Preliminaries: What You Need To Know

Quantum field theory draws heavily from a number of previous courses, both mathematical and physical. Here I summarize the key concepts that you will need throughout this course. This will also serve to set conventions and notations. If you're not comfortable with any of these concepts or terms, you should review a suitable book as soon as possible.

Classical Dynamics

We'll need both the Lagrangian and Hamiltonian formulation of dynamical systems. Recall that a physical system is described by coordinates q_a, $a = 1, \ldots, n$ whose dynamics are governed by a *Lagrangian* $L = L(q_a, \dot{q}_a)$. The *action* is defined by

$$S = \int_{t_1}^{t_2} L(q_a, \dot{q}_a)\, dt \tag{1.1}$$

The true path taken by the system is an extremal of the action, i.e. $\delta S = 0$. This leads to the equations of motion,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial L}{\partial q_a} = 0 \tag{1.2}$$

An important result from classical dynamics that will carry over to field theories is Noether's theorem. Let's review this theorem: it states that every continuous symmetry gives rise to a conserved quantity. Consider an infinitesimal transformation

$$\delta q_a = X_a(q, \dot{q}, t) \tag{1.3}$$

This is a symmetry if $\delta L = 0$ for *all* paths[1]. To prove Noether's theorem, consider an arbitrary deformation δq_a. Then

$$\delta L = \left(\frac{\partial L}{\partial q_a}\,\delta q_a + \frac{\partial L}{\partial \dot{q}_a}\,\delta \dot{q}_a\right) = \left(\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right)\right)\delta q_a + \frac{dQ}{dt} \tag{1.4}$$

where

$$Q = \frac{\partial L}{\partial \dot{q}_a}\,\delta q_a \tag{1.5}$$

When the equations of motion are obeyed, the term in brackets vanishes and we have $\delta L = \dot{Q}$. Since the symmetry transformation $\delta q_a = X_a(q, \dot{q}, t)$ is defined to have $\delta L = 0$, we learn that when the equations of motion are obeyed, $Q = (\partial L/\partial \dot{q}_a) X_a$ is a conserved quantity, i.e. $\dot{Q} = 0$.

[1] In fact we can relax this condition and ask only that $\delta L = dF/dt$, so that the action remains unchanged for all paths.

To move towards the Hamiltonian formalism, we define the momentum p_a conjugate to q_a

$$p^a = \frac{\partial L}{\partial \dot{q}_a} \tag{1.6}$$

The *Hamiltonian* of the system is then defined as

$$H(q,p) = p^a \dot{q}_a - L \tag{1.7}$$

where H is to be considered a function of q_a and p^a rather than q_a and $\dot{q}_a$. The equations of motion are now given by Hamilton's equations

$$\dot{q}_a = \frac{\partial H}{\partial p^a} \quad , \quad \dot{p}^a = -\frac{\partial H}{\partial q_a} \tag{1.8}$$

Finally, let us recall the *Poisson bracket*, an important quantity in classical mechanics which aids the leap to the quantum theory. For any functions $f(q,p)$ and $g(q,p)$ on phase space, the Poisson bracket is defined by

$$\{f,g\} = \frac{\partial f}{\partial q_a}\frac{\partial g}{\partial p^a} - \frac{\partial f}{\partial p^a}\frac{\partial g}{\partial q_a} \tag{1.9}$$

In particular, $\{q_a, q_b\} = \{p^a, p^b\} = 0$ while $\{q_a, p^b\} = \delta_a^b$.

Quantum Mechanics

In classical mechanics, the state of a system is determined by a point in phase space, specified by (q_a, p^a). In contrast, in quantum mechanics, the state of a system is specified by a vector $|\psi\rangle$ in Hilbert space. The path from the classical system to the quantum system proceeds by promoting functions f on phase space to operators $\hat{f}$ acting on the Hilbert space. The map between the two system is given by the relationship between Poisson brackets and commutators,

$$\{\ ,\ \}_{\text{classical}} \leftrightarrow -\frac{i}{\hbar}\ [\ ,\ ]_{\text{quantum}} \tag{1.10}$$

In particular, we have

$$[\hat{q}_a, \hat{q}_b] = [\hat{p}^a, \hat{p}^b] = 0 \quad \text{and} \quad [\hat{q}_a, \hat{p}^b] = i\hbar\ \delta_a^b \tag{1.11}$$

This prescription is known as *canonical quantization*. Note that a generic classical function $f(q,p)$ does not define a unique quantum operator $\hat{f}$ due to ordering ambiguities.

The dynamics of a quantum system is governed by the Hamiltonian operator $\hat{H}$. There are two different ways of viewing the dynamics. In the *Schrödinger picture*, the states $|\psi\rangle$ evolve in time, while the operators $\hat{O}$ are time independent. The dynamics is governed by the Schrödinger equation,

$$i\hbar\frac{d\,|\psi\rangle_S}{dt} = \hat{H}\,|\psi\rangle_S \tag{1.12}$$

where the subscript on $|\psi\rangle_S$ reminds us that we're in the Schrödinger picture. In contrast, in the *Heisenberg picture*, the states are time independent, while the operators now change with time. The relationship between the two is given by

$$|\psi\rangle_H = e^{i\hat{H}t/\hbar}\,|\psi\rangle_S \tag{1.13}$$
$$\hat{O}_H = e^{i\hat{H}t/\hbar}\,\hat{O}_S\,e^{-i\hat{H}t/\hbar} \tag{1.14}$$

and the operators now evolve by

$$\frac{d\hat{O}_H}{dt} = \frac{i}{\hbar}[H,\hat{O}_H] \tag{1.15}$$

Note that $\hat{H}_S = \hat{H}_H \equiv \hat{H}$. The two pictures are entirely equivalent since all correlation functions $\langle\phi|\,\hat{O}\,|\psi\rangle$ agree. In quantum field theory we will jump merrily between these two different pictures. We will also employ a third viewpoint which is something of a hybrid of the first two, known as the *interaction picture.*

In this course our notation will differ slightly from that above. Firstly, we will work in natural units with $\hbar = 1$. Secondly, we will not denote operators with a $\hat{hat}$; it should be clear from the context whether we're talking about classical objects or quantum objects.

Special Relativity

Minkowski space is parameterized by coordinates $x^\mu = (x^0, x^i)$, where $\mu = 0,1,2,3$ and $i = 1,2,3$. We will often denote x^0 as time t, and the spatial three vector x^i as $\vec{x}$. (On the blackboard, the 3-vector will have a wavy line on the bottom rather than an arrow on top). We will work in units in which the speed of light is set to $c = 1$. The Minkowski metric is then given by

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \eta^{\mu\nu} \tag{1.16}$$

so that $\eta_{\mu\nu}\eta^{\nu\rho} = \delta_\mu^\rho$. We use $\eta^{\mu\nu}$ ($\eta_{\mu\nu}$) to raise (lower) indices on vectors and tensors.

The relativistic line interval is given by

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = (dx^0)^2 - d\vec{x} \cdot d\vec{x} \tag{1.17}$$

A relativistic particle of mass m traces out a path $x^\mu(s)$ in Minkowski space. The 4-momentum p^μ is defined by

$$p^\mu = m \frac{dx^\mu}{ds} \tag{1.18}$$

which satisfies $p^2 \equiv p^\mu p_\mu = \eta_{\mu\nu}(dx^\mu/ds)(dx^\nu/ds) = m^2$. So writing the 4-momentum $p^\mu = (E, \vec{p})$ in terms of the energy E and the 3-momentum $\vec{p}$, we arrive at the relativistic dispersion relationship for a massive particle,

$$E^2 - |\vec{p}|^2 = m^2 \tag{1.19}$$

We will denote the energy of a particle of mass m and 3-momentum $\vec{p}$ as

$$E_{\vec{p}} = \sqrt{|\vec{p}|^2 + m^2} \tag{1.20}$$

where we take the +ve square root. This quantity will feature *a lot* in this course! Another quantity that features prominently is the scalar product

$$p \cdot x = \eta_{\mu\nu} p^\mu x^\nu = Et - \vec{p} \cdot \vec{x} \tag{1.21}$$

A note on indices: throughout this course, we will employ the summation convention in which repeated indices are summed over. For the spacetime indices μ it will be crucial to keep track of whether they're up or down: you should never encounter expressions that look like $a^\mu b^\mu$. (For the other indices, such as the $a = 1, \ldots, n$ index which appear on the q_a, we don't need to be as careful, although it's good practice to try!). Also, it's worth stressing that repeated indices are *dummy* indices – it doesn't matter what you call them. But it's very important that you don't use the same pairs of dummy indices twice. For example, the expression $(a^\mu b_\mu)(c^\mu d_\mu)$ makes no sense – even with the brackets! Avoid rampant confusion by writing $a^\mu b_\mu \, c^\nu d_\nu$ to show which pairs are summed over. Mistakenly denoting multiple pairs of dummy indices with the same label will be a very easy trap to fall into in this course.

Fourier Transforms

We'll be frequently changing from position space to momentum space using the Fourier transform. As always, we have to decide where the factors of 2π sit. Our convention will be

$$f(x) = \int \frac{d^n k}{(2\pi)^n} \, \tilde{f}(k) \, e^{ik \cdot x} \tag{1.22}$$

so that the inverse reads

$$\tilde{f}(k) = \int d^n x \, f(x) \, e^{-ik\cdot x} \tag{1.23}$$

There's now no excuse for losing track of factors of 2π: they always accompany momentum integrals, never position integrals. Remember also that if we're doing a Fourier transform over spacetime (as opposed to just space) then $k \cdot x = k^0 x^0 - \vec{k} \cdot \vec{x}$.

Dirac Delta Functions

The Dirac delta function is defined by $\delta(x) = 0$ for all $x \neq 0$ and

$$\int_{-\infty}^{+\infty} dx \, \delta(x) = 1 \tag{1.24}$$

The Fourier transform of the delta-function provides a useful representation that we will make extensive use of:

$$\delta(x) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \, e^{ikx} \tag{1.25}$$

In an n-dimensional space, the delta-function is given by

$$\delta^{(n)}(x) = \int \frac{d^n k}{(2\pi)^n} \, e^{ik\cdot x} \tag{1.26}$$

Complex Functions

There are a few places in the course where we'll need to make use of the properties of complex functions. In particular, we'll need the residue theorem. Let Γ be a positively oriented (i.e. anticlockwise) simple closed contour within (and on) which a function $f(z)$ is analytic except for a finite number of singular points $z_1, \ldots, z_n$ in the interior of Γ. Then

$$\oint_\Gamma f(z) dz = 2\pi i \sum_{i=1}^{n} b_i \tag{1.27}$$

where b_i is the *residue* of $f(z)$ at the singular point z_i. The residue of a function with an isolated singular point z_0 is defined as the coefficient c_1 of the Laurent expansion of $f(z)$ about z_0,

$$f(z) = \sum_{n=0}^{\infty} (z - z_0)^n + \frac{c_1}{z - z_0} + \frac{c_2}{(z - z_0)^2} + \ldots \tag{1.28}$$

for $0 < |z - z_0| < R$, the radius of convergence.