Lecture 1

Topological field theories

1.1 The basic structure

The idea of a topological field theory is due to Witten, but the definition was first written down formally by Atiyah in [1]. From one angle, an n-dimensional topological field theory is a rule which gives one a complex number Ψ_Y for each closed oriented n-manifold Y, the number depending only on the diffeomorphism class of Y. The essential feature is that the assignment $Y \longmapsto \Psi_Y$ is *local* with respect to Y in the sense that when Y is the union $Y = Y_1 \cup Y_2$ of two manifolds Y_1 and Y_2 with a common boundary X the theory gives one a procedure for calculating Ψ_Y from contributions associated to Y_1 and Y_2.

The situation we are trying to formalize here comes from the path-integral approach to quantum field theory. In field theory we consider a space $\mathcal{F}(Y)$ of "fields" which are functions of some kind on the space-time manifold Y. There is supposed to be a measure on the space $\mathcal{F}(Y)$: it is written formally as $e^{-S(\phi)}\mathcal{D}\phi$, with $\phi \epsilon \mathcal{F}(Y)$, where $S : \mathcal{F}(Y) \to \mathbb{R}$ is a functional called the "action". Then we can define a number

$$\Psi = \int_{\mathcal{F}(Y)} e^{S(\phi)}\mathcal{D}\phi \tag{1.1.1}$$

The locality property comes from the fact that a field ϕ on $Y_1 \cup Y_2$ is a pair of fields $\phi_i \in \mathcal{F}(Y_i)$ which agree on the interface $\partial Y_1 = \partial Y_2 = X$, while the action is additive: $S(\phi) = S(\phi_1) + S(\phi_2)$. This permits us to write, very schematically,

$$\Psi_Y = \int_{\mathcal{F}(X)} \Psi_{Y_1}(\psi)\Psi_{Y_2}(\psi)\mathcal{D}\psi, \tag{1.1.2}$$

where, for $\psi \in \mathcal{F}(X)$,

$$\Psi_{Y_i}(\psi) = \int e^{-S(\phi)}\mathcal{D}\phi_i,$$

the integral being over all fields $\phi_i \in \mathcal{F}(Y_i)$ with boundary value ψ.

The physical background suggests that it is appropriate to think of the space-time Y as interpolating between an initial space X_0 and a final space X_1,

and of the path-integral as defining an integral operator $\Psi \longmapsto K\Psi$ from functions on $\mathcal{F}(X_0)$ to functions on $\mathcal{F}(X_1)$:

$$K\Psi(\psi_1) = \int\limits_{\mathcal{F}(X_0)} K(\psi_1, \psi_0)\Psi(\psi_0)\mathcal{D}\psi_0, \tag{1.1.3}$$

where

$$K(\psi_1, \psi_0) = \int e^{-S(\phi)}\mathcal{D}\phi, \tag{1.1.4}$$

the integral being over all $\phi \in \mathcal{F}(Y)$ with $\phi|X_i = \psi_i$.

Path-integral formulae like (1.1.1) — (1.1.4) are the crucial mythological background to everything that will follow, but they do not come into the formal development. We give the axiomatic formulation in two steps, beginning with the $(n-1)$-dimensional interfaces, or time-slices. Thus the first part of the structure of a topological field theory is a functor

$$E : \left\{ \begin{array}{c} \text{closed oriented } (n-1)\text{-manifolds} \\ \text{and diffeomorphisms} \end{array} \right\} \longrightarrow \{\text{complex vector spaces}\}.$$

We think of $E(X)$ as the vector space of functions on some space of fields on X, but the space of fields is *not* part of the structure.

The functor is required to take disjoint unions to tensor products:

$$E(X_1 \amalg X_2) = E(X_1) \otimes E(X_2). \tag{1.1.5}$$

This means, strictly speaking, that for any finite family $\{X_\alpha\}$ of manifolds one is given a natural multilinear map

$$m : \Pi E(X_\alpha) \longrightarrow E(\amalg X_\alpha) \tag{1.1.6}$$

with the universal property which makes $E(\amalg X_\alpha) \cong \otimes E(X_\alpha)$. (There is no ordering of the indices presupposed here.) The maps m are required to be coherent in the obvious sense when a family $\{X_\alpha\}$ is a union of subfamilies. It follows from the universal property of (1.1.6) that the isomorphism

$$E(X)^{\otimes n} \longrightarrow E(\underbrace{X \amalg X \amalg \cdots \amalg X}_{\longleftarrow n \longrightarrow})$$

is equivariant with respect to the symmetric group S_n.

The second part of the structure of the theory assigns a linear map

$$\Psi_Y : E(X_0) \longrightarrow E(X_1)$$

to each cobordism Y from X_0 to X_1. I shall write $Y : X_0 \rightsquigarrow X_1$ to indicate that Y is a cobordism from X_0 to X_1. The maps Ψ_Y are required to be compatible with composition of cobordisms, and also with the maps m. The structure could be expressed concisely by using the notion of a 2-category, but I shall avoid that.

In any case, (1.1.5) implies that $E(X) = \mathbf{C}$ if X is empty. A closed n-manifold Y is a cobordism $Y : \varnothing \rightsquigarrow \varnothing$,and so gives us a number

$$\Psi_Y \in \text{ End } (\mathbf{C}) = \mathbf{C}.$$

For any $(n-1)$-manifold X the cobordism

$$X \times I : X \rightsquigarrow X,$$

where I is the interval [0,1], induces an idempotent operator $E(X) \to E(X)$ which breaks $E(X)$ into the sum of an interesting and a discardable piece. It is therefore natural to add

Assumption 1.1.7 *The cobordism $X \times I$ induces the identity map of $E(X)$.*

Given the definition, some natural questions arise at once.

(*i*) Do the numbers Ψ_Y associated to closed n-manifolds Y determine the theory completely?

(*ii*) Are there restrictions on the kinds of diffeomorphism invariants Ψ_Y which can arise?

(*iii*) Is the vector space $E(X)$ spanned by the vectors Ψ_Y associated to n-manifolds Y with $\partial Y = X$?

(*iv*) Can one describe the totality of n-dimensional theories? For example, is there a moduli space?

We shall return to these questions presently.

The first important formal consequence of the axioms is that the vector spaces $E(X)$ are finite-dimensional, and that $E(X)$ is naturally dual to $E(\overline{X})$, where $\overline{E}$ denotes X with its orientation reversed. To see that, we use the following elementary remark.

Proposition 1.1.8 *Two modules M and N over a commutative ring A are finitely generated, projective, and in duality, if and only if there are homomorphisms $\alpha : A \to M \otimes N$ and $\beta : N \otimes M \to A$ such that the compositions*

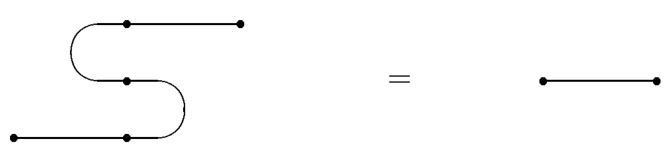
$$M \xrightarrow{\alpha\otimes 1} M \otimes N \otimes M \xrightarrow{1\otimes\beta} M$$

and

$$N \xrightarrow{1\otimes\alpha} N \otimes M \otimes N \xrightarrow{\beta\otimes 1} N$$

are the identity.

The proposition applies directly to the situation at hand, because $X \times I$ can be regarded as a cobordism $\alpha : \varnothing \rightsquigarrow X \amalg \overline{X}$ and also as a cobordism $\beta : X \amalg \overline{X} \rightsquigarrow \varnothing$, and the relation $(1 \otimes \beta) \circ (\alpha \otimes 1) = 1$ is illustrated by



We labour this point because it will recur in a more general context later.

To give a 1-dimensional theory is exactly the same thing as to give a finite-dimensional vector space V, for there is just one connected 0-manifold, and it has two possible orientations P and $\overline{P}$, with $E(P) = V$ and $E(\overline{P}) = V^*$. The number Ψ_{S^1} is necessarily the positive integer $\dim(V)$.

There is a folk theorem which completely describes 2-dimensional theories.

Theorem 1.1.9 *To give a 2-dimensional theory is the same as to give a finite-dimensional commutative algebra A over $\mathbf{C}$ with a 1, together with a linear map $\theta : A \to \mathbf{C}$ such that $(x, y) \mapsto \theta(xy)$ is a nondegenerate bilinear form on A.*

We shall call such a map θ a **nondegenerate trace**, and an algebra with such a trace will be called a (commutative) **Frobenius algebra**.

One half of the proof of (1.1.9) is obvious. There is just one connected oriented 1-manifold S^1, and when a theory E is given we define $A = E(S^1)$. Maps $\mathbb{C} \to A$

and $A \to \mathbb{C}$ corresponding to the unit and to θ are given by the disc D^2 regarded as a cobordism $\varnothing \rightsquigarrow S^1$ or $S^1 \rightsquigarrow \varnothing$. The multiplication $A \otimes A \to A$ is defined by the cobordism $Y : S^1 \amalg S^1 \rightsquigarrow S^1$, where

There is no difficulty in showing that A and θ have the properties claimed. The most interesting point is that the multiplication is commutative because the surface Y admits a diffeomorphism which interchanges the two inner boundary circles while leaving the outer one fixed.

Conversely, if A is a commutative ring with a non-degenerate trace it is not hard to convince oneself that one can assign a linear map

$$\Psi_\Sigma : A^{\otimes p} \to A^{\otimes p}$$

to any surface Σ with p incoming and q outgoing boundary circles so that we have a field theory. All the same, I do not know an illuminating proof of this. By slicing Σ up into discs and "pairs of pants" like Y one gets a candidate for Ψ_Σ, but the point is to show that Ψ_Σ does not depend on the decomposition, i.e. that the changes relating different decompositions are precisely reflected in the algebraic properties of a commutative ring. We shall return to this question later.

The simple concrete description of 2-dimensional theories provides answers to the four questions we raised. But first we need a few elementary formal properties of the theory associated to a Frobenius algebra (A, θ).

(a) There is a distinguished element $\alpha \in A$ defined by a torus with one outgoing boundary circle. If Ψ_g is the invariant for a closed surface of genus g, then

$$\Psi_g = \theta(\alpha^g). \tag{1.1.10}$$

(b) If $\{e_i\}$ is a vector-space basis of A, and $\{e_i^*\}$ is the dual basis (i.e. $\theta(e_i e_j^*) = \delta_{ij}$) then

$$\alpha = \sum e_i e_i^*. \tag{1.1.11}$$

(c) The regular representation $A \to \mathrm{End}(A)$ is an embedding, and for any $\alpha \in A$ we have

$$\theta(a\alpha) = \mathrm{tr}\,(a), \tag{1.1.12}$$

where tr denotes the trace in the regular representation.

(d) The algebra A is semi-simple, i.e. a sum of copies of $\mathbb{C}$, if and only if α is invertible.

(e) If the characteristic polynomial of α in the regular representation is χ, where

$$\chi(t) = \Pi(t - \lambda_i) = t^n + a_1 t^{n-1} + \cdots + a_n$$

then

$$\Psi_g = \sum \lambda_i^{g-1} \quad \text{when } g \geq 1, \tag{1.1.13}$$

so that

$$\sum_{g \geq 1} \Psi_g t^{-g} = \chi'(t)/\chi(t) \tag{1.1.14}$$

for large t. In the semi-simple case, (1.1.13) holds also for $g = 0$.

(f) Rescaling by replacing θ by $\lambda\theta$ changes Ψ_g to

$$\lambda^{1-g}\Psi_g = \lambda^{\frac{1}{2}\chi(X)},$$

where $\chi(X)$ is the Euler number of X.

I shall leave the verification of these properties to the reader. They tell us that in general a 2-dimensional theory is not determined by the number Ψ_Y associated to closed manifolds Y, and that $E(X)$ is not necessarily spanned by the vectors Ψ_Y with $\partial Y = X$. On the other hand, the answer to both questions is yes in the generic case when the polynomial χ has distinct roots.

As to the invariants that can arise, the main restriction is that $\Psi_Y = \Psi_1$ must be a positive integer n when Y is a torus. Apart from that, $\Psi_0, \cdots, \Psi_n$ can be prescribed arbitrarily, and then the rest are uniquely determined. The generic theories with $\Psi_1 = n$ are determined by the monic polynomial χ, so form an n-dimensional complex affine space.

1.2 The 'toy model' for a finite group

In thinking about topological field theories it is useful to have a simple family of concrete examples in mind. Dijkgraaf and Witten [DW] have pointed out that there is a topological field theory (in any number of dimensions) naturally associated to a finite group G. The surprising thing is that a great part of quantum field theory is the study of various generalizations of this toy model in which finite groups are replaced by Lie groups.

In the model, the invariant Ψ_Y for a closed n-manifold Y is the weighted number of principal G-bundles on Y, each bundle being given the weight $1|\mathrm{Aut(P)}|$, where $\mathrm{Aut}(P)$ is the group of automorphisms of P. (I shall comment on this choice of weight in a moment). If Y is connected, then a G-bundle P is determined up to isomorphism by its holonomy, a homomorphism $\rho : \pi_1(Y) \to G$. The group $\mathrm{Aut}(P)$ is the subgroup of elements of G which commute with the image of ρ. So Ψ_Y counts the conjugacy classes of homomorphisms from the fundamental group $\pi_1(Y)$ into G. Taking account of the weighting, we have

$$\Psi_Y = |\mathrm{Hom}(\pi_1(Y); G)|/|G|.$$

The vector space $E(X)$ for a closed $(n-1)$-manifold X is the space of complex-valued functions on the finite set $\mathcal{P}_X$ of isomorphism classes of G-bundles on X. If $\partial Y = X$ then $\Psi_Y : \mathcal{P}_X \to \mathbb{C}$ is defined by

$$\Psi_Y(P) = \sum_Q \frac{1}{|\mathrm{Aut(Q)}|}$$

where Q runs through the G-bundles on Y such that $Q|X = P$.

It is an easy exercise to check that the 2-dimensional version of this theory corresponds to the Frobenius algebra (A, θ), where A is the centre of the group-ring $\mathbb{C}[G]$, and

$$\theta(\sum \lambda_g g) = \frac{1}{|G|}\lambda_1.$$

If we decompose the group-ring

$$\mathbb{C}[G] = \bigoplus \mathrm{End}\ (V),$$

where V runs through the irreducible representations of G, then we find

$$\theta(1_V) = \lambda_V^{-1},$$

where 1_V is the identity-element of $\mathrm{End}(V)$, and $\lambda_V = |G|^2/\dim\ (V)^2$, while

$$\alpha = \sum \lambda_V 1_V.$$

Thus

$$\Psi_\Sigma = |G|^{2g-2} \sum \frac{1}{(\dim V)^{2g-2}}$$

if Σ is a surface of genus g.

A disconcerting feature of this example is that it does not "remember" the group G: for example, two different abelian groups of the some order give rise to indistinguishable theories. Furthermore, for an abelian group the vector Ψ_Y, when $\partial Y = S^1$, is always a scalar multiple of the identity element of $E(S^1) = \mathbb{C}[G]$.

Finally, concerning the weighting with which the G-bundles were counted in this model, I should point out that in combinatorial problems it is often "natural" — or at any rate easiest — to count the objects of a category after weighting them with the reciprocal of their number of automorphisms. A typical example of the advantage gained arises when one has a category in which each object has a positive integral "size", and one is interested in the generating function

$$Z(q) = \sum_{n \geqslant 0} a_n q^n,$$

where a_n is the number of objects of size n. If each object can be decomposed uniquely up to ordering as a sum of "irreducible" or "connected" objects then if we count with weights we have $Z = e^F$, where

$$F(q) = \sum b_n q^n$$

and b_n is the number of connected objects of size n. But this is not true if we count without weights, as we see from the example of counting finite sets.

In the case of counting G-bundles on a space Y, the effect of the weight is that if we pick a *arbitrary* finite subset S of Y, such as the vertices of a triangulation, then

$$\Psi_Y = |\{\text{bundles on } Y \text{ trivialized over } S\}|/|G|^{|S|}.$$

This fact is important for understanding more general gauge theories.

1.3 Open Strings

One way to extend the notion of a 2-dimensional topological field theory is to assign a vector space $E(X)$ to each compact oriented 1-dimensional manifold, with or without boundary. One still requires

$$E(X_1 \amalg X_2) = E(X_1) \otimes E(X_2).$$

Any compact 1-manifold is a union of circles and intervals, so this part of the data amounts to two vector spaces

$$A = E(S^1) \text{ and } B = E(I).$$

Cobordisms $Y : X_0 \rightsquigarrow X_1$ must now be taken to be surfaces Y whose boundary ∂Y is the union of $\overline{X_0} \amalg X_1$ with a "free" part $\partial_f Y$, which is itself a cobordism $\partial_f Y : \partial X_0 \rightsquigarrow X_1$.

Example The surfaces

where the broken curves indicate the free part of the boundary, give us $Y : S^1 \rightsquigarrow I$

and $Z : S^1 \rightsquigarrow S^1$. We also have $\overline{Y} : I \rightsquigarrow S^1$, and $Y \circ \overline{Y} = Z$.

We shall call a theory of this type an *open-string* theory. The map $\Psi_Y : E(S^1) \to E(I)$ induced by the surface Y just illustrated will be denoted by $i : A \to B$.

It is straightforward to check that:

(i) A is a commutative algebra with a 1 and a non-degenerate trace $\theta_A : A \to \mathbb{C}$;

(ii) B is an algebra with a 1 and a non-degenerate trace $\theta_B : B \to \mathbb{C}$;

(iii) the map $i : A \to B$ is an algebra homomorphism such that $i(1) = 1$, and its image is contained in the centre of B.

It is not usually true that $\theta_B \circ i = \theta_A$. In fact $\theta_B = \theta_A \circ i^*$, where $i^* : B \to A$ is the homomorphism of A-modules adjoint to i, and $i^* i$ is multiplication by $i^* 1 \in A$.

The analogue of Theorem (1.1.9) is

Theorem 1.3.1 *To given an open-string theory is the same as to give a collection of data $\{A, B, \theta_A, \theta_B, i\}$ with the properties (i), (ii), and (iii).*

An open-string theory gives us a number $\Psi_{g,n}$ for a surface Σ of genus g with n boundary circles, for we can regard Σ as a cobordism $\varnothing \rightsquigarrow \varnothing$ with a free boundary.

It is clear that

$$\Psi_{g,n} = \theta_A(\alpha^g \beta^n), \tag{1.3.2}$$

where $\beta = i^\star 1 \in A$, so these numbers depend only on the "closed-string" theory (A, θ_A) together with the element β of A.

The 'toy model' associated to a finite group G extends naturally to an open-string theory. For a surface Y with boundary we define Ψ_Y as the weighted number of G-bundles on Y which are trivialized on the boundary ∂Y. The algebra B is then the group ring $\mathbb{C}[G]$, and A is, as before, its centre. The trace $\theta_B : \mathbb{C}[G] \to \mathbb{C}$ is $|G|^{-1}$ times the trace in the regular representation, and it restricts to $|G|\theta_A$. The element β is the scalar $|G|$.

Linear categories

If we are going to consider open-string theories it is natural to allow a little more generality. The number Ψ_Y associated to a surface Y with boundary should be thought of as the value of a path integral over fields on Y which satisfy some "boundary condition" on ∂Y. We can contemplate imposing different boundary conditions on the different boundary components. Let Λ be the set of possible boundary conditions. Then we shall have a number $\Psi_{Y,\lambda}$ for each labelling $\lambda : \amalg_0(\partial Y) \to \Lambda$. In this situation we shall have a vector space

$$E(\lambda_0; \lambda_1) = E(I; \lambda_0, \lambda_1)$$

corresponding to the oriented interval I with its ends labelled by λ_0 and λ_1 in Λ. Maps between these vector spaces will come from cobordisms Y whose "free" boundary components are labelled. Thus the surface

gives us an associative bilinear composition-law

$$E(\lambda_0; \lambda_1) \times E(\lambda_1; \lambda_2) \to E(\lambda_0; \lambda_2).$$

A moment's reflection shows that Λ is the set of objects of a $\mathbb{C}$-linear category in which $E(\lambda_0; \lambda_1)$ is the vector space of morphisms $\lambda_0 \to \lambda_1$. There is a trace $\theta_\lambda : E(\lambda; \lambda) \to \mathbb{C}$ which induces a non-degenerate inner product on $E(\lambda; \lambda)$. Finally, the commutative ring $A = E(S^1)$ acts on each object of the category, and all morphisms in the category commute with this action. Conversely, any $\mathbb{C}$-linear category with non-degenerate traces gives rise to an open-string theory with labels.

In our toy model it is natural to take Λ to be the set of all finite dimensional complex representations of G, and $E(\lambda_0\,; \lambda_1)$ to be the space of G-maps $\lambda_0 \to \lambda_1$. We could, however, use any subset of Λ; if we allow only the regular representation we get the unlabelled theory described earlier.

In any case, in this theory the number $\Psi_{Y,\lambda}$ associated to a surface Y, with boundary circles S_i labelled λ_i, is the sum of a contribution

$$\frac{1}{|\mathrm{Aut}(P)|} \prod_i \chi_{\lambda_i}(h_i(P))$$

for each G-bundle P on Y, where $h_i(P)$ is the holonomy of P around S_i, and χ_{λ_i} is the character of the representation λ_i.

In Lecture 5 we shall meet an important example of this formalism which arises in string theory. We start with a symplectic manifold M, and let Λ be the set of Lagrangian submanifolds of M. We take $E(\lambda_0; \lambda_1)$ to be the Floer cohomology of the space of smooth paths in M which begin in λ_0 and end in λ_1. Then, for a surface Y with labelled boundary components, $\Psi_{Y,\lambda}$ counts the number of pseudoholomorphic maps $Y \to M$ which take the boundary circles into the prescribed Lagrangian submanifolds.

1.4 Area-dependent theories

Witten has pointed out that in any number of dimensions one can modify the definition of a topological field theory by allowing the operators $\Psi_Y : E(X_0) \to E(X_1)$ to depend on a volume-form given on Y. Because any two volume-forms on Y with the same total volume are related by a diffeomorphism of Y (Moser[M]), the effect of this change is simply that the operator $\Psi_Y^{(t)}$ associated to Y depends on a number $t > 0$ which is the volume of Y, and

$$\Psi_{Y_1 \circ Y_2}^{(t_1+t_2)} = \Psi_{Y_1}^{t_1} \circ \Psi_{Y_2}^{t_2}.$$

The axioms no longer imply that the vector spaces $E(X)$ are finite dimensional, and so we shall take them to be locally convex and complete topological vector spaces, and shall interpret the tensor product in (1.1.5) as a completed topological tensor product. On each space $E(X)$ we have a semigroup of operators coming from the cobordism $X \times I :$ we shall write it $\{U_t\}_{t>0}$. The argument which we used before to prove that $E(X)$ was finite dimensional now proves that U_t is of trace class. The analogue of the non-degeneracy assumption (1.1.7) is

Assumption 1.4.1 *$U_t \to 1$ as $t \to 0$, uniformly on compact subsets of $E(X)$.*

The semigroup $\{U_t\}$ defines a "rigging" of $E(X)$, i.e. two complete topological vector spaces $\check{E}(X)$ and $\hat{E}(X)$, with maps

$$\check{E}(X) \to E(X) \to \hat{E}(X)$$

which are injective with dense images. As a set, $\check{E}(X)$ is the union of the images of U_t for $t > 0$, while an element of $\hat{E}(X)$ can be written as a formal expression $U_t^{-1}\xi_t$ for every $t > 0$, with $\xi_t \in E(X)$; more precisely, $\check{E}(X)$ and $\hat{E}(X)$ are the direct and inverse limits respectively of systems $\{E_{(t)}\}$ of copies of $E(X)$ indexed by $t > 0$, the maps $E_{(t)} \to E_{(t')}$ being given, when $t < t'$ or $t > t'$ respectively, by $U_{|t-t'|}$. The analogue of the duality property (1.1.8) is

Proposition 1.4.2 *For any $n-1$ manifold X the spaces $\check{E}(X)$ and $\hat{E}(\bar{X})$ are naturally in duality, by a bilinear form which identifies $\check{E}(X)$ with the dual of $\hat{E}(\bar{X})$ and $\check{E}(\bar{X})$ with the dual of $\hat{E}(X)$.*

Proof: The argument is essentially the same as for (1.1.8): the pairing of $U_s\xi \in \check{E}(X)$ and $U_t^{-1}\eta \in \check{E}(\bar{X})$ for $s > t$ is

$$\Psi_{X\times I}^{(s-t)}(\xi \otimes \eta).$$

There is clearly no harm in assuming $\check{E}(X) = E(X)$, and we shall do so from now on.

Let us now specialize to the two-dimensional case. A moment's reflection shows that the analogue of (1.1.9) is

Proposition 1.4.3 *A two-dimensional area-dependent theory is the same thing as a commutative topological algebra A with a non-degenerate trace $\theta : A \to \mathbb{C}$ and a trace-class approximate unit, i.e. a family $\{\varepsilon_t\}_{t>0}$ in A such that*

(i) $\varepsilon_t \to 1$ *as* $t \to 0$,

(ii) $\varepsilon_s\varepsilon_t = \varepsilon_{s+t}$, *and*

(iii) multiplication by ε_t is a trace-class operator $A \to A$.

The element $\varepsilon_t \in A = E(S^1)$ is associated to the disc with area t, and multiplication $A \otimes A \to A$ is defined by

$$(U_s\xi) \cdot (U_t\eta) = \Psi_\Sigma^{(s+t)}(\xi \otimes \eta),$$

where Σ is a disc with two holes. Thus the operator U_t is multiplication by ε_t. Obvious analogues of the formulae (1.1.10) - (1.1.12) hold for an area-dependent theory: instead of the single element $\alpha \in A$ we have $\alpha_t \in A$ for $t > 0$.

Two ways of rescaling the theory are by multiplying θ by $\lambda \in \mathbb{C}$, and by multiplying ε_t by $e^{-\mu t}$ with $\mu > 0$. The effect, for a closed surface Y of area t is to multiply Ψ_Y by

$$\lambda^{\frac{1}{2}\chi(Y)}e^{-\mu t}.$$

There is an exactly analogous result applying to area-dependent open-string theories. The only change needed in Proposition (1.3.1) is that the algebras A and B can be infinite dimensional, and they have a common trace-class approximate unit $\{\varepsilon_t\}$.

What makes it worthwhile to spell out this unexciting-looking formalism is that there is a beautiful example, discovered, as usual, by Witten []. For a finite group G we have already described the two-dimensional open- and closed-string theories associated with the group-ring $\mathbb{C}[G]$ and its centre. For these, the invariant Ψ_Σ associated to a surface Σ was the weighted number of G-bundles on Σ, trivialized on the boundary if any. For a more general group, Yang-Mills theory provides us with a way of "counting" bundles. I shall first describe the outcome formally, to emphasize its simplicity, and then — in the next section — I shall discuss its significance and applications.

If G is a compact Lie group then it is natural to generalize $\mathbb{C}[G]$ to the ring $\mathcal{F}_G$ of smooth functions on G under convolution. For the trace we take

$$\theta(f) = f(1).$$

The ring $\mathcal{F}_G$ does not have a unit, for its natural unit would be the delta-function δ at the identity element of G. The most obvious choice of approximate unit is to take ε_t to be the heat kernel, i.e. the fundamental solution of the equation

$$\frac{\partial \varepsilon_t}{\partial t} = \Delta \varepsilon_t\,,$$

where Δ is the Laplacian. Thus ε_t is the smooth function to which δ diffuses in time t.

The ring $\mathcal{F}_G$ is a dense subring of the product of the endomorphism rings $\mathrm{End}(V)$ of the irreducible representations V of G, and

$$\theta \;=\; \sum_V (\dim V)\ \mathrm{tr}_V,$$

where $\mathrm{tr}_V : \mathrm{End}(V) \to \mathbb{C}$ is the usual trace. We find at once that the invariant for a closed surface of genus g and area t is

$$\sum_V \frac{e^{-t\lambda_V}}{(\dim V)^{2g-2}},$$

where λ_V is the eigenvalue of the Casimir operator on V, i.e. $-\lambda_V$ is the image of Δ in $\mathrm{End}(V)$.

For each homomorphism $\omega : \pi \to \mathbb{T}$, where $\pi = \pi_1(G)$ is the fundamental group of G there is a natural variant $\mathcal{F}_G^\omega$ of $\mathcal{F}_G$. It gives us another field theory, whose geometric significance I shall explain in the next section. Regarding π as a subgroup of the centre of the simply-connected covering group $\widetilde{G}$ of G, we define $\mathcal{F}_G^\omega$ as the subring of $\mathcal{F}_{\widetilde{G}}$ of function f on $\widetilde{G}$ satisfying

$$f(zg) = \omega(z) f(g)$$

for all $z \in \pi$.

1.5 Yang-Mills theory in two dimensions

In Yang-Mills theory for a compact Lie group G on a surface Σ we consider pairs (P, A), where P is a principal G-bundle on Σ and A is a connection in P. The curvature of (P, A) is described by the *curvature form* $F_A \in \Omega^2(\Sigma; \mathfrak{g}_P)$, a 2-form whose value at $x \in \Sigma$ lies in the Lie algebra $\mathfrak{g}_{P_x}$ of infinitesimal automorphisms of the fibre of P at x. The total amount of curvature is measured by the *Yang-Mills action*

$$S(A) = \frac{1}{2} \int_\Sigma < F_A, *F_A >,$$

defined in terms of a given invariant inner product on the Lie algebra $\mathfrak{g}$ of G. The action $S(A)$ involves the area-element of Σ through the operator $* : \Omega^2 \to \Omega^0$, but we do not need a Riemannian structure on Σ. Multiplying the area of Σ by λ multiplies $S(A)$ by λ^{-1}.

The object of Yang-Mills theory is to calculate

$$\Psi_\Sigma = \int e^{-S(A)} \mathcal{D}A, \tag{1.5.1}$$

the integral being over all isomorphism-classes of pairs (P, A) on Σ. More generally, if Σ has a boundary $\partial\Sigma$, and we give a bundle-with-connection (P_0, A_0) on $\partial\Sigma$, then we want to calculate

$$\Psi_\Sigma(P_0, A_0) = \int e^{-S(A)} \mathcal{D}A, \tag{1.5.2}$$

where now the integral is over all (P, A) which restrict to (P_0, A_0). As a bundle-with-connection on a circle is determined up to isomorphism by the conjugacy-class of its holonomy, the expression (1.5.2) is a function on $G \times \cdots \times G$, with one factor for each boundary circle.

A homomorphism $\omega : \pi_1(G) \to \mathbb{T}$ gives us a characteristic class $\omega(P) \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ for G-bundles on surfaces, for we can think of ω as an element of $H^2(BG; \mathbb{T})$, where BG is the classifying space for G-bundles. I shall write $\omega(P)$ symbolically as

$$\omega(P) = \exp i \int_\Sigma \omega(A).$$

For each ω there is a modified Yang-Mills theory with

$$\Psi^\omega_\Sigma = \int e^{-S(A) + i\omega(A)} \mathcal{D}A.$$

This section has two aims. First, to explain why it is reasonable to believe that the vaguely-defined integrals (1.5.1) and (1.5.2) are given precisely by the simple theory described in §1.4. Secondly, to explain another quite independent remarkable fact, which is that the theory of §1.4 encodes a great deal of information about the geometry of the moduli spaces of flat G-bundles on Σ.

To understand the integral (1.5.1) it may be best to start with the abelian case $G = \mathbb{T}$, for then the integral is essentially Gaussian, and we can calculate it exactly. The connected components of the space of pairs (P, A) correspond to the topological type, i.e. to the isomorphism class of the bundle P, which is given by its first Chern class $n \in H^2(\Sigma; \mathbb{Z}) = \mathbb{Z}$, and is represented by the closed form F_A (which in this case has scalar values). In each component the connections with minimal action are those with $*F_A$ constant, and hence

$$S(A) = n^2/2t,$$

where t is the area of Σ. The connections in a given bundle P are an affine space of $\Omega^1(\Sigma)$, and (P, A) is isomorphic to $(P, \tilde{A})$ if and only if

$$\tilde{A} - A = \frac{1}{2\pi i} f^{-1} df$$

for some $f : \Sigma \to \mathbb{T}$. Thus each component of the space to be integrated over is a copy of

$$\Omega^1(\Sigma)/\text{ image Map}(\Sigma; \mathbb{T}).$$

If we introduce a Riemannian metric on Σ we can write

$$\Omega^1(\Sigma) = P \oplus H \oplus Q,$$

where P are the exact forms, H are the harmonic forms, and $Q = *P$. The image of $\text{Map}(\Sigma; \mathbb{T})$ is $P \oplus H_{\mathbb{Z}}$, where $H_{\mathbb{Z}}$ denotes the harmonic forms with integral periods. So each component is

$$(H/H_{\mathbb{Z}}) \times Q,$$

where $H/H_{\mathbb{Z}} \cong H^1(\Sigma; \mathbb{T}) \cong \text{Hom}(\pi_1(\Sigma); \mathbb{T})$ is the Jacobian torus J_Σ of Σ. The function to be integrated is

$$\exp\{-\frac{n^2}{2t} + \frac{1}{2}\int dq. * dq\}$$

on the n^{th} component.

The result is

$$K\Theta(t) \text{ vol } (J_\Sigma),$$

where K is the Gaussian integral defined by the quadratic form $\frac{1}{2}\int dq \cdot *dq$ on Q, and $\Theta(t)$ is the theta-function

$$\Theta(t) = \sum_{n \in \mathbb{Z}} e^{-n^2/2t}.$$

In fact $k = 1$. The Gaussian integral seems at first to be $(\det d_Q^* d_Q)^{-\frac{1}{2}}$, where d_Q is the isomorphism

$$d : Q \to \Omega_0^2(\Sigma)$$

and $\Omega_0^2(\Sigma)$ is the forms with integral zero. But, as for a finite group G, we weight each bundle with the reciprocal of the volume of its group of automorphisms. In the present situation this volume is interpreted as

$$\mathrm{vol}(\mathbb{T})(\det d_0^* d_0)^{-\frac{1}{2}},$$

where d_0 is the isomorphism given by

$$d : \Omega^0(\Sigma)/\mathbb{R} \to P,$$

for this is the ratio of the volume of the group $\mathrm{Map}(\Sigma; \mathbb{T})$ to the orbit P. The determinants of d_0 and d_Q cancel each other, for the two operators are adjoints.

Now let us turn to the non-abelian case, which was first solved by Migdal. The integral is no longer Gaussian, so the only real way to approach it is by lattice approximations. This means that we triangulate Σ, and replace connections by 1-cochains, i.e. maps $\alpha : \Sigma_1 \to G$, where Σ_1 is the set of 1-simplexes of Σ. The curvature of α is the map $h_\alpha : \Sigma_2 \to G$, where Σ_2 is the set of 2-simplexes, whose value on a 2-simplex $\sigma = (P_0 P_1 P_2)$ is the *holonomy*

$$h_\alpha(\sigma) = \alpha(P_0 P_1)\alpha(P_1 P_2)\alpha(P_2 P_0)$$

of α around the boundary of σ. Instead of the Yang-Mills expression $e^{-S(A)}$ we consider

$$e^{-S(\alpha)} = \prod_{\sigma \in \Sigma_2} \varepsilon_{t(\sigma)}(h_\alpha(\sigma)), \tag{1.5.3}$$

where $\varepsilon_t : G \to \mathbb{C}$ is the heat-kernel on G, and $t(\sigma)$ is the area of σ. This is reasonable, for if $h_\alpha(\sigma)$ is close to the identity, as we expect for a fine triangulation, then we can write

$$h_\alpha(\sigma) = 1 + t(\sigma)F_\alpha(\sigma) + \cdots ,$$

with $F_\alpha(\sigma) \in \mathfrak{g}$, and

(1.5.4)
$$\prod \varepsilon_{t(\sigma)}(h_\alpha(\sigma)) = \exp\left\{-\tfrac{1}{2}\sum_\sigma ||F_\alpha(\sigma)||^2 t(\sigma)\right\}\prod_\sigma (2\pi t(\sigma))^{-\frac{1}{2}}$$

to leading order. The exponent in the first factor on the right is the natural Riemann-integral approximation to the Yang-Mills action. Migdal's striking observation is

Proposition 1.5.5 *The integral*

$$\int e^{-S(\alpha)} d\alpha$$

over all $\alpha : \Sigma_1 \to G$ *is independent of the triangulation of* Σ*, and is equal to the invariant* $\Psi_\Sigma^{(t)}$ *of §1.4.*

Proof. Given a triangulation of Σ, let Σ' be the surface (with a free boundary) obtained from σ by removing a small disc of negligeable area around each vertex of the triangulation. If G were a finite group, and we were using the 'toy' field theory of §1.2, in the open-string version of §1.3 which counts bundles trivialized on the free boundary, then we should have (cf. (1.2.2))

$$\Psi_\Sigma \;=\; |G|^V \Psi_{\Sigma'},$$

where V is the number of vertices of the triangulation. In the corresponding formula for the theory of §1.4 we take the volume of G to be 1, and we have simply $\Psi_\Sigma^{(t)} = \Psi_{\Sigma'}^{(t)}$.

Now let us cut Σ' along all of its 1-simplexes, so that it becomes $\Sigma^\#$, which consists of detached 2-simplexes with nibbled corners. The non-free boundary of $\Sigma^\#$ is $I^{\amalg(2E)}$, where E is the number of 1-simplexes, and the element $\Psi_{\Sigma^\#}^{(t)} \in \mathcal{F}_G^{\otimes 2E}$ is precisely the integrand $e^{-S(\alpha)}$ of (1.5.5), regarded as a function on G^{2E} in the obvious way.

We can reconstruct Σ' from the cobordism $\Sigma^\# : \varnothing \rightsquigarrow I^{\amalg(2E)}$ by composing it with

$$R : I^{\amalg(2E)} \rightsquigarrow \varnothing,$$

where R is the disjoint union of E rectangles R_τ, each of negligeable area, where $R_\tau : I \amalg I \rightsquigarrow \varnothing$ sews together the two edges τ', τ'' into which the 1-simplex τ has been cut. Then $\Psi_R^{(0)}$ is a product of δ-functions

$$\prod_\tau \delta(\alpha(\tau'), \alpha(\tau'')),$$

and

$$\begin{aligned} \Psi_{\Sigma'}^{(t)} &= \langle \Psi_{\Sigma^\#}^{(t)},\ \Psi_{\Sigma_R}^{(0)} \rangle \\ &= \int e^{-S(\alpha)}\, d\alpha\ , \end{aligned}$$

as we want.

∎

Bibliography

[DW] Dijkgraaf and Witten

[G]

[M] Moser