Recitation 1

Wentao Cui

February 12, 2023

Logistics

Contact Information:

• TA:	Wentao Cui	
• Recitation:	Friday 4-5pm,	Room 4-163
• Office Hour:	Monday 6:30-7:30pm,	Room 8-320

Texts:	Peskin and Schroeder:	An Introduction to Quantum Field Theory	$(+,-,-,-)$
	Weinberg, Volume 1:	The Quantum Theory of Fields, Foundations	$(-,+,+,+)$
	Schwartz:	Quantum Field Theory and the Standard Model	$(+,-,-,-)$
	Srednicki:	Quantum Field Theory	$(-,+,+,+)$
		c.f. Hong:	$(-,+,+,+)$

1 Preliminaries

1.1 Relativistic Quantum Mechanics: Why QFT?

Often, it is written that QFT = QM + SR. Physicists say this with nearing a century of experience bourne from confusion and pain due to the limitations of the theories they built to describe physical phenomena. When learning the subject, one is presented with a polished product that is often unmotivated, a big black box that works. Hence, as one is churning the heavy machinery that is QFT to produce some sensible result, it can be difficult to see why we need it at all. For instance, why do we need fields? But QFT is not abstraction for the sake of abstraction, and if there was a simpler theory that described particle physics we would have found it. In light of this, I'd like to begin today by making extremely explicit why quantum mechanics alone fails to describe the physics of very small scales.

Fact 1: Ordinary (non-relativistic) QM breaks down at short distances.
In ordinary QM, let us consider the simplest possible system–a particle in a box. By the uncertainty principle, we know that $\Delta x \Delta p \gtrsim \hbar$. Hence, if we confine our particle to scales $\sim \hbar/c$, then the momentum becomes relativistic, and we cannot ignore relativistic corrections.

Having realized this as a physicist in the mid-1920s, one's programme may be to extend quantum mechanics to include relativity. The naïve extension, which we can call 'relativistic quantum mechanics', would be to generalize the Schrödinger equation using the relativistic dispersion relation. There is a wavefunction giving the probability at time t to find particle 1 at x_1, particle 2 at x_2, and so on.

Fact 2: Relativistic quantum mechanics is inconsistent.
Let us return to our particle in a very small box, where Δp is ultrarelativistic, say many times the mass of our particle. Here we use a principle from SR, namely the dispersion relation. This tells us that ΔE

is also very large. Now, with high probability the energy of the system is much larger than the rest mass of the particle. In fact, there is now enough energy to spontaneously create more particles, so that the particle number is also probabilistic. This can also be seen using the energy-time uncertainty relation: if one stares at a patch of vacuum for a very short time, the notion of how many particles there are is an ill-defined notion. This is impossible to explain in relativistic quantum mechanics where particle number is constant, or maybe at best you might be able to extend the theory so there is some upper bound. We need something more, and this observation begins the high road to quantum field theory.

- Note: in QFT, the first mechanism is explained by the ability to rip particles straight out of the vacuum if the local energy is high enough. The energy-time problem is explained by virtual particles, which only 'exist' for short periods of time, popping in and out of existence. Far be it from static and boring, in QFT the vacuum is a swirling, boiling sea with profoundly rich dynamics.

1.2 Units

Dimensionful Quantities in QFT
In all of physics, there are 4 basic dimensionful units from which any dimensionful unit can be derived. These measure time $[T]$, length $[L]$, mass $[M]$, and temperature $[T]$. In this course we ignore temperature. In Newtonian mechanics, these are all unrelated. Let's see what happens in QFT. For convenience, let's also add p ($[MLT^{-1}]$) and E ($[ML^2T^{-2}]$).

$$\begin{tikzcd}
p \arrow[r, leftrightarrow, "E^2 = (pc)^2 + (mc^2)^2"] \arrow[d, leftrightarrow, "\Delta x \Delta p \gtrsim \hbar" description] & E \arrow[r, leftrightarrow, "E^2 = (pc)^2 + (mc^2)^2"] \arrow[d, leftrightarrow, "\Delta E \Delta t \gtrsim \hbar" description] & m \\
x \arrow[r, leftrightarrow, "ds^2 = (cdt)^2 - dx^2"] & t &
\end{tikzcd}$$

We know that QFT = QM + SR. Each of the theories on the right is characterized by a fundamental dimensionful quantity. In special relativity, we have $c \sim 3 \cdot 10^8 m\, s^{-1}$. It determines the speed of causality relating the structure of space to the structure of time, and determines the dispersion relation relating momentum to mass-energy. In quantum mechanics, $\hbar \sim 1.1 \cdot 10^{-34} kg\, m^2\, s^{-1}$ determines the intensity of quantum probability fluctuations, demarcating the scale at which the sum over all histories in the path integral can be approximated by the principle of least action. From the perspective of the uncertainty relation, it relates position to momentum, or energy to time.

It figures then, that QFT should include both. All these quantities are now related. For instance, generically when we increase the energy E, we increase the momentum p of our particles, and have access to higher mass, heavier particles. By quantum mechanics, this means that characteristic times and distances involved in high energy processes are very short. For this reason, the field surrounding QFT is often called high energy physics.

Natural Units
There's a trick we can use to get these results for free just by introducing some notation. This is the idea of 'natural units', by which we write $\hbar = c = 1$. It's common when starting out to look at this and get confused. How can we set a dimensionful quantity equal to 1? This is not what we're doing, rather it is a shorthand which hides the dimension. What we're really doing is setting two types of measurements equal to one another: time and length. Measuring distance in time units is something all of us are familiar

with–for instance, 10 minutes as the crow flies, or distances between stars in light-years. Here we are doing nothing more than that; all we need is a reference velocity, which special relativity naturally provides (for reasons stated above). By the same argument, using the reference 'action' quantity and the uncertainty relation, we see that we can measure momentum as inverse length, or energy as inverse time.

We started with 3 independent dimensionful quantities, and related 2 pairs. Consequentially, in QFT we can express any quantity in terms of length (e.g. fm $= 10^{-15}m$), or as more often done, a mass scale (e.g. GeV). Working through units one can check that with $c = \hbar = 1$,

$$[M] = [E] = [P] = [L]^{-1} = [T]^{-1} \tag{1}$$

That is, momenta and accessible masses scale with energy, while distance or time scale inversely with energy. This is our result from before, which natural units and dimensional analysis have given us for free.

Note 1: For instance, we can talk about the length scale associated with a particle of mass m just by taking its inverse (in natural units). This is called the Compton wavelength, $\lambda_c = 1/m$. Physically, it's the wavelength of a photon with frequency equal to the rest energy of said particle.

Note 2: Restoring units. Given a generic measurement in terms of some energy scale, how do we return to normal units? This is done by multiplying our answer by some factors of c and $\hbar$ to give the desired dimensions. For instance, say we have a volume $V = 10^{-3}\text{GeV}^{-3} = 10^{-3}\hbar^m c^n \text{GeV}^{-3}$. Since volume has dimensions $[L]^3$, a straightforwards calculation gives $m = n = 3$. Using $1\text{GeV} \sim 10^{-10}J$ and substituting SI values of c and $\hbar$, we get a volume of 10^{-6}fm^3.

2 The Lorentz Group

In this section we review special relativity, at the core of which is the study of Lorentz transformations, which form the Lorentz group. You should think of a group as a bag of verbs, a set whose elements act on a system, such as translations or rotations. For instance, 'turn 30 degrees', or 'move 2 units to the right.' Each element must have an inverse, and you can compose elements by doing one after another.

The Lorentz group can be intimidating, even though as we'll see there's nothing to be intimidated about. To drive this point home we'll start by discussing rotations, which you should all be experts in. Lorentz transformations are just spicy rotations, and all said spice is contained in a single minus sign. If you ever get confused about Lorentz transformations, ask the same question about rotations.

2.1 Rotations

SO(2): Let's start with the simplest case. Consider the space of proper (orientation preserving) rotational symmetries in 2 dimensions. A point in the plane transforms as:

$$\begin{aligned} x &\to x\cos\theta + y\sin\theta \\ y &\to -x\sin\theta + y\cos\theta \end{aligned}$$

I can package this as a column or a row vector:

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} x & y \end{pmatrix} \to \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Index notation: we can write the rotation as a matrix acting on column vectors as $x^i \to R^i{}_j x^j$, and row vectors as $x_i \to x_j (R^T)^j{}_i$.

- Note 1: Einstein summation convention. If an index is repeated, it is assumed to be summed over. This is called a contraction.
- Note 2: Upper and lower indices. Column (contravariant) vectors always carry upper indices, while row (covariant) vectors always carry lower indices. Contractions are only permitted between an upper index and a lower index, reflecting matrix multiplication.
  This is a bookkeeping device that doesn't matter for rotations, and many authors are flippant about it. However, it is reflective of a deep mathematical fact. In general, matrices can act on vectors as a row or a column, and these actions are different. For example, this is why quarks are different from antiquarks. But when the matrix is real, as for rotations, these actions are the same.

We can check explicitly that in 2D, $R^T R = \mathbb{1}$, or in index notation $(R^T)^i{}_j R^j{}_k = \delta^i{}_k$.

SO(N): More generally, rotations act linearly on vectors such that distances between vectors are preserved. In Euclidean space, the distance between x^i and y^i is given by $\delta_{ij}x^i y^j$. Let us see what this means:

$$x_i y^i = \delta_{ij}x^i y^j \to \delta_{kl}R^k{}_i R^l{}_j x^i y^j = (R^T)_{il}R^l{}_j x^i y^j = x_i (R^T)^i{}_l R^l{}_j y^j$$

Since this holds for all x and y, this means that we have $N \times N$ matrices satisfying $R^T R = \mathbb{1}$, which is an equivalent way of characterizing rotations.

The metric: this is an inner product on a space, and provides a notion of distance. It is given by a symmetric matrix g_{ij}, with inner product defined using matrix multiplication, $\langle \mathbf{x}|\mathbf{y}\rangle = h_{ij}x^i y^j$. In Euclidean space, $h_{ij} = \delta_{ij}$. A rotation is a transformation preserving the metric: from the above we have

$$(R^T)_i{}^k \delta_{kl} R^l{}_j = \delta_{ij} \qquad \text{or} \qquad R^T \mathbb{1} R = \mathbb{1}$$

Discrete transformations: another linear transformation that preserves distance is the flip $(x^1, \ldots, x^n) \to (-x^1, x^2, \ldots, x^n)$. It is not orientation preserving, and one cannot reach it by performing strictly orientation preserving transformations. Adding the flip to our existing transformations gives O(N) instead of SO(N). O(2) has 2 connected components, and the component connected to the identity (trivial rotation) is SO(n).

Now, we could now more generally ask what a rotation really is, or study in greater detail the algebraic properties of the group of rotations. For us, this is the wrong question to ask. Rather, we ask the question of what we can rotate, i.e. on what kind of objects rotations can (linearly) act. Such a action is called a representation of SO(N). This may seem like a less fundamental question, but in physics any group that one encounters always acts on a system. A group is a bag of verbs. By introducing some unrelated math jargon, this approach can be summarized by the following maxim,

Groups, like men, will be judged by their actions, not their words.
This is why we should study their representations, not their presentations.

1. Scalars: objects that are invariant under a rotation. For instance, the number 3, the mass of a particle, or the action. They form a 0-dimensional space on which rotations act trivially.
2. Vectors: objects that transform as $v^i \to R^i{}_j v^j$, or $v_i \to v_j (R^T)^j{}_i$. These objects have 1 index, and examples include a position x^i, a derivative ∂_i, or a 3-momentum p^i of a particle.
3. Tensors: multi-index objects, where each index transforms as a row or column vector. For instance,

$$T^{i_1,\ldots,i_m}{}_{j_1,\ldots,j_n} \to R^{i_1}{}_{k_1} \ldots R^{i_m}{}_{k_m} R_{j_1}{}^{l_1} \ldots R_{j_m}{}^{l_n} T^{k_1,\cdots,k_k}{}_{l_1,\cdots,l_n}$$

4. Pauli spinors: these are more exotic objects for which we can define rotations. For SO(3), the rotation matrices are 2×2 acting on a 2D complex space, and look like $R_{ab} = e^{i\theta_i \sigma^i{}_{ab}}$. Here the θ's are constants, and σ^i's are the Pauli matrices.

This is it. A powerful result from representation theory is that any (finite dimensional) mathematical object that is 'rotatable', can be decompose into subobjects that behave as scalars, vectors, tensors, and spinors under rotations.
Note: we call objects that don't change under scalars 'invariant', while objects that transform as representations of SO(3) 'covariant'. We can contract indices of covariant objects (vectors, tensors, etc.) with each other to get something that is invariant, for instance $x_i p^i$ or $x^i \partial_i$.

Finally, we bring up fields. A field is a function that associates an object to each point in space. These too can be rotated, and we can mainly quote our previous results. The point of caution here is that there are two things that can be rotated–the field itself as a function over space, and the point in space at which it is evaluated. We want to see how the field behaves under the former.

1. Scalar fields: under a rotation, $\phi \to \phi'$. Scalar means that if we evaluate the transformed field at the transformed point $x' = Rx$, the field remains the same, $\phi'(x') = \phi(x)$. We can thus isolate the transformation property of the field:
$$\phi(x) \to \phi'(x) = \phi(R^{-1}x)$$
2. Vector fields: $V^\mu \to V'^\mu$. The transformation law is
$$V^i(x) \to V'^i(x) = R^i{}_j V^j(R^{-1}x)$$
The key feature is that it picks up a rotation matrix out front. Again, we need to undo the transformation of the point in space where the field is evaluated.
3. Tensor fields, spinor fields are obtained similarly.

2.2 Lorentz Transformations

We just spent an large amount of time discussing rotations. But this was not in vain; everything we saw can be extended with barely any effort. Lorentz transformations are just rotations, with some extra minus factors we need to keep track of. We denote the (connected part of the) Lorentz group in d spacetime dimensions as $\mathrm{SO}(1, d-1)$, with coordinates $x^\mu = (t, \mathbf{x}) = (t, x^i)$.

The metric: we know from special relativity that the invariant distance between 2 points in spacetime is given by $s^2 = -t^2 + \mathbf{x}^2$. s is the norm of x^μ with respect to the Minkowski metric,
$$g_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$$
The -1 contains all the difference between ordinary rotations and the Lorentz group. The inner product between vectors is $x \cdot y = g_{\mu\nu} x^\mu y^\nu$, and the norm of a vector is $x \cdot x = g_{\mu\nu} x^\mu x^\nu$.

Lorentz transformations: these are the set of symmetries of spacetime preserving the metric, or invariant distance. They act on spacetime vectors as matrices Λ^μ_ν, which as with ordinary rotations, must satisfy
$$\Lambda^\mu{}_\rho g_{\mu\nu} \Lambda^\nu{}_\sigma = g_{\rho\sigma} \tag{2}$$
As with the group of ordinary rotations SO(4) in 4 dimensions, SO(1,3) has 6 independent components, where we can 'rotate' any axis into any different axis: (t,x), (t,y), (t,z), (x,y), (x,z), (y,z). The latter 3 generate the rotations of 3D space. However, because of the -1 in the tt-component of the metric, transformations mixing a time and space axis are phenomenologically different. We call these boosts.

- Example: consider a boost between only x and t. These are transformations of x and t keeping the distance $t^2 - x^2 = \text{cst}$. This generates motion along hyperbolas (as opposed to circles for rotations). Just like 2D rotations, we may write the most general such transformation explicitly as
$$\begin{pmatrix} x \\ t \end{pmatrix} = \frac{1}{\sqrt{1-v^2}} \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \cosh\beta & \sinh\beta \\ \cosh\beta & \sinh\beta \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \tag{3}$$

Here v is just a parameter. We recognize $\frac{1}{\sqrt{1-v^2}} = \gamma = \cosh\beta$, where β is the rapidity.

Aside: parameterizing the Lorentz group. Lie groups, Lie algebras, and the exponential map.

- It is not difficult to write rotations or boosts between any 2 axes. When you compose them, however, it is much easier to use a matrix exponential to write a generic Λ. I could have also done this for ordinary rotations–this provides a profound mathematical alternative to the mess that is the Euler angles. For the Lorentz group one has:

$$\Lambda = \exp(i\beta_i K_i + i\theta_i J_i) = \exp\left(\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}\right) = \lim_{n\to\infty}\left(\mathbb{1} + \frac{1}{n}\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}\right)^n \tag{4}$$

  Here K_i, J_i are antisymmetric 4×4 matrices generating boosts and rotations respectively, with non-vanishing components:

$$(J_i)_{jk} = -i\epsilon_{ijk} \qquad\qquad (K_i)_{0j} = \delta_{ij} = -(K_i)_{j0}$$

  They can be packaged nicely as a tensor of 4×4 matrices $\mathcal{J}^{\mu\nu}$. Furthermore, $\omega_{\mu\nu}$ can be related to the rapidities and rotation angles from before:

$$\mathcal{J}^{\mu\nu} = \begin{pmatrix} 0 & K_1 & K_2 & K_3 \\ -K_1 & 0 & J_3 & -J_2 \\ -K_2 & -J_3 & 0 & J_1 \\ -K_3 & J_2 & -J_1 & 0 \end{pmatrix} \qquad\qquad \omega_{\mu\nu} = \begin{pmatrix} 0 & \beta_1 & \beta_2 & \beta_3 \\ -\beta_1 & 0 & \theta_3 & -\theta_2 \\ -\beta_2 & -\theta_3 & 0 & \theta_1 \\ -\beta_3 & \theta_2 & -\theta_1 & 0 \end{pmatrix}$$

- In the last equality of (4), we have used the exponential map to write a general Lorentz transformation as the composition of infinitesimal Lorentz transformations, meaning they are infinitesimally close to the identity. Intuitively this makes sense, I can add infinitesimal transformations to get something finite. The consequences of this are hefty–in particular, I can know almost everything about the Lorentz group (as in, aside from some topological data) just by studying the group in an infinitesimal neighborhood of the identity.
  A more familiar example might be the space of 2D rotations SO(2), which is just a circle parameterized by the rotation angle. Very close to the identity ($\theta = 0$), this just looks like a line.
- More generally, the rotation and Lorentz groups are special cases of Lie groups, which are groups with continuous parameters. The Lie algebra is the tangent space of the Lie group at the identity, and you can always define an exponential which maps the Lie algebra to the Lie group. These form a core area of study in high energy theory.

Discrete transformations: so far, we have only looked at SO(1, 3) which is the part of the Lorentz group connected to the identity. This called the proper orthochronous subgroup. Just like with the rotation group, the Lorentz group is disconnected, but it has 4 disconnected components instead of 2. From SO(1, 3) one can access the other 3 by acting with parity $P : (t, x, y, z) \to (t, -x, -y, -z)$ or time-reversal $T : (t, x, y, z) \to (-t, x, y, z)$. Together, P and T generate the set $\{1, P, T, PT\}$, which is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

As before, we can consider objects which can be Lorentz transformed–that is, both rotated and boosted. For scalars, vectors, and tensors, the analysis proceeds identically, just replacing Roman indices with Greek ones, and $R^i{}_j$ with $\Lambda^\mu{}_\nu$. Examples of scalars include a particle's mass or the action. Vectors include $x^\mu = (t, x^i)$, $p^\mu = (E, p^i)$, and $\partial^\mu = (\partial^t, \partial^i)$. The discussion for scalar, vector, and tensor fields also proceeds identically. If you're not familiar with it, it's a good exercise to go over everything we did for rotations, and write it for Lorentz transformations. As a final note, the spinors and spinor fields we get are more tricky; the resulting objects that can be rotated and boosted are called Weyl spinors, which differ a bit from Pauli spinors. You'll see much more on spinors in the weeks to come.

3 Why Fields?

I'd like to end off today by answering one of the big questions which is often obscured. This is to address the F in 'QFT'. Where do the fields come from? They are not found in the axioms of either QM or SR. The claim is the following: any theory combining both QM and SR describing even just 1 particle necessitates a description using fields. Here we outline the argument.

1. Special Relativity and Symmetries
   From SR we have the symmetries of our theory: rotations, boosts, and translations. Physics should be the same regardless of whether I do it here or turn 30 degrees walk 2 meters, and jump aboard a train going half the speed of light. The group that combines all of these is called the Poincaré group, which I get by adding translations to the Lorentz group.

2. Describing Particles
   We know that our universe has particles, with momentum, mass, spin, and all sorts of quantum numbers. If we rotate or boost to change frame, only the momenta and spin component (say, S_z) change, and both change in a Poincaré-invariant way. The other quantum numbers are invariant. This motivates the definition of a particle a set of states that mix only among themselves under Poincare transformations.

3. Quantum Mechanics and Representations
   In QM, the states representing a particle form a vector space, and we have argued that the Poincaré group must act naturally on it, $|\psi\rangle \to \mathcal{P}|\psi\rangle$. A set of objects $|\psi\rangle$ which mix under the action of a group is a representation. We saw some representations of this group earlier: scalars, vectors, tensors, and spinors. These are finite-dimensional representations, with dimensions 1, 4, 16, etc. We saw how they transform under rotations and boosts, and they are invariant under rotations. So far this is fine. We can describe particles without fields. The moral here is that 'particles transform under irreducible representations of the Poincare group.' Weinberg uses this as the definition of what a particle is.

4. Unitarity
   But when we add one more condition required from quantum mechanics, we will see that we need fields. This is the notion of unitarity. In particular, we need matrix elements to be invariant under Poincaré transformations:

   $$\mathcal{M} = \langle\psi_1|\psi_2\rangle = \langle\mathcal{P}\psi_1|\mathcal{P}\psi_2\rangle = \langle\psi_1|\mathcal{P}^\dagger\mathcal{P}|\psi_2\rangle$$

   This requires $\mathcal{P}^\dagger\mathcal{P} = \mathbb{1}$, i.e. the matrices which encode the action of Poincaré transformations must be unitary. The problem lies in that there exist no finite-dimensional unitary representations of the Poincaré group. Therefore, if we assume unitarity, particles cannot be described the same way as traditional quantum mechanics.

5. Wigner's classification
   Eugene Wigner classified all the unitary Poincaré representations all the way back in 1939. All of these are infinite dimensional, and have natural descriptions in terms of fields. Neither I nor Hong will go through this proof, you can find it in Weinberg Chapter 2 if you're interested. The key lies in that the group acts on the field's spacetime dependence in a way which restores unitarity. But we have reached the moral I'd like all of you to take away, which is that any theory combining both QM and SR describing particles necessitates a description using fields.

8.323 Relativistic Quantum Field Theory I
Spring 2023