and indeed, in the theoretical counterparts to these fields category theory has become quite common practice cf. the many available textbooks and even undergraduate courses [1].

LOGIC & PROOF THEORY	PROGRAMMING	PHYSICS
Propositions	Data Types	Physical System
Proofs	Programs	Physical Operation

Unfortunately, the standard existing literature on category theory (e.g. [24]) might not be suitable for the audience we want to address in this draft. Category theory literature typically addresses the (broadminded & modern) pure mathematician and as a consequence the presentations are tailored towards them. The typical examples are various categories of mathematical structures and the main focus is on their similarities in terms of mathematical practice. This tendency started with the paper which marked the official birth of category theory [15] in which Samuel Eilenberg and Saunders Mac Lane observe that the collection of mathematical objects of some given kind/type, when equipped with the maps between them, deserves to be studied in its own right as a mathematical structure since this study entails unification of constructions arising from different mathematical fields such as geometry, algebra, topology, algebraic topology etc.

But sometimes going into the area of pure mathematics can be useful exactly to avoid doing to much mathematics. Indeed, an amazing thing of the particular kind of category theory that we need here is that it *formally justifies its own formal absence*, in the sense that at an highly abstract level you can prove that proofs of equational statements in the abstract algebra are equivalent to merely

G. Sica (ed.) *What is Category Theory?*

drawing and manipulating some intuitive pictures [18, 28]. Look for example to how quite sophisticated quantum mechanical calculations can be simplified thanks to category theory in *Kindergarten Quantum Mechanics* [13].

3 Where?

They truly are everywhere! But that's exactly where people start to get confused. (if you are not up for a storm of data just skip this section and go to the next one) We consider some examples from mathematics. A group G is a category with a single object in which every morphism is an isomorphism:

Definition. A morphism $f : A \to B$ is an *isomorphism* (*iso*) if it has an inverse i.e. there exists $f^{-1} : B \to A$ such that

$$f^{-1} \circ f = 1_A \qquad \text{and} \qquad f \circ f^{-1} = 1_B \, .$$

A 'group without inverses' is called a *monoid* and is by definition a category in which there is only one object. Also each partially ordered set P is a category with the elements of this poset as objects, and whenever $a \leq b$ we take $P(a, b)$ to be a singleton, otherwise we take it to be empty. Closedness under composition is guaranteed by transitivity and the identities are provided by reflexivity. Hence a poset is an example of a category with only few morphisms. A preordered set (i.e. 'partial order without anti-symmetry') can be defined as a category in which there is at most one arrow from an object to another one. Still in category theoretic terms, a poset is bounded if it has a terminal and an initial object:

Definition. An object $\top$ is *terminal* if $\mathbf{C}(A, \top)$ is a singleton for all A. An object $\bot$ is *initial* if $\mathbf{C}(\bot, A)$ is a singleton for all A.

It is lattice if it has *products* and *coproducts*, categorical concepts which we will define further below. But on the other hand, we also have the category **Group** which has groups as objects and group homomorphisms as morphisms, and we can also consider the category **Poset** which has posets as objects and order-preserving maps as morphisms. This are two examples of categories with mathematical structure of some kind as objects, and corresponding structure preserving maps as morphisms. Other examples of this sort are topological spaces and continuous maps (**Top**), vector spaces over $\mathbb{K}$ and linear maps ($\mathbf{Vec}_{\mathbb{K}}$), categories and categorical-structure-preserving maps called *functors* (**Cat**), etc.

G. Sica (ed.) *What is Category Theory?*

4 Quantum?

We can also consider two distinct categories which both have sets as objects, but one with functions as morphisms denoted by **Set** and one with relations as morphisms denoted by **Rel**. While you might think that since both have sets as objects they are quite similar, nothing is less true! As a matter of fact, **Rel** much more resembles the category of finite dimensional Hilbert spaces and linear maps **FdHilb** than it resembles **Set**, and here things really start to get interesting. For example, category theory is able to detect the fact that both vector spaces and relations admit a matrix calculus, respectively over the field $\mathbb{K}$ and over the semiring of booleans $\mathbb{B}$.[1] While technically this involves some more sophisticated concepts, we are already able to show that both **Rel** and **FdHilb** admit a notion of *superposition* while **Set** doesn't. We expose this through the categorical notion of *element* i.e. a notion of element which exposes itself at the level of morphisms. First note that for any set X we have a bijection, i.e., categorically, an isomorphism

$$X \times \{*\} \simeq X\,,$$

where $\{*\}$ is just some singleton set, so we can expect $\{*\}$ to play a special role both in **Set** and **Rel**. Similarly, in finite dimensional Hilbert spaces we have

$$\mathcal{H} \otimes \mathbb{C} \simeq \mathcal{H}\,,$$

so we expect the one-dimensional Hilbert space $\mathbb{C}$ to play a special role in **FdHilb**. And indeed, in **Set** we can define X's elements as the functions

$$f_x : \{*\} \to X :: * \mapsto x$$

since in this way each element $x \in X$ arises as $f(*)$ for the function $f_x : * \mapsto x$. Analogously in **FdHilb** we define $\mathcal{H}'s$ elements as linear maps

$$f_{|\psi\rangle} : \mathbb{C} \to \mathcal{H} :: 1 \mapsto |\psi\rangle$$

since by linearity $f_{|\psi\rangle}(1) = |\psi\rangle$ determines the linear map $f_{|\psi\rangle}$ completely. By analogy in **Rel** X's elements are relations

$$\{*\} \xrightarrow{R} X\,,$$

but since relations are 'multi-valued' this means that the elements do not correspond with the elements of X but with the subsets $Y \subseteq X$, and one can think of these subsets as *superpositions* of the singletons. Indeed, setting

$$Y_i := X \text{ iff } i \in Y \qquad \text{and} \qquad Y_i := \emptyset \text{ iff } i \notin Y\,,$$

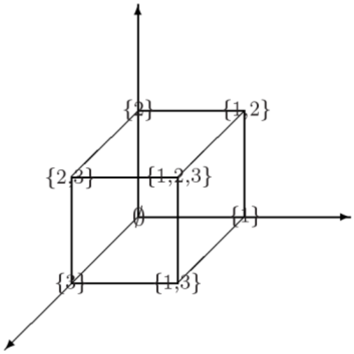
[1] A semiring is a ring in 'without additive inverses'. For a matrix calculus it indeed suffices to be able to add and to multiply scalars, while no substraction is needed.

G. Sica (ed.) *What is Category Theory?*

both in **FdHilb** and **Rel** we can decompose elements over some notion of bases respectively as

$$|\psi\rangle = \sum_{i\in X} \psi_i \cdot |\, i\,\rangle \qquad\qquad Y = \bigcup_{i\in X} Y_i \cap \{i\}\,.$$

Hence the sum becomes a union and the $\mathbb{C}$-valued coefficients become Boolean-valued since $\{\emptyset, X\} \simeq \mathbb{B}$, the Booleans. In other words, we can think of the subsets of a set, i.e. the elements in **Rel**, as being embedded in some vector space:



Very crucial in all this is the fact that we considered the cartesian product $\times$ in **Rel** and the tensor product $\otimes$ in **FdHilb**, while both categories allow to combine their objects in many different other ways (e.g. the direct sum of Hilbert spaces). This shows that it is essential to consider these additional operations as a genuine part of the structure, introducing *monoidal* structure.

5 Which?

The key feature we have seen so far of a category are:

- The structure lives in the space of operations (vs. state space),
- Types enable to distinguish different kinds of systems,
- Composition/application is the primitive ingredient.

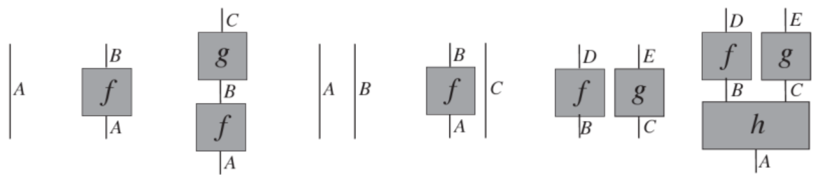
We are still missing something crucial. While not officially part of the basic definition of a category, for any 'operational' situation as discussed in Section 1 it is natural to have, besides (temporal) sequential composition, some notion of parallel composition which allows one to consider two distinct entities as one whole (e.g. the tensor product in quantum mechanics). In abstract category-theoretic terms this means introducing a second dimension.

G. Sica (ed.) *What is Category Theory?*

> **Definition.** A *symmetric monoidal category* is a category with a *symmetric monoidal tensor*, that is, an assignment both for pairs of objects and pairs morphisms
>
> $$\begin{array}{ccc}(A, B) & \mapsto & A \otimes B \\ (A \xrightarrow{f} B\,, C \xrightarrow{g} D) & \mapsto & A \otimes C \xrightarrow{f \otimes g} B \otimes D\end{array}$$
>
> which is *bifunctorioral*, and comes together with *left & right unit natural isos*, a *symmetry natural iso* and an *associativity natural iso*.

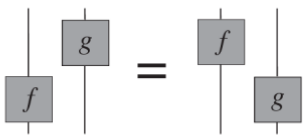
So it remains to explain what bifunctoriality and those natural isos stand for.

To this means we depict morphisms (i.e. physical processes) as square boxes, and we label the inputs and outputs of these boxes by *types* which tell on which kind of system these boxes act cf. one qubit, n-qubits, classical data etc. Sequential composition (in time) is depicted by connecting matching outputs and inputs of these boxes by lines, and parallel composition (cf. tensor) by locating boxes side by side. E.g. $1_A : A \to A$, $f : A \to B$, $g \circ f$ for $g : B \to C$, $1_A \otimes 1_B : A \otimes B \to A \otimes B$, $f \otimes 1_C$, $f \otimes g$ for $f : B \to D$ and $g : C \to E$, and $(f \otimes g) \circ h$ for $h : A \to B \otimes C$ respectively depict as:



We now show that the requirements 'bifunctoriality' and 'existence and naturality for some special isomorphisms' with respect to the operation 'combining systems' are physically so evidently true that they almost seem redundant. (but as we will see further they do have major implications)

Bifunctoriality. In the graphical language *bifunctoriality* stands for:



Bifunctoriality has a very clear conceptual interpretation: If we apply an operation f to one system and an operation g to another system, then the order in which we apply them doesn't matter. Hence bifunctoriality expresses some notion of *locality* but still allows for the quantum type of non-locality. The above

G. Sica (ed.) *What is Category Theory?*

pictorial equation can also be written down in term of a *commutative diagram*:

$$
\begin{array}{ccc}
A_1 \otimes A_2 & \xrightarrow{f \otimes 1_{A_2}} & B_1 \otimes A_2 \\
{\scriptstyle 1_{A_1} \otimes g}\Big\downarrow & & \Big\downarrow{\scriptstyle 1_{B_1} \otimes g} \\
A_1 \otimes B_2 & \xrightarrow[f \otimes 1_{B_2}]{} & B_1 \otimes B_2
\end{array}
$$

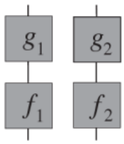
which expresses that both *paths* yield the same result. Actually, taking on a relativistic spirit, $(1 \otimes g) \circ (f \otimes 1) = (f \otimes 1) \circ (1 \otimes g)$ expresses that what is at the left and at the right of the tensor does not temporally compare (cf. are space-like separated) so we can denote them both without any harm by $f \otimes g$, and hence assume the slightly more general condition

$$(g_1 \otimes g_2) \circ (f_1 \otimes f_2) = (g_1 \circ f_1) \otimes (g_2 \circ f_2)$$

from which it easily follows that

$$(1 \otimes g) \circ (f \otimes 1) = (1 \circ f) \otimes (g \circ 1) = (f \circ 1) \otimes (1 \circ g) = (f \otimes 1) \circ (1 \otimes g)\,.$$

This stronger condition was already implicitly present in the picture calculus since the latter explicitly ignores the brackets:

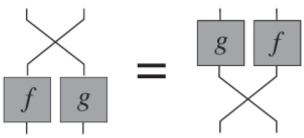


i.e. it doesn't matter if we either first consider the sequential composition or the parallel composition. We read this as: since f_1 is causally before g_1 and f_2 is causally before g_2, the pair (f_1, f_2) is causally before (g_1, g_2) and vice versa, but we do not assume any a priori space-like correlations 'along the tensor'. Finally in addition to the above we also require

$$1_A \otimes 1_B = 1_{A \otimes B}$$

for the tensor, which is again self-evident from an operational perspective.

Symmetry and associativity natural isomorphisms. One can think of *natural isomorphisms* as 'explicitly witnessed' canonical isomorphisms. This is best seen through an example. Consider the following picture:



G. Sica (ed.) *What is Category Theory?*

which, in operational terms, expresses that if we swap the location of two systems then we also have to swap the operations we intend to apply on them in order to get the same result. Diagrammatically it corresponds to commutation of:

$$\begin{array}{ccc} A_1 \otimes A_2 & \xrightarrow{f \otimes g} & B_1 \otimes B_2 \\ {\scriptstyle \sigma_{A_1,A_2}} \downarrow & & \downarrow {\scriptstyle \sigma_{B_1,B_2}} \\ A_2 \otimes A_1 & \xrightarrow[g \otimes f]{} & B_2 \otimes B_1 \end{array}$$

and we call the family of isomorphisms $\{\sigma_{A,B} : A \otimes B \to B \otimes A\}$ which stands for 'swapping the systems' a natural isomorphism. Hence this idea of the existence of morphisms witnessing the fact that two objects are isomorphic is again highly operational. Given two expressions $\Lambda(-, \ldots, -)$ and $\Xi(-, \ldots, -)$ using the bifunctor $(- \otimes -)$, a (restricted[2]) formal notion of natural isomorphism generalizes in terms of the existence of a family

$$\left\{\xi_{A_1,\ldots,A_n} : \Lambda(A_1, \ldots, A_n) \to \Xi(A_1, \ldots, A_n)\right\}$$

for which we have commutation of:

$$\begin{array}{ccc} \Lambda(A_1, \ldots, A_n) & \xrightarrow{\Lambda(f_1, \ldots, f_n)} & \Lambda(B_1, \ldots, B_n) \\ {\scriptstyle \xi_{A_1,\ldots,A_n}} \downarrow & & \downarrow {\scriptstyle \xi_{B_1,\ldots,B_n}} \\ \Xi(A_1, \ldots, A_n) & \xrightarrow[\Xi(f_1, \ldots, f_n)]{} & \Xi(B_1, \ldots, B_n) \end{array} \tag{1}$$

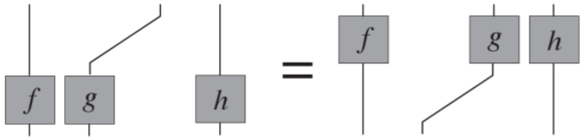
Analogously to 'swapping', we can consider a notion of associating systems to each other e.g. being in the possession of the same agent or being located 'not to far from each other'. The corresponding natural isomorphism which re-associates systems should obviously satisfy:

$$\begin{array}{ccc} (A_1 \otimes A_2) \otimes A_3 & \xrightarrow{(f \otimes g) \otimes h} & (B_1 \otimes B_2) \otimes B_3 \\ {\scriptstyle \alpha_{A_1,A_2,A_3}} \downarrow & & \downarrow {\scriptstyle \alpha_{B_1,B_2,B_3}} \\ A_1 \otimes (A_2 \otimes A_3) & \xrightarrow[f \otimes (g \otimes h)]{} & B_1 \otimes (B_2 \otimes B_3) \end{array}$$

[2] We will present a much more general notion of natural isomorphism/transformation below ones we have the general notion of morphism of categories at our disposal.

G. Sica (ed.) *What is Category Theory?*

that is, in a picture,

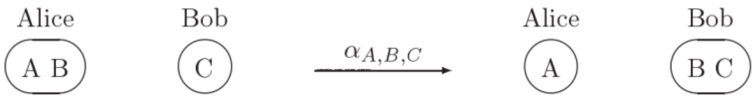


When abandoning the spatial interpretation of associativity, naturality is still implicitly present in the pictures due to the implicit absence of brackets in:

i.e. it makes no difference if we either want to conceive the first two systems or the last two systems as one whole. One can of course always choose to have

$$(A_1 \otimes A_2) \otimes A_3 = A_1 \otimes (A_2 \otimes A_3) \qquad \text{with} \qquad \alpha_{A_1,A_2,A_3} := 1_{A_1 \otimes A_2 \otimes A_3}$$

but in many cases it is very useful to have a non-trivial witness. An example of this is the analysis of quantum teleportation in [2] were it stands for Alice sending a qubit to Bob in the teleportation protocol i.e. 'association' stands for 'spatial colocation':



Unit object and unit natural isomorphisms. Physical operations can destroy a system e.g. measurement of the position of a photon. On the other hand, one can conceive a preparation procedure as the creation of a system from an unspecified source. Therefore it is useful to have an object standing for *no system*, preparation or state then being of the type $\mathrm{I} \to A$ and destruction being of the type $A \to \mathrm{I}$ — in Dirac's notation [14] these respectively are the so-called *kets* and *bras*. Clearly, since I stands for 'no system' we have

$$A \otimes \mathrm{I} \simeq A \simeq \mathrm{I} \otimes A$$

and these *left & right unit natural isomorphisms* obviously should satisfy:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle \lambda_A}\downarrow & & \downarrow{\scriptstyle \lambda_B} \\ \mathrm{I} \otimes A & \xrightarrow[1_{\mathrm{I}} \otimes f]{} & \mathrm{I} \otimes B \end{array} \qquad \begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle \rho_A}\downarrow & & \downarrow{\scriptstyle \rho_B} \\ A \otimes \mathrm{I} & \xrightarrow[f \otimes 1_{\mathrm{I}}]{} & B \otimes \mathrm{I} \end{array} \qquad (2)$$

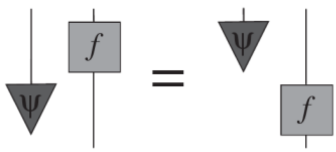
G. Sica (ed.) *What is Category Theory?*

i.e. introducing nothing should not alter the effect of an operation. In other words, the left & right unit natural isomorphisms allow us to introduce or discard such an extra object at any time. Such an object also comes with a notion of *scalar* i.e. a morphism of type $s : \mathrm{I} \to \mathrm{I}$. In particular do these scalars arise when post-composing a state with a costate i.e. when we have a *bra-ket*

$$\pi \circ \psi : \mathrm{I} \xrightarrow{\psi} A \xrightarrow{\pi} \mathrm{I}\,.$$

As we will see below in Section 6, having such a 'no system'-object has much more striking consequences than one would expect at first. We also introduce a graphical symbol for *states* or *elements* $\psi : \mathrm{I} \to A$ (which are now formally defined in the presence of a symmetric monoidal tensor), for *costates* $\pi : A \to \mathrm{I}$, and for *scalars* $s : \mathrm{I} \to \mathrm{I}$, of which $\pi \circ \psi$ is an example:

The above naturality diagram now boils down to:

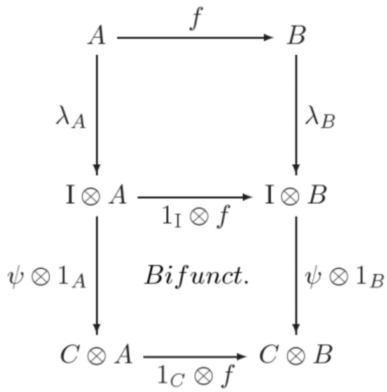


which rewrites as a diagram as:

$$\begin{array}{ccc} A & \xrightarrow{\quad f \quad} & B \\ {\scriptstyle (\psi \otimes 1_A) \circ \lambda_A}\Big\downarrow & & \Big\downarrow{\scriptstyle (\psi \otimes 1_B) \circ \lambda_B} \\ C \otimes A & \xrightarrow[1_C \otimes f]{} & C \otimes B \end{array}$$

G. Sica (ed.) *What is Category Theory?*

and is obtained by *pasting* diagram (2) with bifunctoriality:



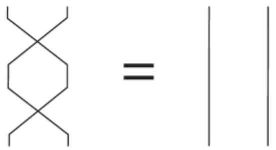
Typical examples of symmetric monoidal categories are $(\mathbf{Set}, \times)$ and $(\mathbf{Rel}, \times)$ with $\{*\}$ as unit object and $(\mathbf{FdHilb}, \otimes)$ with $\mathbb{C}$ as unit object — which we already implicitly referred to when discussing the similarities between their respective elements. But there is for example also $(\mathbf{FdHilb}, \oplus)$ with the 0-dimensional vector space as unit object and $(\mathbf{Set}, +)$ and $(\mathbf{Rel}, +)$ (where $+$ is the 'disjoint union') with the empty set as unit object. Again $(\mathbf{FdHilb}, \oplus)$ and $(\mathbf{Rel}, +)$ are very similar categorically, but still quite different from $(\mathbf{Set}, +)$.

Bases independency. For the particular case of vector spaces over some field $\mathbb{K}$, setting $A_i = B_i := V_i$ and taking $f, g, \ldots$ to be a change of bases for the corresponding vector space, the general naturality diagram (1) exactly expresses *base independency*. Hence in the context of vector spaces *natural concepts* are always *bases independent concepts*.

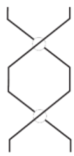
Coherence. We want the different natural isomorphisms introduced above to coexist peacefully and for that reason we need to require some *coherence* conditions e.g. $\sigma_{\mathrm{I},A} \circ \lambda_A = \rho_A$ and $\lambda_{\mathrm{I}} = \rho_{\mathrm{I}}$. We will not spell them out explicitly here. The general theory of coherence in categories is highly non-trivial as a branch of developing category theory (as opposed to using category theory). The reason we mention these coherence conditions here is that the axiomatic algebra of *categorical quantumness* (see Section 11), somewhat surprisingly, first appeared in the context of coherence theory [20, 21].

Braided categories. One coherence condition for a symmetric monoidal tensor is $\sigma^{-1}_{A,B} = \sigma_{B,A}$ i.e. $\sigma_{B,A} \circ \sigma_{A,B} = 1_{A\otimes B}$, which depicts as:

G. Sica (ed.) *What is Category Theory?*



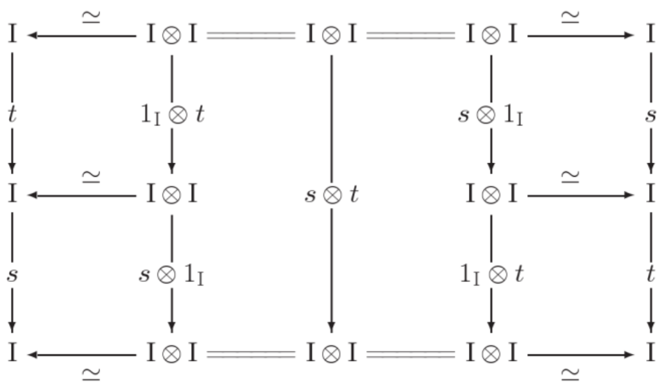
In *braided* monoidal categories this is not true anymore, giving rise to braided structure for $\sigma_{B,A} \circ \sigma_{A,B}$:



We refer to the web pages by John Baez and the books by Louis Kauffman for prose on this body of mathematics research.

6 How much?

So far nothing *quantitative* seems to have been going on here. Not true! Given a category $\mathbf{C}$ we will call $\Sigma_A := \mathbf{C}(\mathrm{I}, A)$ the *state space* of system A and $\mathbb{S} := \mathbf{C}(\mathrm{I}, A)$ the *scalar monoid*. The scalar monoid in $(\mathbf{FdHilb}, \otimes)$ is isomorphic to $\mathbb{C}$ since any linear map $s : \mathbb{C} \to \mathbb{C}$ is by linearity completely determined by the image of $1 \in \mathbb{C}$. Those in $(\mathbf{Rel}, \times)$ are the Booleans, since there are two relations from a singleton to itself, the identity and the empty relation. A remarkable result is that the scalar monoid is always commutative [21] — the big diagram below is indeed a proof, which uses bifunctoriality, left & right unit naturality and $\lambda_\mathrm{I} = \rho_\mathrm{I}$:



This is quite a surprising result. From the very evident operationally motivated assumptions on compoundness we obtain something as strong as a requirement of commutation. This for example implies that if we would want to vary quantum

G. Sica (ed.) *What is Category Theory?*

theory by changing the underlying field of the vector space we need to stay commutative, excluding *quaternionic quantum mechanics* [16]. But there is much more. The left-hand-side of the above diagram expresses

$$s \circ t = \mathrm{I} \xrightarrow{\simeq} \mathrm{I} \otimes \mathrm{I} \xrightarrow{s \otimes t} \mathrm{I} \otimes \mathrm{I} \xrightarrow{\simeq} \mathrm{I}\,.$$

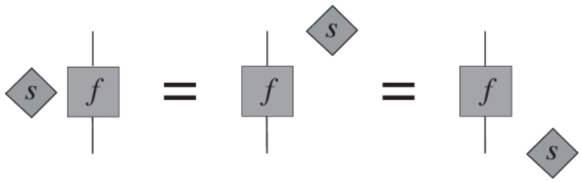
We generalize this and define *scalar multiplication* as

$$s \bullet f := A \xrightarrow{\simeq} A \otimes \mathrm{I} \xrightarrow{f \otimes s} B \otimes \mathrm{I} \xrightarrow{\simeq} B$$

given a scalar s and any morphism f. We think of $s\bullet -$ as being a (probablilistic) weight which is attributed to the operation f. One can prove that (e.g. [12])

$$(s \bullet f) \circ (t \bullet g) = (s \circ t) \bullet (f \circ g) \quad \text{and} \quad (s \bullet f) \otimes (t \bullet g) = (s \circ t) \bullet (f \otimes g)$$

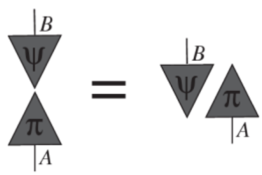
i.e. diamonds can move around freely in 'time' and 'space':



One can also show that states and costates satisfy a similar property (e.g. [12])

$$\psi \circ \pi = A \xrightarrow{\simeq} \mathrm{I} \otimes A \xrightarrow{\psi \otimes \pi} B \otimes \mathrm{I} \xrightarrow{\simeq} B$$

what results in:



Conclusively, at the very basic level of monoidal categories we get a *quantitative notion* of value for free, encoded as scalars (provided the scalar monoid itself is non-trivial), and which arises when a state meets a costate, that is, in Dirac's terminology, *when a ket meets a bra.*

7 Key categorical concepts

The above introduced notions of bifunctoriality and natural iso are instances of the key categorical concepts called functor and natural transformation. In

G. Sica (ed.) *What is Category Theory?*

Eilenberg-Mac Lane functors were introduced as *morphisms* between categories while natural transformations were introduced as morphisms between functors.

> **Definition.** A *functor* $F : \mathbf{C} \to \mathbf{D}$ is a 'structure preserving map of categories' i.e. it maps an object A to an object FA, and a morphism $A \xrightarrow{f} B$ to a morphism $FA \xrightarrow{Ff} FB$, and satisfies
>
> $$F(g \circ f) = Fg \circ Ff \qquad \text{and} \qquad F(1_A) = 1_{FA}\,.$$

Given a category define a new category $\mathbf{C} \times \mathbf{C}$ which has pairs (A, B) as objects, pairs (f, g) as morphisms, pairs $(1_A, 1_B)$ as identities and with composition pairwise defined. Hence a functor $F : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ satisfies

$$F(g_1 \circ f_1, g_2 \circ f_2) = F(g_1 \otimes g_2) \circ (f_1 \otimes f_2)\,.$$

Setting $F(-,-) := - \otimes -$ it follows that a *tensor* is indeed a functor, by bifunctoriality. Another example is a group homomorphism which turns out to be a functor of groups since functoriality implies preservation of inverses:

$$\begin{array}{rcl} a^{-1} \cdot a = e = a \cdot a^{-1} & \Rightarrow & F(a^{-1} \cdot a) = Fe = F(a \cdot a^{-1}) \\ & \Rightarrow & F(a^{-1}) \cdot Fa = e = Fa \cdot F(a^{-1}) \\ & \Rightarrow & (Fa)^{-1} = F(a^{-1})\,. \end{array}$$

This is the case because an inverse is a *categorical property*.

> **Definition.** Given two functors $F, G : \mathbf{C} \to \mathbf{D}$ a *natural transformation* $\xi : F \Rightarrow G$ is a family $\{\xi_A : FA \to GA\}_A$ of morphisms in $\mathbf{D}$ such that for all morphisms $f : A \to B$ in $\mathbf{C}$ we have commutation of
>
> $$\begin{array}{ccc} FA & \xrightarrow{Ff} & FB \\ {\scriptstyle \xi_A}\downarrow & & \downarrow{\scriptstyle \xi_B} \\ GA & \xrightarrow[Gf]{} & GB \end{array}$$

The symmetry isomorphism is indeed a special case of this definition for

$$F : \mathbf{C} \times \mathbf{C} \to \mathbf{C} :: \left\{ \begin{array}{ccc} (A, B) & \mapsto & A \otimes B \\ (f, g) & \mapsto & f \otimes g \end{array} \right.$$

$$G : \mathbf{C} \times \mathbf{C} \to \mathbf{C} :: \left\{ \begin{array}{ccc} (A, B) & \mapsto & B \otimes A \\ (f, g) & \mapsto & g \otimes f \end{array} \right.$$

G. Sica (ed.) *What is Category Theory?*

While this general definition might be non-intuitive, there are some conceptually highly significant examples of it. A *natural diagonal* expresses the process of *copying*. It consists of the family $\{\Delta_A : A \to A \otimes A\}_A$ which again for operational reasons obviously has to satisfy

$$
\begin{array}{ccc}
A & \xrightarrow{\ f\ } & B \\
{\scriptstyle \Delta_A}\big\downarrow & & \big\downarrow{\scriptstyle \Delta_B} \\
A \otimes A & \xrightarrow[f \otimes f]{} & B \otimes B
\end{array}
$$

As a consequence, due to the no-cloning theorem for quantum mechanics [30] we can expect that in **FdHilb** we cannot have a natural diagonal. We can define a map $\mathcal{H} \to \mathcal{H} \otimes \mathcal{H} :: |i\rangle \mapsto |i\rangle \otimes |i\rangle$, but since this map depends on the choice of bases, it cannot be natural. Explicitly, the following diagram *does not commute*:

$$
\begin{array}{ccc}
\mathbb{C} & \xrightarrow{\ 1 \mapsto |0\rangle + |1\rangle\ } & \mathbb{C} \oplus \mathbb{C} \\
{\scriptstyle 1 \mapsto 1 \otimes 1}\big\downarrow & & \big\downarrow{\scriptstyle \begin{array}{l} |0\rangle \mapsto |0\rangle \otimes |0\rangle \\ |1\rangle \mapsto |1\rangle \otimes |1\rangle \end{array}} \\
\mathbb{C} \simeq \mathbb{C} \otimes \mathbb{C} & \xrightarrow[1 \otimes 1 \mapsto (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)]{} & (\mathbb{C} \oplus \mathbb{C}) \otimes (\mathbb{C} \oplus \mathbb{C})
\end{array}
$$

since via one path we obtain the *Bell-state*

$$1 \mapsto |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle$$

while via the other path we obtain a *disentangled state*

$$1 \mapsto (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)\,.$$

Exactly the same phenomenon happens in **Rel**. Recall that a relation between two sets X and Y is a subset $R \subseteq X \times Y$ consisting of the pairs which satisfy the relation. Hence the diagonal function

$$X \to X \times X :: x \mapsto (x, x)$$

can be written as a relation as

$$\{(x, (x, x)) \mid x \in X\}\,.$$

G. Sica (ed.) *What is Category Theory?*

But this relation is not natural since we have non-commutation of:

$$\begin{array}{ccc}
\{*\} & \xrightarrow{\{(*,0),(*,1)\}} & \{0,1\} \\
{\scriptstyle \{(*,(*,*))\}}\big\downarrow & & \big\downarrow{\scriptstyle \{(0,(0,0)),(1,(1,1))\}} \\
\{(*,*)\} & \xrightarrow[\{(*,0),(*,1)\}\times\{(*,0),(*,1)\}]{} & \{0,1\}\times\{0,1\}
\end{array}$$

since via one path we have

$$\{(*,(0,0)),(*,(1,1))\}$$

while the other path yields

$$\{(*,(0,0)),(*,(0,1)),(*,(1,0)),(*,(1,1))\}\,.$$

On the other hand, this example does not carry over to **Set** since we use relations which are properly multi-valued. In fact, in **Set** we do have a natural diagonal:

$$\begin{array}{ccc}
X & \xrightarrow{x\mapsto f(x)} & Y \\
{\scriptstyle x\mapsto (x,x)}\big\downarrow & & \big\downarrow{\scriptstyle f(x)\mapsto (f(x),f(x))} \\
X\times X & \xrightarrow[(x,x)\mapsto (f(x),f(x))]{} & Y\times Y
\end{array}$$

and this is a consequence of the *high-level* fact that in **Set** the cartesian product is a *true product* in the categorical sense.

Definition. A *product* of two objects A and B is a triple consisting of an object and a pair of operations called *projections*

$$(A\sqcap B\ ,\ \ p_1 : A\sqcap B\to A\ ,\ \ p_2 : A\sqcap B\to B)$$

which are such that for every pair operations $f : C\to A$ and $g : C\to B$ there exists a unique operation h such that we have commutation of:

$$\begin{array}{ccccc}
 & & C & & \\
 & {\scriptstyle f}\swarrow & \big\downarrow{\scriptstyle h} & \searrow{\scriptstyle g} & \\
A & \xleftarrow[p_1]{} & A\sqcap B & \xrightarrow[p_2]{} & B
\end{array}$$

G. Sica (ed.) *What is Category Theory?*

The uniqueness of $h : C \to A \sqcap B$ is usually referred to as the *universal property* of the product. But we can reformulate this definition in a manner which gives it a more direct *operational significance* in terms of *pairing* and *unpairing* meta-operations.

Definition. A *product* of two objects A and B is a triple consisting of an object and a pair of operations called *projections*

$$(A \sqcap B\ ,\ \ p_1 : A \sqcap B \to A\ ,\ \ p_2 : A \sqcap B \to B)$$

together with *pairing* and *unpairing* operations

$$[-,-] : \mathbf{C}(C, A) \times \mathbf{C}(C, B) \to \mathbf{C}(C, A \sqcap B)$$

$$p_1 \circ - : \mathbf{C}(C, A \sqcap B) \to \mathbf{C}(C, A) \qquad p_2 \circ - : \mathbf{C}(C, A \sqcap B) \to \mathbf{C}(C, B)$$

which are such that

$$[p_1 \circ h, p_1 \circ h] = h \qquad\qquad p_1 \circ [f, g] = f \qquad\qquad p_2 \circ [f, g] = g\ .$$

The three required equalities essentially say that pairing and unpairing are each other inverse as meta-operations i.e. they allow each operation of type $C \to A \times B$ to be transformed in a pair of operations of respective types $C \to A$ and $C \to B$ and vice versa. If one has such a product structure than one always has a natural diagonal and provide a notion of *copying*. Moreover, also the projections are natural and can be interpreted as a natural notion of *deleting*. (cf. the no-deleting theorem in quantum mechanics [25])

G. Sica (ed.) *What is Category Theory?*

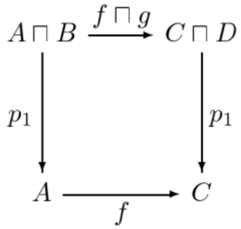
Proposition. Products yield a *monoidal tensor*

$$f \sqcap g := [f \circ p_1, g \circ p_2] : A \sqcap B \to C \sqcap D$$

for $f : A \to C$ and $g : B \to D$ and a *diagonal*

$$\Delta_A := [1_A, 1_A] : A \to A \sqcap A\,.$$

Moreover, *projections are natural* i.e. we have commutation of:



Specifying the idea of pairing and unpairing for states we have that the information encoded in any bipartite state

$$\Psi : \mathrm{I} \to A \otimes B$$

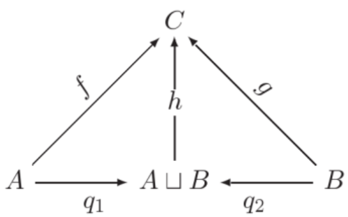
can be equivalently encoded in the pair

$$\psi_1 = p_1 \circ \Psi : \mathrm{I} \to A \qquad \text{and} \qquad \psi_2 = p_2 \circ \Psi : \mathrm{I} \to B$$

which immediately excludes the possibility of *entanglement*. Hence, no-cloning is not a surprise at all in the presence of anything which even remotely behaves like entanglement. But pairing and unpairing are not the only meaningful meta-operations of their kind since there exist also the notions of *copairing* and *co-unpairing*, since there is indeed a dual notion to *product* named *coproduct* which is obtained by reversing all the arrows involved. A *coproduct* of two objects A and B is a triple consisting of an object and a pair of operations called *injections*

$$(A \sqcup B\ ,\ \ q_1 : A \to A \sqcup B\ ,\ \ q_2 : B \to A \sqcup B)$$

which are such that for every pair operations $f : A \to C$ and $g : B \to C$ there exists a unique operation h such that we have commutation of:



G. Sica (ed.) *What is Category Theory?*

From coproducts we can define a *codiagonal*

$$\nabla_A = A \sqcup A \to A$$

analogously as we defined a diagonal given products.

Linear logic. Quantum theory is of course not the only theory in which there are no natural notions of copying and deleting. E.g. in spoken language we have:

$$\mathsf{not\ not} \neq \mathsf{not}\,,$$

a fact which was well-known to one of the main builders of category theory Jim Lambek [22]. Both in computing and proof theory, absence of evident availability of copying and deleting captures *resource sensitivity* i.e. it *counts* how many times a resource is used. While much of the technical machinery was already available due to Jim Lambek, Saunders MacLane, Max Kelly and other category theoreticians, the name and conceptual understanding of linear logic has to be attributed to Jean-Yves Girard [17], and the full identification in category theoretic was provided in [27]. For a useful survey on category theory from the linear logician's perspective we refer to [10].

8 Enriched categories

We will not get in detail on this mathematically highly non-trivial subject and refer the reader for an easy-going introduction to [11]. Here we just want to mention the existence of this particular way of adding more structure to categories, since we will encounter a simple example of it below. Consider the so-called Hilbert-Schmidt correspondence for finite dimensional Hilbert spaces i.e. given two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ there is a *natural* isomorphism in **FdHilb**[3]

$$\mathcal{H}_1^* \otimes \mathcal{H}_2 \simeq \mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_2) \tag{3}$$

between the tensor product of $\mathcal{H}_1^*$ (i.e. the *dual* of $\mathcal{H}_1$) and $\mathcal{H}_2$ and Hilbert space of linear maps between $\mathcal{H}_1$ and $\mathcal{H}_2$. In particular do we have that $\mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_2)$ is itself a Hilbert space. Note also that there exists a linear map

$$\mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_2) \otimes \mathbf{FdHilb}(\mathcal{H}_2, \mathcal{H}_3) \to \mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_3) :: (f, g) \mapsto g \circ f$$

due to the *universal property* of the Hilbert space tensor product i.e. for each triple $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ there exists a particular morphism in **FdHilb** which *internalizes*

[3]Surprisingly enough, in much of the quantum mechanical literature (e.g. [5, 31]) one does not encounter this *natural correspondence* but rather an un-natural one namely $\mathcal{H}_1 \otimes \mathcal{H}_2 \simeq \mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_2)$ which is merely a bijection between sets and which is of course is bases dependent. The same is the case for many other notions used in the quantum physics literature. Life could be made so much easier if physicist would learn about the benefits of naturality.

G. Sica (ed.) *What is Category Theory?*

composition of linear maps. Hence we have a situation where the morphism-sets of a category $\mathbf{C}$ are themselves structured as objects in (possibly another) category $(\mathbf{D}, \otimes)$ in such a manner that composition in $\mathbf{C}$, i.e.

$$- \circ - : \mathbf{C}(A, B) \times \mathbf{C}(B, C) \to \mathbf{C}(A, C)\,,$$

internalizes in $(\mathbf{D}, \otimes)$ as an explicit morphism

$$c_{A,B,C} : \mathbf{C}(A, B) \otimes \mathbf{C}(B, C) \to \mathbf{C}(A, C)\,.$$

Such a category $\mathbf{C}$ is called $\mathbf{D}$-*enriched* or simply a $\mathbf{D}$-*category.* Each category is by definition a **Set**-category. A 2-dimensional category or simply, a *2-category* is defined as a **Cat**-category. Similarly, a *3-category* is a **2-Cat**-category, and a $(n + 1)$-*category* is a **n-Cat**-category, a branch of category which currently intensively studied, and in particular strongly advertised by John Baez. A particular fragment **FdHilb**-enrichment (cf. **FdHilb** is itself **FdHilb**-enriched) is enrichment in commutative monoids **CMon** i.e. linear maps can be added.

9 Logical closure

Categorical enrichment is not the only way to encode the Hilbert-Schmidt correspondence. From eq.(3) and $(\mathcal{H}_1 \otimes \mathcal{H}_2)^* \simeq \mathcal{H}_1^* \otimes \mathcal{H}_2^*$ it follows that

$$\mathbf{FdHilb}(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{H}_2) \simeq (\mathcal{H}_1^* \otimes \mathcal{H}_2^*) \otimes \mathcal{H}_3 \simeq \mathcal{H}_1^* \otimes (\mathcal{H}_2^* \otimes \mathcal{H}_3) \simeq \mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_2^* \otimes \mathcal{H}_3)$$

Hence when defining a new connective between Hilbert spaces by setting

$$\mathcal{H}_2 \Rightarrow \mathcal{H}_3 := \mathcal{H}_2^* \otimes \mathcal{H}_3\,,$$

called *implication*, we obtain

$$\mathbf{FdHilb}(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{H}_2) \simeq \mathbf{FdHilb}(\mathcal{H}_1, \mathcal{H}_2 \Rightarrow \mathcal{H}_3)$$

which is a special case of the general situation of *monoidal closure*:

$$\mathbf{C}(A \otimes B, C) \simeq \mathbf{C}(A, B \Rightarrow C)$$

where we now assume (not necesarily being in a self-enriched context) that the isomorphism is natural in **Set**.[4] This is precisely the deductive content of general *categorical logic*, which states that for each proof

- $f^{\vdash} : A \otimes B \to C$ 'which deduces from A *and* B that C is true'

[4] Actually we have an example of a so-called *adjunction* between the two functors $B \otimes -$ and $B \Rightarrow -$ for each object B of the category. While in many category theory books for very compelling mathematical reasons adjunction will be put forward as the most important mathematical concept of the whole of category theory, we unfortunately won't have the space here to develop it, and it would deviate us too much from our story line.

G. Sica (ed.) *What is Category Theory?*

that there is a corresponding proof

- $f_\vdash : A \to B \Rightarrow C$ 'which deduces from A that B *implies* C'

and vice versa. A particular situation of monoidal closure is *cartesian closure* where the monoidal tensor is a categorical product i.e. we have

$$\mathbf{C}(A \sqcap B, C) \simeq \mathbf{C}(A, B \Rightarrow C)$$

— which of course is *not* the case for $(\mathbf{FdHilb}, \otimes)$ since as we have seen above that a product structure prevents the existence of entanglement. The notion of a *topos* is an even more $(\mathbf{Set}, \times)$-resembling particular case of a cartesian closed category, so, although some have proposed topos theory to be used in the foundations of quantum mechanics, one fact is certain: a topos is manifestly *non-quantum*. Putting this in more technical terms, in **FdHilb** we have another (than cartesian closure) particular case of monoidal closure called $*$-*autonomy* [7], which requires that there exists an operation *negation* denoted by a star and which is such that we can derive the implication from it by setting

$$A \Rightarrow B := (A \otimes B^*)^*$$

which logically makes a lot of sense: A *implies* B when we do *not* have that A is true *and* that *not* B is true, that is, by the *De Morgan* law, that A *implies* B when we *either* have that *not* A is true *or* that B is true.

Proposition. A symmetric monoidal category which is both cartesian closed and $*$-autonomous can only be a preordered set.

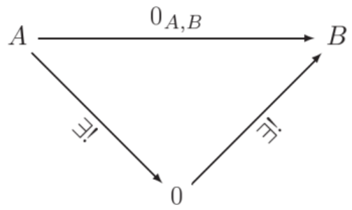
This translates physically in the fact that if a quantum formalism would be cartesian closed then the only operation on a system which preserves it is the identity, which implies that there cannot be any non-trivial notion of unitarity.

But again, **FdHilb** has even more structure than $*$-autonomy, namely the fact that $(A \otimes B)^* \simeq A^* \otimes B^*$, which logically is a bit weird, stating that *not* (A *and* B) is equivalent to (*not* A) *and* (*not* B), hence it follows that *and* is the same as *or*.[5] This kind of logically degenerate monoidal categories in which the tensor is *self-dual* are called *compact closed categories* and where introduced by Max Kelly [20], in terms of a much simpler definition than the above one which we will discuss in the section 'Categorical quantumness'. Surprisingly, they arise in many more contexts than one would expect, for example in linguistics [23], in relativity since cobordism categories turn out to be compact closed [6], in concurrency theory [4], they also enable to formalize the mathematical notion of a knot, and of course, they consitute the key to axiomatizing quantum entanglement [2, 3].

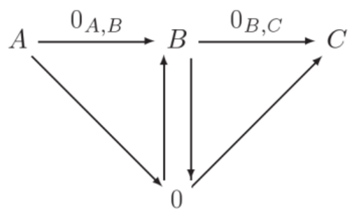
[5] This logical view highly contrasts the Birkhoff-von Neumann proposal in [9] that *quantum logic* is a weak logic in which we can do less than classical logic. In fact, *we can do more*!

10 Categorical matrix calculus

While till now we have focussed on the tensor product of Hilbert spaces, in this section we show how the direct sum of Hilbert spaces carries the matrix calculus. If an object is both terminal and initial we call it a *zero-object* and in that case there is a unique *zero-maps* between any two objects:



and these *zero-maps* are moreover closed under composition:



Assume that in addition to this we have a situation of what we roughly describe as 'coinciding products and coproducts', and which we will denote by $- \oplus -$. Since in this case we both have diagonal Δ and a codiagonal ∇, for each pair $f, g : A \to B$ we can define the following *sum*:

$$f + g := A \xrightarrow{\Delta} A \oplus A \xrightarrow{f \oplus g} A \oplus A \xrightarrow{\nabla} A$$

and by naturality of Δ and ∇ it moreover follows that

$$(f_1 + f_2) \circ g = (f_1 \circ g) + (f_2 \circ g) \qquad f \circ (g_1 + g_2) = (f \circ g_1) + (f \circ g_2)\,.$$

One verifies that we obtain **CMon**-enrichment. But we also both have projections and injections so for each morphism

$$f : A \oplus B \to C \oplus D$$

we can write down a matrix

$$(f_{ij})_{ij} = \begin{pmatrix} p_1 \circ f \circ q_1 & p_1 \circ f \circ q_2 \\ p_2 \circ f \circ q_1 & p_2 \circ f \circ q_2 \end{pmatrix}$$

and it turns out that we obtain a full-blown matrix calculus in which we can add and multiply in the usual linear-algebraic fashion. The exact notion which captures the above situation is that of a *biproduct*. We give two alternative equivalent definitions.

G. Sica (ed.) *What is Category Theory?*

Definition. If **C** has a 0-object, products and coproducts and if all morphisms with matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are isos then **C** has *biproducts.*

Definition. If **C** is **CMon**-enriched and if there are morphisms

$$A \underset{q_1}{\overset{p_1}{\leftrightarrows}} A \oplus B \underset{q_2}{\overset{p_2}{\rightleftarrows}} B$$

with

$$p_i \circ q_j = \delta_{ij} \qquad \sum\nolimits_i q_i \circ p_i = 1_{A\oplus B}$$

then **C** has *biproducts.*

Each such biproduct category admits an additive and multiplicative matrix calculus, and each category with numbers as objects and $n \times m$-matrices in a commutative semiring as morphisms yields a biproduct category. In particular $(\mathbf{Rel}, +)$ and $(\mathbf{Vect}_{\mathbb{K}}, \oplus)$ are biproduct categories.

Distributivity. We have now seen that in **FdHilb** there exist two monoidal structures, namely the $\otimes$-structure which captures entanglement, and the $\oplus$-structure which provides the matrix calculus. But these two are not at all independent since there exists a *distributivity natural isomorphism*:

$$\begin{array}{ccc}
(A_1 \oplus A_2) \otimes C & \xrightarrow{(f_1 \oplus f_2) \otimes g} & (B_1 \oplus B_2) \otimes D \\
\Big\downarrow \mathsf{DIST}_{A_1,A_2,C} & & \Big\downarrow \mathsf{DIST}_{B_1,B_2,D} \\
(A_1 \otimes C) \oplus (A_2 \otimes C) & \xrightarrow[(f_1 \otimes g) \oplus (f_2 \otimes g)]{} & (B_1 \otimes D) \oplus (B_2 \otimes D)
\end{array}$$

Such a distributity isomorphism is a very useful tool which for example can be used to encode *classical communication* between agents [2]:

$$(\mathrm{I} \oplus \mathrm{I}) \otimes Agent \simeq (\mathrm{I} \otimes Agent) \oplus (\mathrm{I} \otimes Agent)\,.$$

However, while the $\otimes$-structure and $\oplus$-structure clearly behave very different, the first is in the case of *finite dimensional objects* derivable from the second. Indeed, given a biproduct category **BP** with an object I such that $\mathbf{BP}(\mathrm{I}, \mathrm{I})$ is commutative, define a new category:

- the objects are the natural numbers $\mathbb{N} \simeq \{\underbrace{\mathrm{I} \oplus \ldots \oplus \mathrm{I}}_{n} \mid n \in \mathbb{N}\}$

G. Sica (ed.) *What is Category Theory?*