MATH 8450 – LECTURE 2 – JAN 23, 2023

ABDELMALEK ABDESSELAM

General introduction to the course (continued):

Regarding the construction of a QFT, one can state the main goal as that of obtaining, for all $n \in \mathbb{N}$, and all smooth and fast decaying (details further below) functions $f_1, \ldots, f_n$, the quantities

$$C_n(f_1, \ldots, f_n) = \lim_{r \to -\infty} \lim_{s \to \infty} \frac{C^{\mathrm{U}}_{n,r,s}(f_1, \ldots, f_n)}{C^{\mathrm{U}}_{0,r,s}} \tag{1}$$

where

$$C^{\mathrm{U}}_{n,r,s}(f_1, \ldots, f_n) := \int_{\mathbb{R}^{\Lambda_{r,s}}} \prod_{x \in \Lambda_{r,s}} \mathrm{d}\phi(x) \quad \prod_{j=1}^{n} \left(\sum_{y_j \in \Lambda_{r,s}} L^{dr} f_j(y_j) \phi(y_j) \right)$$
$$\times \exp\left(- \sum_{x \in \Lambda_{r,s}} L^{dr} \left[\frac{a_{r,s}}{2} \sum_{i=1}^{d} \left(\frac{\phi(x + L^r e_i) - \phi(x)}{L^r} \right)^2 + \frac{m^2_{r,s}}{2} \phi(x)^2 + \frac{\lambda_{r,s}}{24} \phi(x)^4 \right] \right) . \tag{2}$$

Notice, compared to the formulas in Lecture 1, a small difference which is the introduction of $a_{r,s}, m_{r,s}, \lambda_{r,s}$, for more generality and flexibility. The old formula had $a_{r,s} = 1$ and the mass $m_{r,s} = m$ and coupling $\lambda_{r,s} = \lambda$ were constant with respect to the cutoff parameters r, s. The s dependence is not absolutely essential, and one could do with just a_r, m_r, λ_r. However, in order for the limits (1) to exist and avoid being trivial, it is important to allow this r-dependence. Figuring out how the parameters have to depend on r is called renormalization. A particularly trivial situation is to have the C_n exist but be identically zero. A more subtle form of triviality is the following.

First notice that the C_n, if they exist, are symmetric functions, namely, $\forall \sigma \in \mathfrak{S}_n$, and for all $f_1, \ldots, f_n$,

$$C_n(f_{\sigma(1)}, \ldots, f_{\sigma(n)}) = C_n(f_1, \ldots, f_n) .$$

Here $\mathfrak{S}_n$ is the symmetric group, i.e., the set of all bijections (permutations) $\sigma : [n] \to [n]$, with the operation of composition. Of course, as practice for our system of notation, we have $|\mathfrak{S}_n| = n!$.

We have the identity, for any even n and collection of test functions $f_1, \ldots, f_n$,

$$\sum_{P} \prod_{\{i,j\} \in P} C_2(f_i, f_j) = \frac{1}{2^{\frac{n}{2}} \left(\frac{n}{2}\right)!} \sum_{\sigma \in \mathfrak{S}_n} C_2(f_{\sigma(1)}, f_{\sigma(2)}) C_2(f_{\sigma(3)}, f_{\sigma(4)}) \cdots C_2(f_{\sigma(n-1)}, f_{\sigma(n)}) \tag{3}$$

where P is a set partition of $[n]$ with blocks of size 2 only.

Notation/terminology: If X is a set, then the power set of X is the set of all subsets of X and it is denoted $\mathscr{P}(X)$. Clearly, if X is finite, then $|\mathscr{P}(X)| = 2^{|X|}$. A set partition Π of X is a subset of $\mathscr{P}(X)$ such that

- $\forall A \in \Pi, A \neq \varnothing$.
- $\forall A, B \in \Pi, A \neq B \Rightarrow A \cap B = \varnothing$.
- $\bigcup_{A \in \Pi} A = X$.

The elements A of Π are called blocks of the partition. Note that blocks are not numbered. As a result, if say $X = [4]$, then there are exactly 3 set partitions of X with only blocks of cardinality 2, namely:

$$\{\{1,2\},\{3,4\}\}, \quad \{\{1,3\},\{2,4\}\}, \quad \{\{1,4\},\{2,3\}\}.$$

A set partition P with blocks of size 2 as in the identity above is called a Wick contraction scheme in the physics QFT literature, and a perfect matching in the combinatorics literature.

Now an unwanted trivial situation is when for all even $n \geq 4$, and all f's,

$$C_n(f_1, \ldots, f_n) = \sum_P \prod_{\{i,j\} \in P} C_2(f_i, f_j) \ .$$

This corresponds to free/Gaussian/mean field QFT's where the underlying particles do not interact.

For example, for $n = 4$, an example of identity we do not want to see hold is

$$C_4(f_1, f_2, f_3, f_4) = C_2(f_1, f_2)C_2(f_3, f_4) + C_2(f_1, f_3)C_2(f_2, f_4) + C_2(f_1, f_4)C_2(f_2, f_3) \ .$$

Exercise 1. *Prove the identity (3).*

About the test functions: I was a bit vague about the requirements on the so-called test functions f_j. Now let's be more precise. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a C^∞, i.e., infinitely differentiable or smooth function. We will use the multiindex notation for partial derivatives of arbitrary order

$$\partial^\alpha f(x) := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}(x) \ .$$

Here $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$ is called a multiindex, and $|\alpha| := \alpha_1 + \cdots + \alpha_d$ is called the length of the multiindex. Multiindices can be added componentwise, namely, if $\alpha = (\alpha_1, \ldots, \alpha_d)$ and $\beta = (\beta_1, \ldots, \beta_d)$ then

$$\alpha + \beta := (\alpha_1 + \beta_1, \ldots, \alpha_d + \beta_d) \ .$$

For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ we denote the Euclidean norm by

$$|x| := \sqrt{x_1^2 + \cdots + x_d^2} \ ,$$

and the so-called Japanese bracket by

$$\langle x \rangle := \sqrt{1 + |x|^2} \ .$$

The latter as the advantage that $\langle x \rangle \sim |x|$ when x is large but $\langle x \rangle$ never vanishes. For f a smooth function as above, we define the Schwartz seminorm

$$\|f\|_{k,\alpha} := \sup_{x \in \mathbb{R}^d} \langle x \rangle^k |\partial^\alpha f(x)|$$

for all $k \in \mathbb{N}$ and $\alpha \in \mathbb{N}^d$. We now define the set $\mathscr{S}(\mathbb{R}^d)$ of all C^∞ complex-valued functions f on $\mathbb{R}^d$, which are such that

$$\forall k \in \mathbb{N}, \forall \alpha \in \mathbb{N}^d, \|f\|_{k,\alpha} < \infty \ .$$

This is called the Schwartz space and it is a vector space using pointwise sum and scalar multiplication. This is where we will pick our test functions f_j.

Exercise 2. *Check that $\mathscr{S}(\mathbb{R}^d)$ is indeed a $\mathbb{C}$-vector space, and that the maps C_n considered before, if they exist, are n-multilinear maps $\mathscr{S}(\mathbb{R}^d)^{\times n} \to \mathbb{C}$.*

For a point z_j as in Lecture 1, we could take

$$f_j(y) = \frac{1}{(\varepsilon\sqrt{2\pi})^d} e^{-\frac{|y-z_j|^2}{2\varepsilon^2}}$$

with $\varepsilon > 0$ small, as an example of weight profile peaked around z_j. This is an example of function in Schwartz space. Schwartz space is the basic ingredient of the theory of (temperate) Schwartz distributions. We will try to avoid going into it because this is another entire course, but see the partial (Kernel Theorem done in class missing) notes DistributionsMATH7320F17.pdf on Canvas.

Renormalization in a few "easy" steps: In order to make the logic/conceptual structure clear, here is a summary of how it works.

- Step 1: Regularization. Namely, introducing cutoffs/discretizations, in order to have a well defined quantity to start with.
- Step 2: Seeing hidden parameters. Often the most difficult step.
- Step 3: Imposing renormalization conditions. Through an inversion (like finding x in terms of y if we know that $y = f(x)$), this should tell us the value of the parameters, before taking the limit of removing the cutoffs.
- Step 4: Showing that the other quantities of interest (infinitely many of them) not concerned by the renormalization conditions, will have satisfactory limits.

My first example of renormalization procedure: (Calculus 1)

Let's travel back in time to around the mid 1600's and let's put ourselves in the shoes of Leibnitz and Newton working on developing calculus. For the sake of this counterfactual story, suppose the analogues of physics QFT textbooks of the time talked about the mythical object

$$\int_0^1 f(x)\ \mathrm{d}x$$

yet math textbooks only explained the theory of finite sums. Recall that the integral symbol $\int$ is a stylized/artistic rendering of the letter "S" as in the Latin word "Summa" for sum. If taken too seriously/literally then the problem is to define

$$\sum_{x\in[0,1]} f(x) \ .$$

Step 1 or regularization amounts to replacing the very infinite set $[0, 1]$ by a finite discrete set say

$$F_N := \left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}\right\}$$

which looks more and more like the original interval in the limit where the cutoff parameter N is taken to infinity. Now we have a well defined object to work with

$$S_N = \sum_{x \in F_N} f(x) \ .$$

Step 2 amount to noticing that

$$S_N = 1 \times \sum_{x \in F_N} f(x)$$

or rather

$$S_N = c_N \sum_{x \in F_N} f(x) \tag{4}$$

where the hidden parameter c_N just happened to be set equal to 1. We then proceed to allow more flexibility when taking $N \to \infty$ in (4). This flexibility is not carte blanche to do anything we want, because we intend to use the same c_N or N-dependence for all (say continuous) functions f.

Now Step 3 amounts to imposing a renormalization condition, for instance we pick our favorite f given by the constant function equal to 1, and impose that for that particular f,

$$S_N = 1 \ ,$$

before taking the limit. Clearly, this is the equation $Nc_N = 1$ and we can readily invert, i.e., deduce from it the value $c_N = \frac{1}{N}$. Finally, in Step 4 we check that our renormalized definition

$$\int_0^1 f(x)\ \mathrm{d}x := \lim_{N \to \infty} \frac{1}{N} \sum_{x \in F_N} f(x)$$

calibrated using the function $f \equiv 1$, works for all other continuous functions f.

My second example of renormalization procedure: (The Central Limit Theorem)

The highlight of MATH 3100 or introduction to probabilty is the Central Limit Theorem (CLT). Suppose one has an infinite sequence of IID (independent identically distributed) random variables $X_1, X_2, X_3, \ldots$ which can be thought of as independent copies of a basic random variable X, and our goal is to somehow define the distribution of the random variable

$$X_1 + X_2 + \cdots$$

Under a rather mild hypothesis $\mathbb{E}(X^2) < \infty$ and a simple renormalization procedure, one can do that and obtain the CLT. We will denote the mean by $\mu = \mathbb{E}(X)$ and the standard deviation by $\sigma = \sqrt{\mathrm{var}(X)}$ where the variance of X is $\mathrm{var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. In Step 1 we replace the infinite sum by a finite sum

$$S_N = X_1 + \cdots + X_N$$

and ask ourselves about the limit in distribution when $N \to \infty$. Then Step 2 amounts to noticing

$$S_N = 1 \times (X_1 + \cdots + X_N + 0)$$

or rather

$$S_N = a_N \times (X_1 + \cdots + X_N + b_N)$$

where we see the hidden parameters originally set to $a_N = 1$ and $b_N = 0$. We now allow these parameters to vary with the cutoff N, and go ahead and impose in Step 3 the renormalization conditions

$$\mathbb{E}(S_N) = 0$$

or recentering, and

$$\mathbb{E}(S_N^2) = 1$$

which amounts to fixing the standard deviation to be equal to 1. Note that

$$\mathbb{E}(S_N) = a_N(N\mu + b_N)$$

and, given the recentering, independence, and quadratic homogeneity of the variance,

$$\mathbb{E}(S_N^2) = \mathrm{var}(S_N) = a_N^2 N\sigma^2 \ .$$

The renormalization conditions are readily inverted, with the result

$$b_N = -N\mu, \quad a_N = \frac{1}{\sigma\sqrt{N}} \ .$$

In Step 4, we see that the correct definition

$$S_N = \frac{X_1 + \cdots + X_N - N\mu}{\sigma\sqrt{N}}$$

satisfies the CLT, i.e., converges in distribution to the standard Gaussian or $\mathcal{N}(0,1)$ with density

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \ .$$

Namely, for all real numbers $a < b$,

$$\lim_{N\to\infty} \mathbb{P}\{a < S_N < b\} = \int_a^b p(x)\ \mathrm{d}x \ .$$

Note that we also get convergence of all moments, i.e., for all $n \in \mathbb{N}$

$$\lim_{N\to\infty} \mathbb{E}(S_N^n) = C_n := \int_{-\infty}^{\infty} x^n p(x)\ \mathrm{d}x \ .$$

So the nontrivial content of the renormalization procedure here is that fixing the first and second moments $\mathbb{E}(S_N)$ and $\mathbb{E}(S_N^2)$, give us the convergence of all the (infinitely many) other moments of order n.

Exercise 3. *Compute the above C_n's, i.e., the moments of the standard normal random variable, and show they satisfy the identity*

$$C_n = \sum_{P} \prod_{\{i,j\}\in P} C_2$$

(written in a funny way).

The Road Trip: The above formulation of the main problem of defining the n-point correlations $C_n(f_1, \ldots, f_n)$ is a bit like indicating the city we are heading for, say Seattle, when starting a road trip from Charlottesville. However, another reason for the road trip vs. flying to Seattle, is to visit a few interesting places along the way like the Grand Canyon etc. Here too we intend to learn about some interesting mathematical topics and among those: formal power series, tensors, diagrammatic calculus with tensors, Feynman diagrams, a diagram for part of Cardano's famous formula for the solution of cubic equations, some category theory,...

In perturbative renormalization, the parameters $a_{r,s}, m_{r,s}, \lambda_{r,s}$ will be formal power series in some new variable λ_R called the renormalized coupling constant (closer to an elusive physical/measurable coupling). Fixing renormalization conditions, means inverting power series. So the next order of business is to learn more about that. We now move on to the first Chapter/Topic about formal power series. See next page.

FORMAL POWER SERIES

Algebra review:

A group $(G, \cdot)$ is a set G together with an operation or map $\cdot : G \times G \to G$ (we will drop the dot like when writing products) such that the following axioms hold.

- (G1) The operation is associative: $\forall a, b, c \in G, a(bc) = (ab)c$.
- (G2) There is a neutral element: $\exists e \in G, \forall a \in G, ae = a$ and $ea = a$.
- (G3) Every element has an inverse: $\forall a \in G, \exists a^{-1} \in G, a^{-1}a = e$ and $a\ a^{-1} = e$.

Just from (G1) and (G2), it follows that e must be unique and therefore when we get to stating (G3) there is no ambiguity as to who this e might be. Also in (G3) the inverse of a must be unique and we might as well immediately give it a name/notation a^{-1}.
Example: the symmetric group $G = \mathfrak{S}_n$ with the operation given by the composition of (bijective) maps $\circ$.

A group is called commutative or Abelian if the operation is commutative, i.e., if $\forall a, b \in G, ab = ba$.

A commutative ring with unit $(R, +, \times)$ is a set together with two operations or maps $+ : R \times R \to R$ and $\times : R \times R \to R$ such that the following axioms hold.

- (R1) $(R, +)$ is an Abelian group. The neutral element will then be denoted by 0_R or just 0, and the inverse will be denoted $-a$ instead of a^{-1}.
- (R2) $\times$ is commutative.
- (R3) The distributive property holds: $\forall a, b, c \in R, a \times (b + c) = a \times b + a \times c$.
- (R4) There is a neutral element for multiplication denoted by 1_R or just 1 such that $\forall a \in R, 1a = a$.

As a consequence of the previous axioms we have $-a = (-1) \times a$ and $0a = 0$ for all a in R.

As per Diana's question: we do not in general include the existence of multiplicative inverses in the definition of rings. For example $(\mathbb{Z}, +, \times)$ is a commutative ring with unit (we'll just say ring from now on), but the only elements which have multiplicative inverses are 1 and -1. If we want inverses, then we need the notion of field (in the algebra sense).

A field $(K, +, \times)$ is a ring as above which also satifies the following.

- (F1) $\forall a \in K\backslash\{0\}, \exists a^{-1} \in K, a^{-1}a = 1$.
- (F2) $1 \neq 0$.

The axiom (F2) is added to avoid too much triviality. If $0 = 1$ in a ring then R is a singleton. For some reason, mathematicians are willing to live with rings with 1 element, but not with fields with 1 element.

Rings of formal polynomials:

Let R be a ring. Let $X_1, \ldots, X_n$ be a list of (formal) letters or symbols. We will define the ring $R[X_1, \ldots, X_n]$ of formal polynomials in the variables $X_1, \ldots, X_n$ over the (ground) ring R.

As a set

$$R[X_1, \ldots, X_n] := \left\{ (a_\alpha)_{\alpha \in \mathbb{N}^n} \in R^{\mathbb{N}^n} \mid a_\alpha = 0 \text{ for all except finitely many } \alpha'\text{s} \right\}\ .$$

The condition is called almost-finiteness. We will use the notation

$$(a_\alpha)_{\alpha\in\mathbb{N}^n} =: \sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha$$

where as part of the multiindex notation package we also write $X^\alpha = X_1^{\alpha_1}\cdots X_n^{\alpha_n}$.

An element $A = (a_\alpha)_{\alpha\in\mathbb{N}^n}$ or $A(X_1,\ldots,X_n)$, or still (according to the mood/sense of hurry) $A(X)$ of $R[X_1,\ldots,X_n]$ is called a formal polynomial in the X variables with coefficients in R. This element can be used in order to define a polynomial function f_A, i.e., a map $R^n \to R$,

$$x = (x_1,\ldots,x_n) \longmapsto f_A(x_1,\ldots,x_n) = \sum_{\alpha\in\mathbb{N}^n} a_\alpha x_1^{\alpha_1}\cdots x_n^{\alpha_n}$$

where we really mean it when writing the sum (no longer just some funny notation). The sum looks infinite because the summation set $\mathbb{N}^n$ is infinite but, because of the almost-finite condition, this is a finite sum padded with a bunch of zeros.

Keep in mind that formal polynomials are not the same as their associated polynomial functions. The map $A \mapsto f_A$ is not injective for instance if R is a finite field like $\mathbb{Z}/5\mathbb{Z}$. If R is an infinite field like $\mathbb{C}$, as will be the case soon, then the map is injective.

If $A = (a_\alpha)_{\alpha\in\mathbb{N}^n}$ is a formal polynomial, we will use the notation $[X^\alpha]A$ for the coefficient of X^α, i.e., a_α.

We define addition in $R[X_1,\ldots,X_n]$ by

$$\left(\sum_{\alpha\in\mathbb{N}^n} a_\alpha X^\alpha\right) + \left(\sum_{\alpha\in\mathbb{N}^n} b_\alpha X^\alpha\right) := \left(\sum_{\alpha\in\mathbb{N}^n} (a_\alpha + b_\alpha) X^\alpha\right)$$

or equivalently by letting $A + B$ be defined by letting

$$[X^\alpha](A+B) = ([X^\alpha]A) + ([X^\alpha]B)$$

for all $\alpha \in \mathbb{N}^n$.

The product is defined (with intentional variety of notation for practice) by

$$(a_\alpha)_{\alpha\in\mathbb{N}^n} \times (b_\alpha)_{\alpha\in\mathbb{N}^n} := (c_\alpha)_{\alpha\in\mathbb{N}^n}\ ,$$

where, for all $\gamma \in \mathbb{N}^n$,

$$c_\gamma := \sum_{\alpha,\beta\in\mathbb{N}^n} \mathbb{1}\{\alpha+\beta=\gamma\}\ a_\alpha b_\beta\ .$$

The symbol $\mathbb{1}\{\cdots\}$ is by definition equal to 1 (1_R to be precise) if the condition/logical proposition between braces is true, and equal to 0 (namely, 0_R) if the condition is false. For instance the Kronecker delta from matrix theory is just $\delta_{ij} = \mathbb{1}\{i=j\}$.