Topology in Physics 2018 - lecture 8

Marcel Vonk

April 3, 2018

8.1 Chern-Simons theory

8.1.1 The Chern-Simons action

In previous lectures, and in particular in the exercises, we have already ecountered the Chern-Simons form several times. In particular, in exercise 4.3, we have seen that one can write

$$\text{Tr } (F \wedge F) = d\text{Tr } (A \wedge dA + \frac{2}{3} A \wedge A \wedge A) \tag{8.1}$$

where A is a local connection one-form and F is its two-form field strength. In particular, if A is defined globally, and if we have some 4-manifold M with boundary $\partial M = \Sigma$, Stokes' theorem implies that we can write

$$\int_M \text{Tr } (F \wedge F) = \int_\Sigma \text{Tr } (A \wedge dA + \frac{2}{3} A \wedge A \wedge A). \tag{8.2}$$

We would like to think of the right hand side of this equation as the action for some quantum field theory defined on Σ itself – or more generally, for *any* three-manifold Σ, regardless of whether it can be viewed as the boundary of some four-manifold M. (In particular, we will later look at situations where Σ has a boundary itself.)

However, the interpretation of the right hand side of the above equation as an action is not immediately straightforward, as by now we have learned that the 1-form A is not always *globally* defined, but that it is a *local* description of a globally defined connection ∇ on a fibre bundle $E \to \Sigma$. In particular, only on topologically trivial subsets $U \subset \Sigma$ can we write $\nabla = d + A$. Thus, we can really only define

$$S_{CS,U} = \int_U \text{Tr } (A \wedge dA + \frac{2}{3} A \wedge A \wedge A) \tag{8.3}$$

for such patches, and we have to convince ourselves that these partial actions can be consistently "glued together" into a well-defined global action $S_{CS,\Sigma}$.

Fortunately, we already have all the tools at our disposal to see that this can be done. When Σ can actually be seen as a boundary of a 4-manifold, $\Sigma = \partial M$, this is obvious from

(8.2), as we can simply *define* the right hand side (also for topologically nontrivial Σ) to be equal to the left had side, which *is* well-defined globally.

More generally, even if we cannot view Σ as a boundary, we have seen in exercise 7.3 that there is another interpretation of the Chern-Simons form, coming from the concept of a *transgression form* between two connections ∇ and ∇' defined on a bundle $\tilde{E}$ over Σ itself. Recall that to construct this transgression form, we first define a connection ∇^{aff} on $[0,1] \times \Sigma$ as

$$\nabla^{\text{aff}} = d_t + t\nabla' + (1-t)\nabla, \tag{8.4}$$

where t parameterizes $[0,1]$. The above definition essentially "interpolates" between the connections ∇ at $t = 0$ and ∇' at $t = 1$. (The d_t part simply states that parallel transport in the new t-direction is trivial.) Transgression forms are now obtained by integrating invariant polynomials $P(F)$ for the field strength (curvature) of ∇^{aff} over the t-direction:

$$L(\nabla, \nabla') = \int_0^1 P(F(\nabla^{\text{aff}})) \tag{8.5}$$

In particular, for our favorite invariant polynomial $P(F) = \text{Tr}\ (F \wedge F)$, we have seen in exercise 7.3 that for a trivial bundle (where we can always define a connection through a global one-form A), one can write the transgression between the connections $\nabla = d$ and $\nabla' = d + A$ as

$$L(d, d+A) = \text{Tr}\ (A \wedge dA + \frac{2}{3} A \wedge A \wedge A). \tag{8.6}$$

Of course, if Σ is topologically nontrivial then there are nontrivial bundles $\tilde{E} \to \Sigma$ for which the connection d cannot be defined, so the above construction does not quite give us a definition of the Chern-Simons action in those cases yet. However, for such a bundle we can simply pick a "base" connection ∇_0 once and for all, and then define the Chern-Simons action to be

$$S_{CS}(\nabla) = \int_\Sigma L(\nabla_0, \nabla) \tag{8.7}$$

In the case where $\tilde{E}$ *is* trivial and $\nabla_0 = d$, this clearly reduces to our previous definition.

Thus, we see that (after picking a "base connection" if necessary), we can define a global Chern-Simons action that assigns a number to any other connection ∇ on our bundle $\tilde{E}$. Locally, after picking a connection one-form, this action can always be written in the standard Chern-Simons form (8.2). Of course, physicists are often lazy, and simply write

$$S_{CS} = \int_\Sigma \text{Tr}\ (A \wedge dA + \frac{2}{3} A \wedge A \wedge A). \tag{8.8}$$

also in more general cases. There is nothing wrong with this, as long as one remembers that, because of the issues about the global definition of Chern-Simons theory that were discussed above, this expression needs to be evaluated patch by patch, where each patch might require the use of a different 1-form A.

8.1.2 The level of Chern-Simons theory

Let us for a moment go back to the situation where $\Sigma = \partial M$ and have another look at the definition of Chern-Simons theory in that case:

$$S_{CS}(\nabla_{(\Sigma)}) = \int_M \text{Tr } (F \wedge F). \tag{8.9}$$

One may now wonder: is this definition as sound as it seems? If we *start* from a connection $\nabla_{(M)}$ that is defined over all of M and such that

$$\nabla_{(M)}\Big|_\Sigma = \nabla_{(\Sigma)}, \tag{8.10}$$

then everything is of course well-defined, but if we start from a connection $\nabla_{(\Sigma)}$ that is only defined on a bundle over Σ, we may wonder what the right hand side of (8.9) actually means. In particular:

1. Can $\nabla_{(\Sigma)}$ always be extended to a connection $\nabla_{(M)}$ on some bundle over M?
2. Is such an extension $\nabla_{(M)}$ unique?

We will not go into the first question here, but simply state that $\nabla_{(\Sigma)}$ can indeed always be extended to a connection over M. However, the answer to the second question is in general *no*: the extension over M is not unique! This seems to lead to a problem, as this renders the Chern-Simons action ill-defined after all. However, what we learned before about characteristic classes will come to our rescue.

To see this, assume that we have two different extensions $\nabla_{(M)}$ and $\nabla'_{(M)}$ of $\nabla_{(\Sigma)}$. We now use the following trick: we glue together two copies of M (one with its orientation reversed, which we will denote by $-M$) along the boundary Σ – exactly like how one would glue two hemispheres into a sphere by glueing along the equator. Clearly, the resulting manifold $Y = (M \cup -M)/\sim$ ($\sim$ is the equivalence relation that identifies corresponding points on the two boundaries) no longer has a boundary: $\partial Y = 0$. Moreover, we can use this construction to glue the bundles over M and $-M$ into a bundle over Y as well, and because of the orientation reversal, we can also equip the bundle over M with the connection $\nabla_{(M)}$ and the bundle over $-M$ with the orientation reversal of the connection $\nabla'_{(M)}$, in such a way that the result is a smooth connection $\nabla_{(Y)}$. Now, clearly,

$$\int_Y \text{Tr } (F \wedge F) = \int_M \text{Tr } (F \wedge F) + \int'_{-M} \text{Tr } (F \wedge F) \tag{8.11}$$

where the first term on the right hand side is evaluated using $\nabla_{(M)}$ and the second term using $\nabla'_{(M)}$ – indicated by the prime on the integration symbol. Returning to our original orientation in the second term, we find

$$\int_Y \text{Tr } (F \wedge F) = \int_M \text{Tr } (F \wedge F) - \int'_M \text{Tr } (F \wedge F), \tag{8.12}$$

so the intergral over Y tells us exactly what the difference between our two choices for the Chern-Simons action is! But now we can use the facts that Tr $(F \wedge F)$ is a characteristic class and that Y has no boundary to conclude that the left hand side of the above equation must be integral up to a factor[1]:

$$\frac{1}{4\pi}\int_M \text{Tr } (F \wedge F) - \frac{1}{4\pi}\int_M' \text{Tr } (F \wedge F) = 2\pi n, \qquad n \in \mathbb{Z}. \tag{8.13}$$

Thus, the definition of the Chern-Simons action indeed depends on the way we extend $\nabla_{(\Sigma)}$ over M, but we see that the dependence is mild. In particular, if we properly normalize the Chern-Simons action as

$$S_{CS,k}(\nabla_{(\Sigma)}) = \frac{k}{4\pi}\int_M \text{Tr } (F \wedge F), \tag{8.14}$$

with k an arbitrary nonzero integer, then in the path integral, the factor $e^{iS_{CS,k}}$ is *independent* of the extension. As a result, the partition function and all correlation functions for this theory are perfectly well defined and independent of how we extend $\nabla_{(\Sigma)}$. Note that (by rescaling A such that the prefactor only appears in front of the $A\wedge A\wedge A$ term), k plays the role of a *coupling constant* of the theory. Thus, we find that the quantum version of Chern-Simons theory is only well-defined if the coupling constant is integral. This integral coupling constant is called the *level* of the Chern-Simons theory.

8.1.3 Symmetries of the theory

Chern-Simons theory is one of the most studied quantum field theories is theoretical physics. One reason for this is that the theory has an enormous symmetry group. Let us again focus on the case where $\Sigma = \partial M$ is the boundary of a 4-manifold. First of all, the theory is now gauge invariant, just like Yang-Mills theory, for the simple reason that by extending it over M, we can write the action exclusively in terms of the field strength F, which itself is invariant under gauge transformations.

However, there is an important difference between Maxwell theory and Yang-Mills theory. Recall that for Maxwell theory, the action is

$$S_{YM} \sim \int_M \text{Tr } (F \wedge \star F). \tag{8.15}$$

The crucial observation is that in this definition, the Hodge star appears, which depends on a choice of metric $g_{\mu\nu}(x)$ on M. However, in the Chern-Simons action, no metric appears whatsoever, as can be seen both from the local action on patches of Σ and from its global extension over M,

$$S_{CS} \sim \int_M \text{Tr } (F \wedge F). \tag{8.16}$$

[1]Here, we use conventions where F is rescaled by a factor and the wedge product has an additional factor of 2 in it, so that our final expression agrees with the normalizations generally used in physics

Therefore, S_{CS} does not change if we choose a different metric on M or Σ, and since it can be written as an integral of a three-form over a three-manifold in the local descripton, or as the integral of a four-form over a four-manifold in the global description, it also does not depend on a choice of *coordinates* on M.

Thus, we have found that the partition function of Chern-Simons theory only depends on the *topology* of M: $Z_{CS}(M)$ is a topological invariant! This observation has many interesting applications both in physics and in mathematics, which we may come to later in this course if time permits. For example, one may choose the manifold M to be the complement of a one-dimensional knot embedded in (say) S^3, in which case $Z_{CS}(M)$ provides us with a knot invariant.

The above ideas can be made even more interesting by not just studying the partition function $Z_{CS}(M)$, but also some of its correlation functions. However, to write down a correlation function

$$\langle \mathcal{O} \rangle \equiv \int DA \mathcal{O} e^{iS_{CS,k}[A]} \tag{8.17}$$

which has the same nice symmetries as the partition function $\langle 1 \rangle$, we need to introduce very special operators $\mathcal{O}$ that also do not depend on a choice of metric or of coordinates. It is to this choice of operators that we turn next.

8.1.4 Holonomy and Wilson loops

Assume we have some vector bundle $E \to M$ with canonical fiber F, and a topologically trivial subset $U \subset M$ so that we can parameterize points in E by coordinates $(x^\mu, v^i) \in U \times F$. The question we are interested in is the following: if we start from some point (x_0^μ, v_0^μ), where do we end up if we parallel transport this point using a connection $\nabla = d + A$?

Of course, the answer depends on the path along which we parallel transport, so let us assume that we have chosen a path $x^\mu(t)$ with $x^\mu(0) = x_0^\mu$. Moreover, let us begin by parallel transporting an infinitesimal amount, up to $x^\mu(\delta t) \equiv x_0^\mu + \delta x^\mu$. Parallel transport means that $v^i(t)$ must be covariantly constant:

$$\frac{dv^i}{dt} + \frac{dx^\mu}{dt}(A_\mu)^i{}_j v^j = 0. \tag{8.18}$$

Here, we have pulled back the connection one-form A to the real line parameterized by t, explaining the factor $\frac{dx^\mu}{dt}$. In this equation, A_μ of course depends on x and therefore on t, but if we are interested in transporting an infinitesimal amount only, we can consider A_μ to be constant. The same is true for $\frac{dx^\mu}{dt}$, and so we can integrate the above equation to find

$$v^i(\delta t) = \exp\left(-\frac{dx^\mu}{dt}(A_\mu)^i{}_j \delta t\right) v_0^j + \mathcal{O}(\delta t) \tag{8.19}$$

which of course can be written as

$$v^i(\delta t) = \exp\left(-(A_\mu)^i{}_j \delta x\right) v_0^j + \mathcal{O}(\delta t). \tag{8.20}$$

The conclusion of this computation is that we should think of $\exp\left(-(A_\mu)^i{}_j \delta x\right)$ as the Lie group element that implements the action of an infinitesimal parallel transportation along δx.

What about parallel transport over *finite* distances? Of course, we can then chop up such a finite path into infinitesimal pieces: if we want to transport from, say, $t = 0$ to $t = 1$ we simply define

$$x_n = x(n/N) \tag{8.21}$$

for some large integer N, and find that

$$\begin{aligned} v^i(1) &= \exp\left(-(A_\mu)^i{}_j(x_0)\delta x_0\right)\exp\left(-(A_\mu)^i{}_j(x_1)\delta x_1\right)\cdots \\ &\quad \cdots\exp\left(-(A_\mu)^i{}_j(x_{n-1})\delta x_{n-1}\right)v^i(0) + \mathcal{O}(1/N) \end{aligned} \tag{8.22}$$

where we state without proof that indeed, the corrections to the result vanish in the "smooth limit" $N \to \infty$. (A less hand-waving proof of our final result can be found e.g. in section 10.2 of Nakahara.)

In the Abelian case, where A_μ is a number instead of a matrix, one can now easily multiply all the exponentials; in the limit $N \to \infty$ the result is that

$$v_{\text{ab}}(1) = \exp\left(-\int A_\mu(x)dx^\mu\right)v_{\text{ab}}(0). \tag{8.23}$$

In the nonabelian case, things are slightly more complicated, since (by the Baker-Campbell-Hausdorff formula) we cannot simply replace the product of exponentials by an exponential of a sum. In fact, even though the $N \to \infty$ limit of (8.22) is perfectly well-defined, there is not really a shorter expression for it in terms of an integral. The solution is simple: we just introduce a shorter notation to describe the limit of the "path ordered product of exponentials": in analogy with the abelian case, one writes

$$v^i(1) = \mathcal{P}\exp\left(-\int (A_\mu)^i{}_j dx^\mu\right)v^j(0) \tag{8.24}$$

as a shorthand for the $N \to \infty$ limit of (8.22). The group element

$$G_\gamma = \mathcal{P}\exp\left(-\int (A_\mu)^i{}_j dx^\mu\right), \tag{8.25}$$

which of course depends on the path γ parameterized by $x^\mu(t)$ (and in particular on its starting and ending points) is called the *holonomy* of the fiber upon parallel transport along γ.

Now remember our goal: we are interested in constructing observables for Chern-Simons theory that have as many of the symmetries as the theory itself. One thing we can do

to achieve this is to consider *loops* instead of just any paths: if $x^\mu(1) = x^\mu(0)$, we are integrating over a path without boundary, and therefore (by Stokes' theorem) the resulting holonomy does not change if we do a gauge transformation. (In fact, in the nonabelian case one would still have to prove that this is true for an integral in a path ordered exponential, as it is for an ordinary integral, but here we will omit the details of that proof.)

Finally, the holonomy is still a matrix, and of course the form of this matrix will depend on a choice of (say orthonormal) basis for the fiber F. To really compute a basis-invariant number, one can e.g. take the trace of this matrix, and this is what physicists normally do. Collecting all of those ingredients, we arrive at what is called the *Wilson loop* observable in Chern-Simons theory:

$$W_\gamma = \mathrm{Tr}\ \mathcal{P} \exp\left(- \int_\gamma (A_\mu)^i{}_j dx^\mu \right). \tag{8.26}$$

A note on terminology: as we mentioned, the most interesting case is to consider loops with $\gamma(0) = \gamma(1)$. The more general case with different end points is called a Wilson *line*. Wilson lines, and in particular Wilson loops, are the most studied and most interesting observables in Chern-Simons theory. For example, Edward Witten has shown that many interesting knot invariants can be obtained by computing Wilson loop expectation values along the knot in an $SU(N)$ Chern-Simons theory on S^3. This is a great example of the fact that physics can actually help progress in mathematics, and not just the other way around. If time permits, we may come back to this topic later in the course.

8.2 Fermions and path integrals

8.2.1 Motivation: fermions, spin and statistics

The goal of this course is to highlight the many interesting relations between topology and physics. We have already seen several instances of such relations, but to arrive at the most interesting examples, one ingredient is still missing from our story. So far, all the quantum fields we have studied have been *bosonic* fields. However, in physics one also encounters *fermionic* fields, and it turns out that including such fermionic fields into our considerations is a very rich source of topological applications.

The first question we have to face, therefore, is: how do we describe fermions mathematically? Fermions differ from bosons in two important respects: they have a different *spin* and they obey different *statistics*. In fact, it can be shown that these two observations are closely related – the so-called "spin-statistics theorem" – but for now we will view these two statements as separate properties.

Let us make some brief remarks about spin, even though for now we will be mostly be interested in the "statistics" property of fermions. Contrary to bosonic fields, fermionic fields have *half-integer* spin. All fields in nature transform in a specific way if we apply a

spatial rotation: there are scalar fields that do not change at all if we rotate out physical setup (think of temperature as an example), but there are also vector-like fields such as the electric field whose direction changes along with space-time itself if we rotate our setup. In more mathematical terms: (quantum) fields must be valued in a given *representation* of the rotation group.

At first sight, one might guess that this rotation group is $SO(3)$, but $SO(3)$ in fact has the same Lie algebra as its double cover $SU(2)$, and it turns out that in practice it is the latter group that plays a role. Of course, the scalars and vectors we mentioned still form representations of $SU(2)$, but $SU(2)$ has additional representations like its fundamental representation in terms of 2×2 matrices. It turns out that for example the electron field transforms exactly in this representation under rotations: it is a two-component field and applying a rotation to it can be described mathematically by applying the relevant $SU(2)$ matrix. This in particular means that if we rotate an electron by 360 degrees, we do *not* get back its original state: its quantum mechanical wave function obtains a minus sign (which can be measured in certain interference experiments), and only if we rotate by a full 720 degrees do we get back the exact original electron wave function.

More generally, representations of $SU(2)$ can be classified by a nonnegative number called their *spin* s, which is related to the dimension d of the represenation as $d = 2s + 1$, and this spin can be integer or half-integer. (Terminology: by "half-integer", physicists usually denote half of an *odd* integer, so a number of the form $n + 1/2$ with $n \in \mathbb{Z}$.) Fields that transform as integer spin representations are called *bosons*; fields that transform as half-integer spin representations are called *fermions*.

The relation to the "statistics" of fermions arises as follows. Under some reasonable assumptions, it can be shown that the quantum wave function of *two* identical fermionic particles also obtaines a minus sign if one exchanges the properties of the two particles – their postion, momentum, and any other properties they may have:

$$\Psi(x_1, p_1, \ldots, x_2, p_2, \ldots) = -\Psi(x_2, p_2, \ldots, x_1, p_1, \ldots) \tag{8.27}$$

This in particular means that the wave function must vanish if we try to give two fermionic particles exactly the same properties:

$$\Psi(x_1, p_1, \ldots, x_1, p_1, \ldots) = 0. \tag{8.28}$$

Now it is a well-known fact in quantum field theory that the wave function where a particle is in a given state, can be obtained from the ground state (vacuum) wave function, by applying an operator known as the *creation operator*, schematically:

$$\Psi_{\text{1-particle}} = \alpha^\dagger \Psi_{\text{vacuum}}, \qquad \Psi_{\text{2-particle}} = \beta^\dagger \Psi_{\text{1-particle}}, \qquad \ldots \tag{8.29}$$

The dagger here is conventional: it turns out that these "creation operators" are the hermitian conjugates of operators that *annihilate* a particle, and those operators are usually denoted without the dagger.

Now comes the punch line: since creating two of the *same* particles must lead to a vanishing wave function, we must have that these creation operators square to zero:

$$(\alpha^\dagger)^2 = 0. \tag{8.30}$$

In fact, it turns out this observation has an enormous generalization: since creation operators have a linear relationship to quantum fields (they can be viewed as the Fourier modes of those quantum fields), the above observation generalizes to the entire fermionic quantum field: such a field $\psi(x)$ must satisfy

$$\psi(x)^2 = 0. \tag{8.31}$$

This means that clearly, fermionic quantum fields cannot be ordinary functions, or even sections of ordinary bundles: those objects only square to zero. So the main question we have to address now is: how do we describe "function-like" objects that square to zero?

8.2.2 Grassmann numbers

The question that we ended the previous paragraph with can be compared to the question one asks when introducing complex numbers: how does one describe an object that squares to -1? The answer there is: we simply *postulate* a new object "i", and require that it squares to -1, and then see if adding this object to the set of real numbers leads to any useful structure. (Wich of course it does!)

So let us do the same here: we introduce a new object θ and postulate that it squares to 0: $\theta^2 = 0$. In fact, let us be slightly more general here, as we will want to introduce many different fermionic objects that we will call θ_i for now. Each of these objects should square to zero and should moreover *anticommute* with the others (compare this to the property that exchanging two fermions gives the wave function a minus sign):

$$\{\theta_i, \theta_j\} \equiv \theta_i\theta_j + \theta_j\theta_i = 0. \tag{8.32}$$

How should we think of the θ_i? We could think of them as "numbers", like i, but in fact it is a bit more natural to think of them as "variables". For example, we will want to introduce operations such as integration and differentiation with respect to θ_i, just like we would for an ordinary, commuting variable x. Of course, one important difference between θ_i and x is that the commuting variable x can actually take on a *value* (say, 7), but there are no "values" that θ_i can take on. Put differently: whenever we want to compute a physical quantity, it should not depend on any θ_i anymore, just like any real physical quantity should not have the number i in it. Since we will mostly introduce our θ_i as objects to be integrated over inside a path integral, this will be more or less automatic in our setup. Unfortunately, the terminology is still that θ_i is usually called a *Grassmann number* in the literature – though the name "Grassmann variable" also appears.

The θ_i can be thought of as generators of a (noncommutative) algebra over the real numbers. That is: we want to be able to add and multiply Grassmann variables, and also

multiply them with ordinary numbers. Since any product that has the same θ_i in it twice vanishes, this means that when we only have a finite number of Grassmann variables (say labeled by $1 \leq i \leq N$), the resulting algebra is also finite dimensional: any element in it can be written as

$$f(\theta) = f_0 + \sum_{i=1}^{N} f_i \theta_i + \sum_{i<j} f_{ij} \theta_i \theta_j + \ldots + f_{123\ldots N} \theta_1 \cdots \theta_N, \tag{8.33}$$

where the coefficients f are ordinary numbers (real or complex, depending on the setting). As the notation indicates, we should think of $f(\theta)$ as a "function" of the Grassman variables; because of anticommutativity, such a function has a finite (and exact) Taylor expansion.

8.2.3 Differentiation and integration

Just like for ordinary variables, one can define differentiation with respect to a certain θ_i in the algebra generated by the Grassmann variables. The only difference here is that we have to be careful with signs if we differentiate: for example, since $\theta_i \theta_j = -\theta_j \theta_i$ $(i \neq j)$, one may wonder whether the derivative of this product with respect to θ_i is θ_j or $-\theta_j$. The usual convention is to define differentiation as *left* differentiation. That is, to take the derivative of any monomial with respect to θ_i, one first anticommutes θ_i to the left and then removes it from the expression. (Notice that θ_i never occurs twice in a monomial, so differentiation is always simply removing the variable and keeping the "constant" factor that multiplies it.) Moreover, the θ_i derivative of a number is zero, just like the x-derivative of a number vanishes. Some examples:

$$\frac{d}{d\theta} 1 = 0, \qquad \frac{d}{d\theta} \theta = 1, \qquad \frac{d}{d\theta_2} (7\theta_1 \theta_2 \theta_3 + 5\theta_4) = -7\theta_1 \theta_3. \tag{8.34}$$

The next question we want to address is: how do we *integrate* over Grassmann variables? First, we have to decide whether we are interested in the analogue of an indefinite integral (the primitive) or of a definite integral like $\int_{-\infty}^{\infty} f(x)dx$. Since the goal of our integral will be to really "integrate over a variable" – that is, to find an expression which no longer dependent on a certain variable, it is the *definite* integral that we are interested in – despite the fact that "boundary values" like $x = \pm\infty$ do not make sense for a Grassmann variable θ. For this reason, we will always write

$$\int f(\theta) d\theta \tag{8.35}$$

without boundaries, but the reader should keep in mind that Grassmann integerals nevertheless are the analogue of definite integrals.

The reason we stress the above remark is that we want to define Grassmann variable integration in such a way that at has three properties that regular definite integrals also have: we want to define an operator I on the Grassmann algebra that satisfies

- $DI = 0$,
- $ID = 0$,
- If $D(A) = 0$ then $I(AB) = AI(B)$.

Here, $D = \frac{d}{d\theta}$ is the differentiation map with respect to a certain Grassmann variable. The first requirement above is simply the fact that I is a definite integral: its result should no longer depend on θ. The second requirement is that the integral of a total derivative vanishes – that is: we assume that Grassmann integration has no "boundary terms at infinity". The final requiremen is the usual property of integration that constant factors can be taken outside the integral.

There is a surprising map I that satisfies all of the above requirements: one can easily check that $I = D$ works. The first two requirements are satisfied because $D^2 = 0$, and the third one is true because constants can be taken outside differentiation. Perhaps even more surprisingly, it turns out that $I = D$ is essentially the only map that works: in exercise 3 you will show that any map I satisfying the above three requirements must equal $I = cD$ for some constant (ordinary) number c. Clearly, $c = 0$ does not give the map we are looking for, and by rescaling θ one can always set any other c to $c = 1$. Therefore, in what follows we will always work with $I = D$. For Grassmann variables, integration and differentiation are the same!

8.2.4 Fermionic path integrals

Now that we have defined integration over single Grassmann variables, fermionic path integrals are not difficult to understand anymore: if we want to do a path integral over a fermionic field $\psi(x)$, we simply should think of x as a label (just like i was a label in θ_i) and do the definite integral over each $\psi(x)$ separately. However, for future purposes, let us note some differences between fermionic (path) integration and the bosonic case.

1. In (path) integrals, it is often useful to change variables. Let us take the finite case, and change variables from θ_i to $\theta'_i = a_i{}^j\theta_j$. Since integration equals differentiation, we know from
$$\frac{\partial}{\partial\theta_1}\cdots\frac{\partial}{\partial\theta_N}\ f(\theta) = \det(a)\frac{\partial}{\partial\theta'_1}\cdots\frac{\partial}{\partial\theta'_N}\ f(a^{-1}\theta') \tag{8.36}$$
that also
$$\int d\theta_1\cdots d\theta_N\ f(\theta) = \det(a)\int d\theta'_1\cdots d\theta'_N\ f(a^{-1}\theta') \tag{8.37}$$
Thus, the integration measure changes in the oposite way from what we are used to in the bosonic case: instead of multiplying by the Jacobian, one needs to divide by it.

2. Complex conjugation for Grassmann variables is introduced "by hand": for every Grassman variable θ_i, one can simply introduce an additional one labeled θ_i^*, and

require that the operation of complex conjugation acts on these objects by exchanging θ_i and θ_i^*. In fact, in monomials that contain several θ_i, one needs to also exchange the order:

$$(\theta_i \theta_j)^* = \theta_j^* \theta_i \tag{8.38}$$

The reason for this convention is that then $\theta_i \theta_i^*$ equals its own complex conjugate, and therefore can be considered to be a "real" quantity, just like in the bosonic case.

3. Using the previous two points, one can now easily compute the Gaussian integral

$$G = \int d\theta_1^* d\theta_1 \cdots d\theta_N^* d\theta_N \ \exp\left\{ -\sum_{i,j} \theta_i^* M_{ij} \theta_j \right\} \tag{8.39}$$

for example by changing variables to $\theta_i' = \sum_{j=1}^N M_{ij}\theta_j$. Taking care of the signs, one finds that

$$G = \det M \tag{8.40}$$

which once again is the opposite result from the bosonic case, where (up to constant factors) one would find $1/\det M$.

4. Taking the limit $N \to \infty$ to define actual path integrals is not more complicated than in the bosonic case. To derive Euler-Lagrange equations from fermionic path integrals, one needs to define functional derivatives with respect to $\psi(x)$. This is in fact slightly easier than in the bosonic case, as one can simply introduce a new Grassmann variable ϵ and define

$$\frac{\delta F[\psi(t)]}{\delta \psi(x)} \equiv \frac{1}{\epsilon} \left(F[\psi(t) + \epsilon \delta(t-s)] - F[\psi(t)] \right) \tag{8.41}$$

where $\delta(t-s)$ is the ordinary bosonic delta function. Note that no limir is required; the notation "$1/\epsilon$" simply means that the result in brackets is proportional to ϵ, and we remove that auxiliary Grassmann number. (In fact, for generic objects that are not linear in ϵ, division by it is not well-defined.)

If you have never worked with Grassmann variables before, some practice is definitely recommended. You might first want to try exercise 3 in the exercise sheet for this lecture. After that, if you also want to get acquainted with fermionic path integrals, there is a worked out example in section 1.5.10 of Nakahara that computes the partition function of the fermionic harmonic oscillator as a fermionic path integral; it can be very useful to go through this computation step by step.