An Introduction to Topological Field Theory [1]

R.J. Lawrence [2]

Department of Mathematics
University of Michigan
Ann Arbor, Michigan

Abstract. A topological quantum field theory (TQFT) is an, almost, metric independent quantum field theory that gives rise to topological invariants of the background manifold. The most well known example of a 3-dimensional TQFT is Chern-Simons-Witten theory, in which the expectation value of an observable, obtained as the product of the Wilson loops associated with a link, is the generalised Jones invariant of the link. Unfortunately the form for the invariants obtained by this procedure is that of an integral over an infinite dimensional space on which, for a mathematician, a measure is not rigorously defined. Various ways of avoiding this difficulty have been developed. These fall into two main categories, namely, formal manipulations of Witten's path integral into a form which can then be rigorously defined, and axiomatic encapsulations of the properties of TQFTs. In these notes we will be concerned with the second path, demonstrating how complex categorical and algebraic structures appear, from apparently simple geometry. As will be seen in the lecture, these structures are related to the quantum group structures which arise in other approaches.

1: Introduction

In these notes an elementary introduction to some of the basic algebraic structures arising in topological field theory (TFT) will be given. The case of two dimensional topological field theories will be covered in detail. Although the classification of manifolds in two dimensions is very simple, the analysis of possible theories is yet interesting because when correctly interpreted, it contains in a simple form all the main features which arise in higher dimensional theories. In §2, the basic axioms for TFTs will be formulated, stimulated by structures which arise in Quantum Field Theory and Statistical Mechanics. The associated geometrically motivated structure of a domain category is used in §2.4 to express a TFT as a functor from a domain category of manifolds to a domain category of vector spaces. This leads to notions of TFTs on manifolds endowed with various additional structures, while in §2.5 a technique for generating a TFT over a class of such manifolds, from one on manifolds with even more additional structure, is given. In §3, the process of §2 is seen in action, examples being generated of TFTs, both over the collection of topological manifolds and the collection of triangulated manifolds.

The extension of these ideas to higher dimensions are briefly discussed in §4 and §5 where general notions of higher category structure and extended topological

[1] These notes were prepared as preparatory reading for a session in the AMS Short Course on 'Knots and Physics' in San Franscico, to be held January 2–3, 1995.

[2] Current address: IHES, Bures-sur-Yvette, FRANCE.

field theories (ETFTs) are introduced. Some relationships with constructions of invariants of links and 3-manifolds are also mentioned.

Since the field is a huge and rapidly increasing one, these notes necessarily leave many loose ends; however, it is hoped that they will give the reader a feel for the origins of some of the structures naturally arising in TFT.

2: Formulation of axioms

2.1 Motivation

In classical mechanics, there are two ways of describing the possible evolutions of a system. The first is to specify the equations of motion, which will determine the state of a system at any time (as a point in phase space) once it is given at some initial time; this is known as the Hamiltonian approach. The second description proceeds on the assumption that the worldline followed is one which minimises a certain functional, known as the action. This minimisation takes place over all possible paths in the configuration space, beginning and ending at given points, and is known as the Lagrangian approach. In quantum theory, the phase space is replaced by a Hilbert space of possible states, and dynamical variables are replaced by observables, which are operators on the Hilbert space and have expectation values. In a quantum field theory, the states of the system studied are specified by fields on the background manifold. The Hamiltonian approach leads to the consideration of operators on Hilbert space which describe the evolution of the state of the system. In the Lagrangian approach, the basic object which arises is the partition function of the theory which can be expressed as a Feynman integral.

As for any quantum theory, the ouput from a quantum field theory is a collection of expectation values and correlation functions of observables. A *topological* field theory is a theory in which the output is unchanged under a variation of the metric on the background manifold, so that expectation values of observables must give rise to topological invariants of the manifold.

The first interesting topological field theory was introduced by Witten in 1988, see [W].[3] The partition function of the theory supplies invariants of 3–dimensional manifolds in the form of a Feynman integral,

$$Z(M)=\int_{\mathcal{A}} \mathrm{e}^{ikCS(A)}\mathcal{D}A\,, \tag{2.1.1}$$

where $CS(A)=\frac{1}{4\pi}\int_M\langle A, dA+{}^1\!/\!_3[A,A]\rangle$ is the Chern-Simons action.[4] The data for this theory consists of an integer k, called the level, and a Lie group, G. Here

[3] The stationary phase approximation to Chern–Simons theory had already been investigated by A.S. Schwarz, a decade earlier.

[4] Here, A is a $\mathfrak{g}$–valued 1–form on M, $[\cdot,\cdot]$ is the Lie bracket and $\langle\cdot,\cdot\rangle$ is an invariant bilinear form on $\mathfrak{g}$.

the integral for $Z(M)$ takes place over a quotient of the space of all G–connections on a principal G–bundle over M, that is, on $\mathfrak{g}$–valued 1–forms, A, on M. From the same field theory, Witten also generated invariants, $Z(M,L)$, of pairs (M,L), where L is a link embedded in M, as the expectation value of a suitable observable known as a Wilson loop. Additional data of a choice of representation, ρ_i, of G for each component of L is needed, so that,

$$Z(M,L) = \int_{\mathcal{A}} \mathrm{e}^{ikCS(A)} \prod_{L_i} \mathrm{tr}_{\rho_i} \left(P \exp \oint_{L_i} A ds \right) \mathcal{D}A \,, \tag{2.1.2}$$

where the product is over all components L_i of L. The additional term in (2.1.2), which is associated with a component of the link L, may be geometrically expressed as the trace of the holonomy of the connection A around the component.

In the simplest case of $G = SL(2)$, these Witten invariants for links in S^3, all of whose components are labelled by the 2–dimensional representation, reproduce the one–variable Jones' polynomial, an invariant which was discovered in 1984, see [J1]. When $G = SL(m)$ and the components are all labelled by the vector representation, they generate a slice of the 2–variable HOMFLY polynomial, see [HOMFLY], [PT] and [J2]. These invariants were originally constructed combinatorially from presentations of links, either in the form of two–dimensional projections with over and under crossings, or as closures of braids. These are somewhat unsatisfactory formulations since the invariance of the result is not immediately evident. On the other hand, Witten's functional integral can be fairly easily seen, by a formal proof, to define a topological invariant. However, it has the disadvantage of not being defined rigorously because it is unclear what measure, $\mathcal{D}A$, may be placed on the infinite dimensional space $\mathcal{A}$. Various ways of avoiding this difficulty by formal manipulations of Witten's path integral into a form which *can* be rigorously defined, have been developed.

In these notes, however, our basic philosophy is as follows. We wish to forget about the physical origins of these theories and attempt to make sense of the functional integral form of Z. Instead of pusuing the physical connections, we try to extract properties which could be expected to be satisfied by expressions of the form of (2.1.1), if indeed any reasonable meaning could be attached to them. These properties then become our axioms and we attempt to find solutions to this system of axioms by using methods which may be completely unrelated to the physical origins of (2.1.1).

From our perspective, the output of a theory is a topological invariant, $Z(M)$, defined for objects M which, for the moment, we think of as arbitrary $(d+1)$–dimensional oriented manifolds. In practice, objects will often be endowed with extra information, such as a triangulation, a framing or a metric, and indeed, it is not even necessary for M to have any geometric interpretation so long as there are certain formal operations defined, such as boundary and union.

To motivate the initial set of axioms for TFT, we will start by investigating expressions which look formally similar to (2.1.1). Although we have it in mind that $Z(M)$ will depend only on the structures inherent in M, it may be easier to define a scalar from M when it is endowed with extra structure. One way of obtaining a quantity independent of this structure then, is to sum the values obtained over all such structures. Thus we suppose that $Z(M)$ can be expressed in the form,

$$Z(M) = \int_{\mathcal{A}(M)} W_M(A) \mathcal{D}A \, . \tag{2.1.3}$$

Here $\mathcal{A}(M)$ represents the possible set of extra data on M; $W_M(A)$ is a scalar (the *weight*) computed from M and $A \in \mathcal{A}$; while $\mathcal{D}A$ represents a measure on $\mathcal{A}(M)$. If $\mathcal{A}(M)$ happens to be a discrete set then the integral becomes a summation and the measure turns into a weighting.

Suppose further that the allowed additional data that is placed on M has the following properties.

(i) (*Restriction*) Data may be restricted to submanifolds in a well-defined way, so that there are maps $r_{N,M}\colon \mathcal{A}(M) \longrightarrow \mathcal{A}(N)$ for $N \subset M$ with the property that $r_{P,N} \circ r_{N,M} = r_{P,M}$.

(ii) (*Pasting*) Data on two submanifolds M_1 and M_2 may be pasted together so long as it matches on the intersection. Thus if M is split into two parts (see Figure 1), M_1 and M_2, by a codimension 1 submanifold, Σ, we are supposing that for any $A_i \in \mathcal{A}(M_i)$ $(i = 1, 2)$ with $r_{\Sigma,M_1}(A_1) = r_{\Sigma,M_2}(A_2)$, there exists a unique $A \in \mathcal{A}(M)$ such that $A_i = r_{M_i,M}(A)$ for $i = 1, 2$.

(iii) (*Multiplicativity*) The weight can be defined on manifolds with boundary and is multiplicative under pasting. That is, in the situation of (ii), $W_M(A) = W_{M_1}(A_1).W_{M_2}(A_2)$.

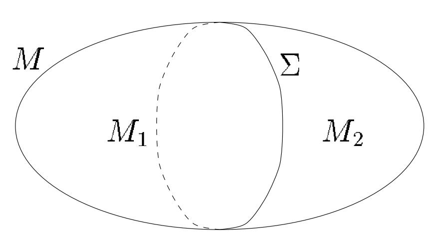

Figure 1: Splitting of a closed manifold

Given boundary data $A_0 \in \mathcal{A}(\Sigma)$, for any manifold M whose boundary contains Σ, we let $\mathcal{A}(M, A_0) \subset \mathcal{A}(M)$, denote the set of all data on M which restricts to A_0 on Σ. The above properties of $\mathcal{A}$ and S enable $Z(M)$ to be written as,

$$Z(M) = \sum_{A_0 \in \mathcal{A}(\Sigma)} Z(M_1, A_0) Z(M_2, A_0) , \tag{2.1.4}$$

where $Z(M_i, A_0)$ is the part of the sum in (2.1.3) consisting of those terms for which $A \in \mathcal{A}(M_i, A_0)$. Equation (2.1.4) may be rewritten as,

$$Z(M) = \langle Z(M_1) \mid Z(M_2) \rangle ,$$

a contraction of two vectors $\langle Z(M_1) \mid$ and $\mid Z(M_2) \rangle$ which lie in dual vector spaces, whose bases are indexed by $\mathcal{A}(\Sigma)$.

Expressions of the form (2.1.3) arise in many areas. In statistical mechanics, $\mathcal{A}(M)$ will represent a set of possible states of the system and $W_M(A)$ will be the relative weight of that state, $\mathbf{e}^{-E_M(A)/kT}$ where $E_M(A)$ is the energy associated with A, T is the temperature and k is Boltzmann's constant. In field theory, $\mathcal{A}(M)$ represents a set of possible fields on M and $W_M(A) = \mathbf{e}^{iS(A)}$, where $S(A)$ represents the action.[5] The physical meaning of the conditions (i)–(iii) above is *locality* of the theory. Thus for a statistical mechanical model on a lattice, it is the requirement that states may be specified by local pieces of data (e.g. spins at vertices), while the energy of a state may be expressed as the sum of local (nearest-neighbour) interaction energies.

2.2 Axioms

The above discussion motivates the following initial definition of a $(d+1)$–dimensional TFT, which is due to Atiyah and was modelled on Segal's axioms for Conformal Field Theory. The theory should assign,

(a) to every closed d–dimensional manifold, Σ, a vector space which will be denoted by $Z(\Sigma)$;

(b) to every $(d+1)$–dimensional manifold, M, a vector $Z(M) \in Z(\partial M)$.

Here the manifolds may be endowed with additional structure and/or constraints; it is assumed that all manifolds are oriented. We generally use the symbols M, Σ and C with appropriate suffices to refer to manifolds of codimension 0,1 and 2 with respect to the dimension of the theory under consideration (here $d+1$). These quantities should satisfy the axioms below.

A1. (VACUUM) $Z(\emptyset) = \mathbf{C}$.

[5] Usually $\mathcal{A}(M)$ is actually a quotient space by the action of an appropriate gauge group. At the level of the present discussion we ignore such complications.

A2. (DUALITY) $Z(\Sigma^*) = Z(\Sigma)^*$.

A3. (MULTIPLICATIVITY) $Z(\Sigma_1 \amalg \Sigma_2) = Z(\Sigma_1) \otimes Z(\Sigma_2)$.

A3′. (GLUING)

(a) $Z(M_1 \amalg M_2) = Z(M_1) \otimes Z(M_2)$ as elements of $Z(\partial M_1) \otimes Z(\partial M_2)$.

(b) If M is a manifold with $\partial M = \Sigma_1 \amalg \Sigma \amalg \Sigma^*$ while $\cup_\Sigma M$ is the manifold obtained from M by identifying the boundary components Σ and Σ^*, then

$$Z(\cup_\Sigma M) = \circ_{Z(\Sigma)} Z(M)\,,$$

where $\circ_{Z(\Sigma)}\colon Z(\Sigma_1) \otimes Z(\Sigma) \otimes Z(\Sigma)^* \longrightarrow Z(\Sigma_1)$ is the natural contraction map on the second and third factors.

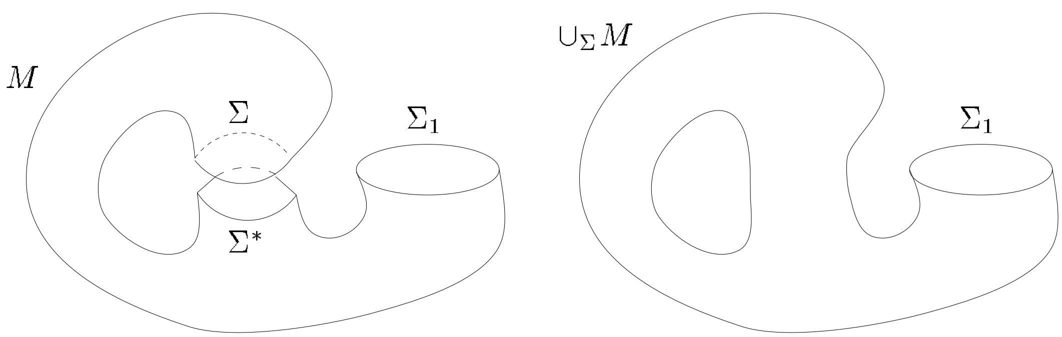

Figure 2: Gluing along a boundary component

Remark 2.2.1

(i) Topological field theories may be defined over any ring, replacing **C** in axiom A1.

(ii) In axiom A2, the symbol '$*$' is used in two different ways on the two sides of the equation. On the left hand side it refers to the operation of reversing the orientation, while on the right hand side it is that of taking the dual on vector spaces.

(iii) As a consequence of axioms A2 and A3, if M, Σ_1 and Σ_2 satisfy $\partial M = \Sigma_1^* \amalg \Sigma_2$ then the theory supplies

$$Z(M) \in Z(\partial M) = Z(\Sigma_1^* \amalg \Sigma_2) = Z(\Sigma_1)^* \otimes Z(\Sigma_2)\,,$$

which may be alternatively viewed as a map $Z(\Sigma_1) \longrightarrow Z(\Sigma_2)$, so that the theory associates linear maps to cobordisms.

(iv) With the notation of the previous remark, one may view M alternatively as a cobordism from Σ_2^* to Σ_1^*. The associated map $Z(\Sigma_2)^* \longrightarrow Z(\Sigma_1)^*$ is now the adjoint of the map $Z(M)$ of the previous remark.

(v) Suppose that M_1 and M_2 are cobordisms from Σ_1 to Σ and from Σ to Σ_2, respectively. Using axioms A3′(a),(b) it follows that

$$Z(M) = Z(M_2) \circ Z(M_1): Z(\Sigma_1) \longrightarrow Z(\Sigma_2)\,,$$

where M is obtained by gluing M_1 and M_2 along Σ. For this reason, axiom A3′ is often referred to as the *associativity* axiom.

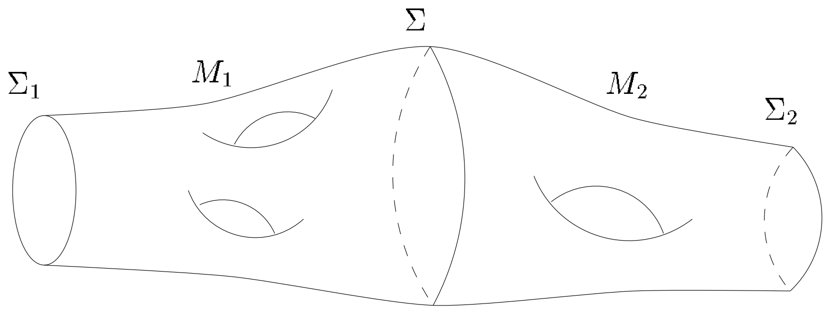

Figure 3: Gluing cobordisms

It is necessary to introduce a further subtlety since it is not really possible to decide when two manifolds are actually *equal*, the most that exists is an isomorphism between them. The same is true for vector spaces and so we arrive at the following additional axioms.

A0. (NATURALITY) Any isomorphism $f: \Sigma \longrightarrow \Sigma'$ of codimension–1 surfaces induces an isomorphism $Z(f): Z(\Sigma) \longrightarrow Z(\Sigma')$ in such a way that $Z(f') \circ Z(f) = Z(f' \circ f)$ for all $f': \Sigma' \longrightarrow \Sigma''$.

A0′. (NATURALITY) For any isomorphism $f: M \longrightarrow M'$, the restriction to ∂M gives an induced isomorphism for which $Z(M') = Z(f|_{\partial M})Z(M)$.

Remark 2.2.2

(i) In axioms A0 and A0′, the isomorphisms, f, are to be thought of as topological equivalences of manifolds, while $Z(f)$ is an isomorphism of vector spaces.

(ii) The same symbol, Z, is here being used to specify the object assigned by the theory to several different object types at the level of manifolds. Thus $Z(*)$ is a vector space, vector, or map according as $*$ is a codimension–1 object, codimension–0 object or morphism specifying the equivalence of codimension–1 objects.

(iii) In Remark 2.2.1(iii), the map $Z(M)$ also depends upon a choice of isomorphism from ∂M to $\Sigma_1^* \amalg \Sigma_2$. Strictly speaking it is therefore necessary to introduce extra isomorphisms in place of equalities in axioms A1, A2, A3 and A3$'$.

In some theories, $Z(\Sigma)$ may possess in addition a *– structure, such as when it is a Hilbert space. In such a situation there is a natural identification of $Z(\Sigma)$ with its dual and it then makes sense to consider a further axiom.

A2$'$. (CONJUGACY) $Z(M^*) = Z(M)^*$ where $Z(M^*) \in Z(\partial M)^*$ is identified with an element of $Z(\Sigma)$.

Those axioms whose labels differ only by the addition of a prime are very similar, the main difference being in the dimension of the base objects — codimension–0 in those axioms with primed label as opposed to codimension–1 in the other axiom. The similarities will become clearer in §2.4 and are extended in §5.

2.3 Consequences

The axioms given above, while they do not determine the theory, Z, do impose quite strong conditions on the structure of $Z(M)$ and $Z(\Sigma)$. Before moving on to a reformulation of the axioms in terms of categories, we give three immediate consequences of the axioms.

Firstly, if it is possible to cut up a manifold M into 'elementary pieces' by closed submanifolds of codimension–1, then the axioms may be used to evaluate $Z(M)$ in terms of the vectors associated with those pieces. This allows theories to be specified completely by a small number of pieces of data, which, as we will see in the next section, may be thought of as constituting an algebraic structure, in an appropriate sense.

Suppose next that Σ is a codimension–1 object. By Remark 2.2.1(iii), the cylinder $\Sigma \times I$ ($I = [0,1]$) supplies a map $i_\Sigma \colon Z(\Sigma) \longrightarrow Z(\Sigma)$. The result of gluing together two copies of this cylinder along a common boundary, is isomorphic to the original cylinder and therefore, by axiom A3$'$,

$$i_\Sigma \circ i_\Sigma = i_\Sigma \ .$$

Thus i_Σ is an idempotent, whose image defines a subspace $Z'(\Sigma)$ of $Z(\Sigma)$. According to Remark 2.2.1(iv), $i_{\Sigma^*} = (i_\Sigma)^*$ and so $Z'(\Sigma^*) \cong Z'(\Sigma)^*$. By axiom A3, $i_{\Sigma_1 \amalg \Sigma_2} = i_{\Sigma_1} \otimes i_{\Sigma_2}$, from which it follows that $Z'(\Sigma_1 \amalg \Sigma_2) = Z'(\Sigma_1) \otimes Z'(\Sigma_2)$. For any manifold, M, it follows from axiom A3$'$(b), by gluing a cylinder on ∂M onto M, that,

$$i_{\partial M}(Z(M)) = Z(M) \in Z(\partial M) \ .$$

We conclude that $Z(M) \in Z'(\Sigma)$ and therefore that the smaller theory Z' defined with $Z'(M) \equiv Z(M)$ also satisfies the axioms for a TFT. As a result, in the classification of TFTs we may assume, without loss of generality, that $Z'(\Sigma) = Z(\Sigma)$ for all Σ. Hence we assume throughout the rest of these notes that $i_\Sigma = \mathrm{id}_{Z(\Sigma)}$.

Finally, suppose that we apply axiom A3′(b) to the cylinder $\Sigma \times I$. The result of identifying the two copies of Σ in the boundary is $\Sigma \times S^1$. Thus $Z(\Sigma \times S^1) \in Z(\emptyset) = \mathbf{C}$ is obtained as the trace of i_Σ, that is,

$$Z(\Sigma \times S^1) = \dim Z(\Sigma)\,.$$

2.4 Interpretation using categories

A neat way of encapsulating the axioms for TFTs is in terms of categories. For a nice account of category theory, see [M]. So that the reader who is unfamiliar with the notion of a category can still follow these notes, we give here the definition of a category and some elementary examples which will be relevant to us.

Definition 2.4.1 *By a category, $\mathcal{C}$, is meant a set of objects, denoted* $\mathrm{Obj}\,(\mathcal{C})$ *along with, for each pair of objects $A, B \in \mathrm{Obj}\,(\mathcal{C})$, a set,* $\mathrm{Morph}_{\mathcal{C}}(A, B)$, *of morphisms and the following additional structures.*

(a) For each $A \in \mathrm{Obj}\,(\mathcal{C})$, an element $\mathrm{id}_A \in \mathrm{Morph}_{\mathcal{C}}(A, A)$, called the identity morphism on the object A.

(b) For each triple $A, B, C \in \mathrm{Obj}\,(\mathcal{C})$, a map,

$$\circ\colon \mathrm{Morph}_{\mathcal{C}}(B, C) \times \mathrm{Morph}_{\mathcal{C}}(A, B) \longrightarrow \mathrm{Morph}_{\mathcal{C}}(A, C)$$

called composition, which is such that $\mathrm{id}_B \circ f = f = f \circ \mathrm{id}_A$ and $f \circ (g \circ h) = (f \circ g) \circ h$, the equalities holding for all objects and morphisms for which both sides make sense.

We will often use the same notation for the set of objects in a category and the category itself, so that '$a \in \mathcal{C}$' means '$a \in \mathrm{Obj}\,(\mathcal{C})$'. Also, when it is clear in which category we are working, we will drop the suffix to Morph. If A and B are two objects in the category $\mathcal{C}$, it is sometimes convenient, and easier on the reader, to represent the fact that $f \in \mathrm{Morph}\,(A, B)$ by $f\colon A \longrightarrow B$.

Example 2.4.2 If G is a group, then one may construct a category with a one-element set of objects, for which the set of morphisms from this object to itself is just G, the identity and multiplication operations in G providing the requisite structures in Definition 2.4.1(a) and (b), respectively. Observe that only the identity and associativity properties of G have been used, so that this construction works equally well for any monoid.

Example 2.4.3 Another way of viewing a category is as a monoid structure in which the multiplication map is only defined on a *subset* of the set of all pairs of elements. Thus, if $\mathcal{C}$ is a category, let $C^0 = \mathrm{Obj}\,(\mathcal{C})$ and C^1 denote the (disjoint) union of the sets $\{A\} \times \{B\} \times \mathrm{Morph}\,(A, B)$ over all pairs of objects A and B. Then there are defined two maps $s, t \colon C^1 \longrightarrow C^0$, called the source and target maps, under which the images of (A, B, f) are A and B, respectively. Elements $a, b \in C^1$ can be composed so long as $s(a) = t(b)$. The identity morphism provides a map $i \colon C^0 \longrightarrow C^1$ for which $s \circ i = t \circ i = \mathrm{id}_{C^0}$, while part (b) of Definition 2.4.1 are identity and associativity constraints.

Example 2.4.4 On any set S, there are two trivial category structures whose object set is S. For the first, $\mathrm{Morph}\,(A, B)$ is a one-element set for all $A, B \in S$, while for the second $\mathrm{Morph}\,(A, B)$ is empty, unless $A = B$, when it contains just one element. The former category will be denoted ALL_S.

Example 2.4.5 The category of vector spaces, $\mathcal{V}^1$, has as objects, finite dimensional vector spaces, and as morphisms, linear transformations. Since any finite dimensional vector space may be determined up to isomorphism by its dimension, one may also consider the category of *coordinatised* vector spaces, $\mathcal{V}^1_c$, whose set of objects is $\mathbf{N} \cup \{0\}$ and for which $\mathrm{Morph}\,(m, n)$ is the set of $n \times m$ matrices with entries in $\mathbf{C}$. Composition is given by matrix multiplication while the identity morphism on n is the $n \times n$ identity matrix.

Example 2.4.6 Suppose that G is any group. Then one may define a category whose set of objects consists of pairs (V, ρ), where V is a vector space and $\rho \colon G \longrightarrow \mathrm{End}(V)$ is a representation of G. For two objects (V, ρ) and (W, σ), the set of morphisms is defined to be the set of linear transformations $\alpha \colon V \longrightarrow W$ for which $\alpha(\rho(g)\mathbf{v}) = \sigma(h)\alpha(\mathbf{v})$ for all $g \in G$ and $\mathbf{v} \in V$. The resulting category is known as the category of representations of G. This same construction can be used for almost any algebraic structure. The smallest structure for which it can be applied is a groupoid, while additional structure on G induces additional functors on the category.

Example 2.4.7 This introduces the fundamental groupoid of a manifold. Suppose that M is a manifold. Construct as follows the category $\mathcal{C}$, whose set of objects is the set of points in M. If $p_0, p_1 \in M$ then $\mathrm{Morph}\,(p_0, p_1)$ consists of all homotopy classes of paths in M from p_0 to p_1. That is, a morphism is an equivalence class of maps $\gamma \colon [0, 1] \longrightarrow M$, with $\gamma(0) = p_0$ and $\gamma(1) = p_1$, under homotopy equivalence.

Definition 2.4.8 *If $\mathcal{C}$ and $\mathcal{C}'$ are two categories, then by a functor from $\mathcal{C}$ to $\mathcal{C}'$ is meant a map $F\colon \mathrm{Obj}(\mathcal{C}) \longrightarrow \mathrm{Obj}(\mathcal{C}')$, along with, for each pair of objects, $A, B \in \mathrm{Obj}(\mathcal{C})$, a map $F\colon \mathrm{Morph}(A,B) \longrightarrow \mathrm{Morph}(F(A),F(B))$ such that,*

(i) $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$;

(ii) $F(f \circ g) = F(f) \circ F(g)$ *for any compatible morphisms f and g in $\mathcal{C}$.*

In terms of the notion of a category, one may now reformulate axioms A1–3′ by stating that a $(d+1)$–dimensional topological field theory defines a functor $Z\colon \mathcal{M} \longrightarrow \mathcal{V}^1$, where $\mathcal{M}$ denotes the category whose objects are closed, oriented, d–dimensional manifolds and whose morphisms are cobordisms. The axioms also imply that this functor preserves various other structures which exist on both these categories, namely $\emptyset$ and $\amalg$ as well as * on $\mathcal{M}_1$.[6]

However, as mentioned immediately following Remark 2.2.1, it is actually necessary to take into account maps induced by isomorphisms between surfaces. Our aim is to reformulate the axioms of TFT so that they just require that Z is a map from an appropriate structure on our collection of manifolds, to a 'linearised' example of this structure, which preserves the whole structure. The required structure will take into account both the category structure in $\mathcal{M}$ and the existence of isomorphisms. To determine what structure we need, it is necessary to go back to the set of axioms for TFT in §2.2, and carefully observe all the operations on manifolds used. Let $\mathcal{M}_0$ denote the collection of $(d+1)$–dimensional oriented manifolds and $\mathcal{M}_1$ denote the collection of d–dimensional oriented closed manifolds, where as usual the word 'manifold' is used to include any additional structures and restrictions. The indices are used to specify the codimension of the objects in these sets. Then isomorphisms between manifolds make $\mathcal{M}_0$ and $\mathcal{M}_1$ into categories. For $i = 0, 1$, there are also special objects $\emptyset \in \mathcal{M}_i$ as well as the operations $\amalg$ and *. Since the latter two operations are consistent with isomorphisms of manifolds, they really define functors,

$$\amalg\colon \mathcal{M}_i \times \mathcal{M}_i \longrightarrow \mathcal{M}_i\,,$$
$$^*\colon \mathcal{M}_1 \longrightarrow \mathcal{M}_1\,.$$

The most important connection between $\mathcal{M}_0$ and $\mathcal{M}_1$ is the boundary operation. As seen in §2.3, there is also another important relation between these sets, namely $\Sigma \longmapsto \Sigma \times I$. This gives two additional functors

$$\partial\colon \mathcal{M}_0 \longrightarrow \mathcal{M}_1\,,$$
$$\times I\colon \mathcal{M}_1 \longrightarrow \mathcal{M}_0\,.$$

The fact that $\partial(\Sigma \times I) = \Sigma \amalg \Sigma^*$ translates into the existence of a morphism between these two objects in $\mathcal{M}_1$. On $\mathcal{M}_0$, in addition to the operation of $\amalg$,

[6] To include axiom A2′ we would require also a * operation on $\mathcal{M}_0$. This gives rise to what we call a *–domain category structure.

the other important construction entering the axioms is that of gluing together boundary components. For this we need a top-dimensional manifold M along with an identification of ∂M with $\Sigma \amalg \Sigma^* \amalg \Sigma'$, for some surfaces Σ, Σ', and the result will be a manifold $\cup_\Sigma M$ whose boundary may be identified with Σ'. When this statement is recast in our categorical notation, the data consists of an element $M \in \mathcal{M}_0$ along with a morphism $f \in \mathrm{Morph}_{\mathcal{M}_1}(\partial M, \Sigma_1 \amalg \Sigma \amalg \Sigma^*)$, while the result is the pair

$$\cup_\Sigma M \in \mathcal{M}_0 \ ,$$
$$\cup_\Sigma f \in \mathrm{Morph}_{\mathcal{M}_0}(\partial(\cup_\Sigma M), \Sigma') \ .$$

We now arrive at the appropriate formal structure on $(\mathcal{M}_0, \mathcal{M}_1)$; it was introduced by Quinn and called a *domain category*. Just as Definition 2.4.1 gave a category as a pair of sets along with a collection of maps which must satisfy a number of conditions, this new structure consists of the pair of categories $(\mathcal{M}_0, \mathcal{M}_1)$, along with objects and maps $(\emptyset, \amalg, ^*, \partial, \times I, \cup_\Sigma)$, which are required to satisfy a fairly long list of properties. These properties are obtained by translating into categorical language, a generating set for the equalities satisfied by the objects and maps just outlined, when considered in their natural setting of manifolds. For a formal definition, the reader is referred to [Q], where there are also given a number of constructions of domain categories without any geometric connotations, coming from purely algebraical data. The domain category based on manifolds will be denoted by $\mathcal{M}$, its dependence on d being omitted from the notation, as throughout, d is assumed to be fixed and clear from the context.

Another example of a domain category, comes from the category of vector spaces over $\mathbf{C}$, say, for definiteness. It is based on the pair of categories $(\mathcal{V}_0, \mathcal{V}_1)$. Here $\mathcal{V}_0$ has as its set of objects, the collection of all pairs $(V, \mathbf{v})$ where V is a vector space and $\mathbf{v} \in V$, while a morphism from $(V, \mathbf{v})$ to $(W, \mathbf{w})$ is an isomorphism $f{:}\, V \longrightarrow W$ for which $f(\mathbf{v}) = \mathbf{w}$. The category $\mathcal{V}_1$ has vector spaces as objects, while the only allowed morphisms are isomorphisms. We now have structures,

$\emptyset$: $\mathbf{C}$ in $\mathcal{V}_1$ and $(\mathbf{C}, 1)$ in $\mathcal{V}_0$;

*: $V \longmapsto V^*$ in $\mathcal{V}_1$;[7]

$\amalg$: $V \amalg W = V \otimes W$ in $\mathcal{V}_1$; $(V, \mathbf{v}) \amalg (W, \mathbf{w}) \equiv (V \otimes W, \mathbf{v} \otimes \mathbf{w})$ in $\mathcal{V}_0$.

The ∂ functor takes $(V, \mathbf{v})$ to V while $\times I$ takes V to $(V \otimes V^*, \mathbf{e})$, where $\mathbf{e} \in V \otimes V^*$ is the vector defined by the identity map $V \longrightarrow V$. Finally, if $M = (V, \mathbf{v}) \in \mathcal{V}_0$ and $f{:}\, V \longrightarrow W_1 \otimes W_0 \otimes W_0^*$, then $\cup_{W_0} M = (W_1, \mathbf{w}) \in \mathcal{V}_0$ and $\cup_{W_0} f = \mathrm{id}$, where $\mathbf{w}$ is the image of $f(\mathbf{v}) \in W_1 \otimes W_0 \otimes W_0^*$ under the contraction of the first two factors.

We can now reformulate the definition of a TFT given in §2.2.

[7] This example can be extended to give a * operation on $\mathcal{V}_0$ by endowing all vector spaces with inner product structures, which morphisms in $\mathcal{V}_i$ are required to preserve. Then $(V,\mathbf{v})^* \equiv (V^*, \mathbf{v}^*)$.

Definition 2.4.9 *A $(d+1)$–dimensional topological field theory is a functor $Z\colon \mathcal{M} \longrightarrow \mathcal{V}$, where now $\mathcal{M}$ and $\mathcal{V}$ refer to the domain categories of manifolds and vector spaces discussed above.*[8]

Whenever we use the word functor, we mean a map between two algebraic structures which preserves all of the structure. In Definition 2.4.9, all the complications of the axioms of TFT have now been wound into the domain category structure. The dimension, $d+1$, of the theory, now only enters into the construction of $\mathcal{M}$. This formulation also allows for considerable generalisation; indeed one can consider a TFT over *any* domain category, not just one arising from manifolds of a certain dimension and this point of view is adopted in [Q]. One may say, that a domain category embodies all the geometry of manifolds we are allowed to know about in TFT.

Bordism categories

In the above discussion of the domain category $\mathcal{M}$, it is not immediately apparent how to reconstruct from $\mathcal{M}$ the basic category structure on the collection of closed codimension–1 objects given by cobordisms, and so we give it here. This process can be carried out quite abstractly on any domain category $\mathcal{M}$. Construct a category $\mathcal{C}$ whose set of objects is $\mathcal{M}_1$, as follows. If $\Sigma_1, \Sigma_2 \in \mathcal{M}_1$ then

$$\mathrm{Morph}_{\mathcal{C}}(\Sigma_1, \Sigma_2) = \{(M, f) | M \in \mathcal{M}_0, f \in \mathrm{Morph}_{\mathcal{M}_1}(\partial M, \Sigma_1^* \amalg \Sigma_2)\} / \sim ,$$

where the equivalence relation $\sim$ is defined by $(M, f) \sim (M', f')$ if, and only if, there exists $F \in \mathrm{Morph}_{\mathcal{M}_0}(M, M')$ such that $f = f' \circ \partial F$. The category $\mathcal{C}$ is known as the *bordism category* associated with $\mathcal{M}$.

2.5 Inverse limit constructions

The definition of TFT given in the last section, allows one to consider theories over domain categories in which the basic objects are manifolds with additional data. Adding further additional data usually makes the construction of TFTs easier. On the other hand, the theory then only produce an invariant of a manifold endowed with this structure, which may be a long way from what one would like to mean by a topological invariant.[9] The compromise often employed is to construct a TFT with additional data and then attempt to eliminate this data using an *inverse limit* construction.

---

[8] Replacing the domain categories $\mathcal{M}$, $\mathcal{V}$ by *–domain categories gives a *–TFT, that is, a TFT satisfying axiom A2$'$. See footnote 7.

[9] For example, in this generalised sense, conformal field theories are 2–dimensional TFTs where the additional data is a metric, given up to conformal equivalence. An axiomatic mathematical definition of CFT was given by Segal in [S], and it is on these axioms that the first axioms for TQFTs [A2] were based.

Suppose that $\mathcal{M}$ and $\mathcal{M}'$ are two domain categories and that there is a functor $r: \mathcal{M}' \longrightarrow \mathcal{M}$. Suppose Z' is a TFT over $\mathcal{M}'$; we wish to construct a TFT over $\mathcal{M}$ in such a way that the following diagram commutes.

$$\begin{array}{ccc} \mathcal{M}' & \xrightarrow{Z'} & \mathcal{V}^1 \\ {\scriptstyle r}\big\downarrow & \nearrow_{Z} & \\ \mathcal{M} & & \end{array}$$

Think of the objects in $\mathcal{M}'$ (that is, elements of $\mathrm{Obj}\,(\mathcal{M}'_i)$ for $i = 0, 1$) as being objects of $\mathcal{M}$ endowed with additional data, such as a triangulation or a metric. This is indicated by writing objects in $\mathcal{M}'$ as pairs (M, α), where M is an object in $\mathcal{M}$ and α represents the additional data. In this notation, $r(M, \alpha) = M$ and $Z'(M, \alpha) \in Z'(\partial M, \partial\alpha)$.

It is first necessary to assume that $Z'(M, \alpha) = Z'(M, \beta)$ for any data α, β agreeing on the boundary, that is for which $\partial\alpha = \partial\beta$. Our aim is, for each Σ, to identify the vector spaces $Z'(\Sigma, \alpha)$ and then to define a particular vector, in this vector space, associated with each M for which $\partial M = \Sigma$. Fix Σ and assume that the possible extra data, $r^{-1}(\Sigma)$ on Σ, forms a category, $\mathcal{C}$, while there is a functor $F: \mathcal{C} \longrightarrow \mathcal{V}^1$ such that on objects it acts according to,

$$F(\Sigma, \alpha) = Z'(\Sigma, \alpha)\,.$$

That is, we are assuming that for each morphism $f: \alpha \longrightarrow \beta$ in $\mathcal{C}$, there is defined a linear transformation $F(f): Z'(\Sigma, \alpha) \longrightarrow Z'(\Sigma, \beta)$ compatible with composition. Using F, we now define a vector space $Z(\Sigma)$, independent of additional data by,

$$Z(\Sigma) = \Big\{ \big(\mathbf{v}: \alpha \longmapsto \mathbf{v}(\alpha)\big) \,\Big|\, \mathbf{v}(\alpha) \in Z'(\Sigma, \alpha)\ \forall\alpha\,, \\ F(f)\big(\mathbf{v}(\alpha)\big) = \mathbf{v}(\beta)\ \forall f: \alpha \longrightarrow \beta \in \mathcal{C} \Big\}\,,$$

the set of all flat sections. An element of $Z(\Sigma)$ is now a choice, for each possible additional structure α on Σ, of an element of $Z'(\Sigma, \alpha)$, in such a way that they are compatible with respect to F.

Suppose that F has 'trivial holonomy', that is, $F(f)$ depends only on the source α and target β of f, and so may be denoted $F(\alpha \longrightarrow \beta)$, while any two objects in $\mathcal{C}$ are connected by a morphism.[10] Then there is an isomorphism,

$$Z(\Sigma) \cong \big\{ \mathbf{v} \in Z'(\Sigma, \alpha_0) \,\big|\, F(f)\mathbf{v} = \mathbf{v}\,, \forall f: \alpha_0 \longrightarrow \alpha_0 \big\}\,,$$

[10] Equivalently, take $\mathcal{C}=\mathrm{ALL}_{r^{-1}(\Sigma)}$.

for each possible additional data α_0 on Σ. To define Z on objects $M \in \mathcal{M}_0$, pick a representative of $r^{-1}(M)$, say $(M,\theta) \in \mathcal{M}'_0$ and define $Z(M)$ to be the choice

$$\left(\alpha \longmapsto F(\partial\theta \longrightarrow \alpha)Z'(M,\theta)\right) .$$

This completes the definition of the inverse limit theory, Z over $\mathcal{M}$.

In some cases, the functor F may be generated in terms of Z' itself. For example, if the extra data is that of a triangulation on a d–dimensional manifold, Σ, then it is known that one may pass between any two triangulations by a sequence of local moves on the simplices. The moves replace a local configuration of j simplices with $d+2-j$ simplices, by replacing the union of a proper subset of the faces of a $(d+1)$–simplex by the union of the complementary set of faces. Such a sequence of moves can be thought of as providing a (possibly singular) triangulation of $\Sigma \times [0,1]$ and therefore Z' provides a map between the vector spaces associated with the two different triangulations of Σ. This process will be seen in action in §3.2.

3: Two dimensional TFT

In this section we shall discuss 2–dimensional TFT, illustrating how the axioms enable a complete classification of the possible theories in terms of simple algebraic structures. In §3.1, decompositions of manifolds into elementary pieces will be used to obtain this classifying structure. In §3.2, a family of TFTs will be constructed, using the inverse limit construction of §2.5, by starting with theories on triangulated manifolds. The examples so obtained, depend upon a choice of data which is similar to, but different from, that arising in the §3.1. Finally, the relation between the data required in §3.2, and the algebraic structure classifying the theory in §3.1 is obtained.

3.1 Classification

Identification of domain category

According to Definition 2.4.9, before we can start to classify 2–dimensional topological field theories, we must understand the structure of the associated domain category, $\mathcal{M}$. Our aim is to construct purely topological invariants, and therefore we take the objects in $\mathcal{M}_i$ to be oriented manifolds given only up to topological equivalence. In particular, an element, Σ, of $\mathcal{M}_1$ should represent a manifold with codimension–1, that is, a disjoint union of oriented copies of the circle, S^1. More precisely, Σ is specified by an integer $N \in \mathbf{N} \cup \{0\}$ (the number of components of Σ), along with an identification of $\amalg^N S^1$ with Σ, given up to orientation-preserving diffeomorphisms. Equivalently, each component of Σ is labelled with an element of $\{1,\ldots,N\}$ and is given an orientation.

Notation 3.1.1 For $N \in \mathbf{N} \cup \{0\}$, let $[N]$ denote the set $\{1, \ldots, N\}$ when $N \neq 0$ and $\emptyset$ when $N = 0$. If M is a manifold, let $c(M)$ denote the number of connected components of M. If M is triangulated, let $n_d(M)$ denote the number of d–simplices in the triangulation.

For our purposes it is sufficient to work with the following domain category, $\mathcal{N}$, which is a slight variation of $\mathcal{M}$. The elements in $\mathcal{N}_1$ will consist of pairs (N, α), where $N \in \mathbf{N} \cup \{0\}$ and $\alpha \in \{+,-\}^N$, that is, $\alpha: [N] \longrightarrow \{+,-\}$. The set of morphisms between two objects (N, α) and (M, β) will be empty unless $M = N$ and then it will consist of those permutations $\sigma \in S_N$ for which $\beta_i = \alpha_{\sigma(i)}$. Let C_n denote the circle in $\mathbf{C}$ of radius $1/3$ and centre n. The elements of $\mathcal{N}_0$ are given by 2–dimensional oriented surfaces, M, embedded in $\mathbf{C} \times \mathbf{R}^-$ along with a choice of orientation of the boundary components, in such a way that $\partial M = \cup_{n \in [N]} C_n \times \{0\}$ for $N = c(\partial M)$. A morphism in $\mathcal{N}_0$ is just an orientation preserving homeomorphism. The full domain category structure on $\mathcal{N}$ is given as follows.

(i) The objects $\emptyset$ are $(0, \emptyset) \in \mathcal{N}_1$ and the empty manifold in $\mathcal{N}_0$.

(ii) The operation * is $(N, \alpha) \longmapsto (N, -\alpha)$ in $\mathcal{N}_1$ and the reversal of all orientations in $\mathcal{N}_0$.

(iii) The operation $\amalg$ is given by $(N, \alpha) \amalg (M, \beta) = (M + N, \alpha\beta)$ in $\mathcal{N}_1$, where $\alpha\beta$ denotes the sequence α followed by the sequence β. On $\mathcal{N}_0$, $M_1 \amalg M_2$ is given by the union of M_1 and M_2, in which M_2 is translated (and possibly also deformed so as to avoid self intersections), so that its boundary is the union of circles $C_i \times \{0\}$ for $i \in c(\partial M_1) + [c(\partial M_2)]$.

(iv) The boundary functor $\partial: \mathcal{N}_0 \longrightarrow \mathcal{N}_1$ takes an object $M \in \mathcal{N}_0$ to the pair $(c(\partial M), \alpha)$ where $\alpha(i)$ is $+$ or $-$ according as the orientation supplied on the i^{th} boundary component matches or otherwise, the orientation induced on the boundary by the orientation on M.

(v) The functor $\times I$ takes (N, α) to a disjoint union of N deformed cylinders whose boundaries are $C_i \cup C_{N+i}$, for $i \in [N]$ and for which the orientations of the $2N$ boundary circles $C_1, \ldots, C_{2N}$, are in accordance with the signs $\alpha(1), \ldots, \alpha(N), -\alpha(1), \ldots, -\alpha(N)$.

(vi) Suppose that $M \in \mathcal{N}_0$ and $f \in \mathrm{Morph}_{\mathcal{N}_1}(\partial M, \Sigma_1 \amalg \Sigma \amalg \Sigma^*)$. Say, $\Sigma = (N, \alpha)$, $\Sigma_1 = (N_1, \alpha_1)$. Then $c(\partial M) = 2N + N_1$ and f is given by a permutation $\sigma \in S_{c(\partial M)}$. Re-embed M in such a way that $\sigma = 1$. Then the manifold $\cup_\Sigma M$ is defined by gluing onto M, N cylinders joining the i^{th} and $(i+N)^{\text{th}}$ boundary components, for each $i = N_1 + 1, \ldots, N_1 + N$.

Fundamental objects

Having now specified the domain category over which to consider topological field theories, we are in a position to consider such a theory, say Z. Denote the image under Z of the object $(1,+) \in \mathcal{N}_0$, consisting of a single positively oriented component, by V. By axioms A2 and A3, it follows that Z is completely determined on $\mathcal{N}_1$ and is given by,

$$Z(N,\alpha) = \bigotimes_{i=1}^{N} (V_{\alpha_i}) ,$$

where $V_+ \equiv V$ and $V_- \equiv V^*$. The morphisms in $\mathcal{N}_1$ provide an action of S_N on the groupoid of tensor products of N vector spaces $V_\pm$. We shall assume that this action is generated by the natural identification of tensor products in opposite orders, $W_1 \otimes W_2$ and $W_2 \otimes W_1$.[11] Any 2–dimensional compact orientable surface with boundary can be obtained by gluing together copies of a disc and a trinion along boundary circles. Thus any element of $\mathcal{N}_0$ may be obtained by appropriately gluing together copies of the three elementary pieces shown in Figure 4, namely a disc with positively oriented boundary, a trinion with all three boundary components positively oriented and a cylinder with its two boundary components oriented negatively.

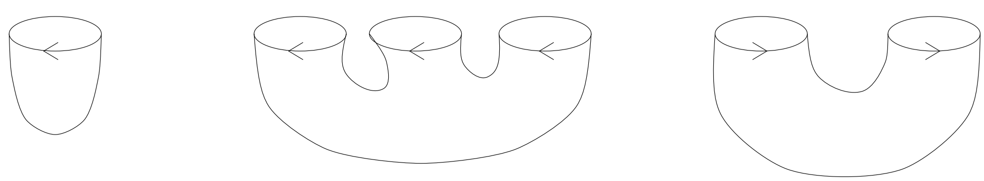

Figure 4: Elementary pieces for 2-d TFT

Under Z, these elementary pieces map to vectors which we denote by $1 \in V$, $t \in V \otimes V \otimes V$ and $\mu \in V^* \otimes V^*$, respectively. By axiom A0′, μ and t must be invariant under the actions of S_2 and S_3, respectively. By the gluing axioms A3 and A3′, it follows that it is possible to determine the image, under Z, of an arbitrary element of $\mathcal{N}_0$, as a contraction of a combination of 1, t and μ. Thus Z is determined by the collection $(V, 1, t, \mu)$.

[11] Allowing more general actions of S_N would provide the structure of a symmetric tensor category on the set of tensor products of copies of V and V^*.

Relations

Suppose that M is an oriented surface with boundary. Choose a collection of oriented simple closed curves, D_i, in Σ which are such that all components, T_j, of $M \backslash \cup_i D_i$ are homeomorphic to one of the forms in Figure 4. To such a decomposition one may correspond an oriented graph Γ with external legs, whose vertices are labelled by the components T_j and whose (internal) edges are in 1–1 correspondence with the curves D_i, while the external legs are labelled by connected components of ∂M. An edge labelled by D_i is oriented from T_j to T_k if the orientation on D_i matches that induced by the orientation of T_j (and consequently is opposite to that induced by the oreintation of T_k). The vertices of Γ will have valencies 1,2 or 3 depending upon the type of the associated piece T_j, while at a vertex of valence v, all edges will be oriented outwards or inwards according as v is odd or even. This process is illustrated in Figure 5.

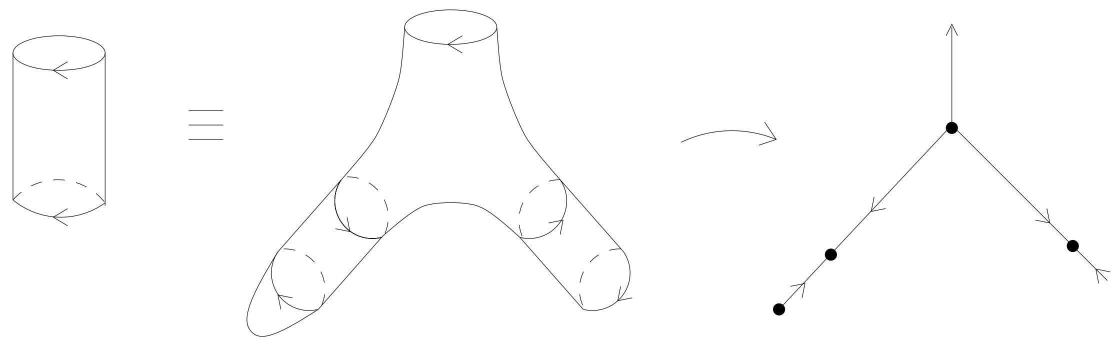

Figure 5: Decomposing a cylinder

Given such a decomposition, $Z(M)$ may be computed from Γ by placing a copy of 1, μ or t at each vertex, according to its valence and then contracting along all internal edges. Indeed, without loss of information, all vertices of Γ of valence 2 may be omitted, while orientations need only be given for the external legs, it being understood that on internal edges the contraction $V \otimes V \longrightarrow \mathbf{C}$ is determined by μ. For example, the contraction represented by Figure 5 is the element of $V \otimes V^*$ given by $i_{S^1} = \mathrm{id}_V : V \longrightarrow V$ so that,

$$(\mu \otimes \mathrm{id} \otimes \mu)(v \otimes t \otimes 1) = v\ , \qquad \text{for all } v \in V. \tag{3.1.2}$$

For a fixed surface, M, there are many different possible decompositions while, by axiom A0, the results of the associated contractions of tensors must be equal. The equality of the images of the two decompositions of a sphere with 4 discs removed shown in the left hand side of Figure 6 yields the relation,

$$\mu_{34}(t \otimes t) \in V^{\otimes 4} \text{ is invariant under the action of } S_4\ . \tag{3.1.3}$$

Except for the sphere, disc and torus any surface can be obtained by gluing trinions without the use of discs, while any two such decompositions of a surface may be obtained from each other, up to homotopies of D_i, by a sequence of local moves of the two types shown in Figure 6. It follows that any quadruple $(V, 1, t, \mu)$ for which t and μ are totally symmetric, while properties (3.1.2) and (3.1.3) hold, may arise as the quadruple associated to a TFT. Such a quadruple forms the structure of a commutative *ambialgebra* in the terminology of [Q].

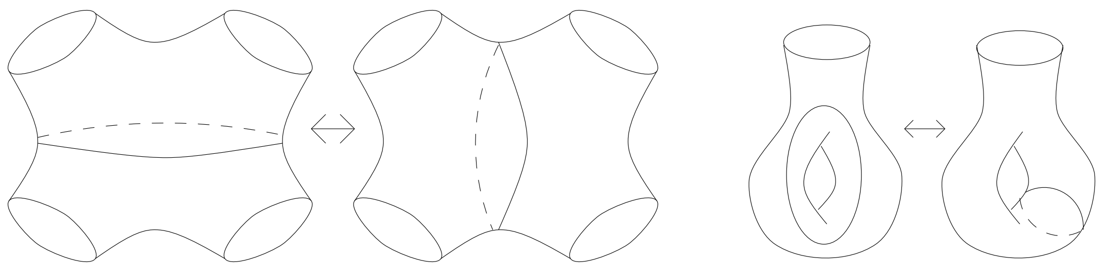

Figure 6: Moves on trinion decompositions

Complete solution

It follows from (3.1.2) that μ defines a non-degenerate inner product on V. Thus μ may be used to identify V with V^*. By contraction with two copies of μ, the element $t \in V \otimes V \otimes V$ gives rise to a map $m: V \otimes V \longrightarrow V$ which may be graphically represented by a trivalent vertex with two incoming arrows and one outgoing arrow. Considering m as a muliplication map, $a \otimes b \longmapsto ab$ makes V into a commutative associative unital algebra. (That $1 \in V$ is a unit, follows from (3.1.2), commutativity follows from the symmetry of t while associativity comes from (3.1.3), which may be graphically depicted as in Figure 7.)

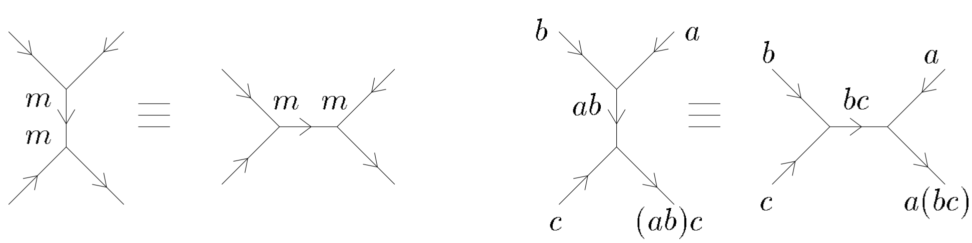

Figure 7: Associativity of m

Assuming that V is finite dimensional, it follows that there exists a basis for V consisting of idempotents $\{w_i\}$, say, for which $w_iw_j=0$ whenever $i\neq j$. The unit here is $1=\sum_i w_i$. Let $\lambda_i=\mu(w_i\otimes w_i)$. Then $t=\sum_i\lambda_i^{-2}w_i\otimes w_i\otimes w_i$. A surface M of genus g and with r boundary components can be decomposed into $2g-2+r$ trinions by $3g-3+r$ interior curves.[12] If ∂M has an orientation which on precisely s components matches that induced by the orientation of M, then $Z(M)$ is given in terms of the basis by,

$$Z(M)=\sum_i(\lambda_i^{-2})^{2g-2+r}\lambda_i^{3g-3+r}w_i^{\otimes r}=\sum_i\lambda_i^{1-g-r}w_i^{\otimes r}\in V^{\otimes r}\cong V^{\otimes s}\otimes(V^*)^{\otimes(r-s)}.$$

This completes the classification of 2–dimensional TFTs and the explicit evaluation of Z.

3.2 Theories derived from triangulations

In this section we construct a particular example of a TFT from algebraic data different from $(V,1,t,\mu)$ of §3.1. This is done by first constructing a TFT over a domain category containing extra data and then applying the inverse limit construction of §2.5.

Identification of domain category $\mathcal{N}'$

The extra data we will employ is that of a triangulation. Thus an object in $\mathcal{N}'_1$ will be a triangulated oriented 1–dimensional closed manifold, Σ. To reduce to combinatorial data, we insist that all our codimension–1 objects be rearranged so as to be unions, $\cup_{i\in[N]}C_i\times\{0\}$, of circles in standard position, as in §3.1. Such objects are specified by a pair (N,a) where $N\in\mathbf{N}\cup\{0\}$ and $a\colon[N]\longrightarrow\mathbf{Z}\backslash\{0\}$. Here $N=c(\Sigma)$ and $a(i)$ is an integer whose absolute value gives the number of vertices (or edges) in the triangulation of the $i^{\rm th}$ boundary component, while the sign gives the orientation. The morphisms in $\mathcal{N}'_1$ are given by permutations of the components.

The objects in $\mathcal{N}'_0$ will be triangulated, oriented, 2–dimensional manifolds with boundary, up to equivalences which preserve the boundary triangulation. More precisely, an object will be an embedding of a triangulated manifold in $\mathbf{C}\times\mathbf{R}^-$, whose boundary is of the form $\cup_{i\in[N]}C_i\times\{0\}$ where $N=c(\partial M)$, up to a change in interior triangulation. The morphisms in $\mathcal{N}'_0$ are topological equivalences of the associated manifolds, up to transformations, which preserve the orientation of the manifold and the combinatorial structure of the boundary triangulation.

[12] For this argument it is necessary to assume that $(g,r)\not\in\{(0,0),(0,1),(1,0)\}$. However it may be separately checked that the result for $Z(M)$ holds in the exceptional cases. Alternatively, for any $u\geq0$, M may be decomposed into $2g-2+r+u$ trinions and u discs by $3g-3+r+2u$ interior curves.

The rest of the domain category structure is much as for $\mathcal{N}$ and so we will not give it in detail except for noting that the operations $*$ and $\amalg$ on $\mathcal{N}'_1$ are given by,

$$(N,a)^* = (N,-a)\ , \qquad (N,a)\amalg(M,b) = (M+N,ab)\ ,$$

where ab denotes the concatenation of the two sequences a and b.

The functor from the domain category $\mathcal{N}'$ to $\mathcal{N}$ is given by forgetting the additional triangulation data. In particular, objects in $\mathcal{N}'_1$ transform according to $(N,a)\longmapsto(N,\alpha)$ where $\alpha(i)=\mathrm{sgn}(a(i))$.

Construction of TFT over $\mathcal{N}'$

Since a triangulated manifold is naturally expressed as the result of gluing[13] together copies of an elementary piece (a triangle), a natural way to construct a theory is to associate a tensor to a triangle and obtain $Z(M)$ by contraction along internal edges.

The data for the theory consists of a vector space, V, along with a tensor $t\in V\otimes V\otimes V$ and a pairing $\mu\colon V\otimes V\longrightarrow \mathbf{C}$. It is assumed that t is invariant under the action of $\mathbf{Z}_3$ by cyclic permutation while μ is symmetric. If Σ is an object in $\mathcal{N}_1$ given by the pair (N,a) then set $Z(\Sigma)=\bigotimes_{i=1}^N V_{\mathrm{sgn}(a(i))}^{\otimes|a(i)|}$. Thus on a positively (negatively) oriented boundary component, there is a copy of V (V^*) assigned to each segment in the triangulation. Again, we make the symmetric group defining the morphisms act by permutation of the tensor product factors.[14]

If M is an object in $\mathcal{N}'_0$, we define $Z'(M)$ as follows. For simplicity, we assume first, that all the boundary components of M are positively oriented. Consider the tensor $t^{\otimes n_2(M)}\in V^{\otimes 3n_2(M)}$. The factors may be put in 1–1 correspondence with pairs (Δ_1,Δ_2) where Δ_1 is an edge in the boundary of the triangle Δ_2 in M.[15] There are, therefore, precisely two factors associated with each internal edge and the result of applying μ on each such pair to $t^{\otimes n_2(M)}$ will be defined to be $Z'(M)$. Observe that this will be an element of $V^{\otimes x}$ where $x=3n_2(M)-2n_1^i(M)=n_1(\partial M)$ and $n_1^i(M)$ denotes the number of internal edges. If ∂M contains components which

[13] Observe that the form of gluing used here involves the identification of part-boundaries, namely along edges of the triangles, and is therefore more general than the gluing entering axiom A3′. Therefore one cannot conclude from the discussion following, that all TFTs can be obtained by the construction of this section, as one could in the case of decompositions into trinions in §3.1. However such a result would exist in the format of 2-ETFTs, see §5.

[14] To be absolutely correct, the morphisms in $\mathcal{N}'_1$ should not only allow permutations of the components but also rotations of the triangulation on each component, so that the correct group is $(\mathbf{Z}_{|a(1)|}\times\cdots\times\mathbf{Z}_{|a(N)|})\rtimes S_N$ rather than S_N. We assume then that $\mathbf{Z}_a$ acts on $V^{\otimes a}$ by cyclic permutation of the factors.

[15] The order of the three factors associated with a triangle is chosen to be consistent with the cyclic order of the edges defined by the orientation on its boundary, induced by that on M. This precisely defines the tensor since t is $\mathbf{Z}_3$ invariant.

are negatively oriented then $Z'(M)$ is defined in the appropriate vector space by using the identification $V \cong V^*$ given by μ.

Since an object in $\mathcal{N}_0$ is only defined up to changes in interior triangulation, it is necessary to check that $Z'(M)$ is well-defined, that is, that it is independent of the choice of triangulation. Any two triangulations of M which agree on the boundary can be obtained from each other by a sequence of local moves, each being of one of the forms depicted in Figure 8. Thus $Z'(M)$ is well-defined so long as the local contractions of tensors remain unchanged under these two moves. Symbolically, this gives the following conditions on t and μ,

$$\begin{aligned} t &= \mu_{24}\mu_{37}\mu_{68}(t\otimes t\otimes t)\in V^{\otimes 3}\,,\\ (\,\mathrm{id}^{\otimes 2}\otimes\mu\otimes\,\mathrm{id}^{\otimes 2})(t\otimes t) &= (\,\mathrm{id}^{\otimes 4}\otimes\mu)P_{15}(t\otimes t)\in V^{\otimes 4}\,,\end{aligned} \tag{3.2.1}$$

where μ_{ij} denotes the action of μ on the $i^{\rm th}$ and $j^{\rm th}$ copies of V and P_{ij} is the transposition of the $i^{\rm th}$ and $j^{\rm th}$ factors. For any triple (V,t,μ) satisfying the above conditions, we have now constructed a TFT over $\mathcal{N}'$.

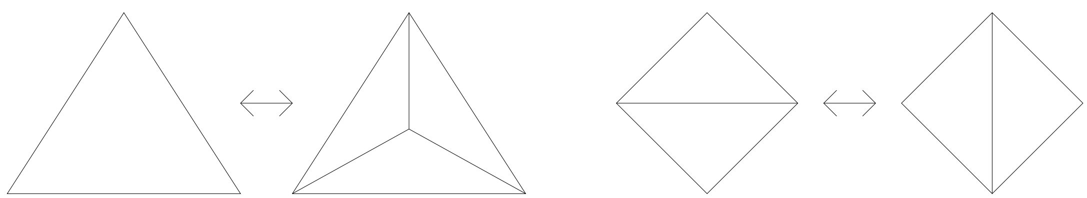

Figure 8: Local moves on triangulations

Graphical presentations

For any triangulated surface, M, the vector $Z'(M)$ has been defined as a contraction of tensors. As in §3.1, this may be graphically represented by a graph, namely the graph, Γ, dual to the triangulation. All internal vertices of Γ have valence 3 (labelled by triangles in the triangulation), while the external legs correspond to the edges in the restriction of the triangulation to ∂M. The tensor t is represented by a trivalent vertex with all arrows directed outwards, while μ is represented by a bivalent vertex with arrows pointing inwards. To obtain $Z'(M)$, a copy of t is placed on each internal vertex, and of μ on each internal edge; external legs have a copy of μ attached if the orientation of the boundary component to which they belong, is negative. Since μ is symmetric while t is only invariant under cyclic permutation of the indices, such graphs must be drawn with attention to the cyclic ordering of the edges eminating from trivalent vertices. Once again, we omit bivalent vertices and internal arrows, since they can be deduced from the vertex types. The constraints given by Figure 8 are now represented by Figure 9.

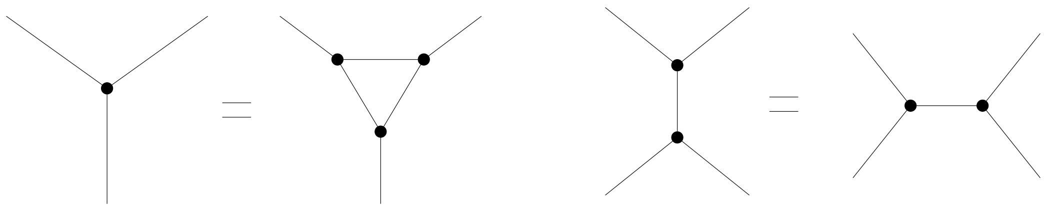

Figure 9: Diagrammatic formulation of relations on t and μ

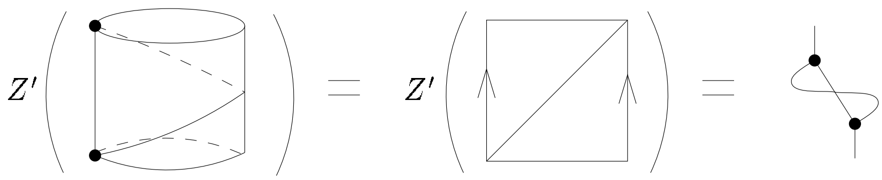

Figure 10: Z'(cylinder)

As an example of the use of this notation, we show graphically that $F_{1,1} \equiv Z'(S^1 \times I)$ is an idempotent. Represent the cylinder by a triangulation in which there are two triangles with precisely one edge in each boundary component, so that $F_{1,1}$ is presented by Figure 10.

Place opposite orientations on the two boundary components, say positive on the upper component, so that $Z'(S^1 \times I)$ describes a map $V \longrightarrow V$ from bottom to top. Then $F_{1,1} \circ F_{1,1}$ is graphically represented by placing two copies of $F_{1,1}$ on top of each other; Figure 11 gives a graphical proof that $F_{1,1}$ is an idempotent, using the relations in Figure 9.

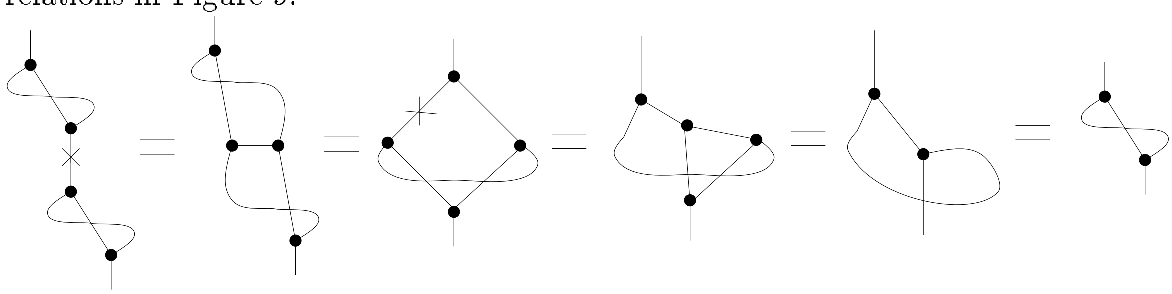

Figure 11: Verification that Z'(cylinder) is an idempotent

The constraints on the triple (V, t, μ) in Figure 9 give rise to an algebraic structure much as in §3.1. Thus, by contracting t with one or two copies of μ, one may obtain maps $\Delta\colon V \longrightarrow V \otimes V$ and $m\colon V \otimes V \longrightarrow V$, see Figure 12.

Figure 12: Construction of m and Δ

It follows from the right hand relation in Figure 9, that m is associative and Δ is coassociative, just as in Figure 7. Since t is only assumed to be invariant under the action of $\mathbf{Z}_3$, m need not be commutative. Also, there is no unit supplied. The two structures m and Δ are not necessarily compatible, in the sense that Δ may not define a homomorphism with respect to m, as can be seen from the example at the end of this section for general finite groups G.

Inverse limit construction

In order to obtain a TFT over $\mathcal{N}$ from Z', it is necessary to apply the inverse limit construction of §2.5. Since all closed codimension–1 objects can be written as disjoint unions of copies of S^1, the images of all such objects under Z are determined by $Z(S^1)$, with positive orientation. An object in $\mathcal{N}'$, representing a positively oriented circle, is specified by the number of segments into whch the circle is subdivided. The image under Z' of a cylinder whose boundary components are oppositely oriented and are subdivided into m and n intervals, the former associated with the negatively oriented boundary component, will be a map $F_{n,m}: V^{\otimes m} \longrightarrow V^{\otimes n}$ and will be independent of the choice of interior triangulation. Picking a simple triangulation gives the presentation of Figure 13, all arrows on external legs pointing upwards. Since Z' defines a TFT over $\mathcal{N}'$, $F_{m,n} \circ F_{n,p} = F_{m,p}$ for all $m, n, p \in \mathbf{N}$.

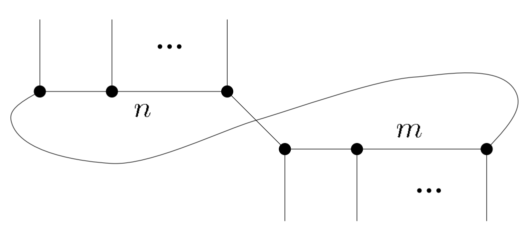

Figure 13: $F_{m,n}$

The vector space, $Z(S^1)$ associated with S^1, generated by the inverse limit construction, is therefore $\ker(F_{1,1} - 1) = \mathrm{Im}\,(F_{1,1}) \subset V$. Suppose that M is an arbitrary two-dimensional oriented surface in $\mathcal{N}_0$, whose boundary components are all positively oriented. Pick a triangulation, $\mathcal{T}$, of M. Then

$$Z(M) = \left(\bigotimes_r F_{n_r,1}\right) Z'(M,\mathcal{T}) \,,$$

where n_r is the number of segments in the r^{th} boundary component, in the restriction of $\mathcal{T}$ to ∂M.

Classification data

We have now constructed a TFT over $\mathcal{N}$, namely Z, from the data of (V,t,μ). By §3.1 any such TFT generates a quadruple $(V',1,t',\mu')$ which describes, in particular, a unital commutative associative algebra. In our case, V' is the image of the map in Figure 10. The element $1 \in V'$ represents the image of a disc under Z. The simplest triangulation of a disc involves just one triangle, giving Figure 14.

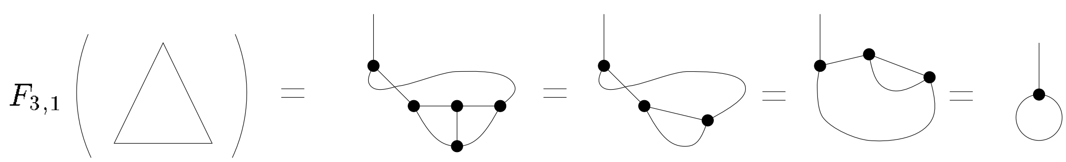

Figure 14: Unit in V'

The element t' comes from the image of a trinion. The simplest triangulation involves five triangles and gives the element of Figure 15.

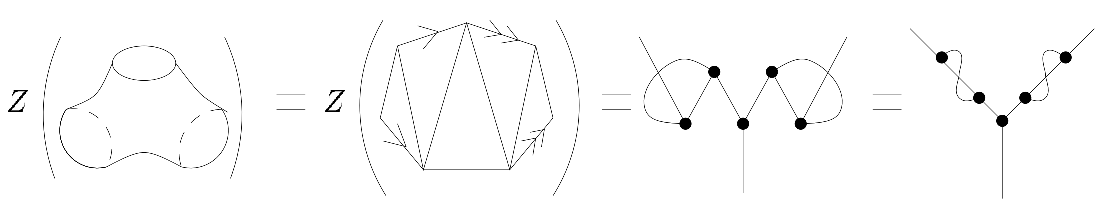

Figure 15: Image of trinion

It may be graphically verified, using manipulations similar to those of Figure 7, that t' defines a commutative algebra, as we know it must, by the analysis of §3.1.

An example

Suppose that G is a finite group and let V denote the vector space spanned by a basis $\{e_g | g \in G\}$, indexed by elements of G. Define,

$$t = |G|^{-2} \sum_{ghk=1} e_g \otimes e_h \otimes e_k \ ,$$
$$\mu(e_g \otimes e_h) = |G| \, \delta_{g=h^{-1}} \ .$$

By the last paragraph, the triple (V, t, μ) defines a TFT, say Z', over $\mathcal{N}'$. Suppose that $\mathcal{T}$ is a triangulation on a manifold M and assume that the orientations on all the boundary components are positive, while $a(i)$ is the number of segments in $\mathcal{T}$ on the i^{th} boundary component. Then $Z'(M) \in V^{\otimes \Sigma a(i)}$. Denote the set of simplices in $\mathcal{T}$ of dimension r by $\mathcal{T}^r$ and the set of edges in the restriction of $\mathcal{T}$ to ∂M by $\mathcal{T}^1_e$. Let $\mathcal{T}^1_{\text{or}}$ denote the set of pairs $\{(u, v) | u \in \mathcal{T}^1, v \in \mathcal{T}^2, u \subset v\}$; this may be thought of equivalently, as the set of oriented edges of $\mathcal{T}$, each internal edge appearing twice, while boundary edges appear only once.[16] Then, by construction,

$$Z'(M) = \sum_{p_e : \mathcal{T}^1_e \longrightarrow G} \Big(\bigotimes_{r \in \mathcal{T}^1_e} e_{p(r)} \Big) |G|^{|\mathcal{T}^1_i| - 2|\mathcal{T}^2|} \ . \ \begin{pmatrix} \#\text{ maps } p \colon \mathcal{T}^1_{\text{or}} \longrightarrow G \\ \text{compatible with } p_e \text{ and } \mathcal{T} \end{pmatrix} .$$

Here, by compatibility with $\mathcal{T}$, we mean that p assigns inverse elements to any two elements of $\mathcal{T}^1_{\text{or}}$ given by the same edge with opposite orientations, while the product of the elements associated with the three pairs (u, v), defined by a triangle v is $1 \in G$. Compatibility with p_e means that the restriction of p to the boundary is p_e. This formula may be rewritten as,

$$Z'(M) = \sum_{p : \mathcal{T}^1_e \longrightarrow G} \Big(\bigotimes_{r \in \mathcal{T}^1_e} e_{p(r)} \Big) |G|^{|\mathcal{T}^2| - |\mathcal{T}^1| + |\mathcal{T}^0_i| - c(M) + c(\partial M)} \prod_{C_i} |[p_{C_i}]|^{-1}$$
$$\cdot \ (\#\text{ maps } \rho \colon \pi_1(M) \longrightarrow G \text{ such that } [\rho(C_i)] = [p_{C_i}] \ \forall i) \ ,$$

where $[a]$ denotes the conjugacy class of $a \in G$, C_i is the i^{th} boundary component and $[p_{C_i}]$ denotes the conjugacy class of the product of the images under p of the segments in C_i.

The maps $F_{n,1} \colon V^{\otimes n} \longrightarrow V$ are given by

$$e_{a_1} \otimes \cdots \otimes e_{a_n} \longmapsto f_{[a_1 \ldots a_n]} \ ,$$

---

[16] An element (u,v) of $\mathcal{T}^1_{\text{or}}$ is represented by the edge u with orientation given by the restriction of the orientation induced by v on its boundary, ∂v.

where $f_\alpha = |\alpha|^{-1} \sum_{a\in\alpha} e_a$. Hence the inverse limit construction generates $Z(S^1) = \langle f_\alpha \rangle$. The resulting theory, Z, has

$$Z(M) = |G|^{\chi(M)-c(M)-c(\partial M)} \sum_b (\otimes_{C_i} e_{b(i)}) \cdot \begin{pmatrix} \#\text{ maps } \rho\colon \pi_1(M) \longrightarrow G \\ \text{such that } [\rho(C_i)] = b_{C_i}\ \forall i \end{pmatrix} .$$

where the sum is over all maps b from the set of components of ∂M to the set of conjugacy classes in G. On closed surfaces, this reduces to $|G|^{\chi(M)-c(M)}$ times the number of homomorphisms $\pi_1(M) \longrightarrow G$.

The commutative unital associative algebra structure associated with this theory by §3.1, is the natural group algebra on G restricted to $\langle f_\alpha \rangle$, with unit f_1. Equivalently, it has a basis whose coordinates with respect to $\{e_g\}$ are given by the characters of irreducible representations of G, so that the algebra is just the representation ring of G. Using the latter basis and the orthogonality of characters, $Z(M)$ may be easily computed for closed surfaces; for a connected closed surface of genus g,

$$Z(M) = \sum_\alpha |\alpha|^{1-g} ,$$

where the sum is over all conjugacy classes α. Hence we obtain the equality,

$$\sum_\alpha |\alpha|^{1-g} = |G|^{1-2g} |\operatorname{Hom}(\pi_1(M), G)| ,$$

between the two different expressions for $Z(M)$, the left hand side obtained from the trinion decomposition and the right hand side from a triangulation. It is instructive to explicitly verify these equalities for small genus surfaces, starting with the sphere and torus.

4: Category structures

The initial motivation for the axioms of §2 comes from physical ideas in field theory, where the image under Z of a cobordism, M from Σ_1 to Σ_2, is thought of as the operator defining the propagation across M of fields on Σ_1 to those on Σ_2. The slices Σ_i can be thought of as time-slices and M is the whole of space-time. In a topological field theory, there is no direction which is specially picked out as labelled by time, and indeed there is no particular reason why one should only consider decompositions by codimension–1 submanifolds.

Consider a topological field theory in dimension d. Suppose that M is a d–dimensional closed manifold and that it is split into two parts M_1 and M_2 by a $(d-1)$–dimensional manifold Σ. Then,

$$Z(M) = \big\langle Z(M_1) \bigm| Z(M_2) \big\rangle ,$$

where $Z(M_1) \in Z(\Sigma)$ and $Z(M_2) \in Z(\Sigma)^*$. Hence, as was seen in §3.1, $Z(M)$ can be completely determined if it is possible to subdivide d–dimensional manifolds into elementary pieces and Z is known on each of these pieces. Suppose now that Σ is split into Σ_1 and Σ_2 by a $(d-2)$–dimensional manifold C. We would like to be able to express $Z(\Sigma)$ as $\langle Z(\Sigma_1)|Z(\Sigma_2)\rangle$, where $Z(\Sigma_1)$ and $Z(\Sigma_2)$ are suitable objects in dual spaces $Z(C)$ and $Z(C)^*$, respectively. However, $Z(\Sigma)$ is to be a vector space, and so the question arises as to what sort of beast $Z(C)$ could be, so that an element of it can be paired with an element of a dual contraption to give a vector space. The answer we employ is that of a 2–vector space, and more generally, an n–vector space for codimension–n manifolds. This is our first motivation for the study of higher categories, since their definition naturally comes in the form of a structure on the collection of all n–vector spaces, analogous to the category structure on the collection of all vector spaces.

Just as domain categories were used in §2 to formalise the possible gluing operations appearing in TFTs, so more complex structures appear when more general gluing operations are allowed and provide a second motivation for the consideration of higher categories. Thus, allowing gluing down to codimension–2 manifolds, provides three types of objects, namely d–dimensional manifolds, $(d-1)$–dimensional manifolds and $(d-2)$-dimensional closed manifolds, forming sets $\mathcal{M}_0$, $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively. There is now a particular object $\emptyset$ in each of these sets, while there are operations * on $\mathcal{M}_1$, $\mathcal{M}_2$ and $\amalg$ on all three sets. The sets are linked via two boundary functors, operations $\times I$ and numerous gluing maps, which allow gluing of objects at one codimension, along objects of higher codimension (up to a maximum of 2). Thus we allow only the operation of disjoint union on manifolds of codimension–2, while manifolds of codimension–1 may be glued along closed parts of their boundaries and top dimensional manifolds may be glued at corners.[17]

4.1 Higher category structures

As discussed in §2.4, a category may be viewed as a pair of sets along with various maps, namely source, target, identity and composition. The simplest way to extend a category structure to higher dimensions, may be obtained by trying to fit it to a generalisation of Example 2.4.7, in which points and paths are generalised to embeddings of closed balls, D^p. Observe that such a closed ball can be decomposed into open cells,

$$A^p \cup A_0^{p-1} \cup A_1^{p-1} \cup \ldots \cup A_0^0 \cup A_1^0 \,, \tag{4.1.1}$$

where $A^s_{\cdot}$ denotes an embedding of a ball, B^s, whose boundary $\partial A^s_{\cdot}$ is given by the union of all A^t_i with $t < s$ and $i = 0, 1$.

[17] In this respect, we view disjoint union as a form of gluing along an empty object at the next higher codimension, although we also allow it on the top codimension.

The formalisation of the geometry just outlined, leads to the consideration of a structure known as a *spheric set*, see [MS]. Here there are objects of multiple types, forming sets $C^0, \ldots, C^n$, along with source, target and identity maps,

$$\begin{aligned} s_k, t_k &: C^k \longrightarrow C^{k-1}\,, &\qquad (1 \le k \le n)\,; \\ i_k &: C^k \longrightarrow C^{k+1}\,, &\qquad (0 \le k < n)\,. \end{aligned}$$

These maps must be compatible in the sense that the following equalities of compositions of maps hold, where, in each equality, k may range over all values for which one side of the equality is meaningful.

(a) $s_{k-1} \circ s_k = s_{k-1} \circ t_k$, $t_{k-1} \circ s_k = t_{k-1} \circ t_k$.

(b) $s_{k+1} \circ i_k = \mathrm{id}_{A_k} = t_{k+1} \circ i_k$.

Such a structure can be visualised in terms of a CW–complex, in which to each element of C^p there is associated a copy of D^p decomposed as a union of open cells, as in (4.1.1). The maps s and t take such an element to the two halves of the spherical boundary, namely

$$A_i^{p-1} \cup A_0^{p-2} \cup A_1^{p-2} \cup \ldots \cup A_0^0 \cup A_1^0\,,$$

with s being given by $i = 0$ and t by $i = 1$. Properties (a) and (b) allow one to construct maps $s_{p,q}, t_{p,q}: C^p \longrightarrow C^q$ whenever $n \ge p > q \ge 0$ by composing s and t maps; geometrically, the images of the object of (4.1.1) under these maps are the closures of A_0^q and A_1^q, respectively.

To get the notion of a higher category from that of a spheric set, it is necessary to require that there also be a collection of composition maps, $\mu_{p,q}$, defined for all $n \ge p > q \ge 0$,

$$\{(u,v) \in C^p \times C^p | s_{p,q}(v) = t_{p,q}(u)\} \longrightarrow C^p\,,$$

which we denote by $(u,v) \longmapsto v \circ_q u$. The source of $v \circ_q u$ is given in terms of those of u and v by,

$$s_p(v \circ_q u) = \begin{cases} s_p(u), & \text{if } q = p-1; \\ s_p(v) \circ_q s_p(u), & \text{if } q < p-1, \end{cases}$$

and similarly for the target. The composition maps are also required to satisfy certain identity and associativity properties generalising those for standard categories. The notion we arrive at, is known as that of a *small n–category*. Its full definition may be found in [MS].

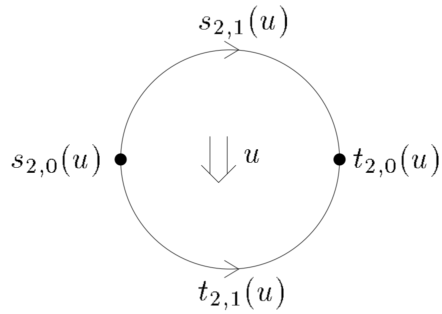

Figure 16: 2–objects in a 2–category

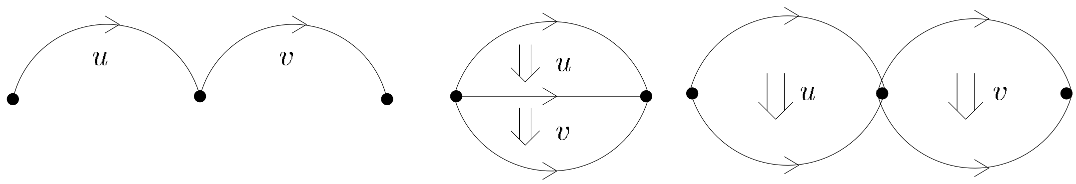

Figure 17: Compositions in a 2–category

In accordance with the above geometrical imagery, compositions of type p objects whose source and target match at the level of type q objects, is represented by gluing along the common q–dimensional disc. Thus in a small 2–category, there are three object types, namely elements of C^0, C^1 and C^2 and the confusing terminology of, objects, 1–morphisms and 2–morphisms, respectively, is often adopted for elements of these sets. They are represented by points, arcs and 2–cells. See Figure 16 for the standard representation of an element of C^2. There are three possible compositions of objects in a 2–category, namely, $\mu_{1,0}$, $\mu_{2,0}$ and $\mu_{2,1}$ which are illustrated in Figure 17.

Independence of the result of a method for evaluating the composition of a collection of compatible objects, upon the order in which the composition maps are applied, is ensured generally by the truth of this statement for the two particular arrangements of objects shown in Figure 18.

The higher category structure we have just constructed, is said to be *strict*, because it is required that the composition maps be absolutely associative. In weaker forms of category structures, it is not required that the equalities of compositions hold strictly, but merely that there are morphisms, given as part of the category structure, between any two such compositions. Thus for example, in a weak 2–category, it is not required that $\mu_{1,0}$ is strictly associative, but rather that, for each

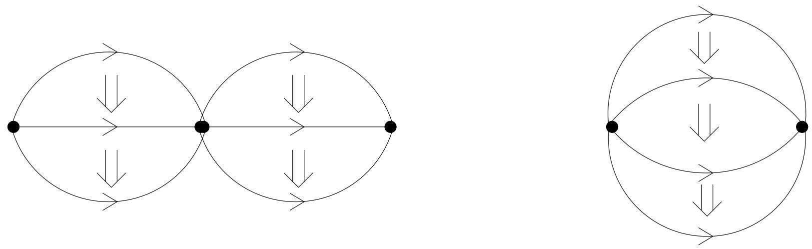

Figure 18: Associativity constraints in a 2–category

triple of compatible 1–objects, u, v and w, an invertible[18] 2–object $c_{u,v,w}$, known as an *associativity constraint*, is given, whose source and target are $(w \circ_0 v) \circ_0 u$ and $w \circ_0 (v \circ_0 u)$, respectively. It is further required that this 2–object behaves 'naturally', in a suitable sense, under morphisms of u, v and w, and that given four compatible 1–objects u, v, w and t, a compatibility condition holds amongst the possible associativity constraints. This is illustrated in Figure 19, where the symbol '$\circ$' has been used to denote $\circ_0$. The structure now obtained is known as a *bicategory*, see [B].

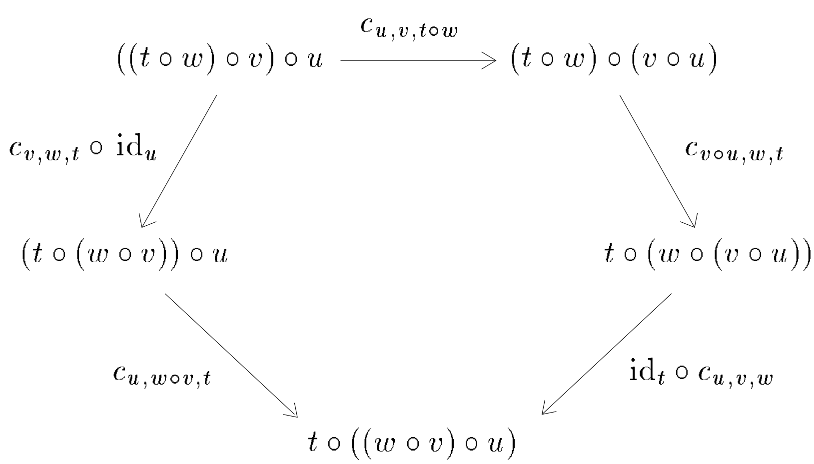

Figure 19: The pentagon relation in a bicategory

[18] The notion of invertibility of an r–object, x, for $r>0$ is the formal analogue in higher categories of requiring a morphism to be an isomorphism in a usual category. Thus, it requires the existence of another r–object, y, whose source and target are those of x reversed and is such that the compositions $x \circ_{r-1} y$ and $y \circ_{r-1} x$ are both isomorphic to identity r–objects on appropriate $(r-1)$–objects. In the case of Example 2.4.6, the notion of invertibility of a morphism reduces precisely to the requirement that the associated map is literally an isomorphism.

Although strict categories suffice[19] for the purposes of constructing higher vector spaces, it is a higher category in the weaker sense which is needed for the higher version of a domain category, just as it was necessary in a domain category to consider isomorphisms between manifolds rather than equalities. Unfortunately, although the notion of a usual category is perfectly well defined, there is at present no unique notion of a maximally weak higher category. A number of different notions of higher category, along with 'categorifications' of other algebraic structures, have been constructed piecemeal over the last 30 years. Although they all seem to fit into the same general framework, it is not currently possible to completely automate the process of categorification, some manual intervention being necessary in each case to decide precisely which of several possible slightly different structures is appropriate. It is, however, instructive to give the general idea of how, starting from the notion of one algebraic structure, another notion 'at the next level of categorification' may be constructed. (I believe this basic idea is due to Grothendieck.)

Suppose that $\mathcal{A}$ denotes a *formal* algebraic structure, that is, it is not a specific example, but rather it is a meta-object defining the notion of that particular type of structure. Hence it consists of a collection of sets, a collection of maps defined on sets generated by the basic sets and a collection of axioms which must be satisfied by these maps. Examples of such meta-objects include the notions of a set, groupoid, group, ring and even of a category. The collection of axioms may be thought of as a generating set with respect to our deduction system, for the set of all universally valid statements in a type $\mathcal{A}$ structure. The categorification process replaces types of object in the definition of $\mathcal{A}$ by other types, according to the list below.

A set A	$\longrightarrow$	a category A.
A map $f\colon A\to B$	$\longrightarrow$	a functor $f\colon A\to B$.
An element $x\in A$	$\longrightarrow$	an object $x\in A$.
An equality $a=b$ with proof P	$\longrightarrow$	an invertible morphism $c_P\in \mathrm{Morph}\,(a,b)$.
Concatenation of proofs	$\longrightarrow$	composition of morphisms.
Two different proofs P,Q of the same statement	$\longrightarrow$	equality of morphisms $c_P=c_Q$.

By a proof is meant a sequence of deductions, each step being an instance of one of the axioms of $\mathcal{A}$, or of the basic deduction system. The relations in the new structure thus arise as relations amongst relations in the old structure. The ambiguity in the new structure comes from the freedom available in deciding exactly what deductions are allowed in a proof, along with the fact that the above process may give inequivalent outputs starting from equivalent initial structures.[20] By this

[19] Actually a complication arises with the identification of the order of the basis of a tensor product of three higher vector spaces.

[20] For example, starting with the notion of a category, as given in Definition 2.4.1, the process generates the notion of a bicategory. However, if we choose to define a category according to Example 2.4.3 then the resulting structure will be slightly different, since a composition will be defined whenever there are given objects $u,v\in C^1$ and a morphism $f\in \mathrm{Morph}_{C^0}(s(u),t(v))$, rather than only when the source of one object matches the target of the other.

type of procedure, category structures have been generated from many different algebraic structures.

set	$\longrightarrow$	category
groupoid	$\longrightarrow$	monoidal category
abelian groupoid	$\longrightarrow$	braided monoidal category
group	$\longrightarrow$	rigid monoidal category
ring	$\longrightarrow$	ring category [21]
module over a ring	$\longrightarrow$	module category over a ring category
category	$\longrightarrow$	bicategory
braided monoidal category	$\longrightarrow$	braided monoidal 2–category [22]
bicategory	$\longrightarrow$	tricategory [23]
Hopf algebra	$\longrightarrow$	Hopf category [24]

4.2 Higher vector spaces

Just as ordinary vector spaces can be considered as objects in an appropriate category, $\mathcal{V}^1$, higher vector spaces appear as objects in an appropriate higher category[25] $\mathcal{V}^n$. To illustrate the structure, we will discuss the case of 2–vector spaces only. Formally, one should think of a 2–vector space as a linear space over the category of vector spaces, just as an ordinary vector space is a linear space over the field of complex numbers. Everywhere that $\mathbf{C}$ previously entered, it is replaced by $\mathcal{V}^1$, while the operations of addition and multiplication in $\mathbf{C}$ are replaced by direct summation and tensor product of vector spaces.[26]

Up to isomorphism, a 2–vector space is specified by its rank, a non-negative integer.[27] Thus, if $\mathcal{V}$ is a 2–vector space of rank n, then an identification with the standard rank n 2–vector space is equivalent to a choice of basis, $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for $\mathcal{V}$. With respect to this basis, any element, V, of $\mathcal{V}$ is specified by its coordinates, a list of n vector spaces, $(V_1, \ldots, V_n)$, and this fact may be formally denoted by,

$$V = \bigoplus_{i=1}^{n} V_i \cdot \mathbf{e}_n \ ,$$

---

[21] See [K], [Lap].

[22] Here there is some ambiguity as to the 'right' definition. See [KV2] for the best list currently known.

[23] See [GPS].

[24] See [CF].

[25] The higher dimensional analogue of the category $\mathcal{V}^1$ of uncoordinatised vector spaces, should be a higher category in the as yet undefined weak sense. To avoid this problem, we only discuss the analogue of $\mathcal{V}^1_c$ in this section. We also take $\mathbf{C}$ as the ground field.

[26] It is actually possible to make this more precise by defining a 2–vector space to be a module category over the ring category of vector spaces. See [KV2].

[27] In categorical language this reads: up to module equivalence, a 2-vector space may be identified with the module category $(\mathcal{V}^1)^n$, of rank n, over $\mathcal{V}^1$, for some non-negative integer n.

where '$\cdot$' is the analogue of scalar multiplication. If $\mathcal{V}$ and $\mathcal{W}$ are two 2–vector spaces, of ranks n and m respectively, then a morphism, S, between them should be a linear transformation. In terms of bases it is specified by an $m \times n$ matrix of vector spaces, $S_{i,j}$, such that if $S(V_1, \ldots, V_n) = (W_1, \ldots, W_m)$ then

$$W_i = \bigoplus_{j=1}^{n} S_{i,j} \otimes V_j \ .$$

Two morphisms S and T from $\mathcal{V}$ to $\mathcal{W}$ cannot be considered as equal, as it is only possible to talk about (2–) morphisms between them. Such a 2–morphism would be given by a family of linear transformations $S_{i,j} \longrightarrow T_{i,j}$ indexed by the pair (i,j). That is, in terms of coordinates, a 2–morphism is given by a matrix of matrices.

To make the form of the 2–category structure, $\mathcal{V}_c^2$, on the collection of totally coordinatised 2–vector spaces clearer, we now give a purely combinatorial list of the sets of objects. A 0–object will be an element $n^0 \in \mathbf{N} \cup \{0\}$. A 1–object will consist of a triple, (n_0^0, n_1^0, n^1) where $n_i^0 \in \mathbf{N} \cup \{0\}$ $(i = 0, 1)$ and n^1 is an $n_1^0 \times n_0^0$ matrix whose entries are elements of $\mathbf{N} \cup \{0\}$.[28] A 2–object will consist of five pieces of data

$$(n_0^0, n_1^0, n_0^1, n_1^1, n^2) \ ,$$

where n_i^0 are non-negative integers, n_i^1 are $n_1^0 \times n_0^0$ matrices with non-negative integer entries and n^2 is an $n_1^0 \times n_0^0$ matrix, whose $(i,j)^{\text{th}}$ entry is an $n_1^1(i,j) \times n_0^1(i,j)$ matrix with complex entries. The source and target maps are given by forgetting parts of the information, while the composition operations are generalisations of matrix multiplication.

This definition can be generalised to totally coordinatised n–vector spaces without any major changes. However, as was mentioned above, in real applications, what is wanted is a category of uncoordinatised higher vector spaces, which must necessarily form only a weak higher category.

5: Extended Topological Field Theories

The axioms of §2 where formulated by considering the properties that would be expected of Feynman integrals over manifolds with boundary, under various gluing operations. As noted at the start of §4, if objects are to be associated with codimension-1 manifolds with boundary, in such a way that they behave reasonably under gluing, then these objects must be elements of a 2–vector space.

Let us fix the top dimension, d, of the theories under consideration (that is, the dimension at which manifolds are assigned vectors). By an r–ETFT (extended

---

[28] This is because, up to isomorphism, a vector space is specified by its dimension, so that a matrix of vector spaces turns into a matrix of dimensions.

topological field theory) we mean an assignment of objects to manifolds of codimension up to r, which behaves naturally under gluing operations. This assignment will give,

codimension k closed manifold, Σ	$\longrightarrow$	k–vector space $\overline{Z}_k(M)$, for $0 \leq k < r$.
codimension k arbitrary manifold, M	$\longrightarrow$	element $Z_k(M)$ of the $(k+1)$–vector space $\overline{Z}_{k+1}(\partial M)$, for $0 \leq k \leq r$;

The only gluing operations allowed are those in which all parts of the gluing are expressed in terms of manifolds of codimension at most r, while those manifolds which appear at codimension r are all closed.

Notice that in §4.2, the collection of k–vector spaces was considered as forming a k–category, $\mathcal{V}^k$, while a particular k–vector space was considered as a linear space over $\mathcal{V}^{k-1}$ in place of $\mathbf{C}$. An element of a coordinatised k–vector space of rank m is specified by a list of m, $(k-1)$–vector spaces, namely, the *coordinates* of the element with respect to a basis. This statement at $k = 2$ translates into the fact that a 2–vector space is a (1–) category and that, with respect to a basis, an element of a 2–vector space is given by a sequence of vector spaces. Pushing down to $k = 1$, we see that a 1–vector space is an ordinary vector space and, with respect to a basis, an element is specified by a sequence of complex numbers. In this sense we think of a 0–vector space as being a complex number, while the collection of 0–vector spaces forms the set (= 0–category) $\mathbf{C}$.

When $r = 1$, the structure of an r–ETFT therefore associates complex numbers and vector spaces to closed codimension–0 and 1 manifolds, while associating vectors to arbitrary codimension 1–manifolds; thus a 1–ETFT is just an ordinary TFT. For ordinary TFTs, these three structures are not all independent, in the sense that the vector associated to a codimension–1 manifold, M, in the case when M is closed, is an element of $Z(\partial M) = Z(\emptyset) = \mathbf{C}$. In a similar way, the structures associated with arbitrary manifolds and with closed manifolds are identified, by the requirement that $\overline{Z}_k(\emptyset) = \mathbf{1}_k$, the one–dimensional k–vector space, that is, it has rank one over $\mathcal{V}^{k-1}$. Thus, if M is a closed codimension k manifold, then $Z_k(M) \in \mathbf{1}_{k+1}$, so that it is a k–vector space, and we require that it be precisely $\overline{Z}_k(M)$.

In principle, it is fairly straightforward to define the notion of an r–domain category, to be the generalisation of the notion of a domain category (for $r = 1$), which includes all the object types above and the allowed gluing operations which may be performed upon them. A d–dimensional r–ETFT would then just be a functor $\mathcal{M}^r \longrightarrow \mathcal{V}^r$, where $\mathcal{M}^r$ is the r–domain category of manifolds of codimension up to r, with base dimension d, and $\mathcal{V}^r$ is the r–domain category of r–vector spaces. By an ETFT we will mean a d–ETFT, that is, it will assign structures to submanifolds all the way down to a point. Since, when gluing at arbitrary corners is allowed, it is possible to subdivide any manifold into fixed basic pieces, namely simplices, an ETFT can be expressed in terms of the image of a simplex and the gluing rules. This approach shifts all the complication into the gluing rules, since unlike the case of TFTs, the gluing operations may now involve corners and it is important to know the local structure of the manifolds being glued near such corners.

If $\mathcal{T}$ is a triangulation of a manifold M, then a subdivision which shows all the complexity of the gluing operations in the neighbourhood of subsimplices of $\mathcal{T}$ is obtained by a construction known as the dual barycentric subdivision. It produces a subdivision $\mathcal{T}^*$, for which there are 1–1 correspondences,

$$\begin{aligned} \text{top dimensional cells} &\longleftrightarrow \text{simplices and subsimplices,} \\ \text{vertices} &\longleftrightarrow \text{complete flags of simplices,} \\ \text{edges} &\longleftrightarrow \text{partial flags of simplices,} \end{aligned}$$ [29]

in which the left hand column represents structures in $\mathcal{T}^*$ and the right hand column is relevant to $\mathcal{T}$. For each subsimplex, Δ, in $\mathcal{T}$, the geometry of the associated top dimensional cell in $\mathcal{T}^*$, precisely specifies the complexity of the gluing operation required in the neighbourhood of Δ. By decomposing these cells into a small number of elemental pieces, it is possible to specify an ETFT by giving the structures associated with only a finite number of geometric types of cell. See [L2].

Writing out all of the gluing data in terms of bases, one obtains the following coordinatised description of an ETFT. It consists of a large quantity of data, sets and numbers. It is simplest to think in terms of cells that may arise in the decomposition $\mathcal{T}^*$, for a generic $\mathcal{T}$, since these describe all the required gluing operations. The basic set-data is recursively given as a set of labels that may be assigned to vertices, along with, for each geometric type of face, of dimension $k < d$, a set of labels depending on a prior choice of labels on all the facets of the face. Finally, for each possible d–dimensional cell type, there is a weight supplied by the data, whenever all its subcells have been labelled by elements of the appropriate labelling sets. From such data, the scalar assigned to a triangulated top dimensional manifold is,

$$Z(M) = \sum_{\substack{\text{labellings} \\ \sigma}} \prod_{\substack{\text{subsimplices} \\ \Delta \in \mathcal{T}}} (\text{weight assigned to cell labelled } \Delta \text{ in } \mathcal{T}^*) \,,$$

where σ ranges over all allowed labellings of the subcells of $\mathcal{T}^*$. This expression may be viewed as a discretised form of a Feynman integral.

The example of a TFT constructed in §3.2 may be viewed in this framework as follows. To vertices of $\mathcal{T}^*$, there is associated a labelling set with only a single

[29] A flag of simplices is a sequence $\Delta_0 \subset \Delta_1 \subset \cdots \subset \Delta_k$, for which each Δ_i is a subsimplex of the complex $\mathcal{T}$. A flag is called complete if $k = \dim M$ and $\dim(\Delta_i) = i$ for all i. The type of a flag refers to the list of dimensions of the simplices involved.

element. To edges, there is associated a labelling set with $\dim V$ elements, while the weights on 2–dimensional cells are the matrix elements of t or μ, according to the cell type. Of course, just as there were conditions on t and μ in §3.2, in order for a set of combinatorial data to define an ETFT, certain conditions must be satisfied which guarantee the independence of the result upon the choice of triangulation. Fortunately, there is a simple set of local moves on triangulations which generates the equivalence (see [Ma], [P]) and therefore the constraints on the data may be written down as a finite set of conditions similar to (3.2.1). It is also possible to recover a full topologically invariant ETFT by using an inverse limit construction similar to that employed in §3.2. For more details of small dimensional ETFTs, the reader is referred to [L2].

6: Conclusion

What we have indicated in these notes is how complicated algebraic and categorical structures can arise out of the consideration of gluing rules in higher dimensions. The study of topological field theory is really the study of space, and to this extent, any structure appearing in a classification of TFTs may be thought of as dual to an appropriate structure on higher dimensional geometric objects.

We have constructed our algebraic structures in a fairly naive way, throwing in all the structure that must be present and then constraining it by the obvious conditions. The artful part comes next, namely extracting that part of the structure which is essentially interesting. Thus, just as in §3.1 it was seen that the essential algebraic structure in 2–dimensional TFT was that of a commutative unital associative algebra, in 3–dimensional TFT, it is a quasi-Hopf algebra.[30]

The relation between algebraic and geometric structures is nicely illustrated with reference to the Yang–Baxter equation. It plays a central role in the theory of quantum groups and so one would expect to be able to realise its geometrical significance. Indeed, all compact connected 3–manifolds may be obtained by surgery round suitable links in S^3. Any link can be represented as the closure of an appropriate braid, and braids on a given number of strands form a group whose main relation is a variant on the Yang–Baxter equation. In fact, historically, it is this latter connection which came first.[31]

In the search for the construction of interesting higher dimensional TFTs, this mirroring of the geometry in algebra has been exploited to try to guess at appropriate algebraic structures, though the 'right' one has yet to be found. The simplest higher dimensional geometric objects are simplices and hypercubes, and various approaches at algebraic and categorical constructions to encapsulate their

---

[30] See [F] for a differential geometric approach to the construction of the quasi-Hopf algebra associated to a TFT.

[31] See [RT] and Reshetikhin's session in this Short Course.

combination along codimension–1 faces have been investigated.[32] There have been various attempts at the construction of higher dimensional versions of the Yang–Baxter equation and braid groups, some of which arise both in connection with physics and the geometry of knots in higher dimensions.[33] Using purely categorical constructions, it is possible to construct from a suitably constrained braided monoidal tensor category, a four–dimensional TFT, but unfortunately the invariants it defines are expressible in terms of simple classical invariants of manifold such as the signature.[34] There have also been attempts, using the general language of §4.1, to construct category structures relevant in four dimensions, from those known to be relevant in three dimensions, such as [KV2] for braided monoidal 2–categories and [CF] for Hopf categories.

Acknowledgements The author would like to thank IHES for their generous hospitality during the writing of these notes and for support from a Raymond and Beverly Sackler Fellowship.

References

[A1] M.F. Atiyah, 'The Geometry and Physics of Knots', *Lezione Lincee* Cambridge University Press (1990).

[A2] M.F. Atiyah, 'Topological quantum field theory', *Publ. Math. IHES* **68** (1989) p.175–186.

[B] J. Bénabou, 'Introduction to bicategories', *Lect. Notes in Math.* **47** (1967) p.1–77.

[CS] J.S. Carter, M. Saito, 'Reidemeister moves for surface isotopies and their interpretation as moves to movies', *J. Knot Th. Ramif.* **2** (1993) p.251–284.

[CY] L. Crane, D. Yetter, 'A categorical construction of 4D topological quantum field theories', *Quantum Topology* World Scientific (1993) p.120–130.

[CF] L. Crane, I.B. Frenkel, 'Hopf categories', *preprint* (1994).

[Fi] J.E. Fischer, '2–Categories and 2–Knots', *preprint* (1993).

[F] D.S. Freed, 'Higher algebraic structures and quantization', *Commun. Math. Phys.* **159** (1994) p.343–398.

[32] See for example [KV1], [L1] and [St].

[33] See [Z1], [Z2] for the physical background behind the Zamolodchikov tetrahedron equation. In [MS], a family of higher braid groups is constructed. In [KV2], [L1] notions of a braided monoidal 2–category and a 3–algebra, respectively, are introduced and related to different forms of generalisation of the Yang–Baxter equation. See [CS], [Fi] for category structures arising from 2–knots in 4–dimensions.

[34] See [Y], [KR] for purely categorical constructions of 3–manifold invariants. Using similar methods, [CY] and [Y2] construct 4–manifold invariants, while [R] gives the basic proof that these invariants are 'trivial'.

[GPS] R. Gordan, A.J. Power, R. Street, 'Coherence for tricategories', *Memoirs A.M.S., to appear.*

[HOMFLY] P. Freyd, D. Yetter, J. Hoste, W.B.R. Lickorish, K.C. Millett, A. Ocneanu, 'A new polynomial invariant of knots and links', *Bull. A.M.S.* **12** (1985) p.239–246.

[J1] V.F.R. Jones, 'A new knot polynomial and von Neumann algebras', *Notices A.M.S.* **33** (1986) p.219–225.

[J2] V.F.R. Jones, 'Hecke algebra representations of braid groups and link polynomials', *Ann. of Math.* **126** (1987) p.335–388.

[KR] D. Kazhdan, N.Yu. Reshetikhin, *In preparation.*

[KV1] M.M. Kapranov, V.A. Voevodsky, 'The free n-category generated by a cube, oriented matroids and higher Bruhat orders', *Funct. Anal. Appl.* **25** (1991) p.62–65.

[KV2] M.M. Kapranov, V.A. Voevodsky, '2-Categories and Zamolodchikov tetrahedra equations', *Proc. Symp. Pure Math.* **56(2)** (1994) p.177–259.

[K] G. Kelly, 'Coherence theorems for lax algebras and distributive laws', *Lect. Notes in Math.* **420** (1974) p.281–375.

[Lap] M. Laplaza, 'Coherence for distributivity', *Lect. Notes in Math.* **281** (1972) p.29–65.

[L1] R.J. Lawrence, 'On algebras and triangle relations', *Topological and geometrical methods in field theory* World Scientific (1992) p.429–447.

[L2] R.J. Lawrence, 'Triangulations, categories and extended topological field theories', *Quantum Topology* World Scientific (1993) p.191–208.

[M] S. MacLane, 'Categories for the working mathematician', *GTM* **5** (1971).

[MS] Yu.I. Manin, V.V. Schechtman, 'Arrangements of hyperplanes, higher braid groups and higher Bruhat orders', *Adv. Studies in Pure Math.* **17** (1989) p.289–308.

[Ma] S. Matveev, 'Special spines of piecewise linear manifolds', *Mat. Sbornik* **92** (1973) p.282–293.

[P] U. Pachner, 'PL Homeomorphic manifolds are equivalent by elementary shellings', *Europ. J. Comb.* **12** (1991) p.129–145.

[PT] J.H. Przytycki, P. Traczyk, *preprint* (1985).

[Q] F. Quinn, 'Lectures on axiomatic topological field theory', *preprint* (1992).

[R] J. Roberts, 'Skein theory and Turaev-Viro invariants', *preprint* (1993).

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

[S] G.B. Segal, 'The definition of conformal field theory', *preprint*.
[St] R. Street, 'The algebra of oriented simplices', *J. Pure Appl. Alg.* **49** (1987) p.283–335.
[W] E. Witten, 'Quantum Field Theory and the Jones Polynomial', *Commun. Math. Phys.* **121** (1989) p.351–389.
[Y] D. Yetter, 'State-sum invariants of 3-manifolds associated to Artinian semisimple tortile categories', *Topology Appl.* **58** (1994) p.47–80.
[Y2] D. Yetter, *In preparation.*
[Z1] A.B. Zamolodchikov, 'Tetrahedra equations and integrable systems in three-dimensional space', *Sov. JETP* **52** (1980) p.325–336.
[Z2] A.B. Zamolodchikov, 'Tetrahedron equations and the relativistic S-matrix of straight strings in $2+1$-dimensions', *Commun. Math. Phys.* **79** (1981) p.489–505.

E-mail: lawrence@math.lsa.umich.edu

© 1994 R.J. Lawrence