A Brief Overview of Topological Quantum Field Theory

Andrew Poelstra*

March 2013

1 Introduction

Beyond the horizon of the place we lived when we were young,
in a world of magnets and miracles,
our thoughts strayed constantly and without boundary.
—Pink Floyd

Topological quantum field theories are elegant, general, expansive mathematical theories which hold great promise as tools for setting quantum field theory on solid ground. They were originally created as an abstraction of the path integral formalism [1, 23] which sought to avoid the infinities plaguing Feynmanology. Michael Atiyah [1] suggested an explicit axiomatization for a TQFT, which has been realized in low dimensions (c.f. [6]).

The essential motivating idea behind topological field theories is that the modern physical theories are defined in terms of invariance under certain group actions (e.g. gauge groups in particle physics, diffeomorphism groups in general relativity, unitary operator groups in quantum mechanics). Related to this is the idea that a system can be characterized by some number, an invariant under the group — for example, a four-vector in relativity or a vacuum expectation value in a field theory — or a "relative invariant" as seen in symmetry-breaking theories such as the Higgs mechanism. In topological field theory, we are concerned with *topological invariants*, which are objects computed from a topological space (usually a smooth manifold) without respect to any metric [24]. Concretely, topological invariance means invariance under the diffeomorphism group of the manifold.

Mathematically, enormous strides were made in geometry in the 19th and 20th centuries. Important milestones were René Thom's theory of cobordism [22], de Rham cohomology (and cohomology in general), and knot theory. Through theories such as the Chern-Weil theory linking differential geometry and algebraic topology, abstract formalisms found powerful geometric applications. These were applied to physics starting in the 70's (c.f. the original Chern-Simons paper [3]), and largely through the work of Witten and Atiyah, flourished in the 80's and 90's.

More recently, they have been taken up by Louis Crane [7] as a strategy for unifying gravitation and quantum physics. (Diffeomorphism invariance, the signature of general relativity, is translated beautifully by Atiyah's axioms into cobordism equivalence, which is a more completely understood theory and also much more suitable to quantization.)

The purpose of this document is twofold: first, to provide an accessible overview and introduction to topological quantum field theories to a beginning graduate student of mathematics or physics; second, it is a term paper for a course in traditional field theory. As this puts serious time constraints on the writing of this paper (and I am already pushing my luck by running so far afield!), it is necessarily and unfortunately brief.

Required prerequisites are a familiarity with differential geometry, category theory and traditional quantum field theory. For the most part, we cover a lot of ground on a high level and leave detailed construction to the references, so no intimate mathematical knowledge is required.

While the presentation here is new, no claim to originality is made of the content of this paper.

2 Mathematical Preliminaries

2.1 Category Theory

The broader the brush, the bigger the fool wielding it.
—Folk wisdom[1]

We will briefly cover the essential features of category theory needed for this paper. For a thorough exposition, consult MacLane's classic text [15]; for a more readable introduction, the the text [2] by Barr and Wells is excellent.

We will use category theory mainly as it applies to algebraic topology; a good introductory overview is [16].

Definition 1. A *category* $\mathscr{C}$ consists of a collection of objects[2] along with a collection of *morphisms* or *maps*, which are arrows between the objects. Given objects A, B in $\mathscr{C}$, a morphism f between the two is denoted $f : A \to B$, where A is the *domain* of f and B is the *codomain*.

Let A, B, C, D be objects in $\mathscr{C}$. Given two arrows $f : A \to B$ and $g : B \to C$, we require there exist a *composition* $f \circ g : A \to C$. Further, to each object A, there exists an *identity* Id_A which satisfies $\mathrm{Id}_A \circ f = f$ for all $f : A \to D$ and $g \circ \mathrm{Id}_A = g$ for all $g : D \to A$. (Here the identity depends on A but not D.)

Definition 2. A *functor* is a map between categories $\mathscr{C}$ and $\mathscr{D}$ which carries objects to objects, arrows to arrows, and preserves composition (i.e., $F(f \circ g) = F(f) \circ F(g)$).

Definition 3. A *monoidal* or *tensor category* is a category $\mathscr{C}$ with a map $\otimes : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ which is a functor in both coordinates.

2.2 Differential Geometry and Cobordism

We assume familiarity with the theory of manifolds. For a mathematical introduction, see [13]; for physical interpretation, the first chapter of [10].

Definition 4. Let M, N be d-dimensional oriented manifolds. If there exists a $(d+1)$-dimensional manifold B such that $\partial B = M \cup N^*$, we say B is a *cobordism* between M and N and that M and N

[1] Thanks to `TallTim` on `#bitcoin` for this quote.

[2] Strictly speaking, a *class* of objects. A class is essentially a generalized set, used to allow constructions like "the class of all sets" without raising the spectre of Russell's paradox. Every set is a class. A category whose class of objects is a set is called a *small category*, though we will not need this notion.

are *cobordant*.

Throughout this paper, N^* denotes N with orientation reversed.

It is not hard to see that cobordism is an equivalence relation and generalizes both homeomorphism and diffeomorphism. By chaining cobordisms we obtain a groupoid (identityless group).

The theory of cobordism took off with René Thom's classic paper [22], in which he gives a complete construction of cobordism groups. The analogous problems for homeomorphism and diffeomorphism are unsolved (in fact, it can be shown that these groups cannot be computed in dimension ≥ 4). The generalization to cobordisms gives a clean, beautiful theory well-developed enough to build on.

A general model for topological quantum field theories, first described by Michael Atiyah in [1], leans heavily on cobordism. The idea is further exploited, to connect spaces whose dimension differs by 1, by Louis Crane in his proposal [7] to bootstrap higher-dimensional TQFT's from lower-dimensional ones.

2.3 Lie Groups

The reader will need a familiarity with Lie groups. We will briefly cover the essential results, focusing in particular on compact Lie groups because of their simplified theory. The reader interested in pursuing this further should read, for example, [12].

Definition 5. A *compact Lie group* G is a compact differentiable manifold with a group structure such that the mapping $(g,h) \mapsto gh^{-1}$ of $G \times G$ into G is diffeomorphic.

It can be shown [20] that any compact Lie group can be identified by some subset of $\mathrm{GL}(n,\mathscr{C})$, the group of $n \times n$ complex matrices; the tangent space to the identity can then also be described as $n \times n$ matrices, and therefore has a group structure of its own.

We can define a skew-symmetric bilinear map on this tangent space by $[X,Y] = XY - YX$, where the multiplication used is ordinary matrix multiplication. An alternate approach suitable for general Lie groups is given in [12].

Definition 6. The *Lie algebra* of a compact Lie group is defined as the tangent space of the identity element, along with the above bilinear map.

Definition 7. Given some vector space V, a *representation* of a Lie group G is a homomorphic map from G into $\mathrm{Aut}(V) = \mathrm{GL}(V)$. A *representation* of a Lie algebra $\mathfrak{g}$ is a map from the Lie algebra into $\mathfrak{gl}(V)$ which preserves the bracket, where $\mathfrak{gl}(V)$ is the space of automorphisms on V (with bracket $[X,Y] = XY - YX$).

Notice that $\mathrm{GL}(n)$ is itself a Lie group; we refer to its Lie algebra, the set of all $n \times n$ matrices, as $\mathfrak{gl}(n)$. When we consider an arbitrary compact Lie group as being embedded in $\mathrm{GL}(n)$ (and thus its algebra embedded in $\mathfrak{gl}(n)$), this is the *standard representation*.

When speaking of the standard representation, we will forget the embedding map and just treat

G as a subgroup of $\mathrm{GL}(V)$.

The *Adjoint representation* of a Lie group on its Lie algebra is given by $G \to \mathfrak{g}$; $g \mapsto d(ghg^{-1})$. That is, each group element is mapped to the differential of its inner automorphism. By chasing definitions we see that the action of such a differential on a tangent vector X is $\mathrm{Ad}(g) : X \mapsto gXg^{-1}$.

The *adjoint representation* of a Lie algebra on itself is given by the differential of the Adjoint representation. For a compact Lie subgroup of $\mathrm{GL}(n,\mathbb{C})$, a tangent vector X acts under this representation by $\mathrm{ad}(X) : Y \mapsto \frac{d}{dt}(\mathrm{Ad}(\exp(tX))Y)\big|_{t=0}$.

Consider some element $X \in \mathfrak{g}$. It is well-known [13] that there exists a unique geodesic $\gamma : \mathbb{R} \to G$ with $\gamma_X(0)$ equal the identity and $\gamma_X'(0) = X$. The mapping $\mathfrak{g} \to G$ given by $X \to \gamma(1)$ is the *exponential map* exp. This is locally diffeomorphic and gives a canonical way to relate a Lie group to its algebra.

3 The Atiyah-Segal Axioms

One geometry cannot be more true than another; it can only be more convenient.
Geometry is not true, it is advantageous.
—Robert Pirsig

In this section, we state (in more mathematical terms) the axioms given by Atiyah in [1], which in turn were based on axioms for conformal theory given by Graeme Segal[3].

3.1 Notation

Fix some base field Λ, say, $\mathbb{R}$ or $\mathbb{C}$. (There exist generalizations in which Λ is a ring, but we will not discuss these.) We denote by $\mathscr{V}_\Lambda$ the category of all vector spaces on Λ, with linear operators as morphisms. We denote by $\mathscr{V}_\Lambda^*$ the category of vector spaces on Λ with basepoint. (Here "basepoint" does *not* refer to a zero vector, just some privileged vector in each space. Morphisms must preserve basepoint.)

These categories are tensor categories with duality under the ordinary vector tensor product $\otimes$ and vector duality.

We denote by $\mathscr{M}^d$ the category whose objects are d-dimensional manifolds and whose morphisms are orientation-preserving diffeomorphisms. This is a tensor category with duality under disjoint union and orientation reversal. Then to each $M \in \mathscr{M}^d$, we write $\mathscr{B}(M)$ for M's *bordisms*: the set of manifolds in $\mathscr{M}^{d+1}$ which have M as their boundary.

For example, $\mathscr{B}(\varnothing)$ is the set of closed $(d+1)$-dimensional manifolds.

[3] I could not track these down. Atiyah cites a paper listed as *to appear*, and other authors simply assume familiarity.

3.2 Mathematical Formalism

We introduce functors $Z^d : \mathscr{M}^d \to \mathscr{V}_\Lambda$ and $Z^{d+1} : \mathscr{M}^{d+1} \to \mathscr{V}^*_\Lambda$; these functors (along with Λ and the manifolds) are called a *topological field theory (of dimension d)* provided they satisfy the *Atiyah-Segal axioms*, which are

1. Whenever $M \in \mathscr{M}^d$, $B \in \mathscr{B}(M)$, we have $Z^d(M) = Z^{d+1}(B)$ as spaces[4]. **Therefore, from here on, by $Z^{d+1}(B)$ we mean its basepoint.** We can then write $Z^{d+1}(B) \in Z^d(M)$, and if $f : M \to M'$ is a morphism in $\mathscr{M}^d$ which extends to a morphism $\tilde{f} : B \to B'$ (where B, B' are bordisms of M, M'), then $Z^{d+1}(\tilde{f})$ carries $Z^{d+1}(B)$ to $Z^{d+1}(B')$.

2. Z^d and Z^{d+1} are multiplicative with respect to tensor product:

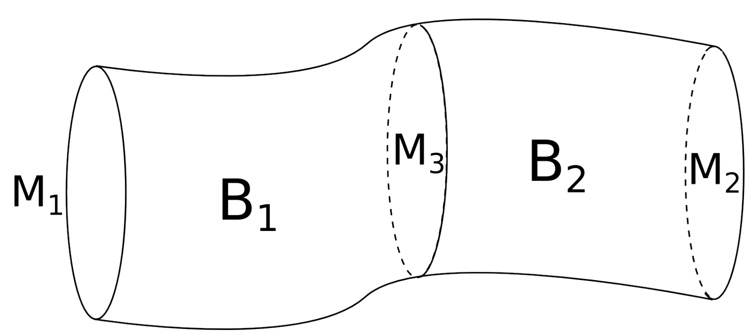
   (a) With $\dot\cup$ denoting disjoint union, $Z^d(M_1 \dot\cup M_2) = Z^d(M_1) \otimes Z^d(M_2)$. Similarly, for the basepoints, $Z^{d+1}(B_1 \dot\cup B_2) = Z^{d+1}(B_1) \otimes Z^{d+1}(B_2)$. This has an important consequence: if $B \in \mathscr{M}^{d+1}$ and we decompose its boundary into two components M_1^* and M_2, then

   $$Z^{d+1}(B) \in Z^d(M_1^*) \otimes Z^d(M_2) = \mathrm{Hom}(Z^d(M_1), Z^d(M_2))$$

   Here $\mathrm{Hom}(X,Y)$ refers to the set of linear operators from X to Y. A proof of the second equality can be found in any text on representation theory, such as [12, 18].

   This says, explicitly, that any cobordism between manifolds M_1 and M_2 in $\mathscr{M}^d$ is carried to a linear transformation by the theory.

   (b) Further, for any tensor of the form $u \otimes u^*$, we contract to get an element of Λ. This is a transitivity axiom, in the following sense: when $B = B_1 \cup B_2$ is a cobordism between manifolds M_1 and M_2 obtained by gluing together B_1 and B_2 along a shared boundary (i.e., $\partial B_1 = M_1 \cup M_3^*$ and $\partial B_2 = M_3 \cup M_2$), then



   $$Z^{d+1}(B) = \langle Z^{d+1}(B_1), Z^{d+1}(B_2) \rangle$$

[4] It is sometimes said that Z^d and Z^{d+1} are the same object, a *bifunctor*. I think this is unnecessary and confusing.

with $\langle\cdot,\cdot\rangle$ the usual contraction map from $Z^d(M_1)\otimes Z^d(M_3)^*\otimes Z^d(M_3)\otimes Z^d(M_2)$ to $Z^d(M_1)\otimes Z^d(M_2)$.

3. Since $M\otimes\varnothing = M$ for $M\in\mathscr{M}^d$ or $M\in\mathscr{M}^{d+1}$, the previous axiom requires $Z^d(\varnothing)$ and $Z^{d+1}(\varnothing)$ to be idempotent. To exclude the trivial theory, we therefore impose

$$Z^d(\varnothing)=\Lambda \qquad\qquad Z^{d+1}(\varnothing)=1_\Lambda$$

Transitivity of cobordisms gives that $Z^{d+1}(M\times I)$, considered as a linear operator between $Z^d(M)$ and itself, is also an idempotent which acts as the identity on the span of $Z^{d+1}(\mathscr{B}(M))$, which is a subspace of $Z^d(M)$. We identify $Z^d(M)$ with this subspace, or in other words, impose

$$Z^{d+1}(M\times I)=\mathrm{Id}_{Z^d(M)}$$

From these axioms, we can derive some immediate and useful results:

1. If $M\in\mathscr{M}^d$, the M-shaped torus $M\times S^1$ can be considered equally well to have boundary $M\dot\cup M^*$ or $\varnothing$. Then by the multiplicative axiom we have

$$\Lambda = Z^d(\varnothing)=Z^d(M\dot\cup M^*)=Z^d(M)\otimes Z^d(M^*)$$

It follows that $Z^d(M^*)=Z^d(M)^*$, where $Z^d(M)^*$ is the vector dual space of $Z^d(M)$.

In general, $Z^{d+1}(B^*)$ does not necessarily equal $Z^{d+1}(B)^*$.

2. If we construct $M\times S^1$ by identifying the ends of the cylinder $M\times I$, we obtain the identity on $Z^d(M)$ contracted with itself; i.e., $Z^{d+1}(M\times S^1)=\mathrm{Tr}(\mathrm{Id}) = \dim(Z^d(M))$.

In fact, if we identify the ends of $M\times I$ by any diffeomorphism f, we obtain a torus-style manifold M_f, and $Z^{d+1}(M_f)=\mathrm{Tr}(Z^d(f))$.

3.3 Physical Interpretation

We take d to be the the spacial dimension of the universe; then $(d+1)$ dimensions are needed to include time. Our manifolds in $\mathscr{M}^d$ thus represent physical systems, while the cobordisms of $\mathscr{M}^{d+1}$ give a transport through “time”. The vector spaces $Z^d(M)$ will be Hilbert spaces with a Hamiltonian H; then the special cobordism $M\times I$ (M with “imaginary time”) will induce a time-evolution operator e^{-itH}.

If $H\equiv 0$, there are superficially no dynamics in time; in fact, we typically force $H\equiv 0$ to preclude the existence of some preferred time axis, which would certainly be incompatible with relativity. However, interesting behavior can still occur via cobordisms other than $M\times I$. Such cobordisms can describe physical processes (with some nonzero probability amplitude), giving rise to a dynamics in a topologically-invariant way.

The basepoint vectors $Z^{d+1}(B) \in Z^d(M)$ are states (spins in Chern-Simons, lengths in Crane's gravity model, etc.) in $Z^d(M)$ corresponding to B; if M is the empty manifold, then $Z^{d+1}(B) \in \Lambda$ is typically a vacuum expectation value.

The vector spaces themselves are state spaces corresponding to different observers. If these spaces are isomorphic, the linear maps induced by the cobordisms are simply change-of-basis matrices; however, this is not a necessary feature, and topological theories can describe models of reality in which variables have no global definition (c.f. [19]).

4 Examples

Reality leaves a lot to the imagination.
—John Lennon

We begin by describing zero-dimensional theories, loosely based on Atiyah's exposition in [1]. From there, we give a cursory overview of the Chern-Simons theory based on Wilson loops and associated invariants. Though we will not discuss them, many theories have been constructed in dimensions less than 4; however, there is no satisfactory theory for the case $d = 4$[5]. An idea for doing so is provided by Crane's "dimensional ladder" [7, 8] for bootstrapping high-dimensional theories from lower-dimensional ones, which is described in the next section.

4.1 The Zero-Dimensional Theory

In this case, our manifolds of $\mathscr{M}^d$ are single points, and our diffeomorphism group is a symmetry group. The impact of this is that we can describe the representation theory of compact Lie groups (with the symmetry actions then appearing as adjoint actions on the weight spaces of the Lie algebra, c.f. [12, Lectures 12,16]) in terms of a TQFT. We will give a brief overview of this representation theory.

Specifically, the functor Z^0 assigns a vector space V to each point; disjoint unions let us map n points into the tensor space $\underbrace{V \otimes \cdots \otimes V}_{n \text{ times}}$. (Since the symmetry group S^n is transitive, we must have n copies of the same vector space.)

The cobordisms are all diffeomorphisms and therefore also live in S^n. (In Atiyah's words, there is no interesting topology.)

For our vector space V, we can a Hilbert space by quantizing some symplectic (i.e., possessing a closed non-degenerate 2-form) manifold. Examples of such manifolds are the phase space of some classical or quantum Hamiltonian theory [5].

[5]Here "satisfactory" means "describes gravity".

Alternately, starting with a Lie group, there is a algebraic-geometric way to construct a symplectic manifold (from a line bundle) which preserves much of the group structure. From this, we will find that the Borel-Weil theorem gives a method of quantization — a physical interpretation of a purely mathematical result!

We outline this process here, starting by describing the requisite theory, then following [12, Lecture 23] (simplified to the compact case):

1. It is well known [20, Theorems 3.28, 2.15] that compact Lie groups are isomorphic to closed unitary subgroups of $\mathrm{GL}(n,\mathbb{C})$; then Cartan subgroups will be maximal torii[6]. Choose a Lie group G with algebra $\mathfrak{g}$. Let T be a maximal torus in G with subalgebra $\mathfrak{t}$. Write $\mathfrak{t}_{\mathbb{C}}$ for the complexification $\mathfrak{t}\otimes\mathbb{C}$, where the tensor product is taken over $\mathbb{R}$. All the Lie algebra structure is extended by $\mathbb{C}$-linearity from $\mathfrak{t}$ to $\mathfrak{t}_{\mathbb{C}}$.

2. For any representation V' of M, there is a finite set $\Delta(V')\subset\mathfrak{t}_{\mathbb{C}}^*$ (the set of complex linear functionals on $\mathfrak{t}_{\mathbb{C}}$) called *weights*, and V' can be written

$$V' = \bigoplus_{\lambda\in\Delta(V')} V_\alpha$$

   where each $v\in V_\lambda$ is acted on by each $X\in\mathfrak{t}_{\mathbb{C}}$ by $Xv=\lambda(X)v$. This decomposition is unique.

   Weights of the adjoint representation[7] are called *roots*, and corresponding weight spaces are called *root spaces*. Notice that zero is always a root, and has root space $\mathfrak{t}_{\mathbb{C}}$.

   We can choose a (non-unique) collection Φ^+ of *positive roots* characterized by: for every root λ, exactly one of $\pm\lambda$ is in Φ^+; if $\lambda,\gamma\in\Phi^+$ and $\lambda+\gamma$ is a root, then $\lambda+\gamma\in\Phi^+$.

   Letting $(\cdot,\cdot)$ be the Killing form[8] for $\mathfrak{g}$, we define a *dominant weight* Λ as one such that $(\Lambda,\gamma)\geq 0$ for every positive root γ.

3. Since $\mathfrak{t}_{\mathbb{C}}^*$ is defined by linearity, it is essentially two copies of $\mathfrak{t}$, and we can identify

$$\mathfrak{t}_{\mathbb{C}}^* \simeq \mathfrak{t}^* \simeq (i\mathfrak{t})^*$$

   With this in mind, a weight $\lambda\in(i\mathfrak{t})^*$ is called *analytically integral* if $\lambda(X)\in 2\pi i\mathbb{Z}$ for all $X\in\mathfrak{t}$ with $\exp(X)$ equal the identity. (Notice that we say the analytically integral weights live in $(i\mathfrak{t})^*$ even though they are defined by their behavior as members of $\mathfrak{t}^*$.)

[6] A *torus* is a Lie group isomorphic to some $(S^1)^k$, where $S^1=\mathbb{R}/\mathbb{Z}$. Notice that while we treat compact Lie groups as complex linear spaces, torii are real.

[7] More correctly, weights of the *extension* of the adjoint representation by $\mathbb{C}$-linearity to the complexification of $\mathfrak{g}$. This complexification is well-defined since the members of G are unitary matrices; thus $\mathfrak{g}\subseteq\mathfrak{u}(n)$; by writing matrices as sums of Hermitian and skew-Hermitian parts, we can obtain $\mathfrak{gl}(n,\mathbb{C})=\mathfrak{u}(n)\otimes i\mathfrak{u}(n)$ and therefore write $\mathfrak{g}_{\mathbb{C}}=\mathfrak{g}\otimes i\mathfrak{g}\subseteq\mathfrak{gl}(n,\mathbb{C})$.

[8] The Killing form is a symmetric bilinear form on a Lie algebra most simply defined by $(X,Y)=\mathrm{Tr}(\mathrm{ad}(X)\mathrm{ad}(Y))$. It can be extended by linearity to the complexification of an algebra.

It can be shown that an element of $(i\mathfrak{t})^*$ is analytically integral if and only if there exists some homomorphism $\xi_\lambda : T \to \mathbb{C} \setminus \{0\}$ which satisfies

$$\xi_\lambda(\exp(H)) = e^{\lambda(H)}$$

for all $H \in \mathfrak{t}$.

4. So, choose some analytically integral weight λ; this gives a one-dimensional representation of T in $\mathbb{C}$ by ξ_λ. Call this representation $\mathbb{C}_\lambda$, and write

$$L_\lambda = (G \times \mathbb{C}_\lambda)/\sim \qquad\qquad (gh, v) \sim (g, hv)$$

This can be projected into G/T by $(g,z) \mapsto gT$ and is acted on by G by $g(g',z) = (gg',z)$; from this it can be verified that L_λ is a holomorphic line bundle on G/T whose space of global sections are acted on homogeneously (i.e., continuously and transitively) by G.

A natural question is to ask is: can we obtain irreducible representations in this way? As a compact (thus semisimple) group, any finite-dimensional representation of G can be decomposed as a direct sum of irreducible representations. How can we characterize these? (A reducible representation is one that can be broken up without breaking G-invariance. Physically, this would correspond to a theory of two worlds, unable to communicate with each other via the action of G. This is both physically unreasonable and mathematically uneconomical.)

The answer to this question is given by the Borel-Weil theorem:

Theorem 1. (Borel-Weil) The irreducible unitary representations of G are exactly the global section spaces of L_λ corresponding to dominant weights λ. Further, each such representation is the highest weight representation for λ.

Now, a line bundle (or section of one) by itself is not a physical theory. To describe physics, we need a symplectic manifold (a symplectic manifold is one with some closed nondegenerate 2-form; this may describe classical restraints as a Lagrangian, or directly create a geometrical theory as in GR.) Using the orbit method we obtain such a thing. Specifically, there is a canonical symplectic structure on coadjoint orbits[9] [14], which is exactly what we want.

After all this mathematical work, the physical reader should be asking *what have we accomplished by all this nonsense?* The answer comes from the theory of *geometric quantization*, one of several mathematical theories intended to give a concrete description of the quantization process. A concise overview and references galore are given in [14, Section 4.6.1], which we paraphrase here:

A classical physical system is given by a symplectic manifold; present symmetries appear as a group G for which the manifold is a G-set. Quantum systems, on the other hand, are described by

[9] For any representation (π, V) of a Lie group G, there is a dual representation (π^*, V^*) in the dual space of V. This representation is given by $(\pi^*(g)f)(v) = f(\pi(g^{-1})v)$ for $g \in G$, $f \in V^*$, $v \in V$. Then the *coadjoint representation* of a Lie group is the dual in this sense to the adjoint representation. The orbits come from the action of G on this representation.

projection lattices on Hilbert spaces [4]; symmetry groups appear as unitary representations. These representations ought to be irreducible for the above reasons. Therefore, a natural definition of "quantization" should be a correspondence between G-invariant symplectic manifolds and unitary Hilbert space representations of G.

For technical reasons (related to quantum theories having a well-defined ground energy while classical ones leave this free), it is more appropriate to seek a *Poisson* structure rather than a symplectic one. A Poisson structure consists of a manifold with a Poisson bracket (a Lie bracket which is also a derivation). There is a canonical Poisson bracket on the space of real-valued smooth functions over any symplectic manifold; the generalization comes from the fact that Poisson brackets may be degenerate.

For our purposes, it turns out [14] that Poisson manifolds are essentially the same thing as coadjoint orbits!

With this context, we see that Borel-Weil, far from being an abstract muddle, gives a useful description of the quantization of symmetric classical theory in terms of integral roots[10].

Now, there was little need to use a TQFT here; we started with some compact Lie group, noted that its root space decomposition had a symmetric structure which could be described as a TQFT, then didn't mention this again. However, this is not just a toy example: in Section 5 we will see a connection between TQFT's and TQFT's of lower dimension, and the theory developed here will wind up deeply embedded in the background of any higher-dimensional TQFT. Further, this example provides intuition on the topological invariants related to unitary representations of compact groups; this is a very common theme in topological field theory [1, 6, 7, 24].

For now, we conclude that in 0-D, while there is perhaps nothing topological to say, our axioms can still be used to describe much of the representation theory of compact Lie groups.

4.2 Three-Dimensional Theory: Chern-Simons

The Chern-Simons theory is the one that started it all. Its history and motivation are elegantly described by Edward Witten in his paper [24]. It is tied up with knot theory (haha!), which is a deep and technically difficult subject. Therefore we will give only a cursory overview of the theory in this paper.

A *knot* is a piecewise smooth embedding of the circle S^1 in 3-space. A collection of knots is called a *link*. For a 2-knot link, we can define the *linking number* as the number of times one knot wraps around the other. A theorem of Gauss states that if you have two interlinked loops, and you ran a current through one, the line integral of the magnetic field around the other gives the linking number! This is actually the $U(1)$ Chern-Simons theory[11].

[10] Integral roots are well-understood; in fact, root systems form a rich mathematical theory unto themselves and are studied outside of the context of Lie theory.

[11] This, and other stories, can be found on John Baez' web page `http://math.ucr.edu/home/baez/symmetries.html`.

Chern-Simons theory was originally created as a physical realization of the Jones polynomial, a generalization of the linking number. (The Jones polynomial is a knot invariant and is mathematically fascinating, but does not admit a simple concrete description — for example, the Jones polynomial on the "unknot", or circle, is 1, but whether the unknot is the *only* knot with this value is an open question. Witten's paper gave physical insight and new computational results for this polynomial.)

So, let's consider 3D Yang-Mills theory. Let M be an oriented three-manifold with line bundle (e.g., tangent space) E, and G a gauge group which acts on E. Choose a connection A_μ^a on E (a is the Lie algebra index) with associated covariant derivative D_μ. Write the *field strength tensor* as

$$F_{ij} = [D_i, D_j] = \partial_i A_j - \partial_j A_i + [A_i, A_j]$$

If we had metric g, we could use the standard Yang-Mills action

$$\int_M \sqrt{g} g^{ik} g^{jl} \mathrm{Tr}(F_{ij} F_{kl}).$$

However, in seeking topological invariance, we decidedly do *not* have a metric. Fortunately, in 3D there is a topologically invariant quantity we can use instead: the *Chern-Simons action*

$$\begin{aligned}\mathscr{L} &= \frac{k}{4\pi} \int_M \mathrm{Tr}(A \wedge dA + \frac{3}{2} A \wedge A \wedge A) \\ &= \frac{k}{8\pi} \int_M \varepsilon^{ijk} \mathrm{Tr}(A_i(\partial_j A_k - \partial_k A_j) + \frac{3}{2} A_i [A_j, A_k]).\end{aligned}$$

(The origin and history of this quantity, which was originally described purely for geometrical reasons [3], are summarized in [24].) The stationary points of this action are described by $F_{ij} = 0$, the "flat connections".

Now, this action is not (quite) Gauge invariant — under a transformation with nonzero winding number the Lagrangian will be increased by some constant. However, in quantum field theory we care only about the complex exponential $\exp i\mathscr{L}$, so if we can force this constant to be a multiple of 2π we are okay. This can always be done by restricting the free parameter k (e.g., in [11, Section III] the $SU(2)$ case is considered and the restriction is $k \in \mathbb{Z}$).

Next, choose some oriented closed curve C on M — since M is a 3-manifold, C may be a knot, and even in the case that M is simply connected there are non-trivial embeddings to be considered. Denote by P_C the parallel transport map via A along C. (For any point $x \in M$, the collection of transport maps along closed curves based at x forms the *holonomy group* at x.)

Given some irreducible representation R of G (c.f. the previous Example), we define the *Wilson loop* of C as

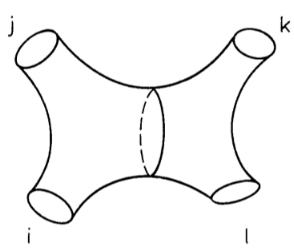
$$W_R(C) = \mathrm{Tr}_R P_C$$

These are the observables of our theory, and the problem of solving the theory amounts to calculating their expectation values. Notice that all quantities are topologically invariant.

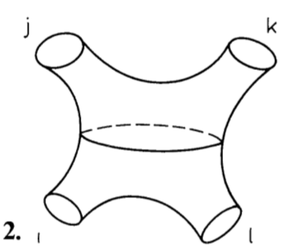
This theory can be solved on the 3-sphere S^3, but for a general manifold M the situation is more complicated. The strategy is to cut up M into several simpler (i.e., genus 1) manifolds, find the desired expectation values on these, then sew them back together and add up the resulting values to get a "topological state sum".

Let us stop for a moment to reframe this problem in terms of our general topological field theory axioms. Suppose we cut M along some 2D manifold Σ; Σ will become a boundary for the two resulting components of M. That is, M will be a bordism of Σ. So the state we seek on M is (in our axiomatic language) $Z^{2+1}(\Sigma)$, which will live in the space $Z^2(\Sigma)$. For the theory to make sense, both of these must be independent of the choice of Σ.

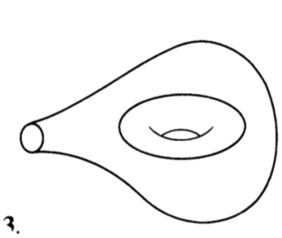
In cutting the manifold, we will also cut through Wilson loops, leaving marked points on the new boundary. Associated to these marked points are the loops' representations of our gauge group G; the tensor product of these representations will be the vector space $Z^2(\Sigma)$. To make this independent of Σ, we must (a) decompose carefully, and (b) use a "modular tensor category" for our vector spaces rather than an ordinary tensor category. The correct decomposition appears to be one involving *trinions*, or three-holed-spheres (or "pairs of pants"), and is described in [17], from which we lift the following image:



1.



2.


3.

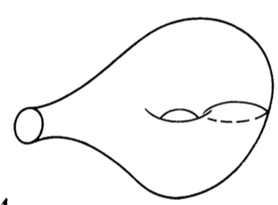

4.

Fig. 1. One sewing of the four-point function on the sphere

Fig. 2. Another sewing of the four-point function on the sphere

Fig. 3. One sewing of the one-point function on the torus

Fig. 4. Another sewing of the one-point function on the torus

Physically, what does this all mean? In the case that $G = U(1)$, we obtain Gauss' linking numbers, though this is perhaps not so interesting. In the case $G = SU(2)$, we can think of the Wilson loops as the paths of charged particles; our decompositions are then a sort of Feynman diagram in which the labelled points are virtual particles and the invariant sums on cuts correspond to conservation laws. Louis Crane has suggested that the $SU(2)$ case (with lengths instead of spins) may also describe quantum gravity, as explained in the next section, though the details are sketchy.

5 The Dimensional Ladder and Quantum Gravity

Time will perfect matter.
—Terrence McKenna

As our axiomatic definition uses manifolds of dimension d and $(d+1)$ in different ways, a natural question to ask is *what is the relationship between a d-dimensional TQFT and a $(d+1)$-dimensional one?*. We notice that the d-dimensional contains a "contracted" copy of a $(d+1)$-dimensional theory. Conversely, a $(d+1)$-dimensional theory can be found by "tracing out" a d-dimensional one.

This vague notion can be made rigorous, by the process of *categorification*, a term invented by Louis Crane. Categorification is an inverse process of *decategorification*, which means to take a category and identify all isomorphic objects.

Before describing how categorification applies to TQFT's, we make some general notes about the process [7]. First, it is intuitively similar to quantization, in the sense that (a) it is an ill-defined[12] inverse of a well-defined process, (b) it creates a more complicated theory based on a simpler one, and (c) it does so by "zooming in" on the components of the original theory to reveal their inner structure. A (very technical) mathematical definition is given in [9], though we will not need it here.

The idea to use categorification as a method of extending a 3D theory (in this case, Chern-Simons) to a 4D one including gravity came from the observation that categorification of TQFT's seems to produce a new TQFT more often than not [8]. Crane coined the phrase "dimensional ladder" to describe this phenomenon.

While the goal of topological quantum gravity has not been fully realized, Crane gives evidence that categorification is a fruitful direction to pursue in [7]. We summarize his arguments here.

First, we have eight principles for a theory of quantum gravity:

1. All observations require an observer. This means that a closed universe cannot be observed; no Hilbert space is associated to it. (Notice the philosophical similarity to Rovelli's relational quantum mechanics [19], which was published less than a year later.)

2. All observations are local. We consider observers as cells and associate Hilbert spaces to the surfaces between cells.

3. Observers observe each other: the Hilbert spaces associated to surfaces within the same 3-manifold are homomorphic to each other and are "consistent" in some sense. (Again, this is a principle of relational QM.)

[12] What is meant by "ill-defined" is that categorifications are neither guaranteed to exist nor to be unique. A precise definition does exist.

4. States of quantum gravity can be described by embedded graphs or knots. This reflects the (discrete) topological nature of gravity, and is consistent with both string-theoretic and loop-gravitational pictures.

5. General relativity is diffeomorphism invariant. This says that GR retains its fundamental significance; in fact, since the Atiyah-Segal axioms are defined to produce invariants under diffeomorphism groups, this principle is pretty-much required to describe gravity using a TQFT. (A relativist should read this as a strong argument supporting the topological picture.)

6. General relativity is a theory of geometry: length and time intervals appear as quantum states, subject to the same consistency and locality requirements as any other observable.

7. The Hamiltonian for general relativity is 0. This means no global time evolution; observation of time can only be local. (A null Hamiltonian is again a characteristic of a TQFT, signaling that we are on the right track. But notice that a TQFT does not necessarily have any notion of "local clocks".)

8. The classical limit is a global Hilbert space. (Topologically, this limit is taken by using the linear maps between local observers to link them, ignoring their interactions amongst themselves, and zooming out.)

The first three principles describe a relational picture of quantum mechanics. This is a radical extension of the principle of relativity, espoused by Carlo Rovelli, which essentially says: not only are time and space undefined except with respect to an observer; *all observables are*. The price of this view (objective reality) may seem unpalatable to some; however, as Rovelli argues in [19, 21], the seeming paradoxes of quantum mechanics disappear in a relational context, replaced instead by the mathematics of information theory. It is the author's opinion that relational QM will be required to make sense of quantum gravity.

Note that while mathematical conditions for "consistency" are given in [9], these may not be strong enough to ensure that observers always agree. (The importance of observers agreeing cannot be understated — the entire program of science depends on it!) It is not known what the correct consistency axioms should look like [19].

The next three principles contain the core of the theory, describing how gravity ought to be quantized. In the context of topological quantum field theory, they are not terribly remarkable: they say that the invariants and relative invariants assigned to manifold bordisms (and preserved under maps between manifolds) corresponds to the diffeomorphism invariance of general relativity. In other words, diffeomorphism invariance is a central tenet of Crane's theory, rather than "just another gauge theory".

The last principle is the least controversial. All it says is that a theory of gravity ought to match experiment. Probably it is only there to remind mathematicians, who like to forget it.

It is the seventh principle, the existence of local clocks, is the most interesting. After listing these eight proposals Crane drops a bombshell: *except for local clocks, the Chern-Simons TQFT satisfies all of these principles*. In other words, we "almost" already know the theory of quantum gravity.

Indeed, if we take the manifolds of the Chern-Simons TQFT as our "observer cells", the (unitary) linear observation maps as the maps associated to cobordisms between the cells, and the loop invariants to give probability amplitude for lengths (as opposed to spins), it is clear on a high level that we have a good match.

Unfortunately, this statement is not as dramatic as it seems. After all, local clocks are what we use to measure the results of experiment, and without them, we cannot make predictions. (Of course, one could argue that compared to the string theorists' lament "we need new mathematics", this is an excellent state of affairs.)

6 Closing Remarks

Not all who wander are lost.
—J. R. R. Tolkien

Mathematics and physics have a long and tightly interwoven history. Theories of Hilbert spaces, functional analysis and integration all have their roots in finding a rigorous foundation for physicists' need to handle massive quantities of data. Conversely, abstract logical leaps in geometry led to breakthroughs in our understanding of the potential structure of space and time. Calculus itself was invented by Newton and Liebniz for physical purpose. Throughout the centuries, appeal to physical intuition would give mathematics direction while mathematical rigour would ensure that physical models were consistent and sensible.

However, the 21st century has found physicists using the Feynman calculus, by far the most experimentally successful theory in history, with no sure footing in mathematics. This situation has persisted for some seventy years. Meanwhile, mathematicians have spent the same period recovering from devastating limitative results in set theory and logic, retreating fearfully to stranger and deeper abstractions.

The natural outcome of these trends is the application of algebraic topology to quantum physics. This is motivation behind topological quantum field theory, and as we have seen, it holds great promise for providing a rigorous backbone to field theory and for answering unknown questions about the nature and behavior of singularities. As experimentalists probe to higher and higher energies, we can expect topological language to provide a framework for describing the plethora of new results, much as the theories of gauge and quantum groups provided a framework for the zoo of particle physics.

Further still, algebraic topology is the natural setting for unifying the algebraic invariants of quantum mechanics with the observer-dependent philosophy of general relativity. While there is still much to be done, topological field theory holds great promise and may well prove to be a critical shift in our understanding of physics and reality.

References

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

[2] Michael Barr and Charles Wells, *Toposes, triples and theories*, Springer-Verlag, New York, 1985, Available online at http://www.tac.mta.ca/tac/reprints/articles/12/tr12abs.html.

[3] S.-S. Chern and J. Simons, *Characteristic forms and geometric invariants*, The Annals of Mathematics, Second Series **99** (1974), 48–69.

[4] David W. Cohen, *An introduction to Hilbert space and quantum logic*, Problem books in mathematics, Springer-Verlang, 1989.

[5] Henry Cohn, *Why symplectic geometry is the natural setting for classical mechanics*, http://research.microsoft.com/en-us/um/people/cohn/thoughts/symplectic.html.

[6] Louis Crane, *2-D physics and 3-D topology*, Commun. Math. Phys. **135** (1991), 615–640.

[7] ______, *Clock and category: Is quantum gravity algebraic?*, J. Math. Phys. **36** (1995), 6180–6193.

[8] Louis Crane and Igor B. Frenkel, *Four dimensional topological quantum field theory, Hopf categories, and the canonical bases*, J. Math. Phys. **35** (1994), 5136.

[9] Louis Crane and David N. Yetter, *Examples of categorification*, Cahiers de topologie et gémotétrie différentielle catégoriques **39** (1998), 3–25.

[10] Anadijiban Das and Andrew DeBenedictis, *The general theory of relativity: a mathematical exposition*, Springer, 2012.

[11] S. Deser, R. Jackiw, and S. Templeton, *Topologically massive gauge theory*, Ann. Phys. NY **281** (2000), no. 1, 409.

[12] William Fulton and Joe Harris, *Representation theory. a first course*, Graduate Texts in Mathematics, Readings in Mathematics, vol. 129, Springer-Verlag, New York, 1991.

[13] Sigurdur Helgason, *Differential geometry, lie groups, and symmetric spaces*, Academic Press, 1978.

[14] A.A Kirillov, *Lectures on the orbit method*, Graduate Studies in Mathematics, vol. 64, American Mathematical Society, 1936.

[15] Saunders Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Mathematics, vol. 5, Springer-Verlag, 1998.

[16] J.P. May, *A concise course in algebraic topology*, University of Chicago Press, Chicago, 1999.

[17] G. Moore and N. Seiberg, *Classical and quantum conformal field theory*, Commun. Math. Phys. **123** (1989), 177.

[18] Claudio Procesi, *Lie groups: An approach through invariants and representations*, Universitext, Springer, 2007.

[19] Carlo Rovelli, *Relational quantum mechanics*, International Journal of Theoretical Physics **35** (1996), 1637–1678.

[20] Mark Sepanski, *Compact lie groups*, Graduate Texts in Mathematics, vol. 235, Springer, 2007.

[21] Matteo Smerlak and Carlo Rovelli, *Relational EPR*, Foundations of Physics **37** (2007), 427 – 445.

[22] René Thom, *Quelques propriétés globales des variétés différentiables*, Comment. Math. Helv. **28** (1954), 17–86.

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

[24] ______, *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. **121** (1989), 351–399.