Quantum Field Theory

quinten tupker

October 7 2020 - October 25, 2020

Introduction

These notes are based on the course lectured by Professor Nicholas Dorey in Michaelmas 2020. This was lectured online due to measures taken to counter the spread of Covid-19 in the UK. These are not necessarily an accurate representation of what was lectures, and represent solely my personal notes on the content of the course, combinged with probably, very very many personal notes and digressions... Of course, any corrections/comments would be appreciated.

But, let's actually introduce the content of this course. What is quantum field theory? Quantum field theory (QFT) essentially succeeds in merging special relativity and quantum mechanics. Why is this so difficult? The first relativistic theory was electromagnetism, and the biggest idea that was introduced, and what truly set it apart from older theories was the idea of fields. In Newtonian theory, and much of what followed, the protagonist of the theory was a particle, or if not a particle, at least a body of some kind. Field theory changed. Now, there was a new protoganist: the field. What difference does that make, architectually? Firstly, the field theory is often simpler. But more importantly, the biggest structural difference is that field theory excels in describing "delayed interactions." When the particle is the progagonist, it is very difficult describe theories where forces are not instantaneous. Field theory avoids this. All interactions are made through the field, through which they propogate through space. Now, delayed responses become natural. In electromagnetism, the simplest expression thereof is electromagnetic waves: light.

What does this have to do with relativity? Well, as soon as high speeds become relevant, forces can no longer be considered instantaneous. As such, it is difficult to keep using particles as the protoganist of these theories. Consequently, the natural step is to make, instead of particles, fields the protoganist of this new quantum theory we are developing.

That is the goal of QFT. There is one important consequence though, once particles are no longer the protagonists of the theory. That is that particle number no longer has to be conserved. In the most elegant fashion, by removing the supremacy of the "particle" in our theory, and replacing it with the more powerful notion of the field, particles merely become phenomenon to be observed, and tools of analysis. In this context, it is only natural that particle number is no

longer conserved. Whereas before, the wavefunction was often associated with a wave-particle like object, now the wavefunction (which is a field) describes a multiparticle state. Well, really it describes the field, and the multiparticle state is something that can be deduced from it. Somehow, although this is just the beginning of the course, I feel that that's not that important anymore. It is deeply intriguing though, how the imposition of boundary conditions somehow forces a degree of discreteness onto this theory...

Well then, the overall architecture is more or less the same as standard quantum theory. It is probabilistic, and we assume a degree of symmetry under boosts, and rotations (isotropy and translation invariance). The fundamental approach to making predictions still boils down to the same calculation: evaluating

$$A_{i\to f} = \langle f| \, e^{iHT} \, |i\rangle$$

for probability amplitude A, initial state i, final state f, Hamiltonian (time translation generator) H, and time interval T.

There are two caveates with most of these field theories, though. Firstly, they have not been mathematically formalised, so often there are areas that are somewhat ambiguous. Secondly, contributing to this ambiguity, many of the sums are divergent, so the meaning of some calculations can really be somewhat ambiguous... How curious! I'd like to think about this a bit more...

1 Preliminaries

Anyways, getting down to business. We'll be mostly using natural units during this course. That means that $c = \hbar = 1$, and these can be added back into the calculation using dimensional analysis. The effect of this, is that the only unit used throughout all calculations is really a unit of mass-energy. As such, all quantities scale by a power of the unit of energy.

Definition 1 (dimension of X)**.** Denoted $[X]$, this is δ, such that for unit of mass-energy M, X scales as M^δ. δ may also be called the scaling or the engineering dimension of X.

Also, for special relativity we use the convention that we are working on Minkowsky space-time $\mathbb{R}^{3,1}$, with metric tensor

$$\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1).$$

2 Classical Field Theory

Fortunately for me, we are starting with a description of classical field theory, and we are starting, very simply, with scalar fields.

Definition 2 (Scalar Field)**.** $\phi(t, \underline{x}) = \phi(x) : \mathbb{R}^{3,1} \to \mathbf{R}$ is a sclar field if it is Lorentz invariant, meaning that it follows the transformation rule

$$\phi(x) \to \phi(\Lambda^{-1}(x))$$

Here the domain is called the spacetime, and the codomain is called the **field space**. These may, and will be replaced with other spaces.

Changing the spacetime here corresponds to implementing gravity in some way or another, by changing the manifold we are working on. The field space corresponds to the complexity of what is being descibed. Since we will be describing multi-particle states with our wavefunction, this will become significantly more complex. And, I intentionally left the description of what it means to be Lorentz invariant a bit vague, since, well, in our case, it just means being invariant under the Lorentz transformations, which is the group of linear transformations that preserves the Minkowsky metric (ie. the set of matrices such that $\Lambda\eta\Lambda^T = \eta$). But really, while linearity makes a lot of sense in the context of linear Minkowsky space, I doubt (though I have no familiarity with this area) this remains the case when we are on an arbitrary manifold, which happens when we consider general relativity. As such, I prefer to think of Λ as any arbitrary invertible map on the manifold, corresponding to the symmetries we impose.

The difficulties that arise in comibing quantum theory with general relativity are also quite clear. It does seem tremendously difficult. If we do not simply assuming that general relativity simply bends space, which I will assume is not entirely the case, or else I feel a theory reconciling the two would have already been developed long ago, in spite of the tremendous difficulty of the calculations involved, and it also does not seem to make much sense of Hawking radiation, since in general relativity, the space at the centre of a black hole truly is cut off... Nevertheless, from what I've heard, if you want to turn gravity into a quantised force with mediator particle, then somehow the protagonist of that theory would be not only be a field, but somehow span the space of possible manifolds as well. Purely intuitively, I would imagine that we would be getting fields of the form $\phi : \mathcal{D} \to V$ where $\mathcal{D}$ is an object that stitches many manifolds together. Brrr... I have not thought too deeply about this, but that does seem like a truly terrifying object indeed! Or perhaps not quite. Hm, it might be worth thinking a bit more about this...

Anyways, I also wanted to remark that having fields transform as $\phi(\Lambda^{-1}x)$ is more or less an arbitrary definition that is called the "active" definition of field transformations. Oh well, you can assign some intuition to it, but it is more or less convention to use the inverse of the matrix instead of the matrix itself.

Extending our notion of fields to vector fields, first note that notation wise, we use $\partial_\mu \phi = \frac{\partial \phi}{\partial x^\mu}$, and we define

Definition 3 (ϕ^μ transforms as a vector field)**.** if

$$\phi \to \Lambda^\mu_\nu \phi^\nu(\lambda^{-1}x)$$

This is just the transformation rule for rank 1 tensors, so is nothing particularly remarkable. The only remarkable part is that, as a result $\partial^\mu \phi$ transformas as vector, and so the following becomes a rank 0 tensor (ie, a scalar field) $\partial^\mu \phi \partial_\mu \phi$.

Well, that ends [lecture 1], for those of you interested.

2.1 The Lagrangian

We will review some of the tools from classical mechanics that we are using. The first is the Lagrangian. There are three advantages to using the lagrangian here:

- they are independent of coordinates
- the symmetries of the system can be easily expressed
- the path integral formulation follows immediately

although the Lagrangian is used much more heavily in the Advanced Quantum Field Theory course than the current course. Anyways, we review the Lagrangian, except that now we implement it on fields, so $L = L(\phi, \partial_\mu \phi)$ for (scalar) fields ϕ, and we make three requirements of our action/Lagrangian

- The action S is Lorentz invariant
- we require locality (see below)
- we have at most a second order time derivative

Locality essentially means that $L = \int dx^3 \, \mathcal{L}(\phi(x), \partial_\mu \phi(x))$, so the Lagrangian has an associated local Lagrangian density $\mathcal{L}$. From hereon, this $\mathcal{L}$ will usually be referred to as the Lagrangian instead. We note that the action (over an infinite time scale) may now be expressed as

$$S = \int dx^4 \mathcal{L}(\phi, \partial_\mu \phi)$$

Now, Lorentz invariance essentially means that the Lagrangian density is a scalar field (a tensor really), meaning that it transforms as $\mathcal{L}(x) \mapsto \mathcal{L}(\Lambda^{-1} x)$. This corresponds to the action being Lorentz invariant since the Jacobian of a Lorentz transformation (the determinant) is 1, so the relevant change of variables, the action does not change.

Finally, using no more than a second order time derivative in the context of relativity means using no more than a second order derivative of any kind, and since we also are rank-0 tensors, we in fact, cannot have first order derivatives either. Consequently, we can write the general form of the Lagrangian as

$$\mathcal{L} = \frac{1}{2} F(\phi) \partial_\mu \phi \partial^\mu \phi - V(\phi)$$

(note that $\phi\partial_\mu\partial^\mu\phi$ is related to the above by integration by parts so can be safely ignored). Also, in practice in quantum theory we may neglect $F(\phi)$ leaving us with quite a simple general form.

To make all this work, we apply the principle of least action, which gives us the Euler Lagrange equation

$$\partial_\mu\partial_{\partial_m u\phi}\mathcal{L} = \partial_\phi\mathcal{L}$$

An important special case to consider is the case when the potential is quadratic, so $V(\phi) = \frac{1}{2}M^2\phi^2$, which means the Euler-Lagrange case is linear, leaving us with the so-called **Klein-Gordon Equation**

$$\partial_\mu\partial^\mu\phi + M^2\phi = 0$$

As one might expect, since in Minkowski spacetime $\partial_\mu\partial^\mu = \partial_t^2 - \nabla^2$ we get wavelike solutions

$$\phi \sim e^{ix\cdot p}$$

where $x\cdot p = \omega t - \vec{k}\cdot\vec{x}$, and the dispersion relation requires $\omega_k = \sqrt{k^2 + M^2}$.

On a philosophical note, I was wondering what difference using the principle of least action makes compared to just Newton's equations (on a philosophical level - it is obviously more practical), and as such I was wondering to what extent the Lagrangian formulation in a sense is just a tensor formulation of Newton's equations?

I also was wondering what locality means. When discussing life, I have remarked that really as long as one is not starving, etc., reality is little more than a medium for communication between people. This is certainly the case for particles, and objects, which do not have to worry about starvation, etc. But then I wonder, do particles really not need to worry about starvation? Many particles decay after all, although I doubt that has to do with a kind of starvation of any kind... [End of lecture 2]

2.2 Maxwell's Theory

Maxwell's theory uses a 4-vector potential $A^\mu = (\phi, \vec{A})$. This being a rank 1 tensor means it transforms as

$$A^\mu(x) \mapsto \Lambda^\mu_\nu A^\nu(\Lambda^{-1}x)$$

In electromagnetism, the **field strenght tensor** is given by

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

and as a tensor, it transforms as appropriate. But it satisfies a further condition that it is invariant under Gauge transformations $A^\mu \mapsto A^\mu + \partial^\mu\lambda$. This leads to the **Bianchi identity**

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0$$

which captures 2 out of Maxwell's 4 equations. The other two arise from the principle of least action applied to the **Maxwell Lagrangian**

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

We should note that there really are not that many options of Lagrangians here once one introduces the standard assumptions in addition to assuming it must be expressed in terms of physical quantities (the only candidate is the field strength vector used above)... Writing

$$\mathcal{L} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu + \frac{1}{2}(\partial_\mu A^\mu)^2\eta^{\mu\nu}$$

gives us (after Euler-Lagrange)

$$\partial_\mu F^{\mu\nu} = 0$$

which provides us with the rest of the Maxwell equations.

2.3 Symmetries in QFT

Symmetries (see the Symmetries, Fields, and Particles (SFP) course) are variations of fields that leave the action invariant. Symmetries have several effects in physics. Firstly, by Noether's theorem, they lead to conservation laws. Secondly, they restrict the form of the Lagrangian. In particular, Gauge symmetry is very powerful due to strong restrictions it puts on the form of the Lagrangian.

Common symmetries include (See SFP) translation invariance and Lorentrz transformation invariance. Another symmetry occurs when either $m = 0, V = 0$ or the action is propoertional to the field. In this case we get an additional "scale invariance" where $x^\mu \mapsto \lambda x^\mu$ and $\phi \mapsto \lambda^{-\Delta}\phi(\lambda^{-1}x)$ for engineering constant Δ.

Internal symmetries like charge, flavour or colour conservation also occur, further restricting the theory. These are not related to the space, and are somehow better expressed by the quantity they conserve. Another small curious symmetry is the following:

Example 1. For complex scalar fields with $\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi - V(|\psi|^2)$ the following is also a symmetry: $\psi \mapsto e^{i\alpha}\psi$ for real α. The lecturer does not go into any more detail about this.

Finally, we note that for continuous symmetries, which form Lie groups, we can take the set of elements of the group "near" the identity, $g = e^{\alpha X}$ where X sits in the Lie algebra of G. Working to 1st order we now may write

$$g \cdot \psi \mapsto \psi + \delta\psi = \psi + \alpha X\psi$$

to first order. [End of lecture 3]

Just as a remark on terminology, a Lorentz transformation is a linear transformation preserving the Minkowski metric. On the other hand, a **proper Lorentz transformation** is a Lorentz transformation with determinant 1 (which

ensures causality). Our goal now is to work towards Noether's theorem. Noether's theorem states that

Theorem 1 (Noether's Theorem)**.** For every continuous symmetry, there exists a conserved current, given by

$$j^\mu = \partial_{\partial_\mu \phi} \mathcal{L} \delta\phi - F^\mu(\phi(x))$$

where $\delta\phi$ is the variation in ϕ and $\delta\lambda = F^\mu$. Conservation here means that $\partial_\mu j^\mu = 0$.

The conserved quantity associated with the conserved current then can be shown to be $Q = \int dx^3 j^0$ since

$$\begin{aligned}
\frac{d}{dt}Q &= \int dx^3 \partial_t j^0 \\
&= -\int dx^3 \nabla \cdot J \\
&= \int_S dS \cdot J \\
&= 0
\end{aligned}$$

A good example is $j^\mu = (\rho, J)$ for charge density and current in electromagnetism. Anyways, here is our "proof" of Noether's theorem:

Proof. Using the Euler-Lagrange equation applied to the field ϕ we see that

$$\begin{aligned}
\delta\mathcal{L} &= \partial_\phi \mathcal{L} \delta\phi + \partial_{\partial_\mu \phi} \mathcal{L} \delta \partial_\mu \phi \\
&= (\partial_\phi \mathcal{L} - \partial_\mu(\partial_{\partial_\mu \phi} \mathcal{L}))\delta\phi + \partial_\mu(\partial_{\partial_\mu \phi} \mathcal{L} \delta\phi) \\
&= \partial_\mu(\partial_{\partial_\mu \phi} \mathcal{L} \delta\phi)
\end{aligned}$$

Then using our definition of j^μ as above we see that $\partial_\mu j^\mu = 0$. □

Here it is assumed that $\delta\mathcal{L} = \partial_\mu F^\mu$ which means we can apply Stokes' theorem and integrate on the surface only. When we do so, F^μ becomes that surface term we integrate with. But Euler-Lagrange also means that our expression above, for the first term in the conserved current, is also a surface term (often determined by the boundary conditions) for $\delta\mathcal{L}$. So what exactly is the difference? Is it that the first term depends on the variation whereas the second does not?

[End of lecture 4 - this bit needs rewriting]

[Start of Noether's theorem rewrite]

Alright, so now to the main topic of this lecture: Noether's theorem. Noether's theorem states that for every continuous symmetry we get a conserved current (and quantity), which she gives an equation for. Now how do we get that?

Firstly, we observe that for a Lorentz transformation

$$\Lambda \approx I + s\Omega$$

where Ω is antisymmetric (to ensure we preserve the Minkowski metric). If that the case, we see that if we tree $\mathcal{L}$ as a scalar field $\mathcal{L}(x)$ instead of a function of the field, then

$$\delta\mathcal{L} = -s\Omega^{\mu}_{\nu}x^{\nu}\partial_{\mu}(\mathcal{L}(x)) = -s\Omega^{\mu}_{\nu}\partial_{\mu}(x^{\mu}\mathcal{L}(x)).$$

This last equality only holds because Ω is asymmetric (so its diagonal is all zeros). But importantly, we see that $\delta\mathcal{L}$ can be written as a total derivative when x is varied.

Now, we vary the same $\mathcal{L}$ but instead of varying with respect to x, we vary with respect to the field ϕ. Doing so, and assuming Euler-Lagrange we find that:

$$\begin{aligned}\delta\mathcal{L} &= \partial_{\phi}\mathcal{L}\delta\phi + \partial_{\partial_{\mu}\phi}\mathcal{L}\delta\partial_{\mu}\phi \\ &= (\partial_{\phi}\mathcal{L} - \partial_{\mu}(\partial_{\partial_{\mu}\phi}\mathcal{L}))\delta\phi + \partial_{\mu}(\partial_{\partial_{\mu}\phi}\mathcal{L}\delta\phi) \\ &= \partial_{\mu}(\partial_{\partial_{\mu}\phi}\mathcal{L}\delta\phi)\end{aligned}$$

Remarkably, we find once again that we may write $\delta\mathcal{L}$ as a total derivative. Let us now denote these variations as $\delta_x\mathcal{L}$ for the first, and $\delta_{\phi}\mathcal{L}$ for the second. Now, since we have a symmetry, we know that the action is unchanged when we apply a Lorentz transformation. In particular, we know that the "partial variation" of $\mathcal{L}$ with respect to x alone, but ignoring ϕ should be "constant." What do we mean by constant? The partial derivative with respect to space is a 4-vector, so really we mean that the Minkowski metric is unchanged. In other words, to first order the quantity

$$v_M \cdot \partial_x\delta\mathcal{L} = 0$$

where we use $\partial_x\delta$ to denote this "pure" partial derivative and

$$v_M = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is the "Minkowski vector" (the first order derivative of the Minkowski metric in a sense). This would mean that the Minkowski metric is unchanged under a pure variation of spacetime. So how do we write this pure variation of space time? We notice that $\delta\partial_x\mathcal{L} = \delta_x\mathcal{L} - \delta_{\phi}\mathcal{L}$ since in a sense $\delta_x\mathcal{L}$ is the total derivative of $\mathcal{L}$ with respect to x, and $\delta_{\phi}\mathcal{L}$ is the total derivative of $\mathcal{L}$ with respect to ϕ, and so, by the chain rule, what is left over is the partial variation with respect to x.

This is exactly what Noether's theorem refers to. Noether's theorem defines the **conserved current** to be

$$j^\mu = \partial_{\partial_\mu \phi} \mathcal{L} \delta\phi - F^\mu$$

where $\delta_x \mathcal{L} = \partial_\mu F^\mu$, and $\delta_\phi \mathcal{L} = \partial_\mu(\partial_{\partial_\mu \phi} \mathcal{L} \delta\phi)$, and so we see that our partial variation of $\mathcal{L}$ with respect to only x is

$$\delta \partial_x \mathcal{L} = -\frac{\partial}{\partial x_\mu} j^\mu$$

(no summation convention - this is a vector here), and importantly

$$\partial_\mu j^\mu = v_M \cdot \delta \partial_x \mathcal{L} = 0$$

as required since $\partial_\mu j^\mu = \delta\mathcal{L} - \delta\mathcal{L} = 0$.

That is Noether's theorem. I do not fully understand it, and to me it seems weird that the "generators" of these variations (the terms like F^μ inside the derivative) are not the same. But anyhow, Noether's theorem states that j^μ is conserved, so $\partial_\mu j^\mu = 0$ (which means the Minkowski metric stays the same in the way we described). Aside from that we will remark that we assume in general that for any symmetry $\partial_x \mathcal{L} = \partial_\mu F^\mu$ for some F^μ,

Now where does a conserved quantity arise from. Here we note that if $Q = \int dx^3 j^0$ then it satisfies

$$\begin{aligned} \frac{d}{dt} Q &= \int dx^3 \partial_t j^0 \\ &= -\int dx^3 \nabla \cdot J \\ &= \int_S dS \cdot J \\ &= 0 \end{aligned}$$

and os is conserved. One can check that in the case of electromagnetism, Q corresponds precisely to electromagnetic charge.

[End of Noether's theorem rewrite]

Today we will look at certain examples of Noether's theorem. In particular, we will define the energy-momentum constant as the Noether current arising from translation invariance. Under the symmetry $x^\mu \mapsto x^\mu + \epsilon^\mu$, we get $\phi \mapsto \phi(x - \epsilon) \approx \phi(x) - \epsilon^\mu \partial_\mu \phi + \dots$. We also get $\delta\mathcal{L} = -\epsilon^\mu \partial_\mu \mathcal{L}$. Consequently we get energy-momentum tensor

$$T^\mu_\nu = j^\mu_\nu = \partial_{\partial_\mu \phi} \mathcal{L} \partial_\nu \phi - \delta^\mu_\nu \mathcal{L}$$

As Noether current, this is conserved, meaning $\partial_\mu T^\mu_\nu = 0$. We consequently get conserved quantities:

$$E = \int d^3 x T^{00}$$

$$p^i = \int d^3x T^{0i}$$

or equivalently, the 4-momentum is conserved:

$$p^\nu = \int d^3x T^{0\nu}.$$

[Lecturer computes example in Klein-Gordon case.] In many cases, such as the Klein-Gordon case we find the energy-momentum tensor to be symmetric, but this need not be the case in general. However, just as the Lagrangian gives rise to the same physics if we add a total derivative, one can show that the energy-momentum tensor has a similar symmetry, and if a particular choice of total derivative is made, we find that we can always make $T^{\mu\nu}$ symmetric.

More specifically, we find that if we use the formula from general relativity for the energy-momentum tensor:

$$T_{\mu\nu}(x) = \frac{-2}{\sqrt{-g}} \partial_{g^{\mu\nu}} (\sqrt{-g}\tilde{\mathcal{L}})$$

for $g = \det(g_{\mu\nu})$ then we always get a symmetric result. By using the Minkowski metric we can recover our original result.

3 Canonical Quantisation

In the Advanced Quantum Field Theory course in Lent, we will use the Lagrangian directly to formulate the path integral version of quantum mechanics. Here, however, we will stick to the Hamiltonian approach. For that we do a small review of Hamiltonian mechanics by starting with the Lagrangian

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^2 - V(q)$$

We then define the **momentum conjugate to** q to be $p = \partial_{\dot{q}} L$. Performing the Legendre transform we then can define the **Hamiltonian** to be

$$H = p\dot{q} - \mathcal{L} = \frac{1}{2}p^2 + V(q) = H(p, q)$$

since H is really seen as a functino of p and q. It is also clear that Hamiltonian can be seen as the total energy of a point in state space (p, q).

Now, we generalise our system to N particles as

$$H = \sum_{i=1}^{N} p_i q_i - \mathcal{L}$$

and introduce the **Poisson bracket** (which is related to the commutator) as

$$\{F, G\} = \sum_{i=1}^{N} \partial_{q_i} F \partial_{p_i} G - \partial_{p_i} F \partial_{q_i} G$$

for functions F, G of p, q. We can then get the following important result: the Hamiltonian is the generator of time evolution, meaning that $\forall F(p, q)$ we find

$$\dot{F} = \{H, F\}.$$

An important special case of this is that we can write the principle of least action by applying the time evolution property to p and q asking

$$\begin{aligned} \dot{q}_i &= \{H, q_i\} = \partial_{p_i} H \\ \dot{p}_i &= -\partial_{q_i} H \end{aligned}$$

and for any conserved Q, we get $\{H, Q\} = 0$. Finally, we note that these also satisfy the property that

$$\{q_i, p_j\} = \delta_{ij}$$

[End of lecture 5]

How do translate this Hamiltonian to field theory? Firstly we define the **momentum conjugate to the field** as

$$\pi = \partial_{\dot{\partial}_\phi} \mathcal{L}$$

and we apply the Legendre transform to get the **Hamiltonian density**

$$\mathcal{H} = \pi \dot{\phi} - \mathcal{L} = \mathcal{H}(\phi, \pi)$$

and, as expected the normal Hamiltonian can then be calculated as

$$H = \int d^3x \mathcal{H}.$$

As a check, and example, if we use $\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi)$ (a "suitably" general Lagrangian) we get

$$\mathcal{H} = \frac{1}{2} \dot{\phi}^2 + \frac{1}{2} |\nabla \phi|^2 + V(\phi) = T^{00}$$

the energy density, as one might hope.

3.1 The Canonical Quantisation

Now to actually quantise things we convert the Poisson bracket $\{\}$ to the commutator $\frac{1}{i\hbar}[]$, with overall quantities staying the same. That means that for π, ϕ we get that

$$[\pi_i, \pi_j] = [\phi_i, \phi_j] = 0$$

but that

$$[\pi_i, \phi_j] = i\hbar\delta(x - y),$$

as one might expect. The time evolution property of H becomes

$$H |\psi\rangle = -i\hbar\partial_t |\psi\rangle .$$

In particular that means that for any conserved quantity (commutes with H), we can simultaneously diagonalise these to get a simultaneous basis H and the conserved quantity. Also, note here that we later set $\hbar = 1$.

Anyways, using the standard Hamiltonian approach we then find that

$$H = \int d^3x \frac{1}{2}\pi^2 + \frac{1}{2}|\nabla\phi|^2 + V(\phi)$$

In practice, this is very hard to calculate, so not very helpful (one would have to diagonalise this function to find the energy eigenstates...). Instead we find some other ways.

Finally, on interpretation, we note that the eigenvalues have become functions $f : \mathbb{R}^3 \to \mathbb{R}$, which we assume span the Hilbert space as usual. However, this has never been formalised mathematically, so it is a bit of a grey area... [End of lecture 6]

In general, solving this Hamiltonian works out to be a functional differential equation, which is predictably hard. Instead, we tend to work with special cases. That is how we will proceed.

3.2 Free Field Theory

The simplest case is to assume everything is non-interacting, and that the Lagrangian obeys the Klein-Gordon equation. That means we are left with

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 \implies \partial_\mu\partial^\mu\phi + m^2\phi = 0$$

We can solve this using a Fourier transform to get

$$(\partial_t^2 + |p|^2 + m^2)\tilde{(\phi)} = 0$$

which has solutions for $\omega_p = \sqrt{|p|^2 + m^2}$ of the form

$$\tilde{\phi} = A_p e^{-i\omega_p t} + B_p e^{-i\omega_p t}$$

which form the modes of a Harmonic oscillator. In fact, calculating the action we find

$$S = \frac{1}{2}\int dt \int \frac{d^3p}{(2\pi)^2}\tilde{\phi}^*(-\partial_t^2 - |p|^2 - m^2)\tilde{\phi}$$

which means really do get a infinite number of decoupled complex simple harmonic oscillators. This is essential, since it is this simplicity locally that gives us a chance at solving the problem at all!

3.2.1 The Quantum Simple Harmonic Oscillator (review)

Firstly, though, let's review the quantum simple harmonic oscillator (SHO) so that we can use it for comparison. In this case

$$L = \frac{1}{2}\dot{q}^2 - \frac{\omega^2}{2}q^2 \implies H = \frac{p^2}{2} + \frac{\omega^2}{2}q^2$$

We can quantise this as usual for p, q, and then solve it using general methods. The unique aspect of the SHO though is that we can use purely its algebra to gain a solution by considering the raising and lowering/ladder operators

$$a = \sqrt{\frac{\omega}{2}}q + \frac{i}{\sqrt{2\omega}}p, a^\dagger = \sqrt{\frac{\omega}{2}}q - \frac{i}{\sqrt{2\omega}}p$$

where we find $[a, a^\dagger] = \hbar$, which crucially means that

$$H = \frac{1}{2}\omega(aa^\dagger + a^\dagger a) = \omega(a^\dagger a + \frac{1}{2}\hbar)$$

where one can identify $N = a^\dagger a$ as the number operator. Even more importantly we see that

$$[H, a] = -\omega\hbar a, [H, a^\dagger] = \omega\hbar a^\hbar$$

The fact that the commutator is still proportional to $a, a^\dagger$ means that when applied to an energy eigenstate, $a, a^\dagger$ either annihilate the state, keep the same state, or create a new state. As such, we create the ladder operators as usual, and we find a state $|0\rangle$ such that $a|0\rangle = 0$, but we can define an inifite ladder of states $|n\rangle = (a^\dagger)^n|0\rangle$. The energies of these states are given by $E_n = \hbar\omega(n + \frac{1}{2})$.

3.2.2 Back to Field Theory

To generalise to field theory we might define our ladder operators as:

$$\phi = \int \frac{d^3p}{(2\pi)^2}\frac{1}{\sqrt{2\omega_p}}\left(a_p e^{ip\cdot x} + a_p^\dagger e^{-ip\cdot x}\right)$$

$$\pi = \int \frac{d^3p}{(2\pi)^2}(-i)\sqrt{\frac{\omega_p}{2}}\left(a_p e^{ip\cdot x} - a_p^\dagger e^{-ip\cdot x}\right)$$

If we gain the same commutation relations, from these definitions then we can call ourselves satisfied that these definitions have the properties what we want. As such, we wish to show that the following commutation relations

$$[\phi(x), \phi(y)] = [\pi(x), \pi(y)] = 0, [\phi(x), \pi(y)] = i\delta(x-y)$$

if and only ifoddpage

$$[a_p, a_q] = [a_p^\dagger, a_q^d agger] = 0, [a_p.a_q^\dagger] = i\delta(x-y)$$

Some algebra shows that indeed this is the case.

Having shown that these definitions can make sense, we want to see if we can calculate the energy eigenstates from here, as we do in the non-field theory case. As such we calculate

$$H = \frac{1}{2}\int d^3x(\pi^2 + |\nabla p|^2 + m^2\phi^2)$$

by writing every term in terms of $a, a^\dagger$ and find that as expected

$$H = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega_p(a_p a_p^\dagger + a_p^\dagger a_p)$$

[End of lecture 7] This calculation can be simplified by creating and eliminating delta functions as soon as possible. So our first step here would be to calculate the ground state energy, which in the context of field theory is called the "vacuum" (or the energy of the vacuum). Doing so for $|0\rangle$ such that $a|0\rangle = 0$ we find that

$$H|0\rangle = \frac{1}{2}\int d^3p\omega_p\delta(0)|0\rangle = \infty|0\rangle$$

which points us to a general aspect of field theory: many things diverge. How do we get around this? There are two types of divergence present here, which are called **infrared divergence** (IR divergence) and **ultraviolet divergence** (UV divergence). IR divergence refers to divergence that occurs due to the large distances (but possibly low energies) involved. Ultraviolet divergence on the other hand occurs due to the high energies (but often short distances) involved. The infrared divergence is relatively simple to solve: we consider the energy density $\mathcal{E}$ instead of the total energy, which makes sense since almost everything else we're working with is a density. Formally this can be expressed as restricting the energy to a box of limited size.

That solves the IR divergence, leaving us with

$$\mathcal{E}_0 = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2}\omega_p \sim \int |p|^3 d|p| = \infty$$

(where exactly does this come from?), but we still have UV divergence. To resolve this, we use a rather crude UV cutoff, which is an energy cutoff, limiting our integrals to only energies lower than $\Lambda >> 1/m$ where $1/m$ is the only

natural length scale available to us in this context. Calculating $\mathcal{E}_0$ in this limit for Λ gives us

$$\mathcal{E}_0^\Lambda = \frac{1}{16\pi^2}\Lambda^4(1 + O(m^2/\Lambda^2)).$$

An alternative, and more intuitive approach would be to replace continuous spacetime with a lattice. This does work, and gives a similar result in fact, but is much harder to do. The equivalence between this relates to a somewhat profound relationship between small distances and high energies, and in fact the lattice spacing a is inversely proportional to our energy cutoff in a certain sense.

Some interpretations can be offered for the situations we get here. One involves the fact that we don't really know what happens at high energies, and so we focus on where we know our field theory works (referred to as **effective field theory**). Another approach is that the method that we quantise our theory is more or less arbitrary and so we could instead use a different quantisation called the "natural ordering" which forces all terms involving $a, a^\dagger$ to put as in front, leaving $\mathcal{E}_0 = 0$. I don't know how well this works in general.

Finally, another approach to this issue is simply not to care, since why should inifinite energies really make a difference (closest to my attitude to the situation). After all, we can only really measure energy differences in experiment. That's not entirely true though, since once one starts to incorporate gravity, one sees that through the energy-momentum tensor relativity does introduce a certain absolute measure of energy... This relation remains a relatively mysterious aspect of physics. [End of lecture 8]