Frobenius algebras and 2D topological quantum field theories (short version)

Joachim Kock[1]
Université de Nice Sophia-Antipolis

Abstract

These notes centre around notions of Frobenius structure which in recent years have drawn some attention in topology, physics, algebra, and computer science. In topology the structure arises in the category of 2-dimensional oriented cobordisms (and their linear representations, which are 2-dimensional topological quantum field theories) — this is the subject of the first section. The main result here (due to Abrams [1]) is a presentation in terms of generators and relations of the monoidal category ***2Cob***. In algebra, the structure manifests itself simply as Frobenius algebras, which are treated carefully in Section 2. The main result is a characterisation of Frobenius algebras in terms of comultiplication which goes back to Lawvere [21] and was rediscovered by Quinn [25] and Abrams [1]. The main result of these notes is that these two categories are equivalent (cf. Dijkgraaf [12]): the category of 2D topological quantum field theories and the category of commutative Frobenius algebras. More generally, the notion of Frobenius object in a monoidal category is introduced, and it is shown that ***2Cob*** is the free symmetric monoidal category on a commutative Frobenius object. This generalises the main result.

The present text is the bare skeleton of an extensive set of notes prepared for an undergraduate minicourse on the subject. The full text [20] will appear elsewhere.

Contents

[1]Currently supported by a Marie Curie Fellowship from the European Commission

0 Introduction

0.1 Main themes

0.1.1 Frobenius algebras. A Frobenius algebra is a finite-dimensional algebra equipped with a nondegenerate bilinear form compatible with the multiplication. Examples are matrix rings, group rings, the ring of characters of a representation, and artinian Gorenstein rings (which in turn include cohomology rings, local rings of isolated hypersurface singularities…)

In algebra and representation theory such algebras have been studied for a century.

0.1.2 Frobenius structures. During the past decade, Frobenius algebras have shown up in a variety of topological contexts, in theoretical physics and in computer science. In physics, the main scenery for Frobenius algebras is that of topological quantum field theory, which in its axiomatisation amounts to a precise mathematical theory. In computer science, Frobenius algebras arise in the study of flowcharts, proof nets, circuit diagrams…

In any case, the reason Frobenius algebras show up is that it is essentially a topological structure: it turns out the axioms for a Frobenius algebra can be given completely in terms of graphs — or as we shall do, in terms of topological surfaces.

The goal of these notes is to make all this precise. We will focus on topological quantum field theories — and in particular on dimension 2. This is by far the best picture of the Frobenius structures since the topology is explicit, and since there is no additional structure to complicate things. In fact, the main theorem of these notes states that there is an equivalence of categories between that of 2D TQFTs and that of commutative Frobenius algebras.

(There will be no further mention of computer science in these notes.)

0.1.3 Topological quantum field theories. In the axiomatic formulation (due to M. Atiyah [3]), an n-dimensional topological quantum field theory is a rule $\mathscr{A}$ which to each closed oriented manifold Σ (of dimension $n-1$) associates a vector space $\Sigma\mathscr{A}$, and to each oriented n-manifold whose boundary is Σ associates a vector in $\Sigma\mathscr{A}$. This rule is subject to a collection of axioms which express that topologically equivalent manifolds have isomorphic associated vector spaces, and that disjoint unions of manifolds go to tensor products of vector spaces, etc.

0.1.4 Cobordisms. The clearest formulation is in categorical terms: first one defines a category of cobordisms $\boldsymbol{nCob}$: the objects are closed oriented $(n-1)$-manifolds, and an arrow from Σ to Σ' is an oriented n-manifold M whose 'in-boundary' is Σ and whose 'out-boundary' is Σ'. (The cobordism M is defined up to diffeomorphism rel the boundary.) Composition of cobordisms is defined by gluing together the underlying manifolds along common boundary components; the cylinder $\Sigma \times I$ is the identity arrow on Σ. The operation of taking disjoint union of manifolds gives this category monoidal structure. Now the axioms amount to saying that a TQFT is a monoidal functor from $\boldsymbol{nCob}$ to $\boldsymbol{Vect}_{\Bbbk}$.

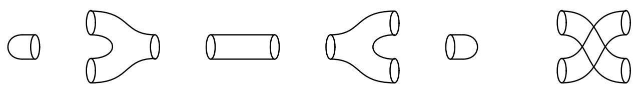
0.1.5 Physical interest in TQFTs comes from the observation that TQFTs possess certain features one expects from a theory of quantum gravity. It serves as a baby model in which one can do calculations and gain experience before embarking on the quest for the full-fledged theory. Roughly, the closed manifolds represent *space*, while the cobordisms represent *space-time*. The associated vector spaces are then the *state spaces*, and an operator associated to a space-time is the time-evolution operator (also called transition amplitude, or Feynman path integral). That the theory is topological means that the transition amplitudes do not depend on any additional structure on space-time (like riemannian metric or curvature), but only on the topology. In particular there is no time-evolution along cylindrical space-time. That disjoint union goes to tensor product expresses the common principle in quantum mechanics that the state space of two independent systems is the tensor product of the two state spaces.

(No further explanation of the relation to physics will be given — the author of these notes recognises he knows nearly nothing of this aspect. The reader is referred to Dijkgraaf [13] or Barrett [8], for example.)

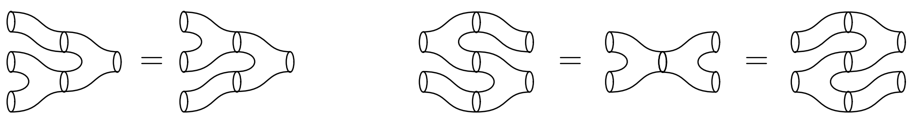
0.1.6 Mathematical interest in TQFTs stems from the observation that they produce invariants of closed manifolds: an n-manifold without boundary is a cobordism from the empty $(n-1)$-manifold to itself, and its image under $\mathscr{A}$ is therefore a linear map $\Bbbk \to \Bbbk$, i.e., a scalar. It was shown by E. Witten how TQFT in dimension 3 is related to invariants of knots and the Jones polynomial — see Atiyah [4].

The viewpoint of these notes is different however: instead of developing TQFTs in order to describe and classify manifolds, we work in dimension 2 where a complete classification of surfaces already exists; we then use this classification to describe TQFTs!

0.1.7 Cobordisms in dimension 2. In dimension 2, 'everything is known': since surfaces are completely classified, one can also describe the cobordism category completely. Every cobordism is obtained by composing the following basic building blocks (each with the in-boundary drawn to the left):



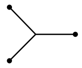
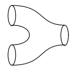
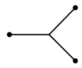
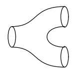
Two cobordisms are equivalent if they have the same genus and the same number of in- and out-boundaries. This gives a bunch of relations, and a complete description of the monoidal category ***2Cob*** in terms of generators and relations. Here are two examples of relations which hold in ***2Cob***:



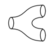
(0.1.8)

These equations express that certain surfaces are topologically equivalent rel the boundary.

0.1.9 Topology of some basic algebraic operations. Some very basic principles are in play here: 'creation', 'coming together', 'splitting up', 'annihilation'. These principles have explicit mathematical manifestations as algebraic operations:

Principle	Feynman diagram	2D cobordism	Algebraic operation (in a $\Bbbk$-algebra A)	
merging			multiplication	$A \otimes A \to A$
creation			unit	$\Bbbk \to A$
splitting			comultiplication	$A \to A \otimes A$
annihilation			counit	$A \to \Bbbk$

According to this dictionary, the left-hand relation of (0.1.8) is just the topological expression of associativity!

0.1.10 Frobenius algebras. In order to relate this to Frobenius algebras, the definition given in the beginning of this Introduction is not the most convenient. The principal result of Section 1 establishes an alternative characterisation of Frobenius algebras (which goes back at least to Lawvere [21]), namely as algebras (with multiplication denoted ) which are simultaneously *coalgebras* (with comultiplication denoted ) subject to a certain compatibility condition between  and  — this compatibility condition is exactly the right-hand relation drawn in (0.1.8). In fact, the basic relations valid in ***2Cob*** correspond precisely to the axioms of a commutative Frobenius algebra. This comparison leads to the main theorem:

There is an equivalence of categories

$$\boldsymbol{2TQFT}_{\Bbbk} \simeq \boldsymbol{cFA}_{\Bbbk},$$

given by sending a TQFT to its value on the circle (the unique closed connected 1*-manifold).*

0.1.11 Aftermath. Rather than leaving the result at that, it is rewarding to try to place the theorem in its proper context. The second part of Section 3 is devoted to this. The theorem is revealed to be a mere variation of a much more fundamental result: there is a monoidal category Δ (the simplex category) which is quite similar to ***2Cob*** (in fact it is a subcategory) such that giving a monoidal functor from Δ to $\boldsymbol{Vect}_{\Bbbk}$ is the same as giving a $\Bbbk$-algebra. More generally, there is a 1–1 correspondence between monoids in any monoidal category $\boldsymbol{V}$ and monoidal functors from Δ to $\boldsymbol{V}$. This amounts to saying that Δ is the free monoidal category on a monoid.

This result also has a variant for Frobenius algebras: we define a notion of (commutative) Frobenius object in a general (symmetric) monoidal category, such that a (commutative) Frobenius object in $\boldsymbol{Vect}_{\Bbbk}$ is precisely a (commutative) Frobenius algebra. This leads to the result that ***2Cob*** is in fact the universal symmetric Frobenius structure, in the sense that every commutative Frobenius object (in any symmetric monoidal category) arises as the image of a unique symmetric monoidal functor from ***2Cob***.

0.2 The context of these notes

0.2.1 The original audience. These notes originate in an intensive two-week mini-course for advanced undergraduate students, given in the Recife Summer School, January 2002. The requisites for the mini-course were modest: the students were expected only to have some familiarity with the basic notions of differentiable manifolds; basic notions of rings and groups; and familiarity with tensor products.

At an immediate level, the aim was simply to expose some delightful and not very well-known mathematics where a lot of figures can be drawn: a quite elementary and very nice interaction between topology and algebra — and quite different in flavour from what one learns in a course in algebraic topology. On a deeper level, the aim was to convey an impression of unity in mathematics, an aspect which is often hidden from the students until later in their mathematical apprenticeship. Finally, perhaps the most important aim was to use this as motivation for category theory, and specifically to serve as an introduction to monoidal categories.

The main theorem — that 2D TQFTs are just commutative Frobenius algebras — is admittedly not a particularly useful theorem in itself. What the lectures were meant to give the students were rather some techniques and viewpoints. A lot of emphasis was placed on universal properties, symmetry, distinction between structure and property, distinction between identity and natural isomorphism, the interplay between graphical and algebraic approaches to mathematics — as well as reflection on the nature of the most basic operations of mathematics: multiplication and addition.

In order to achieve to these goals, a large (perhaps exaggerated) amount of details, explanations, examples, comments, and pictures were provided. All this is recorded in the 'Schoolbook' for the minicourse, the *Recife notes on Frobenius algebras and 2D topological quantum field theories* [20].

0.2.2 The present text is a very condensed version of [20], prepared upon request from the organisers for inclusion in this volume. It is targeted at more experienced readers, who perhaps would feel impatient reading the schoolbook. The task of adapting the text to the requisition has consisted in merciless deletion of details, examples, and most of the figures, and drying up all chatty explanations and digressive comments. In this way, the exposition has become closer to the original sources (Quinn [25] and Abrams [1]), but it is my hope there is still a place for it, being more consistent about the categorical context and viewpoint — and there are still plenty of details here which the standard sources are silent about, especially with respect to symmetry.

0.2.3 Further reading. My big sorrow about these notes is that I don't understand the physical background or interpretation of TQFTs. The physically inclined reader must resort to the existing literature, for example Atiyah's book [4] or the notes of Dijkgraaf [13]. I would also like to recommend John Baez' web site [5], where a lot of references can be found.

Within the categorical viewpoint, an important approach to Frobenius structures which has not been touched upon is the 2-categorical viewpoint, in terms of monads and adjunctions. This has recently been exploited to great depth by Müger [24]. Again, a pleasant introductory account is given by Baez [5], TWF 174 (and 173).

Last but not least, I warmly recommend the lecture notes of Quinn [25], which are detailed and go in depth with concrete topological quantum field theories.

0.2.4 Acknowledgements. The idea of these notes goes back to a workshop I led at KTH, Stockholm, in 2000, whose first part was devoted to understanding the paper of Abrams [1] (corresponding more or less to Section 1 and 2 of this text). I am indebted to the participants of the workshop, and in particular to Dan Laksov. I have also benefited very much from discussions and e-mail correspondence with José Mourão, Peter Johnson, and especially Anders Kock. Finally, I thank the organisers of the Summer School in Recife — in particular Letterio Gatto — for the opportunity to giving this mini-course, and for warm reception in Recife.

0.2.5 General conventions. We consistently write composition of functions (or arrows) from the left to the right: given functions (or arrows)

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

we denote the composition fg. Similarly, we put the symbol of a function to the right of its argument, writing for example

$$\begin{aligned} f : X &\longrightarrow Y \\ x &\longmapsto xf. \end{aligned}$$

1 Cobordisms and TQFTs

1.1 Cobordisms

For the basic notions from differentiable topology, see for example Hirsch [18]. (A more elementary introduction which emphasises the concepts used here is Wallace [28].)

1.1.1 Terminology. Throughout the word *manifold* means smooth manifold (i.e., differentiable of class C^∞), and unless otherwise specified we always assume our manifolds to be compact and equipped with an orientation, but we do not assume them to be connected. *Closed* means (compact and) without boundary. For convenience, we consistently denote manifolds with boundary by capital Roman letters (typically M) while manifolds without boundary (and typically in dimension one less) are denoted by capital Greek letters, like Σ. Maps between manifolds are always understood to be smooth maps, and maps between manifolds of the same dimension are required to preserve orientation.

Contrary to custom in books on differentiable topology, we let submanifolds (e.g., the boundary) come equipped with an orientation on their own (instead of letting the ambient manifold induce one on it). This leads to the convenient notion of

1.1.2 In-boundaries and out-boundaries. Let Σ be a closed submanifold of M of codimension 1 — both equipped with an orientation. At a point $x \in \Sigma$, let $\{v_1, \ldots, v_{n-1}\}$ be a positively oriented basis for $T_x\Sigma$. A vector $w \in T_xM$ is called a *positive normal* if $\{v_1, \ldots, v_{n-1}, w\}$ is a positively oriented basis for T_xM. Now suppose Σ is a connected component of the boundary of M; then it makes sense to ask whether the positive normal w points inwards or outwards compared to M. If a positive normal points inwards we call Σ an *in-boundary*, and if it points outwards we call it an *out-boundary.* This notion does not depend on the choice of positive normal (nor on choice of point $x \in \Sigma$). Thus *the* boundary of a manifold M is the union of various in-boundaries and out-boundaries.

For example, if Σ is a submanifold of codimension 1 in M which divides M into two parts, then Σ is an out-boundary for one of the parts and an in-boundary for the other.

1.1.3 Cobordisms. Intuitively, given two closed $(n-1)$-manifolds Σ_0 and Σ_1, a cobordism from Σ_0 to Σ_1 is an oriented n-manifold M whose in-boundary is Σ_0 and whose out-boundary is Σ_1. However, in order to allow cobordisms from a given Σ to itself, we need a more relative description:

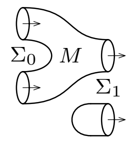
An *(oriented) cobordism* from Σ_0 to Σ_1 is an (oriented) manifold M together with maps

$$\Sigma_0 \to M \leftarrow \Sigma_1$$

such that Σ_0 maps diffeomorphically onto the in-boundary of M, and Σ_1 maps diffeomorphically onto the out-boundary of M. We will write it

$$\Sigma_0 \overset{M}{\Rightarrow} \Sigma_1.$$

Here is an example of a cobordism from a pair of circles Σ_0 to another pair of circles Σ_1:



1.1.4 Cylinders. Take a closed manifold Σ and cross it with the unit interval I (with its standard orientation). The boundary of $\Sigma \times I$ consists of two copies of Σ: one which is an in-boundary, $\Sigma \times \{0\}$, and another which is an out-boundary, $\Sigma \times \{1\}$. So we get a cobordism from Σ to itself by taking the obvious maps

$$\begin{array}{rcl} \Sigma & \xrightarrow{\sim} & \Sigma \times \{0\} \subset \Sigma \times I \\ \Sigma & \xrightarrow{\sim} & \Sigma \times \{1\} \subset \Sigma \times I. \end{array}$$

The same construction serves to give a cobordism between any pair of $(n-1)$-manifolds Σ_0 and Σ_1 both of which are diffeomorphic to Σ; just take

$$\begin{array}{rcl} \Sigma_0 \xrightarrow{\sim} \Sigma & \xrightarrow{\sim} & \Sigma \times \{0\} \subset \Sigma \times I \\ \Sigma_1 \xrightarrow{\sim} \Sigma & \xrightarrow{\sim} & \Sigma \times \{1\} \subset \Sigma \times I. \end{array}$$

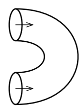
Any diffeomorphism $\Sigma \times I \xrightarrow{\sim} M$ will also define a cobordism $M : \Sigma_0 \Rightarrow \Sigma_1$. So in conclusion: for any two diffeomorphic manifolds Σ_0 and Σ_1 there exists a cobordism from Σ_0 to Σ_1, and in fact there are MANY! Those produced in this way are all *equivalent* cobordisms in the sense we now make precise:

1.1.5 Equivalent cobordisms. Two cobordisms from Σ_0 to Σ_1 are called *equivalent* if there is a diffeomorphism from M to M' making this diagram commute:

$$\begin{tikzcd} & M' & \\ \Sigma_0 \arrow[ur] \arrow[dr] & & \Sigma_1 \arrow[ul] \arrow[dl] \\ & M \arrow[uu, "\simeq"'] & \end{tikzcd}$$

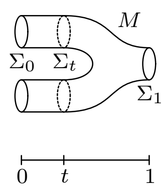
(Note that the source and target manifolds Σ_0 and Σ_1 are completely fixed, not just up to diffeomorphism.) In the next subsection we will divide out by these equivalences, and consider equivalence classes of cobordisms, called *cobordism classes*.

1.1.6 'U-tubes'. U-tubes are cylinders with reversed orientation on one of the boundaries. Precisely take a closed manifold Σ and map it onto the ends of the cylinder $\Sigma \times I$, in such a way that both boundaries are in-boundaries (then the out-boundary is empty). We will often draw such a cylinder like this:



just to keep the convention of having in-boundaries on the left, and out-boundaries on the right.

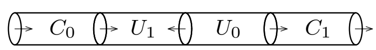
1.1.7 Decomposition of cobordisms. An important feature of a cobordism M is that you can *decompose* it: this means introducing a submanifold Σ which splits M into two parts, with all the in-boundaries in one part and all the out-boundaries in the other; Σ must be oriented such that its positive normal points toward the out-part. To arrange such a submanifold, take a smooth map $f : M \to [0,1]$ such that $f^{-1}(0) = \Sigma_0$ and $f^{-1}(1) = \Sigma_1$, and let Σ_t be the inverse image of a regular value t, oriented such that the positive normal points towards the out-boundaries, just as the positive normal of $t \in [0,1]$ points towards 1.



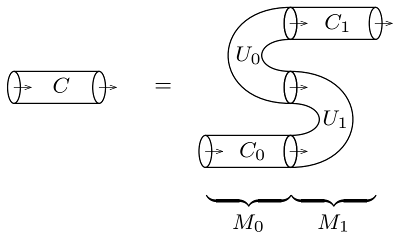
The result is two new cobordisms: one from Σ_0 to Σ_t given by the piece $M_{[0,t]} := f^{-1}([0,t])$, and another from Σ_t to Σ_1 given by the piece $M_{[t,1]} := f^{-1}([t,1])$.

In a minute, we will reverse this process and show how to compose two cobordisms, provided they have compatible boundaries.

1.1.8 'Snake decomposition' of a cylinder. Starting with a cylinder $C = \Sigma \times I$ over a closed manifold Σ, we can decompose it by cutting along three copies of Σ, with orientation as indicated:

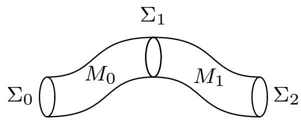


This is a true decomposition, provided we interpret the three pieces in the correct way: the in-part of the decomposition is $M_0 := C_0 \coprod U_0 \;:\; \Sigma \Rightarrow \Sigma \coprod \Sigma \coprod \Sigma$; the out-part of the decomposition is $M_1 := U_1 \coprod C_1 \;:\; \Sigma \coprod \Sigma \coprod \Sigma \Rightarrow \Sigma$. If we draw the pieces U_0 and U_1 as U-tubes as in 1.1.6 it is easier to grasp the decomposition:



1.2 The category of cobordism classes

We will now assemble cobordisms into a category: the objects should be closed oriented $(n-1)$-manifolds, and the arrows should be oriented cobordisms. So we need to show how to compose two cobordisms (and check associativity), and we need to find identity arrows for each object. Given one cobordism $M_0 : \Sigma_0 \Rightarrow \Sigma_1$ and another $M_1 : \Sigma_1 \Rightarrow \Sigma_2$ then the composition $M_0 M_1 : \Sigma_0 \Rightarrow \Sigma_2$ should be obtained by gluing together the manifolds M_0 and M_1 along Σ_1. This is a manifold with in-boundary Σ_0 and out-boundary Σ_2, and Σ_1 sits inside it as a submanifold.



However this construction $M_0 M_1 := M_0 \coprod_\Sigma M_1$ is not well defined in the category of smooth manifolds. It is well-defined as a topological manifold, but there is no canonical choice of smooth structure near the gluing locus Σ_1: the smooth structure turns out to be well-defined only up to diffeomorphism. (And not even unique diffeomorphism.)

Concerning identity arrows, the identity ought to be a cylinder of height zero, but such a 'cylinder' is not an n-manifold!

Both problems are solved by passing to diffeomorphism classes of cobordisms rel the boundary. Precisely, we identify cobordisms which are equivalent in the sense of 1.1.5 and let the arrows of our category be these equivalence classes, called *cobordism classes*.

The pertinent result is this — see Milnor [23], Thm. 1.4:

1.2.1 Theorem. *Let $M_0 : \Sigma_0 \Rightarrow \Sigma_1$ and $M_1 : \Sigma_1 \Rightarrow \Sigma_2$ be two cobordisms, then there exists a smooth structure on the topological manifold $M_0 M_1 := M_0 \coprod_{\Sigma_1} M_1$ such that the embeddings $M_0 \hookrightarrow M_0 M_1$ and $M_1 \hookrightarrow M_0 M_1$ are diffeomorphisms onto their images. This smooth structure is unique up to diffeomorphism fixing Σ_0, Σ_1, and Σ_2.*

We will not go into the proof of this result, but only mention the reason why at least the smooth structure exists. (In the 2-dimensional case the uniqueness then follows from the well-known result that two smooth surfaces are diffeomorphic if and only if they are homeomorphic.)

1.2.2 Gluing of cylinders. The first step is to provide smooth structure on the gluing of two cylinders $\Sigma \times [0, 1]$ and $\Sigma \times [1, 2]$. The gluing is simply $\Sigma \times [0, 2]$, and it has an obvious smooth structure namely the product one. (Clearly this structure is diffeomorphically compatible with the two original structures as required.)

This innocent-looking observation becomes important in view of the following result:

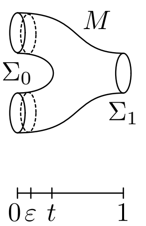
1.2.3 Regular interval theorem. (See Hirsch [18], 6.2.2.) *Let $M : \Sigma_0 \Rightarrow \Sigma_1$ be a cobordism and let $f : M \to [0,1]$ be a smooth map without any critical points at all, and such that $\Sigma_0 = f^{-1}(0)$ and $\Sigma_1 = f^{-1}(1)$. Then there is a diffeomorphism from the cylinder $\Sigma_0 \times [0,1]$ to M compatible with the projection to $[0,1]$ like this:*

$$\begin{tikzcd} \Sigma_0 \times [0,1] \arrow[r, "\sim"] \arrow[dr] & M \arrow[d, "f"] \\ & {[0,1]} \end{tikzcd}$$

(And similarly, there is a diffeomorphism $\Sigma_1 \times [0,1] \xrightarrow{\sim} M$ compatible with the projection.)

1.2.4 Corollary. *Let $M : \Sigma_0 \Rightarrow \Sigma_1$ be a cobordism. Then there is a decomposition $M = M_{[0,\varepsilon]} M_{[\varepsilon,1]}$ such that $M_{[0,\varepsilon]}$ is diffeomorphic to a cylinder over Σ_0. (And similarly there is another decomposition such that the part near Σ_1 is diffeomorphic to a cylinder over Σ_1.)*

Indeed, take a Morse function $f : M \to [0,1]$, and let t be the first critical value. Then for $\varepsilon < t$ the interval $[0,\varepsilon]$ is regular, so if we cut M along the inverse image $f^{-1}(\varepsilon)$, by the regular interval theorem we get the required decomposition.



1.2.5 Gluing of general cobordisms. Given cobordisms $M_0 : \Sigma_0 \Rightarrow \Sigma_1$ and $M_1 : \Sigma_1 \Rightarrow \Sigma_2$. Take Morse functions $f_0 : M_0 \to [0,1]$ and $f_1 : M_1 \to [1,2]$, and consider the topological manifold $M_0 \coprod_{\Sigma_1} M_1$, with the induced continuous map $M_0 \coprod_{\Sigma_1} M_1 \to [0,2]$. Choose $\varepsilon > 0$ so small that the two intervals $[1-\varepsilon, 1]$ and $[1, 1+\varepsilon]$ are regular for f_0 and f_1 respectively; then the inverse images of these two intervals are diffeomorphic to cylinders. So within the interval $[1-\varepsilon, 1+\varepsilon]$ we are in the situation of 1.2.2, and we can take the smooth structure to be the one coming from the cylinder.

Theorem 1.2.1 shows that the composition of two cobordisms is a well-defined cobordism class. Clearly this class only depends on the classes of the two original cobordisms, not on the cobordisms themselves, so we have a well-defined composition for cobordism classes. This composition is associative since gluing of topological spaces (the pushout) is associative.

Also, it is easy to prove that the class of a cylinder is the identity cobordism class for the composition law: it amounts to two observations: (i) every cobordism has a part near the boundary where it is diffeomorphic to a cylinder (cf. 1.2.4), (ii) the composition of two cylinders is again a cylinder (cf. 1.2.2).

1.2.6 The category *nCob*. The objects of ***nCob*** are $(n-1)$-dimensional closed oriented manifolds. Given two objects Σ_0 and Σ_1, the arrows from Σ_0 to Σ_1 are equivalence classes of oriented cobordisms $M : \Sigma_0 \Rightarrow \Sigma_1$, in the sense of 1.1.5.)

Henceforth, the term 'cobordism' will be used meaning 'cobordism class'.

1.2.7 Diffeomorphisms and their induced cobordism classes. It was mentioned in 1.1.4 how a diffeomorphism $\phi : \Sigma_0 \xrightarrow{\sim} \Sigma_1$ induces a cobordism $C_\phi : \Sigma_0 \Rightarrow \Sigma_1$, via the cylinder construction. In fact this construction is functorial: given two diffeomorphisms $\Sigma_0 \overset{\phi_0}{\xrightarrow{\sim}} \Sigma_1 \overset{\phi_1}{\xrightarrow{\sim}} \Sigma_2$, then we have

$$C_{\phi_0} C_{\phi_1} = C_{\phi_0 \phi_1}.$$

In particular, a cobordism induced from a diffeomorphism is invertible. Also, the identity diffeomorphism $\Sigma \xrightarrow{\sim} \Sigma$ induces the identity cobordism. In other words, the cylinder construction 1.1.4 defines a functor from the category of $(n-1)$-manifolds and diffeomorphisms to the category ***nCob***.

The next thing to note is that

Two diffeomorphisms $\psi_0, \psi_1 : \Sigma_0 \xrightarrow{\sim} \Sigma_1$ induce the same cobordism class $\Sigma_0 \Rightarrow \Sigma_1$ if and only if they are (smoothly) homotopic. Indeed, a smooth homotopy between the diffeomorphisms is a map $\Sigma_0 \times I \to \Sigma_1$ compatible with ψ_0 and ψ_1 on the boundary; such a map induces also a map $\Sigma_0 \times I \to \Sigma_1 \times I$ which is the diffeomorphism rel the boundary witnessing the equivalence of the induced cobordisms — and conversely.

In particular, a cobordism $\Sigma \Rightarrow \Sigma$ induced by a diffeomorphism $\psi : \Sigma \xrightarrow{\sim} \Sigma$ is the identity if and only if ψ is homotopic to the identity. An important example of a diffeomorphism which is not homotopic to the identity is the twist diffeomorphism $\Sigma \coprod \Sigma \to \Sigma \coprod \Sigma$ which interchanges the two copies of Σ.

1.2.8 Monoidal structure. If Σ and Σ' are two $(n-1)$-manifolds then the disjoint union $\Sigma \coprod \Sigma'$ is again an $(n-1)$-manifold, and given two cobordisms $M : \Sigma_0 \Rightarrow \Sigma_1$ and $M' : \Sigma'_0 \Rightarrow \Sigma'_1$, their disjoint union $M \coprod M'$ is naturally a cobordism from $\Sigma_0 \coprod \Sigma'_0$ to $\Sigma_1 \coprod \Sigma'_1$. The empty n-manifold $\emptyset_n$ is a cobordism $\emptyset_{n-1} \Rightarrow \emptyset_{n-1}$. Clearly these structure make $(\boldsymbol{nCob}, \coprod, \emptyset)$ into a monoidal category.

The cobordism induced by the twist diffeomorphism $\Sigma \coprod \Sigma' \xrightarrow{\sim} \Sigma' \coprod \Sigma$ will be called the *twist cobordism* (for Σ and Σ'), denoted $T_{\Sigma,\Sigma'} : \Sigma \coprod \Sigma' \Rightarrow \Sigma' \coprod \Sigma$. It will play an important rôle in the sequel. It is straight-forward to check that the twist cobordisms satisfy the axioms for a symmetric structure on the monoidal category $(\boldsymbol{nCob}, \coprod, \emptyset)$ — this is an easy consequence of the fact that the twist diffeomorphism is a symmetry structure in the monoidal category of smooth manifolds. So $(\boldsymbol{nCob}, \coprod, \emptyset, T)$ is a symmetric monoidal category — see 3.1.5 for some more details on these definitions.

1.2.9 Topological quantum field theories. Roughly, a quantum field theory takes as input *spaces* and *space-times* and associates to them *state spaces* and *time evolution operators*. The space is modelled as a closed $(n-1)$-manifold, while space-time is an n-manifold whose boundary represents time 0 and time 1. The state space is a vector space (over some ground field $\Bbbk$), and the time evolution operator is simply a linear map from the state space of time 0 to the state space of time 1.

In mathematical terms (cf. Atiyah [3]), it is a rule $\mathscr{A}$ which to each closed $(n-1)$-manifold Σ associates a vector space $\Sigma\mathscr{A}$, and to each cobordism $M : \Sigma_0 \Rightarrow \Sigma_1$ associates a linear map $M\mathscr{A}$ from $\Sigma_0\mathscr{A}$ to $\Sigma_1\mathscr{A}$.

The theory is called *topological* if it only depends on the topology of the space-time. This means that equivalent cobordisms must have the same image:

$$M \cong M' \Rightarrow M\mathscr{A} = M'\mathscr{A}$$

It also means that 'nothing happens' as long as time evolves cylindrically, i.e., the cylinder $\Sigma \times I$, thought of as a cobordism from Σ to itself, must be sent to the identity map of $\Sigma\mathscr{A}$. Furthermore, given a decomposition $M = M'M''$ then

$$M\mathscr{A} = (M'\mathscr{A})(M''\mathscr{A}) \quad \text{(composition of linear maps).}$$

These requirements together amount to saying that $\mathscr{A}$ is a functor from $\boldsymbol{nCob}$ to $\boldsymbol{Vect}_{\Bbbk}$.

Next there are two requirements which account for the word 'quantum': Disjoint union goes to tensor product: if $\Sigma = \Sigma' \coprod \Sigma''$ then $\Sigma\mathscr{A} = \Sigma'\mathscr{A} \otimes \Sigma''\mathscr{A}$. This must also hold for cobordisms: if $M : \Sigma_0 \Rightarrow \Sigma_1$ is the disjoint union of $M' : \Sigma_0' \Rightarrow \Sigma_1'$ and $M'' : \Sigma_0'' \Rightarrow \Sigma_1''$ then $M\mathscr{A} = M'\mathscr{A} \otimes M''\mathscr{A}$. This reflects a standard principle of quantum mechanics: that the state space for two systems isolated from each other is the tensor products of the two state spaces. Also, the empty manifold $\Sigma = \emptyset$ must be sent to the ground field $\Bbbk$. (It follows that the empty cobordism (which is the cylinder over $\Sigma = \emptyset$) is sent to the identity map of $\Bbbk$.)

Finally we will require that the monoidal functor be symmetric. More remarks on this requirement will be given in 3.1.14.

Summing up these axioms we can say that

An n-dimensional topological quantum field theory is a symmetric monoidal functor from $(\boldsymbol{nCob}, \coprod, \emptyset, T)$ *to* $(\boldsymbol{Vect}_{\Bbbk}, \otimes_{\Bbbk}, \Bbbk, \sigma)$.

Let us see how the axioms work, and extract some important consequences of them.

1.2.10 Nondegenerate pairings and finite-dimensionality. Take any closed manifold Σ, and let $V := \Sigma\mathscr{A}$ be its image under a TQFT $\mathscr{A}$. Consider the U-tube over Σ with two in-boundaries (cf. 1.1.6); its image under $\mathscr{A}$ is then a pairing $\beta : V \otimes V \to \Bbbk$. Similarly, the U-tube with two out-boundaries is sent to a copairing $\gamma : \Bbbk \to V \otimes V$ (see 2.1.4). Now consider the snake decomposition of the cylinder $\Sigma \times I$ described in 1.1.8. The axioms imply that the composition of linear maps

$$V \xrightarrow{\mathrm{id}_V \otimes \gamma} V \otimes V \otimes V \xrightarrow{\beta \otimes \mathrm{id}_V} V \tag{1.2.11}$$

is the identity map. This is to say that the pairing β is nondegenerate, according to Definition 2.1.4). In particular, by Lemma 2.1.5, V is of finite dimension. In other words, the axioms for a TQFT automatically imply that the image vector spaces are finite-dimensional.

1.3 Generators and relations for *2Cob*

The categories $\boldsymbol{nCob}$ are very difficult to describe for $n \geq 3$. But the category $\boldsymbol{2Cob}$ can be described explicitly in terms of generators and relations, and that is the goal of this subsection. (The reason for this difference is of course that while there is a complete classification theorem for surfaces, no such result is known in higher dimensions.)

1.3.1 Generators for a monoidal category. Recall that a *generating set* for a monoidal category $\boldsymbol{C}$ is a set S of arrows such that every arrow in $\boldsymbol{C}$ can be obtained the arrows of S by composition and the monoidal operation.

1.3.2 Skeletons of *2Cob*. To give meaning to the concept of generators and relations in our case, we take a skeleton of the category ***2Cob***, i.e., a full subcategory comprising exactly one object from each isomorphism class.

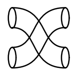
The first observation is that every closed 1-manifold is diffeomorphic to a finite disjoint union of circles — remember that closed implies compact. Each diffeomorphism induces an invertible cobordism via the cylinder construction, and conversely, it is not hard to show that in fact all the invertible cobordisms arise this way. In other words, two objects in ***2Cob*** are in the same isomorphism class, if and only if they are diffeomorphic.

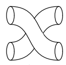
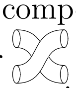
So we get a skeleton of ***2Cob*** as follows. Let $\mathbf{0}$ denote be the empty 1-manifold; let $\mathbf{1}$ denote a given circle Σ, and let $\mathbf{n}$ denote the disjoint union of n copies of Σ. Then the full subcategory $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots\}$ is a skeleton of ***2Cob***.

Henceforth we let ***2Cob*** denote this skeleton.

Notice that the chosen skeleton is closed under the operation of taking disjoint union, and that in fact disjoint union is the main principle of its construction.

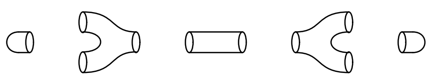
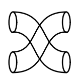
1.3.3 The twist. Since disjoint union is a generating principle, we can largely concentrate on cobordisms which are connected. But not completely: for each object $\mathbf{n}$, ($n \geq 2$), there are cobordisms $\mathbf{n} \Rightarrow \mathbf{n}$ which are not the identity. For example, for $\mathbf{2}$ (the disjoint union of two circles), we have the *twist* cobordism T (induced by the twist diffeomorphism):



(The drawing is not meant to indicate that the two components intersect: the reason for drawing it like this instead of something like  or  is to avoid any idea of crossing over or under. Since we are talking about abstract manifolds, not embedded anywhere, it has no meaning to talk about crossing over or under.)

Is is important to note that the twist cobordism T is *not* equivalent to the disjoint union of two cylinders, since the diffeomorphisms realising an equivalence are required to respect the boundary. Another argument for this fact is that T is induced by the twist diffeomorphism, which certainly is not homotopic to the identity map, and therefore T cannot be equivalent to the identity cobordism.

1.3.4 Proposition. *The monoidal category* **2Cob** *is generated under composition (serial connection) and disjoint union (parallel connection) by the following six cobordisms.*

 *and* 

It should be noted that the identity arrow is not usually included when listing generators for a category, because identity arrows is a concept which actually comes before the notion of a category. Here we include mostly for graphical convenience. The price to pay is of course some corresponding extra relations, (1.3.10).

We give two proofs of the Proposition, since they both provide some insight. In any case some nontrivial result about surfaces is needed. The first proof relies directly on the classification of surfaces (quoted below): the connected surfaces are classified by some topological invariants, and we simply build a surface with given invariants! To get the non-connected cobordisms we use disjoint union and permutation of the factors of the disjoint union. Since every permutation can be written as a composition of transpositions, the sixth generator suffices to do this. The drawback of this first proof is that it doesn't say so much about how a given surface relates to this 'normal form' — this information is hidden in the quoted classification theorem.

The second proof relies on a result from Morse theory (which is an ingredient in one of the possible proofs of the classification theorem (cf. Hirsch [18])), and here we do exactly what we missed in the first proof: start with a concrete surface and cut it up in pieces; now identify each piece as one of the generators.

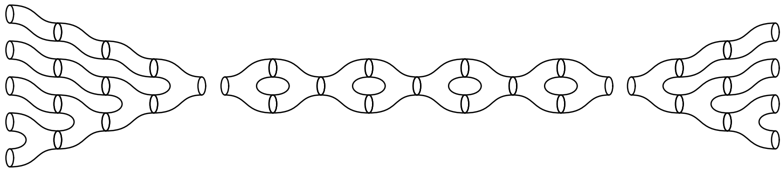
1.3.5 Reminders on genus, Euler characteristic, and classification of surfaces. For a surface with boundary, the genus is defined to be the genus of the closed surface obtained by sewing in discs along each boundary component. Alternatively, using the Euler characteristic, we have

$$\chi(M) + k = 2 - 2g$$

where k is the number of missing discs (i.e., the number of boundary components). *Two connected, compact, oriented surfaces with oriented boundary are diffeomorphic if and only if they have the same genus and the same number of in-boundaries and the same number of out-boundaries.*

1.3.6 'Normal form' of a connected surface. It is convenient to introduce the *normal form* of a connected surface with m in-boundaries, n out-boundaries, and genus g. It is actually a decomposition of the surface into a number of basic cobordisms. The normal form has three parts: the first part (called the in-part) is a cobordism $\mathbf{m} \Rightarrow \mathbf{1}$ of genus 0; the middle part (referred to as the topological part) is a cobordism $\mathbf{1} \Rightarrow \mathbf{1}$ of genus g; and the third part (the out-part) goes $\mathbf{1} \Rightarrow \mathbf{n}$ (again genus 0).

Rather than describing the normal form formally, we content ourselves with the following figure of the normal form in the case $m = 5$, $g = 4$, and $n = 4$.



(In the case $m = 0$, instead of having any pair-of-pants, the whole in-part just consists of a single . In case $n = 0$, the out-part consists of a single copy of .)

1.3.7 First proof of Proposition 1.3.4. The normal form is at the same time a recipe for constructing any connected cobordism from the generators — this proves 1.3.4 for connected cobordisms.

Assume now that M is disconnected — for simplicity say with two components M_0 and M_1. In other words, as a manifold, M is the (disjoint) union of M_0 and M_1. However, this is not sufficient to invoke 'disjoint union as generating principle' because that principle refers to the special notion of *disjoint union of cobordisms*, and being a cobordism involves some labelling or ordering of the boundaries. The easiest example of this distinction is the twist: we explained in 1.3.3 that although the twist (as a manifold) is the (disjoint) union of two cylinders, it is *not* the disjoint union (as cobordism) of two identity cobordisms.

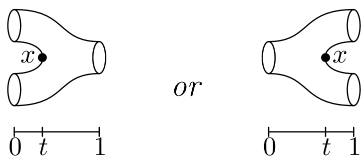
The only problem is ordering of the boundaries, and we can permute the boundaries by composing with twist cobordisms. Since the symmetric group is generated by transposition of two neighbour letters, every permutation of the boundaries can be realised using twists. □

The Morse theoretic proof is a bit different in spirit. The key is to characterise the generators in terms of their critical points. Part of this task was done in 1.2.3, where we saw that a cobordism admitting a Morse function without critical points is diffeomorphic to a cylinder rel one of the boundaries. But such a cobordism $\Sigma_0 \Rightarrow \Sigma_1$ is induced by a diffeomorphism $\psi : \Sigma_0 \xrightarrow{\sim} \Sigma_1$, and therefore equivalent to a permutation cobordism.

So in conclusion, if a cobordism admits a Morse function without critical points then it is equivalent to a permutation cobordism, and thus it can be constructed with twist cobordisms (and the identity).

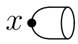
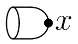
The next ingredient is this lemma:

1.3.8 Lemma. *(See Hirsch [18], 4.4.2.) Let M be a compact connected orientable surface with a Morse function $M \to [0,1]$. If there is a unique critical point x, and x has index* 1 *(i.e., is a saddle point) then M is diffeomorphic to a disc with two discs missing (these boundaries are over* 0 *and* 1*). In other words we have*

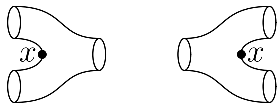


1.3.9 Morse theoretic proof of Proposition 1.3.4. Consider a cobordism $M : \Sigma_0 \Rightarrow \Sigma_1$, and take a Morse function $f : M \to [0,1]$ with $f^{-1}(0) = \Sigma_0$ and $f^{-1}(1) = \Sigma_1$. Take a sequence of regular values $a_0, a_1, \ldots, a_k$ in such a way that there is (at most) one critical value in each interval $[a_i, a_{i+1}]$; consider one of these intervals, $[a,b]$. We can assume there is at most one critical point x in the inverse image $M_{[a,b]}$. The piece $M_{[a,b]}$ may consist of several connected components: (at most) one of them contains x; the others are equivalent to permutation cobordisms, as we have argued above (based on the regular interval theorem 1.2.3), so these pieces can be chopped up further until they are a composition of twist cobordisms and identities.

So we can assume $M_{[a,b]}$ is connected and has a unique critical point x. Now if x has index 0 (resp. 2) then we have a local minimum (resp. maximum), and then $M_{[a,b]}$ is a disc like this:

 (resp. ).

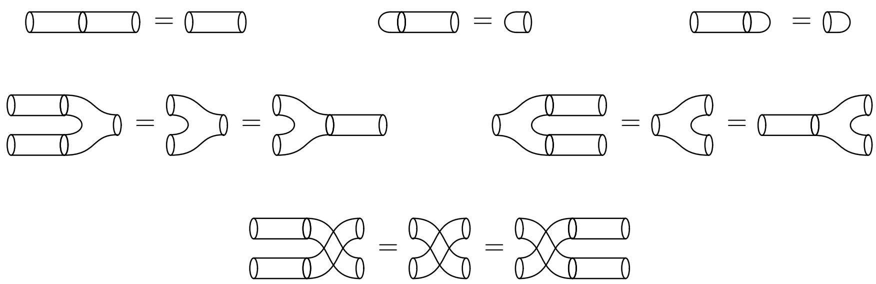
And finally if the index is 1 then we have a saddle point, and by Lemma 1.3.8, $M_{[a,b]}$ is then topologically a disc with two holes, so in our picture is looks like one of those:



$\square$

We will first list all the relations. Afterwards we prove that they hold, and provide some more comments on each relation.

1.3.10 Identity relations. First of all, we have already shown that the cylinders are identities. This gives a bunch of relations:



(Note that  is the identity for **2**.)

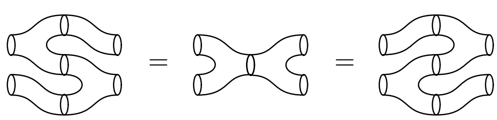
1.3.11 Sewing in discs. *The following relations hold.*

1.3.12 'Associativity' and 'coassociativity'. *These relations hold:*

1.3.13 'Commutativity' and 'cocommutativity'. *We have:*

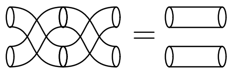
And finally:

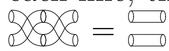
1.3.14 'The Frobenius relation' *holds:*

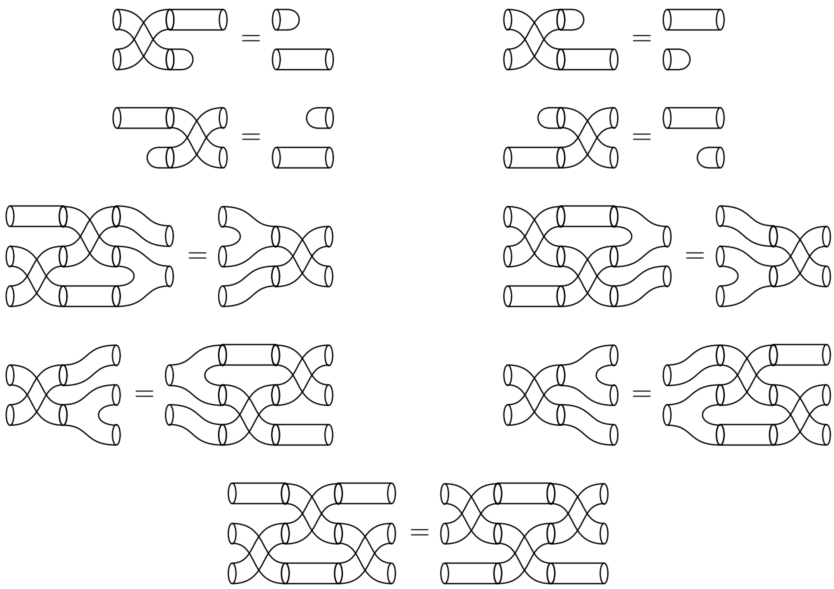


1.3.15 Easy proof of all the above relations. Simply note that in each case the surfaces have the same topological type, so according to the classification theorem they are diffeomorphic. $\square$

1.3.16 Relations involving the twist. The statement that the twist cobordism makes $(\boldsymbol{2Cob}, \coprod, \emptyset)$ into a symmetric monoidal category amounts to a set of relations involving the twist. The basic relation is the fact that the twist is its own inverse



The naturality of the twist cobordism means that for any pair of cobordisms, it makes no difference whether we interchange their input or their output. With our list of generators this boils down to the following relations. (In each line, the two listed relations are actually dependent modulo the basic twist relation  .)



The last relation is the relation for the symmetric group, as generated by transpositions (see Coxeter-Moser [10], 6.2.). (The above relations are analogous to the famous Reidemeister moves in knot theory (see Kassel [19], p.248), but they are much simpler because in our context there is no distinction between passing over and under.)

1.3.17 Sufficiency of the relations. To show that the listed relations suffice, we first consider connected surfaces. We show that every decomposition of a connected surface can be brought on normal form (1.3.6). First we treat the case without twist cobordisms, next we eliminate any eventual twists by an induction argument.

1.3.18 Counting the pieces. Given an arbitrary decomposition of a connected surface M with m in-boundaries, of genus g, and with n out-boundaries, then the Euler characteristic is $\chi(M) = 2-2g-m-n$. Let a be the number of  pieces in the decomposition; let b be the number of  ; let p be the number of  ; and let q be the number of  . If $M = M_0M_1$ is a decomposition then by the additive property of the Euler characteristic,

$\chi(M) = \chi(M_0) + \chi(M_1)$ (valid only in dim 2), so we can also write $\chi(M) = p + q - a - b$. Thus we have the equation

$$2 - 2g - m - n = p + q - a - b.$$

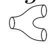
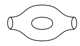
On the other hand we have the distinction between in- and out-boundaries. Summing up what each piece contributes to the number of boundaries we get the equation

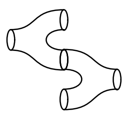
$$a + q + n = b + p + m.$$

Combining the two equations we can solve for a and b to get

$$\begin{aligned} a &= m + g - 1 + p \\ b &= n + g - 1 + q, \end{aligned}$$

and all the involved symbols are non-negative integers.

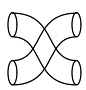
1.3.19 Moving  pieces left. Our strategy is to take the $m - 1 + g + p$  pieces and move them to the left: g of them will get stuck when they meet a  like this:  ; p of them will meet a  and vanish (due to the unit relation 1.3.11); and the remaining $m - 1$ will pass all the way through and make up the in-part of the normal form. On its way left, a  may also meet  like this

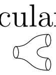


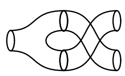
But it can move past this one, due to relation 1.3.14. Similarly, if a left-moving  meets a handle it can pass through: use first associativity and then the Frobenius relation.

Similarly we move all the  to the right until they form the out-part of the normal form. The middle part will then consists exactly in g handles, as wanted.

1.3.20 Eliminating twist maps. Suppose now there are twist cobordisms in the decomposition; pick one T, and let A, B, C, D denote the rest of the surface as indicated here:

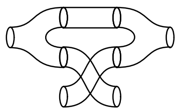


Since the surface is assumed to be connected, some of the regions A, B, C, D must be connected with each other. Suppose A and C are connected. Then together they form a connected surface involving strictly fewer twist than the original, so by induction we can assume it can be brought on normal form using the relations. In particular only the out-part of the normal form of A-with-C touches T, and we can shuffle the  pieces up and down until there is a piece which matches exactly with T like this

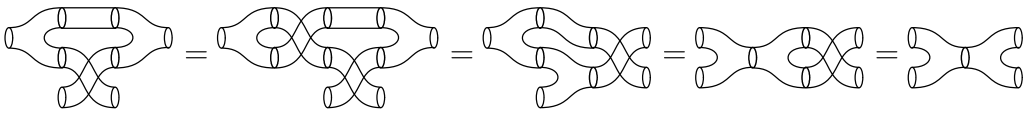


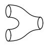
and then we use the cocommutativity relation 1.3.13 to remove T and we are done. If B and D are interconnected the same argument applies.

So we can assume that A and B are connected. Now each region A and B comprises fewer twist maps than the whole, so we can assume they are on normal form, and therefore the situation close to T is this:



Now we can eliminate the twist with this series of moves:



(The first move is cocommutativity 1.3.13, the second move is naturality of the twist map (moving past a ) cf. 1.3.16; next comes the Frobenius relation 1.3.14, and finally we use cocommutativity again.)

1.3.21 Unravelling disconnected surfaces. It remains to notice that any disconnected surface can be brought to be a disjoint union of connected surfaces by permuting the boundaries, which can be done by twist maps. (The fact that the relations listed in 1.3.16 are sufficient for this purpose follows from the fact that the corresponding relations for the symmetric group form a complete set of relations.)

Finally it should be noted that the listed set of relations is not minimal: we will see in 2.4.11 that The Frobenius relation 1.3.14 together with the unit and counit relations 1.3.11 imply the associativity and coassociativity relations 1.3.12.

1.3.22 Remarks. The first explicit description of the monoidal category ***2Cob*** in terms of generators and relations was given in Abrams [1], but the proof is already sketched in Quinn [25], and perhaps goes further back.

2 Frobenius algebras

2.1 Algebras, modules, and pairings

Throughout let $\Bbbk$ be a field. All vector spaces will be $\Bbbk$-vector spaces, and all algebras will be $\Bbbk$-algebras; all tensor products are over $\Bbbk$. When we write $\operatorname{Hom}(V, V')$ we will always mean the space of $\Bbbk$-linear maps (even if the spaces happen to be algebras or modules).

2.1.1 $\Bbbk$-algebras. A $\Bbbk$-*algebra* is a $\Bbbk$-vector space A together with two $\Bbbk$-linear maps

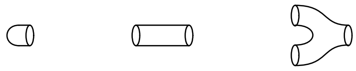
$$\mu : A \otimes A \to A, \qquad \eta : \Bbbk \to A$$

(multiplication and unit map) satisfying the associativity law and the unit law

$$(\mu \otimes \mathrm{id}_A)\mu = (\mathrm{id}_A \otimes \mu)\mu \qquad (\eta \otimes \mathrm{id}_A)\mu = \mathrm{id}_A = (\mathrm{id}_A \otimes \eta)\mu.$$

In other words, a $\Bbbk$-algebra is precisely a monoid in the monoidal category $(\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk)$, cf. 3.3.1.

2.1.2 Graphical language. The maps above are all linear maps between the tensor powers of A — note that $\Bbbk$ is naturally the 0'th tensor power of A. We introduce the following graphical notation for these maps, inspired by Section 1:

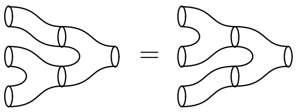


η unit id_A identity μ multiplication

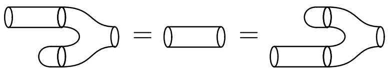
These symbols are meant to have status of formal mathematical symbols, just like the symbols $\to$ or $\otimes$. The symbol corresponding to each $\Bbbk$-linear map $\phi : A^{\otimes m} \to A^{\otimes n}$ has m boundaries on the left (input holes): one for each factor of A in the source, and ordered such that the first factor in the tensor product corresponds to the bottom input hole and the last factor corresponds to the top input hole. If $m = 0$ we simply draw no in-boundary. Similarly there are n boundaries on the right (output holes) which correspond to the target $A^{\otimes n}$, with the same convention for the ordering.

The tensor product of two maps is drawn as the (disjoint) union of the two symbols — one placed above the other, in accordance with our convention for ordering. This mimics $\coprod$ as monoidal operator in the category ***2Cob***. Composition of maps is pictured by joining the output holes of the first figure with the input holes of the second.

Now we can write down the axioms for an algebra like this:

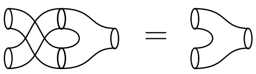


associativity



unit axiom

The canonical twist map $A \otimes A \to A \otimes A$, $v \otimes w \mapsto w \otimes v$ is depicted  , so the property of being a *commutative* algebra reads



2.1.3 Pairings of vector spaces. A bilinear pairing — or just a *pairing* — of two vector spaces V and W is by definition a linear map $\beta : V \otimes W \to \Bbbk$. By the universal property of the tensor product, giving a pairing $V \otimes W \to \Bbbk$ is equivalent to giving a bilinear map $V \times W \to \Bbbk$ (i.e., a map which is linear in each variable.) For this reason we will allow ourselves to write like

$$\begin{array}{rcl} \beta : V \otimes W & \longrightarrow & \Bbbk \\ v \otimes w & \longmapsto & \langle\, v | w \,\rangle \,. \end{array}$$

2.1.4 Nondegenerate pairings. A pairing $\beta : V \otimes W \to \Bbbk$ is called *nondegenerate in the variable* V if there exists a linear map $\gamma : \Bbbk \to W \otimes V$, called a *copairing*, such that the following composition is equal to the identity map of V:

$$\begin{array}{ccccc} V & & (V \otimes W) \otimes V & \xrightarrow{\beta \otimes \mathrm{id}_V} & \Bbbk \otimes V \\ \| & & \| & & \| \\ V \otimes \Bbbk & \xrightarrow{\mathrm{id}_V \otimes \gamma} & V \otimes (W \otimes V) & & V \end{array}$$

Similarly, β is called nondegenerate in the variable W if there exists a copairing $\gamma : \Bbbk \to W \otimes V$, such that the following composition is equal to the identity map of W:

$$\begin{array}{ccccc} \Bbbk \otimes W & \xrightarrow{\gamma \otimes \mathrm{id}_W} & (W \otimes V) \otimes W & & W \\ \| & & \| & & \| \\ W & & W \otimes (V \otimes W) & \xrightarrow{\mathrm{id}_W \otimes \beta} & W \otimes \Bbbk \end{array}$$

These two notions are provisory (but convenient for Lemma 2.1.5 below); the important notion is this: the pairing $\beta : V \otimes W \to \Bbbk$ is simply called *nondegenerate* if it is simultaneously nondegenerate in V and in W. In that case the copairing is unique.

2.1.5 Lemma. *The pairing $\beta : V \otimes W \to \Bbbk$ is nondegenerate in W if and only if W is finite-dimensional and the induced map $W \to V^*$ is injective. (Similarly, nondegeneracy in V is equivalent to finite dimensionality of V plus injectivity of $V \to W^*$.)* □

The finite-dimensionality comes about because the copairing γ singles out an element in $W \otimes V$: this is a *finite* linear combination $\sum_i w_i \otimes v_i$. Now the image of the map $W \to V^* \to W$ is seen to lie in the span of those w_i...

2.1.6 Lemma. *Given a pairing*

$$\begin{aligned} \beta : V \otimes W &\longrightarrow \Bbbk \\ v \otimes w &\longmapsto \langle\, v \,|\, w \,\rangle\,, \end{aligned}$$

between finite-dimensional vector spaces, the following are equivalent.
(i) *β is nondegenerate.*
(ii) *The induced linear map $W \to V^*$ is an isomorphism.*
(iii) *The induced linear map $V \to W^*$ is an isomorphism.*

If we already know for other reasons that V and W are of the same dimension, then nondegeneracy can also be characterised by each of the following a priori weaker conditions:
(ii′) $\langle\, v \,|\, w \,\rangle = 0 \ \forall v \in V \ \Rightarrow \ w = 0$
(iii′) $\langle\, v \,|\, w \,\rangle = 0 \ \forall w \in W \ \Rightarrow \ v = 0$
which is perhaps the most usual definition of nondegeneracy.

2.1.7 Pairings of A-modules. Suppose now M is a right A-module (i.e., a vector space M equipped with a right action $M \otimes A \to M$), and let N be a left A-module. A pairing $\beta : M \otimes N \to \Bbbk$, $x \otimes y \mapsto \langle\, x \,|\, y \,\rangle$ is said to be *associative* when

$$\langle\, xa \,|\, y \,\rangle = \langle\, x \,|\, ay \,\rangle \qquad \text{for every } x \in M, a \in A, y \in N.$$

2.1.8 Lemma. *For a pairing $M \otimes N \to \Bbbk$ as above, the following are equivalent:*
(i) *$M \otimes N \to \Bbbk$ is associative.*
(ii) *$N \to M^*$ is left A-linear.*
(iii) *$M \to N^*$ is right A-linear.*

2.2 Definition and basic properties of Frobenius algebras

Given a linear functional $\Lambda : A \to \Bbbk$, we call the hyperplane $\text{Null}(\Lambda) := \{x \in A \mid x\Lambda = 0\}$ the *nullspace*.

2.2.1 Definition of Frobenius algebra. A *Frobenius algebra* is a $\Bbbk$-algebra A of finite dimension, equipped with a linear functional $\varepsilon : A \to \Bbbk$ whose nullspace contains no nontrivial left ideals. The functional $\varepsilon \in A^*$ is called a *Frobenius form.*

2.2.2 Remarks. The Frobenius form is part of the structure. We will see in 2.2.7 that a given algebra may allow various distinct Frobenius forms. Equivalent characterisations of Frobenius algebras will be given shortly (in 2.2.5 and 2.2.6).

2.2.3 Functionals and associative pairings on A. Every linear functional $\varepsilon : A \to \Bbbk$ (Frobenius or not) determines canonically a pairing $A \otimes A \to \Bbbk$, namely $x \otimes y \mapsto (xy)\varepsilon$. Clearly this pairing is associative (cf. Definition 2.1.7). Conversely, given an associative pairing $A \otimes A \to \Bbbk$, denoted $x \otimes y \mapsto \langle\, x \,|\, y \,\rangle$, a linear functional is canonically determined, namely

$$\begin{array}{rcl} A & \longrightarrow & \Bbbk \\ a & \longmapsto & \langle\, 1_A \,|\, a \,\rangle \;=\; \langle\, a \,|\, 1_A \,\rangle\,. \end{array}$$

This gives a one-to-one correspondence between linear functionals on A and associative pairings. The following lemma is now immediate from 2.1.6:

2.2.4 Lemma. *Let $\varepsilon : A \to \Bbbk$ be a linear functional and let $\langle\ \,|\ \,\rangle$ denote the corresponding associative pairing $A \otimes A \to \Bbbk$. Then the following are equivalent:*
(i) *The pairing is nondegenerate.*
(ii) $\text{Null}(\varepsilon)$ *contains no nontrivial left ideals.*
(iii) $\text{Null}(\varepsilon)$ *contains no nontrivial right ideals.*

In particular this shows that in the definition of Frobenius algebra we could have used right ideals instead of left ideals.

Since the data of a associative pairing and a linear functional completely determine each other as above, we can give the following

2.2.5 Alternative definition of Frobenius algebra. A *Frobenius algebra* is a $\Bbbk$-algebra A of finite dimension, equipped with an associative nondegenerate pairing $\beta : A \otimes A \to \Bbbk$. We call this pairing the *Frobenius pairing.*

This second definition of Frobenius algebras quickly leads to a couple of other characterisations. A nondegenerate pairing $A \otimes A \to \Bbbk$ induces two isomorphisms $A \xrightarrow{\sim} A^*$ (2.1.6) — they do not in general coincide. Associativity means that one of these maps is left A-linear and the other right A-linear (2.1.8) Thus we get a

2.2.6 Third definition of Frobenius algebra. A *Frobenius algebra* is a finite-dimensional $\Bbbk$-algebra A equipped with a left A-isomorphism to its dual. Alternatively (and equivalently) A is equipped with a right A-isomorphism to its dual.

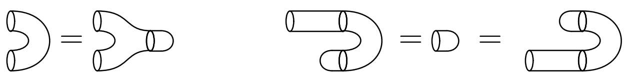
2.2.7 About the choice of structure. The four different versions of Frobenius structure are canonically determined by each other, and therefore we think of them as being one and the same structure. But this structure is not unique: if $\varepsilon : A \to \Bbbk$ is a Frobenius form, and $u \in A$ is invertible, then the functional $x \mapsto (xu)\varepsilon$ is also a Frobenius form. Precisely,

2.2.8 Lemma. *If A is a $\Bbbk$-algebra with Frobenius form ε, then every other Frobenius form on A is given by precomposing ε with multiplication by an invertible element of A. Equivalently, given a fixed left A-isomorphism $\theta : A \xrightarrow{\sim} A^*$, then the elements in A^* which are Frobenius forms are precisely the images of the invertible elements in A.* $\square$

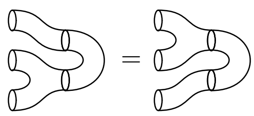
2.2.9 Graphical expression of the Frobenius structure. According to our principles we draw like this:

ε β
Frobenius form Frobenius pairing

The relations $\langle\, x \mid y \,\rangle = (xy)\varepsilon$ and $\langle\, 1_A \mid x \,\rangle = x\varepsilon = \langle\, x \mid 1_A \,\rangle$ of 2.2.3 then get this graphical expression:

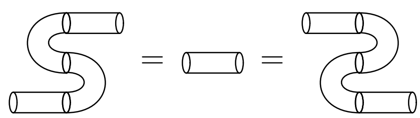


It is trickier to express the axioms which ε and β must satisfy in order to be a Frobenius form and a Frobenius pairing, respectively. The axiom for a Frobenius form $\varepsilon : A \to \Bbbk$ (that its nullspace contains no nonzero ideals) is not expressible in our graphical language because we have no way to represent an ideal. In contrast, it is easy to write down the two axioms for the Frobenius pairing. The associativity condition reads



associativity
of the pairing β

And the nondegeneracy condition is this:

There exists  *such that* 

This is really the crucial property — we will henceforth refer to this as *the snake relation.*

2.3 Examples

In each example, A is assumed to be a $\Bbbk$-algebra of finite dimension over $\Bbbk$.

2.3.1 Algebraic field extensions. Let A be a finite field extension of $\Bbbk$. Since fields have no nontrivial ideals, any nonzero $\Bbbk$-linear map $A \to \Bbbk$ will do as Frobenius form.

2.3.2 Division rings. Let A be a division ring (of finite dimension over $\Bbbk$). Since just like a field a division ring has no nontrivial left ideals (or right ideals), any nonzero linear form $A \to \Bbbk$ will make A into a Frobenius algebra over $\Bbbk$. For example, the Hamiltonians $\mathbb{H}$ is a Frobenius algebra over $\mathbb{R}$.

2.3.3 Matrix rings. The ring $\mathrm{Mat}_n(\Bbbk)$ of all n-by-n matrices over $\Bbbk$ is a Frobenius algebra with the usual trace map as Frobenius form:

$$\begin{array}{rcl} \mathrm{Tr} : \mathrm{Mat}_n(\Bbbk) & \longrightarrow & \Bbbk \\ (a_{ij}) & \longmapsto & \sum_i a_{ii}. \end{array}$$

2.3.4 Group algebras. (See for example Curtis-Reiner [11], § 10.) Let $G = \{t_0, \ldots, t_n\}$ be a finite group written multiplicatively, and with $t_0 = 1$. The *group algebra* $\Bbbk G$ is defined as the set of formal linear combinations $\sum c_i\, t_i$ (where $c_i \in \Bbbk$) with multiplication given by the multiplication in G. It can be made into a Frobenius algebra by taking the Frobenius form to be the functional

$$\begin{array}{rcll} \varepsilon : \Bbbk G & \longrightarrow & \Bbbk & \\ t_0 & \longmapsto & 1 & \\ t_i & \longmapsto & 0 & \text{for } i \neq 0. \end{array}$$

Indeed, the corresponding pairing $g \otimes h \mapsto (gh)\varepsilon$ is nondegenerate since $g \otimes h \mapsto 1$ if and only if $h = g^{-1}$.

2.3.5 The ring of group characters. (See Curtis-Reiner [11], §§ 30–31.) Assume the ground field is $\Bbbk = \mathbb{C}$. Let G be a finite group of order n. A *class function* on G is a function $G \to \mathbb{C}$ which is constant on each conjugacy class; the class functions form a ring denoted $R(G)$. In particular, the characters (traces of representations) are class functions, and in fact every class function is a linear combination of characters. There is a bilinear pairing on $R(G)$ defined by

$$\langle\, \phi \,|\, \psi \,\rangle := \frac{1}{n} \sum_{t \in G} \phi(t)\psi(t^{-1}).$$

Now the *orthogonality relations* (see [11], (31.8)) state that the characters form a orthonormal basis of $R(G)$ with respect to this bilinear pairing, so in particular the pairing is nondegenerate and provides a Frobenius algebra structure on $R(G)$.

2.3.6 Gorenstein rings. (See Eisenbud [15], Ch. 21.) Let A be a commutative artinian local ring with maximal ideal $\mathfrak{m}$. The *socle* of A, denoted $\mathrm{Soc}(A)$, is the annihilator of $\mathfrak{m}$. The ring A is *Gorenstein* if $\mathrm{Soc}(A)$ is a simple A-module, meaning that there are no nontrivial submodules in $\mathrm{Soc}(A)$. Since A is a local ring this just means $\mathrm{Soc}(A) \simeq A/\mathfrak{m}$. Now we claim that if A is Gorenstein then $\mathrm{Soc}(A)$ is contained in every nonzero ideal of A. To establish this, we must show that $\mathrm{Soc}(A)$ lies inside the ideal (x) for every nonzero $x \in \mathfrak{m}$. Since $\mathrm{Soc}(A)$ is a 1-dimensional vector space (over $K := A/\mathfrak{m}$), it is enough to show that the two ideals intersect nontrivially. Now if x is already in $\mathrm{Soc}(A)$, then we are done. Otherwise there exists an element $y \in \mathfrak{m}$ such that xy is nonzero. But then (xy) is an ideal *strictly* smaller than (x) (by Nakayama's lemma). Now repeat the argument with xy in place of x, and continue iteratively. Since A is artinian, we cannot continue forever like that: eventually we arrive at a nonzero element in $\mathrm{Soc}(A)$, and we are done.

Now if A happens to be a finite-dimensional vector space over $\Bbbk$, then it follows that A can be made into a Frobenius algebra simply by taking any linear form which is nonzero on the socle. Indeed, since the nullspace of such a form does not contain the socle, it contains no nontrivial ideals at all.

In fact, conversely, every local Frobenius algebra is Gorenstein.

2.3.7 Jacobian algebras. (See Griffiths-Harris [17], Ch. 5.1.) Let f be a polynomial in n variables, and suppose the zero locus $Z(f) \subset \mathbb{C}^n$ has an isolated singularity at $0 \in \mathbb{C}^n$. Put $f_i := \frac{\partial f}{\partial z_i}$ and let $I = (f_1, \ldots, f_n) \subset \mathcal{O}_0$ (the local ring at the origin). The local ring $\mathcal{O}_0/I$ is called a Jacobian algebra. Since I is generated by n elements which is also its codimension, $\mathcal{O}_0/I$ is a complete intersection ring and in particular Gorenstein. But more interestingly, there is a canonical Frobenius form on it, defined by integrating around the singularity along a real n-ball. Precisely, let $B = \{z \mid f_i(z) = \rho\}$ (for some small $\rho > 0$), and let the functional be the residue

$$\begin{array}{rcl} \mathrm{res}_f : \mathcal{O}_0/I & \longrightarrow & \mathbb{C} \\ g & \longmapsto & \left(\frac{1}{2\pi i}\right)^{2n} \displaystyle\int_B \frac{g(z) \cdot dz_1 \wedge \cdots \wedge dz_n}{f_1(z) \cdots f_n(z)} \end{array}$$

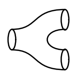
Now *local duality* (see Griffiths-Harris [17], p.659) states that the corresponding bilinear pairing is nondegenerate.

2.3.8 Cohomology rings. (See for example Bott-Tu [9], Ch. 1, or Fulton [16], 24.32.) To be concrete, let X be a compact oriented manifold of dimension n, and let $H^*(X) = \oplus_{i=0}^n H^i(X)$ denote the de Rham cohomology ($H^i(X)$ = closed differentiable i-forms modulo the exact ones). It is a ring under the wedge product. Integration over X (with respect to a chosen volume form) provides a linear map $H^*(X) \to \mathbb{R}$, and *Poincaré duality* states that the corresponding bilinear pairing $H^*(X) \otimes H^*(X) \to \mathbb{R}$ is nondegenerate; precisely, $H^i(X)$ is dual to $H^{n-i}(X)$. Thus, $H^*(X)$ is a Frobenius algebra over $\mathbb{R}$.

In fact, if X is connected then $H^*(X)$ is a (graded-commutative) Gorenstein ring (2.3.6): the maximal ideal is $\bigoplus_{i>0} H^i(X)$, and the socle is $H^n(X) \simeq \mathbb{R}$. By *graded-commutative* we mean that classes of odd degree anti-commute: given $\alpha \in H^p(X)$ and $\beta \in H^q(X)$ then $\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha$.

2.4 Frobenius algebras and comultiplication

2.4.1 Coalgebras. Recall that a coalgebra over $\Bbbk$ is a vector space A together with two $\Bbbk$-linear maps

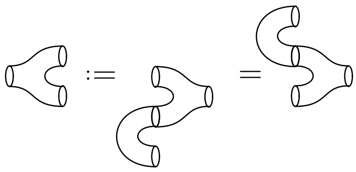
 

$$\delta : A \to A \otimes A, \qquad \varepsilon : A \to \Bbbk$$

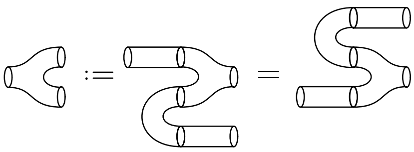
satisfying axioms dual to the algebra axioms (2.1.1). The map δ is called *comultiplication*, and $\varepsilon : A \to \Bbbk$ is called the *counit*.

Our goal is to show that a Frobenius algebra (A, ε) has a natural coalgebra structure for which ε is the counit. The idea is to use the copairing  to turn around an input hole of the multiplication:

2.4.2 Comultiplication. Define a comultiplication map $\delta : A \to A \otimes A$ by



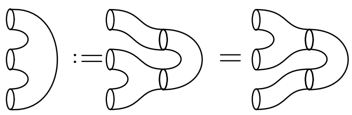
Here and in the sequel we suppress identity maps. What is meant is actually

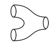


That the two expressions in the definition agree will follow from the next lemma. First some notation:

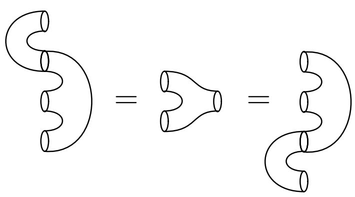
2.4.3 The three-point function $\phi : A \otimes A \otimes A \to \Bbbk$ is defined by

$$\phi := (\mu \otimes \mathrm{id}_A)\beta = (\mathrm{id}_A \otimes \mu)\beta,$$

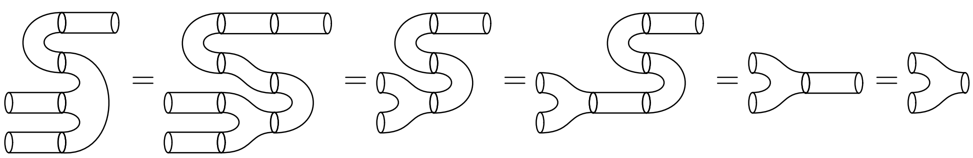


Associativity of β says that the two expressions coincide. Conversely, using the snake relation we can express  in terms of the three-point function:

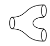
2.4.4 Lemma. *We have*

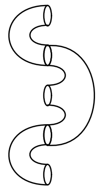


Proof. (As illustration, all the identity maps are provided in this proof. In the sequel they will be suppressed).

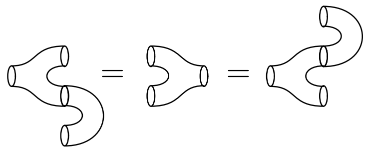


Here the first step was to use the definition of the three-point function; then remove some identity maps and insert one; next, apply the the snake relation; and finally remove an identity map. ☐

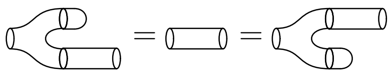
Now it is clear that the two expression of the definition of  agree: using this lemma they are both seen to be equal to



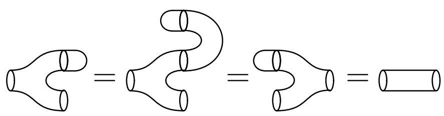
2.4.5 Multiplication in terms of comultiplication. Conversely, turning some holes back again, using β, and then using the snake relation, we also get the relations dual to 2.4.2:



2.4.6 Lemma. *The Frobenius form ε is counit for δ:*

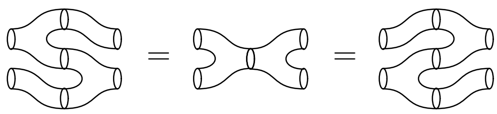


Proof. Suppressing the identity maps, write

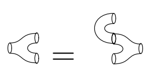


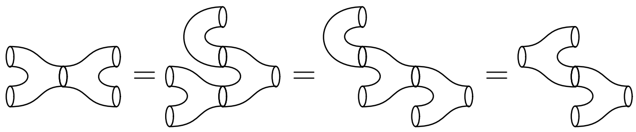
Here the first step was to use the expression 2.2.9 for  . The next step was to use relation 2.4.5. Finally we used that  is neutral element for the multiplication (cf. 2.1.2). (The right-hand equation is analogous.) ☐

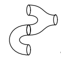
2.4.7 Lemma. *The comultiplication satisfies the following compatibility condition with respect to the multiplication, called the Frobenius relation.*



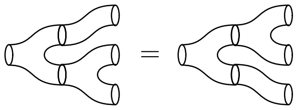
The right-hand equation expresses right A-linearity of δ; the left-hand equation expresses left A-linearity.

Proof. For the left-hand equation, use ; then use associativity, and finally use the relation back again:

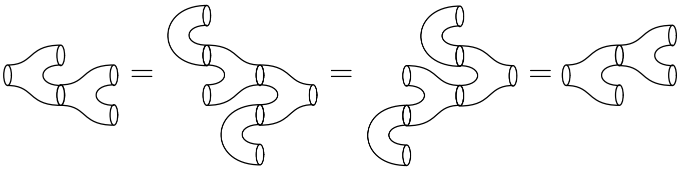


The right-hand equation is obtained using  = . □

2.4.8 Lemma. *The comultiplication is coassociative:*

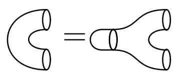


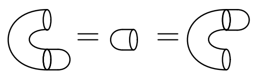
Proof. Use the definition of δ (2.4.2); then the associativity, and finally the definition again:



□

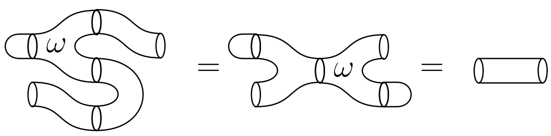
2.4.9 Remark. The relation between the copairing and the unit is analogous (dual) to the relation between the Frobenius form and the Frobenius pairing (relation 2.2.9):



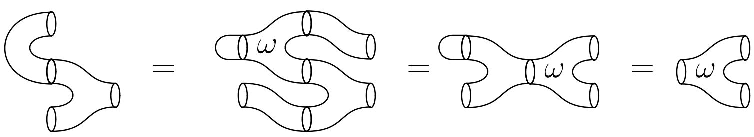


2.4.10 Proposition. *Given a Frobenius algebra (A, ε), there exists a unique comultiplication δ whose counit is ε and which satisfies the Frobenius relation, and this comultiplication is coassociative.*

Proof. We have already constructed such a comultiplication (2.4.2, 2.4.6, 2.4.7), and established its coassociativity (2.4.8). The uniqueness is a consequence of the fact that the copairing corresponding to a nondegenerate pairing is unique. In detail: suppose that ω is another comultiplication with counit ε and which satisfies the Frobenius relation. Putting caps on the upper input hole and the lower output hole of (the left-hand part of) the Frobenius relation we see that $\eta\omega$ satisfies the snake equation:



by the unit and counit axioms. So by the uniqueness of copairing we have $\eta\omega = \gamma$. Using this, if instead we put only the cap η on, then we get

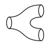
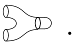


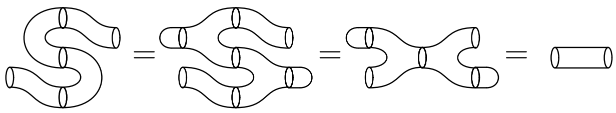
That is: ω is nothing but μ with an input hole turned around, just like δ was defined. $\square$

The reason why the relation of 2.4.7 is called the Frobenius condition is that it characterises Frobenius algebras, as the next result shows. In fact, not only it characterises Frobenius algebras among the associative algebras of finite dimension, but also among general vector spaces equipped with unitary multiplication. Precisely,

2.4.11 Proposition. *Let A denote a vector space equipped with a multiplication map $\mu : A \otimes A \to A$ with unit $\eta : \Bbbk \to A$, and a comultiplication $\delta : A \to A \otimes A$ with counit $\varepsilon : A \to \Bbbk$, and suppose the Frobenius relation holds. Then*

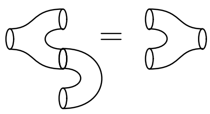
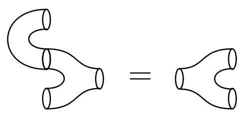
(i) *The vector space A is of finite dimension.*
(ii) *The multiplication μ is associative (and thus A is a finite-dimensional $\Bbbk$-algebra).*
(iii) *The counit ε is a Frobenius form (and thus (A, ε) is a Frobenius algebra).*

Proof. We use the graphical notation  $= \mu$,  $= \eta$,  $= \delta$,  $= \varepsilon$. Set $\beta := \mu\varepsilon$, that is:  $=$  . We will show that β is nondegenerate, i.e., establish the snake relation, with $\gamma = \eta\delta$. Put caps on the left-hand part of the Frobenius relation like this:

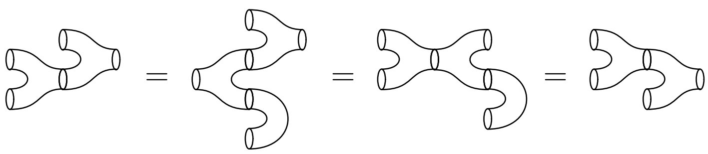


by the unit and counit axioms. This is the left-hand part of the snake relation; similarly, the right-hand side of the Frobenius relation gives the right-hand side of the snake relation, so β is nondegenerate. This in particular implies that A is of finite dimension (cf. 2.1.5).

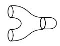
To get associativity, note that if we put only one cap on the Frobenius relation (left-hand relation) we get these two identities:

 and 

Now we can write

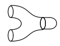


So  is associative.

Finally, since  is associative, clearly the pairing  $=$  is associative as well, so (A, β) is a Frobenius algebra. $\square$

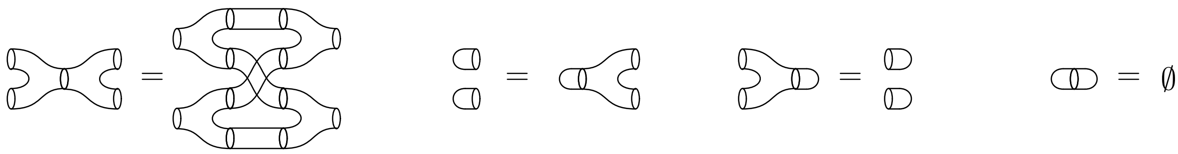
2.4.12 Uniqueness of the comultiplication. Since the comultiplication is defined explicitly in terms of the multiplication and the copairing, it follows from the uniqueness of the copairing that also the comultiplication is unique.

2.4.13 Historical remarks. The characterisation of Frobenius algebras in terms of comultiplication goes back (at least) to Lawvere [21], (1967) where it is a parenthetical remark at the end of the paper. In a very general categorical context (which we will take up in Section 3, notably 3.4.4) he describes a *Frobenius standard construction* (standard construction meaning monad) as being a combined monoid/comonoid object with this compatibility requirement (which we reproduce in graphical language):

These are 2.4.2 and 2.4.5 above. The Frobenius relation is an immediate consequence, cf. 2.4.7. The nondegenerate pairing  =  is mentioned explicitly, but the Frobenius relation is not.

The first explicit mention of the Frobenius relation, and a proof of 2.4.11, were given in 1991 by Quinn [25], unaware of [21]. Independently, Abrams gave the commutative case of the new characterisation in [1] (1995), and the noncommutative case appeared in [2] (1998).

2.4.14 Digression on bialgebras. Bialgebras are also algebras which are at the same time coalgebras — here the compatibility condition is different however: the comultiplication is required to be an algebra homomorphism (cf. Kassel [19], Ch. III). Bialgebras are not necessarily of finite dimension. The graphical version of the bialgebra axioms are:



— none of which hold in a Frobenius algebra. In fact, it is not difficult to prove that

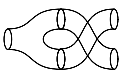
If $A \neq \Bbbk$ is a bialgebra of finite dimension, with structure maps $\eta, \mu, \delta, \varepsilon$ as above, then ε is not a Frobenius form.

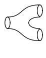
Hopf algebras are particular examples of bialgebras. It is proven in Sweedler's book [27], Ch. V that finite-dimensional Hopf algebras admit Frobenius structure (this generalises Example 2.3.4). It should be stressed that the Frobenius comultiplication is then distinct from the bialgebra comultiplication.

2.4.15 Duality. It is particularly clear from the pictures that there is a complete symmetry between μ and η on one side and δ and ε on the other side. As a consequence, if (A, ε) is a Frobenius algebra then the dual vector space A^*, equipped with the linear form $\varepsilon^* = \eta$ is canonically a Frobenius algebra as well.

2.4.16 Proposition. *The comultiplication of a Frobenius algebra is cocommutative if and only if the multiplication is commutative.*

Proof. (Sketch.) Suppose the multiplication is commutative. The proof amounts to checking that this map



has ε as counit and satisfies the Frobenius relation. (Showing this relies on the naturality of the twist map.) Then by uniqueness of the comultiplication, this map must coincide with  . The converse implication follows from duality. $\square$

2.4.17 The category of Frobenius algebras. A *Frobenius algebra homomorphism* $\phi : (A, \varepsilon) \to (A', \varepsilon')$ between two Frobenius algebras is an algebra homomorphism which is at the same time a coalgebra homomorphism. In particular it preserves the Frobenius form, in the sense that $\varepsilon = \phi\varepsilon'$. Let $\boldsymbol{FA}_{\Bbbk}$ denote the category of Frobenius algebras over $\Bbbk$ and Frobenius algebra homomorphisms.

2.4.18 Lemma. *A Frobenius homomorphism is always invertible. (In other words, the category $\boldsymbol{FA}_{\Bbbk}$ is a groupoid.)*

Proof. If an algebra homomorphism $\phi : (A, \varepsilon) \to (A', \varepsilon')$ between two Frobenius algebras satisfies $\varepsilon = \phi\varepsilon'$, then necessarily it is injective — indeed, the kernel of ϕ is an ideal contained in $\mathrm{Null}(\varepsilon)$. Now if ϕ is furthermore a Frobenius homomorphism, then by duality, the dual map is also a Frobenius algebra homomorphism, and thus injective. Since A is a finite-dimensional vector space this implies that ϕ is surjective. □

3 Monoids and Frobenius structures

3.1 Monoidal categories

The category of vector spaces $\boldsymbol{Vect}_{\Bbbk}$, with tensor product $\otimes$ and 'neutral space' $\Bbbk$ is the key example of a *monoidal category*, also called tensor category. Our second example is of course the category of cobordisms $\boldsymbol{2Cob}$ with disjoint union $\coprod$ and empty manifold $\emptyset$. Both these categories are furthermore equipped with a symmetry structure — they are symmetric monoidal categories.

At first, the tensor product is an operation which to two vector spaces associates a new one. At a second level, once we have this structure, it is used as background for defining structures on individual vector spaces: we defined a multiplication map as a map $A \otimes A \to A$. Algebra structure (in the category of vector spaces) has its generalisation in the notion of *monoid* in a monoidal category, and as a corresponding generalisation of *Frobenius algebra* we will introduce the notion of *Frobenius object* in a monoidal category.

Let us briefly recall the basic notions of monoidal categories. For more details see Mac Lane [22] or Kassel [19].

3.1.1 Monoidal categories. A *(strict) monoidal category* is a category $\boldsymbol{V}$ together with two *functors*

$$\mu : \boldsymbol{V} \times \boldsymbol{V} \to \boldsymbol{V}, \qquad \eta : \mathbf{1} \to \boldsymbol{V}$$

satisfying the associativity axiom and the neutral object axiom. Precisely we require the following three identities of functors:

$$(\mu \times \mathrm{id}_{\boldsymbol{V}})\mu = (\mathrm{id}_{\boldsymbol{V}} \times \mu)\mu \qquad (\eta \times \mathrm{id}_{\boldsymbol{V}})\mu = \mathrm{id}_{\boldsymbol{V}} = (\mathrm{id}_{\boldsymbol{V}} \times \eta)\mu.$$

In order to be able to use these structures as background for further operations, we need infix notation, so we write

$$\begin{array}{rcl} \boldsymbol{V} \times \boldsymbol{V} & \xrightarrow{\mu} & \boldsymbol{V} \\ (X, Y) & \longmapsto & X \Box Y \\ (f, g) & \longmapsto & f \Box g. \end{array}$$

Also, let I denote the object which is the image of $\eta : \mathbf{1} \to \boldsymbol{V}$ — it is the 'neutral object'. We refer to a monoidal category by specifying the triple $(\boldsymbol{V}, \Box, I)$.

3.1.2 Monoidal functors. A *(strict) monoidal functor* between two (strict) monoidal categories $(\boldsymbol{V}, \Box, I)$ and $(\boldsymbol{V}', \Box', I')$ is a functor $F : \boldsymbol{V} \to \boldsymbol{V}'$ that commutes with all the structure. Precisely, these two diagrams are required to commute:

$$\begin{tikzcd}
\boldsymbol{V} \times \boldsymbol{V} \arrow[r, "F \times F"] \arrow[d, "\mu"'] & \boldsymbol{V}' \times \boldsymbol{V}' \arrow[d, "\mu'"] \\
\boldsymbol{V} \arrow[r, "F"'] & \boldsymbol{V}'
\end{tikzcd}
\qquad
\begin{tikzcd}
\boldsymbol{V} \arrow[r, "F"] & \boldsymbol{V}' \\
\mathbf{1} \arrow[u, "\eta"] \arrow[r, phantom, "="] & \mathbf{1} \arrow[u, "\eta'"']
\end{tikzcd}$$

3.1.3 Not-necessarily-strict monoidal categories. There is a weaker notion of monoidal categories, usually simply called *monoidal categories* where the axioms hold only up to coherent isomorphisms instead of holding strictly. These are the monoidal categories that mostly occur in practice — both $(\boldsymbol{Vect}, \otimes, \Bbbk)$ and the non-skeletal version of $(\boldsymbol{2Cob}, \coprod, \emptyset)$ are actually non-strict — but we will always pretend that they are strict. This abuse is justified by Mac Lane's coherence theorem which *grosso modo* asserts that making identifications along the said coherent isomorphisms does not lead to contradictions. See Mac Lane [22], Ch. XI.

3.1.4 Categories with products or coproducts. Let $\boldsymbol{C}$ be a category which admits products (resp. coproducts) $\Box$, and let I denote the terminal (resp. initial) object. Then $(\boldsymbol{C}, \Box, I)$ is a monoidal category.

This observation provides a long list of important monoidal categories, including for example $(\boldsymbol{Set}, \times, 1)$ and $(\boldsymbol{Set}, \coprod, \emptyset)$.

3.1.5 Symmetric monoidal categories. We are going to use monoidal categories as context for defining monoids. In order to be able to talk about *commutative* monoids, we need a notion of *twist map* in the category, generalising the twist map we know in the category of sets. Intuitively this new twist map will be 'interchange of factors'. When such a twist map is specified with properties similar to those of the twist map in $\boldsymbol{Set}$ then we have a *symmetric monoidal category.* (Specifying twist maps with slightly weaker properties leads to the notion of braided monoidal categories, which we will not need in this text, but which are important in knot theory (see Kassel [19]).)

A (strict) monoidal category $(\boldsymbol{V}, \Box, I)$ is called a *symmetric monoidal category* if for each pair of objects X, Y there is given a *twist map*

$$\tau_{X,Y} : X \Box Y \to Y \Box X$$

subject to the following four axioms:

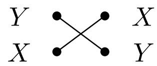
(i) The maps are natural.
(ii) We have $\tau_{X,Y} \tau_{Y,X} = \mathrm{id}_{X \Box Y}$.
(iii) For every object X, these two diagrams commute:

$$\begin{tikzcd}
X \Box I \arrow[rr, "\tau_{X,I}"] \arrow[dr] & & I \Box X \arrow[dl] \\
& X &
\end{tikzcd}
\qquad
\begin{tikzcd}
I \Box X \arrow[rr, "\tau_{I,X}"] \arrow[dr] & & X \Box I \arrow[dl] \\
& X &
\end{tikzcd}$$

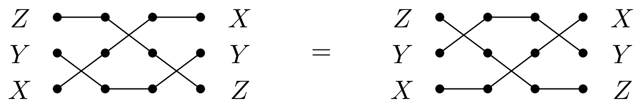
(iv) For every triple of objects X, Y, Z, these two diagrams commute:

$$\begin{tikzcd}[column sep=small] X\Box Y\Box Z \arrow[rr, "\tau_{X,Y\Box Z}"] \arrow[dr, "\tau_{X,Y}\Box\,\mathrm{id}_Z"'] & & Y\Box Z\Box X \\ & Y\Box X\Box Z \arrow[ur, "\mathrm{id}_Y\Box\,\tau_{X,Z}"'] & \end{tikzcd} \qquad \begin{tikzcd}[column sep=small] X\Box Y\Box Z \arrow[rr, "\tau_{X\Box Y,Z}"] \arrow[dr, "\mathrm{id}_X\Box\,\tau_{Y,Z}"'] & & Z\Box X\Box Y \\ & X\Box Z\Box Y \arrow[ur, "\tau_{X,Z}\Box\,\mathrm{id}_Y"'] & \end{tikzcd}$$

We picture the twist map like this:



Conditions (iii) and (iv) imply the 'symmetric group relation':



(3.1.6)

The collection of twist maps (one for each pair of objects) is a *structure* that must be specified (not a property of a given monoidal category), so a symmetric monoidal category is a quadruple $(\boldsymbol{V}, \Box, I, \tau)$. In the monoidal category of vector spaces $(\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk)$ we will always refer to the canonical symmetry σ given by interchange of factors:

$$\begin{aligned} V \otimes W &\longrightarrow W \otimes V \\ v \otimes w &\longmapsto w \otimes v. \end{aligned}$$

So we write $(\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk, \sigma)$.

3.1.7 Canonical symmetries. If $(\boldsymbol{V}, \Box, I)$ is a monoidal category given by product or coproduct (as in 3.1.4) then the axioms for a symmetry are so strong that in fact there is a unique symmetric structure on $(\boldsymbol{V}, \Box, I)$, namely the interchange of factors. This follows from the naturality of the twist with respect to the projection maps from the product (together with axiom (iii)).

As an example of a monoidal category which admits more than one significant symmetric structure we have

3.1.8 The category of graded vector spaces. A *graded vector space* is a direct sum $V = \bigoplus_{n\in\mathbb{Z}} V_n$, and a graded linear map is one that respects the grading. The tensor product of two graded vector spaces V and W is again a graded vector space, with grading $(V \otimes W)_n = \bigoplus_{p+q=n} (V_p \otimes W_q)$. The ground field $\Bbbk$ is a graded vector space concentrated in degree 0. So we have a monoidal category $(\boldsymbol{grVect}_{\Bbbk}, \otimes, \Bbbk)$ of graded vector spaces. Now of course we have the canonical symmetry σ just as above: $v \otimes w \mapsto w \otimes v$. But there is also another important possibility for defining a twist map, namely via Koszul's sign change:

$$v \otimes w \mapsto (-1)^{pq}\, w \otimes v$$

where $\deg(v) = p$ and $\deg(w) = q$. One checks that this symmetry κ does indeed satisfy the axioms, making $(\boldsymbol{grVect}_{\Bbbk}, \otimes, \Bbbk, \kappa)$ into a symmetric monoidal category.

3.1.9 Symmetric monoidal functors. A monoidal functor F from a symmetric monoidal category $(\boldsymbol{V}, \Box, I, \tau)$ to another $(\boldsymbol{V}', \Box', I', \tau')$, is called a *symmetric monoidal functor* if it preserves the symmetry maps. That is, for every pair of objects X, Y in $\boldsymbol{V}$ we have

$$\tau_{X,Y}F = \tau'_{XF,YF}.$$

3.1.10 Monoidal natural transformations. Let $(\boldsymbol{V}, \Box, I)$ and $(\boldsymbol{V}', \Box', I')$ be two monoidal categories, and let

$$\boldsymbol{V} \underset{F}{\overset{G}{\rightrightarrows}} \boldsymbol{V}'$$

be two monoidal functors. For a natural transformation $u : F \Rightarrow G$ we denote by $u_X : XF \to XG$ its component on each object X in $\boldsymbol{V}$. The natural transformation u is called a *monoidal natural transformation* if for every two objects X, Y in $\boldsymbol{V}$ we have $u_X \Box' u_Y = u_{X\Box Y}$, and also $u_I = \mathrm{id}_{I'}$.

There is a category $\boldsymbol{MonCat}(\boldsymbol{V}, \boldsymbol{V}')$ whose objects are the monoidal functors from $\boldsymbol{V}$ to $\boldsymbol{V}'$, and whose arrows are the monoidal natural transformations between such functors.

In case $\boldsymbol{V}$ and $\boldsymbol{V}'$ are both symmetric, it makes sense to consider only symmetric monoidal functors: there is a category $\boldsymbol{SymMonCat}(\boldsymbol{V}, \boldsymbol{V}')$ whose objects are the symmetric monoidal functors from $\boldsymbol{V}$ to $\boldsymbol{V}'$, and whose arrows are the monoidal natural transformations between such functors.

By definition, a *linear representation* of a symmetric monoidal category $(\boldsymbol{V}, \Box, I, \tau)$ is a symmetric monoidal functor $(\boldsymbol{V}, \Box, I, \tau) \to (\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk, \sigma)$. So the set of all linear representations of $\boldsymbol{V}$ are the objects of a category which we denote

$$\boldsymbol{Repr}_{\Bbbk}(\boldsymbol{V}) := \boldsymbol{SymMonCat}(\boldsymbol{V}, \boldsymbol{Vect}_{\Bbbk}).$$

3.1.11 TQFTs. By definition 1.2.9, a topological quantum field theory is a symmetric monoidal functor from $\boldsymbol{nCob}$ to $\boldsymbol{Vect}_{\Bbbk}$. Such functors form the objects of a category

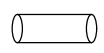
$$\boldsymbol{nTQFT}_{\Bbbk} = \boldsymbol{Repr}_{\Bbbk}(\boldsymbol{nCob}) = \boldsymbol{SymMonCat}(\boldsymbol{nCob}, \boldsymbol{Vect}_{\Bbbk});$$

the arrows being the monoidal natural transformations.

3.1.12 2D TQFTs and Frobenius algebras. So a 2-dimensional TQFT is a linear representation of the symmetric monoidal category $(\boldsymbol{2Cob}, \coprod, \emptyset, T)$. This is a category we fully control, because we are given a presentation of it in terms of generators and relations. Recall from 1.3 that the objects of $\boldsymbol{2Cob}$ are $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots\}$ where $\mathbf{n}$ is the disjoint union of n circles, and that the generating arrows are  ,  ,  ,  ,  , and that the relations which hold (cf. 1.3.10–1.3.14) express topological equivalence of different ways of cutting surfaces into pieces.

In general, a monoidal functor is determined completely by its values on the generators of the source category. In our case, to give a symmetric monoidal functor $\mathscr{A} : \boldsymbol{2Cob} \to \boldsymbol{Vect}_{\Bbbk}$, we must specify a vector space A as image of $\mathbf{1}$, and a linear map for each of the generators. The fact that the functor is monoidal implies in particular that the image of $\mathbf{2}$ is $A \otimes A$, and so on. To ease the notation, put $A^n := A \otimes \ldots \otimes A$ (with n factors). The fact that $\mathscr{A}$ is a *symmetric* monoidal functor means that the image of  must be the usual twist for the tensor product, so the following is automatic once we have fixed the vector space A:

$$\begin{array}{rcl} \textbf{\textit{2Cob}} & \longrightarrow & \textbf{\textit{Vect}}_{\Bbbk} \\ \mathbf{1} & \longmapsto & A \\ \mathbf{n} & \longmapsto & A^n \end{array}$$

 $\longmapsto [\mathrm{id}_A : A \to A]$

 $\longmapsto [\sigma : A^2 \to A^2]$

Let the images of the generators be denoted like this:

$$\textbf{\textit{2Cob}} \longrightarrow \textbf{\textit{Vect}}_{\Bbbk}$$

 $\longmapsto [\eta : \Bbbk \to A]$

 $\longmapsto [\mu : A^2 \to A]$

 $\longmapsto [\varepsilon : A \to \Bbbk]$

 $\longmapsto [\delta : A \to A^2]$

So A is a vector space equipped with certain linear maps among its tensor powers; now the relations that hold in ***2Cob*** translate into relations among these linear maps. With the graphical notation used in Section 2 we anticipated this comparison: it is easy to see that the relations translate exactly into the axioms for a commutative Frobenius algebra. So in conclusion, given a 2D TQFT $\mathscr{A}$, then the image vector space $A = \mathbf{1}\mathscr{A}$ is a commutative Frobenius algebra.

Conversely, starting with a commutative Frobenius algebra (A, ε) (whose multiplication is denoted μ, etc.), we can construct a TQFT $\mathscr{A}$ by using the above description as definition. In other words, $\mathscr{A}$ is defined by sending $\mathbf{1} \mapsto A$ (the underlying vector space),  $\mapsto \mu$, etc. For this to make sense we must check that the relations are respected, but again, since the relations in ***2Cob*** correspond precisely to the axioms for a commutative Frobenius algebra, this is automatic, so the symmetric monoidal functor $\mathscr{A}$ is indeed well-defined. It is clear that these two constructions are inverse to each other, so we have established a one-to-one correspondence between 2D TQFTs and commutative Frobenius algebras.

This correspondence works also for arrows. The arrows in ***2TQFT*** are the monoidal natural transformations. Given two TQFTs $\mathscr{A}$, $\mathscr{B}$, i.e., two symmetric monoidal functors ***2Cob*** $\to$ ***Vect***$_{\Bbbk}$, then a natural transformation u between them consists of linear maps $A^n \to B^n$ for each $n \in \mathbb{N}$. That u is a *monoidal* natural transformation means that the map $A^n \to B^n$ is just the n'th tensor power of the map $A^1 \to B^1$, so u is determined completely by specifying this linear map. The *naturality* of u means that all these maps are compatible with arrows in ***2Cob***. Now every arrow in ***2Cob*** is built up from the generators, so naturality boils down to four commutative diagrams which amounts precisely to the statement that $A \to B$ is simultaneously a $\Bbbk$-algebra homomorphism and a $\Bbbk$-coalgebra homomorphism, and thus a Frobenius algebra homomorphism (2.4.17).

Conversely, given a Frobenius algebra homomorphism between two commutative Frobenius algebras, we can use the above arguments in the reverse direction to construct a monoidal natural transformation between the TQFTs corresponding to A and B.

All together we have proven this:

3.1.13 Theorem. *There is a canonical equivalence of categories*

$$\mathbf{2TQFT}_{\Bbbk} \simeq \mathbf{cFA}_{\Bbbk}.$$

(If we regard $\mathbf{2TQFT}$ as the category of representations of the *skeletal* version of $\mathbf{2Cob}$, then in fact we have an *isomorphism* of categories. See 3.3.7 for further discussion.) Note that $\mathbf{2TQFT}_{\Bbbk}$ is a groupoid, cf. 2.4.18.

3.1.14 Question about symmetry — a non-example. It is natural to ask whether our requirement that the monoidal functor defining a TQFT be symmetric is really necessary — it might turn out to be symmetric automatically?!

Here is an example which illuminates this question, and in particular shows that the symmetry of the functor is not automatic. Consider a *graded-commutative* Frobenius algebra H, for example the cohomology ring of a smooth compact manifold (cf. 2.3.8). Now define a non-symmetric TQFT by sending $\mathbf{1} \mapsto H$. Concerning the generators for $\mathbf{2Cob}$: send them to multiplication and unit, comultiplication and counit, just as usual, but send the twist cobordism to Koszul's sign-changed twist (cf. 3.1.8)

$$a \otimes b \mapsto (-1)^{pq}\, b \otimes a,$$

where $\deg(a) = p$ and $\deg(b) = q$. This works! Even though H is not a *commutative* Frobenius algebra in the usual sense, there is no contradiction because the twist cobordism is not sent to the usual twist!

Unfortunately, our current definition of TQFT is too narrow to include this example, and simply dropping the symmetry requirement in the definition would not be satisfactory for two reasons: First, the symmetries exist and it would be morally wrong just to deny them; there is a structure, and thus our functors should be required to preserve it. Second, Theorem 3.1.13 would simply be wrong if there were no symmetry requirement!

The problem of the example is actually the target category $(\mathbf{Vect}_{\Bbbk}, \otimes, \Bbbk, \sigma)$. If we put H into another monoidal category, like for example $(\mathbf{grVect}, \otimes, \Bbbk, \kappa)$, then the monoidal functor *is* symmetric! Also, from the viewpoint of quantum theory, there is no reason for favouring precisely the category $(\mathbf{Vect}_{\Bbbk}, \otimes, \Bbbk, \sigma)$ — in fact, in the list of examples that M. Atiyah gives in [3], half of the examples really use mod-2 graded vector spaces instead of plainly vector spaces — such creatures abound in quantum physics…

3.1.15 Historical remarks. The observation that 2-dimensional TQFTs are essentially the same thing as commutative Frobenius algebras was first made by R. Dijkgraaf in his Ph.D. thesis [12]. More precise proofs have been given by Dubrovin [14], Quinn [25] Sawin [26], and Abrams [1] — this is at the same time the chronological order and the order according to the amount of details presented. However, all these sources are silent on the questions of symmetry…

3.1.16 What's next? The rest of these notes is devoted to placing the above theorem in its proper context. We will show that it is just a variation of a more basic result: there is a monoidal category Δ (the simplex category) which is quite similar to $\mathbf{2Cob}$ (in fact it is a subcategory) with the property that giving a monoidal functor from Δ to $\mathbf{Vect}_{\Bbbk}$ is the same as giving an algebra:

$$\mathbf{MonCat}(\Delta, \mathbf{Vect}_{\Bbbk}) \simeq \mathbf{Alg}_{\Bbbk}.$$

This in turn is just a special case of a general result which states that monoidal functors from Δ to any monoidal category $\boldsymbol{V}$ correspond to monoids in $\boldsymbol{V}$ (cf. Mac Lane [22]).

This result also has a variant for Frobenius algebras: we will define a notion of *(commutative) Frobenius object* in a general (symmetric) monoidal category, such that a (commutative) Frobenius object in $\boldsymbol{Vect}_{\Bbbk}$ is precisely a (commutative) Frobenius algebra. We will see that ***2Cob*** does for commutative Frobenius objects what Δ does for monoids: every commutative Frobenius object (in any symmetric monoidal category) arises as the image of a unique symmetric monoidal functor from ***2Cob***. In other words, ***2Cob*** *is the free symmetric monoidal category on a commutative Frobenius object* (Theorem 3.4.14).

Our main motivation for striving for this generality is to place the above theorem in its natural context, isolating those properties of $\boldsymbol{Vect}_{\Bbbk}$ that we used. Example 3.1.14 hints at the importance of this generality.

3.2 The simplex categories Δ and Φ

3.2.1 The category Δ of finite ordinals — also called the simplex category. For each $n \in \mathbb{N}$, let $\mathbf{n}$ denote the ordered set $\{0, \dots, n-1\}$ — these are called finite ordinals. (So $\mathbf{0}$ is the empty set and $\mathbf{1}$ is the one-element set $\{0\}$.) The category Δ has as objects the finite ordinals

$$\Delta = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots\},$$

and as arrows the order preserving maps between those sets. In other words, the arrows are functions $f : \mathbf{m} \to \mathbf{n}$ such that $i \leq j$ in $\mathbf{m}$ implies $if \leq jf$ in $\mathbf{n}$. This is a skeleton of the category ***FinOrd*** of finite ordered sets.

Δ is a (strict) monoidal category under *ordinal sum*, which to two ordinals $\mathbf{m}$ and $\mathbf{n}$ associates the ordinal $\mathbf{m} + \mathbf{n}$ corresponding to the natural numbers sum. The set $\mathbf{m} + \mathbf{n}$ already comes with an order, but we must specify how the two original sets inject into it: these (order preserving) injections are given by

$$\begin{aligned} \mathbf{m} &\longrightarrow \mathbf{m} + \mathbf{n} \\ i &\longmapsto i \end{aligned} \qquad\qquad \begin{aligned} \mathbf{n} &\longrightarrow \mathbf{m} + \mathbf{n} \\ j &\longmapsto m + j. \end{aligned}$$

3.2.2 Graphical version of Δ. The *objects* of Δ are interpreted as finite sets of dots arranged in a column

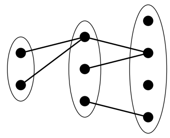
$$\Delta = \{\emptyset,\ \bullet\,,\ \begin{smallmatrix}\bullet\\ \bullet\end{smallmatrix}\,,\ \begin{smallmatrix}\bullet\\ \bullet\\ \bullet\end{smallmatrix}\,, \dots\}$$

An *arrow* from one column to another is a collection of strands starting at the dots in the source column and ending at dots in the target column, subject to the following two rules:

(i) for each dot in the source column there is exactly one strand coming out (and going to the target column)

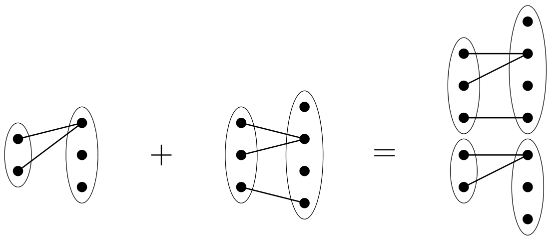
(ii) the strands are not allowed to cross each other, but they are allowed to merge — in other words, two or more strands may share a single dot in the target.

The *composition* of two arrows is given by joining the input ends of the second collection of strands to the output ends of the first collection of strands. Here is an example of a composition $\mathbf{2} \to \mathbf{3} \to \mathbf{4}$:



For each object there is an *identity arrow* given by taking a strand from each dot to itself. It is easy to check that these definitions satisfy the axioms for a category, since the first rule amounts to saying that f is a *function* from the set of dots in the source column to the set of dots in the target column. The fact that the dots are arranged in columns is just to say that the sets are ordered, and rule (ii) says exactly that the functions are order preserving.

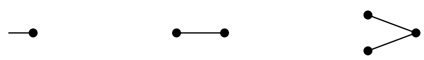
The monoidal structure is easy to grasp graphically:



3.2.3 Δ is not symmetric. Symmetry would mean that for every pair of objects $\mathbf{m}, \mathbf{n}$ we should have an invertible arrow $\mathbf{m} + \mathbf{n} \to \mathbf{n} + \mathbf{m}$, natural with respect to all other arrows. Now the only invertible arrows are the identities, and they fail to be natural in this sense. As an easy example, the identity map $\mathbf{2} + \mathbf{1} \to \mathbf{1} + \mathbf{2}$ is not natural with respect to the maps $\mu : \mathbf{2} \to \mathbf{1}$ and $\mathrm{id}_{\mathbf{1}} : \mathbf{1} \to \mathbf{1}$ — this diagram does *not* commute:

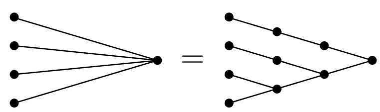
$$\begin{tikzcd}
\mathbf{2}+\mathbf{1} \arrow[r, "\mathrm{id}_{\mathbf{3}}"] \arrow[d, "\mu+\mathrm{id}_{\mathbf{1}}"'] & \mathbf{1}+\mathbf{2} \arrow[d, "\mathrm{id}_{\mathbf{1}}+\mu"] \\
\mathbf{1}+\mathbf{1} \arrow[r, "\mathrm{id}_{\mathbf{2}}"'] & \mathbf{1}+\mathbf{1}
\end{tikzcd}$$

3.2.4 Generators and relations for Δ. Clearly $\mathbf{1}$ is terminal object in Δ. Let the arrows be denoted $\mu^{(n)} : \mathbf{n} \to \mathbf{1}$. Now we claim that $(\Delta, +, \mathbf{0})$ is generated (monoidally) by $\mu^{(0)} : \mathbf{0} \to \mathbf{1}$ and $\mu^{(2)} : \mathbf{2} \to \mathbf{1}$. The drawings of these maps (and the identity map on $\mathbf{1}$) are:

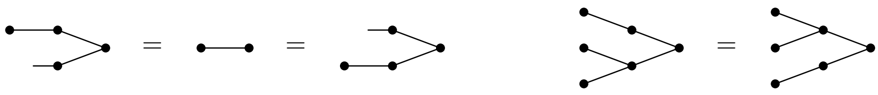


$\mu^{(0)}$ $\qquad$ $\mu^{(1)} = \mathrm{id}_{\mathbf{1}}$ $\qquad$ $\mu^{(2)}$

To establish the claim, note first that every arrow $\mathbf{m} \to \mathbf{n}$ in Δ is the ordinal sum of n arrows to $\mathbf{1}$. Indeed, each dot in the target column has an inverse image, so this splits up the graph into its connected components. Since the strands are not allowed to cross over, this partition is in fact the ordinal sum of arrows. Now observe that for every $n \geq 2$ the map $\mu^{(n)} : \mathbf{n} \to \mathbf{1}$ can be obtained as a composition of ordinal sums of $\mu^{(1)} = \mathrm{id} : \mathbf{1} \to \mathbf{1}$ and $\mu^{(2)} = \mu : \mathbf{2} \to \mathbf{1}$ — this is clear from the pictures:

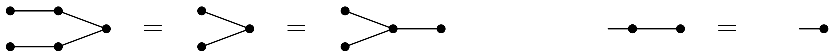


It is straightforward to check that the following relations suffice:



(3.2.5)

Since we have used the identity map although it is not considered a generator, we'd better list the relations it is subject to:



We now come to the symmetric case. From the graphical viewpoint what we do (compared to the construction of Δ) is simply to allow 'crossing-over'. In terms of finite sets it means there is no longer any order relation in play.

3.2.6 Finite cardinals. The category Φ of finite cardinals is defined as a skeleton of the category of finite sets ***FinSet***: Let $\mathbf{n}$ denote the set $\{0, 1, 2, \ldots, n-1\}$, then Φ is the full subcategory of ***FinSet*** given by

$$\Phi = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots\}$$

The objects of Φ are called *finite cardinals*. In 3.2.1 we used the symbol $\mathbf{n}$ to denote the *ordered* set $\{0, 1, 2, \ldots, n-1\}$, which we called an *ordinal* (object in Δ). As *sets* they are the same (i.e., forgetting the ordering), so we allow ourselves to say that Δ and Φ have the same objects. More precisely, the functor $\Delta \hookrightarrow \Phi$ which forgets the order is a bijection on objects.

The *Cardinal sum* is defined in the same way as ordinal sum. But in spite of this similarity, an important property of Φ (which is not shared by Δ) is that *cardinal sum is the coproduct in* Φ, just as disjoint union is in ***FinSet*** and in ***Set***. Hence (cf. 3.1.7) there is a unique symmetric structure on the monoidal category $(\Phi, +, \mathbf{0})$, which we will denote τ.

3.2.7 Remark. The category Φ also has products namely the *cardinal product* which to two cardinals $\mathbf{m}$ and $\mathbf{n}$ associates the cardinal corresponding to the natural number product mn. Thus the category Φ is is the *categorification* of $\mathbb{N}$, in the sense that all important set-theoretic properties of $\mathbb{N}$ are in fact reflections of category-theoretic properties of Φ. For a pleasant introduction to the philosophy of categorification, see Baez-Dolan [7].

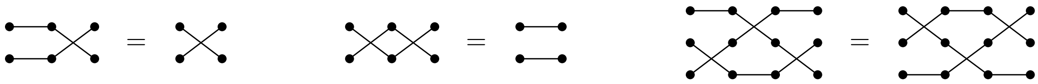
3.2.8 Graphical description of Φ. The *objects* are finite sets of dots arranged in a column. An *arrow* from one column to another is a collection of strands starting at the dots in the source column and ending at dots in the target column, subject to the following single rule:

(i) for each dot in the source column there is exactly one strand coming out (and going to the target column)

Compared to Δ, the difference is that the strands are now allowed to cross over each other.

3.2.9 The category of finite sets and bijections. Let $\boldsymbol{FinSet}_0$ denote the category whose objects are the finite sets, and whose arrows are the bijections between finite sets. Let Σ be the skeleton of $\boldsymbol{FinSet}_0$ defined by $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots\}$. Since there are no bijections between sets of different cardinality, the graph of the category Σ is disconnected, with a connected component for each object $\mathbf{n}$. The arrows from $\mathbf{n}$ to $\mathbf{n}$ are precisely the permutations of the elements in the set $\mathbf{n} = \{0, 1, \dots, n-1\}$, so the category Σ is the disjoint union of the symmetric groups $\mathfrak{S}_n$, via the identification of groups with one-object categories.

Now $(\Sigma, +, \mathbf{0})$ is a monoidal category. As such it is generated by the transposition , subject to the relations



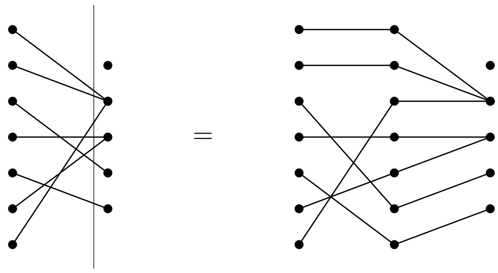
(3.2.10)

(Compare with the relations for the symmetric groups, found e.g. in Coxeter-Moser [10].)

3.2.11 Lemma. *Every arrow in Φ can be factored as a permutation followed by an order-preserving map. In other words,*

$$\Phi = \Sigma\Delta.$$

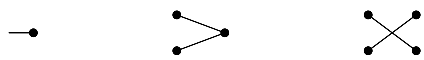
Proof. From the graphical viewpoint this is clear: take any arrow and factor it by taking a vertical cut so far to the right that all the cross-overs occur on the left-hand side of it:



□

Observe that the factorisation is not unique. The order-preserving part f *is* unique, but for the permutation part, we have the freedom to permute those elements which have the same image under f.

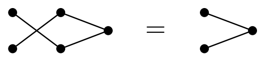
3.2.12 Generators and relations for Φ. It follows immediately from the lemma that this is a complete set of generators for Φ:



$\mu^{(0)}$ $\mu^{(2)}$ τ

In addition to the relations coming from Δ (cf.3.2.5) and the relations coming from Σ (cf. 3.2.10), there are the relations expressing the naturality of the twist map (cf. 3.1.5):

Finally there is one more relation, namely the commutativity relation



valid since **1** is terminal object. More generally, this last relation in combination with the associativity relation implies that it has no effect to permute the input dots within a 'connected component', i.e., dots which have the same image.

It is not difficult to show that these relations suffice.

3.3 Monoids, and monoidal functors on Δ

3.3.1 Monoids (in a monoidal category). Let $(\boldsymbol{V}, \square, I)$ be a monoidal category. A *monoid* in $\boldsymbol{V}$ is an object M together with two arrows

$$\mu : M \square M \to M, \qquad \eta : I \to M$$

subject to the associativity axiom

$$(\mu \square \operatorname{id}_M)\mu = (\operatorname{id}_M \square \mu)\mu$$

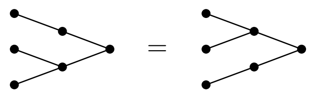
and the unit axiom

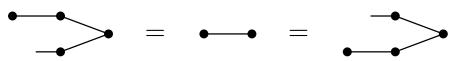
$$(\eta \square \operatorname{id}_M)\mu = \operatorname{id}_M = (\operatorname{id}_M \square \eta)\mu.$$

We picture M itself as a single dot, $M \square M$ as two dots, etc. and draw the two structure maps like this:

μ η

Then the axioms read





3.3.2 Monoid homomorphisms. A *monoid homomorphism* in $\boldsymbol{V}$ between two monoids (M, μ, η) and (M', μ', η') is an arrow $\phi : M \to M'$ that commutes with all the monoid structure. Precisely,

$$\begin{tikzcd}
M \square M \arrow[r, "\phi \square \phi"] \arrow[d, "\mu"'] & M' \square M' \arrow[d, "\mu'"] \\
M \arrow[r, "\phi"'] & M'
\end{tikzcd}
\qquad
\begin{tikzcd}
M \arrow[r, "\phi"] & M' \\
1 \arrow[u, "\eta"] \arrow[r, phantom, "="] & 1 \arrow[u, "\eta'"']
\end{tikzcd}$$

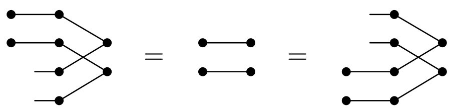
Let $\boldsymbol{Mon}(\boldsymbol{V})$ denote the category whose objects are the monoids in $\boldsymbol{V}$ and whose arrows are the monoid homomorphisms in $\boldsymbol{V}$.

3.3.3 $\Bbbk$-algebras. *A monoid in* $(\boldsymbol{Vect}_\Bbbk, \otimes, \Bbbk)$ *is precisely a* $\Bbbk$*-algebra.* (In fact we defined $\Bbbk$-algebras just like this in 2.1.1.):

$$\boldsymbol{Mon}(\boldsymbol{Vect}_\Bbbk) = \boldsymbol{Alg}_\Bbbk.$$

3.3.4 Monoids in Δ. $(\Delta, +, \mathbf{0})$ *contains a single nontrivial monoid, namely* $(\mathbf{1}, \mu, \eta)$.

Indeed, the structure maps $\mathbf{0} \to \mathbf{1} \leftarrow \mathbf{2} = \mathbf{1} + \mathbf{1}$ for the monoidal category Δ define monoid structure on $\mathbf{1}$. To see that there are no other monoids consider for example $\mathbf{2}$. The only imaginable map $\mathbf{4} \to \mathbf{2}$ satisfying the unit axiom



does not exist in Δ, since crossing-over is not allowed (In the bigger category Φ, the above map does exist, so $\mathbf{2}$ will be a monoid in Φ, as we shall see in 3.3.10.)

3.3.5 Monoidal functors on $(\Delta, +, \mathbf{0})$. The monoidal category Δ with its monoid $\mathbf{1}$ enjoy a universal property:

3.3.6 Theorem. *Given a monoid category $(\boldsymbol{V}, \Box, I)$ there is a one-to-one correspondence between monoidal functors $(\Delta, +, \mathbf{0}) \to (\boldsymbol{V}, \Box, I)$ and monoids in $\boldsymbol{V}$. This gives a canonical equivalence of categories*

$$\begin{array}{rcl} \boldsymbol{MonCat}(\Delta, \boldsymbol{V}) & \xrightarrow{\simeq} & \boldsymbol{Mon}(\boldsymbol{V}) \\ [\mathscr{M} : \Delta \to \boldsymbol{V}] & \mapsto & (\mathbf{1}\mathscr{M}, \mu\mathscr{M}, \eta\mathscr{M}). \end{array}$$

sending a monoidal functor $\mathscr{M}$ to its value on the monoid $\mathbf{1}$.

In the other direction the correspondence goes like this: given a monoid M in $(\boldsymbol{V}, \Box, I)$, define a monoidal functor

$$\begin{array}{rcl} \mathscr{M} : \Delta & \longrightarrow & \boldsymbol{V} \\ [\mathbf{0} \to \mathbf{1} \leftarrow \mathbf{2}] & \longmapsto & [I \to M \leftarrow M \Box M]. \end{array}$$

This makes sense because the relations that hold among the generators in Δ correspond exactly to the monoid axioms for M. (The details of the proof can be copied over from the proof of 3.1.13.)

The functor $\mathscr{M}$ 'points to' a specified monoid in $\boldsymbol{V}$, just like a monoid homomorphism $\mathbb{N} \to M$ points to an element of the monoid M. So the theorem is a sort of categorisation of the bijection between elements in a monoid M and monoid homomorphisms $\mathbb{N} \to M$.

3.3.7 Remark on equivalence versus isomorphism. The way we have set things up, with everything strict, the equivalence is in fact an *isomorphism* of categories. The reason for stating it as just an equivalence is to get a more robust statement, which remains true if the involved notions are weakened, i.e., if we consider non-strict monoidal categories and non-strict monoidal functors.

3.3.8 Corollary. *There is a canonical equivalence of categories*

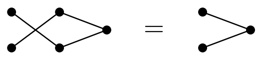
$$\boldsymbol{MonCat}(\Delta, \boldsymbol{Vect}_{\Bbbk}) \simeq \boldsymbol{Mon}(\boldsymbol{Vect}_{\Bbbk}) = \boldsymbol{Alg}_{\Bbbk}.$$

Thinking of Δ as a certain category of non-crossing strands, just like ***2Cob*** is described in terms of certain tubes, we see that this result is analogous to Theorem 3.1.13. A more precise analogy arises when we pass to the symmetric case:

3.3.9 Commutative monoids. A monoid M in a symmetric monoidal category $(\boldsymbol{V}, \Box, I, \tau)$ is called *commutative* if the multiplication $\mu : M\Box M \to M$ is compatible with the twist map. That is, we have a commutative diagram

$$\begin{tikzcd} M\Box M \arrow[rr, "\tau"] \arrow[dr, "\mu"'] & & M\Box M \arrow[dl, "\mu"] \\ & M & \end{tikzcd}$$

Graphically,



Note that the notion of commutativity depends on the symmetric structure. For example a monoid in $(\boldsymbol{grVect}, \otimes, \Bbbk)$ is just a graded algebra. Let H be a graded-commutative algebra (i.e., $ab = (-1)^{pq}\, ba$, where $\deg(a) = p$ and $\deg(b) = q$). Then H is *not* commutative as monoid in $(\boldsymbol{grVect}, \otimes, \Bbbk, \sigma)$, but it *is* commutative as monoid in $(\boldsymbol{grVect}, \otimes, \Bbbk, \kappa)$.

3.3.10 Monoids in Φ. Just as in Δ, there is a unique monoid structure on $\mathbf{1}$ in Φ, and now this monoid structure is commutative (with respect to the unique symmetric structure on Φ). But in contrast to the situation in Δ: every object in Φ carries a unique monoid structure. This is a consequence of the general fact that *In a monoidal category given by coproduct, every object carries a unique monoid structure, and this monoid structure is commutative (with respect to the unique symmetry).*

3.3.11 Theorem. *There is a canonical equivalence of categories*

$$\boldsymbol{SymMonCat}(\Phi, \boldsymbol{V}) \simeq \boldsymbol{cMon}(\boldsymbol{V}).$$

This is just a variation of 3.3.6.

3.3.12 Corollary. *There is a canonical equivalence of categories*

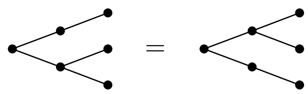
$$\boldsymbol{Repr}_{\Bbbk}(\Phi) = \boldsymbol{SymMonCat}(\Phi, \boldsymbol{Vect}_{\Bbbk}) \simeq \boldsymbol{cMon}(\boldsymbol{Vect}_{\Bbbk}) = \boldsymbol{cAlg}_{\Bbbk}.$$

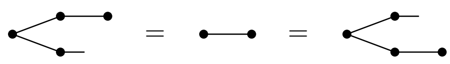
3.4 Frobenius structures

3.4.1 Comonoids. A *comonoid* in a monoidal category $(\boldsymbol{V}, \Box, I)$ is an object M equipped with a *comultiplication* $\delta : M \to M\Box M$ and a *counit* $\varepsilon : M \to I$,

δ $\qquad\qquad$ ε

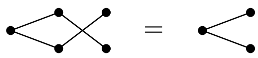
satisfying the axioms dual to the monoid axioms of 3.3.1:





A *comonoid homomorphism* between two comonoids in $\boldsymbol{V}$ is one which preserves the comonoid structure. There is a category $\boldsymbol{Comon}(\boldsymbol{V})$ of comonoids and comonoid homomorphisms in $\boldsymbol{V}$.

A comonoid in a symmetric monoidal category $(\boldsymbol{V}, \square, I, \tau)$ is called *cocommutative* if it satisfies the axiom dual to the commutativity axiom 3.3.9:

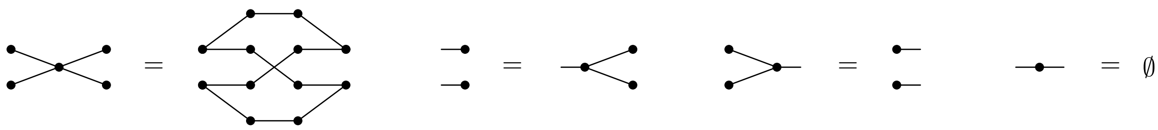


The category of cocommutative comonoids in $\boldsymbol{V}$ is denoted $\boldsymbol{cComon}(\boldsymbol{V})$.

3.4.2 Coalgebras. Dual to 3.3.3 we find that comonoids in $\boldsymbol{Vect}_{\Bbbk}$ are coalgebras:

$$\begin{aligned} \boldsymbol{MonCat}(\Delta^{\mathrm{op}}, \boldsymbol{Vect}_{\Bbbk}) &\simeq \boldsymbol{Comon}(\boldsymbol{Vect}_{\Bbbk}) = \boldsymbol{Coalg}_{\Bbbk} \\ \boldsymbol{Repr}_{\Bbbk}(\Phi^{\mathrm{op}}) = \boldsymbol{SymMonCat}(\Phi^{\mathrm{op}}, \boldsymbol{Vect}_{\Bbbk}) &\simeq \boldsymbol{cComon}(\boldsymbol{Vect}_{\Bbbk}) = \boldsymbol{cCoalg}_{\Bbbk}. \end{aligned}$$

3.4.3 Digression on bimonoids and bialgebras. A *bimonoid* in a symmetric monoidal category $(\boldsymbol{V}, \square, I, \tau)$ is an object B which is simultaneously a monoid and a comonoid, and such that the following compatibility conditions hold



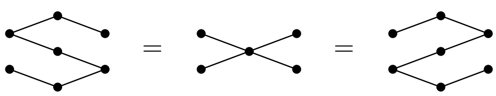
Thus a bimonoid in $(\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk, \sigma)$ is exactly a bialgebra (cf. 2.4.14).

3.4.4 Frobenius objects. A *Frobenius object* in a monoidal category $(\boldsymbol{V}, \square, I)$ is an object A equipped with four maps:

$\eta : I \to A$ $\qquad$ $\mu : A \square A \to A$ $\qquad$ $\delta : A \to A \square A$ $\qquad$ $\varepsilon : A \to I$

satisfying the unit and counit axioms:

as well as the *Frobenius relation*:



3.4.5 Frobenius algebras. A Frobenius object in $(\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk)$ is precisely a Frobenius algebra. This is the content of 2.4.11. Note that a priori we do not require vector spaces to be finite dimension, but we have shown in 2.1.5 that the Frobenius condition on a vector space implies finite dimension.

3.4.6 Lemma. *The multiplication μ of a Frobenius object A is associative, and the comultiplication δ is coassociative. In other words, A is at the same time a monoid and a comonoid.*

Proof. This was proved for Frobenius algebras in 2.4.11. Since the proof was graphical, it is valid for Frobenius objects in any monoidal category. $\square$

3.4.7 Frobenius homomorphisms. A Frobenius homomorphisms between two Frobenius object in $\boldsymbol{V}$ is a map which preserves all the structure. That is, a map which is at the same time a monoid homomorphism and a comonoid homomorphism. There is a category $\boldsymbol{Frob}(\boldsymbol{V})$ of Frobenius objects and Frobenius homomorphisms in $\boldsymbol{V}$.

3.4.8 Question. We saw in 2.4.18 that a morphisms of Frobenius algebras is always invertible. Is the same true for a Frobenius homomorphism in a general monoidal category?

This seems difficult because the argument given there used kernels and dimensions and other vector space specific features. . .

3.4.9 Commutative Frobenius objects. A Frobenius object in a symmetric monoidal category $(\boldsymbol{V}, \Box, I, \tau)$ is said to be *(co)commutative* if it is (co)commutative as a (co)monoid. Commutative Frobenius objects in $(\boldsymbol{Vect}_{\Bbbk}, \otimes, \Bbbk, \sigma)$ are precisely commutative Frobenius algebras.

Now, *a Frobenius object is cocommutative if and only if it is commutative.* Indeed, we proved this statement for algebras (cf. 2.4.16), and since the proof was graphical it carries over to the general case. Let $\boldsymbol{cFrob}(\boldsymbol{V})$ denote the category of commutative Frobenius objects in $\boldsymbol{V}$.

3.4.10 Example. In the symmetric monoidal category $(\boldsymbol{2Cob}, \coprod, \emptyset, T)$, every object has a canonical structure of commutative Frobenius object. In particular, $\mathbf{1}$ is a commutative Frobenius object — this is more or less the content of the description in terms of generators and relations given in Section 1.

(More generally, the $(n-1)$-sphere is a commutative Frobenius object in $\boldsymbol{nCob}$.)

3.4.11 The free monoidal category on a Frobenius object. Now we look for a universal Frobenius structure — the free monoidal category on a Frobenius object. It is straight-forward to describe it in terms of generators and relations: just take $\boldsymbol{X}$ to be the monoidal category generated by the four arrows listed in 3.4.4, subject to the three relations — then clearly $\bullet$ is a Frobenius object in $\boldsymbol{X}$. Now the proof of 3.1.13 carries over directly to establish

3.4.12 Theorem. *There is a canonical equivalence of categories*

$$\boldsymbol{MonCat}(\boldsymbol{X}, \boldsymbol{V}) \simeq \boldsymbol{Frob}(\boldsymbol{V}).$$

To give a more geometric description of $\boldsymbol{X}$, we can interpret it as the category whose *objects are column of dots*, and whose arrows are certain strands between them, which are not allowed to cross over each other. Compared to the description of Δ it is no longer required that precisely one strand emanate from each input dot, neither do they necessarily end up at an output dot (they may double back to an input dot); therefore we can no longer regard the strands as 'functions' of any kind. The precise description of when we consider two such collections of strands to define the same arrow is this: *the arrows are isotopy classes of planar graphs.* The relations express exactly isotopy invariance of certain modifications of planar graphs. The isotopy classes of connected graphs are classified by the genus of the graph and the number of input- and output-dots. To classify non-connected graphs requires some invariant to express nesting.

3.4.13 The category of 2-cobordisms. Finally we can apply the same arguments to the symmetric case. The free symmetric monoidal category on a (co)commutative Frobenius object is clearly the one generated by the four maps in 3.4.4 together with the twist map, and subject to the three relations plus all the relations involving the twist map — including commutativity. But this category was amply described in Section 1: it is precisely (the skeleton of) ***2Cob***! The arguments of the proof of Theorem 3.1.13 now carry over word for word to prove more generally:

3.4.14 Theorem. *The skeletal cobordism category* $(\boldsymbol{2Cob}, \coprod, \emptyset, T)$ *is the free symmetric monoidal category containing a (co)commutative Frobenius object. In other words, given a (co)commutative Frobenius object A in a symmetric monoidal category* $(\boldsymbol{V}, \Box, \boldsymbol{I}, \tau)$ *then there is a unique symmetric monoidal functor* $\mathscr{A} : \mathbf{2Cob} \to \boldsymbol{V}$ *such that* $\mathbf{1} \mapsto A$*. This gives a canonical equivalence of categories*

$$\boldsymbol{SymMonCat}(\boldsymbol{2Cob}, \boldsymbol{V}) \simeq \boldsymbol{cFrob}(\boldsymbol{V}).$$

References

[1] Lowell Abrams. *Two-dimensional topological quantum field theories and Frobenius algebras.* J. Knot Theory and its Ramifications **5** (1996), 569–587. Available at http://home.gwu.edu/~labrams/docs/tqft.ps.

[2] Lowell Abrams. *Modules, comodules, and cotensor products over Frobenius algebras.* J. Alg. **219** (1999), 201–213. Available at http://home.gwu.edu/~labrams/docs/cotensor.ps.

[3] Michael Atiyah. *Topological quantum field theories.* Inst. Hautes Études Sci. Publ. Math. **68** (1989), 175–186.

[4] Michael Atiyah. *The geometry and physics of knots.* Cambridge University Press, Cambridge, 1990.

[5] John Baez' web site, http://math.ucr.edu/home/baez/.

[6] John Baez and James Dolan. *Higher-dimensional algebra and topological quantum field theory.* J. Math. Phys. **36** (1995), 6073–6105. (q-alg/9503002).

[7] John Baez and James Dolan. *From finite sets to Feynman diagrams.* In B. Engquist and W. Schmid, editors, *Mathematics unlimited—2001 and beyond*, pp. 29–50. Springer-Verlag, Berlin, 2001. (math.QA/0004133).

[8] John W. Barrett. *Quantum gravity as topological quantum field theory.* J. Math. Phys. **36** (1995), 6161–6179. (gr-qc/9506070).

[9] Raoul Bott and Loring W. Tu. *Differential forms in algebraic topology.* No. 82 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1982.

[10] Harold S. M. Coxeter and William O. Moser. *Generators and relations for discrete groups.* No. 14 in Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, 1957.

[11] Charles W. Curtis and Irving Reiner. *Representation theory of finite groups and associative algebras.* Interscience Publishers, New York, 1962.

[12] Robbert Dijkgraaf. *A geometric approach to two dimensional conformal field theory.* PhD thesis, University of Utrecht, 1989.

[13] Robbert Dijkgraaf. *Fields, strings and duality.* In *Symétries quantiques (Les Houches, 1995)*, pp. 3–147. North-Holland, Amsterdam, 1998. (hep-th/9703136).

[14] Boris Dubrovin. *Geometry of 2D topological field theories.* In *Integrable systems and quantum groups*, no. 1620 in Lecture Notes in Mathematics, pp. 120–348. Springer-Verlag, New York, 1996. (hep-th/9407018).

[15] David Eisenbud. *Commutative algebra with a view toward algebraic geometry.* No. 150 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1995.

[16] William Fulton. *Algebraic topology: A first course.* No. 153 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1995.

[17] Phillip Griffiths and Joe Harris. *Principles of Algebraic Geometry.* John Wiley and Sons, 1978.

[18] Morris Hirsch. *Differential Topology.* No. 33 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1976.

[19] Christian Kassel. *Quantum groups.* No. 155 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1995.

[20] Joachim Kock. *Frobenius algebras and 2D topological quantum field theories.* Currently available at http://www-math.unice.fr/˜kock/TQFT.html.

[21] F. William Lawvere. *Ordinal sums and equational doctrines.* In B. Eckmann, editor, *Seminar on triples and categorical homology theory, ETH 1966/67*, no. 80 in Lecture Notes in Mathematics, pp. 141–155. Springer-Verlag, New York, 1967.

[22] Saunders Mac Lane. *Categories for the working mathematician, second edition.* No. 5 in Graduate Texts in Mathematics. Springer-Verlag, New York, 1998.

[23] John Milnor. *Lectures on the h-cobordism theorem.* Princeton University Press, Princeton, N.J., 1965.

[24] Michael Müger. *From subfactors to categories and topology I. Frobenius algebras in and Morita equivalence of tensor categories.* Preprint, math.CT/0111204.

[25] Frank Quinn. *Lectures on axiomatic topological quantum field theory.* In *Geometry and quantum field theory (Park City, UT, 1991)*, pp. 323–453. Amer. Math. Soc., Providence, RI, 1995.

[26] Stephen Sawin. *Direct sum decompositions and indecomposable TQFTs.* J. Math. Phys. **36** (1995), 6673–6680. (q-alg/9505026).

[27] Moss E. Sweedler. *Hopf algebras.* W.A. Benjamin, Inc., New York, 1969.

[28] Andrew H. Wallace. *Differential topology: First steps.* W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[29] Edward Witten. *Topological quantum field theory.* Comm. Math. Phys. **117** (1988), 353–386.