

Grassmann numbers and the fermion path integral

Resources: Zee pp. 125-127; Schwartz pp. 269-272; Peskin pp. 298 - 301.

- **Grassmann numbers** θ, η are anticommuting numbers, $\theta\eta = -\eta\theta$. We saw in the previous problem set that they’re very important to describing fermions, as **fermion fields are Grassmann-valued!** Generally we’ll consider n Grassmann variables $\{\eta_i\}_{i=1}^n$, which then satisfy the anticommutative algebra:

$$\eta_i\eta_j = -\eta_j\eta_i. \quad (1)$$

Note that this implies a single Grassmann variable *must square to zero*:

$$\eta_i^2 = 0. \quad (\text{no sum on } i) \quad (2)$$

We can add and scalar multiply Grassmann variables like normal, so they form a vector space over \mathbb{C} .

- Functions of one Grassmann variable: These are easy to describe, since Grassmann variables square to zero. We’ll assume that any function f of a Grassmann variable is analytic, in the sense that it can be expanded in a Taylor series. This shows us that the most general form of a function $f(\eta)$ is in fact quite restrictive, and the expansion for f can be truncated at second order in η :

$$f(\eta) = a + b\eta + c\eta^2 + \dots = a + b\eta. \quad (3)$$

Thus, **any function of one Grassmann variable is completely specified by two complex scalar coefficients a and b** .

- Functions of multiple Grassmann variables. When we have multiple Grassmann variables, things get more complicated. The general structure is still there, but it’s not as simple as just specifying two scalars to specify the function. That’s because now we have to worry about cross terms. For example, for three variables $\{\eta_1, \eta_2, \eta_3\}$, an arbitrary function is specified by $8 = 2^3$ parameters:

$$f(\eta_1, \eta_2, \eta_3) = a + b_1\eta_1 + b_2\eta_2 + b_3\eta_3 + c_1\eta_2\eta_3 + c_2\eta_3\eta_1 + c_3\eta_1\eta_2 + d\eta_1\eta_2\eta_3. \quad (4)$$

This generalizes to N Grassmann variables. Note $f(\eta_1, \eta_2, \eta_3)$ is an arbitrary element of the Grassmann algebra on three variables: by the “Grassmann algebra”, we mean the space of all possible products of $\{\eta_1, \eta_2, \eta_3\}$.

- The algebra generated by N independent Grassmann variables $\{\eta_1, \dots, \eta_N\}$ is $\mathcal{G}_N \equiv \mathbb{C}[\eta_1, \dots, \eta_N]$. A general element $\theta \in \mathcal{G}_N$ can be written as:

$$\theta = \sum_{k=0}^N \sum_{i_1 > \dots > i_k} a_{i_1 \dots i_k} \eta_{i_1} \dots \eta_{i_k}. \quad (5)$$

There are $\binom{N}{k}$ linearly independent degree k monomials in \mathcal{G}_N , and so:

$$\dim \mathcal{G}_N = \sum_{k=0}^N \binom{N}{k} = 2^N. \quad (6)$$

Thus if we write out a function $f(\eta_1, \dots, \eta_N)$ of N Grassmann variables, since $f(\eta_1, \dots, \eta_N) \in \mathcal{G}_N$, it may be specified by 2^N coefficients (compare this to the 3D case).

The exterior algebra

The math that Grassmann developed is also called **exterior calculus**. Let’s look at the mathematical construction behind this! Note that you really don’t need to know any of the stuff in this box, I just think it’s a fun opportunity to inject some math into these notes. Assume we have a vector space V . The exterior product is a construction that we put onto a tensor product of a vector space to create an anticommuting product. Let’s denote $T^k V$ as the space of k -tensors of V , and TV the **tensor algebra** of V , which is the graded algebra generated by all the $T^k V$ spaces:

$$T^k V = \underbrace{V \otimes V \otimes \dots \otimes V}_{k \text{ times}} = V^{\otimes k} \quad TV \equiv \{0\} \oplus V \oplus (V \otimes V) + \dots = \bigoplus_{k=0}^{\infty} T^k V. \quad (7)$$

The idea is that $T^k V$ allows us to formally study k -tensors of V , and TV lets us study any linear combination of vectors and their tensor products in V . To construct the **exterior algebra** ΛV , we mod $T^k V$ out by the space generated by $x \otimes x$ and sum the corresponding spaces:

$$\Lambda^k V \equiv T^k V / \underbrace{\langle \{x \otimes x : x \in V\} \rangle}_{\equiv I} \quad \Lambda V \equiv \sum_{k=0}^{\infty} \Lambda^k V \quad (8)$$

where $\langle A \rangle$ denotes the subalgebra generated by A (i.e. all elements which are linear combinations or products of the elements in A). As for notation, when we pass $v \otimes w$ to the quotient space $\Lambda^k V$, we denote it with a wedge product \wedge :

$$[v] \wedge [w] \equiv [v \wedge w] \quad (9)$$

where $[\cdot]$ denotes equivalence class mod I . The nice part about this is that \wedge is a *well-defined, antisymmetric product*, in the sense that:

$$[v] \wedge [w] = -[w] \wedge [v]. \quad (10)$$

To see this for $x, y \in V$, note $(x + y) \otimes (x + y), x \otimes x, y \otimes y \in I$ by definition (these are products of vectors with themselves). Passing $(x + y) \otimes (x + y) = x \otimes x + x \otimes y + y \otimes x + y \otimes y$ to ΛV , we have:

$$\begin{aligned} 0 &= [x + y] \wedge [x + y] = [(x + y) \otimes (x + y)] = \cancel{[x] \wedge [x]} + [x] \wedge [y] + [y] \wedge [x] + \cancel{[y] \wedge [y]} \\ \implies [x] \wedge [y] &= -[y] \wedge [x]. \end{aligned} \quad (11)$$

This generalizes to higher order wedge products, and thus we have constructed an algebra with an antisymmetric product, starting with V ! If you’ve seen **differential forms**, you may be familiar with this construction already. In the case of differential forms, V here is the cotangent space $T_p^* M$, and we look at sections $\Gamma(\cdot)$ (continuous maps) on the cotangent bundle $\Lambda^k(T^* M) \equiv \sqcup_p \Lambda^k(T_p^* M)$:

$$\Omega^k(M) \equiv \Gamma(\Lambda^k(T^* M)) \quad \Omega^* = \bigoplus_{k=0}^{\infty} \Omega^k(M). \quad (12)$$

- The **Grassmann integral** is defined as the unique linear functional $\int d\eta : \mathcal{G}_1 \rightarrow \mathbb{C}$ which is shift-invariant and normalized. By shift-invariant and normalized, we mean that:

$$\int d\eta f(\eta) = \int d\eta f(\eta + \theta) \quad \int d\eta \eta = 1 \quad (13)$$

These conditions imply that the Grassmann integral of an arbitrary function $a + b\eta$ is:

$$\int d\eta (a + b\eta) = b. \quad (14)$$

Note the integral equals the derivative of $a + b\eta$, where we differentiate with respect to η the same way we would a normal c -number. For a function $f(\eta_1, \dots, \eta_n) \in \mathcal{G}_n$, the integral $\int d\eta_i$ and $\partial/\partial\eta_i$ are treated as Grassmann numbers, so for example $\int d\eta_1 d\eta_2 \eta_1 \eta_2 = -(\int d\eta_1 \eta_1)(\int d\eta_2 \eta_2)$.

- **Gaussian integrals of Grassmann numbers:** A Gaussian integral of $2N$ Grassmann variables (η_1, \dots, η_N) and $(\eta_1^*, \dots, \eta_N^*)^1$ can be Taylor expanded and the N th order term integrated to yield:

$$\int d^n \eta^* d^n \eta e^{-\eta^* A \eta} = \int d\eta_1^* d\eta_1 \dots d\eta_N^* d\eta_N e^{-\bar{\eta}_i A_{ij} \eta_j} = \det(A) \quad (15)$$

Note the main difference between the c -number path integral and the Grassmann integral Eq. (15) is the power the determinant is raised to. The generating functional for Grassmann variables is:

$$\mathcal{Z}(\theta, \theta^*) = \int d^n \eta^* d^n \eta e^{-\eta_i^* A_{ij} \eta_j + \theta^* \eta + \eta^* \theta} = \det(A) e^{\theta_i^* A_{jk}^{-1} \theta_k} \quad (16)$$

- A useful identity for the homework to prove Eq. (15) is:

$$\mathcal{I} = \int \prod_{i=1}^N (d\theta_i^* d\theta_i) \theta_{j_1} \theta_{j_2} \dots \theta_{j_N} \theta_{k_1}^* \theta_{k_2}^* \dots \theta_{k_N}^* = (-1)^{N(N-1)/2} \epsilon_{j_1 j_2 \dots j_N} \epsilon_{k_1 k_2 \dots k_N}. \quad (17)$$

The easiest way to show this is to note a few facts:

1. If any $j_m = j_n$ (and like wise for k_m and k_n), then the integral vanishes, because somewhere you have two of $\theta_{j_m} = \theta_{j_n}$, so after the appropriate amount of anticommutation, they'll hit one another and give you 0.
2. Since anticommuting variables around gives negative signs, we can relate a polynomial to the ϵ tensor as:

$$\theta_{j_1} \dots \theta_{j_N} = \epsilon_{j_1 \dots j_N} \theta_1 \dots \theta_N \quad (18)$$

(this is pretty much the definition of ϵ).

3. Pairs of Grassmann values commute, i.e. if I'm looking at $d\theta_j^* d\theta_j$, I can treat this like a commuting c -number.

¹Note that we haven't defined complex conjugation on Grassmann variables! You can think of η_1^* and η_1 as two independent Grassmann values, where η_1^* "looks" like the complex conjugate of η_1 . However, formally they're just two independent Grassmann variables.

4. We can reorder our products of Grassmann variables as:

$$(\theta_1 \theta_1^*) \dots (\theta_N \theta_N^*) = (-1)^{N(N-1)/2} (\theta_1 \dots \theta_N) (\theta_1^* \dots \theta_N^*). \quad (19)$$

The $(-1)^{N(N-1)/2}$ factor comes from $(-1)^N (-1)^{N-1} \dots (-1) = (-1)^{\sum_{k=1}^N (N-k)} = (-1)^{N(N-1)/2}$, as we incur a $(-1)^{N-1}$ factor when we shuffle the k th factor around.

So, we can write out integral as:

$$\begin{aligned} \mathcal{I} &= \int (d\theta_1^* d\theta_1) \dots (d\theta_N^* d\theta_N) (\theta_{j_1} \theta_{j_2} \dots \theta_{j_N}) (\theta_{k_1}^* \theta_{k_2}^* \dots \theta_{k_N}^*) \\ &= \epsilon_{j_1 \dots j_N} \epsilon_{k_1 \dots k_N} \int (d\theta_N^* d\theta_N) \dots (d\theta_1^* d\theta_1) (\theta_1 \theta_2 \dots \theta_N) (\theta_1^* \theta_2^* \dots \theta_N^*) \\ &= \epsilon_{j_1 \dots j_N} \epsilon_{k_1 \dots k_N} \underbrace{\int (d\theta_N^* d\theta_N) \dots (d\theta_1^* d\theta_1) (\theta_1 \theta_1^*) (\theta_2 \theta_2^*) \dots (\theta_N \theta_N^*)}_1 \underbrace{(-1)^{N-1} (-1)^{N-2} \dots (-1)}_{(-1)^{N(N-1)/2}} \\ &= (-1)^{N(N-1)/2} \epsilon_{j_1 \dots j_N} \epsilon_{k_1 \dots k_N}. \end{aligned} \quad (20)$$

In the third line, each factor of $(-1)^{N-k}$ is incurred from moving θ_k^* from the grouping of conjugate terms to be next to θ_k . Then, in the line after that we just did the integral from inside out, since $\int d\theta \theta = 1$, so every time we see that pair we can remove it from the equation! So, that proves Eq. (17): you might find this useful for Problem (3)a this week!

Chirality in electroweak theory

Now that we've discussed some details about chiral symmetry, we can get into some of the weird parts about the Standard Model. We can't go too far in detail, but I want to discuss some qualitative details about neutrinos and their role in the Standard Model. Specifically, we'll be talking about electroweak theory.

- The electroweak sector of the Standard Model is a **chiral gauge theory** described by the gauge group $SU(2)_L \times U(1)_Y$. The "gauge" theory part of this is not important for now, since we haven't studied gauge symmetry yet: needless to say, all you basically need to know is that the electroweak Lagrangian is invariant under local $SU(2) \times U(1)$ transformations. The important part for us here is the "L" subscript on $SU(2)$, which stands for "left". This means that the $SU(2)_L$ term only acts on left-handed particles, and this is why we call the Standard Model a **chiral theory**: it treats left-handed particles different than right-handed particles.
- **Toy model (symmetry structure)**: To put this in terms that we've seen before, let's consider a theory which is invariant under two global symmetries: a left-handed symmetry $U(1)_L$ is parameterized by a parameter α , and rotates Weyl fields as:

$$\psi_L \longrightarrow e^{i\alpha Q} \psi_L \quad \psi_R \longrightarrow \psi_R \quad (21)$$

and a global "hypercharge" symmetry $U(1)_Y$, which is parameterized by a parameter β and left-and right-handed fermions equally:

$$\psi_L \longrightarrow e^{i\beta Y} \psi_L \quad \psi_R \longrightarrow e^{i\beta Y} \psi_R. \quad (22)$$

We have experience with the $U(1)_Y$ global symmetry (I’ve just labeled it slightly differently), but the chiral rotation of $U(1)_L$ is new. Here Q and Y are charge operators, and let us assign different charges to different particles that we study. Let’s work with a theory of electrons and neutrinos. It’s been observed that there are only left-handed neutrinos (more on this later) and

- What effect does this symmetry structure have on the fields that are in the Standard Model? One immediate consequence from the chiral gauge structure of the SM that’s rather surprising is that **fermionic matter in the Standard Model is not described as Dirac spinors, and instead is described by Weyl spinors**. To see this, let’s consider a Dirac spinor which is charged under $U(1)_L$ with $Q = 1$:

$$\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \quad (23)$$

The problem with the Dirac spinor is that the mass term is not invariant under the $U(1)_L$ symmetry:

$$\bar{\Psi}\Psi = \psi_R^\dagger \psi_L + \psi_L^\dagger \psi_R \rightarrow e^{i\alpha} \psi_R^\dagger \psi_L + e^{-i\alpha} \psi_L^\dagger \psi_R. \quad (24)$$

This isn’t surprising, since the mass term couples together the LH and RH components of the fermion, and the LH and RH components transform differently. Since there’s no mass term, there’s nothing to couple together ψ_L with ψ_R , so instead of having one Dirac fermion Ψ , our Lagrangian has two independent Weyl fermions ψ_L and ψ_R !

- **Toy model (matter structure):** We need to work with Weyl fermions. We can add in a scalars and other types of particles as well, but we cannot have Dirac spinors, since massless Dirac spinors are equivalent to Weyl spinors. The interesting part here is that the Weyl spinors that we add in are decoupled, and are interpreted as distinct particles. The fermion fields that we’ll work with are a LH electron, RH electron, and LH neutrino:

$$e_L \qquad e_R \qquad \nu_L. \quad (25)$$

Note that here e_L and e_R are decoupled fields: at this point, we have no reason to treat them as different chiral components of the same field.

We need to specify the charges of each field under the symmetry transformations of Eq. (21) and Eq. (22). This is typically done in a table (the "Lorentz" section specifies if the particle is LH Weyl, RH Weyl, or a scalar):

Field	$U(1)_L$	$U(1)_Y$	Lorentz
e_L	$-1/2$	$-1/2$	$(1/2, 0)$
e_R	0	-1	$(0, 1/2)$
ν_L	$1/2$	$-1/2$	$(1/2, 0)$

Table 1: Charge assignments to the toy model.

From this we can see the mass terms that we can assign these particles. We can enumerate the Lorentz invariant ways to combine these fields together into a mass term. First, we have Majorana masses (and their hermitian conjugates):

$$e_L^T \sigma^2 e_L \qquad e_R^T \sigma^2 e_R \qquad \nu_L^T \sigma^2 \nu_L \qquad e_L^T \sigma^2 \nu_L \quad (26)$$

These are all not invariant under one or more of the $U(1)_L$ and $U(1)_Y$ symmetries, and so they aren’t valid terms in the Lagrangian. Next, we can consider the Dirac masses (and their hermitian conjugates):

$$e_L^\dagger e_R \qquad \qquad \qquad \nu_L^\dagger e_R \qquad (27)$$

Note that when we \dagger a field, we swap its quantum numbers, so e_L^\dagger has $U(1)_L$ charge $1/2$ and $U(1)_Y$ charge $1/2$. Both of these Dirac mass terms aren’t invariant under $U(1)_L$. For example, the electron Dirac mass transforms as

$$e_L^\dagger e_R + e_R^\dagger e_L \rightarrow e^{\frac{i}{2}\alpha} e_L^\dagger e_R + e^{-\frac{i}{2}\alpha} e_R^\dagger e_L, \qquad (28)$$

which is not invariant. Therefore, **any mass we construct is not invariant under the toy electroweak symmetry $U(1)_L \times U(1)_Y$.**

- **The Higgs mechanism:** What gives? This is supposed to be a toy theory of the Standard Model, which has a massive electron. Why can’t we add mass to the theory, and why are e_L and e_R separate Weyl fermions instead of being packaged together into one Dirac fermion? The answer to this ended up being worthy of winning a Nobel prize, and is one of the most pop-sciency particles that exist: the answer lies with the Higgs boson. Essentially, the Higgs boson H is a scalar particle whose existence allows e_L and e_R to couple together. In the electroweak sector of the Standard Model, we find the following extra coupling (called a **Yukawa coupling**):

$$\mathcal{L}_{EW} \supset \lambda(H^\dagger e_R^\dagger e_L + e_L^\dagger e_R H) \qquad (29)$$

From this, we can deduce its transformation properties under $U(1)_L$ and $U(1)_Y$, since \mathcal{L}_{EW} has to be invariant under these symmetries. They are:

Field	$U(1)_L$	$U(1)_Y$	Lorentz
H	$-1/2$	$1/2$	$(0,0)$

Table 2: Charge assignments to the toy Higgs boson.

Now, unfortunately we can’t go too far into details on the Higgs mechanism right now, but I can give a basic description of how this coupling gives the electron mass. Through a process called **spontaneous symmetry breaking**, below the energy $v \equiv 247$ GeV, the Higgs boson takes on a “vacuum expectation value” (vev) and can be parameterized as:

$$H(x) = \frac{1}{\sqrt{2}}(v + h(x)) \qquad (30)$$

Here v is this constant energy scale, and $h(x)$ is the **physical Higgs boson**, which is a real field. Note that parameterization is **only valid when we work at energies $E \ll v$; when we get to energy scales on the order of $\gtrsim v$, this parameterization does not hold**. However, 247 GeV is a HUGE energy scale; for context, the LHC currently probes energies around 14 TeV. The Higgs mass is around this scale too (it’s $m_h = v/2 \approx 125$ GeV), and it’s one of the reasons we took so long to find the Higgs: it was predicted to exist in the 1960s, but it wasn’t actually found until 2012.

Let’s then assume the parameterization of Eq. (30) and see where we can go from there. If we plug this into \mathcal{L}_{EW} , we get:

$$\mathcal{L}_{EW} \Big|_{E \ll v} \supset \frac{\lambda}{\sqrt{2}} ((v+h)^\dagger e_R^\dagger e_L + e_L^\dagger e_R (v+h)) \quad (31)$$

$$= \frac{\lambda v}{\sqrt{2}} \underbrace{(e_R^\dagger e_L + e_L^\dagger e_R)}_{\text{Dirac mass for } e} + \frac{\lambda}{\sqrt{2}} \underbrace{(h^\dagger e_R^\dagger e_L + e_L^\dagger e_R h)}_{\text{3-point couplings}} \quad (32)$$

$$= \frac{\lambda v}{\sqrt{2}} \bar{e}e + \frac{\lambda}{\sqrt{2}} h \bar{e}e \quad (33)$$

where now e is a Dirac spinor field for the electrons:

$$e = \begin{pmatrix} e_L \\ e_R \end{pmatrix}. \quad (34)$$

What happens here is that **the Higgs has given the electrons mass**. By taking on a vev, it allows us to write a mass term in the Lagrangian, and treat the two Weyl fields e_L and e_R , which we once thought were independent fields, as two chiral components of the same Dirac fermion. The actual Higgs mechanism is slightly more complicated than this, but this is the main idea: the Higgs taking a vev allows for a way to give particles mass that respects chiral gauge symmetry.

So, let’s sum the cases up for the different energy scales we can work at:

1. $E \gtrsim v$ (very high energy): In this case, spontaneous symmetry breaking has not occurred. In this regime, the LH and RH electrons e_L and e_R are massless, independent particles.
 2. $E \ll v$ (“low” energies²)
- Why are all neutrinos left-handed? The short answer: **we don’t know**. There are some thoughts that at higher energies, the electroweak gauge group expands to include right-handed rotations:

$$SU(2)_L \times U(1)_Y \rightarrow SU(2)_L \times SU(2)_R \times U(1)_Y \quad (35)$$

but we haven’t seen any indication of this yet.

²I don’t want to say low energy here because they aren’t really that low! That’s how large the scale v is. These energy scales are what we experience in everyday life, and when we do science, unless we’re actively working at a particle collider.