

Introduction to Quantum Field Theory

Darsh KODWANI

May 31, 2021

Abstract

These are notes i have made as i started learning QFT in 2013 during my undergraduate. ¹

¹These notes are strongly inspired by the Perimeter Scholars International course of QFT

Contents

1	Introduction	4
1.1	Units and conventions in QFT	5
1.2	Notation	6
2	Field theory	12
2.1	Lagrangian and equation of motion	13
2.2	Application of Lagrangian and EL equation	14
2.3	Symmetries and Noethers theorem	17
3	Quantisation of Fields	24
3.1	Canonical Quantisation	24
3.2	Free fields	25
3.3	Complex Scalar Fields	33
3.4	Heisenberg Picture and KG field in space-time	35
4	Propagators	39
4.1	Introduction to propagators	39
4.2	Feynmann propagator	40
4.3	Chapter 2, Peskin and Schroeder Problems	43
5	Interacting Theories	45
5.1	Introduction to interactions	45
5.2	Interaction picture	46
5.3	Introduction to Scattering	48
5.4	Wick's Theorem	51
5.5	Feynmann Diagrams	55
6	Dirac Equation	59
6.1	Spin	59
6.2	Spinor representation	61
6.3	Action	65
6.4	The Dirac Equation	67
6.5	Notation	68
6.6	γ^5	69
6.7	Symmetries of the Dirac equation	71
7	Quantisation of Dirac Field	74
7.1	Hamiltonian	75
7.2	Fermionic Quantisation	76
7.3	Feynmann Propagator	78
7.4	Dimensional analysis	79
7.5	Feynmann rules for Yukawa theory	80

8	Quantum-Electro-Dynamics (QED)	81
8.1	Maxwell's Equations	81
8.2	Lorentz gauge quantisation	83
8.3	Feynmann rules for QED	89
9	More on symmetries	91
9.1	Review of Noether's theorem	91

1 Introduction

From Quantum mechanics (QM) we know that particles behave like waves under certain circumstances and the same is true for waves, this is what we know as the concept of wave-particle duality. Classically we think of electrons as fundamental particles and photons as ripples or disturbances in electro-magnetic (E-M) fields. So to combine these views with wave-particle duality two possible theories could be true. Firstly, we may postulate the fields are the fundamental object in nature and that particle like behavior comes from the quantization of the ripples in the field or the other possibility could be that particles are infact the fundamental objects in nature and fields are generated by the particles.

The second view would suggest that each particle would have a field associated with it, so there would be an electron field, an up quark field, a neutrino field etc. This picture of fundamental fields is what is accepted and is the subject of Quantum field theory.

Why QFT?

Locality

This is one of the most fundamental reasons for the acceptance of QFT, it is the same reason for classical field theories such as E-M and gravitational fields. It is introduced to avoid the problem of having "action at a distance", that is the fact that without a field we need an explanation of how a force is felt on objects that are a large distance away. Another way to think about it, is that causality has to be conserved according to special relativity.

Indistinguishable particles

All particles of the same type are indistinguishable. This is a remarkable fact, that every electron ever created is *exactly the same* as every other one. So postulating that an electron field exists from which every electron is made seems a reasonable explanation to this fact that all the electrons are the same (as they are made of the same thing).

In QM we are simply told that fermions have half integer spin and bosons have integer spin and therefore fermions obey the Pauli exclusion principle, QFT proves these facts and their relations to Fermi-Dirac statistics and Bose-Einstein statistics.

Special Relativity(SR) + Quantum Mechanics(QM)

We know from special relativity $E = mc^2$ so at high energies the phenomena of pair production takes place which naturally means that the number of particles in relativistic systems is conserved. An example is in particle accelerators, two particles traveling at high velocities can collide and generate a shower of 1000s of particles.

Now if we try to combine QM with SR and form a relativistic Schrodinger equation, it won't work for a simple reason that the Schrodinger equation describes the motion of a single particle wave-function, however in the relativistic case, this can never be true as we do not have a single

particle. Now consider a particle trapped in a box of size L , from uncertainty principle we have:

$$\Delta p \geq \frac{\hbar}{L} \tag{1.0.1}$$

However in under relativity we have an energy of:

$$\Delta E \geq \frac{\hbar c}{L} \tag{1.0.2}$$

When $\Delta E = 2mc^2$, one gets pair production of the particles with mass m . If a particle is localised in a region of size L that is $\leq \frac{\hbar}{mc}$, then it is not even possible to say there is just one particle. The length scale

$$\lambda = \frac{\hbar}{mc} \tag{1.0.3}$$

is called the *Compton wavelength*. This basically means that at length scales smaller than this we cannot talk about a single particle².

What is QFT?

Conceptually there is nothing new in QFT that isn't already in QM. In QM one takes the classical degrees of freedom (d.o.f) and promotes them to operators which act on some Hilbert space (general complete vector space). Exactly the same is true for QFT, the classical d.o.f are fields and one takes this field and promotes it to operators which act on some Hilbert space, so we have a field that is an *operator valued function* over space.

Classically every position in space has a field value associated with it, in QFT every position in space has an operator associated with it. This however raises a major problem with QFT, the problem is that there are infinite number of d.o.f, as space is a continuum (from what we know now) (so it contains an infinite number of points). The major work done since the framework for QFT was first laid down in the early 1930's has been to resolve this mystery of the infinitely many d.o.f's.

1.1 Units and conventions in QFT

There are three fundamental constants in QFT, with the following units:

$$[c] = LT^{-1} \tag{1.1.1}$$

$$[\hbar] = L^2 M^2 T^{-1} \tag{1.1.2}$$

$$[G] = L^3 M^{-1} T^{-2} \tag{1.1.3}$$

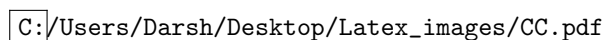
In QFT we set $\hbar = c = 1$, sometimes called natural units. To get back the SI units one has to put back the \hbar and c .

²The other familiar length scale is the de Broglie wavelength, $\lambda = \frac{\hbar}{p}$. This is in the non-relativistic case, one can think about these two length scales as follows, the de Broglie wavelength is when a particle starts to exhibit wave like properties and the Compton wavelength is when it doesn't make sense to think about a particle anymore

1.2 Notation

The first thing one notices when being introduced to QFT³ is the indicies that are in the subscript (lower) or in the superscript (raised) of various algebraic terms such as x^μ . So I shall start of by explaining them in as much detail as possible as they are probably the most important notation (and one cannot go any further without understanding them).

A raised indicy such as x^μ represents a set of Contravariant coordinates and a lower indicy represents Covariant coordinates, the meaning of which is easily shown by a diagram.

 C:/Users/Darsh/Desktop/Latex_images/CC.pdf

The figure above shows the two axis labeled ϕ_1 and ϕ_2 respectively (I have chosen these particularly to emphasise that even though, the diagram shows them as axis, in general they are simply algebraic notation showing two basis vectors of a vector space, say ϕ). The dotted red lines show the *contravariant* coordinates, hence they have the raised indicies. It might easier to think about them in terms of a vector such as: x^μ led ϕ_1 and ϕ_2 respectively (I have chosen these particularly to emphasise that even though, the diagram shows them as axis, in general they are simply algebraic notation showing two basis vectors of a vector space, say ϕ). The dotted red lines show the *contravariant* coordinates, hence they have the raised indicies. It might easier to think about them in terms of a vector such as:

$$x^\mu = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

Of course this is just for the two dimensional case, in general it can be for any number of dimensions and each dimension will have a coordinate in the matrix. Similarly for *covariant* coordinates one has:

$$x_\mu = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

This can also simply be extended to higher dimensions. This simple illustration shows the representation of the same vector in contravariant and covariant coordinates (note that the basis vectors are the same, $\hat{\phi}_1$ and $\hat{\phi}_2$. Now I shall discuss the relations between the two coordinates system, to that I have to lay down some formal definitions for the contrvariant and covariant coordinates:

$$x_{1,2} \equiv \mathbf{x} \cdot \hat{\phi}_{1,2}$$

³Infact this first course in quantum mechanics did not even involve the Dirac notation as everything was worked out in terms of wave-functions, however I found that notation relatively trivial and therefore I will not dwell on it here.

This is the expression for covariant coordinates (lower indicies) and there is a similar equation that defines the contravariant coordinates (simply raising the indicies). Note that there is no difference between covariant and contravariant coordinates when one has Euclidean space⁴ (i.e flat space). Now we consider the scalar product of two vectors:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot (y^1 \hat{\phi}_1 + y^2 \hat{\phi}_2) \quad (1.2.2)$$

Here I have chosen the y coordinates to be contravariant, this can be rearranged as (using the definitions in equation (1)):

$$\mathbf{x} \cdot \mathbf{y} = y^1 \mathbf{x} \cdot \hat{\phi}_1 + y^2 \mathbf{x} \cdot \hat{\phi}_2 = y^1 x_1 + y^2 x_2 \quad (1.2.3)$$

This shows that the scalar product is obtained by pairing upper with lower indicies, infact this is always true. The relation between contrvariant and covariant coordinates is straightforward to derive. Consider the contravariant coordinates for \mathbf{x} :

$$\mathbf{x} = x^1 \hat{\phi}_1 + x^2 \hat{\phi}_2 \quad (1.2.4)$$

One can obtain the covariant coordinates x_i by dotting \mathbf{x} with the unit vectors in the covariant system, $\hat{\mathbf{u}}_i$ (for clarity, $\hat{\phi}_i$ are *contravariant basis unit vectors* and $\hat{\mathbf{u}}_i$ are *covariant basis unit vectors*):

$$x_i = \mathbf{x} \cdot \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i \quad (1.2.5)$$

$$x_i = (x^1 \hat{\phi}_1 + x^2 \hat{\phi}_2) \cdot \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i \quad (1.2.6)$$

$$x_i = x^j (\hat{\mathbf{u}}_i \hat{\mathbf{u}}_j \cdot \hat{\phi}_i) \quad (1.2.7)$$

Now we can define the metric tensor, which is crucial in quantum field theory:

$$g_{ij} = \hat{\mathbf{u}}_i \hat{\mathbf{u}}_j \cdot \hat{\phi}_i \quad (1.2.8)$$

Which immediately leads to:

$$x_i = g_{ij} x^j \quad (1.2.9)$$

Now we use the Einstein convention that repeated indicies are summed over. This is a notation that was introduced by Einstein in his development of Relativity and it streamlines many common algebraic expressions. Instead of using the traditional sigma for sums, the strategy is to allow the repeated subscript to become itslef the designation for the summation:

$$a_1 x_1 + a_2 x_2 + a_3 x_3 + \dots + a_n x_n \equiv \sum_{i=1}^n a_i x_i \quad (1.2.10)$$

Instead of writing down the sum symbol one would simply write $a_i x_i$, where $1 \leq i \leq n$ is adopted as the universal range of summation. Consider the expression $a_{ij} x_k$ does not indicate summation, but both $a_{ii} x_k$ and $a_{ij} x_j$ do so over the respective ranges $1 \leq i \leq n$ and $1 \leq j \leq n$. If $n = 4$, then:

$$a_{ii} x_k \equiv a_{11} x_k + a_{22} x_k + a_{33} x_k + a_{44} x_k \quad (1.2.11)$$

⁴In quantum field theory one generally deals with Euclidean geometry so the two coordinate systems are equivalent.

$$a_{ij}x_j \equiv a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + a_{i4}x_4 \quad (1.2.12)$$

Free and Dummy Indices

In the example above the expression $a_{ij}x_j$ involves two sorts of indices. The index of summation, j , which ranges over the integers $1, 2, 3, \dots, n$, cannot be preempted. However at the same time, it is clear that the use of the particular character j is inessential; e.g., the expressions $a_{ir}x_r$ and $a_{iv}x_v$ represent exactly the same sum as $a_{ij}x_j$ does. For this reason, j is called a *dummy* index. The index i , which may take on any particular value $1, 2, 3, \dots, n$ independently, is called a *free* index. Note that, although we call the i "free" in the expression $a_{ij}x_j$, that "freedom" is limited in the sense that generally, unless $i = k$:

$$a_{ij}x_j \neq a_{kj}x_j \quad (1.2.13)$$

Special tensors and four vectors

I mentioned the metric tensor above, now that we have talked about the summation rule over repeated indices I can define it in another way:

$$g_{ij} \equiv g_{ik}g_j^k \equiv g_{ik}g_j^l g^{kl} \quad (1.2.14)$$

Here we have made use of the matrix form of the Kronecker delta function, δ which is also defined in terms of the metric tensor:

$$\begin{aligned} g_i^j &\equiv \delta_i^j \\ \delta_i^j &= 0, i \neq j \\ &= 1, i = j \end{aligned} \quad (1.2.15)$$

The flat Minkowski space metric, $g_{\mu\nu}$ (sometime also written as $\eta_{\mu\nu}$) is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Moving on to four vectors, they are a very simple mathematical construct⁵, once again best explained through an example, consider the contravariant vector x^μ . It has coordinates (t, x, y, z) and μ generally represents $0, 1, 2, 3$. This can now be used along with the metric tensor and the relation in Eq. (9) to show new properties. Consider the multiplication (as shown in the equation) of $x^\mu g_{\mu\nu}$. Firstly we notice that these are two matrices of dimensions (1×4) and (4×4) therefore the result from the matrix multiplication should be another matrix of dimension (1×4) . So computing the result:

⁵The momentum four vector, $P^\mu = (E, \mathbf{p}) = (E, p_x, p_y, p_z)$. The energy and momentum combine to form the momentum four vector, analogous to the spatial and time four vector

$$\begin{pmatrix} t & x & y & z \\ & x^\mu & & \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} t & -x & -y & -z \\ & x_\nu & & \end{pmatrix} \quad (1.2.16)$$

$g_{\mu\nu}$

This shows that the co-variant coordinates are negative in position and will give rise to other important properties. One of the definitions of a vector space is that there is a function known as a scalar product that basically maps the component of one vector onto another vector in the vector space. Since vectors can be represented as matrices, I shall now show how to compute a scalar product of two four vectors. Consider two vectors:

$$\mathbf{V} = \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix} \quad \mathbf{U} = (U^0 \quad U^1 \quad U^2 \quad U^3)$$

Scalar product is written as:

$$\mathbf{U} \cdot \mathbf{V} = \mathbf{U}^\mu \eta_{\mu\nu} \mathbf{V}^\nu \quad (1.2.17)$$

The R.H.S term can be written as:

$$\begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} (U^0 \quad U^1 \quad U^2 \quad U^3)$$

By multiplying out the second and third matrices we get:

$$(U^0 \quad -U^1 \quad -U^2 \quad -U^3) \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix}$$

Which just equals:

$$U^0 V^0 - U^1 V^1 - U^2 V^2 - U^3 V^3 \quad (1.2.18)$$

This matches the result we obtained in Eq.(16) which means the scalar product can be written in general as:

$$a_\mu b^\mu = a^\mu b_\mu = a^\mu g_{\mu\nu} b^\nu \quad (1.2.19)$$

This proves to be a very important result that links the metric tensor and the vectors. Another important tensor is the *anti-symmetric tensor* (sometimes referred to as the *Levi-Civita* tensor). It is defined as:

$$\begin{aligned} \epsilon^{\mu\nu\alpha\beta} &\equiv 1, \text{ if the indices have an even permutation}(0, 1, 2, 3) \\ &-1, \text{ if the indices have an odd permutation}(0, 1, 2, 3) \\ &0, \text{ if the indices is of no permutation} \end{aligned} \quad (1.2.20)$$

- *Even permutation* means that the indices follow *cyclic permutations*, that is $0 \rightarrow 1, 1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 0$. For example $\epsilon^{0123}, \epsilon^{1230}, \epsilon^{2301}$ are all even permutations
- *Odd permutation* means that the indices follow *inverse-cyclic permutations*, that is $1 \rightarrow 0, 2 \rightarrow 1, 3 \rightarrow 2, 0 \rightarrow 3$. For example $\epsilon^{3210}, \epsilon^{2103}, \epsilon^{1032}$ are all odd permutations
- *no permutation* means there is no order to the indices, for example, $\epsilon^{1302}, \epsilon^{0231}, \epsilon^{0312}$

Lorentz transformations

Under a Lorentz transformation a four-vector transforms according to matrix multiplication:

$$x^\mu = \Lambda_\nu^\mu x^\nu \quad (1.2.21)$$

Where $\Lambda_\nu^\mu = 4 \times 4$ matrix and it defines the Lorentz transformation.

Special cases of Λ_ν^μ :

- Rotation; Λ_ν^μ (about z axis):
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
- Boost (simply meaning the changing of coordinates under a Lorentz transformation); Λ_ν^μ (boost in x direction):
$$\begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Where γ is the Lorentz factor (remember the units are $c = 1$):

$$\gamma = \sqrt{(1 - v^2)} \quad (1.2.22)$$

The set of Lorentz transformations may be defined as transformations which leave the metric tensor *invariant*:

$$g_{\mu\nu} \equiv g_{\alpha\beta} \Lambda_\mu^\alpha \Lambda_\nu^\beta \quad (1.2.23)$$

Derivative transformations under Lorentz transforms

I shall define here how derivatives transform under the Lorentz transforms specifically. Firstly consider the scalar field ϕ , and a small variation in the field:

$$\delta\phi = \frac{\partial\phi}{\partial x^\mu} \delta x^\mu \quad (1.2.24)$$

Since this is a scalar quantity we want to write it as:

$$\delta\phi = \partial_\mu \phi \delta x^\mu \quad (1.2.25)$$

This equation shows the variation in this field is the covariant derivative of the field times the contravariant variation in x^μ four-vector. So we define:

$$\begin{aligned}\partial_\mu &= \frac{\partial}{\partial x^\mu} \\ &= \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)\end{aligned}\tag{1.2.26}$$

$$\begin{aligned}\partial^\mu &= \frac{\partial}{\partial x_\mu} \\ &= \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right)\end{aligned}\tag{1.2.27}$$

So we see that ∂_μ transforms as a *covariant* four-vector and opposite for ∂^μ , so we note that from Eq.(26,27):

$$\partial_\mu A^\mu = \partial_0 A^0 + \partial_j A^j\tag{1.2.28}$$

For example if $A^\mu = x^\mu$ we get:

$$\partial_\mu x^\mu = \partial_0 x^0 + \partial_1 x^1 + \partial_2 x^2 + \partial_3 x^3 + \partial_4 x^4$$

which can also be written as:

$$\partial_\mu x^\mu = \left(\frac{\partial t}{\partial t}, \frac{\partial x}{\partial x}, \frac{\partial y}{\partial y}, \frac{\partial z}{\partial z} \right)\tag{1.2.29}$$

The final identity is the relation between ∂^μ and ∂_μ :

$$\partial^\mu \partial_\mu = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right)$$

multiplying this put gives:

$$\partial^\mu \partial_\mu = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\tag{1.2.30}$$

which can be written in operator terms as:

$$\partial^\mu \partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2\tag{1.2.31}$$

where ∇^2 is just the del operator:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\tag{1.2.32}$$

2 Field theory

As stated before, the main purpose of QFT is to describe the dynamics of fields. Field is a physical quantity defined at every point in space and time, there can be vector fields (e.g. EM fields, can be thought of arrows at each point in space) or scalar fields (e.g. temperature, can be thought of as a number associated with each point in space).

When discussing fields in classical mechanics, we have a finite number of d.o.f's. Dynamical variables are denoted by:

$$q_a(t)$$

where q represents the coordinate of the variable, like position, a represents a label (could be whether the coordinate is negative or positive, for example). In QFT we apply the same logic, however the dynamic variable is now a field so we get:

$$\phi_a(\mathbf{x}, t)$$

Note here that the position \mathbf{x} is itself a label, as supposed to being a variable. This means that the uncertainty principle no longer holds for position and momentum in QFT. A simple example of this are the electric and magnetic fields, $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$ are both spatial three vectors. However these can be derived from a single 4-vector⁶:

$$A^\mu(\mathbf{x}, t) = (\phi, \mathbf{A}) \tag{2.0.1}$$

where $\mu = 0, 1, 2, 3$. Maxwell's equations are:

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} \tag{2.0.2}$$

and similarly for \mathbf{B} :

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.0.3}$$

Now using the vector calculus identities:

$$\text{div}(\text{curl}\mathbf{A}) = 0$$

$$\text{curl}(\text{curl}\mathbf{A}) = 0$$

It is easy to show that $\mathbf{A}^\mu(\mathbf{x}, t)$ does in fact lead to the Maxwell equations.

⁶Very similar to the problem in PeskinSchroeder chapter 2 question 1

2.1 Lagrangian and equation of motion

In particle mechanics the Lagrangian is a function of the velocities and positions, therefore carrying on the same logic, in QFT the Lagrangian is a function of the dynamical variable and its time derivative as well (which happen to be fields); $\phi(\mathbf{x}, t)$ and $\dot{\phi}(\mathbf{x}, t)$. However, as I said before the field now also depends on position, \mathbf{x} , so we are at liberty to include terms like; $\nabla\phi$, $\nabla^2\phi$, $\nabla^3\phi$ etc⁷. Since there is no classical analogue one is not forced to limit the derivatives to be first order (as is the case with time derivatives).

The Lagrangian is written in the following way:

$$\mathbf{L}(t) = \int d^3x L(\phi_a, \partial_\mu\phi_a) \quad (2.1.1)$$

Here L is called the Lagrangian density and to obtain the actual Lagrangian one needs to integrate L over all space⁸. The action, S , is defined as:

$$S = \int dt \int d^3x L = \int d^4x L \quad (2.1.2)$$

As stated above the Lagrangian can depend on $\nabla\phi$ in any way as there is no classical analogy. However since we are trying to combine SR and QM. we require space and time be to be of an equal footing and for that reason, we only consider Lagrangians that are functions of *equal* derivatives of space and time. We determine the equation of motion by the principle of least action ($\delta S = 0$) keeping the end points fixed. Small change in action is given by using the equation above:

$$\delta S = \int d^4x \left[\frac{\partial L}{\partial \phi_a} \delta \phi_a + \frac{\partial L}{\partial (\partial_\mu \phi_a)} (\partial \partial_\mu \phi_a) \right] \quad (2.1.3)$$

Usually in CM, the initial conditions are something like, the particle starts at position, \mathbf{x} and has velocity, \mathbf{v} . However in the case of field theory what the initial conditions are slightly different (strictly speaking they are not initial conditions, but boundary conditions). We take the initial value of the field and the final value of the field and we see how the field can evolve into its final value, such that the action is an extremum (generally a minimum). Returning to the equation above, it can be rewritten using integration by parts as:

$$\delta S = \int d^4x \left[\frac{\partial L}{\partial \phi_a} - \partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi_a)} \right) \right] \delta \phi_a + \partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) \quad (2.1.4)$$

The last term is a boundary term, and it goes to zero under the following initial conditions (boundary conditions):

$$\delta \phi_a(\mathbf{x}, t_1) = 0 \quad (2.1.5)$$

$$\delta \phi_a(\mathbf{x}, t_2) = 0 \quad (2.1.6)$$

$$\delta \phi_a(\mathbf{x} \rightarrow \infty, t) = 0 \quad (2.1.7)$$

⁷We will only consider Lagrangian that depend on $\nabla\phi$, the reason for this lies in Lorentz invariance which will be discussed later

⁸in general everybody calls the Lagrangian density the Lagrangian which is what I shall be doing as well from now on

Now remember we are trying to find describe the dynamics of the field where action is minimised, so the remaining equation without the boundary term is:

$$\delta S = \int d^4x \left[\frac{\partial L}{\partial \phi_a} - \partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi_a)} \right) \right] \delta \phi_a$$

So for δS to be zero we can the simple, but very important relation:

$$\partial \left(\frac{\partial L}{\partial (\delta_\mu \phi_a)} \right) - \frac{\partial L}{\delta \phi_a} = 0 \quad (2.1.8)$$

This is the *equation of motion* for fields, it is called the *Euler-Lagrange* equation (EL equation). So for a given Lagrangian that is a function of a set of fields, the equation of motion can now be calculated.

2.2 Application of Lagrangian and EL equation

In this section I shall apply the basic method of finding the equation of motion for different Lagrangians. The steps are as follows:

- Form a Lagrangian using the time and spatial derivatives of a field
- Compute:

$$\partial \left(\frac{\partial L}{\delta (\delta_\mu \phi_a)} \right)$$

$$\frac{\partial L}{\delta \phi_a}$$

- Then substitute into EL equation to form the equation of motion

Klein- Gordon equation

The Lagrangian for this system is:

$$L = \frac{1}{2} \eta^{\mu\nu} \delta_\mu \phi \delta_\nu \phi - \frac{1}{2} m^2 \phi^2 \quad (2.2.1)$$

Substituting the metric tensor and the partial derivatives identities from the introduction section into the Klein-Gordon equation:

$$L = -\frac{1}{2} m^2 \phi^2 + \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi \right) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

multiplying out the terms gives:

$$L = \frac{1}{2} \dot{\phi}^2 - \left(\frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 \right) \quad (2.2.3)$$

Comparing this to the classical Lagrangian:

$$L = T - V$$

Where T is the kinetic energy and V is the potential energy, now we have to remember that the L we have been calling the Lagrangian is actually the density, so to get the Lagrangian we have to integrate over all space. Hence the equivalent terms for the kinetic energy and potential energy are:

$$T = \int d^3x \frac{1}{2} \dot{\phi}^2 \quad (2.2.4)$$

$$V = \int d^3x \frac{1}{2} (\nabla\phi)^2 + \frac{1}{2} m^2 \phi^2 \quad (2.2.5)$$

The first term in the potential expression is called the gradient energy, while the last term is simply called the potential. To determine the equation of motion from the Lagrangian we compute the following the terms for the EL equation:

$$\frac{\partial L}{\partial \theta} = -m^2 \phi \quad (2.2.6)$$

$$\frac{\partial L}{\partial(\partial_\mu \phi)} = \partial^\mu \phi = (\dot{\phi} - \nabla\phi) \quad (2.2.7)$$

Therefore substituting the ingredients into the EL equation:

$$\ddot{\phi} - \nabla^2 \phi + m^2 \phi = 0 \quad (2.2.8)$$

this can also be rearranged and written concisely as:

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0 \quad (2.2.9)$$

This is the *Klein-Gordon equation*⁹. This is the equation of motion that shows how a scalar field varies in time.

First-Order Lagrangian

Consider a Lagrangian linear in time derivative, rather than quadratic being quadratic which was the case in the Klein-Gordon equation. We take a complex scalar field defined by the real Lagrangian:

$$L = \frac{i}{2} (\psi^* \dot{\psi} - \dot{\psi}^* \psi) - \nabla\psi^* \cdot \nabla\psi - m\psi^* \psi \quad (2.2.10)$$

Once again to obtain the equation of motion we compute the usual factors:

$$\frac{\partial L}{\partial \psi^*} = \frac{i}{2} \dot{\psi} - m\psi \quad (2.2.11)$$

$$\frac{\partial L}{\partial \dot{\psi}^*} = -\frac{i}{2} \psi \quad (2.2.12)$$

⁹In fact this was first written down by Schrodinger in an attempt to find a relativistic version of the "Schrodinger equation"

$$\frac{\partial L}{\partial \Delta \psi^*} = -\nabla \psi \quad (2.2.13)$$

Finally substituting into the EL equation gives:

$$i \frac{\partial \psi}{\partial t} = -\nabla^2 \psi + m\psi \quad (2.2.14)$$

Yup, this looks exactly like the Schrodinger equation! However it is not, because the function ψ is not a wavefunction here, it is a field. It does not have any of the properties of the wavefunction (i.e, it is not subject to the Born interpretation of probability). I shall mention briefly the initial conditions of the two examples above; When the Lagrangian is quadratic in the field time derivatives we need to know ψ and $\dot{\psi}$ at a specific time to determine future time evolution; when the Lagrangian was a function of a single power of the time derivative of the field, we only need the value of the field at a given time to determine time evolution.

Maxwell's Equation¹⁰

The Lagrangian is:

$$L = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A^\mu)^2 \quad (2.2.15)$$

Once again we compute the ingredients:

$$\frac{\partial L}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + (\partial_\mu A^\mu)\eta^{\mu\nu} \quad (2.2.16)$$

therefore:

$$\partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu A_\mu)} \right) = -\partial^2 A^\nu + \partial^\nu (\partial_\mu A^\mu) = -\partial_\mu (\partial^\mu A^\mu - \partial^\nu A^\mu) \quad (2.2.17)$$

Here one can define the field strength as:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.2.18)$$

So we can succinctly write Eq.(2.28) as:

$$\partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu A_\mu)} \right) \equiv -\partial_\mu F^{\mu\nu} \quad (2.2.19)$$

This can be used along with vector calculus identities to obtain Maxwell's equations. Using the notation of the field strength, we may rewrite the Maxwell Lagrangian as:

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (2.2.20)$$

Lorentz Invariance

¹⁰Once again this relates to problem 1 in chapter 2 of Peskin and Schroeder

The notation of Lorentz transformations was discussed in the introduction section here we shall emphasise its importance. the goal is to construct a relativistic field theory with time and space on equal footing. This just means that the theory should be invariant under Lorentz transformations. The following equation is the Lorentz transformation:

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu \quad (2.2.21)$$

Examples are rotations and Lorentz boosts as defined in the notations section. This shows how Lorentz transformations act on space, however we want to know how the transformations affect the *fields* in space. The effect of a Lorentz transformation on a scalar field is of the form:

$$\phi(x) \rightarrow \phi'(x) = \phi(\Lambda^{-1}x) \quad (2.2.22)$$

This is an active transformation. The way to think about it is; consider you have a temperature field and at a given position we have a hotspot. Now a Lorentz transformation is performed on the temperature field. Suppose the transformation is a rotation, so we take the entire field and move it (active transformation, as supposed to simply changing coordinates) so now the hotspot is in another position. If now one wants to write the new field in terms of the old field, the coordinates of the hotspot in the old field will have to be converted back and this is the reason for having the Λ^{-1} as supposed to just Λ . This is how Λ acts on any field, we are just interested in how they act on fields that are *invariant* after the transformation.

If a solution to a particular equation of motion is subjected to a Lorentz transformation and the resulting function is also a solution to that equation of motion, the theory to which the equation of motion belongs, is said to be Lorentz invariant. This can be insured by saying that the action is invariant, therefore theories in which Lagrangians have dependence of time and space, *not* on equal footing will not be Lorentz invariant (i.e first order in time and second order in space will not lead to theories being Lorentz invariant).

Next consider a Lorentz transformation on a vector field, $A_\mu(x)$:

$$A_\mu(x) \rightarrow A'_\mu(x) = \Lambda^\nu_\mu A_\nu(\Lambda^{-1}x) \quad (2.2.23)$$

The x label transforms in exactly the same way, but now there is an index μ aswell in the vector field which, under Lorentz transformation, goes to ν , which is the reason for the transformation matrix having the ν index aswell. This is expected as the indicy represents the direction of a value in the field (best thought of as the direction of an arrow, in a space having arrows at every point in space).

2.3 Symmetries and Noethers theorem

The role of symmetries in field theory is just as important (if not more so) as it is in classical particle mechanics. There are Lorentz symmetries, gauge symmetries, internal symmetries and supersymmetries. Here we shall briefly look at Lorentz symmetries and internal symmetries, but first we visit the topic of Noethers theorem.

Noethers theorem

I shall start with the definition of the theorem; *Every continuous symmetry of the Lagrangian gives rise to a conserved current, $j^\mu(x)$, such that the equations of motion imply:*

$$\partial_\mu j^\mu(x) = 0 \quad (2.3.1)$$

Or in a more familiar form (atleast it is for me), this just represents the equation of continuity (look at definition of ∂_μ):

$$\frac{\partial}{\partial t} j^0 + \nabla \cdot \mathbf{j} = 0 \quad (2.3.2)$$

If the current is conserved, we have a conserved quantity (usually called charge) which is defined as:

$$Q = \int \partial^3 x j^0 \quad (2.3.3)$$

Here we are integrating over all space, the zeroth component of the current to obtain the charge, which is quantity that does not change with time. To see why it is zero consider the following equation for the derivative of the charge:

$$\dot{Q} = \int_{R^3} \frac{\partial j^0}{\partial t} \partial^3 x \quad (2.3.4)$$

Where the integration limit R^3 means the integration is over all space. Now using the continuity equation 2.3.2, we can rewrite this as:

$$\dot{Q} = - \int_{R^3} \nabla \cdot \mathbf{j} \partial^3 x \quad (2.3.5)$$

This integral will go to zero if we assume (as usual) that $\mathbf{j} \rightarrow 0$ as $x \rightarrow \infty$ and hence we have shown that \dot{Q} is zero and therefore a conserved quantity. The series of equations I have written above also hint at local conservation of the quantity, which is best seen by the equation above. The best way to prove local conservation is to invoke Gauss's law upon this equation. First consider the charge in a fixed volume, \mathbf{V} , we have:

$$\dot{Q} = - \int_{\mathbf{V}} \nabla \cdot \mathbf{j} \partial^3 x \quad (2.3.6)$$

Now using Gauss's divergence theorem we have:

$$- \int_{\mathbf{V}} \nabla \cdot \mathbf{j} \partial^3 x = - \int_{\mathbf{A}} \mathbf{j} \cdot d\mathbf{S} = \dot{Q} \quad (2.3.7)$$

This shows that the change in a conserved quantity is equal to the *flux* of the quantity out of a given surface.¹¹

Proof of Noethers theorem

Returning to the definition of Noethers theorem given at the very beginning, I shall now attempt to prove it. First we take a symmetry and consider an infinitesimally small rotation (or any transformation), since the statement given in the beginning is of a continuous symmetry we have:

¹¹This is exactly the same for the electric charge and the more familiar Gauss's law in EM

$$\phi_a(x) \rightarrow \phi'_a(x) \cdot \phi_a(x) + \epsilon \delta \phi_a \quad (2.3.8)$$

So a field is subject to a Lorentz transformation of an infinitesimal amount, ϵ , in general there can be higher order terms as well such as, ϵ^2 , ϵ^3 , etc but for now we shall consider them to be negligible and ignore them. The definition of a symmetry is that the Lagrangian does not change under a symmetrical transformation. So consider making an arbitrary transformation:

$$\phi_a \rightarrow \phi_a + \delta \phi_a \quad (2.3.9)$$

A small change in the Lagrangian can be modelled as:

$$\delta L = \frac{\partial L}{\partial \phi_a} \delta \phi_a + \frac{\partial L}{\partial (\partial_\mu \phi_a)} \delta (\partial_\mu \phi_a) \quad (2.3.10)$$

This can be rewritten as¹²:

$$\delta L = \left[\frac{\partial L}{\partial \phi_a} - \partial_\mu \frac{\partial L}{\partial (\partial_\mu \phi_a)} \right] \delta \phi_a + \partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) \quad (2.3.11)$$

The term in brackets is zero when the equations of motion are obeyed (as the action is always minimised, $\delta S = 0$). So we are left with:

$$\delta L = \partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) \quad (2.3.12)$$

Now if we suppose that $\delta \phi_a$ is a symmetry then δL is zero, therefore:

$$\partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) = 0 \quad (2.3.13)$$

Comparing this to Eq 2.3.12 we see that:

$$\frac{\partial L}{\partial (\partial_\mu \phi_a)} \delta \phi_a = j^\mu \quad (2.3.14)$$

This is a very important result as it shows the power of Noether's theorem, not only can we find out whether a quantity is conserved, we can also find the current immediately by the relation above.

The theorem can be further generalised by stating that for a given symmetry, instead of $\delta L = 0$ we have:

$$\delta L = \partial_\mu F^\mu \quad (2.3.15)$$

So under the symmetry, the Lagrangian changes by the *derivative* of some function F^μ . In this case the action is (now make sure to remember that what we are calling the Lagrangian is really the Lagrangian density so we must integrate over all space to get the Lagrangian):

$$S = \int dt \int L \partial^3 x = \int \partial_\mu F^\mu \partial^3 x \int dt \quad (2.3.16)$$

¹²an analogous step was taken when we were considering an small change in the action, check that for more details if unsure

However the R.H.S of Eq 2.3.16 is a total derivative and therefore zero! This can be reconciled with Noethers theorem to get a conserved current by simply subtracting $\partial_\mu F^\mu$ from the original equation for j^μ :

$$j^\mu = \frac{\partial L}{\partial(\partial_\mu \phi_a)} \delta \phi_a - F^\mu \quad (2.3.17)$$

So this is just another way of saying that under a *symmetric transformation* the *action is invariant*.

Application of translations and Noethers theorem

Classically, invariance under spatial translation of a system gives rise to the conservation of momentum and inavriance under temporal translations gives rise to the conservation of energy. We will now see a similar concept in terms of fields. Consider the infinitesimal translation:

$$x^\nu \rightarrow x^\nu - \epsilon^\nu \quad (2.3.18)$$

This gives a change in a field of:

$$\phi_a(x) \rightarrow \phi_a(x) + \epsilon^\nu \partial_\nu \phi_a(x) \quad (2.3.19)$$

If the Lagrangian has no explicit x dependence (just depends of fields, that depend of x):

$$L(x) \rightarrow L(x) + \epsilon^\nu \partial_\nu L(x) \quad (2.3.20)$$

The indicies $\nu = 0, 1, 2, 3$ show that their are four different translations taking place, resulting in four conserved currents according to Noethers theorem (as the Lagrangian is a total derivative):

$$(j^\mu)_\nu = \frac{\partial I}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta_\nu^\mu L \equiv T_\nu^\mu \quad (2.3.21)$$

Just to explain the notation; (j^μ) is the current that we have seen before. The index ν is for the four different current that we are computing. T_ν^μ is called the *energy-momentum tensor* (sometimes called the stress-energy tensor, or stress tensor) and it follows the relation:

$$\partial_\mu T_\nu^\mu = 0 \quad (2.3.22)$$

Now we have a very important result which shows four conserved quantites:

$$E = \int d^3x T^{00} \quad (2.3.23)$$

This is the total energy of the entire field. The other conserved quantity is momentum (in all three dimensions):

$$P^i = \int d^3x T^{0i} \quad (2.3.24)$$

An example: KG Lagrangian

The simplest form of Lagrangian:

$$L = \frac{1}{2}\partial_\mu\phi\partial^\nu\phi - \frac{1}{2}m^2\phi^2 \quad (2.3.25)$$

Now we can use the definition of the energy-momentum tensor we have:

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\nu\mu}L \quad (2.3.26)$$

For this example, the conserved energy and momentum are given by:

$$E = \int d^3x \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 \quad (2.3.27)$$

$$P^i = \int d^3x \dot{\phi} \partial^i\phi \quad (2.3.28)$$

These are two conserved quantities for that transcend the entire field. For this example the energy-momentum tensor happens to be symmetric i.e.:

$$T^{\mu\nu} = T^{\nu\mu}$$

This is not always the case. Generally we can make sure that the energy-momentum tensor of any theory is symmetric by adding an extra term of the form:

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\rho\Gamma^{\rho\mu\nu} \quad (2.3.30)$$

Here we have defined $\Gamma^{\rho\mu\nu}$ to be some function of the fields that is anti-symmetric in the first two indicies such that:

$$\Gamma^{\rho\mu\nu} = -\Gamma^{\mu\rho\nu} \quad (2.3.31)$$

This guarantees that $\partial_\mu\partial_\rho\Gamma^{\rho\mu\nu} = 0$ so that the new energy-momentum tensor, $\Theta^{\mu\nu}$ is also conserved quantity.

Internal symmetries (Global symmetries)

Consider a Lagrangian which depends a complex scalar field ψ and its derivatives:

$$L = \partial_\mu\psi^{*\mu}\psi - V(\psi^*\psi) \quad (2.3.32)$$

we have a symmetry such that:

$$\begin{aligned} \psi &\rightarrow e^{i\alpha}\psi \\ \delta\psi &\rightarrow i\alpha\psi \\ \delta\psi^* &\rightarrow -i\alpha\psi \end{aligned}$$

So we can calculate the current j using Noethers theorem:

$$j^\mu = i(\delta^\mu\psi^*)\psi - i\psi^*(\partial^\mu\psi) \quad (2.3.34)$$

This current shows that there is no dependence on the potential energy terms of the Lagrangian. The main difference between internal symmetries and other symmetries like rotation and translation symmetries is that the internal symmetries provide an insight into how the values of the field changes under transformations as supposed to the others in which the space changes and then one sees how the fields are affected.

We will see later that conserved quantities arising from internal symmetries have the interpretation of electric charge or particle number (this means lepton number, baryon number etc. **not** the number of particles, as the whole point of developing QFT is the fact that the number of particles is not conserved in relativistic conditions!)

How to manipulate symmetries

Suppose we have an initial symmetry such that $\delta L = 0$. In general one has a set of parameters (like α above) that provide a measure of change in a field, i.e the amount a field is rotated by. Now we redo the transformation with α not being a constant, but instead a function of x , $\alpha(x)$. Therefore for this general case $\delta L \neq 0$, however we know that when $\alpha(x)$ is constant $\delta L = 0$. Therefore we can logically conclude that δL must be proportional to the derivative of α :

$$\delta L = (\delta_\mu \alpha) h^\mu(\phi) \tag{2.3.35}$$

for some function $h^\mu(\phi)$. Now one can compute the action for this system:

$$\begin{aligned} \delta S &= \int d^4x \delta L \\ &= - \int d^4x \alpha(x) \partial_\mu h^\mu \end{aligned} \tag{2.3.36}$$

when the equation of motion is obeyed, the action is always minimised, so $\delta S = 0$ for any transformation. The only way this can be true is if the derivative of the new function h^μ is zero:

$$\partial_\mu h^\mu = 0 \tag{2.3.37}$$

So in this case h^μ is the conserved current! $h^\mu = j^\mu$.

Hamiltonian Theory

Define conjugate momenta:

$$\pi(x) = \frac{\partial L}{\partial \dot{\phi}} \tag{2.3.38}$$

The Hamiltonian density is defined as:

$$H = \pi(x) \dot{\phi}(x) - L(x) \tag{2.3.39}$$

Here $L(x)$ is the Lagrangian density. This should be viewed as a function of momentum $\Pi(x)$ and fields $\phi(x)$ not of $\dot{\phi}(x)$. Note that in the Lagrangian theory we had included derivatives of space as well, however this is not the case in Hamiltonian theory and therefore it is not Lorentz invariant (space and time do not have equal footing).

Example: Real Scalar Field

The Lagrangian for this system (as we have computed many times before) is:

$$L = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi) \quad (2.3.40)$$

First calculate the conjugate momentum:

$$\pi(x) = \frac{\partial L}{\delta\dot{\phi}} = \dot{\phi} \quad (2.3.41)$$

Therefore the Hamiltonian density is:

$$H = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi) \quad (2.3.42)$$

And the overall Hamiltonian \mathbf{H} is the Hamiltonian density integrated over all space:

$$\mathbf{H} = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi) \right] \quad (2.3.43)$$

Notice that this exactly the same as the equation for energy obtained by the Noethers theorem (no surprise there!).

3 Quantisation of Fields

So far we have discussed classical fields. By following the same process that takes us from classical particle mechanics to quantised particle mechanics (i.e Quantum mechanics) we will build a quantum field theory from a classical field theory. So the idea is to start with classical field theory (the theory of a classical scalar field governed by the Lagrangian) and then quantize it, that is, reinterpret the dynamical variables as operators that obey *canonical* commutation relations.

3.1 Canonical Quantisation

Recall, in QM canonical quantisation tells us to take coordinates of momenta p^a and position, q^a , and promote them to operators. The Poisson bracket structure of classical mechanics morphs into the structure of commutation relations between operators, such that¹³:

$$[q_a, p^b] = i\delta_a^b \tag{3.1.1}$$

$$[q_a, q_b] = [p^a, p^b] = 0 \tag{3.1.2}$$

In field theory we will do the same for $\phi(x)$ and $\Pi(x)$. A quantum field is an operator valued function over space (not space-time), obeying the commutation relations:

$$[\phi_a(\mathbf{x}), \phi_b(\mathbf{y})] = [\pi^a(\mathbf{x}), \pi^b(\mathbf{y})] = 0 \tag{3.1.3}$$

$$[\phi(\mathbf{x}), \phi_b(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} - \mathbf{y})\delta_a^b \tag{3.1.4}$$

The R.H.S of Eq 3.1.4 has a dirac delta function, $\delta^{(3)}(\mathbf{x} - \mathbf{y})$, aswell as the usual δ_a^b . The dirac delta comes from the fact that space is just a label for fields and since space is a continuum we have a dirac delta as supposed to a Kronecker delta (the three in the power simply indicates the three dimensions of space).

We are working in the Schrodinger picture where $\phi(x)$ depends on space but not on time. All the time dependence is in the state and how it evolves, and it evolves according to the SE:

$$i\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle \tag{3.1.5}$$

So we have a Hamiltonian, which is some function of operators. We want to know, what the eigenvalue spectrum of this Hamiltonian looks like. However here lies the major problem; the Hamiltonian has an infinite number of operators as space has an infinite number of d.o.f (as it is a continuum). Therefore it is very to compute the eigenvalues. There is a set of field theories, known as *free fields*, where we can write the dynamics in a way that each d.o.f evolves indepenently from the others and these type of theories are relatively easier to do calculations in.

¹³remember we are working in units of $\hbar = 1$

3.2 Free fields

Lagrangians that contain only quadratic terms of fields are known as being Lagrangians of free fields. A system with quadratic fields in Lagrangian will always have an equation of motion (given by the EL equation) that is linear in the fields. It is this linearity in the equation on motion that makes these fields alot easier to work with as one can invoke the principle of superposition in this case and therefore sum over all the disturbances.

The simplest free field theory is the Klein-Gordon(KG) theory:

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0 \quad (3.2.1)$$

We want the spectrum of eigenvalues for the Hamiltonian, however since there is no obvious way to do this, one seeks guidance by taking the fourier transform of the KG field¹⁴:

$$\phi(\mathbf{x}, t) = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \phi(\mathbf{p}, t) \quad (3.2.2)$$

where:

$\phi(\mathbf{x}, t)$ = field in ordinary space

$\phi(\mathbf{p}, t)$ = field in FT (p) space

Now the KG equation in momentum space is:

$$\left(\frac{\partial^2}{\partial t^2} + (\mathbf{p}^2 + m^2)\right)\phi(\mathbf{p}, t) = 0 \quad (3.2.3)$$

This is just the equation for a Harmonic oscillator with frequency:

$$\omega_{\mathbf{p}} \equiv \sqrt{\mathbf{p}^2 + m^2} \quad (3.2.4)$$

The solution to the classical KG equation is a superposition of the infinite number of SHO, therefore we can add these oscillators together to give the solution. So to quantise the field we need to quantise the SHO. We shall do this in the Schrodinger notation of QM, so we have the Hamiltonian:

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2 \quad (3.2.5)$$

Where the usual commutation relations hold:

$$[q, p] = i \quad (3.2.6)$$

We define creation(a^\dagger and annihilation(a) operators as:

$$a = \sqrt{\frac{\omega}{2}}q + \frac{i}{\sqrt{2\omega}}p \quad (3.2.7)$$

and

$$a^\dagger = \sqrt{\frac{\omega}{2}}q - \frac{i}{\sqrt{2\omega}}p \quad (3.2.8)$$

¹⁴I am still not sure how are why we compute this step, I am simply following Peskin and Schroeder for this step.

which can be easily inverted to give:

$$q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger) \quad (3.2.9)$$

and

$$p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger) \quad (3.2.10)$$

Substituting into the above expressions we find the commutation relation:

$$[a, a^\dagger] = 1 \quad (3.2.11)$$

and the Hamiltonian can be rewritten as:

$$H = \omega(a^\dagger a + \frac{1}{2}) \quad (3.2.12)$$

The commutation relation between these operators and the Hamiltonian are as follows:

$$\begin{aligned} [H, a^\dagger] &= \omega a^\dagger \\ [H, a] &= -\omega a \end{aligned} \quad (3.2.13)$$

This means that if we have one eigenstate of H we can apply a and a^\dagger and get another eigenstate of H . So if we have the SE as follows:

$$H|E\rangle = E|E\rangle \quad (3.2.14)$$

Applying the annihilation and creation operators:

$$Ha^\dagger|E\rangle = |E + \omega\rangle a^\dagger|E\rangle \quad (3.2.15)$$

$$Ha|E\rangle = (E - \omega)a|E\rangle \quad (3.2.16)$$

So spectrum bounded below implies that there exists a ground state, $|0\rangle$ such that

$$a|0\rangle = 0 \quad (3.2.17)$$

And a is defined as the annihilation operator, therefore:

$$H|0\rangle = \frac{1}{2}\omega|0\rangle \quad (3.2.18)$$

as $H = \omega(a^\dagger a + \frac{1}{2})$ and in the ground state $aa^\dagger = 0$ therefore $H = \frac{1}{2}\omega$. Excited states are given by:

$$|n\rangle = (a^\dagger)^n|0\rangle \quad (3.2.19)$$

So for the n^{th} excited state, we apply the ground state with n annihilation operators. Therefore:

$$H|n\rangle = (n + \frac{1}{2})\omega|n\rangle \quad (3.2.20)$$

Where $(n + \frac{1}{2})\omega$ are the energy eigenvalues.

First application of QFT

The Hamiltonian of the free field theory is:

$$H = \int d^3x \pi^2 + (\nabla\phi)^2 + m^2\phi^2 \quad (3.2.21)$$

Remember the fields are now operators and are functions of just space, not of time because by going from the Lagrangian formalism to Hamiltonian formalism we have specifically broken Lorentz invariance (space and time are not on equal footing). Another reason for doubting the Lorentz invariance is because we are in the Schrodinger picture and as in QM, these operators are independent of time (the states on the other hand evolve in time according to the SE but are independent of position). The commutation relation we have is:

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (3.2.22)$$

In the last section we solved the Harmonic oscillator from QM, using the creation and annihilation operators, here we are going to do exactly the same thing and then interpret the solution that is obtained. First we define $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ as an infinite set of annihilation and creation operators (therefore labelled by \mathbf{p} subscript). The field operators are now:

$$\phi(x) \equiv \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} [a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} + a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}] \quad (3.2.23)$$

$$\Pi(x) \equiv \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{\omega_{\mathbf{p}}}{2}} [a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}] \quad (3.2.24)$$

$$\omega_{\mathbf{p}} = \sqrt{p^2 + m^2} \quad (3.2.25)$$

These can be considered to be the definitions of a and a^\dagger operators. The definition of $\omega_{\mathbf{p}}$ motivated by realising that doing a Fourier transform on the KG equation gives rise to the equation of the quantum harmonic oscillator. Now we claim the following commutation relations:

$$[\phi(\mathbf{x})\phi(\mathbf{y})] = [\pi(\mathbf{x})\pi(\mathbf{y})] = 0 \quad (3.2.26)$$

So now we substitute the ϕ 's and π 's into the commutation relations above to obtain the following commutation relation as well:

$$[a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (3.2.27)$$

The other commutation relation is:

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (3.2.28)$$

Once again substitute for ϕ 's and π 's:

$$[a_{\mathbf{p}}, a_{\mathbf{q}}] = [a_{\mathbf{p}}^\dagger, a_{\mathbf{q}}^\dagger] = 0 \quad (3.2.29)$$

Here is the proof:

First we assume the relation 3.2.27 and since we are saying it comes from 3.2.26, let us compute $[\phi(\mathbf{x})\phi(\mathbf{y})]$:

$$[\phi(\mathbf{x})\phi(\mathbf{y})] = \int \frac{d^3p d^3q}{(2\pi)^6} \frac{-i}{2} \sqrt{\frac{\omega_{\mathbf{q}}}{\omega_{\mathbf{p}}}} (-[a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] e^{i\mathbf{p}\cdot\mathbf{x} - i\mathbf{q}\cdot\mathbf{y}} + [a_{\mathbf{p}}^\dagger, a_{\mathbf{q}}] e^{-i\mathbf{p}\cdot\mathbf{x} + i\mathbf{q}\cdot\mathbf{y}}) \quad (3.2.30)$$

Now substitute Eq 3.2.27:

$$[\phi(\mathbf{x})\phi(\mathbf{y})] = \int \frac{d^3p}{(2\pi)^3} \frac{-i}{2} (-e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} - e^{i\mathbf{p}\cdot(\mathbf{y}-\mathbf{x})}) \quad (3.2.31)$$

This simply gives:

$$[\phi(\mathbf{x})\phi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (3.2.32)$$

Which completes the proof.

The Hamiltonian

Now we can compute the Hamiltonian in terms of $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$, initially we have:

$$H = \frac{1}{2} \int d^3x \pi^2 + (\nabla\phi)^2 + m^2\phi^2 \quad (3.2.33)$$

substituting the π and ϕ operators gives this ugly looking expression:

$$\begin{aligned} H &= \frac{1}{2} \int \frac{d^3x d^3p d^3q}{(2\pi)^3} \left[-\frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}}{2} (a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}) \right. \\ &+ \frac{1}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}} (i\mathbf{p}a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - i\mathbf{p}a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}) \cdot (i\mathbf{q}a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} - i\mathbf{q}a_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{x}}) \\ &+ \left. \frac{m^2}{2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}} (a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}) \cdot (a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} - a_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{x}}) \right] \end{aligned} \quad (3.2.35)$$

To simplify we can integrate over d^3x which will give produce delta functions of the form, $\delta^{(3)}(\mathbf{p} \pm \mathbf{q})$ which in turn allow us to perform the d^3p integral. So after all these calculations one obtains:

$$H = \frac{1}{4} \int \frac{d^3x}{(2\pi)^3} \frac{1}{\omega_{\mathbf{p}}} [(-\omega_{\mathbf{p}}^2 + \mathbf{p}^2 + m^2)(a_{\mathbf{p}}a_{-\mathbf{p}} + a_{\mathbf{q}}^\dagger a_{-\mathbf{q}}^\dagger) + (\omega_{\mathbf{p}}^2 + \mathbf{p}^2 + m^2)(a_{\mathbf{p}}a_{\mathbf{p}}^\dagger + a_{\mathbf{p}}^\dagger a_{\mathbf{p}})]$$

However we know:

$$\omega_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2$$

Therefore the first term in the integral above just goes to zero! and we are left with:

$$\begin{aligned}
 H &= \frac{1}{2} \int \frac{d^3x}{(2\pi)^3} \omega_{\mathbf{p}} [a_{\mathbf{p}} a_{\mathbf{p}}^\dagger + a_{\mathbf{p}}^\dagger a_{\mathbf{p}}] \\
 &= \int \frac{d^3x}{(2\pi)^3} \omega_{\mathbf{p}} [a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0)]
 \end{aligned} \tag{3.2.36}$$

The first line of the equation above shows that the free field is just the sum of an infinite number of harmonic oscillators (imagine crossing out the integral, we are just left with the Hamiltonian of a quantum harmonic oscillator in terms of creation and annihilation operators). The second line is just obtained by using the commutator relation between the creation and annihilation operators (we get the delta function evaluated at 0 as in this case $q = p$). Now the delta function is evaluated at zero, so we have a problem as there is an infinite spike inside the integral!

It is easy to see where this infinity is coming from. As stated above we are summing up a number of harmonic oscillators, however the space is a continuum and we are integrating over all of space so we are integrating over an infinite number of oscillators. Let us think about the vacuum state (ground state of the field), suppose we define $|0\rangle$ as the vacuum state, such that it is killed by the annihilation operator:

$$a_{\mathbf{p}} |0\rangle = 0 \tag{3.2.37}$$

and this must hold for all annihilation operators (\mathbf{p} of them). The energy of the ground state is given by applying the Hamiltonian operator to it:

$$H|0\rangle = E_0|0\rangle \tag{3.2.38}$$

For the ground state the term $a_{\mathbf{p}}^\dagger a_{\mathbf{p}}$ is just zero, so we have:

$$H|0\rangle = \left[\int d^3p \frac{1}{2} \omega_{\mathbf{p}} \delta^{(3)}(0) \right] |0\rangle = \infty |0\rangle \tag{3.2.39}$$

Here we see that there are in fact two infinities, one that is already present in the delta function and the other that comes from integrating over all of space. The infinity from space being infinitely large are called infra-red divergences. To deal with this divergence, consider the definition of a delta function in terms of the limiting case of a box function. Consider a box of size L :

$$(2\pi)^3 \delta^{(3)}(0) = \lim_{L \rightarrow \infty} \int_{-\frac{L}{2}}^{\frac{L}{2}} d^3x e^{ipx} \tag{3.2.40}$$

Note that the factors of 2π are still here simply because of the way I have chosen to define the Fourier transform. Now if we evaluate this at $p = 0$ (ground state), then the integral is simply:

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} d^3x = Volume(L^3) \tag{3.2.41}$$

So the ground state energy is simply:

$$E_0 = V \int \frac{d^3p}{(2\pi)^3} \frac{\omega}{2} \quad (3.2.42)$$

So the problem is that we are calculating the total energy, which becomes infinite as volume goes to infinity. What we should be doing is working in terms of an energy density, to get the energy density we simply divide by the volume:

$$\epsilon_0 = \frac{E_0}{V} = \int \frac{d^3p}{(2\pi)^3} \frac{\omega}{2} \quad (3.2.43)$$

However this is still infinite as the limit of p is still infinity. This infinity is called UV divergence (high frequency, or short distance, can be thought of mathematically by remembering that space is a continuum, so as one tries to sum up all the infinitesimally small space regions, the fourier transform of that, which happens to be the momentum will diverge).

There are two ways to think about this infinity, firstly we may argue that the theory doesnt work for arbitrarily large energies (short distances) or we way argue that since in physics we are just concerned with energy differences, and this infinity is present for every state, this infinity can just be subtracted from each state and this leaves the theory unchanged. So we simply redefine the Hamiltonian by subtracting of this infinity (this method is called *normal ordering*):

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} \quad (3.2.44)$$

From this defination, the vacuum enegy is zero (ground state):

$$H|0\rangle = 0 \quad (3.2.45)$$

Definition: We write the *normal order* string of operators $\phi_1(\mathbf{x}_1)\dots\phi_n\mathbf{x}_n$ with colons either side:

$$: \phi_1(\mathbf{x}_1)\dots\phi_n\mathbf{x}_n : \quad (3.2.46)$$

In particular, the Hamiltonian we chosen could be written as:

$$: H := \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} \quad (3.2.47)$$

From now all operators will be normal order.

Gravity

Coming back to the point about the infinities, we said that they can be ignored as we only care about differences in energy. However this cannot possibly be true as gravity couples directly to the energy, so the question is, why doesnt gravity see this infinite amount of energy. One might say that we only know for sure that QFT works only upto energies of TeV's (as that is the energy that the LHC can go upto to test the theory). So surely that means we should have have ground state energies upto the TeV scale, so the cosmological constant or vacuum energy, should be about the TeV scale in energy aswell, but its measured value is far from it, $\approx 10^{-12}eV$. The reason for this

is unknown even today.

Recovering particles

Using the equation for the Hamiltonian it is easy to show the following commutation relations:

$$[H, a_{\mathbf{p}}^\dagger] = \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger \quad (3.2.48)$$

$$[H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}} \quad (3.2.49)$$

So by acting on the Hamiltonian with a creation operator one can create an excitations in the field. If we have a state:

$$|\mathbf{p}\rangle = a_{\mathbf{p}}^\dagger |0\rangle \quad (3.2.50)$$

The energy of this state is:

$$H|\mathbf{p}\rangle = \omega_{\mathbf{p}}|\mathbf{p}\rangle \quad (3.2.51)$$

Remember:

$$\omega_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2 \quad (3.2.52)$$

But this is just the relativistic energy dispersion relation for a particle of mass m and 3-momentum \mathbf{p} , in Minkowski space. This motivates us to interpret $|\mathbf{p}\rangle$ as the state of a single particles of momentum \mathbf{p} and mass m . Note that up until now m has not been defined as the mass, it was simply a free parameter. However, quantising the energy and the comparison with relativistic momentum and mass energy relations, we can define m to be the *mass of a quantum particle*. To check this further we can compute its momentum and see if it agrees with that of a particle of mass m . Recall the classical momentum operator for this theory (derived by considering symmetry and translational invariance):

$$P^i = \int d^3x T^{0i} = \int d^3x \dot{\phi} \partial^i \phi \quad (3.2.53)$$

This can be turned into an operator, such that for KG theory we get:

$$\mathbf{P} = \int d^3x \pi x \nabla \phi(x) = \int \frac{d^3(p)}{(2\pi)^3} \mathbf{p} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \quad (3.2.54)$$

So applying the operator on the state we get:

$$\mathbf{P}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle \quad (3.2.55)$$

Therefore the state $|\mathbf{p}\rangle$, who's energy is the energy of a relativistic particle and the momentum is, \mathbf{p} , which is the momentum three vector of a relativistic particle, so we can be pretty sure that this object is a particle. The final thing one can check is the angular momentum of the particle, more specifically if we measure the angular momentum of a particle with zero momentum, then we will be getting its rotational angular momentum, i.e. its spin! The angular momentum operator is given by:

$$J^i = \epsilon^{ijk} \int d^3x (J^0)^{jk} \quad (3.2.56)$$

Applying this operator on the state which has a particle with zero momentum, say $|\mathbf{p} = 0\rangle$ we have:

$$J^i |p = 0\rangle = 0 \quad (3.2.57)$$

which we interpret as telling us that the particle carries no internal angular momentum. In other words, quantizing a scalar field gives rise to a spin 0 particle.

Fock Space

We can create multi-particle states by acting with many creation operators. So if we operate on the vacuum with n creation operators we get a n particle state:

$$|\mathbf{p}_1, \dots, \mathbf{p}_n\rangle = a_{\mathbf{p}_1}^\dagger \dots a_{\mathbf{p}_n}^\dagger |0\rangle \quad (3.2.58)$$

Since all the a^\dagger 's commute, the state of particles is symmetric under particle exchange:

$$|\mathbf{p}, \mathbf{q}\rangle = |\mathbf{q}, \mathbf{p}\rangle \quad (3.2.59)$$

This shows that the particles here are bosons (integer spin particles, hence the spin is zero as well as we found in the last section). The full Hilbert space of a theory is spanned by acting on the vacuum with all possible combinations of creation operators, this is called *fock space*. There is an important operator that counts the number of particles there, creatively named the *number* operator:

$$N = \int \frac{d^3p}{(2\pi)^3} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \quad (3.2.60)$$

Applying this operator gives:

$$N |\mathbf{p}_1, \dots, \mathbf{p}_n\rangle = n |\mathbf{p}_1, \dots, \mathbf{p}_n\rangle \quad (3.2.61)$$

Another property of this operator is that it commutes with the Hamiltonian, so when can know the energy and the number of particles at the same time. Recalling Ehrenfest's theorem, any operator that commutes with the Hamiltonian gives rise to a conserved quantity, therefore the number of particles in this case are conserved (just means that this is a very dull theory in which nothing really happens). This is only true for free field theories, in general this is not true.

Relativistic Normalisation

This comes from the fact that we don't know whether the theory is Lorentz invariant or not. Suppose we define the ground state is defined as $|0\rangle$ and is normalised as $\langle 0|0\rangle = 1$. The one-particle states $|\mathbf{p}\rangle = a_{\mathbf{p}}^\dagger |0\rangle$ (remember acting on a state with a creation operator generates a particle, since we have acted on using just one creation operator it is a one particle state), this obeys the following relation:

$$\langle \mathbf{p} | \mathbf{q} \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (3.2.62)$$

The question is, whether, $\delta^{(3)}(\mathbf{p} - \mathbf{q})$ is Lorentz invariant. It turns out that it isn't Lorentz invariant as \mathbf{q} and \mathbf{p} are three vectors as supposed to four vectors (hence time and space will not transform equally).

Now we claim that:

$$\int \frac{d^3p}{2E_{\mathbf{p}}} \quad (3.2.63)$$

is Lorentz invariant ($E_{\mathbf{p}} = \sqrt{p^2 + m^2}$)

Proof

$\int d^4p$ is Lorentz invariant since it is a four vector. The relativistic dispersion relation for a massive particle is:

$$E_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2 \quad (3.2.64)$$

is also Lorentz invariant. So combining all the factors we have the following invariant quantity:

$$\int d^4p \delta(p_0^2 - \mathbf{p}^2 - m^2) \quad (3.2.65)$$

After some thought it can be seen that this integral can be rewritten as:

$$\int \frac{d^3p}{2p_0} \quad (3.2.66)$$

but $p_0 = E_{\mathbf{p}}$ which shows that our claim was indeed Lorentz invariant. Now we make a second claim that:

$$2E_{\mathbf{p}}\delta^3(\mathbf{p} - \mathbf{q})$$

is a Lorentz invariant function.

Proof

If we integrate using the result obtained we get:

$$\int \frac{d^3p}{2E_{\mathbf{p}}} \times 2E_{\mathbf{p}}\delta^3(\mathbf{p} - \mathbf{q}) = 1 \quad (3.2.67)$$

And one is Lorentz invariant!

As it turns out we can use this result not just for this example but for everything else. Therefore the relativistically states are normalised as:

$$|p \rangle = \sqrt{2E_{\mathbf{p}}}|E_{\mathbf{p}} \rangle \quad (3.2.68)$$

The overall of these new normalised states also gives a Lorentz invariant quantity:

$$\langle p|q \rangle = (2\pi)^3 2E_{\mathbf{p}}\delta^3(\mathbf{p} - \mathbf{q}) \quad (3.2.69)$$

Notice the change in notations, $|p \rangle$ is only different from $|\mathbf{p} \rangle$ in the fact that it is relativistically normalised in a way that keeps it Lorentz invariant.

3.3 Complex Scalar Fields

One might expect that since the complex scalar field is just made of two scalar fields, we will just get two different types of particles and that is exactly what we get.

Consider first the Lagrangian:

$$L = \partial_\mu \psi^* \partial^\mu \psi - M^2 \psi^* \psi \quad (3.3.1)$$

For a complex scalar field there is no factor of $\frac{1}{2}$ in front of the Lagrangians, as to compute the equation of motions (from EL equation) we can treat ψ and ψ^* as separate fields, the equations of motion are:

$$\partial_\mu \partial^\mu \psi + M^2 \psi = 0 \quad (3.3.2)$$

$$\partial_\mu \partial^\mu \psi^* + M^2 \psi^* = 0 \quad (3.3.3)$$

We want to quantize the Lagrangian in the same way we quantised the real scalar fields. So we expand the complex field operator (reminder: As I stressed before, fields are operator valued functions) in some Fourier basis:

$$\psi(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} (b_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} + c_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}) \quad (3.3.4)$$

and

$$\psi^\dagger = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} (b_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}} + c_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}}) \quad (3.3.5)$$

Now recall the classical conjugate momenta:

$$\pi = \frac{\partial L}{\partial \dot{\psi}} = \dot{\psi}^* \quad (3.3.6)$$

This can be written as a quantum operator valued field as:

$$\pi = \int \frac{d^3 p}{(2\pi)^3} (i) \sqrt{\frac{E_{\mathbf{p}}}{2}} (b_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}} - c_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}}) \quad (3.3.7)$$

$$\pi^\dagger = \int \frac{d^3 p}{(2\pi)^3} (-i) \sqrt{\frac{E_{\mathbf{p}}}{2}} (b_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - c_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}) \quad (3.3.8)$$

Here we have defined the new operators b and c , they are not the same as the a 's in the real field, however individually they are creation and annihilation operators. The commutation relations between these fields are:

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} - \mathbf{y})$$

$$[\psi(\mathbf{x}), \pi^\dagger(\mathbf{y})] = 0$$

$$[\psi(\mathbf{x}), \psi(\mathbf{y})] = 0$$

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = 0$$

$$\begin{aligned}
[b_{\mathbf{p}}, b_{\mathbf{q}}^\dagger] &= (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \\
[c_{\mathbf{p}}, c_{\mathbf{q}}^\dagger] &= (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \\
[b_{\mathbf{p}}, b_{\mathbf{q}}] &= 0 \\
[c_{\mathbf{p}}, c_{\mathbf{q}}] &= 0 \\
[b_{\mathbf{p}}, c_{\mathbf{q}}] &= 0 \\
[b_{\mathbf{p}}, c_{\mathbf{q}}] &= 0 \tag{3.3.9}
\end{aligned}$$

We get two particles of mass M and spin zero (as the fields, when separated, is just a scalar field like the one in the previous section) by applying the creation operator $c_{\mathbf{p}}^\dagger$. The theory also has a conserved charge given by Noether's theorem:

$$Q = \int \frac{d^3p}{(2\pi)^3} (c_{\mathbf{p}}^\dagger c_{\mathbf{p}} - b_{\mathbf{p}}^\dagger b_{\mathbf{p}}) \tag{3.3.10}$$

Look at this carefully this is just the Number operator and it simply gives $N_c - N_b$, that is, the number of particles of type c minus the number of particles of type b . In this case the number of particles does not change as N_c commutes with the Hamiltonian and so does N_b (as this is just two fields as supposed to 1 field as was in the last section). However when we talk about interacting fields the individual particle numbers will not be conserved. But the value of $N_c - N_b$ is always conserved (as long as the symmetry is maintained. Recall the symmetry was:

$$\psi = e^{i\alpha} \psi \tag{3.3.11}$$

So now we have two particles that have the same mass and are conserved in total, hence we can call of the conserved charges a particle and the other an antiparticle! This gives the result that we have always know, the particles and antiparticles are conserved. For real scalar fields we only got one particle and therefore the particle is said to be its own antiparticle. Typically in QFT at relativistic energies there is no way to have the number of particles conserved (this was one of the reasons for developing QFT as stated in the introduction).

3.4 Heisenberg Picture and KG field in space-time

In the last part we switched from the Lagrangian formalism to the Hamiltonian formalism and then in Schroedinger theory we solved the Hamiltonian and interpreted the resulting theory in terms of relativistic particles.

In quantisation of the field we introduced a preferred time coordinate, t . It is not at all obvious that now the theory is still Lorentz invariant after quantisation as the field operator $\phi(x)$ only depends on position and not on time, so space and time are not on the same footing. The state on

the other hand does not depend on space (it is simply defined at every point in space) and it varies in time according to the Schroedinger equation:

$$i \frac{d|\mathbf{p}(t)\rangle}{dt} = H|\mathbf{p}(t)\rangle \quad (3.4.1)$$

where the state is given by:

$$|\mathbf{p}(t)\rangle = e^{-iE_{\mathbf{p}}t}|\mathbf{p}\rangle \quad (3.4.2)$$

However the equality between space and time can be resolved by going from the Schroedinger picture to the Heisenberg picture. At the moment, the operator $\phi(x)$ depends on space only. We can rewrite it in the Heisenberg picture as:

$$\Theta_H(t) = e^{iHt}\Theta_s e^{-iHt} \quad (3.4.3)$$

The new notation is; Θ_s is the operator in the Schroedinger picture and Θ_H is the operator in the Heisenberg picture. We also have the relation:

$$\frac{d\Theta_H}{dt} = i[H, \Theta_H] \quad (3.4.4)$$

This is the equation of motion in the Heisenberg picture. In general we will specify which picture one is working in by $\phi(\mathbf{x})$, meaning Schroedinger picture and $\phi(\mathbf{x}, t)$ meaning the Heisenberg picture. Since the operator in the two pictures agree at a fixed time, say at $t = 0$, the commutation relations become equal to each other at this point:

$$[\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)] = 0$$

$$[\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = 0$$

$$[\phi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (3.4.5)$$

Since the operator is time dependent, we can study how it evolves in time. First consider:

$$\begin{aligned} \dot{\phi} &= i[H, \phi] = \frac{i}{2} \left[\int d^3y \pi(y)^2 + \nabla\phi(y)^2 + m^2\phi(y)^2, \phi(x) \right] \\ \dot{\phi} &= i \int d^3y \pi(y) (-i)\delta^{(3)}(\mathbf{y} - \mathbf{x}) = \pi(x) \end{aligned} \quad (3.4.6)$$

Meanwhile the equation of motion for π reads:

$$\dot{\pi} = i[H, \pi] = \frac{i}{2} \left[\int d^3y \pi(y)^2 + \nabla\phi(y)^2 + m^2\phi(y)^2, \pi(x) \right] \quad (3.4.7)$$

This equation can be separated out by using the commutation relations involving the operators, 3.4.5 :

$$\frac{i}{2} \int d^3y (\nabla_y[\phi(y)\pi(x)])\nabla\phi(y) + \nabla\phi(y)\nabla_y[\phi(y), \pi(x)] + 2im^2\phi(y)\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (3.4.8)$$

Integrating by parts gives:

$$\dot{\pi} = -\left(\int d^3y(\nabla_y^{(3)}(\mathbf{x} - \mathbf{y})\nabla_y\phi(\mathbf{y})) - m^2\phi(x)\right) = \nabla^2\phi - m^2\phi \quad (3.4.9)$$

Combining with we get:

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0 \quad (3.4.10)$$

Which is just the KG equation (remember ϕ is not a wavefunction here). We want the mode (Fourier) expansion for $\phi(\mathbf{x}, t)$ in the Heisenberg picture, for which we need the following relations involving the creation and annihilation operators:

$$e^{iHt}a_{\mathbf{p}}e^{-iHt} = e^{-iE_{\mathbf{p}}t}a_{\mathbf{p}} \quad (3.4.11)$$

$$e^{iHt}a_{\mathbf{p}}^\dagger e^{-iHt} = e^{iE_{\mathbf{p}}t}a_{\mathbf{p}}^\dagger \quad (3.4.12)$$

These follow from the commutation relations:

$$[H, a_{\mathbf{p}}] = -E_{\mathbf{p}}a_{\mathbf{p}}$$

$$[H, a_{\mathbf{p}}^\dagger] = -E_{\mathbf{p}}a_{\mathbf{p}}^\dagger \quad (3.4.13)$$

Therefore we get:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} (a_{\mathbf{p}}e^{-ipx} + a_{\mathbf{p}}^\dagger e^{ipx}) \quad (3.4.14)$$

New notation denotes Heisenberg picture:

$$\phi(x) = \phi(\mathbf{x}, t)$$

This is similar to the expression for $\phi(x)$ in the Schroedinger picture except that the exponents have a sign change. This comes about due to the minkowski metric contraction and we have a momentum four vector now (it was a 3 vector in the Schroedinger equation), now we should have Lorentz invariance in the theory:

$$p \cdot x = E_{\mathbf{p}}t - \mathbf{p} \cdot \mathbf{x} \quad (3.4.15)$$

When we work in the Schroedinger picture we are working with operators that are independent of time, the Heisenberg picture includes the time dependence. So the question is, how do the commutation relations change in the Heisenberg picture. We have seen in Eq 3.4.5 that the commutation relations are the same at a fixed point (as expected), but what about at an arbitrary space-time separation.

To examine this, first imagine space-time diagram with an operator, Θ_1 at the origin ($t = 0$ line), now if Θ_2 (another operator) was on the same line (i.e $t = 0$), then we expect the operators to commute. Now suppose the second operator was not on the $t = 0$ line but at some arbitrary space-time point that lies *outside* the light cone of the first operator. But for a different observer, it is possible that both the operators are on the $t = 0$ line (or any line of constant time). Since we expect all the physics to be the same in any reference frame, it had better be that Θ_1 and Θ_2 commuted at every position in space-time:

$$[\Theta_1(x), \Theta_2(y)] = 0 \quad (3.4.16)$$

for any space-time separation (x and y represent arbitrary space time coordinates). This is another statement of causality, saying that any measurement made by Θ_2 (outside the light cone) cannot affect any measurement made inside the light cone where Θ_1 is. To check weather this is true, let's define:

$$\Delta(x - y) = [\phi(x), \phi(y)] \quad (3.4.17)$$

The R.H.S is:

$$[\phi(x), \phi(y)] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} (e^{-ip \cdot (x-y)} - e^{ip \cdot (x-y)}) \quad (3.4.18)$$

Therefore:

$$\Delta(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} (e^{-ip \cdot (x-y)} - e^{ip \cdot (x-y)}) \quad (3.4.19)$$

Things to note about this relation:

- The R.H.S will integrate to give a c-number function, that means just an integer! (this is because it is normalised with the factor of $E_{\mathbf{p}}$)
- It is Lorentz invariant as we have a four vector in the exponent (not a three vector! and the fact that the integral is just a number)

It doesn't vanish for time-like seperated events:

$$(x - y) = (t, 0, 0, 0) \quad (3.4.20)$$

This gives:

$$[\phi(\mathbf{x}, 0), \phi(\mathbf{x}, t)] \approx (e^{-imt} - e^{imt}) \quad (3.4.21)$$

- It vanishes for space-like seperations. As \mathbf{p} is simply an integration variable and we know that $\Delta(x - y) = 0$ for equal times so we can flip the sign of \mathbf{p} .

4 Propagators

4.1 Introduction to propagators

The final section here is going to be on another important concept of operators. Consider the following question:

We disturb the field at point y and ask what is the amplitude:

$$\langle 0|\phi(x)\phi(y)|0\rangle \quad (4.1.1)$$

This is basically a statement of measuring the field at point y and x , and then finding the amplitude for the ground state. To compute we plug in the mode expansions:

$$\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3p d^3p'}{(2\pi)^6} \frac{1}{\sqrt{4E_{\mathbf{p}}E_{\mathbf{p}'}}} \langle 0|a_{\mathbf{p}}a_{\mathbf{p}'}^\dagger|0\rangle e^{ipx+ip'y} \quad (4.1.2)$$

To be clear about the integral, it has $a_{\mathbf{p}}$ and $a_{\mathbf{p}'}^\dagger$ terms. However $a|0\rangle = 0$ and $\langle 0|a^\dagger = 0$, so we are only left with the $|a_{\mathbf{p}}a_{\mathbf{p}'}^\dagger|$ term. To compute this expression we commute $a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger$ which gives a delta function, over which we can integrate over to give:

$$\langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-p(x-y)} \equiv D(x-y) \quad (4.1.3)$$

- $D(x-y)$ is the propagator, it depends only on the difference between x and y , which is a four-vector difference.
- So for space like separation, $(x-y)^2 < 0$, this integral does not vanish:

$$D(x-y) \approx e^{-m|\mathbf{x}-\mathbf{y}|} \quad (4.1.4)$$

So it can be shown that it is just an exponential decay, but not zero outside the light cone. It can be thought of as the field leaks outside the light cone (almost analogous to the phenomena of tunneling and a particle looking through a potential). I must stress that this does not violate causality, because we are not talking about a measurement here. In fact we have previously shown that the space-time measurement commutes and the theory does not violate causality. The commutator relation can be written as:

$$[\phi(x), \phi(y)] = D(x-y) - D(y-x) = 0 \quad (4.1.5)$$

Four space-like separations but not for time like separations. So to put everything we have discussed so far in words, we can think of $\phi(y)$ representing a particle being made at y and $\phi(x)$ representing a particle being detected at x . Therefore $D(x-y)$ represents the particle travelling from $y \rightarrow x$ and $D(y-x)$ represents another particle with the same mass travelling in the opposite direction. This is therefore called the anti-particle. Therefore the Amplitude of one particle travelling in one direction, is exactly canceled by the amplitude the other particle travelling in the opposite direction, hence there is no probability that a particle is found outside the light cone (and causality is preserved).

4.2 Feynmann propagator

The important quantity for the interacting theories is the Feynmann propagator, defined as:

$$\begin{aligned}\nabla_F(x-y) &= \langle 0|T\phi(x)\phi(y)|0 \rangle = D(x-y), x^0 > y^0 \\ \nabla_F(x-y) &= \langle 0|T\phi(x)\phi(y)|0 \rangle = D(y-x), x^0 < y^0\end{aligned}\tag{4.2.1}$$

Here T stands for a function called time ordering, basically it orders operators evaluated at later times towards the left and the operators evaluated at earlier times to the right. Now we claim:

The Feynman propagator can be written as a 4-momentum integral:

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip \cdot (x-y)}\tag{4.2.2}$$

For the first time, so far, p_0 is not a fixed term, until now everything we called p_0 has been equal to $E_{\mathbf{p}} = \sqrt{p^2 + m^2}$, the relativistic energy, but now we shall integrate over all values of p_0 (it appears in the p 's in the equation above which are the four-momentum vector). The problem you might notice is that the integral here is not well defined. For a fixed \mathbf{p} (3 momentum), if we integrate over p_0 , at some point the term, $\frac{1}{p^2 + m^2}$ will diverge.

This singularity is sometimes called a *pole* and it occurs when $p_0 = \pm\sqrt{p^2 + m^2}$. So we need a method to make this a well defined integral. To do this we will integrate along the contour in which p_0 is along the vertical imaginary axis. We define the integral by the contour which has p_0 on an imaginary axis perpendicular to the real axis (look at figure on top of page 31 of Schroeder and Peskin).

The integral we are trying to compute is just the real axis of p_0 . There are two points over which I can run into singularities on the real axis and they are the point where $p_0 = -E_{\mathbf{p}}$ and when $p_0 = E_{\mathbf{p}}$. The idea now is to go along the real axis and dip below the $-E_{\mathbf{p}}$ pole and go above the $E_{\mathbf{p}}$ pole and then finish off the integral.

To understand Contour integrals¹⁵ I find it easiest to think about them in terms of vector calculus, where we integrate over different loops and surfaces. This is just the same concept, if we just ignore the fact that the axis is not complex just the same axis in a plane, we are simply integrating over a different path.

Residue/Cauchy's theorem

Consider a function $f(z)$, that depends on a complex number z . It is an analytic function for all z except at a few singular points. Let Γ be a closed anti-clockwise contour in a given space, the integral of the contour is the sum of the residues inside the contour:

$$\oint_{\Gamma} f(z) dz = 2\pi i \sum_{\infty} \text{Residues}\tag{4.2.3}$$

¹⁵This is the first time I have come across contour integrals

Where the integral is over the contour in an anticlockwise direction, residues are singularities. To define them correctly, first we consider the Laurent expansion of the function $f(z)$:

$$f(z) = \sum_0^{\infty} (z - z_0)^n + \frac{c_1}{z - z_0} + \frac{c_2}{(z - z_0)^2} + \dots \quad (4.2.4)$$

for $0 < |z - z_0| < R$, the radius of convergence. The coefficient, c_1 , is defined as the residue.

Now we can finally return to proving the claim that was made about writing the Feynmann propagator as an integral. Firstly we write:

$$\frac{1}{p^2 - m^2} = \frac{1}{(p^0)^2 - E_{\mathbf{p}}^2} = \frac{1}{(p^0 - E_{\mathbf{p}})(p^0 + E_{\mathbf{p}})} \quad (4.2.5)$$

So going back to the contour integral in the complex plane, we want to integrate from before $-E_{\mathbf{p}}$ pole and after $E_{\mathbf{p}}$ pole (look at the diagram in Peskin and Schroeder as given above). But we know that if we can somehow get a closed loop to integrate over then we can just use Cauchy's theorem to just add up the residues inside the contour. So we choose a contour that either goes all the way to $-\infty$ or all the way to $+\infty$, and the contour is chosen in such a way that it adds zero to the integral (because we can't just change the integral!). To remind you, the integral we are trying to evaluate is in Eq 4.2.2.

and we are doing the integral over p_0 . First let's look at the point when $x_0 > y_0$, we close the contour in the lower-half plane, $p_0 \rightarrow i\infty$ therefore the exponential term becomes:

$$e^{-ip_0(x^0 - y^0)} \rightarrow e^{-\infty} \rightarrow 0 \quad (4.2.6)$$

therefore it adds zero integral, as we required. So now the integral over \mathbf{p} is:

$$\Delta_F(x - y) = \int \frac{d^3p}{(2\pi)^4} \frac{-2i\pi}{2E_{\mathbf{p}}} i e^{-iE_{\mathbf{p}}(x^0 - y^0)} e^{ip(\mathbf{x} - \mathbf{y})} \quad (4.2.7)$$

Now we can flip the sign of \mathbf{p} which is valid since we integrate over d^3p and all other quantities depend only on \mathbf{p}^2 , so we get:

$$\int \frac{d^3}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-ip \cdot (x - y)} \equiv D(x - y) \quad (4.2.8)$$

Which is indeed the Feynmann propagator, so we have proved the claim we made in Eq 4.2.2. A similar method can be applied when $y^0 > x^0$. This time we close the contour in the anti-clockwise direction in the upper half of the plane to counter the change in sign of the exponent, and following the same method we get the following equation:

$$\Delta_F(x - y) = D(y - x), \quad y^0 > x^0 \quad (4.2.9)$$

Once again we reproduce the Feynmann operator, in this case for a particle going from x to y . Instead of worrying about defining what contour is being chosen, a new notation is generally implemented. The Feynmann propagator is generally written as (including in most text books like Peskin and Schroeder):

$$\Delta(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{-ip \cdot (x-y)}}{p^2 - m^2 + i\epsilon} \quad (4.2.10)$$

x

Here $\epsilon > 0$, and is an infinitesimal. This basically shows that the poles are slightly off the real axis and has the same effect as contour integration. This is called the "i ϵ prescription".

Another way to think about the Feynmann propagator is that it is a green function of the KG operator:

$$\begin{aligned} (d_t^2 - \nabla^2 + m^2)\Delta_F(x-y) &= \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} (-p^2 + m^2) e^{-ip \cdot (x-y)} \\ (d_t^2 - \nabla^2 + m^2)\Delta_F(x-y) &= -i \int \frac{d^4p}{(2\pi)^4} e^{-ip \cdot (x-y)} \\ (d_t^2 - \nabla^2 + m^2)\Delta_F(x-y) &= -i\delta^{(4)}(x-y) \end{aligned} \quad (4.2.11)$$

We didn't use the contour integral here; we could have taken different contours in getting to here by going over both poles (known as *retarded* contour) or below them (*advanced* contour). When we take the contour that goes above both poles, we get a *retarded* Green function. This green function representation is useful when we are applying an operator to an unknown function but the result of applying the operator is known; so we can simply integrate the operator over the known function to obtain the unknown function.

4.3 Chapter 2, Peskin and Schroeder Problems

Problem 2.1

Electromagnetism properties:

Action:

$$S = \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (4.3.1)$$

Now we want to derive Maxwell's equations. These are just the equations of motion that come from the EL equations, the dynamical variable is $A_\mu(x)$. I will follow the procedure that I have stated in the very beginning of section 2.2 here, so first we compute:

$$\frac{\partial F_{\mu\nu}}{\partial(\partial_\lambda A_k)} = \delta_\mu^\lambda \delta_\nu^k - \delta_\nu^\lambda \delta_\mu^k \quad (4.3.2)$$

$$\partial F_{\mu\nu} \partial A_\lambda = 0 \quad (4.3.3)$$

For clarity on how the derivatives transform according to the indices look in the notation section. Now applying Eq 4.3.2 we have:

$$\frac{\partial}{\partial(\partial_\lambda A_k)} (F_{\mu\nu} F^{\mu\nu}) = 4F^{\lambda k} \quad (4.3.4)$$

Now we substitute into EL to obtain the equation of motion for the field:

$$\partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu A_\nu)} \right) - \frac{\partial L}{\partial A_\nu} = -\partial_\mu F^{\mu\nu} \quad (4.3.5)$$

This shows how the EM field propagates in free space and is therefore equivalent to Maxwell's equation. To obtain the classical Maxwell equations we can rewrite Eq 4.3.1 as (this is known as the Bianchi Identity):

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\mu\lambda} = 0 \quad (4.3.6)$$

If take the electric field to be:

$$E^i = -F^{0i} = \mathbf{E} \quad (4.3.7)$$

and the magnetic field is:

$$\epsilon^{ijk} B^k = -F^{ij} = \mathbf{B} \quad (4.3.8)$$

Eq 4.3.6 gives the following two Maxwell equations:

$$\partial^i B^i = 0$$

$$\epsilon^{ijk} B^k - \partial^0 E^i = 0 \quad (4.3.9)$$

and the equations of motion give the following relations:

$$\partial^i E^i = 0$$

$$\epsilon^{ijk} \partial^j E^k = 0 \quad (4.3.10)$$

Together these comprise to give the full set of Maxwell's equations in free space with no sources.

Now we want to construct the energy-momentum tensor for this theory. We use Noethers theorem for this, (section 2.1). The energy momentum tensor is a conserved current, therefore obeys:

$$\partial_\nu T_\nu^\mu = 0 \quad (4.3.11)$$

Applying Noethers theorem we have:

$$\tilde{T}^{\mu\nu} = \frac{\partial L}{\partial(\partial_\mu A_\lambda)} \partial^\nu A_\lambda - \eta^{\mu\nu} L \quad (4.3.12)$$

But we know that the Lagrangian for an EM field is:

$$L = -\frac{1}{4} F_{\lambda k} F^{\lambda k} \quad (4.3.13)$$

So we can rewrite the energy-momentum tensor as:

$$\tilde{T}^{\mu\nu} = -F^{\mu\nu} \partial^\nu A_\lambda + \frac{1}{4} \eta^{\nu\mu} F_{\lambda k} F^{\lambda k} \quad (4.3.14)$$

Clearly this is not a symmetric tensor, so now we add a term $\partial_\lambda K^{\lambda\mu\nu}$ (this is discussed in the section Application of Noethers theorem above):

$$K^{\lambda\mu\nu} = -K^{\mu\lambda\nu} \quad (4.3.15)$$

So now the energy-momentum tensor is:

$$T^{\mu\nu} = \tilde{T}^{(\mu\nu)} + \partial_\lambda K^{\lambda\mu\nu} = F^{\mu\nu} F_\lambda^\nu + \frac{1}{4} \eta^{\mu\nu} F_{\lambda k} F^{\lambda k} \quad (4.3.16)$$

This is also called the Belinfante tensor, it can be rewritten as:

$$T^{00} = \frac{1}{2} (E^i E^i + B^i B^i) \quad (4.3.17)$$

Which is equivalent to :

$$\epsilon = \frac{1}{2} (E^2 + B^2) \quad (4.3.18)$$

and:

$$T^{i0} = \epsilon^{ijk} E^j B^k \quad (4.3.19)$$

which is equivalent to:

$$\mathbf{S} = \mathbf{E} \times \mathbf{B} \quad (4.3.20)$$

5 Interacting Theories

So far we have only studied free field theories, which are basically theories that have a Lagrangian that is *quadratic* in terms of the fields and its derivatives. This leads to equations of motion that are *linear* in the fields, these equations can be solved exactly, which lead to multi-particle states with fixed momenta and mass.

5.1 Introduction to interactions

To introduce interactions we can add higher order terms to the Lagrangian, like ϕ^4 , this leads to interactions between particles. In general, higher order terms can be introduced into a Lagrangian as follows:

$$L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \sum_{n=3}^{\infty} \frac{\lambda_n}{n!} \phi^n \quad (5.1.1)$$

These λ_n are called coupling constants, so when we say we want small perturbations, it means something like saying the coupling constants are small. To explore exactly what small perturbations are we do some dimensional analysis:

$[S] = 0$ since S has dimensions of \hbar which we have set to 1, which is just a pure number. But the action is the integral over all space of the Lagrangian density:

$$S = \int d^4x L \quad (5.1.2)$$

Since d^4x has dimensions of -4 and S has dimensions of 0 , the Lagrangian density must naturally have units of 4 . Therefore, $[\partial_\mu] = 1$. Therefore comparing to the Lagrangian above, the dimensions of ϕ must be 1 and m must be 1 (as expected from $E = mc^2$ and $c = 1$). So for $L = 4$, λ_n must have units of $4 - n$. Now there is a problem, we can't simply say that lets make λ_n small, i.e $\lambda_n \ll 1$ since λ_n is not just a pure number. So to make it small, the λ_n has to be given in relation to another quantity with the same dimensions (it will become more clear as we discuss some examples next):

$[\lambda_3] = 1$ has dimensions on mass. So we want to make, $\frac{\lambda_3}{E}$ small. The real requirement is that the new perturbation function is small compared to the first two terms. This type of coupling is known as *relevant*.

$\frac{\lambda_4 \phi^4}{4!}$, $[\lambda_4] = 0$, therefore for this case the coupling constant is just a number and can be made arbitrarily small. These type of perturbations are called *marginal* perturbations.

$\frac{\lambda_n \phi^n}{n!}$, $[\lambda_n] < 0$, in this case the coupling constant has units of negative mass, so it is small for low energies and high for high energies and is therefore known as the *irrelevant* perturbations.

These are all weakly interacting perturbations that we are dealing with now, however a lot of interesting physics goes on when we have strongly interacting systems (the extra perturbation terms added to the Lagrangian are of the same order of magnitude as the first two terms), an example is the fractional quantum hall effect.

5.2 Interaction picture

This is a useful trick to deal with small perturbations. We know that in the Schroedinger picture, it is the states that evolve in time:

$$\frac{d|\psi_s\rangle}{dt} = H|\psi_s\rangle \quad (5.2.1)$$

Here we have chosen s to denote that we are working in the Schroedinger picture. While the operators Θ_s are independent of time. The Heisenberg picture is the opposite, states are fixed in time and operators evolve. The interaction picture is a hybrid of the two:

$$H = H_0 + H_{int} \quad (5.2.2)$$

In the Lagrangian, H_0 is the free field Hamiltonian that we have solved already and H_{int} consists of the interaction pieces. The time dependence of the operator is governed by H_0 , which means without the H_{int} we will just be in the Heisenberg picture. We define:

$$|\psi_I(t)\rangle = e^{iH_0 t} |\psi_s(t)\rangle \quad (5.2.3)$$

$$\Theta_I(t) = e^{iH_0 t} \Theta_s e^{-iH_0 t} \quad (5.2.4)$$

The I index corresponds to the interaction picture:

$$H_I \equiv (H_{int})_I \quad (5.2.5)$$

This is the interaction part of the interaction picture:

$$H_I \equiv e^{iH_0 t} (H_{int})_s e^{-iH_0 t} \quad (5.2.6)$$

Now the SE in the Schroedinger picture is just:

$$i \frac{d|\psi_s\rangle}{dt} = H_s |\psi_s\rangle \quad (5.2.7)$$

Now substitute equations defined above into the $|\psi_s\rangle$:

$$i \frac{d|\psi_s\rangle}{dt} = i \frac{d}{dt} (e^{iH_0 t} |\psi_I\rangle) \quad (5.2.8)$$

Now it is easily shown that:

$$i \frac{d|\psi_I\rangle}{dt} = H_I(t) |\psi_I\rangle \quad (5.2.9)$$

This is the SE in the interaction picture and it is solved by Dyson's formula:

$$|\psi(t)_I\rangle = U(t, t_0) |\psi(t_0)_I\rangle \quad (5.2.10)$$

So we have the time evolution operator in the interaction picture. The relations involving the unitary time operator $U(t, t_0)$ are:

$$U(t_1, t_2) U(t_2, t_3) = U(t_1, t_3)$$

$$U(t, t) = 1 \quad (5.2.11)$$

Substituting $|\psi(t)_s\rangle$ in the Schroedinger equation:

$$i \frac{dU}{dt} = H_I(t)U \quad (5.2.12)$$

The question now is, how do we solve an equation that involves just operators (they are all matrices). If U and H were functions, we could solve this by:

$$U(t, t_0) = \exp(-i \int_0^t H_I(t') dt') \quad (5.2.13)$$

This will not work for our problem, as we have operators that are represented by matrices. Another thing is that the operators are generally non-commutative:

$$[H_I(t), H_I(t')] = 0, t \neq t' \quad (5.2.14)$$

Since $i \int_0^t H_I(t') dt'$ is a matrix, it doesn't make sense to raise it to a power. The way we deal with this is, expanding the exponential by a Taylor expansion, just to remind you of the Taylor expansion for a general function, $f(x)$, expanded about a point, a :

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \dots \quad (5.2.15)$$

Now we expand the exponential around 0:

$$U(t, t_0) = 1 + [-i \int_0^t H_I(t') dt'] + \dots \quad (5.2.16)$$

The problem arises when we have H 's in the quadratic term of the expansion as typically they are evaluated at different times and do not commute. The best way to solve 5.2.12 is to first state the solution and then prove it.

Claim: The correct solution is given by Dyson's formula:

$$U(t, t_0) = T \exp(-i \int_0^t H_I(t') dt') \quad (5.2.17)$$

This T we have seen before and it is just the time ordering operator (operators evaluated at later times are moved to the left). This means:

$$U(t, t_0) = 1 - i \int_{t_0}^t H_I(t') dt' + (-i)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' H_I(t') H_I(t'') + \dots \quad (5.2.18)$$

Proof: Under the T operator, everything commutes, as the T operator just reorders everything! Therefore:

$$i \frac{d}{dt} T \exp(-i \int_{t_0}^t dt' H_I(t')) = T [H_I(t) \exp(-i \int_{t_0}^t dt' H_I(t'))]$$

The $H_I(t)$ is clearly going to be evaluated at the latest time, so it can be moved outside the integral, which completes the proof:

$$i \frac{d}{dt} T \exp(-i \int_{t_0}^t dt' H_I(t')) = H_I(t) T \exp(-i \int_{t_0}^t dt' H_I(t')) \quad (5.2.19)$$

In general time order integrals are very difficult to evaluate and are therefore very rarely used.

5.3 Introduction to Scattering

Firstly we choose a theory in which we shall examine scattering. The first theory we shall use is *scalar - Yukawa theory*. Two fields make up the Lagrangian:

$$L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \partial_\mu \psi^\dagger \partial^\mu \psi - \frac{1}{2} m^2 \phi^2 - M^2 \psi \psi^\dagger \quad (5.3.1)$$

ϕ is the real field

ψ is the complex field

m is the mass of the particle that comes from the quantisation of the scalar field that is real

M is the mass of the particle arising from the quantisation of the complex field.

Now we shall add a term that makes the particles from ϕ interact with particles from ψ , the term is:

$$- g \psi^\dagger \psi \phi$$

So the Lagrangian becomes:

$$L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \partial_\mu \psi \partial^\mu \psi - \frac{1}{2} m^2 \phi^2 - M^2 \psi^\dagger \psi - g \psi^\dagger \psi \phi \quad (5.3.2)$$

Notice that according to the perturbation terms that were described previously this additional term is a relevant perturbation (it has units of m). So to make this term small, g has to be smaller than a particular energy scale. We take $g \ll \frac{M}{m}$ to ensure that we have a weakly interacting theory. This means our Hamiltonian will now have a H_0 which will be the first four terms of the Lagrangian above (they will be the free terms, corresponding to the usual free field picture). The last term will be, what we called the interaction term:

$$H_{int} = g \psi^\dagger \psi \phi \quad (5.3.3)$$

The sign change is simply due to the change from a Lagrangian (T-V) to the Hamiltonian (T+V). We just saw in Dyson's formula, that if we have a state, the time evolution of that state, is given by acting on the state by the exponential (of H_{int}). The expansion of the unitary operator shows that the particle number is not conserved as, $\phi \approx a + a^\dagger$, this can create and destroy ϕ particles (mesons).

$\psi \approx b + c^\dagger$, this can destroy ψ particles (b operator) and create ψ^\dagger particles (c^\dagger), *wecallthesenucleons*.

$\psi^\dagger \approx b^\dagger + c$, so it can create nucleons through b^\dagger and destroy anti-nucleons through.

So we see that the number of mesons will not be conserved, but N_c (Number of particles created by c) - N_b (Number of particles created by b), is always conserved. At first-order in perturbation theory we have:

$$\psi^\dagger \psi \phi \approx c^\dagger b^\dagger a$$

c^\dagger comes from ψ , b^\dagger comes from ψ^\dagger and a comes from ϕ . There are eight terms in total for the first order perturbation, however the rest of the terms give zero terms, so can be ignored. This particular operator will kill the meson (a) and create a nucleon (c^\dagger) and an anti-nucleon (b^\dagger), (makes an nucleon, anti-nucleon pair). At second order in perturbation theory (so the Taylor expansion of the exponential is expanded out to the quadratic term) we get terms like:

$$(c^\dagger b^\dagger a)(cba^\dagger)$$

This gives $\psi + \psi' \rightarrow \phi \rightarrow \psi + \psi'$ scattering. This is because (cba^\dagger) kills a nucleon and an anti-nucleon and creates a meson, and $(c^\dagger b^\dagger a)$ creates a nucleon and an anti-nucleon and kills a meson.

To summarise the perturbation effects:

- The interaction Hamiltonian, becomes the time evolution operator
- Then we expand out this operator using the Taylor expansion, so get a terms like: $1 + H + H^2 + \dots$
- Each of these terms causes interactions.

What we want to do is compute the amplitude and the probability of these scattering events actually happening. Suppose we have a state that consists of a ψ particle and an ψ particle and I throw them towards each other at a given momentum and I want to know, what is the probability that they come at with a different momentum, and in particular what is the probability that they come out with different momentum by first annihilating into a meson and then being re-emitted.

Alternatively I could have a meson and I want to know the quantum amplitude, that it will disappear and turn into a nucleon-antinucleon pair, or better still, I want to know the Half-life of the meson. To calculate amplitudes for these processes, we first need an important (and slightly dodgy) assumptions. The assumption is; The initial and final state look like non-interacting particles. This means that the initial state $|i\rangle$ at $t \rightarrow -\infty$ and the final state $|f\rangle$ at $t \rightarrow +\infty$ are eigenstates of H_0 .

It seems very hard to deal with bound states, for example; if initially I have an electron and a proton and I bring them close together. Suppose they combine and form a hydrogen atom, then at any time in the future they are certainly not non-interacting!

A much more worrying issue is the fact that in QFT no particles are ever isolated, instead they are surrounded by a cloud of virtual particles. This problem is dealt with renormalisation, so for now we accept this assumption and move forward.

The amplitude is given by:

$$\langle f|U(t_+, t_-)|i\rangle \equiv \langle f|S|i\rangle \tag{5.3.4}$$

The operator $U(t_+, t_-)$ is called the unitary operator and is defined as the S-matrix (S stands for scattering)

Meson Decay

We are going to compute the amplitude for a meson in the initial state to decay into two nucleons in the final state.

Initial state:

$$|i\rangle = \sqrt{2E_p} a_{\mathbf{p}}^\dagger |0\rangle \quad (5.3.5)$$

Final state:

$$|f\rangle = \sqrt{4E_{\mathbf{q}_1} E_{\mathbf{q}_2}} b_{\mathbf{q}_1}^\dagger c_{\mathbf{q}_2}^\dagger |0\rangle \quad (5.3.6)$$

So I create a single nucleon and an anti-nucleon with momentum \mathbf{q}_1 and \mathbf{q}_2 . We are trying to compute:

$$\langle f|S|i\rangle \quad (5.3.7)$$

This is the amplitude for a meson to decay into two nucleons.

Finally, we replace S with the exponential term that comes from the interaction of the Hamiltonian. The expansion of the exponential, firstly gives a 1, but this has no amplitude since $\langle f|i\rangle = 0$ as the initial and final states have no overlap (by the assumption we made). The terms we are going to be interested in are the ones after the first term, the seconde term in the expansion is $\approx \psi^\dagger \psi \phi$:

$$\langle f|S|i\rangle = -ig \langle f| \int d^4x \psi^\dagger(x) \psi(x) \phi(x) |i\rangle \quad (5.3.8)$$

The interaction Hamiltonian is infact:

$$\int d^3x \psi(x) \psi^\dagger(x) \phi(x) \quad (5.3.9)$$

The reason we have a d^4x is because we are dealing with the S matrix, which is evaluated over time, from $t \rightarrow -\infty$ to $t \rightarrow +\infty$. There will be other terms in the expression for the S matrix, but they will all be smaller in g , so we will just ignore the other terms. Now we use the mode expansions of ψ 's and ϕ 's to try and evaluate this expression. Firstly substitute for ϕ and the state $|i\rangle$:

$$\langle f|S|i\rangle = -ig \langle f| \int d^4x \psi^\dagger(x) \psi(x) \int \frac{d^3k}{(2\pi)^3} \frac{\sqrt{2E_{\mathbf{p}}}}{\sqrt{2E_{\mathbf{k}}}} a_{\mathbf{k}} a_{\mathbf{p}}^\dagger e^{ikx} |0\rangle \quad (5.3.10)$$

Now if we commute $a_{\mathbf{k}}, a_{\mathbf{p}}^\dagger$ we get $\delta^{(3)}(\mathbf{p} - \mathbf{k})$ and then we integrate over the delta function to give:

$$\langle f|S|i\rangle = -ig \langle f| \int d^4x \psi^\dagger(x) \psi(x) e^{-i\mathbf{p}\cdot\mathbf{x}} |0\rangle \quad (5.3.11)$$

Now we expand out ψ and ψ^\dagger . To get non-zero overlap with $\langle f|$, only the b^\dagger and c^\dagger contribute as $b^\dagger|0\rangle = \psi^\dagger$ and $c^\dagger|0\rangle = \psi$, so we have:

$$\langle f|S|i\rangle = -ig \langle 0| \iint \frac{d^4x d^3k_1 d^3k_2}{(2\pi)^3} \sqrt{\frac{E_{\mathbf{q}_1} E_{\mathbf{q}_2}}{E_{\mathbf{k}_1} E_{\mathbf{k}_2}}} c_{\mathbf{q}_2} b_{\mathbf{q}_1} c_{\mathbf{k}_1}^\dagger b_{\mathbf{k}_2}^\dagger e^{i(k_1+k_2-p)\cdot x} |0\rangle$$

The $c_{\mathbf{q}_2}$ and $b_{\mathbf{q}_1}$ terms can be ignored as they act on $|0\rangle$ to annihilate them. So we are left with:

$$\langle f|S|i\rangle = -ig \langle f| \int d^4x \int \frac{d^3k_1 d^3k_2}{(2\pi)^6} \frac{1}{\sqrt{2E_{k_2}k_1}} c_{\mathbf{k}_1}^\dagger b_{\mathbf{k}_2}^\dagger e^{i(k_1+k_2-p)x} |0\rangle \quad (5.3.12)$$

Now we substitute for δ — which gives delta functions, as we had above, and we integrate over them to give:

$$\langle f|S|i\rangle = -ig \langle 0| \int d^4x e^{i(q_1+q_2-p)\cdot x} |0\rangle \quad (5.3.13)$$

Computing this final integral we get our final expression:

$$\langle f|S|i\rangle = -ig(2\pi)^4 \delta^{(4)}(q_1 + q_2 - p) \quad (5.3.14)$$

Here k_1 and k_2 have become q 's because the $\langle f|$ function gave a delta function, which was integrated over k 's to pick out the q 's. This is the first, non-trivial thing we have calculated in QFT, and it's the amplitude of a meson to decay into a nucleon and an anti-nucleon and we see that it is proportional to g which is the coupling constant for this particular decay.

This is the amplitude for a meson of momentum p , to decay into two nucleons with momenta q_1 and q_2 . The delta function is basically showing that the initial momentum is equal to the final momentum ($q_1 + q_2$), otherwise the amplitude is 0, therefore it is a statement of conservation of momentum.

5.4 Wick's Theorem

Recall that the unitary operator, that gives the time evolution, is the time ordered exponential:

$$U = T \exp(-i \int d^4x H_I(x)) \quad (5.4.1)$$

of the interaction Hamiltonian in the interaction picture, this comes from Dyson's formula (the solution to the SE in the interaction picture). We generally need to compute things like:

$$\langle f|T\{H_I(x_1)\dots H_I(x_n)\}|i\rangle \quad (5.4.2)$$

The final state $\langle f|$, multiplied by a product of many Hamiltonians multiplied together, acting on an initial state, $|i\rangle$ (integrated over all x 's). Each of the interaction Hamiltonians contains some fields, we then expand these fields in a mode expansion in terms of the creation and annihilation operators, and then we hope that a bunch of these product terms are going to vanish by annihilating the initial states.

We saw this in the meson decay, when we expanded in terms of a 's, b 's, c 's operators, we could have eight different terms in the first order expansion, but only one of them was relevant (the rest would simply annihilate the initial states). But that example was a simple theory, in general there can be many annihilation and creation operators in the product of the interaction Hamiltonian and one will have to move them through all the other operators to the right. So what we get is a bunch of delta functions along the way, to finally reach the initial state.

What we would like to do, is come up with some formal way that computes the trivial way of going through the calculation. Something like an algorithm. We want all the annihilation operators to be towards the initial state and all the creation operators towards the final state. In any calculation we'd like to rewrite the time ordered operator as a normal ordered operator, where all the annihilation operators are to the right.

Example: Real scalar field

Field is $\phi(x)$, which can be written as:

$$\phi(x) = \phi^+(x) + \phi^-(x) \tag{5.4.3}$$

Here the ϕ 's are defined to be:

$$\phi^+(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} a_{\mathbf{p}} e^{-ip \cdot x} \tag{5.4.4}$$

$$\phi^-(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} a_{\mathbf{p}}^\dagger e^{ip \cdot x} \tag{5.4.5}$$

These are simply the mode expansions that we have used over and over again, they are simply redefined in terms of ϕ^+ and ϕ^- . The p and x in the exponent are four vectors (as we are in the Heisenberg picture, free field theory, which is the same as the interaction picture without the H_{int} term). Look at time ordered operator:

$$T\phi(x)\phi(y) = \phi(x)\phi(y), x^0 > y^0 \tag{5.4.6}$$

$$\phi(x)\phi(y) = (\phi^+(x) + \phi^-(x))(\phi^+(y) + \phi^-(y)) \tag{5.4.7}$$

Multiplying the terms out:

$$\phi(x)\phi(y) = \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^+(x)\phi^-(y) + \phi^-(x)\phi^-(y) \tag{5.4.8}$$

But as I stated above, we want to get this normal ordered, which can be done by changing around the third term (as the fourth term can't be normal ordered since it has no annihilation operators and the first two are already normal ordered), but to swap the field terms we have to introduce a commutator as well:

$$\phi(x)\phi(y) = \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^-(y)\phi^+(x) + [\phi^+(x), \phi^-(y)] + \phi^-(x)\phi^-(y) \tag{5.4.9}$$

Now there is a clear difference between a time ordered product and the normal ordered product:

$$T\phi(x)\phi(y) =: \phi(x)\phi(y) : + [\phi^+(x), \phi^-(y)] \tag{5.4.10}$$

The commutator is given by the familiar expression:

$$[\phi^+(x), \phi^-(y)] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} e^{-ip(x-y)} \tag{5.4.11}$$

Which is what we meet previously; and it is a propagator $D(x-y)$. Similarly when $y^0 > x^0$ we get:

$$T\phi(x)\phi(y) =: \phi(x)\phi(y) : + D(y-x) \quad (5.4.12)$$

$D(x-y)$ and $D(y-x)$ were previously defined as the Feynmann propagators. In general:

$$T\phi(x)\phi(y) =: \phi(x)\phi(y) : + \Delta_p(x-y) \quad (5.4.13)$$

$$\Delta_p(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{ie^{ik(x-y)}}{k^2 - m^2 + i\epsilon} \quad (5.4.14)$$

Now we define a *contraction*, of a pair of fields, in a string of operators:

$$\dots \overbrace{\phi_1(x)\dots\phi(x_2)} \dots \quad (5.4.15)$$

to mean replacing the operators with the Feynmann propagator:

$$\overbrace{\phi(x_1)\phi(x_2)} = \Delta_F(x_1 - x_2) \quad (5.4.16)$$

For a complex scalar fields we have:

$$\overbrace{\phi(x)\phi^\dagger(y)} = \Delta_F(x-y) \quad (5.4.17)$$

$$\overbrace{\phi(x)\phi(y)} = \overbrace{\psi^\dagger(x)\psi^\dagger(y)} = 0 \quad (5.4.18)$$

The contraction should be thought of as the difference between the normal ordered product and the time ordered product.

Wicks theorem statement

For any collection of fields we define:

$$\phi_1 = \phi(x_1)$$

$$\phi_2 = \phi(x_2)$$

Wick's theorem tells us what the difference between a time ordered product of fields and normal ordered product of fields:

$$T(\phi_1, \dots, \phi_n) =: \phi_1 \dots \phi_n : + : \text{all possible contractions} : \quad (5.4.19)$$

Example: Take a field evaluated at four different points

$$\begin{aligned} T(\phi_1\phi_2\phi_3\phi_4) =: & \phi_1\phi_2\phi_3\phi_4 : + \overbrace{\phi_1\phi_2} : \phi_3\phi_4 : + \overbrace{\phi_1\phi_3} : \phi_2\phi_4 : + \overbrace{\phi_1\phi_4} : \phi_2\phi_3 : + 3 \text{ similar terms} \\ & + \overbrace{\phi_1\phi_2\phi_3\phi_4} + \overbrace{\phi_1\phi_3\phi_2\phi_4} + \overbrace{\phi_1\phi_4\phi_2\phi_3} \end{aligned}$$

These are not operators because everything is contracted (which just gives the Feynmann propagator, which is just a number). The proof for Wicks theorem can be seen from the fact that we should it works for the case of $n = 2$ and then the rest of the proof follows from induction.

Scattering of Nucleons, $\psi + \psi \rightarrow \psi + \psi$

I have an initial state $|i\rangle$, which consists of two nucleons, which have momenta $\mathbf{p}_1, \mathbf{p}_2$, so we define:

$$|i\rangle = \sqrt{4E_{\mathbf{p}_1}E_{\mathbf{p}_2}} b_{\mathbf{p}_1}^\dagger b_{\mathbf{p}_2}^\dagger |0\rangle \equiv |p_1, p_2\rangle \quad (5.4.20)$$

And a final state, $|f\rangle$, which also has two nucleons with momenta p'_1 and p'_2 :

$$|f\rangle = \sqrt{4E_{\mathbf{p}'_1}E_{\mathbf{p}'_2}} b_{\mathbf{p}'_1}^\dagger b_{\mathbf{p}'_2}^\dagger |0\rangle \equiv |p'_1, p'_2\rangle \quad (5.4.21)$$

We want the probability for the initial state to turn into the final state. If the momentum does not change, i.e. $p_1 = p'_1$ and $p_2 = p'_2$ it simply means the particles haven't done anything!, since we are not interested in the situation in which nothing happens we are only interested in the $S - 1$ matrix (as supposed to the S-matrix). The S-1 matrix can be understood as follows:

The S matrix includes terms of the Taylor expansion of the unitary operator, the first term is just 1, which gives reaction in which nothing happens! (final state is the same as the initial state) so the S-1 matrix is the same expansion but without the 1 term.

Now, the first order term is simply H_0 and this will not turn the initial state into the final state, as a single H always has a single creation and annihilation operator for a meson, but neither of our states has mesons, so it will just kill it. So it is the second order term in the expansion that we use:

$$\frac{(-ig)^2}{2} \int d^4x_1 d^4x_2 T(\psi^\dagger(x_1)\psi(x_1)\phi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi(x_2)) \quad (5.4.22)$$

Now we use Wicks theorem, which states that the time ordered part, is equal to the normal ordered part plus all possible normal ordered contractions.¹⁶ We look for terms in Wicks theorem, in which the ϕ terms contract (as the ϕ terms remove mesons and there are no mesons in this state):

$$:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2): \overbrace{\phi(x_1)\phi(x_2)} \quad (5.4.23)$$

So we have:

$$\langle p'_1 p'_2 | : \psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2) : | p_1 p_2 \rangle \quad (5.4.24)$$

Multiplied by the contraction, which we can put in at the end. Now if we plug in the mode expansions of the fields and get them to cancel with the creation and annihilation operators from the initial and final states we get the following:

$$(e^{ip'_1 \cdot x_1 + ip'_2 \cdot x_2} + e^{ip'_1 \cdot x_2 + ip'_2 \cdot x_1})(e^{-ip_1 \cdot x_1 - ip_2 \cdot x_2} + e^{-ip_1 \cdot x_2 - ip_2 \cdot x_1}) \quad (5.4.25)$$

¹⁶Remember we are working in the interaction theory: $H_I = g\psi^\dagger\psi\phi$.

This can be multiplied out to give four terms, which are extremely long and trivial so I shall just call the E terms. So the overall amplitude is given by the integral over these E terms and the contraction, which is just the Feynmann propagator:

$$\langle f|S - 1|i \rangle = \frac{(-m)^2}{2} \int d^4x_1 d^4x_2 [E \text{ terms}] \times \int \frac{d^4k}{(2\pi)^4} \frac{ie^{ik(x_1-x_2)}}{k^2 - m^2 + i\epsilon} \quad (5.4.26)$$

Now the E terms are exactly the same under the integral as x_1 and x_2 are under the footing, so that cancels the factor of $\frac{1}{2}$ outside the integral. Then we integrate over x_1 and x_2 , which both give a delta function, which just leaves the propagator, which we can simply integrate over the delta functions:

$$i(-ig)^2 \left[\frac{1}{(p_1 - p'_1)^2 - m^2 + i\epsilon} + \frac{1}{(p_1 - p'_2)^2 - m^2 + i\epsilon} \right] (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p'_1 - p'_2)$$

I have missed out a few steps in the derivation here check green notebook! simply because they would not fit properly on the page!, however the calculation is not very difficult at all.

5.5 Feynmann Diagrams

Instead of writing all these long equations that come from working with Wicks theorem (that dont even fit on my latex page!), we can draw pretty pictures that represent the same mathematical meaning. The terms that arise in computing the non-trivial parts of the S-matrix, $\langle f|S - 1|i \rangle$ are in one to one correspondence with these diagrams. The kind of processes we are interested in have some number of particles in the initial state and some number of particles in the final state.

The idea is to draw an *external* line, for particle in the initial and the final states, if the particle has a charge associated to it, then draw an arrow on the line. A convention is that we choose incoming arrows for ψ particles in initial states and opposite for final state. The lines are joined together at vertices, if we have mesons; they are represented by a dotted line, if we have nucleons, then we have a solid line (another convention).

In general, the number of lines joining at a vertex is the number of interacting fields (i.e for three fields there will be three lines at each vertex).

Feynman Rules

Add a four momentum k to each internal line

For each vertex, write down a factor of:

$$(-g)(2\pi)^4 \delta^{(4)}\left(\sum_i k_i\right) \quad (5.5.1)$$

For each internal line, with momentum k , write down a factor of:

$$\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \quad (5.5.2)$$

For dotted internal lines we use mass m , for solid lines we use mass, M . It is trivially easy now to write down these terms for any reaction and compute the integral (without worry about all the prerequisite steps from Wicks theorem).

An interesting point is that the initial and final momentum p are such that $p^2 = m^2$, but the internal momenta k do not obey this rule. The particles that obey this rule are said to be on-shell and the ones that do not are said to be off-shell. Leading order diagrams (g^2 terms) are called tree level diagrams, the g^4 term diagrams are called loop diagrams. I must also stress that Feynmann rules are different theories!

Returning to the nucleon scattering events that we were describing before, the expression of the amplitude involving the $S - 1$ matrix has a $i\epsilon$ term, which can infact be ignored in this case. This is because we introduced it take into account the problem of singularities arising in the integral. However we are not integrating here and therefore there are never any singularities in the denominator.

Infact it turns out that this true for all tree level diagrams. It is only in the loop diagrams that we need to worry about this (as we have integrate a momentum over the loop). Now I shall introduce another convention. All the amplitudes that we compute will have a $(2\pi)^4\delta((p_1 + p_2) - (p'_1 + p'_2))$ and due to length of the calculations I will implicity assume that these terms are there, but I shall not write them out in full for every step, so the general term is:

$$\langle f|S - 1|i \rangle = (iA)(2\pi)^4\delta(p_i - p_f) \quad (5.5.3)$$

So from now on we will only be computing the A term in the equation above, which I will call the amplitude (the i is in the equation simply from convention).

Other processes

Nucleon to Meson: $\psi + \bar{\psi} \rightarrow \phi + \phi$

Now we shall compute the Amplitude for this process to occur, using Feynmann diagram and Feynmann rules for this theory. This is a tree level process so for the Feynmann rules give the following terms:

- The two vertices give an $(-ig)^2$ term each.
- Now there is a propagator for the nucleon rather than a meson, so we have:

$$\left[\frac{i}{(p_1 - p'_1)^2 - m^2} + \frac{i}{(p_1 - p'_2)^2 - m^2} \right] \quad (5.5.4)$$

Therefore the amplitude is simply the product of these terms (just shows how much easier it is to work with Feynmann diagrams, as supposed to working with Wicks theorem):

$$iA = (-ig)^2 \left[\frac{i}{(p_1 - p'_1)^2 - m^2} + \frac{i}{(p_1 - p'_2)^2 - m^2} \right] \quad (5.5.5)$$

Meson scattering: $\phi + \phi \rightarrow \phi\phi$

There is no tree level diagram for this scattering, the first term is a comes from a loop diagram. If we label the momenta of incoming particles with "p's" and the momenta of outgoing particles as "P's", then I have to pick an arbitrary momentum "k" and attach it to one of the lines of the loop, the rest of the momenta for the lines, follows from the conservation of momentum. Applying the same Feynmann rules we get the amplitude as:

$$iA = (-ig)^4 \int \frac{d^4k}{(2\pi)^4} \frac{i^4}{(k^2 - m^2 + i\epsilon)((k + p_1)^2 - m^2 + i\epsilon)} \frac{1}{((k + P_1 - p_1)^2 - m^2 + i\epsilon)((k - P_2)^2 - m^2 + i\epsilon)} \quad (5.5.6)$$

This integral does not look very nice, most of these integrals are divergen, however when we look closely at this integral, we see that the integral goes as:

$$\int \frac{d^4k}{k^8}$$

which will converge as $\frac{1}{k^4}$ as $k \rightarrow \infty$. Therefore this specific integral converges, however in general these integrals do not converge, especially for loop diagrams (as supposed to tree level diagram). What is happening here, is that a meson is moving and constantly splitting up into nucleon, antinucleon pairs and this constant process changes what we observe as the mass of the meson, and this process is dealt with in *Renormalisation*.

ϕ^4 Theory

The Lagrangian is given by:

$$L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4!} \phi^4 \quad (5.5.7)$$

Single and scalar field theory, with a *marginal* perturbation, therefore λ is a dimensionless number. This time each vertex of a Feynmann diagram has four legs (as there are four fields). Each vertex has a term, $-i\lambda$. Notice that there is no $\frac{1}{4!}$ term, remember $\frac{\lambda}{4!} \phi^4$, is an interaction term in the Hamiltonian and that means that when are use Dyson's formula, we are sandwiching products of this interaction Hamiltonian between, initial and final states (and time ordering them).

Example of ϕ^4 scattering: $\phi + \phi \rightarrow \phi + \phi$

We have an initial state: $|p_1, p_2 \rangle$

Final state: $\langle p'_1, p'_2 |$

So the amplitude is:

$$\frac{-i\lambda}{4!} \langle p'_1, p'_2 | : \phi(x)\phi(x)\phi(x)\phi(x) : |p_1, p_2 \rangle \quad (5.5.8)$$

Now each of the ϕ terms has a creation and annihilation operator sitting in it, so what we want to do is, take a creation operator from two ϕ terms and act on $\langle p'_1, p'_2 |$ and take an annihilation

operator from two ϕ 's and act on $|p_1, p_2\rangle$. This tells us what will happen to the $4!$ term, as the number of ways to pick out these operators from the ϕ 's is given by $4!$, therefore it just cancels the other term.

6 Dirac Equation

So far we have only considered scalar fields. Under a Lorentz transformation, the field rotates or boosts and the new field is just the old field evaluated at the old position in space:

$$\phi(x) = \phi'(x) = \phi(\Lambda^{-1}x) \quad (6.0.1)$$

The problem is that the fundamental particles that we see around us are not described by scalar fields, as they have a property of spin¹⁷. To describe particles with spin, we look at fields which have non-trivial transformations under the Lorentz group (A trivial transformation is one that occurs because the space itself changes, nothing to do with the field). An example of a non-trivial field is the gauge field of E-M:

$$A^\mu(x) \rightarrow \Lambda_\nu^\mu A^\nu(\Lambda^{-1}x) \quad (6.0.2)$$

6.1 Spin

We start the discussion by spin $\frac{1}{2}$ particles. In general a field can transform as:

$$\phi^a(x) \rightarrow D_b^a(\Lambda)\phi^b(\Lambda^{-1}x) \quad (6.1.1)$$

Where the ϕ 's represent a whole set of fields labeled by some index a . Under a transformation, these set of fields could transform into each other, the transformation is given by a matrix, $D_b^a(\Lambda)$. The matrix depends on what Lorentz transformations one is doing, and it should have various properties like:

$$D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2] \quad (6.1.2)$$

$$D(\Lambda^{-1}) = D[\Lambda]^{-1} \quad (6.1.3)$$

$$D[1] = 1 \quad (6.1.4)$$

Mathematically we would say that, it is a *representation* of the Lorentz group. A representation simply means the transformations have the same properties as the Lorentz group. A group basically means that we do one transform and then another transform and the result is still a Lorentz transformation. There is a structure to Lorentz transformations, it is this thing called a group structure. So we want matrices, D , which have the same properties of matrix Λ .

To find representations, we look at the *Lie algebra*, this means we look at *infinitesimal* Lorentz transformations:

$$\Lambda_\nu^\mu = \delta_\nu^\mu + \epsilon\omega_\nu^\mu + \Theta(\epsilon^2) \quad (6.1.5)$$

where ϵ is an infinitesimal. The properties of the Lorentz transformation, should also hold for the properties of the infinitesimal transformation. Remember the definition of the Lorentz transformation:

$$\Lambda_\sigma^\mu \Lambda_\sigma^\nu \eta^{\sigma\rho} = \eta^{\mu\nu} \quad (6.1.6)$$

¹⁷This has now changed with the discovery of the Higgs, which arises from a scalar field

Substitute the relations above into the definition of the Lorentz transformation:

$$(\delta_\sigma^\mu + \epsilon\omega_\nu^\mu)(\delta_\sigma^\nu + \epsilon\omega_\nu^\mu)\eta^{\sigma\rho} = \eta^{\mu\nu} + \Theta(\epsilon^2) \quad (6.1.7)$$

We only look at the terms linear in ϵ (the delta terms simply reorder the matrix and it cancels with the η term on the left and the ϵ^2 term is very small and can be neglected), and equate the coefficients on both sides of the equations, to obtain the simple result:

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0 \quad (6.1.8)$$

This basically states that the matrix has to be antisymmetric. These are 4×4 matrices and we know they must be antisymmetric, therefore the total number of possible configurations are $4 \times \frac{3}{2} = 6$, which agrees with the 6-components of the Lorentz group (3 rotations and 3 boosts).

So to summarise, infinitesimal Lorentz transformations correspond to anti-symmetric matrices. Now we are going to introduce a basis of infinitesimal Lorentz transformations, so any anti-symmetric matrix can be written as a linear combination of 6 anti-symmetric matrices:

$$(M^{\rho\sigma})_\nu^\mu = \eta^{\rho\mu}\delta_\nu^\rho - \eta^{\rho\nu}\delta_\nu^\rho \quad (6.1.9)$$

The index of (ν) in front of $(M^{\rho\sigma})$ labels the rows and columns of the 4×4 matrix.

The $\rho\sigma$ index in the power of M represents ρ and σ going from $0 \rightarrow 3$, they are anti-symmetric as seen from the delta functions (we could equally have used a single index that goes from $1 \rightarrow 6$). So this labels which of the 6 matrices I am interested in.

For example:

$$(M^{01})_\nu^\mu = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \eta^{1\mu}\delta_\nu^1 - \eta^{2\mu}\delta_\nu^0$$

$$(M^{12})_\nu^\mu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \eta^{1\mu}\delta_\nu^2 - \eta^{2\mu}\delta_\nu^1$$

$(M^{01})_\nu^\mu$ is an infinitesimal boost in the x direction.

$(M^{12})_\nu^\mu$ is an infinitesimal rotation around the z direction.

We can now write any infinitesimal Lorentz transformation as a linear combination of the 6 known transformations:

$$\omega_\nu^\mu = \frac{1}{2}\Omega_{\rho\sigma}(M^{\rho\sigma})_\nu^\mu \quad (6.1.12)$$

Any finite Lorentz transformation can be written as:

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) \quad (6.1.13)$$

$M^{\rho\sigma}$ represents a set of 6, 4×4 matrices (there is a μ index after this and after the Λ , however they have been left out for clarity), these are called generators of the Lorentz transformations.

$\frac{1}{2}\Omega_{\rho\sigma}$ is a set of 6 numbers.

The entire term is therefore a 4×4 matrix.

It is easy to see that if $\Omega_{\sigma\rho}$ is small, the exponential term can be expanded out into the infinitesimal terms that we saw before. For this we need to understand some aspects of Lie algebra as the generators of these transformations follow:

$$[M^{\rho\sigma}, M^{\Gamma\nu}] = \eta^{\sigma\Gamma} M^{\rho\nu} - \eta^{\rho\Gamma} M^{\sigma\nu} + \eta^{\sigma\nu} M^{\sigma\Gamma} - \eta^{\rho\nu} M^{\rho\Gamma} \quad (6.1.14)$$

Each of these terms are a 4×4 matrix (the row-column labels have been left out for clarity). This equation above involves the commutation relation contains all the important information about the Lorentz group.

6.2 Spinor representation

Now what we want to do, is to find other representations of the Lorentz group. That means we want to find matrices that have the same multiplication properties as the Lorentz group. This is called the spinor representation. The Clifford algebra is defined by:

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbf{1} \quad (6.2.1)$$

$$\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu, \quad \mu \neq \nu \quad (6.2.2)$$

$$(\gamma^0)^2 = \mathbf{1} \quad (6.2.3)$$

$$(\gamma^i)^2 = -\mathbf{1} \quad (6.2.4)$$

here $\mathbf{1}$ represents a unit identity matrix and $i = 1, 2, 3$. We want a set of matrices that obey these relations. The simplest representation is in terms of 4×4 matrices:

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$

where each element is itself a 2×2 matrix, with σ^i being the Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

These are not the only solutions to these equations, for example:

$$U\gamma^0U^{-1} \tag{6.2.10}$$

U is a matrix and U^{-1} is its inverse, this will also satisfy the clifford algebra. So there are many other representations of the form:

$$U\gamma^\mu U^{-1} \tag{6.2.11}$$

for any U . However there is a theorem that shows, that upto the equivalence of $U\gamma U^{-1}$, there is a unique irreducible representation of the clifford algebra, the particular solution we have is called the *chiral representation*.

We just pulled these γ matrices from thin air, now what we want to do, is to see what they have to do with the Lorentz group. Consider the commutator relations of the γ matrices:

$$S^{\rho\sigma} = \frac{1}{4}[\gamma^\rho, \gamma^\sigma] = \frac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma} \tag{6.2.12}$$

Now we consider the commutation relation between $S^{\mu\nu}$ and γ^ρ , when $\mu \neq \nu$:

$$[S^{\mu\nu}, \gamma^\rho] = \frac{1}{2}[\gamma^\mu\gamma^\nu, \gamma^\rho]$$

$$[S^{\mu\nu}, \gamma^\rho] = \frac{1}{2}\gamma^\mu\gamma^\nu\gamma^\rho - \frac{1}{2}\gamma^\rho\gamma^\mu\gamma^\nu$$

$$[S^{\mu\nu}, \gamma^\rho] = \frac{1}{2}\gamma^\mu\{\gamma^\nu\gamma^\rho\} - \frac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu - \frac{1}{2}\{\gamma^\rho\gamma^\mu\}\gamma^\nu + \frac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu$$

Now using the relation in Eq 6.2.1 we get:

$$[S^{\mu\nu}, \gamma^\rho] = \gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\rho\mu} \tag{6.2.13}$$

Now I shall prove an important point by taking the commutator of $S^{\mu\nu}$ and $S^{\rho\sigma}$ when $\rho \neq \sigma$:

$$[S^{\mu\nu}, S^{\rho\sigma}] = \frac{1}{2}[S^{\mu\nu}, \gamma^\rho\gamma^\sigma]$$

$$[S^{\mu\nu}, S^{\rho\sigma}] = \frac{1}{2}[S^{\mu\nu}, \gamma^\rho]\gamma^\sigma + \frac{1}{2}\gamma^\rho[S^{\mu\nu}, \gamma^\sigma]$$

$$[S^{\mu\nu}, S^{\rho\sigma}] = \frac{1}{2}\gamma^\mu\gamma^\sigma\eta^{\nu\rho} - \frac{1}{2}\gamma^\nu\gamma^\sigma\eta^{\rho\mu} + \frac{1}{2}\gamma^\rho\gamma^\mu\eta^{\nu\sigma} - \frac{1}{2}\gamma^\rho\gamma^\nu\eta^{\sigma\mu}$$

Now we use the relation in Eq 6.2.12:

$$[S^{\mu\nu}, S^{\rho\sigma}] = S^{\mu\sigma}\eta^{\nu\rho} - S^{\nu\sigma}\eta^{\rho\mu} + S^{\rho\mu}\eta^{\nu\sigma} - S^{\rho\nu}\eta^{\sigma\mu} \tag{6.2.14}$$

This equation matches the Eq 6.1.14, and therefore it shows that $S^{\mu\nu}$ matrices form a representation of the Lorentz algebra. Now we need a field for the $S^{\mu\nu}$ matrices to act upon. So we introduce the Dirac spinor field, which is a complex valued object $\psi^\alpha(x)$. The field spans over all of space, $\alpha = 1, 2, 3, 4$. The important point here is that the α indices have nothing to do with the 0, 1, 2, 3 i.e. they do not represent space-time components.

Under a Lorentz transformation we have:

$$\psi^\alpha(x) \rightarrow S[\Lambda]^\alpha_\beta \psi^\beta(\Lambda^{-1}x) \quad (6.2.15)$$

where:

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) \quad (6.2.16)$$

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \quad (6.2.17)$$

Where the $S[\Lambda]$ and Λ are both representations of the Lorentz group:

$$S[\Lambda_1]S[\Lambda_2] = S[\Lambda_1\Lambda_2] \quad (6.2.18)$$

As we have shown that the S matrices follow the same commutation relations as the Lorentz group. Λ and $S[\Lambda]$ are both 4×4 matrices, so to make sure that we have not wasted all this time finding a new representation for $S[\Lambda]$ that isn't just the same as Λ , we look at the explicit $S[\Lambda]$ in the chiral representation.

Rotations

$$S^{ij} = \frac{1}{4}[\gamma^i, \gamma^j] \quad (6.2.19)$$

where $i, j = 1, 2, 3$, now substituting for the gamma matrices (which have Pauli matrices in there):

$$S^{ij} = \frac{1}{2} \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix}$$

$$S^{ij} = \frac{-i}{2} \epsilon^{ijk} \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}$$

Define:

$$\Omega_{ij} = -\epsilon_{ij} = -\epsilon_{ijk}\phi^k \quad (6.2.22)$$

such that:

$$\Omega_{12} = -\phi^3$$

$$\Omega_{23} = -\phi^1$$

$$\Omega_{31} = -\phi^2$$

So that the rotation matrix is:

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \quad (6.2.23)$$

$$S[\Lambda] = \begin{pmatrix} e^{i\phi\cdot\frac{\sigma}{2}} & 0 \\ 0 & e^{i\phi\cdot\frac{\sigma}{2}} \end{pmatrix}$$

This tells us how to act on a spin matrix to rotate by an angle ϕ . Consider a rotation of 2π along the z-axis:

$$\phi = (0, 0, 2\pi) \quad (6.2.25)$$

So plugging this relation into the rotation matrix:

$$S[\Lambda] = \begin{pmatrix} e^{i\pi\sigma^3} & 0 \\ 0 & e^{i\pi\sigma^3} \end{pmatrix} = -1$$

Therefore under a 2π rotation:

$$\psi^\alpha(x) \rightarrow -\psi^\alpha(x) \quad (6.2.27)$$

This is not what happens for a vector, as a rotation by 2π leads to exactly the same vector. This is the most striking difference between the Lorentz representation and spinor representation. This is something we are familiar with in terms of Fermions, as they are antisymmetric particles. This is always said to be due to some quantisation effects, however this is not true as it holds for classical fields as well. The anti-symmetry comes up due to the representation of the Lorentz group and finding that there are new objects in the fields that have this anti-symmetric property, so it is in fact a property of the Lorentz group (not of QM).

Boosts

Consider the matrix:

$$S^{0i} = \frac{1}{4}[\gamma^0, \gamma^i]$$

$$S^{ij} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$

$$S^{ij} = \frac{1}{2} \begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}$$

We have $\Omega_{i0} = -\Omega_{0i} = \chi_i$:

$$S[\Lambda] = \begin{pmatrix} e^{\chi\cdot\frac{\sigma}{2}} & 0 \\ 0 & e^{-\chi\cdot\frac{\sigma}{2}} \end{pmatrix}$$

These are the familiar S matrices, with exponentials with powers of the sigma matrices. There is a factor of i for rotations but not for boosts. So for rotations we have:

$$S[\Lambda]^\dagger S[\Lambda] = 1 \quad (6.2.31)$$

This is called being *unitary*. However boosts do not follow this as there is no minus sign in the exponent of the conjugate to cancel with (as it is real):

$$S[\Lambda]^\dagger S[\Lambda] \neq 1 \quad (6.2.32)$$

This is infact a general statement; there are no finite dimensional unitary representations of the Lorentz group.

6.3 Action

Now that we have this new field of the Dirac spinor, which transforms in an interesting way under a Lorentz transformation, we would like to know what the equations of motion this could obey that are also invariant (covariant) under Lorentz transformations.

The trick to find this, is to write down an action (or Lagrangian) which is invariant under a Lorentz transformation. We have this spinor index floating around, so we will have to contract spinor indicies together to get something that is invariant, the obvious thing to contract is ψ with ψ^\dagger and then contract the indicies.

Define:

$$\psi^\dagger(x) = (\psi^*)^T(x) \quad (6.3.1)$$

Under a Lorentz transformation:

$$\psi(x) \rightarrow S[\Lambda]\psi(\Lambda^{-1}x) \quad (6.3.2)$$

$$\psi^\dagger(x) \rightarrow \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger \quad (6.3.3)$$

So this is just the conjugate of the one above, but this representation is not unitary:

$$S[\Lambda]^\dagger S[\Lambda] \neq 1 \quad (6.3.4)$$

Therefore contracting ψ, ψ^\dagger does not give a Lorentz scalar. Recall that $S[\Lambda]$ was the exponential of the generators:

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \quad (6.3.5)$$

To get $S[\Lambda]$ to be unitary, we require:

$$(S^{\mu\nu})^\dagger = -S^{\mu\nu} \quad (6.3.6)$$

But we know that:

$$(S^{\mu\nu})^\dagger = \frac{1}{4}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger] \quad (6.3.7)$$

So we need all the γ 's to be Hermitian, or all the γ 'a to be anti-Hermitian:

$$(\gamma^\mu)^\dagger = -\gamma^\mu \quad (6.3.8)$$

$$(\gamma^0)^2 = 1 \quad (6.3.9)$$

This has real eigenvalues, therefore can be made Hermitian. But any other γ is equal to -1, when squared. Which means they are imaginary eigenvalues and hence can always be made anti-Hermitian, but never Hermitian:

$$(\gamma^i)^2 = -1 \quad (6.3.10)$$

But we need all of them to be the same if this was a unitary representation. The reason it was not unitary is the minus sign in the Minkowski metric. However this also gives a hint as to how to get out of this problem. It γ^0 is Hermitian:

$$(\gamma^0)^\dagger = \gamma^0 \quad (6.3.11)$$

and γ^i is anti-Hermitian:

$$(\gamma^i)^\dagger = -\gamma^i \quad (6.3.12)$$

Then we see that:

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger \quad (6.3.13)$$

It means that if we take any γ matrix and multiply it on the left and right by γ^0 , we get the Hermitian conjugate of that γ matrix. If we use this along with the definitions of the $S^{\mu\nu}$ matrices:

$$(S^{\mu\nu})^\dagger = \frac{1}{4}[(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] = \gamma^0 S^{\mu\nu} \gamma^0 \quad (6.3.14)$$

Therefore we have:

$$S[\Lambda]^\dagger = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\dagger\right) = \gamma^0 S[\Lambda]^{-1} \gamma^0 \quad (6.3.15)$$

Now we introduce the *Dirac conjugate*:

$$\bar{\psi}(x) = \psi^\dagger(x) \gamma^0 \quad (6.3.16)$$

Now consider the following calculation:

$$\bar{\psi}(x) \psi(x) = \psi^\dagger(x) \gamma^0 \psi(x)$$

Introducing the transformations:

$$\bar{\psi}(x) \psi(x) = \psi^\dagger(\Lambda^{-1}x) S[\Lambda]^\dagger \gamma^0 S[\Lambda] \psi(\Lambda^{-1}x)$$

$$\bar{\psi}(x) \psi(x) = \psi^\dagger(\Lambda^{-1}x) \gamma^0 \psi(\Lambda^{-1}x)$$

$$\bar{\psi}(x) \psi(x) = \bar{\psi}(\Lambda^{-1}x) \psi(\Lambda^{-1}x) \quad (6.3.17)$$

This shows that $\bar{\psi}(x) \psi(x)$ is a Lorentz scalar. Now consider:

$$\bar{\psi} \gamma^\mu \psi$$

Once again we introduce the Lorentz transformations:

$$\bar{\psi} \gamma^\mu \psi \rightarrow \bar{\psi} S[\Lambda]^{-1} \gamma^\mu S[\Lambda] \psi \quad (6.3.18)$$

If $\bar{\psi} \gamma^\mu \psi$ is to transform as a vector, we have:

$$S[\Lambda]^{-1} \gamma^\mu S[\Lambda] = \Lambda_\nu^\mu \gamma^\nu \quad (6.3.19)$$

Since we are working infinitesimally:

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma}\right) \approx 1 + \frac{1}{2}\Omega_{\rho\sigma}M^{\rho\sigma} + \dots \quad (6.3.20)$$

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) \approx 1 + \frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma} + \dots \quad (6.3.21)$$

Now we substitute these terms into Eq 6.3.19:

$$-[S^{\rho\sigma}, \gamma^\mu] = (M^{\rho\sigma})^\mu_\nu \gamma^\nu \quad (6.3.22)$$

where we have suppressed the α and β indicies on the γ^μ and $S^{\mu\nu}$, but otherwise left all other indicies explicit. Infact, this equation follows from:

$$[S^{\rho\sigma}, \gamma^\mu] = \gamma^\rho \eta^{\sigma\mu} - \gamma^\sigma \eta^{\mu\rho} \quad (6.3.23)$$

We can see this explicitly by expanding ($M^{\rho\sigma}$):

$$(M^{\rho\sigma})^\mu_\nu \gamma^\nu = (\eta^{\rho\mu} \delta_\nu^\sigma - \eta^{\sigma\mu} \delta_\mu^\rho) \gamma^\nu \quad (6.3.24)$$

By acting on the γ matrices with the delta functions we get:

$$(M^{\rho\sigma})^\mu_\nu \gamma^\nu = \eta^{\sigma\mu} \gamma^\sigma - \eta^{\sigma\mu} \gamma^\rho \quad (6.3.25)$$

Therefore:

$$-[S^{\rho\sigma}, \gamma^\mu] = \eta^{\rho\mu} \gamma^\sigma - \eta^{\sigma\mu} \gamma^\rho \quad (6.3.26)$$

6.4 The Dirac Equation

We want to write down a Lagrangian that is a Lorentz invariant. We have already constructed a Lorentz scalar and a Lorentz vector in the previous section. The following action is invariant under Lorentz transformation:

$$S = \int d^4x \bar{\psi}(x) (i\gamma^\mu \partial_\mu - m) \psi(x) \quad (6.4.1)$$

We have already proven that $\bar{\psi}(x)\psi(x)$ is a Lorentz scalar, and we are integrating it over d^4x so the action is invariant. The $\bar{\psi}(x)\gamma^\mu\psi(x)$, is a Lorentz vector, so to get something that is Lorentz invariant, we need to contract it with another vector. This happens to be in the Dirac equation in the form of the differential operator, ∂_μ . The equation of motion is:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (6.4.2)$$

$i\gamma^\mu \partial_\mu$ is 4×4 matrix.

m is a unit 4×4 matrix.

ψ is now a four component, complex column vector.

This is the **Dirac equation**. The amazing thing about this equation (among many other things) is the fact that it is linear in derivatives and yet is Lorentz invariant. The key to this is the properties of the γ matrices.

Suppose we tried to do this for a scalar field (no indices) and we wanted to write down an equation linear derivatives, you would have a μ index in the derivative and therefore you would also require the $i\gamma$ term to have a μ index, but there is nothing natural that has a μ index, you would need something that points somewhere. It would therefore always have a preferred direction in spacetime and would not be Lorentz invariant. So what this equation shows is that there are natural vectors floating around in the universe and these are the γ matrices, that rotate naturally like a vector under Lorentz transformations.

We derived the action for a spinor field and we did this by looking for representation of the Lorentz group, we found a new representation, the spinor representation. We introduced a field, a four vector complex object called a Dirac spinor, $\psi(x)$, and this was defined in a particular way as to how it would transform under the Lorentz transformations to give an action that was Lorentz invariant. It also depends on m as supposed to m^2 , therefore the mass can be positive or negative.

6.5 Notation

Any vector, A_μ , contracted with the γ matrices we end up with a 4×4 matrix:

$$A_\mu \gamma^\mu \neq \not{A} \tag{6.5.1}$$

So the Dirac equation in this notation is:

$$(i\not{\partial} - m)\psi = 0 \tag{6.5.2}$$

Chiral spinors

For any choice of the γ matrices:

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

In this representation, the spinor rotation transformation is:

$$S[\Lambda_{rot}] = \begin{pmatrix} e^{i\phi \cdot \frac{\sigma}{2}} & 0 \\ 0 & e^{i\phi \cdot \frac{\sigma}{2}} \end{pmatrix}$$

and the boost transformation is:

$$S[\Lambda] = \begin{pmatrix} e^{\chi \cdot \frac{\sigma}{2}} & 0 \\ 0 & e^{-\chi \cdot \frac{\sigma}{2}} \end{pmatrix}$$

These have come out to be block diagonal. This means that we have not found the irreducible representation of the Lorentz group. It decomposes into two irreducible representations:

$$\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix}$$

where the u 's are each a two component spinor. These are called *Weyl* spinors (chiral spinors). Under rotations:

$$u_{\pm} \rightarrow e^{i\phi \cdot \frac{\sigma}{2}} u_{\pm} \quad (6.5.8)$$

ϕ is just a vector here, it states how much the angle should be rotated by in a given rotation. Under boosts:

$$u_{\pm} \rightarrow e^{\chi \cdot \frac{\sigma}{2}} u_{\pm} \quad (6.5.9)$$

The Lagrangian can be written in terms of the Weyl spinors:

$$L = \bar{\psi}(i\not{\partial} - m)\psi$$

$$L = iu_-^\dagger \sigma^\mu \partial_\mu u_- + iu_+^\dagger \bar{\sigma}^\mu \partial_\mu u_+ - m(u_+^\dagger u_-^\dagger u_- u_+) = 0 \quad (6.5.10)$$

where:

$$\sigma^\mu = (1, \sigma^i)$$

$$\mu = 0, 1, 2, 3$$

and σ^0 is a unit matrix and $\sigma^{1,2,3}$ are the Pauli matrices.

The Hermitian conjugate of σ^μ is:

$$\bar{\sigma}^\mu = (1, -\sigma^i)$$

The Weyl spinors are 2×2 objects, and when we write out the Lagrangian in terms of Weyl spinors, the kinetic terms are fine, since they just have u_+ and u_- separately, however the mass term, couples together the u_+ and the u_- . So for massive fermions we need both u_+ and u_- , but for massless fermions, we can describe them in terms of single Weyl spinors or double Weyl spinors:

$$i\bar{\sigma}^\mu \partial_\mu u_+ = 0 \quad (6.5.11)$$

And the same expression is true for u_- . These are called the Weyl equations. So a massive fermion is described by Dirac spinors and massless fermions are described by Weyl spinors. Every fermion in the standard model is described by a Weyl spinor! This is because every fermion has a Lagrangian that has no mass term. The mass is introduced by another mechanism, known as the Higgs mechanism.

6.6 γ^5

In the previous section, we wrote down that $\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix}$ if we trace back why this was sensible; it was because of our particular choice of γ matrices. The chiral representation, which then gave rise to that particular choice of Lorentz transformations, but was block diagonal and hence reducible.

If for the same reason U want to work with a different set of γ matrices, my Lorentz transformations will not be block diagonal and everything will be mixed up. We introduce a "fifth" γ matrix, that is simply the product of all the γ matrices:

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (6.6.1)$$

This matrix satisfies:

$$\{\gamma^5, \gamma^\mu\} = 0 \quad (6.6.2)$$

$$(\gamma^5)^2 = 1 \quad (6.6.3)$$

Now γ^5 and γ^μ obey the Clifford algebra, that is in 5 dimensions, rather than 4 dimensions. This is something that haolds in all even dimensions. From Eq 6.6.3, we can define a projection operator:

$$P_\pm = \frac{1}{2}(1 \pm \gamma^5) \quad (6.6.4)$$

which also follows the following relations:

$$P_+^2 = P_+$$

$$P_-^2 = P_-$$

$$P_+P_- = 0$$

So these are two Projection operators, each of which spans over half the projection space. So we define the chiral spinors as:

$$\psi_\pm = P_\pm\psi \quad (6.6.5)$$

This thing is analogous to the Weyl spinors, but in an arbitrary basis. We find that in the chiral representation:

$$\psi_+ = \begin{pmatrix} u_+ \\ 0 \end{pmatrix}$$

$$\psi_- = \begin{pmatrix} u_- \\ 0 \end{pmatrix}$$

We had previously worked very hard to come up with objects that are scalars under the Lorentz group. Now we can show that:

$$\bar{\psi}\gamma^5\psi$$

is a Lorentz (psuedo) scalar.

$$\bar{\psi}\gamma^5\gamma^\mu\psi$$

is a Lorentz (psuedo) vector. The psuedo is included to describe how these things transform under parity. A psuedo vector or scalar means they pick up a minus sign after the transformation.

6.7 Symmetries of the Dirac equation

Recall the action:

$$S = \int d^4x \bar{\psi}(i\cancel{\partial} - m)\psi \quad (6.7.1)$$

There are a bunch of symmetries, that this Lagrangian has, the first two we have already seen for the scalar field, that is the invariance under space-time translations and also invariant under the 6 Lorentz transformations. From Noether's theorem we know that each of these symmetries will lead to a conserved current and a conserved charge. From translations we get the stress-energy tensor using Eq 2.3.17:

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi \quad (6.7.2)$$

The Lorentz transformations provide a current:

$$(J^\mu)^{\rho\sigma} = x^\sigma T^{\mu\rho} - x^\rho T^{\mu\sigma} - i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi \quad (6.7.3)$$

the $\rho\sigma$ index represents which Lorentz transformation one is doing (they are anti-symmetric). When we quantise the scalar field, we want to understand the spin of the scalar field, which is nothing but how things change under rotation, in that case you compute the conserved quantities due to rotations, you then turn that into an operator and with that operator on the state that corresponds to a state of a stationary particle and it goes zero.

We shall do exactly the same thing here and we will find that we get a $\frac{1}{2}$, which obviously comes from the last term in the current as it is the only difference between the scalar field and the vector field.

Internal Symmetry

$$\psi \rightarrow e^{i\alpha}\psi \quad (6.7.4)$$

The fact that ψ is complex, means we can change the phase of ψ . This has a current associated to it, sometimes called the *vector current*, j_V^μ .

$$j_V^\mu = \bar{\psi}\gamma^\mu\psi \quad (6.7.5)$$

V stands for vector. If $m = 0$, then we have a new symmetry:

$$\psi \rightarrow e^{i\alpha\gamma^5}\psi \quad (6.7.6)$$

And a similar equation stands for $\bar{\psi}$. This is because for zero mass the operators u_+ and u_- are not coupled together. So the γ^5 is making u_+ transform as $e^{i\alpha}$ and u_- transform as $e^{-i\alpha}$, if we are in the chiral representation. So the conserved current is:

$$j_A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi \quad (6.7.7)$$

A stands for axial. This is infact a symmetry of the classical theory, and it does not survive to the quantum theory. So its an example of what is called an anomoly; you think when you look at your equations you have a symmetry, but actually its not there. This happens because of the infinities that arise in QFT, and in any way that we try to get around these infinities, ends up breaking this

symmetry, as we have to introduce a mass, somewhere along the line.

Solutions to the Dirac equation

We want to solve:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (6.7.8)$$

First we think about positive frequency solutions. We take the solution to be of the form:

$$\psi = u_{\mathbf{p}} e^{-ip \cdot x} \quad (6.7.9)$$

ψ is some 4 vector (spinor) and $u_{\mathbf{p}}$ is a particular fixed 4 spinor, which depends on which \mathbf{p} we are looking at. Now we claim that this solution works when:

$$u_{\mathbf{p}} = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi \\ \sqrt{p \cdot \bar{\sigma}} \xi \end{pmatrix}$$

where \mathbf{p} is the particular four-momentum that is part of the solution and it is dotted with the four 2×2 matrices σ . So p_0 is dotted with the unit matrix and p_i 's are dotted with the Pauli matrices (same is true for $\bar{\sigma}$, except they have a minus sign). ξ is a 2-component spinor such that:

$$\xi^\dagger \xi = 1 \quad (6.7.11)$$

To prove this we write:

$$(u_{\mathbf{p}})^T = (u_1, u_2) \quad (6.7.12)$$

Then the Dirac equation gives:

$$(p \cdot \sigma)u_2 = mu_1 \quad (6.7.13)$$

$$(p \cdot \bar{\sigma})u_1 = mu_2 \quad (6.7.14)$$

Either one of these equations implies the other, a fact which follows from the identity:

$$(p \cdot \sigma)(p \cdot \bar{\sigma}) = p_0^2 - p_i p_j \sigma^i \sigma^j = p_0^2 - p_i p_j \delta^{ij} = p_\mu p^\mu = m^2 \quad (6.7.15)$$

Lets start with the anstaz:

$$u_1 = (p \cdot \sigma)\xi' \quad (6.7.16)$$

for some spinor ξ' . Then Eq 6.7.14 immediately tells us that:

$$u_2 = m\xi'$$

So we learn that any spinor of the form:

$$u_{\mathbf{p}} = A \begin{pmatrix} p \cdot \sigma \xi' \\ m \xi' \end{pmatrix}$$

with constant A is a solution to the Dirac equation. Now we can choose:

$$A = \frac{1}{m}$$

$$\xi' = \sqrt{p \cdot \bar{\sigma}} \xi$$

with constant ξ . Then

$$u_1 = (p \cdot \sigma) \sqrt{p \cdot \bar{\sigma}} \xi = m \sqrt{\rho \cdot \bar{\sigma}} \xi \quad (6.7.18)$$

Therefore we have completed the proof of the solution to the Dirac equation.

Now we introduce a basis ξ^s for $s = 1, 2$ such that:

$$\xi^{r\dagger} \xi^s = \delta^{rs}$$

$$\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Now we can show the negative solutions in a similar way:

$$\psi = v_{\mathbf{p}} e^{ip \cdot x} \quad (6.7.21)$$

$$v_{\mathbf{p}} = \begin{pmatrix} \sqrt{p \cdot \bar{\sigma}} \eta \\ \sqrt{p \cdot \sigma} \eta \end{pmatrix}$$

$$\eta^\dagger \eta = 1$$

7 Quantisation of Dirac Field

To quantise the Dirac field we shall follow the same method as we did for a scalar field. Recall the action:

$$S = \int d^4x \bar{\psi}(x)(i\not{\partial} - m)\psi(x) \quad (7.0.1)$$

The first thing to do, is to think of this as a classical theory and figure out what the conjugate momentum of ψ . The momentum π is defined as usual:

$$\pi = \frac{\partial L}{\partial \dot{\psi}} = i\bar{\psi}\gamma^0 = i\psi^\dagger \quad (7.0.2)$$

For a classical field, if we want to know how it will evolve for all times in the future, we have to know what the initial data of the field is and its initial velocities. Here this is not true, you need to specify the field and its momentum (notice that π is not proportional to $\dot{\psi}$ here), but we see that the momentum is just the conjugate of the field, this is expected for an equation of motion that is first order in time (rather than second order).

For the scalar field, we have a field, which is a function of space and time and we have a momentum that is a function of space and time. What we then did for the scalar field was introduce canonical commutation relations between position and momentum. First we shall work in the Schroedinger picture, as supposed to the Heisenberg picture, so now in quantum theory, ψ and π become functions of just space (not of space-time). To quantise the theory we promote ψ and ψ^\dagger to operators; which have the following canonical commutation relations:

$$[\psi_\alpha(\mathbf{x}), \psi_\beta(\mathbf{y})] = [\psi_\alpha^\dagger(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})] = 0 \quad (7.0.3)$$

all of the \mathbf{x} 's and \mathbf{y} 's are three vectors, because we are in the Schroedinger picture.

$$[\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})] = \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (7.0.4)$$

We need a mode expansion for these fields ψ , just as we did before in terms of creation and annihilation operators:

$$\psi(\mathbf{x}) = \sum_{s=1}^2 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} [b_{\mathbf{p}}^s u_{\mathbf{p}}^s e^{i\mathbf{p}\cdot\mathbf{x}} + c_{\mathbf{p}}^{s\dagger} v_{\mathbf{p}}^s e^{-i\mathbf{p}\cdot\mathbf{x}}] \quad (7.0.5)$$

$$\psi^\dagger(\mathbf{x}) = \sum_{s=1}^2 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} [b_{\mathbf{p}}^{s\dagger} u_{\mathbf{p}}^{s\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}} + c_{\mathbf{p}}^s v_{\mathbf{p}}^{s\dagger} e^{i\mathbf{p}\cdot\mathbf{x}}] \quad (7.0.6)$$

ψ and ψ^\dagger are four operators, as we have suppressed the α and β indicies on them. The $b_{\mathbf{p}}^{s\dagger}$ and $c_{\mathbf{p}}^{s\dagger}$ are operators that create particles associated with spinors $u_{\mathbf{p}}^s$ and $v_{\mathbf{p}}^s$ respectively (similar for $b_{\mathbf{p}}^s$ and $c_{\mathbf{p}}^s$, are the annihilation operators for their respective spinors. Now we claim the following commutation relations:

$$[b_{\mathbf{p}}^r, b_{\mathbf{p}}^{s\dagger}] = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (7.0.7)$$

$$[c_{\mathbf{p}}^r, c_{\mathbf{p}}^{s\dagger}] = -(2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (7.0.8)$$

with all other mixed commutators vanishing. To prove this, first consider:

$$\begin{aligned}
[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] &= \sum_{r,s} \int \frac{d^3 p d^3 q}{(2\pi)^6} \frac{1}{\sqrt{4E_{\mathbf{p}} E_{\mathbf{q}}}} ([b_{\mathbf{p}}^r, b_{\mathbf{q}}^s] u_{\mathbf{p}}^r u_{\mathbf{p}}^{s\dagger} e^{i(\mathbf{x}\cdot\mathbf{p}-\mathbf{y}\cdot\mathbf{q})} \\
&\quad + [c_{\mathbf{p}}^{r\dagger}, c_{\mathbf{q}}^s] v_{\mathbf{p}}^r v_{\mathbf{q}}^{s\dagger} e^{-i(\mathbf{x}\cdot\mathbf{p}-\mathbf{y}\cdot\mathbf{q})}) \\
&= \sum_s \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} (u_{\mathbf{p}}^s \bar{u}_{\mathbf{p}}^s \gamma^0 e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} + v_{\mathbf{p}}^s \bar{v}_{\mathbf{p}}^s \gamma^0 e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})})
\end{aligned}$$

Now we can use the outer product formulae:

$$\sum_s u_{\mathbf{p}}^s \bar{u}_{\mathbf{p}}^s = \not{p} + m \quad (7.0.9)$$

$$\sum_s v_{\mathbf{p}}^s \bar{v}_{\mathbf{p}}^s = \not{p} - m \quad (7.0.10)$$

Therefore we have:

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} ((\not{p} + m) \gamma^0 e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} + (\not{p} - m) \gamma^0 e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}) \quad (7.0.11)$$

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} ((p_0 \gamma^0 + p_i \gamma^i + m) \gamma^0 + (p_0 \gamma^0 - p_i \gamma^i - m) \gamma^0) e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \quad (7.0.12)$$

We know that $p_0 = E_{\mathbf{p}}$, so we have:

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} = \delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (7.0.13)$$

which proves Eq 7.0.4

7.1 Hamiltonian

Now we shall construct the Hamiltonian for the Dirac theory. Using the momentum $\pi = i\psi^\dagger$, we have:

$$H = \pi \dot{\psi} - L = \bar{\psi} (-i\gamma^i \partial_i + m) \psi \quad (7.1.1)$$

Notice that the time derivative, disappears completely from the Hamiltonian. The actual Hamiltonian is the Hamiltonian density, H , integrated over all space and this should agree with the conserved energy computed using Noether's theorem. Now we shall turn the Hamiltonian into an operator¹⁸. After normal ordering we have:

$$H = \int \frac{d^3 p}{(2\pi)^3} E_{\mathbf{p}} \sum_{s=1}^2 (b_{\mathbf{p}}^{s\dagger} b_{\mathbf{p}}^s - c_{\mathbf{p}}^s c_{\mathbf{p}}^{s\dagger}) \quad (7.1.2)$$

¹⁸I have not provided the proof, but it is not very difficult and is in Peskin and Schroeder

$$E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2} \quad (7.1.3)$$

The minus sign in front of the $c_{\mathbf{p}}^s c_{\mathbf{p}}^{s\dagger}$ means we have a problem because it has no stable ground state. as we create new particles with $c_{\mathbf{p}}^{s\dagger}$, however this is subtracted, so every time it is used we get a state with a Lower energy! and we can keep going for as long as we want. Infact there is no way yo make sense of spinor fields quantised in this way, we must treat them as **fermions**. This is the first glimpse of the spin statistics theorem in QFT.

7.2 Fermionic Quantisation

To get a well defined theory, we need to define anti-commutation relations. The notation for anti-commutators is:

$$\{A, B\} = AB + BA$$

We require the spinor fields to satisfy:

$$\{\psi_{\alpha}(\mathbf{x}), \psi_{\beta}(\mathbf{y})\} = \{\psi_{\alpha}^{\dagger}(\mathbf{x}), \psi_{\beta}^{\dagger}(\mathbf{y})\} = 0 \quad (7.2.1)$$

$$\{\psi_{\alpha}(\mathbf{x}), \psi_{\beta}^{\dagger}(\mathbf{y})\} = \delta_{\alpha\beta} \delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (7.2.2)$$

Now we translate the field commutation relations into relations between the creation and annihilation operators:

$$\{b_{\mathbf{p}}^r, b_{\mathbf{p}}^{s\dagger}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (7.2.3)$$

$$\{c_{\mathbf{p}}^r, c_{\mathbf{p}}^{s\dagger}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (7.2.4)$$

all the other anti-commutation relations are zero. Now we can compute the Hamiltonian, before normal ordering:

$$H = \int \frac{d^3p}{(2\pi)^3} E_{\mathbf{p}} [b_{\mathbf{p}}^{s\dagger} b_{\mathbf{p}}^s + c_{\mathbf{p}}^{s\dagger} c_{\mathbf{p}}^s - (2\pi)^3 \delta^{(3)}(0)] \quad (7.2.5)$$

Note that here the infinity that comes from the delta function has a minus sign, the infinity we got from bosons came with a positive sign. So in a theory with equal number of bosons and fermions these infinities would just cancel¹⁹.

Fock space

The vacuum states are defined in the Fock space as:

$$b_{\mathbf{p}}^s |0\rangle = c_{\mathbf{p}}^s |0\rangle = 0 \quad (7.2.6)$$

for all p and $s = 1, 2$. The Hamiltonian has the following commutation relations with b and c operators:

$$[H, b_{\mathbf{p}}^r] = -E_{\mathbf{p}} b_{\mathbf{p}}^r$$

$$[H, b_{\mathbf{p}}^{r\dagger}] = E_{\mathbf{p}} b_{\mathbf{p}}^{r\dagger}$$

¹⁹This is what supersymmetry suggests

$$[H, c_{\mathbf{p}}^r] = -E_{\mathbf{p}} c_{\mathbf{p}}^r$$

$$[H, c_{\mathbf{p}}^{r\dagger}] = E_{\mathbf{p}} c_{\mathbf{p}}^{r\dagger}$$

We have one particles states:

$$b_{\mathbf{p}}^s |0\rangle = \text{particle} \quad (7.2.7)$$

$$c_{\mathbf{p}}^s |0\rangle = \text{anti-particle} \quad (7.2.8)$$

This is the same as we had the bosonic case, the new thing is the S index and it will determine the spin of the particle. The S index is linked to which spinor we are thinking about; and that is a little bit arbitrary. Notice that:

$$b_{\mathbf{p}_1}^{s_1\dagger} b_{\mathbf{p}_2}^{s_2\dagger} |0\rangle \equiv |\mathbf{p}_1, r_1 : \mathbf{p}_2 r_2\rangle = -|\mathbf{p}_2, r_2 : \mathbf{p}_1 r_1\rangle \quad (7.2.9)$$

The minus sign came about as the operators anti-commute and this shows that particles obey Fermi-Dirac statistics.

Dirac's hole interpretation

The Dirac equation we have seen is:

$$(i\not{\partial} - m)\psi = 0$$

when Dirac came up with his equation, he wrote it in the following form:

$$i\frac{\partial\psi}{\partial t} = -i\boldsymbol{\alpha} \cdot \nabla\psi + m\beta\psi = \hat{H}\psi \quad (7.2.10)$$

where:

$$\boldsymbol{\alpha} = -\gamma^0\boldsymbol{\gamma}$$

$$\beta = \gamma^0$$

Originally Dirac interpreted ψ as a wavefunction for a single particle moving in space (and not a field), so it doesn't anti-commute or anything, basically it is not an operator. This is then identified as the Schroedinger equation. The Hamiltonian in the equation above, \hat{H} is not the Hamiltonian that we have been considering. In this interpretation of ψ as a one particle wavefunction, the solutions are thought of as Energy eigenstates:

$$u(\mathbf{p}) = e^{-ip \cdot x} \rightarrow i\frac{\partial\psi}{\partial t} = E_{\mathbf{p}}\psi \quad (7.2.11)$$

$$v(\mathbf{p}) = e^{ip \cdot x} \rightarrow i\frac{\partial\psi}{\partial t} = -E_{\mathbf{p}}\psi \quad (7.2.12)$$

These two solutions we called positive and negative frequency, respectively. So we see that this wavefunction representation leads to positive and negative energies. So it leads to the same problem we had when we quantised the field as a boson, as it leads to an unbounded theory. To get around this, Dirac argued that the particles are fermions (which they are!) and for fermion the

Pauli Exclusion Principle holds, so I can't put a fermion twice in the same state. He suggested that all the states with negative energies had fermions in them already and all the excitations had to exist in the positive energy region.

But then he suggested that suppose one the particles that was in the $E < 0$ state, gets excited and goes upto the positive enegy state, so that now there is a particle missing, from the $E < 0$ state, and this particle would have the property of an actual electron with a positive charge and it would annihilate with an actual electron. However this is not the current interpretation, as bosons have anti-particles aswell so it cant be explained by this theory. Basicly what this is telling us, is what was discussed in the very beginning, in the introduction, that when special relativity and QM are combined we cannoy have single particle states on their own.

When we take ψ to be a field as supposed to a wavefunction, the theory has a better foundation and this process is called second quantisation. Now we follow the same steps that we took in scalar fields. So we go from the Schroedinger picture, to the Heisenberg picture, by defining the spinors $\psi(\mathbf{x}, t)$ at every point in space-time (not just space), such that they satisfy:

$$\frac{\partial \psi}{\partial t} = i[\mathbf{H}, \psi] \quad (7.2.13)$$

The solutions to this equation are:

$$\psi(x) = \sum_{s=1}^2 \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} [b_{\mathbf{p}}^s u^s(\mathbf{p}) e^{-ip \cdot x} + c_{\mathbf{p}}^s v^s(\mathbf{p}) e^{ip \cdot x}] \quad (7.2.14)$$

$$\psi^\dagger(x) = \sum_{s=1}^2 \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} [b_{\mathbf{p}}^{s\dagger} u^{s\dagger}(\mathbf{p}) e^{ip \cdot x} + c_{\mathbf{p}}^s v^{s\dagger}(\mathbf{p}) e^{-ip \cdot x}] \quad (7.2.15)$$

One can also show that:

$$\{\psi_\alpha(x), \bar{\psi}_\beta(y)\} = 0, \quad \text{when } (x - y)^2 < 0 \quad (7.2.16)$$

Here we have two operators that do not commute at space-like seperations. This would suggest that the theory is non-causal. The best that we can say about this is that fermionic operators are not observable therefore this non-causality is not a problem.

7.3 Feynmann Propagator

It is defined as:

$$(S_F)_{\alpha\beta}(x - y) \equiv \langle 0 | T \psi(x) \bar{\psi}(y) | 0 \rangle \quad (7.3.1)$$

Which is also the same as:

$$(S_F)_{\alpha\beta}(x - y) \equiv \langle 0 | \psi(x) \bar{\psi}(y) | 0 \rangle \quad x^0 > y^0 \quad (7.3.2)$$

$$(S_F)_{\alpha\beta}(x - y) \equiv \langle 0 | -\bar{\psi}(x) \psi(y) | 0 \rangle \quad y^0 > x^0 \quad (7.3.3)$$

This minus is there because the operators anti-commute (and it is a necessity for Lorentz invariance). We can show that:

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip \cdot (x-y)} \frac{\not{p} + m}{p^2 - m^2 + i\epsilon} \quad (7.3.4)$$

This satisfies:

$$(i\not{\partial}_x - m)S_F(x-y) = i\delta^{(4)}(x-y) \quad (7.3.5)$$

Therefore $S_F(x-y)$ is a Green function for the Dirac operator.

Wicks theorem

We define the contraction between the two operators as usual:

$$\psi(x)\bar{\psi}(y) = T(\psi(x)\psi(y)) - : \psi(x)\psi(y) : = S_F(x-y) \quad (7.3.6)$$

Yukawa Theory

Single real scalar field ϕ of mass μ interacts with Dirac fermion of mass m :

$$L = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\mu^2\phi^2 + \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \lambda\phi\psi\bar{\psi} \quad (7.3.7)$$

We looked at very similar things when we looked at complex scalar fields interacting with a real scalar field. There are interactions of exactly this type in the standard model, the ψ 's represent the Fermions (electrons, quarks etc) and the ϕ is the Higgs field.

7.4 Dimensional analysis

We know that the Lagrangian should have dimension 4, so that the action has dimensions of zero. The first term involves two derivatives, therefore they have dimension 2 hence:

$$[\phi] = 1$$

Now we look at the third term that involves the ψ 's. The derivative has dimension 1, the mass is the same, so to make it dimensionally correct the ψ 's would need to have dimensions of $4 - 1 = 3$, but there are two ψ 's, so it is $\frac{3}{2}$:

$$[\psi] = \frac{3}{2}$$

Finally, the last term in the Lagrangian shows that λ has to be a pure number (dimensionless):

$$[\lambda] = 0$$

So for weak coupling λ must be small (in reality it is about 10^{-6} for electron coupling to the Higgs)

7.5 Feynmann rules for Yukawa theory

- To each incoming fermion with momentum p and spin s , we associate, $u_{\mathbf{p}}^s$
- For outgoing fermions we write down, $\bar{u}_{\mathbf{p}}^s$
- If we have incomin/outgoing anti-fermions, we have $\bar{v}_{\mathbf{p}}^s/v_{\mathbf{p}}^s$
- There is a symmetry:

$$\psi \rightarrow e^{i\alpha} \psi \quad (7.5.1)$$

which gives rise to a current:

$$j^\mu = \bar{\psi} \gamma^\mu \psi \quad (7.5.2)$$

this has a charge:

$$Q = \int d^3x \bar{\psi} \gamma^0 \psi = N_b - N_c \quad (7.5.3)$$

Which shows the conservation of anti-particle and particle number.

- Each vertex gets a factor of $(-i\lambda)$.
- Internal lines carry propagator terms with them:

$$\frac{i}{p^2 - u^2 + i\epsilon} \quad \text{Scalar propagator} \quad (7.5.4)$$

$$\frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \quad \text{Fermion propagator} \quad (7.5.5)$$

Impose momentum conservation at each vertex and integrate over undetermined loop momenta (Note that there are extra minus signs in this theory as we have anti-commutation relations as supposed to commutation relations). An example is the nucleon scattering:

Nucleon scattering

Scattering amplitude is simply given the Feynmann rules for this theory which we described above:

$$A = (-i\lambda)^2 \left(\frac{[\bar{u}_{\mathbf{p}}^s \cdot u_{\mathbf{p}}^s][\bar{u}_{\mathbf{q}}^s \cdot u_{\mathbf{q}}^s]}{(p - p')^2 - \mu^2} - \frac{[\bar{u}_{\mathbf{p}}^{s'} \cdot u_{\mathbf{q}}^r][\bar{u}_{\mathbf{q}}^{r'} \cdot u_{\mathbf{p}}^s]}{(p - p')^2 - \mu^2} \right) \quad (7.5.6)$$

where:

$$u_{\mathbf{p}}^s = \left(\frac{\sqrt{p} \cdot \sigma \xi^1}{\sqrt{p} \cdot \sigma \xi^2} \right) \quad (7.5.7)$$

ξ 's represent the usual spinors.

8 Quantum-Electro-Dynamics (QED)

QED is theory of light interacting with matter (in this case we just mean electron). We have seen how to quantise spin zero fields and spin half fields. Now we want to quantise light (which was the motivation for field theory, Maxwell's equations).

8.1 Maxwell's Equations

We start with the Lagrangian:

$$L = -\frac{1}{4}F_\mu F^\mu \quad (8.1.1)$$

where

$$F_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (8.1.2)$$

The equation of motion for this Lagrangian come from the usual E-L equation:

$$\partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu A_\nu)} \right) = -\partial_\mu F^{\mu\nu} = 0 \quad (8.1.3)$$

Field strength follows the Bianchi identity:

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \quad (8.1.4)$$

These are Maxwell's equations. We want to quantise this Lagrangian, therefore we follow the same steps that we have followed so far for scalar fields and spinor fields. But there will be a problem because, the A_μ which is the field we want to quantise has an index which has four values (0,1,2,3). If we were quantise naively we would say there are four fields and each of them will give rise to a different type of particle, so this would mean we would have a photon with four d.o.f However this is not what is observed as a photon only has d.o.f (polarisation up or down).

So somehow when we quantise A_μ we have to loose two d.o.f, which is the difficulty in the theory.

Comment 1: A_0 is not dynamical, i.e there is no \dot{A}_0 terms in the action. This means that if we are given A_i and \dot{A}_i , then A_0 is fixed by:

$$\underline{\nabla} \cdot \mathbf{E} = 0 \rightarrow \nabla^i A_0 + \underline{\nabla} \cdot \dot{\mathbf{B}} = 0 \quad (8.1.5)$$

This gives a solution for A_0 to be:

$$A_0(x) = \int d^3x' \frac{(\underline{\nabla} \cdot \dot{\mathbf{A}})(\mathbf{x}')}{4\pi|\mathbf{x} - \mathbf{x}'|} \quad (8.1.6)$$

Comment 2: The theory has a very large symmetry group, acting on the vector potential as:

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \lambda(x) \quad (8.1.7)$$

for any function of x , $\lambda(x)$. The only requirement on $\lambda(x)$ is that it dies of quickly as $x \rightarrow \infty$. To see that this is a symmetry of the Lagrangian, we substitute it into F:

$$F_\mu \rightarrow \partial_\mu(A_\mu + \partial_\nu \lambda) - \partial_\nu(A_\mu + \partial_\mu \lambda) \quad (8.1.8)$$

Therefore $F_m u$ is invariant and since the action only depends on the field strength, it does not change either. Previously we have had symmetries that remain symmetric depending on a number (i.e, rotate by x amount of degrees or boost by yms^{-1}). Here we have a symmetry that holds even when we act on the field with an entire function. So now we have a theory that has an infinite number of symmetries as we can choose an infinite number of functions. Now we might think that each of the symmetries, obey Noethers theorem and lead to conserved quantities. The symmetries we have depend of space (all the ones we have encountered so far depend on a parameter). These symmetries are called *gauge symmetries* or *local symmetries*, however these are not really symmetries as they do not take one physical state to a different physical state. Instead they are to be viewed as redundancies in our description of the system. To see this we could notice that Maxwell's equations do not determine the evolution of A_μ :

$$[\eta_{\mu\nu}(\partial^\rho \partial_\rho - \partial_\mu \partial_\nu)]A^\nu = 0 \quad (8.1.9)$$

But the operator $[\eta_{\mu\nu}(\partial^\rho \partial_\rho - \partial_\mu \partial_\nu)]$ is not invertible as it annihilates any function of the form $\partial_\mu \lambda$. This means we cannot figure out uniquely what A is doing as the operator will annihilate $A_m u$ and $A_m u + \partial_\mu \lambda$ and therefore we cannot distinguish between different states. So we can either say that our theory is not well defined, or we could say that actually this is okay, because if we identify physical states A and $A + \partial \lambda$ as the same physical state, then the evolution operator is telling us the truth and the physics in both states is the same. So the picture we have of the phase space of the theory is that each point in the space has a different value of A_i and \dot{A}_i . There will be lines that correspond to specific states that are called gauge orbits, therefore all functions that lie on these lines correspond to the same physics. To calculate we often have to make a gauge choice.

Examples

Lorentz gauge

$$\partial_\mu A^\mu = 0 \quad (8.1.10)$$

Suppose we started with some $A'_m u$ such that:

$$\partial_\mu (A')^\mu = f \quad (f \neq 0) \quad (8.1.11)$$

Act with $A_\mu = A'_\mu + \partial_\mu \lambda$ such that:

$$\partial_\mu \partial^\mu \lambda = -f \quad (8.1.12)$$

This gauge is Lorentz invariant (hence the name!). The equation $\partial_\mu \partial^\mu \lambda = -f$, always has a solution. This does not uniquely pick a configuration along the orbit as we can always act with:

$$A_\mu \rightarrow A_\mu + \partial_\mu \lambda$$

$$\partial_\mu \partial^\mu \lambda = 0$$

Coulomb gauge

We can get to this by making use of the residual gauge symmetry in Lorentz gauge:

$$\underline{\nabla} \cdot \mathbf{A} = 0 \rightarrow A_0 = 0 \tag{8.1.13}$$

The physical d.o.f are immediately apparent in this gauge because we have set A_0 to zero, so we just forget about it and A satisfies $\underline{\nabla} \cdot \mathbf{A}$ which basically means that there are no longitudinal polarisations leaving us with the 2 d.o.f that we require and it is also Lorentz invariant.

Now we have to make a choice as to which gauge we should use.

8.2 Lorentz gauge quantisation

First we use the Lorentz gauge:

$$\partial_\mu A^\mu = 0 \tag{8.2.1}$$

The equations of motion for A_μ are:

$$(\partial_\mu \partial^\mu) A^\mu = 0 \tag{8.2.2}$$

So each component of A^μ obeys the massless Klein-Gordon equation. To quantise this theory we shall write down a Lagrangian that produce the equations of motion above. After quantisation we impose the residual gauge symmetry condition and $\partial_\mu A^\mu = 0$. The Lagrangian we work with is:

$$L = -\frac{1}{4} F_\mu F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2 \tag{8.2.3}$$

This Lagrangian gives 8.2.2, but it only agrees with Maxwell's theory when we impose the conditions I stated above. Firstly we compute the conjugate momenta of A :

$$\pi^0 = \frac{\partial L}{\partial \dot{A}_0} = -\partial_\mu A^\mu \tag{8.2.4}$$

$$\pi^i = \frac{\partial L}{\partial \dot{A}_i} = \partial^i A^0 - \dot{A}^i \tag{8.2.5}$$

To quantise we impose canonical commutation relations:

$$[A_\mu(\mathbf{x}), A_\nu(\mathbf{y})] = [\pi^\mu(\mathbf{x}), \pi^\nu(\mathbf{y})] = 0 \quad (8.2.6)$$

$$[A_\mu(\mathbf{x}), \pi_\nu(\mathbf{y})] = i\eta_{\mu\nu}\delta^{(3)}(\mathbf{x} - \mathbf{y}) \quad (8.2.7)$$

The \mathbf{x} shows that we are currently working in the Schroedinger picture. Now we write down the mode expansions:

$$A_\mu(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\mathbf{p}|}} \sum_{\lambda=0}^3 \epsilon_\mu^\lambda(\mathbf{p}) [a_\mathbf{p}^\lambda e^{i\mathbf{p}\cdot\mathbf{x}} + a_\mathbf{p}^{\lambda\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}}] \quad (8.2.8)$$

$$\pi^\mu(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{|\mathbf{p}|}{2}} i \sum_{\lambda=0}^3 (\epsilon^\mu)^\nu(\mathbf{p}) [a_\mathbf{p}^\lambda e^{i\mathbf{p}\cdot\mathbf{x}} - a_\mathbf{p}^{\lambda\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}}] \quad (8.2.9)$$

These are similar to previous mode expansions, but there are a few subtalties:

- There used to be $E_\mathbf{p}$ in the equation before, but this is a massless particle, therefore it has no $E_\mathbf{p}$'s instead its energy is just the modulus of its momentum, $|\mathbf{p}|$. There is a change in sign, but that is conventially introduced to make things simpler.
- Only a 's and a^\dagger 's appear as this is a real field (not complex).
- The one new thing is the $(\epsilon^\mu)^\lambda$ which are the polarisation vectors. These polarisation vectors have 4 vectors ($\lambda = 0, 1, 2, 3$) spanning the vector space. We want to define ϵ^0 to be time like:

$$(\epsilon^0)^\lambda \rightarrow \lambda = 1, 2, 3 \quad (8.2.10)$$

and $\epsilon^{1,2,3}$ to be spacelike. We choose an orthonormal basis:

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \quad (8.2.11)$$

We will further pick ϵ^3 to be the longitudinal polarisation and ϵ^1 and ϵ^2 to lie transverse to the momentum:

$$\epsilon^1 \cdot p = \epsilon^2 \cdot p = 0 \quad (8.2.12)$$

So these polarisation vectors depend on the momentum p of the photon.

Example: $p \approx (1, 0, 0, 1)$

We have:

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Now we translate the field commutation relations into relations involving creation and annihilation operators:

$$[a_{\mathbf{p}}^\lambda, a_{\mathbf{q}}^{\lambda'}] = [a_{\mathbf{p}}^{\lambda\dagger}, a_{\mathbf{q}}^{\lambda'\dagger}] = 0 \quad (8.2.17)$$

$$[a_{\mathbf{p}}^\lambda, a_{\mathbf{q}}^{\lambda'\dagger}] = -\eta^{\lambda\lambda'} (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (8.2.18)$$

The problem that we have is the Minkowski metric, $\eta^{\lambda\lambda'}$, it has +'s and -'s. For space-like polarisation states ($\lambda = 1, 2, 3$):

$$[a_{\mathbf{p}}^\lambda, a_{\mathbf{q}}^{\lambda'\dagger}] = (2\pi)^3 \delta^{\lambda\lambda'} \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (\lambda, \lambda' = 1, 2, 3) \quad (8.2.19)$$

For time like annihilation and creation operators, we have:

$$[a_{\mathbf{p}}^0, a_{\mathbf{q}}^{0\dagger}] = -(2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (8.2.20)$$

If we define the vacuum as:

$$a_{\mathbf{p}}^\lambda |0\rangle = 0 \quad (8.2.21)$$

Then the usual one particle state we get by acting with creation operators on the vacuum:

$$a_{\mathbf{p}}^{\lambda\dagger} |0\rangle = |\mathbf{p}, \lambda\rangle \quad (8.2.22)$$

For space-like polarisation states ($\lambda = 1, 2, 3$) all is well. But for the time-like polarisation $\lambda = 0$, the state $|\mathbf{p}, 0\rangle$ has a negative norm:

$$\langle \mathbf{p}, 0 | \mathbf{q}, 0 \rangle = \langle 0 | a_{\mathbf{p}}^0 a_{\mathbf{q}}^{0\dagger} | 0 \rangle = -(2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (8.2.23)$$

This is something very strange; we have a Hilbert space, and it does not have a positive norm, and this leads to concepts like negative probabilities. What has happened is that commutation relation gave rise to the Minkowski metric of space-time and this consists either space to be

negative or time. These negative norm states are called *ghosts*. We have quantised this theory but we still have the constraint:

$$\partial_\mu A^\mu = 0 \tag{8.2.24}$$

So now we shall impose this constraint together with this residual gauge symmetry (which is left over in the Lorentz gauge), which should eventually give us a sensible Hilbert space for our theory. There are three ways to impose this constraint:

Option 1

We can turn A into an operator in the Heisenberg picture and consider $\partial_\mu A^\mu = 0$ as a restriction on operators. This constraint is too strong as it is not consistent with the commutation relations ($\pi^0 = -\partial_\mu A^\mu \neq 0$).

Option 2

The problem we are finding in this theory is that the Hilbert space is not as well behaved as we would like. So we could try to define sensible states by the requirement that:

$$\partial_\mu A^\mu |\psi\rangle = 0 \tag{8.2.25}$$

This does not work, to see why consider:

$$A_\mu(x) = A_\mu^+(x) + A_\mu^-(x) \tag{8.2.26}$$

where A_μ^+ is an annihilation operator and A_μ^- is the creation operator. But if we impose this constraint on $|0\rangle$:

$$\partial_\mu A_\mu^+ |0\rangle = 0 \tag{8.2.27}$$

$$\partial^\mu A_\mu^- |0\rangle \neq 0 \tag{8.2.28}$$

So not even the vacuum is a physical state according to this constraint.

Option 3

We'll define the physical states within our Hilbert space to obey:

$$\partial^\mu A_\mu^+ |\psi\rangle = 0 \tag{8.2.29}$$

This insures that for any two physical states $|\psi\rangle$ and $|\bar{\psi}\rangle$:

$$\langle \bar{\psi}' | \partial_\mu A^\mu |\psi\rangle = 0 \tag{8.2.30}$$

This is known as the *Gupta-Bleuler* condition. Now let's think about this physical Hilbert space, H_{phys} , looks like (another way of saying which states of the Fock space satisfy our constraint). Now we can decompose the state into:

$$|\psi\rangle = |\psi_T\rangle \times |\phi\rangle \tag{8.2.31}$$

$|\psi_T\rangle$ = contains the transverse photons of the state

$|\phi\rangle$ = contains the time-like and longitudinal photons of the state

The $|\psi_T\rangle$ state is created by $a_{\mathbf{p}}^{1,2\dagger}$ and $|\phi\rangle$ is created by $a_{\mathbf{p}}^{0\dagger}$ and $a_{\mathbf{p}}^{3\dagger}$.

Therefore applying this to $\partial^\mu A_\mu^\dagger|\psi\rangle = 0$ we have:

$$(a_{\mathbf{p}}^3 - a_{\mathbf{p}}^0)|\phi\rangle = 0 \quad (8.2.32)$$

This equation is telling us that any physical state can only have equal numbers of time-like and longitudinal photons. So this condition does eliminate negative norm states but it replaces them with zero norm states:

$$\langle\phi|\phi\rangle = 0 \quad (8.2.33)$$

as the time-like photons are equal and opposite to the Longitudinal photons and they just cancel. An example of a zero norm state:

$$|\phi\rangle = a_{\mathbf{p}}^{0+} + |0\rangle + a_{\mathbf{p}}^{3+}|0\rangle \quad (8.2.34)$$

In the Lorentz gauge we also have some residual symmetry. In this case all the states of the form:

$$|\psi\rangle = |\psi_T\rangle|\phi\rangle \quad (8.2.35)$$

with any $|\phi\rangle$ with zero norm are considered states of the Hilbert space. We treat the zero norm states as gauge equivalent to zero. Two states which differ just in their longitudinal and timelike polarisation states are said to be physically equivalent. To show that this is true we want to evaluate a physical observable in both these states and show that it is indeed the same. The observable could be energy, so we consider the Hamiltonian:

$$H = \int \frac{d^3p}{(2\pi)^3} |\mathbf{p}| \left(\sum_{i=1}^2 a_{\mathbf{p}}^{i\dagger} a_{\mathbf{p}}^i - a_{\mathbf{p}}^{0\dagger} \right) \quad (8.2.36)$$

But:

$$(a_{\mathbf{p}}^3 - a_{\mathbf{p}}^0)|\psi\rangle = 0 \quad (8.2.37)$$

therefore we get:

$$\langle\psi|a_{\mathbf{p}}^{3\dagger} - a_{\mathbf{p}}^3|\psi\rangle = \langle\psi|a_{\mathbf{p}}^{0\dagger} - a_{\mathbf{p}}^0|\psi\rangle \quad (8.2.38)$$

so the contribution from the timelike and longitudinal photons cancel amongst themselves in the Hamiltonian. There is no restriction on how many transverse photons and this is the only thing that $|\psi\rangle$ depends on.

Photon propagator

Now we can finally compute the photon propagator (also could be called the Feynmann propagator for QED):

$$\langle 0|T A_\mu(x) A_\nu(y)|0\rangle = \int \frac{d^4 p}{(2\pi)^3} - \frac{i\eta_{\mu\nu}}{p^2 + i\epsilon} e^{-p \cdot (x-y)} \quad (8.2.39)$$

Coupling to Fermions (matter)

We want an interacting Lagrangian:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A_\mu j^\mu \quad (8.2.40)$$

j^μ depends on other fields. The equations of motion are:

$$\partial_\mu F^\mu = j^\mu \quad (8.2.41)$$

These are Maxwell's equations in the presence of sources, where j^μ is a current. There is a consistency requirement that has to be met:

$$\partial_\mu \partial_\nu F^{\mu\nu} = 0 = \partial_\nu j^\nu \quad (8.2.42)$$

This is just the definition of the current that we saw in Noethers theorem!.

The Dirac Lagrangian is:

$$L = \bar{\psi}(i\cancel{D} - m)\psi \quad (8.2.43)$$

and we know it has the symmetry:

$$\psi \rightarrow e^{i\alpha} \psi \quad (8.2.44)$$

And using Noethers theorem, the current can be computed:

$$j^\mu = \bar{\psi} \gamma^\mu \psi \quad (8.2.45)$$

This tells us that we can write down a theory of light interacting with Fermions by coupling this current with the Gauge field. So we consider the Lagrangian:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\cancel{D} - m)\psi - e\bar{\psi}\gamma^\mu A_\mu \psi \quad (8.2.46)$$

The e is the coupling constant which is some number. To see that this Lagrangian is gauge invariant, write:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\cancel{D} - m)\psi \quad (8.2.47)$$

where:

$$D_\mu \psi = \partial_\mu \psi + ie A_\mu \psi \quad (8.2.48)$$

D is called the covariant derivative (remember the slash just indicates that the term is contracted with γ indicies). Under the gauge transformations we have:

$$A_\mu \rightarrow A_\mu + \partial_\mu \lambda \quad (8.2.49)$$

$$\psi \rightarrow e^{-ie\lambda}\psi \quad (8.2.50)$$

for an arbitrary function $\lambda(x)$ that decays suitably quickly as $x \rightarrow \infty$. To see that this works for $D_\mu\psi$:

$$\begin{aligned} D_\mu\psi &= \partial_\mu\psi + ieA_\mu\psi \\ &= \partial_\mu(e^{-ie\lambda}\psi) + ie(A_\mu + \partial_\mu\lambda)(e^{-ie\lambda}\psi) \\ &= e^{-ie\lambda}D_\mu\psi \end{aligned} \quad (8.2.51)$$

So the covariant derivative just picks up a phase under the gauge transformations, with the derivative of $e^{-ie\lambda}$ cancelling the transformation of the gauge fields. This ensures that the whole Lagrangian is invariant, since:

$$\bar{\psi} \rightarrow e^{ie\lambda(x)}\bar{\psi} \quad (8.2.52)$$

This is also the way gravity couples to matter. Now returning to the Lagrangian 8.2.47, the e that sits in \not{D} is related to the fine structure constant:

$$\alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137} \quad (8.2.53)$$

The coupling constant actually turns out to be the charge of the electron.

8.3 Feynmann rules for QED

The Feynmann rules are given by using Wick's theorem and Dyson's formula ²⁰.

– Each vertex has:

$$(-ie\gamma^\mu) \quad (8.3.1)$$

– The photon propagator is:

$$\frac{-i\eta_{\mu\nu}}{p^2 + i\epsilon} \quad (8.3.2)$$

– The fermion propagator is:

$$\frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \quad (8.3.3)$$

– For external lines we add:

Photons: $\epsilon_{in}^\mu/\epsilon_{out}^\mu$ for incoming/outgoing photons (polarisation vectors)

Fermions: we add information about the spin; $u_{\mathbf{p}}^r/\bar{u}_{\mathbf{p}}^r$ for incoming/outgoing fermions.
We add a spinor $\bar{v}_{\mathbf{p}}^r/v_{\mathbf{p}}^r$ for incoming/outgoing anti-fermions/

²⁰I had done this for one the early theories, however due to the length of the equations I will not go through Wicks theorem for each theory, I will simply quote the results

As a final example we can compute the scattering amplitude of the electron.

Electron scattering: $e^-e^- \rightarrow e^-e^-$ (tree-level diagram)

The amplitude, A , is just given by the product of the appropriate Feynmann rules:

$$A = -i(-ie)^2 \left(\frac{[\bar{u}_{\mathbf{p}'}^s \gamma^\mu u_{\mathbf{p}}^s][\bar{u}_{\mathbf{q}'}^r \gamma_\mu u_{\mathbf{q}}^r]}{(p' - p)^2} - \frac{[\bar{u}_{\mathbf{p}'}^s \gamma^\mu u_{\mathbf{q}}^r][\bar{u}_{\mathbf{q}'}^r \gamma_\mu u_{\mathbf{p}}^s]}{(p - q')^2} \right) \quad (8.3.4)$$

This is almost exactly the same as was found for Yukawa theory, the difference are that the propagator is mass-less, as the exchange particle is a photon and there is a γ^μ that multiplies the spin vectors and it tells you that interactions between spinning particles are going to be different depending on the spin of the mediating particle.

9 More on symmetries

So we have seen that symmetries have proven to be very useful to guide our physical theories. The key symmetries built into all the QFT's so far the spacetime time symmetries that we continue to call Lorentz symmetry. There are 10 of these symmetries corresponding to the 4 translations, 3 rotations and 3 boosts. The conservation laws associated with these symmetries ensure that linear momentum (translation invariance), energy (time translation), angular momentum (rotational invariance) are conserved. If we are in D dimensions we can generalise to say that there are D translations, $d = D - 1$ boosts and $d(d - 1)/2$ rotations. (Note: the number of independent planes of rotations is $[d/2]$ where $[\]$ denotes the integer part e.g. the (x, y) and (z, w) planes in $d = 4$ define two independent angular momenta.) There are other internal symmetries that act on the fields in the so called *field space*. For example there is a $U(1)$ gauge symmetry in QED. The charge carried by a field can be thought of as like a handle pointing in field space onto which a gauge boson can hold on.

Gauge fields can be in three distinct phases of physical behaviour. The most familiar is the Coulomb phase, resulting in an inverse square law in $D = 4$ as per intuition. The other phase is confinement which is seen in QCD. The final phase is known as the *Higgs phase* which involves spontaneous symmetry breaking (SSB).

9.1 Review of Noether's theorem

We have seen before that Noether's theorem implies that for every continuous symmetry there exists a conservation law along with a conserved charge. For example translational symmetry $x^\mu \rightarrow x^\mu + \delta x^\mu$ gives rise to conserved linear momentum and energy. The internal symmetries that act on fields are said to carry representations of the gauge group. A general symmetry transformation may act on both the fields and the coordinates. Consider a general collection of fields $\{\phi^a\}$ where a represents a general index; our general symmetry transformation acts as,

$$\Delta\phi^a = \phi^{a'}(x') - \phi^a(x) \tag{9.1.1}$$

which we can rewrite as,

$$\begin{aligned} \Delta\phi^a &= \phi^{a'}(x') + (-\phi^{a'}(x) + \phi^{a'}(x)) - \phi^a(x) \\ &= (\phi^{a'}(x') - \phi^{a'}(x)) + (\phi^{a'}(x) - \phi^a(x)) \\ &= (\partial_\mu\phi^a)\delta x^\mu + \delta\phi^a(x) \end{aligned} \tag{9.1.2}$$

where in the last step we used the chain rule. Notice that we have obtained the second term from the direct variation of the field while the first, known as the transport term, arose more indirectly through the dependence of the field on coordinates. To prove Noether's theorem, consider the action for the fields $\{\phi^a\}$,

$$S = \int d^D x L \tag{9.1.3}$$

then the variation is

$$\delta S = \int (\delta(d^D x)L + d^D x(\delta L)) \quad (9.1.4)$$

To see how the measure of integration $d^D x$ vary under such a symmetry transformation we can look at the Jacobian, $J(x'|x) : dx : d^D x' \equiv d^D x J(x'|x)$. We have,

$$\begin{aligned} \frac{\partial x^{\mu'}}{\partial x^\nu} &= \frac{\partial}{\partial x^\nu} \{x^\mu + \delta x^\mu\} \\ &= \delta_\nu^\mu + \partial_\nu(\delta x^\mu) \end{aligned} \quad (9.1.5)$$

dropping second order terms,

$$J(x'|x) = 1 + \partial_\mu(\delta x^\mu) \quad (9.1.6)$$

The variation of the Lagrangian is given by the chain rule,

$$\delta L = \frac{\partial L}{\partial \phi^a} \delta \phi^a + \frac{\partial L}{\partial(\partial_\mu \phi^a)} \delta(\partial_\mu \phi^a) + \frac{\partial L}{\partial x^\mu} \delta x^\mu \quad (9.1.7)$$

where we have use $\delta(\partial_\mu \phi^a) = \partial_\mu(\delta \phi^a)$. Therefore, counting the variation of both the measure and the integrand we have,

$$\begin{aligned} \delta S &= \int d^D x \left((\partial_\mu \delta x^\mu) L + \frac{\partial L}{\partial \phi^a} \delta \phi^a + \frac{\partial L}{\partial \partial_\mu \phi^a} \partial_\mu(\delta \phi^a) + \partial_\mu L \delta x^\mu \right) \\ &= \int d^D x \left(\left(\frac{\partial L}{\partial \phi^a} - \partial_\mu \left(\frac{\partial L}{\partial \partial_\mu \phi^a} \right) \right) \delta \phi^a + \partial_\mu \left(\frac{\partial L}{\partial \partial_\mu \phi^a} \delta \phi^a + L \delta x^\mu \right) \right) \\ &= \int d^D x \partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu \phi^a)} (\delta \phi^a + (\partial_\lambda \phi^a) \delta x^\lambda) - \left(\frac{\partial L}{\partial(\partial_\mu \phi^a)} (\partial_\lambda \phi^a) - \delta_\lambda^\mu L \right) \delta x^\lambda \right) \\ &= \int d^D x \partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu \phi^a)} \Delta \phi^a - T_\lambda^\mu \delta x^\lambda \right) \end{aligned} \quad (9.1.8)$$

where we used the fact that there is no surface term and used the EL equations and also defined the stress energy tensor,

$$T_\lambda^\mu \equiv \frac{\partial L}{\partial(\partial_\mu \phi^a)} (\partial_\lambda \phi^a) - \delta_\lambda^\mu L \quad (9.1.9)$$