Self-Study Notes on Quantum Field Theory

Based on David Tong's Lectures

Keanan Scarbro

June 1, 2025

Contents

1 Introduction

Fields were introduced in classical physics to resolve "spooky action at a distance" – the idea that one object can affect another object instantly, regardless of how far apart they are. In a local theory, like Maxwell's electrodynamics, an electron doesn't instantly feel the force from a distant proton. Instead, the proton modifies the electromagnetic field in its immediate vicinity, and changes in the field then propagate outward at the speed of light. The electron feels the force from the proton only when the changes in the field reach the electron. In more mathematically-precise language, a local theory is one wherein the equations governing an object's motion depend only on the fields evaluated at the particle's position, not on the position of any other object. For example, the force felt by the electron depends only on the electromagnetic field at the electron's position, not on the instantaneous position of the proton.

It's important to note that the fields aren't just a bookkeeping trick to hide non-locality. For a theory to be truly local, the fields themselves need to evolve locally – in electrodynamics, this is ensured by Maxwell's equations. Changes in the proton's position affect the field, but that influence propagates through the field at a finite speed. The field at the electron's position at time t depends on the proton's state only at an earlier time. By contrast, Coulomb's law describes an instantaneous interaction and is therefore non-local. A field defined by Coulomb's law alone simply hides this non-locality; it is only through the ∂_t terms in Maxwell's equations that genuine locality and finite propagation speed are enforced.

From a philosophical perspective, it might seem natural to you that the force a particle feels should only depend on its immediate surroundings, not some arbitrary point some distance away. How would an electron "know" about a proton located arbitrarily far away? Einstein's special relativity reinforces this viewpoint: if causal influences *could* propagate instantly, some reference frames would see the causal influence travel backwards in time, violating causality.

Over time, it became clear that fields are not merely devices used to enforce locality – they are physical, dynamical entities. One major piece of evidence for this is that electromagnetic waves can propagate in vacuum, carrying energy and momentum[†]. During the quantum revolution, it was also discovered that particles exhibit both wave-like and particle-like behavior. If particles like electrons (once thought of as fundamental) and waves like light (originally viewed as disturbances in a field) are to be treated on equal footing, there are two philosophical approaches:

- Treat particles as fundamental, and view fields as effective descriptions of a large number of particles
- Treat the fields as fundamental, where wave behavior is natural, and interpret particles as

[†]Of course, this energy and momentum can only be measured once the waves interact with something. The term "existence" is inherently slippery: what does it really mean for a field to be a "real, physical object"? In practice, we define the existence of the field through its effects on other things. But the fact that these fields can propagate indefinitely, even in the absence of sources, gives good reason to treat them as dynamical entities in their own right. After all, how is a particle any more "real"? Both particles and fields are considered real only insofar as they influence other objects. If a field can exist independently and exert observable effects, is there any meaningful philosophical distinction between a field and a particle?

quantized excitations of these fields

It turns out that the second option is far more powerful. This leads to quantum field theory – a framework in which particles are manifestations of underlying quantum fields.

2 Classical Field Theory

2.1 The Dynamics of fields

A field is a quantity defined at every point in spacetime. For example, a complex scalar field is a function

$$\phi : \mathbb{R}^4 \to \mathbb{C}, \quad \phi(x^\mu) = \phi(\vec{x}, t),$$

and a real vector field is a set of functions

$$A_\mu : \mathbb{R}^4 \to \mathbb{R}, \quad A_\mu(x^\nu) = A_\mu(\vec{x}, t), \quad \forall\ \mu = 0, 1, 2, 3.$$

That is, a vector field $A_\mu(x^\nu)$ assigns a value to each spacetime point x^ν for each vector component μ. The full field is then a collection of the four functions:

$$A(x^\nu) = (A_0(x^\nu), A_1(x^\nu), A_2(x^\nu), A_3(x^\nu)) \in \mathbb{R}^4.$$

2.1.1 Euler Lagrange Equations

The dynamics of the fields are governed by the principle of stationary action. Consider a set of fields ϕ_a. We construct the action as follows:

$$S : \{\phi_a\} \to \mathbb{R}, \quad S[\phi_a] = \int d^4x\ \mathcal{L}(\phi_a, \partial_\mu \phi_a),$$

where $\mathcal{L}$ is called the Lagrangian density and is, usually, a function of just the fields and their first derivatives (higher order terms can lead to issues, as we'll see later). The Lagrangian is given by the spatial integral of this density

$$L = \int d^3x\ \mathcal{L}(\phi_a, \partial_\mu \phi_a).$$

In field theory, the Lagrangian density is what appears in the action rather than the Lagrangian. This is because, in classical mechanics, our field was the position, and our only path parameters was time; in field theories, the fields depend on each point in spacetime (infinite degrees of freedom), so we have four path parameters, and we need to integrate over all of them. How else could we find a scalar action given fields that depend on all spacetime points. We assert that the physical field configuration is defined to be the one for which the action is stationary under variations of the fields.

As such, consider a set of fields ϕ_a to be the configuration for which the action is stationary. Consider infinitesimal variations of the fields ϕ_a written as: $\phi'_a(x^\mu) = \phi_a(x^\mu) + \epsilon\eta_a(x^\mu)$ for small ϵ and field variation η_a. The variations η_a are arbitrary except that they vanish at the boundary of the spacetime volume Ω we're considering (effectively making this a boundary value problem). To say that ϕ_a extremizes the action means:

$$\delta S = \frac{dS}{d\epsilon}\bigg|_{\epsilon=0} = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_\Omega d^4x \; \mathcal{L}(\phi_a + \epsilon\eta_a, \partial_\mu(\phi_a + \epsilon\eta_a))$$

$$= \int_\Omega d^4x \; \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} \mathcal{L}(\phi_a + \epsilon\eta_a, \partial_\mu\phi_a + \epsilon\partial_\mu\eta_a)$$

$$= \int_\Omega d^4x \; \left(\frac{\partial \mathcal{L}}{\partial\phi_a}\eta_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu\eta_a\right)^\dagger$$

$$= \int_\Omega d^4x \; \left(\frac{\partial \mathcal{L}}{\partial\phi_a}\eta_a - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\eta_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\eta_a\right)\right)$$

$$= \int_\Omega d^4x \; \left(\frac{\partial \mathcal{L}}{\partial\phi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\eta_a + \int_{\partial\Omega} d\Sigma_\mu\eta_a\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)},$$

where, in the last step, the divergence theorem is used to convert the last term to an integral over the boundary of the spacetime volume Ω. By construction, η_a vanishes at the boundary, so the last term vanishes. Thus,

$$\delta S = \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial\phi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\eta_a = 0.$$

Because η_a is arbitrary (subject only to vanishing boundary conditions), the integrand itself must vanish pointwise. Thus,

$$\frac{\partial \mathcal{L}}{\partial\phi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} = 0$$

must hold independently. These are the Euler-Lagrange equations for fields. Any field configuration that satisfies the variational principle must obey these equations.

For future reference, it's helpful to introduce some cleaner notation. The functional derivative of S is defined as

$$\frac{\delta S}{\delta\phi_a} \quad \text{such that} \quad \delta S = \int_\Omega d^4x \; \frac{\delta S}{\delta\phi_a}\delta\phi_a.$$

This derivative yields the unique function that determines the change in S under an arbitrary infinitesimal variation of the field ϕ_a.

Now, we've defined a general structure for a field theory: our physical objects are a set of fields ϕ_a. Given this set of fields, we compute the action S (an integral of a function of the fields, $\mathcal{L}$, over spacetime). The dynamics of ϕ_a are governed by the requirement that S remains stationary under small variations of ϕ_a. We have our physical objects, and we know how they evolve. Are

† This is a straightforward application of the chain rule, but, if you're like me, it looks a little odd compared to how you've maybe seen it before. If it helps, consider viewing $\mathcal{L}(\phi_a, \partial_\mu) \circ (\phi_a + \epsilon\eta_a, \partial_\mu\phi_a + \epsilon\partial_\mu\eta_a)$ as a composition of maps in function space. The Lagrangian density symbolically depends on any arbitrary field configuration. We're inserting a path in field space (a specific field configuration). This is analogous to evaluating a Lagrangian on a specific path in classical mechanics. The derivatives $\frac{\partial\mathcal{L}}{\partial\phi_a}$ and $\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}$ are formal derivatives of the arbitrary parameters of the Lagrangian density.

there any restrictions on what types of fields and evolution (inevitably governed by $\mathcal{L}$) we allow? Yes; there are two more properties of our theory that we require: locality and Lorentz invariance.

2.2 Locality

As mentioned earlier, locality – the principle that no influence can propagate instantly – is fundamental to modern physics and is one of the major motivations for introducing fields in the first place. The dynamics of a field at one point in spacetime should depend only on what's happening in its immediate vicinity.

Luckily, enforcing locality in a field theory is straightforward: it suffices to ensure that the Lagrangian density depends only on the fields and their derivatives at a single spacetime point. In contrast, a non-local theory might involve an action of the form:

$$S[\phi_a] = \int d^4x \int d^4y \; \phi_a(x^\nu) K(x^\nu, y^\nu) \phi_a(y^\nu),$$

where $K(x^\nu, y^\nu)$ is a function that couples field values at two distinct spacetime points x^ν and y^ν. This explicit coupling of different spacetime points immediately makes it clear that the theory isn't local, but the non-locality becomes even clearer when one derives the Euler-Lagrange equations for this theory:

$$\int d^4y \; K(x^\nu, y^\nu) \phi_a(y^\nu) = 0.$$

This is an integral equation. The behavior of ϕ_a at spacetime point x^ν depends on its value at all other spacetime points y^ν. Thus, the theory requires global information to determine local dynamics.

By contrast, consider the theory with action

$$S[\phi_a] = \int d^4x \; \mathcal{L}(\phi_a, \partial_\mu \phi_a)$$

yielding the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} = 0.$$

Here, the Lagrangian density depends only the fields and their derivatives evaluated at a single spacetime point x^ν. As a result, the Euler-Lagrange equations are differential equations, which are inherently local: they determine the behavior of ϕ_a at x^ν based solely on how the fields change in an infinitesimally small neighborhood around that point. So, as long as we don't introduce terms that couple two distinct spacetime points in the Lagrangian density (we have no motivation to do this anyway), then our theory is manifestly local.

2.3 Lorentz Invariance

The principle of relativity states that all physical laws should be frame independent. From a philosophical perspective, this seems reasonable: if the universe truly is homogeneous and isotropic, it shouldn't at all favor a specific frame. Mathematically, this idea is expressed by ensuring that the equations of motion are frame-independent. The equations of motion for my field in my frame better be the same equations of motion for your field in your frame. How are two frames related? Lorentz transformations. $x'^{\nu} = \Lambda^{\nu}_{\mu} x^{\mu}$. There There are two ways to look If the field is truly to be a physical object, then $\phi'(x'^{\nu}) = \phi(x^{\nu})$. That is, my field at spacetime point x^{ν} must equal your field at your spacetime point x'^{ν} (where x'^{ν} is the spacetime point that you see matching the spacetime point that I see). Further, both of our fields must evolve the same. This isn't just recasting the equation in a different set of coordinates (that would be just shifting $x^{\nu} \to x'^{\nu}$. Instead, we're fundamentally treating no frame as fundamental. Measure ϕ from any frame, and you'll see the same result. ϕ doesn't exist in any one frame. Now, we ask: how do frames shift? How is x'^{ν} related to x^{ν}. The answer is the Lorentz transformation. Special relativity tells us that, to preserve the constancy of the speed of light, spacetime coordinates must be transformed in a certain way, so that $x'^{\nu} = \Lambda^{\nu}_{\mu} x^{\mu}$. If I measure an event at x^{μ}, you'll see that event at x'^{ν}.

To ensure that a theory is Lorentz invariant, it only requires that we ensure that the action is invariant. If the action is invariant under Lorentz transformations, then of course the dynamical equations will be Lorentz invariant (its a plus that the action is itself a Lorentz scalar. Quantities that remain completely unchanged in different frames are of importance. The action remains completely unchanged). In our case, we'd like to see how a field transforms in a Lorentz transform. An active transformation is one that directly shifts the coordinates. We're looking

Passive transformation

Active transformation

So far, I've been using Lorentz notation without really explaining what's going on. To be precise, the Minkowski metric is ...; contraction of indices is defined as $a^{\mu} b_{\nu} = a^{\mu} \eta^{\rho\nu} b_{\rho}$. The Minkowski metric defines the geometry of spacetime. Through contraction, the Minkowski metric defines an inner product that is different from that in Euclidean physics.

2.4 Symmetries and Noether's Theorem

2.4.1 Internal Symmetries

2.5 Hamiltonian Formalism

2.5.1 Legendre Transform

2.5.2