

NOTES ON ATIYAH'S TQFT'S

J.P. MAY

As an example of “categorification”, I presented Atiyah’s axioms [1] for a topological quantum field theory (TQFT) to undergraduates in the University of Chicago’s summer 2002 REU. These notes are written for the use of the participants and to test my own understanding. The aim is just to make very categorically precise what the basic definitions are. Also, there are some details in Atiyah’s original definition that are subsumed in the more recent and by now standard categorical definition, and it seemed worthwhile to present a comparison. I have in mind hitting TQFT’s with infinite loop space machines, which may perhaps provide a little unusual motivation. However, these notes are written to be readable by the students attending the REU, although standard categorical background that was explained in the talks is not repeated here. Needless to say, we have very good students at Chicago!!!

1. SYMMETRIC MONOIDAL CATEGORIES

The notion of a symmetric monoidal category \mathcal{C} is standard: \mathcal{C} must have a unit object S and a product $\otimes_{\mathcal{C}} \times \mathcal{C} \rightarrow \mathcal{C}$ that is unital, associative, and commutative up to coherent natural isomorphism. When such a category is essentially small (has a set of isomorphism classes of objects), it can be replaced by an equivalent permutative category, which is a symmetric monoidal category that is strictly associative and unital. One cannot make the commutativity isomorphism, or “braiding”, an isomorphism, and we use the letter γ to denote it. See [?] for symmetric monoidal categories and, for example, [?] for permutative categories.

For symmetric monoidal categories \mathcal{C} and \mathcal{D} and a functor $F : \mathcal{C} \rightarrow \mathcal{D}$, there are three obvious choices of what it means for F to be “symmetric monoidal”. In all of them, we require a map $\lambda : R_{\mathcal{D}} \rightarrow F(R_{\mathcal{C}})$ relating the unit objects and a natural transformation

$$\phi : F(X) \otimes_{\mathcal{D}} F(Y) \rightarrow F(X \otimes_{\mathcal{C}} Y)$$

relating the products, and in all of them we require all coherence diagrams relating the associativity, unit, and commutativity isomorphisms of \mathcal{C} and \mathcal{D} to commute. We would refer to a “monoidal functor”, if we only had such coherence for the associativity and unit isomorphisms, not for the commutativity isomorphisms. We say that F is *strict*, *strong*, or *lax* symmetric monoidal if λ and ϕ are both *identity maps*, both *isomorphisms*, or both just morphisms. In practice, while it is often true that λ is the identity, it is rarely true that ϕ is the identity, and it is often the case that ϕ is not even an isomorphism. In our definition of a TQFT, we understand strong symmetric monoidal, but with identities on unit objects. Replacement of symmetric monoidal categories by equivalent permutative categories is functorial: it transforms strong symmetric monoidal functors between symmetric monoidal categories to strict symmetric monoidal functors between permutative categories.

2. SYMMETRIC MONOIDAL CATEGORIES WITH DUALITY

We say that a symmetric monoidal category \mathcal{C} has *duality* if every object is “dualizable”. This means that there is an object DX and maps $\eta : S \rightarrow X \otimes DX$ and $\varepsilon : DX \otimes X \rightarrow S$ such that the following composites are identity maps.

$$X \cong S \otimes X \xrightarrow{\eta \otimes \text{id}} X \otimes DX \otimes X \xrightarrow{\text{id} \otimes \varepsilon} X \otimes S \cong X$$

and

$$DX \cong DX \otimes S \xrightarrow{\text{id} \otimes \eta} DX \otimes X \otimes DX \xrightarrow{\varepsilon \otimes \text{id}} S \otimes DX \cong DX$$

This notion is often discussed in the context of subcategories of closed categories, which have an internal Hom functor. In such a category, DX is isomorphic to $\text{Hom}(X, S)$ when X is dualizable, but not all objects X need be dualizable; see [?] for a recent discussion. The properties proven in that context carry over directly to categories that are not equipped with an internal Hom functor. We record some of the relevant properties.

Dual map

Adjunction

Uniqueness of X^*

Symmetric monoidal maps preserves duality, on both objects and morphisms.

Euler characteristics and traces

3. HILBERT SPACES AND *-CATEGORIES

A $*$ -category is a category \mathcal{C} with a contravariant involution $*$: $\mathcal{C} \rightarrow \mathcal{C}$ that is the identity on objects. That is, we must have $X^* = X$, $(gf)^* = g^*f^*$, and $f^{**} = f$. In many sources, $*$ is written for both duals and an involution. This can be a source of real confusion, especially when the duality and the involution are closely related. This is very much the case in the context of TQFT’s. The natural target for a TQFT is a symmetric monoidal $*$ -category with duality.

It is usual to take the target category to be the category Hilb of finite dimensional Hilbert spaces, but we can equally well start from any field with an involution. Write $x \mapsto \bar{x}$ for the involution on elements and write \bar{V} for V with the conjugate action of Λ , $(x, v) \mapsto \bar{x}v$. Write \bar{v} for v regarded as an element of \bar{V} . Write $\langle v, v' \rangle$ for hermitian forms. For present purposes a Hilbert space is a finite dimensional complex vector space equipped with a nondegenerate hermitian form. The morphisms of \mathcal{C} are all the linear transformations $L : V \rightarrow W$, and the adjoint $L^* : W \rightarrow V$ of L is characterized by $\langle Lv, w \rangle = \langle v, L^*w \rangle$ for $v \in V$ and $w \in W$. The contravariant functor $*$ is an involution that gives Hilb a structure of $*$ -category.

The tensor product of Hilbert spaces gives Hilb a structure of symmetric monoidal category. Its unit is \mathbb{C} . It has duals $DX = \text{Hom}(X, \mathbb{C})$. The map $\varepsilon : DX \otimes X \rightarrow \mathbb{C}$ is evaluation. We can define $\eta : \mathbb{C} \rightarrow V \otimes DV$ by $\eta(1) = \sum v_i \otimes \alpha_i$, where $\{v_i\}$ is a basis for V with dual basis $\{\alpha_i\}$. It is a pleasant exercise to verify that this does specify a duality and that the trace of a linear transformation $L : V \rightarrow V$ defined as in the previous section is precisely the trace of the matrix of L with respect to the basis $\{v_i\}$. In particular, $\chi(V) = \dim(V)$.

We say that L is an isometry if $\langle Lv, Lv' \rangle = \langle v, v' \rangle$ for $v, v' \in V$. We have an isomorphism $\zeta : \bar{V} \rightarrow DV$ specified by $\zeta(\bar{v})(v) = \langle v, w \rangle$ for $v, w \in V$. It is contravariantly functorial with respect to isometric isomorphisms $f : V \rightarrow W$, in the sense that $Df = \zeta \bar{f}^{-1} \zeta^{-1}$. This relates duality to the involution $*$.

4. THE DEFINITION OF A TQFT

Throughout these notes, manifolds, with or without boundary, are understood to be smooth, compact, and oriented. Diffeomorphisms are understood to be orientation-preserving. Fix an integer d . We define the symmetric monoidal category with duality and $\text{Cob}(d+1)$. With this convention, the objects of $\text{Cob}(d+1)$ are the manifolds of dimension d , without boundary. We write Σ^* for Σ with the opposite orientation. We have two kinds of morphisms between objects. First, we have the “ordinary morphisms”, namely the diffeomorphisms $f : \Sigma_0 \rightarrow \Sigma_1$, with the evident composition and identity maps.

Second, we have the “cobordism morphisms” $[M] : \Sigma_0 \rightarrow \Sigma_1$. These are the equivalence classes of (smooth, compact, oriented) manifolds M of dimension $d+1$, with boundary $\partial M = \Sigma_0^{\text{op}} \amalg \Sigma_1$. Two such “cobordisms” M and M' are equivalent if there is a diffeomorphism $M \rightarrow M'$ that restricts to the identity map on their common boundary. We allow Σ_0, Σ_1 , or both, to be empty, with $\emptyset^{\text{op}} = \emptyset$, and we allow M to be empty. Composition is by gluing. That is, if $\partial N = \Sigma_1^{\text{op}} \amalg \Sigma_2$, then the composite $[N] \circ [M] : \Sigma_0 \rightarrow \Sigma_2$ is the equivalence class of the manifold $M \cup_{\Sigma_1} N$ with boundary $\Sigma_0^{\text{op}} \amalg \Sigma_2$ that is obtained by gluing M and N together along Σ_1 . We have $\partial(\Sigma \times I) = \Sigma^* \amalg \Sigma$, and it is immediate from the smooth boundary collar theorem that $(\Sigma_1 \times I) \cup_{\Sigma_1} M$ and $M \cup_{\Sigma_0} (\Sigma_0 \times I)$ are diffeomorphic to M , relative boundary, and therefore represent the same cobordism morphism. Thus the $\Sigma \times I : \Sigma \rightarrow \Sigma$ give the identity cobordism morphisms. Composition is associative since we have evident diffeomorphisms between the relevant manifold-level composites of gluing operations.

We also have “2-morphisms” $M \rightarrow M'$ between cobordism morphisms $M : \Sigma_0 \rightarrow \Sigma_1$ and $M' : \Sigma'_0 \rightarrow \Sigma'_1$, namely equivalence classes of diffeomorphisms $g : M \rightarrow M'$, where g and g' are equivalent if they restrict to the same diffeomorphisms $f_0^{\text{op}} : \Sigma_0^{\text{op}} \rightarrow \Sigma_0'^{\text{op}}$ and $f_1 : \Sigma_1 \rightarrow \Sigma_1'$.

Finally, we have a symmetric monoidal structure on $\text{Cob}(d+1)$ given by disjoint union of d -manifolds and of $(d+1)$ -manifolds with boundary. The empty d -manifold gives the unit object for $\text{Cob}(d+1)$ regarded as a category with its ordinary morphisms, and the empty $(d+1)$ -manifold gives the unit for the symmetric monoidal category under disjoint union whose objects are the cobordism morphisms and whose morphisms are the 2-morphisms.

Now fix a symmetric monoidal category \mathcal{C} with product \otimes and unit object R . We are thinking of the category of finitely generated modules over a commutative ring R under the tensor product. We think of \mathcal{C} as a symmetric monoidal 2-fold 2-category in which the “ordinary morphisms” and the “cobordism morphisms” coincide and in which the only 2-morphisms are identity morphisms.

Definition 4.1. A TQFT Z in dimension d over \mathcal{C} is a map of symmetric monoidal categories $Z : \text{Cod}(d) \rightarrow \mathcal{C}$.

We spell out precisely what this means, leaving the phrase “without duality” for discussion in the next section. For each d -manifold Σ , we have an object $Z(\Sigma)$ of \mathcal{C} . For each diffeomorphism $f : \Sigma_0 \rightarrow \Sigma_1$, we have an isomorphism $Z(f) : Z(\Sigma_0) \rightarrow Z(\Sigma_1)$. For each $(d+1)$ -manifold M with $\partial M = \Sigma_0^{\text{op}} \amalg \Sigma_1$, we have a morphism $Z(M) : Z(\Sigma_0) \rightarrow Z(\Sigma_1)$, and equivalent cobordisms give the same morphism $Z(M)$. These assignments give us two functors with values in \mathcal{C} . In

particular, since the $\Sigma \times I$ are identity morphisms,

$$(4.2) \quad Z(\Sigma \times I) = id : Z(\Sigma) \longrightarrow Z(\Sigma).$$

If $\partial M' = \Sigma'^{\text{op}} \amalg \Sigma'_1$ and if diffeomorphisms $f_0 : \Sigma_0 \longrightarrow \Sigma'_0$ and $f_1 : \Sigma_1 \longrightarrow \Sigma'_1$ extend to a diffeomorphism $g : M \longrightarrow M'$, then the following diagram commutes.

$$(4.3) \quad \begin{array}{ccc} Z(\Sigma_0) & \xrightarrow{Z(M)} & Z(\Sigma_1) \\ Z(f_0) \downarrow & & \downarrow Z(f_1) \\ Z(\Sigma'_0) & \xrightarrow{Z(M')} & Z(\Sigma'_1) \end{array}$$

This is forced by the requirement that the only 2-morphisms of \mathcal{C} are identity maps, but those unfamiliar with the relevant categorical language should take this as part of what we mean by definition of a “map” in Definition 4.1.

Note in particular that properties (4.2) and (4.3) give an expected homotopy invariance property of a TQFT. To see this, take $M = M' = \Sigma \times I$. Then g is a homotopy $f_0 \simeq f_1$, and the cited properties give that $Z(f_0) = Z(f_1)$.

The question of the meaning of a “map” becomes more serious when we consider the symmetric monoidal structure. For symmetric monoidal categories \mathcal{C} and \mathcal{D} and a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$, there are three obvious choices of what it means for F to be “symmetric monoidal”. In all of them, we require a map $\lambda : R_{\mathcal{D}} \longrightarrow F(R_{\mathcal{C}})$ relating the unit objects and a natural transformation

$$\phi : F(X) \otimes_{\mathcal{D}} F(Y) \longrightarrow F(X \otimes_{\mathcal{C}} Y)$$

relating the products, and in all of them we require all coherence diagrams relating the associativity, unit, and commutativity isomorphisms of \mathcal{C} and \mathcal{D} to commute. We would refer to a “monoidal functor”, if we only had such coherence for the associativity and unit isomorphisms, not for the commutativity isomorphisms. We say that F is *strict*, *strong*, or *lax* monoidal or symmetric monoidal if λ and ϕ are both *identity maps*, both *isomorphisms*, or both just morphisms. In practice, while it is often true that λ is the identity, it is rarely true that ϕ is the identity, and it is often the case that ϕ is not even an isomorphism. In our definition of a TQFT, we understand strong symmetric monoidal, but with identities on unit objects.

We spell this out explicitly. We require that $Z(\emptyset) = R$, where \emptyset is the empty d -manifold, and we require that $Z(\emptyset) = id : R \longrightarrow R$, where \emptyset is the empty $(d+1)$ -manifold regarded as a cobordism morphism from the empty d -manifold to itself. We require a binatural isomorphism

$$(4.4) \quad \phi : Z(\Sigma) \otimes Z(\Sigma') \xrightarrow{\cong} Z(\Sigma \amalg \Sigma')$$

that makes all coherence diagrams commute. Here “binatural” means that it is natural with respect to both diffeomorphisms and cobordisms. In particular, for $M : \Sigma_0 \longrightarrow \Sigma_1$ and $M' : \Sigma'_0 \longrightarrow \Sigma'_1$, the following diagram commutes

$$(4.5) \quad \begin{array}{ccc} Z(\Sigma_0) \otimes Z(\Sigma'_0) & \xrightarrow[\cong]{\phi} & Z(\Sigma_0 \amalg \Sigma'_0) \\ Z(M) \otimes Z(M') \downarrow & & \downarrow Z(M \amalg M') \\ Z(\Sigma_1) \otimes Z(\Sigma'_1) & \xrightarrow[\cong]{\phi} & Z(\Sigma_1 \amalg \Sigma'_1) \end{array}$$

5. ATIYAH'S DEFINITION OF A TQFT

We insist in this section that \mathcal{C} be the category of modules over a commutative ring R . We could just as well work with a general closed symmetric monoidal category, with internal hom functor Hom , but we desist. We write V^* for the dual $\text{Hom}(V, R)$ of an R -module V and we write ε , or $\langle \nu, v \rangle$ on elements, for the evaluation pairing $V^* \otimes V \rightarrow R$. For R -modules U, V , and W , we later write $\langle u \otimes v, \nu \otimes w \rangle$ for the pairing $\text{id} \otimes \varepsilon \circ \tau \otimes \text{id} : U \otimes V \otimes V^* \otimes W \rightarrow U \otimes W$, where τ interchanges V and V^* .

If we think of a manifold M with boundary ∂M as a cobordism $\emptyset \rightarrow \partial M$, then $Z(M)$ is a morphism $R = Z(\emptyset) \rightarrow Z(\partial M)$ in \mathcal{C} . It is determined by its value on 1, which is an element $z(M)$ of $Z(\partial M)$. In [1], both the element $z(M)$ and the morphism $Z(M) : Z(\Sigma_0) \rightarrow Z(\Sigma_1)$ when $\partial M = \Sigma_0^{\text{op}} \amalg \Sigma_1$ are denoted $Z(M)$. I find this a little confusing, and it would obscure our comparison of definitions, hence I have changed notation to $z(M)$. On the other hand, we can view M^{op} as a cobordism $\partial M \rightarrow \emptyset$. Then $Z(M^{\text{op}})$ is a morphism $Z(\partial M) \rightarrow Z(\emptyset) = R$. It may be viewed as an element $z^{\text{op}}(M^{\text{op}})$ of $Z(\partial M)^*$.

When M is closed, $\partial M = \emptyset$, $z(M) = z^{\text{op}}(M)$ is an element of $R = Z(\emptyset)$ and is thus a numerical invariant of M . In general, the elements $z(M)$ and $z^{\text{op}}(M)$ are invariants of the pair $(M, \partial M)$. We regard the axiomatization of a TQFT as a *categorification* of the notion of such relative invariants of manifolds with boundary. An essential point is to understand how to compute $z(M)$ when M is cut into two parts M_1 and M_2 along an embedded d -manifold Σ , so that $M = M_1 \cup_{\Sigma} M_2$ with $\partial M_1 = \Sigma$ and $\partial M_2 = \Sigma^{\text{op}}$. Regarding M_1 as a cobordism $\emptyset \rightarrow \Sigma$ and M_2 as a cobordism $\Sigma \rightarrow \emptyset$, we have $Z(M) = Z(M_2) \circ Z(M_1)$ and thus, on elements, $z(M) = \langle z^{\text{op}}(M_2), z(M_1) \rangle$.

Clearly, it is natural to insist on a relationship between the values of Z on manifolds and their opposites. Atiyah's notion of a TQFT is a restriction of ours that accomplishes this. He requires R to be a field, denoted Λ , and he requires the $Z(\Sigma)$ to be finite dimensional vector spaces. Actually, his definition works just as well for general commutative rings R , provided that the $Z(\Sigma)$ are required to be finitely generated projective R -modules. Thus we now take \mathcal{C} to be the category of finitely generated projective R -modules. This allows the following definition.

Definition 5.1. A TQFT without duality is *involutory* if there is an isomorphism

$$(5.2) \quad \delta : Z(\Sigma^{\text{op}}) \xrightarrow{\cong} Z(\Sigma)^*$$

that is natural with respect to diffeomorphisms in the sense that the following diagram commutes for $f : \Sigma_0 \rightarrow \Sigma_1$.

$$(5.3) \quad \begin{array}{ccc} Z(\Sigma_0^{\text{op}}) & \xrightarrow{Z(f^{\text{op}})} & Z(\Sigma_1^{\text{op}}) \\ \delta \downarrow & & \downarrow \delta \\ Z(\Sigma_0)^* & \xrightarrow{Z(f^{-1})} & Z(\Sigma_1)^* \end{array}$$

We can now reconcile our definition with Atiyah's original definition.

Proposition 5.4. *A TQFT in the sense of Atiyah is the same structure as an involutory TQFT without duality.*

Proof. In Atiyah's paper [1], ϕ in (4.4) and δ in (5.2) are written as equalities rather than isomorphisms, and the naturality diagram (5.3) is left implicit. However, in pedantically rigorous terms, it is clear that this is what Atiyah had in mind. Even after making these points precise, his definition looks just a little different than ours, and we indicate how to compare them. As already noted, Atiyah requires an element $Z(M)$ in $Z(\partial M)$ for each $(d+1)$ -manifold M , and we are rewriting his element as $z(M)$ to avoid confusion with our linear transformations $Z(M)$. Our $Z(M)$ determines his $z(M)$ as the image of 1 under $Z(M) : R = Z(\emptyset) \rightarrow \partial M$. Conversely, when $\partial M = \Sigma_0^{op} \amalg \Sigma_1$, his $z(M)$ determines our $Z(M)$ as the image of $z(M)$ under the composite isomorphism

$$(5.5) \quad Z(\partial M) = Z(\Sigma_0^{op} \amalg \Sigma_1) \cong Z(\Sigma_0)^* \otimes Z(\Sigma_1) \cong \text{Hom}(Z(\Sigma_0), Z(\Sigma_1)).$$

Observe that the last displayed isomorphism requires that the $Z(\Sigma)$ be finite dimensional, or finitely generated projective when working over a general commutative ring. Atiyah's key axiom, [1, 3b], can be formulated as follows. Take manifolds M_1 and M_2 with $\partial M_1 = \Sigma_0 \amalg \Sigma_1$ and $\partial M_2 = \Sigma_1^{op} \amalg \Sigma_2$ and let $M = M_1 \cup_{\Sigma_1} M_2$. Then

$$(5.6) \quad z(M) = \langle z(M_1), z(M_2) \rangle,$$

where the isomorphisms ϕ and δ of (4.4) and (5.2) are used to identify $Z(\partial M)$, $Z(\partial M_1)$ and $Z(\partial M_2)$ with $Z(\Sigma_0) \otimes Z(\Sigma_2)$, $Z(\Sigma_0) \otimes Z(\Sigma_1)$ and $Z(\Sigma_1)^* \otimes Z(\Sigma_2)$. Using (5.5), it is easily verified that this axiom is equivalent to the transitivity of our $Z(M)$. Together with (4.2), which is Atiyah's axiom (4c), this verifies our requirement that Z be functorial with respect to cobordism morphisms. The rest is an immediate comparison of definitions. \square

6. THE DEFINITION OF A TQFT WITH DUALITY

As Atiyah remarks, his involutory axiom is conceptually lacking in that it ignores the cobordism morphisms. In other words, although it allows a definition entirely in terms of the $Z(\Sigma)$ and the $z(M)$, it fails to relate $z(M)$ to $z(M^{op})$. This correlates with our introduction of $z^{op}(M^{op}) \in Z(\partial M)^*$. There are important examples where these invariants are genuinely different, so the extra generality is desirable. This also makes it plausible that a definition such as ours that does away entirely with the involutory axiom will find applications. However, duality is of fundamental importance in many examples, and the more structured version of a TQFT *with duality* appropriately sharpens the involutory axiom. The obvious point is that, since the maps $Z(M)$ need not be isomorphisms, we need additional structure to obtain the appropriate analogue of (5.3). We could describe such structure axiomatically on a general closed symmetric monoidal category V , but we follow Atiyah in using standard field theory instead.

Henceforward, let Λ be a field with an involution, denoted $x \mapsto \bar{x}$ on elements. We write \bar{V} for V with the conjugate action of Λ , $(x, v) \mapsto \bar{x}v$, and we write \bar{v} for v regarded as an element of \bar{V} . We write $\langle v, v' \rangle$ for hermitian forms, and we take \mathcal{C} to be the category of finite dimensional vector spaces V equipped with nondegenerate hermitian forms. The morphisms of \mathcal{C} are the linear transformations $L : V \rightarrow W$, and the adjoint $L^* : W \rightarrow V$ of L is characterized by $\langle Lv, w \rangle = \langle v, \text{Ad}L w \rangle$ for $v \in V$ and $w \in W$. For $K : U \rightarrow V$, $\text{Ad}(L \circ K) = \text{Ad}K \circ \text{Ad}L$. We say that L is an isometry if $\langle Lv, Lv' \rangle = \langle v, v' \rangle$ for $v, v' \in V$. As usual, the essential point is that we have a natural isomorphism $\zeta : \bar{V} \rightarrow V^*$ specified by $\zeta(\bar{v})(v) = \langle v, v \rangle$ for $v, w \in V$, and this allows us to interpret V^* as a covariant functor of V .

Definition 6.1. Let Z be an involutory TQFT in \mathcal{C} such that $Z(f)$ is an isometry for each diffeomorphism f . We say that Z is a TQFT *with duality* if

$$(6.2) \quad Z(M^{op}) = \text{Ad}(Z(M)) : Z(\Sigma_1) \longrightarrow Z(\Sigma_0)$$

for all cobordisms $M : \Sigma_0 \longrightarrow \Sigma_1$.

When $\Sigma_0 = \emptyset$, this says that $z^{op}(M^{op}) = z$

Applying the isomorphisms $Z(\Sigma^{op}) \cong Z(\Sigma)^* \cong \overline{Z(\Sigma)}$ to $\Sigma = \partial M$, this can be reinterpreted elementwise as $z(M^{op}) = \overline{z(M)}$ in $\overline{Z(\partial M)}$.

The double $M \cup_{\partial M} M$ of a manifold M is closed, and $z(M \cup_{\partial M} M) = z(M)\overline{z(M)}$, which is $|z(M)|^2$.

REFERENCES

- [1] M. Atiyah. Topological quantum field theories. Atiyah, Michael(4-OX) Topological quantum field theories. Inst. Hautes tudes Sci. Publ. Math. No. 68, (1988), 175–186 (1989).