

Quantum Field Theory I

Instructor: Dr. Suvrat Raju

Adithya A. Rao

This is the lecture notes for Prof. Suvrat Raju's course on Quantum Field Theory I, given at ICTS Bangalore in the semester Jan-April 2024. The course is available on youtube [here](#).

There are a few sections in the notes which are in red. These are statements and calculations that I have done and concluded on my own and are not stated or performed in the lecture.

If you find any errors in the text, please send an email to me at adithyarao3132001@gmail.com

Contents

Introduction — The Machinery of QFT	1
1 The Need for Fields	1
1.1 ANY Dynamics Consistent with STR Demand Infinite Degrees of Freedom	1
1.2 Quantizing the STR Dynamics	2
1.3 Another Reason Why We Need Infinite DOFs	2
1.4 Regular QM Breaks Causality	2
2 Classical Field Theory	4
2.1 Dynamics in Classical Field Theory	4
2.2 Parameterizing the Phase Space	5
3 Quantizing the Fields	9
3.1 Hamiltonian in terms of the creation and annihilation operators	10
3.2 Momentum of Fields	10
3.3 Why a^\dagger and a are Creation and Annihilation Operators	11
3.4 Constructing the Hilbert Space	11
4 Detour — Schrodinger Representation	16
4.1 The Schrodinger Equation	16

	4.2	The Vacuum in Schrodinger Formulation	17
5		Causality	19
	5.1	Observables in QFT	19
	5.2	Does QFT Preserve Causality	19
	5.3	Localized Excitations in QFT	22
6		Interlude — Symmetries	24
	6.1	Noether’s Theorem	24
	6.2	Applications	26
7		Interactions	31
	7.1	The S –matrix	31
	7.2	The Interaction Hamiltonian	32
	7.3	The Interaction Picture	33
	7.4	The Time Evolution Operator	35
8		Wick’s Theorem	39
	8.1	Contractions	39
	8.2	The theorem	41
	8.3	The Feynman Propagator	43
	8.4	The retarded propagator	46

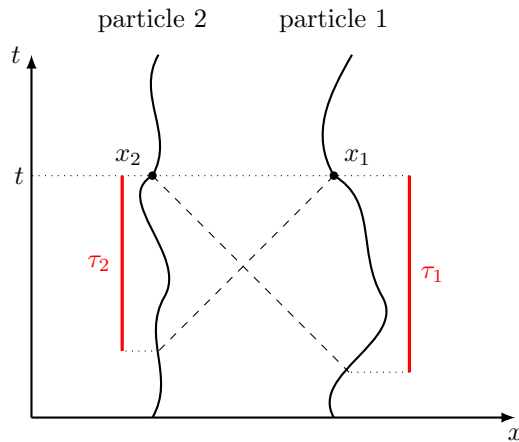
1 The Need for Fields

Quantum field theory is your normal quantum mechanics generalised to infinite number of degrees of freedom, where the dynamics of interactions between these infinite degrees of freedom is consistent with the postulates of Special Theory of Relativity (STR).

The question is then why these infinite number of degrees of freedom are needed. In non-relativistic classical mechanics, one sees that there is a requirement of only $6N$ number of degrees of freedom at-most, when describing the dynamics involving N particles, but when transitioning to QFT, why does describing even 2 particles require infinite degrees of freedom?

1.1 ANY Dynamics Consistent with STR Demand Infinite Degrees of Freedom

Consider two particles, at the same time slice t in some reference frame at positions x_1 & x_2 interacting with each other. Consistency with STR requires that the force on particle 1 due to 2 at this instant should depend on the retarded position of the particle 2 and not the position at that instant.



$$F_{12} = F(x_2(t - \tau_2), \dot{x}_2(t - \tau_2), \dots, x_1(t), \dot{x}_1(t), \dots) \quad (1.1)$$

and similarly

$$F_{21} = F'(x_1(t - \tau_1), \dot{x}_1(t - \tau_1), \dots, x_2(t), \dot{x}_2(t), \dots) \quad (1.2)$$

In such an equation, the τ_1 and τ_2 themselves are dynamically determined, and depend on the entire past history of the two particles, and therefore in order to determine the dynamics of the particles, one needs to keep track of functions, which have infinite degrees of freedom.

That said, even in classical dynamics, imposing STR postulates demands infinite number of degrees of freedom.

1.2 Quantizing the STR Dynamics

In dealing with such systems, we encounter *retarded differential equations*, or also called as *functional differential equations*. And nobody knows how to quantize the system starting from the functional differential equations.

Passage from CM to QM relies strongly on the idea that we are given an initial surface and we have a Hamiltonian/Lagrangian on it.

In a Hamiltonian system we have a phase space and we see how stuff evolves in that space. In the above discussed case, we have to deal with an infinite number of degrees of freedom and these degrees of freedom are not in the form that is convenient to put in the Hamiltonian formalism.

Fields solve this problem by rather than saying that particle 1 interacts with particle 2, we say particle 1 interacts with a field, and particle 2 too interacts with the field. The field and the particles influence each other only locally. In introducing this formalism we paid a price, we introduced infinite number of degrees of freedom in the form of fields, at the effect of removing the necessity of having to track all particle histories. The field can be thought of as keeping track of the history of the particles. In the language of fields, nothing depends on retarded times.

With the introduction of fields, we now have everything in the language of phase space. We are given an initial surface, and data on it, and we can evolve everything into future. But rather than having only finite degrees of freedom we still need to keep track of infinite degrees of freedom of the fields, but they are now in a form convenient to quantize.

1.3 Another Reason Why We Need Infinite DOFs

Heisenberg's uncertainty principle — $\Delta x \Delta p \geq 1$

If we have small Δx , then Δp becomes large, such that $\Delta E \geq m$. In this case we are now uncertain if we are dealing with the original particle, or if we have created a new particle. It is convenient therefore, to discuss the particles as excitations of fields, using which one can describe variable number of particles.

ALSO fields explain why all electrons are identical.

1.4 Regular QM Breaks Causality

This is a very generic calculation that can be found in almost all lectures/texbooks. But still included here for completeness.

The regular, single particle quantum mechanics that we all know of and are familiar with breaks causality. Causality requires that a particle can never move faster than the speed of light, i.e. a particle initially localised around the position \mathbf{x} should not have any amplitude to move out of the future light cone. Now in the case of single particle QM, consider a particle localised about \mathbf{x} , described by state $|\mathbf{x}\rangle$.

Suppose the particle follows the relativistic Hamiltonian, $\hat{H} = \sqrt{\hat{\mathbf{p}}^2 + m^2}$, the state of the particle after time t would be $\exp(-iHt)|\mathbf{x}\rangle$. Relativity requires that this state should have zero overlap with $|\mathbf{x}'\rangle$ where (t, \mathbf{x}) and (t, \mathbf{x}') are spacelike separated.

Let $G(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}' | \exp(-iHt) | \mathbf{x} \rangle$. The explicit value of G can be calculated by inserting the identity in $|\mathbf{p}\rangle$ as follows.

$$G(\mathbf{x}, \mathbf{x}', t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \langle \mathbf{x}' | \exp(-i\sqrt{\mathbf{p}^2 + m^2}t) | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{x} \rangle = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \exp(-i\sqrt{\mathbf{p}^2 + m^2}t) \langle \mathbf{x}' | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{x} \rangle \quad (1.3)$$

$$G(\mathbf{x}, \mathbf{x}', t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \exp(-i\sqrt{\mathbf{p}^2 + m^2}t) \exp(-i\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}')) \quad (1.4)$$

Write $\mathbf{x} - \mathbf{x}' = \mathbf{r}$, and align the z axis to be about \mathbf{r} . Then we can convert the above integral into polar coordinates as (where we use $p = |\mathbf{p}|$ and $r = |\mathbf{r}|$)

$$G(\mathbf{r}, t) = \int \frac{p^2}{(2\pi)^2} \exp(-i\sqrt{p^2 + m^2}t) \exp(-ipr \cos \theta) dp d(\cos \theta) \quad (1.5)$$

$$= \int \frac{p^2}{(2\pi)^2} \exp(-i\sqrt{p^2 + m^2}t) \left(\frac{\exp(-ipr) - \exp(ipr)}{pr} \right) dp \quad (1.6)$$

$$= \frac{2}{(2\pi)^2} \frac{1}{r} \int \exp(-i\sqrt{p^2 + m^2}t) p \sin(pr) dp \quad (1.7)$$

You can look up this integral in some Russian textbook and find out that this is equal to

$$G \propto K_2(m\sqrt{r^2 + t^2}) \quad (1.8)$$

This is some finite non-zero number (although it might be small for some values) for all r and t . Since we require causality to be exactly held, and not approximately held, single particle quantum mechanics breaks causality.

2 Classical Field Theory

Everything we discussed previously does not logically conclude that the *only* candidates are fields, but fields provide a convenient formalism, and end up solving the problems we were facing.

In general, field is something associated with spacetime. That is, for every spacetime point, it spits out some mathematical entity. For the field, we would like to write a Lagrangian, as an integral of some Lagrangian density over all space.

$$L = \int \mathcal{L} d^3\mathbf{x} \quad (2.1)$$

We consider Lagrangian densities of the form

$$\mathcal{L}(\partial_\mu\phi(x), \phi(x)) \quad (2.2)$$

(where we assume the notation $x = (t, \mathbf{x})$). In principle, the Lagrangian density can be anything, but mostly we consider the most simple forms. We usually start with something that is quadratic in fields and their derivatives. Linear term is ignored since it can be absorbed into the quadratic term by considering a shift by constant factor of the fields leading to the a quadratic Lagrangian.

Example — suppose

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - \alpha\phi \quad (2.3)$$

consider the transformation $\phi(x) \rightarrow \phi(x) + \frac{\alpha}{m^2}$, with α being some spacetime independent constant.

$$\mathcal{L}' = \frac{1}{2} \left(\partial_\mu \left(\phi' - \frac{\alpha}{m^2} \right) \right)^2 - \frac{1}{2}m^2 \left(\phi'^2 - 2\frac{\alpha}{m^2}\phi' + \left(\frac{\alpha}{m^2} \right)^2 \right) - \alpha\phi = \frac{1}{2}(\partial_\mu\phi')^2 - \frac{1}{2}m^2\phi'^2 \quad (2.4)$$

(dynamics doesn't change when a constant term is added to the Lagrangian) which is a Lagrangian density with only quadratic terms.

2.1 Dynamics in Classical Field Theory

Lagrangian Dynamics

Action — $S = \int \mathcal{L} d^4x$

Nature is lazy

Classical equations of motion — field configurations such that variation of S under $\phi \rightarrow \phi + \delta\phi$ is zero.

Under such change,

$$\begin{aligned} \delta S &= \int \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi} \delta(\partial_\mu\phi) + \frac{\delta\mathcal{L}}{\delta\phi} \delta(\phi) d^4x \\ &= \int -\delta\phi \partial_\mu \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi} + \delta\phi \frac{\delta\mathcal{L}}{\delta\phi} + \partial_\mu \left(\delta\phi \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi} \right) d^4x \end{aligned} \quad (2.5)$$

Considering variations which vanish at infinity, the last term (which is only a boundary term) vanishes. The other two functions also should vanish for every possible values of $\delta\phi$, therefore we get the equations

of motions to be

$$\partial_\mu \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \quad (2.6)$$

Hamiltonian Dynamics

In order to make a passage to QFT, we need a Hamiltonian formulation. In going to Hamiltonian formulation, we trade the *time derivative* with the *conjugate momentum*.

The basic degrees of freedom in classical mechanics were the spatial points \mathbf{x} . In the case of the field theory, the basic degrees of freedom are just the fields at a certain instant of time (which we conveniently call $t = 0$) $\psi(\mathbf{x})$. The $\partial_i \phi(\mathbf{x})$ are nothing but linear combinations of the $\phi(\mathbf{x})$, and in going to the Hamiltonian formulation, we treat the time derivative as a different degree of freedom.

$$\Pi(t = 0, \mathbf{x}) = \dot{\phi}(t = 0, \mathbf{x}) \quad (2.7)$$

The Hamiltonian is then

$$H(t) = \int \Pi \dot{\phi} - \mathcal{L} d^3 \mathbf{x} = \int \mathcal{H} d^3 \mathbf{x} \quad (2.8)$$

and one can write equations similar to the Hamiltonian equations of motion for the field and its conjugate momenta.

2.2 Parameterizing the Phase Space

Given the phase space, the usual parameterization is using Π and Φ . But we can also parameterize the phase space using different variables — the set of solutions to the equations of motion. This is true because given a point in phase space, we can use it as initial conditions and obtain a solution to the equations of motion. Therefore, there is a one-to-one map between points in phase space and solutions to the equations of motion, and we can see the set of solutions to the equations of motion as parameterizing the phase space. That is, a point corresponds to a solution, and a solution provided a section corresponds to a point.

In other words, the *parameterization* of all solutions to the equations of motion is the same as parameterization of the whole phase space.

For concreteness, let us start considering a specific theory — the scalar free field theory.

$$\mathcal{L} = \frac{1}{2} ((\partial_\mu \phi)^2 - m^2 \phi^2) \quad (2.9)$$

The equation of motion arising from this lagrangian is simply

$$(\partial_\mu \partial^\mu + m^2) \phi(x) = (\partial_t^2 - \nabla^2 + m^2) \phi(x) = 0 \quad (2.10)$$

We want to find all possible solutions to this equations of motion in order to parameterize the phase space. To find a parameterization for the solution of the equations of motion, the easiest way is to go to the

fourier space.

Writing $\phi(t, \mathbf{x}) = \int d^3\mathbf{p} \phi(t, \mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{x})$, we get

$$\int (\partial_t^2 + \mathbf{p}^2 + m^2) \phi(t, \mathbf{p}) d^3\mathbf{p} = 0 \quad (2.11)$$

We are keeping time as it is because in Hamiltonian formulation we require solutions at a given slice of time.

The solutions to the above equation of motion is (with $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$)

$$\phi(t, \mathbf{p}) = a(\mathbf{p}) \exp(-i\omega_{\mathbf{p}}t) + b(\mathbf{p}) \exp(i\omega_{\mathbf{p}}t) \quad (2.12)$$

Since we require $\phi(t, \mathbf{x})^* = \phi(t, \mathbf{x})$, we need $\phi(t, \mathbf{p})^* = \phi(t, -\mathbf{p})$, meaning

$$a^*(\mathbf{p}) \exp(i\omega_{\mathbf{p}}t) + b^*(\mathbf{p}) \exp(-i\omega_{\mathbf{p}}t) = a(-\mathbf{p}) \exp(-i\omega_{\mathbf{p}}t) + b(-\mathbf{p}) \exp(i\omega_{\mathbf{p}}t)$$

Comparing, we get $b(\mathbf{p}) = a^*(-\mathbf{p})$, therefore, the solutions are

$$\phi(t, \mathbf{p}) = a(\mathbf{p}) \exp(-i\omega_{\mathbf{p}}t) + a^*(-\mathbf{p}) \exp(i\omega_{\mathbf{p}}t) \quad (2.13)$$

and

$$\phi(t, \mathbf{x}) = \int a(\mathbf{p}) \exp(-i\omega_{\mathbf{p}}t + i\mathbf{p} \cdot \mathbf{x}) + a^*(\mathbf{p}) \exp(i\omega_{\mathbf{p}}t - i\mathbf{p} \cdot \mathbf{x}) \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \frac{d^3\mathbf{p}}{(2\pi)^3} \quad (2.14)$$

where in second term, we did the change $\mathbf{p} \rightarrow -\mathbf{p}$ since the integral is invariant under this change, and $\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}$ is simply a choice of normalisation.

We can also write this as

$$\phi(x) = \int a(p) \exp(-ip \cdot x) + a^*(p) \exp(ip \cdot x) \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \frac{d^3\mathbf{p}}{(2\pi)^3} \quad (2.15)$$

where now p and x are 4-vectors.

The coordinates on the phase space are $\phi(t=0, \mathbf{x})$ and $\Pi(t=0, \mathbf{x})$, and in this case the parameterization is very clear that

$$\phi(t=0, \mathbf{x}) = \int a(\mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{x}) + a^*(\mathbf{p}) \exp(-i\mathbf{p} \cdot \mathbf{x}) \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \frac{d^3\mathbf{p}}{(2\pi)^3} \quad (2.16)$$

and

$$\Pi(t=0, \mathbf{x}) = \int -i (a(\mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{x}) - a^*(\mathbf{p}) \exp(-i\mathbf{p} \cdot \mathbf{x})) \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \frac{d^3\mathbf{p}}{(2\pi)^3} \quad (2.17)$$

Rather than parameterizing the phase space using the field and conjugate momentum, we can parameterize

using the $a(\mathbf{p})$ and $a^*(\mathbf{p})$. We can also invert these relations to write a and a^* in terms of field and momenta.

$$a(\mathbf{p}) = \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \int \phi(0, \mathbf{x}) e^{-i\mathbf{p} \cdot \mathbf{x}} d^3\mathbf{x} + \frac{i}{\sqrt{2\omega_{\mathbf{p}}}} \int \Pi(0, \mathbf{x}) e^{-i\mathbf{p} \cdot \mathbf{x}} d^3\mathbf{x} \quad (2.18)$$

$$a^*(\mathbf{p}) = \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \int \phi(0, \mathbf{x}) e^{i\mathbf{p} \cdot \mathbf{x}} d^3\mathbf{x} - \frac{i}{\sqrt{2\omega_{\mathbf{p}}}} \int \Pi(0, \mathbf{x}) e^{i\mathbf{p} \cdot \mathbf{x}} d^3\mathbf{x} \quad (2.19)$$

In order to understand this better, let us do the same analysis for the simpler classical mechanical system of a harmonic oscillator. The equation of motion is

$$(\partial_t^2 + \omega^2)x = 0$$

This is solved simply by

$$x(t) = \frac{1}{\sqrt{2\omega}} (A \exp(-i\omega t) + A^* \exp(i\omega t)) \quad \because x \text{ is real}$$

where the prefactor is simply a choice of normalisation, and

$$p(t) = -i\sqrt{\frac{\omega}{2}} (A \exp(-i\omega t) - A^* \exp(i\omega t))$$

The coordinates on the phase space are $x(t=0)$ and $p(t=0)$, which in this case are

$$x = \frac{1}{\sqrt{2\omega}} (A + A^*) \quad \& \quad p = -i\sqrt{\frac{\omega}{2}} (A - A^*)$$

For every x, p , there is a solution A (where A is complex) that has (x, p) as its initial conditions, and therefore the set of A can be thought of as parameterizing the entire phase space. This is exactly what is done in introducing the ladder operators for quantizing the harmonic oscillator. Instead of imposing the quantization condition on the usual coordinates x and p , one uses the coordinates A and A^* and quantizes these degrees of freedom.

Important! — There can be confusion in what one calls a solution to the equation of motion. One might be tempted to assume that, for example in this case, the set of ellipses in the phase space to be the set of solutions to the equations of motions. This is wrong, in the sense that the solutions for motion with initial state x_1, p_1 and with initial state x_2, p_2 , even if both lie on the same ellipse, are different. Therefore, given an initial point in phase space, we determine a unique equation that governs this initial point's evolution and this equation is termed a solution. Therefore, corresponding to every point on the phase space, there is a unique solution, and therefore a bijection.

We can further extend the above discussion by looking at the commutator relations of A and A^* , which is

given by

$$\{x, p\} = 1 \implies \{A^*, A\} = -i, \{A, A\} = 0, \{A^*, A^*\} = 0$$

We can also construct the Hamiltonian in terms of the variables A and A^* , which is given as

$$H = \frac{1}{2}(p^2 + \omega^2 x^2) = \frac{1}{2}\omega(A^* A + A A^*) \quad (\star)$$

This reduces to

$$H = \omega A^* A$$

The commutators of Hamiltonian with the variables A^* and A are

$$\{H, A^*\} = \omega A^*, \quad \{H, A\} = -\omega A$$

Notice that all of these are classical calculations, and invokes no quantum mechanics.

3 Quantizing the Fields

To quantize a theory, we take a set of *coordinates* on phase space, and promote them to operators and impose the commutation relations.

In case of multiple particles in the regular quantum mechanics, the commutation relation that we used to impose was

$$[x_i, p_j] = i\delta_{ij}$$

Similarly, the $\phi(0, \mathbf{x})$ and $\Pi(0, \mathbf{x})$ behave like multiple degrees of freedom, this time instead of labeled by a discrete index i , they are labeled by a continuous index x . Therefore, we impose a similar commutation relation

$$[\phi(t, \mathbf{x}), \Pi(t, \mathbf{y})] = i\delta^3(\mathbf{x} - \mathbf{y}) \quad (3.1)$$

Notice that these are equal time commutation relations, i.e., this is imposed on one time slice. This is true even in the single particle quantum mechanics, where the canonical commutation relations hold only at equal times, i.e.

$$[x(t), p(t')] \neq i, \quad [x(t), x(t')] \neq 0, \quad [p(t), p(t')] \neq 0$$

As an example, consider the harmonic oscillator,

$$x(t) = \frac{1}{\sqrt{2\omega}}(A \exp(-i\omega t) + A^* \exp(i\omega t))$$

$$p(t) = -i\frac{\sqrt{\omega}}{2}(A \exp(-i\omega t) - A^* \exp(i\omega t))$$

The equal time commutation relation implies

$$[x(0), p(0)] = i[A, A^*] = i \implies [A, A^*] = 1, \quad \text{and} \quad [A, A] = [A^*, A^*] = 0$$

Using this,

$$[x(t), x(t')] = \frac{1}{2\omega} \left(e^{i\omega(t'-t)} [A, A^*] + e^{i\omega(t-t')} [A^*, A] \right) = \frac{1}{2\omega} \cos(\omega(t' - t)) \neq 0 \quad (\diamond)$$

In terms of $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$ (where we are now calling $a^* \equiv a^\dagger$ since we have promoted it to an operator), the commutation relations are

$$\begin{aligned} [a(\mathbf{p}_1), a^\dagger(\mathbf{p}_2)] &= \frac{-i}{2} \int [\phi(0, \mathbf{x}_1), \Pi(0, \mathbf{x}_2)] e^{-i\mathbf{p}_1 \cdot \mathbf{x}_1} e^{i\mathbf{p}_2 \cdot \mathbf{x}_2} d^3\mathbf{x}_1 d^3\mathbf{x}_2 + \frac{i}{2} \int [\Pi(0, \mathbf{x}_1), \phi(0, \mathbf{x}_2)] e^{-i\mathbf{p}_1 \cdot \mathbf{x}_1} e^{i\mathbf{p}_2 \cdot \mathbf{x}_2} d^3\mathbf{x}_1 d^3\mathbf{x}_2 \\ &= \int \delta^3(\mathbf{x}_1 - \mathbf{x}_2) e^{-i\mathbf{p}_1 \cdot \mathbf{x}_1} e^{i\mathbf{p}_2 \cdot \mathbf{x}_2} d^3\mathbf{x}_1 d^3\mathbf{x}_2 = (2\pi)^2 \delta^3(\mathbf{p}_1 - \mathbf{p}_2) \end{aligned} \quad (3.2)$$

It also holds that

$$[a(\mathbf{p}), a(\mathbf{q})] = [a^\dagger(\mathbf{p}), a^\dagger(\mathbf{q})] = 0 \quad (3.3)$$

3.1 Hamiltonian in terms of the creation and annihilation operators

The expression for Hamiltonian of the theory is

$$H = \frac{1}{2} \int \Pi(0, \mathbf{x})^2 + (\nabla \phi(0, \mathbf{x}))^2 + m^2 \phi(0, \mathbf{x})^2 d^3 \mathbf{x} \quad (3.4)$$

Calculating this term wise (note that for brevity, we are now calling $a(\mathbf{p}) \equiv a_{\mathbf{p}}$. See that $a_{\mathbf{p}}$ is not a number, but a function of \mathbf{p}).

$$\frac{1}{2} \int \Pi(0, \mathbf{x})^2 d^3 \mathbf{x} = \frac{1}{2} \int d^3 \mathbf{x} \frac{d^3 \mathbf{p} d^3 \mathbf{q}}{(2\pi)^6} \frac{\sqrt{\omega_{\mathbf{p}} \omega_{\mathbf{q}}}}{2} \left(-a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger e^{i(\mathbf{p}+\mathbf{q}) \cdot \mathbf{x}} + a_{\mathbf{p}}^\dagger a_{\mathbf{q}} e^{i(\mathbf{p}-\mathbf{q}) \cdot \mathbf{x}} + a_{\mathbf{p}} a_{\mathbf{q}}^\dagger e^{-i(\mathbf{p}-\mathbf{q}) \cdot \mathbf{x}} - a_{\mathbf{p}} a_{\mathbf{q}} e^{-i(\mathbf{p}+\mathbf{q}) \cdot \mathbf{x}} \right)$$

The integral over \mathbf{x} gives delta functions, and then we can perform an integral over \mathbf{q} to remove the delta function, leading to the simplification

$$\frac{1}{2} \int \Pi(0, \mathbf{x})^2 d^3 \mathbf{x} = \frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{\omega_{\mathbf{p}}}{2} (-a_{\mathbf{p}}^\dagger a_{-\mathbf{p}}^\dagger + a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + a_{\mathbf{p}} a_{\mathbf{p}}^\dagger - a_{\mathbf{p}} a_{-\mathbf{p}}) \quad (3.5)$$

$$\frac{1}{2} \int (\nabla \phi(0, \mathbf{x}))^2 d^3 \mathbf{x} = \frac{1}{2} \int d^3 \mathbf{x} \frac{d^3 \mathbf{p} d^3 \mathbf{q}}{(2\pi)^6} \frac{\mathbf{p} \cdot \mathbf{q}}{2\sqrt{\omega_{\mathbf{p}} \omega_{\mathbf{q}}}} \left(-a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger e^{-i(\mathbf{p}+\mathbf{q}) \cdot \mathbf{x}} + a_{\mathbf{p}}^\dagger a_{\mathbf{q}} e^{i(\mathbf{p}-\mathbf{q}) \cdot \mathbf{x}} + a_{\mathbf{p}} a_{\mathbf{q}}^\dagger e^{-i(\mathbf{p}-\mathbf{q}) \cdot \mathbf{x}} - a_{\mathbf{p}} a_{\mathbf{q}} e^{i(\mathbf{p}+\mathbf{q}) \cdot \mathbf{x}} \right)$$

which again simplifies to

$$\frac{1}{2} \int (\nabla \phi(0, \mathbf{x}))^2 d^3 \mathbf{x} = \frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{\mathbf{p}^2}{2\omega_{\mathbf{p}}} (a_{\mathbf{p}}^\dagger a_{-\mathbf{p}}^\dagger + a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + a_{\mathbf{p}} a_{\mathbf{p}}^\dagger + a_{\mathbf{p}} a_{-\mathbf{p}}) \quad (3.6)$$

and similarly

$$\frac{1}{2} \int m^2 \phi(0, \mathbf{x})^2 d^3 \mathbf{x} = \frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \frac{m^2}{2\omega_{\mathbf{p}}} (a_{\mathbf{p}}^\dagger a_{-\mathbf{p}}^\dagger + a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + a_{\mathbf{p}} a_{\mathbf{p}}^\dagger + a_{\mathbf{p}} a_{-\mathbf{p}}) \quad (3.7)$$

Adding the above three equations, using the fact that $\omega_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2$, we get

$$H = \frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} (a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + a_{\mathbf{p}} a_{\mathbf{p}}^\dagger) \quad (3.8)$$

3.2 Momentum of Fields

Claim — Momentum (not the conjugate momentum, but rather the physical momentum) has the form

$$\mathbf{P} = - \int d^3 \mathbf{x} \Pi(0, \mathbf{x}) \nabla \phi(0, \mathbf{x}) \quad (3.9)$$

Why?

We can see that

$$[\mathbf{P}, \phi(0, \mathbf{x})] = -i \nabla \phi(0, \mathbf{x})$$

Given this commutation relation, we can write, therefore, that

$$\phi(0, \mathbf{x} + \epsilon) = e^{-i\mathbf{P} \cdot \epsilon} \phi(0, \mathbf{x}) e^{i\mathbf{P} \cdot \epsilon} = \phi(0, \mathbf{x}) - i[\epsilon \cdot \mathbf{P}, \phi(0, \mathbf{x})] = \phi(0, \mathbf{x}) + \epsilon \cdot \nabla \phi(0, \mathbf{x}) \quad (3.10)$$

Therefore the operator \mathbf{P} generates translations.

In terms of a and a^\dagger (the calculation is simple, except that at the end we drop some terms owing to *normal ordering* which will be discussed later)

$$\mathbf{P} = \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \mathbf{p} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \quad (3.11)$$

3.3 Why a^\dagger and a are Creation and Annihilation Operators

Given the form of Hamiltonian in terms of a and a^\dagger , we immediately see that

$$[H, a_{\mathbf{p}}^\dagger] = \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger, \quad [H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}} \quad (3.12)$$

which means that for a stationary state $|E\rangle$ with energy E the operators $a_{\mathbf{p}}^\dagger$ and $a_{\mathbf{p}}$ map it to another state with energy $E + \omega_{\mathbf{p}}$ and $E - \omega_{\mathbf{p}}$

$$H(a_{\mathbf{p}}^\dagger |E\rangle) = (E + \omega_{\mathbf{p}})(a_{\mathbf{p}}^\dagger |E\rangle), \quad H(a_{\mathbf{p}} |E\rangle) = (E - \omega_{\mathbf{p}})(a_{\mathbf{p}} |E\rangle) \quad (3.13)$$

This behavior is exactly the same as that of a harmonic oscillator, and therefore the free Klein-Gordon theory is exactly a decoupled set of infinite harmonic oscillators.

3.4 Constructing the Hilbert Space

Vacuum

In order to construct the Hilbert space, we need to start with a vacuum, which is the state with the lowest energy.

Since vacuum has the lowest energy, operating on it with any $a_{\mathbf{p}}$ should annihilate the state. Therefore, the state vacuum is defined as the state

$$a_{\mathbf{p}} |0\rangle = 0 \quad \forall \mathbf{p}$$

The energy of vacuum is

$$H |0\rangle = \frac{1}{2} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}} a_{\mathbf{p}}^\dagger |0\rangle = \frac{1}{2} \int d^3 \mathbf{p} \delta^3(0) \omega_{\mathbf{p}} |0\rangle$$

where we picked the $\delta(0)$ by commuting the a and a^\dagger . The result states that the vacuum has infinite energy, which is expected since there are infinite number of harmonic oscillators and each of the oscillator contribute to a finite zero-point energy.

We don't have to worry about this so much because we are able to physically measure only the energy differences, in which case we can set, by-hand, the energy of vacuum to be 0. Another reason why we need not worry is because we are not even sure if the expression for the Hamiltonian is right. This is because when going from classical to quantum theory, there is an inherent ambiguity in operator ordering, and it is upto us to choose what ordering we use. As an example, consider a term xp in classical physics. How do you write the corresponding operator? There is an ambiguity in whether it will be $\hat{x}\hat{p}$ or $\hat{p}\hat{x}$ or even $\frac{\hat{x}\hat{p} + \hat{p}\hat{x}}{2}$ or any other combination. For an example in the case of harmonic oscillator see equation (\star), where the classical variables in the Hamiltonian commute, but there will be an ambiguity in the ordering when promoting the variables to operators.

One convenient choice is to place all the annihilation operator to the right and all the creation operators to the left, and it is called normal ordering, and is denoted as $\hat{\mathcal{O}}:$.

The normal ordered Hamiltonian is

$$:H:= \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}}^{\dagger} a_{\mathbf{p}} \quad (3.14)$$

and with this choice of H , we get

$$:H: |0\rangle = 0$$

First Excited States

There are an infinite number of first excited states given by

$$a_{\mathbf{p}}^{\dagger} |0\rangle$$

The energy of these states is $\omega_{\mathbf{p}}$ (since $|0\rangle$ has energy 0), and momentum

$$\mathbf{P}(a_{\mathbf{p}}^{\dagger} |0\rangle) = \int \frac{d^3\mathbf{q}}{(2\pi)^3} \mathbf{q} a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}} a_{\mathbf{p}}^{\dagger} |0\rangle = \int \frac{d^3\mathbf{q}}{(2\pi)^3} \mathbf{q} a_{\mathbf{q}}^{\dagger} (a_{\mathbf{p}}^{\dagger} a_{\mathbf{q}} + \delta^3(\mathbf{p} - \mathbf{q})) |0\rangle = \mathbf{p} a_{\mathbf{p}}^{\dagger} |0\rangle$$

The energy and momentum indeed satisfy the dispersion relation $\omega_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2$, and therefore these are to be interpreted as single particle states.

The normalisation of these states is given as

$$\langle \mathbf{p} | \mathbf{q} \rangle = \langle 0 | a_{\mathbf{p}} a_{\mathbf{q}}^{\dagger} | 0 \rangle = \langle 0 | a_{\mathbf{q}}^{\dagger} a_{\mathbf{p}} + (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{q}) | 0 \rangle = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{q}) \quad (3.15)$$

Delta function normalisation does not make any sense physically. Therefore in order to produce a normalisable state, we need to smear the operator, i.e. create a wavepacket

$$|f\rangle = \int f(\mathbf{p}) |\mathbf{p}\rangle \frac{d^3\mathbf{p}}{(2\pi)^3}$$

and for this state the normalisation is given by

$$\langle f'|f\rangle = \int f'(\mathbf{q}) |\mathbf{q}\rangle f'(\mathbf{p}) |\mathbf{p}\rangle \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6} = \int f'(\mathbf{p})f(\mathbf{p}) \frac{d^3\mathbf{p}}{(2\pi)^3}$$

These are not the only states in the Hilbert space of the system, there are other states too.

Higher Excited States

We can act with multiple creation operators to create higher excited states which can also be interpreted as n -particle states. As an example, let us check the state $a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle$.

$$H(a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle) = \omega_{\mathbf{q}} + \omega_{\mathbf{p}}$$

and

$$\begin{aligned} P(a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle) &= \int \frac{d^3\mathbf{r}}{(2\pi)^3} \mathbf{r} a_{\mathbf{r}}^\dagger a_{\mathbf{r}} a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle \\ &= \int \frac{d^3\mathbf{r}}{(2\pi)^3} \mathbf{r} a_{\mathbf{r}}^\dagger (a_{\mathbf{q}}^\dagger a_{\mathbf{r}} + \delta^3(\mathbf{q} - \mathbf{r})) a_{\mathbf{p}}^\dagger |0\rangle \\ &= \int \frac{d^3\mathbf{r}}{(2\pi)^3} \mathbf{r} a_{\mathbf{r}}^\dagger \delta^3(\mathbf{q} - \mathbf{r}) a_{\mathbf{p}}^\dagger |0\rangle + \int \frac{d^3\mathbf{r}}{(2\pi)^3} \mathbf{r} a_{\mathbf{r}}^\dagger a_{\mathbf{q}}^\dagger (a_{\mathbf{p}}^\dagger a_{\mathbf{r}} + \delta^3(\mathbf{p} - \mathbf{r})) |0\rangle \\ &= (\mathbf{p} + \mathbf{q})(a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle) \end{aligned}$$

Therefore this state has total energy equal to the sum of the energies of one particle states with momentum \mathbf{p} and \mathbf{q} , and total momentum equal to the vector sum of the two momenta. Therefore this state can be interpreted as a two particle state with the two particles having the same mass m and different momenta \mathbf{p} and \mathbf{q} .

The two particle states obey bosonic statistics. We can see this by noticing that $a_{\mathbf{p}}^\dagger$ commutes with $a_{\mathbf{q}}^\dagger$, and therefore the states $\langle \mathbf{q}, \mathbf{p} \rangle \equiv a_{\mathbf{q}}^\dagger a_{\mathbf{p}}^\dagger |0\rangle$ and $\langle \mathbf{p}, \mathbf{q} \rangle \equiv a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger |0\rangle$ are the same states. Therefore, these particles automatically satisfy Bose-Einstein statistics, and there is no need to introduce these statistics by hand.

The normalisation of the two particle states can be found as

$$\begin{aligned} \langle \mathbf{p}_2, \mathbf{q}_2 | \mathbf{p}_1, \mathbf{q}_1 \rangle &= \langle 0 | a_{\mathbf{p}_2} a_{\mathbf{q}_2} a_{\mathbf{p}_1}^\dagger a_{\mathbf{q}_1}^\dagger |0\rangle \\ &= \langle 0 | a_{\mathbf{p}_2} a_{\mathbf{p}_1}^\dagger a_{\mathbf{q}_2} a_{\mathbf{q}_1}^\dagger |0\rangle + (2\pi)^3 \delta^3(\mathbf{q}_2 - \mathbf{p}_1) \langle 0 | a_{\mathbf{p}_2} a_{\mathbf{q}_1}^\dagger |0\rangle \\ &= \langle 0 | a_{\mathbf{p}_2} a_{\mathbf{p}_1}^\dagger |0\rangle (2\pi)^3 \delta^3(\mathbf{q}_2 - \mathbf{q}_1) + (2\pi)^6 \delta^3(\mathbf{q}_2 - \mathbf{p}_1) \delta^3(\mathbf{p}_2 - \mathbf{q}_1) \\ &= (2\pi)^6 \delta^3(\mathbf{q}_2 - \mathbf{q}_1) \delta^3(\mathbf{p}_2 - \mathbf{p}_1) + (2\pi)^6 \delta^3(\mathbf{q}_2 - \mathbf{p}_1) \delta^3(\mathbf{p}_2 - \mathbf{q}_1) \end{aligned}$$

For proper normalisation, we can look at states of the kind

$$|f\rangle = \int f(\mathbf{p}, \mathbf{q}) |\mathbf{p}, \mathbf{q}\rangle \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6}$$

The smearing function should be symmetric in \mathbf{p} and \mathbf{q} , because if there were an antisymmetric part, the bosonic statistics of the state will lead to a zero under the integral.

Similarly we can construct n particle states, with n going all the way to ∞ , by repeated application of creation operators of the respective momenta.

The Fock Space

With this, we have done all the possible things we could have done on the vacuum, and have generated all the possible states of the theory. Therefore the Hilbert space of the theory is the direct sum

$$\mathbb{H} = \oplus_{n=0}^{\infty} \mathbb{H}_n \quad (3.16)$$

where \mathbb{H}_n is the Hilbert space with n particles. This is called a Fock Space. In this we are summing over spaces which are orthogonal to each other.

Check that the one particle states are orthogonal to two particle states.

$$\langle \mathbf{p}, \mathbf{q} | \mathbf{r} \rangle = \langle 0 | a_{\mathbf{p}} a_{\mathbf{q}} a_{\mathbf{r}}^{\dagger} | 0 \rangle = \langle 0 | a_{\mathbf{p}} a_{\mathbf{r}}^{\dagger} a_{\mathbf{q}} | 0 \rangle + (2\pi)^3 \delta^3(\mathbf{q} - \mathbf{r}) \langle 0 | a_{\mathbf{p}} | 0 \rangle = 0$$

Similarly one can show that the individual spaces are orthogonal to each other.

We can therefore write projectors onto different spaces.

$$P_0 = |0\rangle \langle 0|$$

$$P_1 = \int \frac{d^3\mathbf{p}}{(2\pi)^3} |\mathbf{p}\rangle \langle \mathbf{p}|$$

Why is this a projector?

$$P_1^2 = \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6} |\mathbf{p}\rangle \langle \mathbf{p}| |\mathbf{q}\rangle \langle \mathbf{q}| = \int \frac{d^3\mathbf{p}}{(2\pi)^3} |\mathbf{p}\rangle \langle \mathbf{p}| = P_1$$

The projector on 2-particle states is naively

$$P_2 = \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6} |\mathbf{p}, \mathbf{q}\rangle \langle \mathbf{p}, \mathbf{q}|$$

This is not the complete picture since,

$$P_2 |\mathbf{k}_1, \mathbf{k}_2\rangle = 2 |\mathbf{k}_1, \mathbf{k}_2\rangle$$

(use the normalisation derived above, and the fact that the particles are bosons.)

Therefore, we need a prefactor

$$P_2 = \frac{1}{2!} \int \frac{d^3\mathbf{p} d^3\mathbf{q}}{(2\pi)^6} |\mathbf{p}, \mathbf{q}\rangle \langle \mathbf{p}, \mathbf{q}|$$

and therefore, the projector to n -particle state would be

$$P_n = \frac{1}{n!} \int \frac{d^3\mathbf{p}_1 \cdots d^3\mathbf{p}_n}{(2\pi)^{3n}} |\mathbf{p}_1, \cdots, \mathbf{p}_n\rangle \langle \mathbf{p}_1, \cdots, \mathbf{p}_n|$$

The identity therefore would be

$$\mathbb{I} = \sum_{i=0}^{\infty} P_i$$

4 Detour — Schrodinger Representation

So far we discussed *one* representation of the Hilbert Space. Here we will discuss another representation called the Schrodinger representation. (Hattfeld — QFT of particles and strings, and Polchinski — String theory).

4.1 The Schrodinger Equation

Initially we told that field theory is not different from QM, but so far we have not discussed any wavefunctions or Schrodinger equation governing the evolution of the wavefunction. It is indeed possible to do so, just that it is not a convenient picture to speak about the fields in terms of. Here we discuss the picture a bit.

We started our quantization procedure through the commutation relations

$$[\phi(0, \mathbf{x}), \Pi(0, \mathbf{x}')] = i\delta^3(\mathbf{x} - \mathbf{x}') \quad (4.1)$$

There is also a natural way to also represent states, which is to think of wavefunctions on the different degrees of freedom, i.e. in this case it will become a wavefunctional

$$\psi[\phi(0, \mathbf{x})] \quad (4.2)$$

In normal QM, if you give a value of \mathbf{x} at an instant of time, it spits out a c-number. In the case of QFT, when you give a field configuration in space at an instant of time, the wavefunctional gives a c-number.

An example of such a wavefunction is

$$\psi[\phi] = e^{-\int \phi(0, \mathbf{x})^2 d^3\mathbf{x}} \quad (4.3)$$

The constraint on the possible functionals is that it should be normalisable.

How does the momentum acts on the wavefunction?

Claim — The momentum operator is given by the following operator on the wavefunction

$$\Pi(0, \mathbf{x}) = -i \frac{\delta}{\delta\phi(0, \mathbf{x})} \quad (4.4)$$

To check that this is indeed the form of momentum operator, check

$$\begin{aligned} [\phi(0, \mathbf{x}), \Pi(0, \mathbf{y})]\psi &= -i \left(\phi(0, \mathbf{x}) \frac{\delta}{\delta\phi(0, \mathbf{y})} \psi - \frac{\delta}{\delta\phi(0, \mathbf{y})} (\phi(0, \mathbf{x}) \psi) \right) \\ &= i\delta^3(\mathbf{x} - \mathbf{y})\psi \end{aligned}$$

as required.

We can also check that the momentum operator acts as translation operator on the wavefunctional when we perform a shift on the field configuration. An infinitesimal shift in the field configuration is of the form

$$\phi(0, \mathbf{x}) \rightarrow \phi(0, \mathbf{x}) + \epsilon\varphi(0, \mathbf{x}) \quad (4.5)$$

Under such a transformation, the wavefunctional transforms as

$$\psi[\phi(0, \mathbf{x}) + \epsilon\varphi(0, \mathbf{x})] \approx \psi[\phi(0, \mathbf{x})] + \epsilon\varphi(0, \mathbf{x}) \frac{\delta}{\delta\phi(0, \mathbf{x})} \psi[\phi(0, \mathbf{x})] = e^{i\epsilon\varphi(0, \mathbf{x}) \left(-i \frac{\delta}{\delta\phi(0, \mathbf{x})}\right)} \psi[\phi(0, \mathbf{x})] \quad (4.6)$$

If we require momentum operator to generate these translations, then the momentum operator must be indeed of the form

$$\Pi(0, \mathbf{x}) = -i \frac{\delta}{\delta\phi(0, \mathbf{x})} \quad (4.7)$$

Therefore, the theory has three momentums which one can talk about — the conjugate momentum to \mathbf{x} - \mathbf{p} which generates translations in space, the physical momentum operator \mathbf{P} which generates translations of fields in space, i.e. $\phi(0, \mathbf{x}) \rightarrow \phi(0, \mathbf{x} + \mathbf{a})$, and has as eigenvalues the total momentum of the excitations, and the momentum conjugate to the fields, Π which generates the translations of wavefunctional under shifts in field configurations, i.e. $\psi[\phi] \rightarrow \psi[\phi + \varphi]$.

In a similar language as QM, we can also write wavefunctions in other spaces such as momentum space in configuration space, by making a fourier transformation of $\psi[\phi]$ with respect to Π and converting to $\psi[\Pi]$ given as

$$\psi[\Pi] = \int \mathcal{D}\phi \psi[\phi] e^{-i \int \phi \Pi d^3\mathbf{x}} \quad (4.8)$$

but at this point, such a fourier transform has no use to it in our discussions.

We can also write a Schrodinger equation given as

$$i \frac{\partial}{\partial t} \psi[\phi] = \int -\frac{\delta^2}{\delta\phi^2} \psi[\phi] + (\nabla\phi)^2 \psi[\phi] + m^2 \phi^2 \psi[\phi] d^3\mathbf{x} \quad (4.9)$$

which governs the time evolution of the wavefunctional.

4.2 The Vacuum in Schrodinger Formulation

The defining property of vacuum is that the annihilation operators annihilate the vacuum. It is convenient to introduce the fourier transform

$$\phi(0, \mathbf{p}) = \int \phi(0, \mathbf{x}) e^{-i\mathbf{p}\cdot\mathbf{x}} d^3\mathbf{x} \quad (4.10)$$

$$\Pi(0, \mathbf{p}) = \int \Pi(0, \mathbf{x}) e^{-i\mathbf{p}\cdot\mathbf{x}} d^3\mathbf{x} \quad (4.11)$$

In terms of these, the commutation relation is

$$[\phi(0, \mathbf{p}), \Pi(0, \mathbf{q})] = \int [\phi(0, \mathbf{x}), \Pi(0, \mathbf{y})] e^{-i\mathbf{p}\cdot\mathbf{x}} e^{-i\mathbf{q}\cdot\mathbf{y}} d^3\mathbf{x} d^3\mathbf{y} = i(2\pi)^3 \delta^3(\mathbf{p} + \mathbf{q}) \quad (4.12)$$

The action of $\Pi(0, \mathbf{p})$ on ψ is given by

$$\Pi(0, \mathbf{p})\psi[\phi(0, \mathbf{q})] = -i(2\pi)^3 \frac{\delta}{\delta\phi(0, -\mathbf{p})} \psi[\phi(0, \mathbf{q})] \quad (4.13)$$

To check that this is true, see that

$$[\psi(0, \mathbf{p}), \Pi(0, \mathbf{q})]\psi = -i(2\pi)^3 \left(\phi(0, \mathbf{p}) \frac{\delta}{\delta\phi(0, -\mathbf{q})} \psi - \frac{\delta}{\delta(\phi(0, -\mathbf{q}))} (\phi(0, \mathbf{p})\psi) \right) = i(2\pi)^3 \delta^3(\mathbf{p} - (-\mathbf{q})) \quad (4.14)$$

which is indeed the required relation.

We had

$$a_{\mathbf{p}} = \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \int \phi(0, \mathbf{x}) e^{-i\mathbf{p}\cdot\mathbf{x}} d^3\mathbf{x} + \frac{i}{\sqrt{2\omega_{\mathbf{p}}}} \int \Pi(0, \mathbf{x}) e^{-i\mathbf{p}\cdot\mathbf{x}} d^3\mathbf{x} \quad (4.15)$$

which now becomes

$$a_{\mathbf{p}} = \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \phi(0, \mathbf{p}) + \frac{i}{\sqrt{2\omega_{\mathbf{p}}}} \Pi(0, \mathbf{p}) \quad (4.16)$$

The vacuum is the state that is annihilated by $a_{\mathbf{p}}$ which gives the equation

$$\sqrt{\frac{\omega_{\mathbf{p}}}{2}} \phi(0, \mathbf{p}) \psi_0[\phi] - i(2\pi)^3 \frac{i}{\sqrt{2\omega_{\mathbf{p}}}} \frac{\delta}{\delta\phi(0, -\mathbf{p})} \psi_0[\phi] = 0 \quad (4.17)$$

This is solved by

$$\psi_0 = e^{-\frac{1}{2} \int \omega_{\mathbf{p}} \phi(0, \mathbf{p}) \phi(0, -\mathbf{p}) \frac{d^3\mathbf{p}}{(2\pi)^3}} \quad (4.18)$$

Since $\phi(0, \mathbf{x})$ is real field, we need that $\psi(0, -\mathbf{p}) = \psi(0, \mathbf{p})^*$. Therefore the solution is an exponential of squares, which means it is a Gaussian, as was in the case of harmonic oscillator. Notice that this ground state is not normalised. In such a formalism, the expectation values are to be written as

$$\langle A \rangle = \frac{\int \mathcal{D}\phi \psi[\phi]^2 A[\phi]}{\int \mathcal{D}\phi \psi[\phi]^2} \quad (4.19)$$

and therefore the normalisation doesn't matter as it gets cancelled when taking the ratios.

Similarly we can also write the wavefunctionals for higher excited states by acting upon with $a_{\mathbf{p}}^\dagger$ s.

5 Causality

Before discussing causality, we need a notion of what are the observables in a QFT.

5.1 Observables in QFT

In quantum mechanics all Hermitian operators are observables. When transitioning to QFT, we need to modify our notion of the observables. In QFT, observables are tied to spacetime points.

Observables are of the form $O(x)$ (where x is now a spacetime point). We can define observables in a region R as

$$\int_R O(x) f(x) d^4x \quad (5.1)$$

We can also construct operator like

$$\int_{\bar{R}} O(x) g(x) d^4x \quad (5.2)$$

where \bar{R} is the region complement to R , and this operator is NOT an observable in R .

5.2 Does QFT Preserve Causality

Suppose R and R' are spacelike separated regions. Then causality requires that observations in R should be independent of observations in R' . What this means is that observations in R should not affect, and should not be affected by observations in R' . What this implies is that

$$[O(R), \tilde{O}(R')] = 0 \quad (5.3)$$

This doesn't mean that observation in R should be uncorrelated with observations in R' . They can be correlated as in entanglement, but they should not be able to affect each other. Notice that we have always used this notion in our regular quantum mechanics, where we say that when measuring one observable does not affect the value of another should imply that they have a simultaneous set of eigenvalues and therefore the operators must commute. Here we are running a similar argument.

All observables in our theory are polynomials in ϕ . Therefore causality requires that

$$[\phi(x), \phi(y)] = 0, \text{ if } (x - y)^2 < 0 \quad (5.4)$$

On a single time slice, this is true and is a canonical commutation relation, since every \mathbf{y} is disconnected from \mathbf{x} . But causality requires this to be true also for any x and y at different times, but spacelike separated. (The commutator relation above being satisfied at $t = 0$ doesn't guarantee that it will be satisfied at some other t as can be seen in the example (\diamond))

Is this criteria satisfied in our theory?

We can explicitly calculate this commutator as follows

$$\begin{aligned}
\phi(x) &= \int a_{\mathbf{p}} e^{-ip \cdot x} + a_{\mathbf{p}}^\dagger e^{ip \cdot x} \frac{d^3 \mathbf{p}}{(2\pi)^3 \sqrt{2\omega_{\mathbf{p}}}} \\
\Rightarrow [\phi(x), \phi(y)] &= \int [a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] e^{-ip \cdot x + iq \cdot y} + [a_{\mathbf{q}}^\dagger, a_{\mathbf{p}}] e^{ip \cdot x - iq \cdot y} \frac{d^3 \mathbf{p} d^3 \mathbf{q}}{(2\pi)^6 2\sqrt{\omega_{\mathbf{p}} \omega_{\mathbf{q}}}} \\
&= \int e^{-ip \cdot (x-y)} - e^{ip \cdot (x-y)} \frac{d^3 \mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}}
\end{aligned}$$

The LHS is Lorentz invariant, but on RHS, we see that the integral is over $d^3 \mathbf{p}$. The RHS should still be Lorentz invariant. To see how we can write the integral in a covariant form, consider the following integral first

$$\int d^4 p \delta^4(p^2 - m^2) \theta(p^0) \quad (5.5)$$

Now this integral is manifestly Lorentz invariant, because it has an integral over $d^4 p$, a δ function in $p^2 - m^2$ and a θ function in p^0 . The θ function is also Lorentz invariant since δ function fixes that $p^2 = m^2 > 0$, making p a timelike vector, and there is no Lorentz transformation that can change the sign of the time component of a timelike vector. The *theta* function is also needed since we initially started with a specific sign of ω , that is $\omega_p > 0$. If we choose the other sign, then the result will not be that of the integral we are trying to evaluate.

We can do this integral as

$$\begin{aligned}
&\int d^3 \mathbf{p} dp^0 \delta(p_0^2 - \mathbf{p}^2 - m^2) \delta(p^0) \\
&= \int d^3 \mathbf{p} dp^0 \frac{\delta(p_0 - \omega_p)}{2p_0} \\
&= \int \frac{d^3 \mathbf{p}}{2\omega_p}
\end{aligned}$$

Therefore, the commutator can be written as a Lorentz invariant integral

$$[\phi(x), \phi(y)] = \int e^{-ip \cdot (x-y)} - e^{ip \cdot (x-y)} \frac{1}{(2\pi)^3} d^4 p \delta^4(p^2 - m^2) \theta(p^0) \quad (5.6)$$

When x and y are spacelike separated, there is a Lorentz transformation that takes $x - y$ to $-(x - y)$. An explicit example would be to first boost to a frame where x and y are simultaneous, making $x - y = (0, \mathbf{\Delta})$, and then a rotation of the 3-vector by 180 degree taking it to $(0, -\mathbf{\Delta}) = -(x - y)$. Notice that this is not possible when x and y are timelike separated, since there is no frame of reference where they are simultaneous. Therefore, when x and y are timelike separated, the second integral can be Lorentz transformed to match the first integral, and therefore the difference between them, which gives the commutator, is zero.

We also need to check that the commutator is not identically zero when the separation is timelike. To

check this, we need to explicitly evaluate the integral by converting it into polar coordinates. Consider the integral

$$\begin{aligned}
f(q) &= \int e^{-i\mathbf{p}\cdot\mathbf{q}} \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}} = e^{-i\omega_p q^0} \int e^{i\mathbf{p}\cdot\mathbf{q}} \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}} \\
&= e^{-i\omega_p q^0} \int e^{ipq \cos \theta} p^2 dp d\cos \theta \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}} \text{ where now } p \text{ \& } q \text{ stand for } |\mathbf{p}| \text{ \& } |\mathbf{q}| \\
&= e^{-i\omega_p q^0} \int_0^\infty \frac{1}{iq} (e^{ipq} - e^{-ipq}) p dp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}}
\end{aligned}$$

Since the integral is invariant under the transformation $p \rightarrow -p$, we can do the transformation for the second piece, and then flip the integration limits, to obtain

$$\begin{aligned}
f(q) &= e^{-i\omega_p q^0} \int_0^\infty \frac{1}{iq} e^{ipq} p dp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}} + e^{-i\omega_p q^0} \int_{-\infty}^0 \frac{1}{iq} e^{ipq} p dp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}} \\
&= e^{-i\omega_p q^0} \int_{-\infty}^\infty \frac{1}{iq} e^{ipq} p dp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}} \\
&= \frac{-i}{4\pi q} \frac{\partial}{\partial q} \int_{-\infty}^\infty e^{-i\omega_p q^0 + ipq} \frac{dp}{2\pi\omega_p}
\end{aligned}$$

Now, we can make the substitutions

$$\begin{aligned}
\omega_p &= m \cosh \phi \quad \& \quad p = m \sinh \phi \implies \frac{dp}{\omega_p} = d\phi \\
q^0 &= \alpha \cosh \phi_0 \quad \& \quad q = \alpha \sinh \phi_0
\end{aligned}$$

This substitution also preserves the relations between ω_p & p and q^0 & q . Also we are taking q^0 with cosh because we are looking at the case where q^0 is large compared to q .

With this, we can write the integral as

$$f(q) = \frac{-i}{4\pi q} \frac{\partial}{\partial q} \int_{-\infty}^\infty e^{-im\alpha \cosh(\phi - \phi_0)} \frac{d\phi}{2\pi}$$

This integral has the value

$$f(q) = \frac{1}{4\pi} \frac{\partial}{\partial q} \left(-\frac{i}{2} \text{sgn}(q^0) J_0(m\alpha) - N_0(m\alpha) \right)$$

(there seems to be some errors in the calculation. If you are able to rectify please send your calculations by email!)

The commutator in the case $x - y$ is timelike is therefore

$$[\phi(x), \phi(y)] = \frac{i}{4\pi} \text{sgn}(x^0 - y^0) \frac{m}{\alpha} J_1(m\alpha)$$

We get this as, differentiation of J_0 gives J_1 and when taking the Hermitian conjugate of the above integral

and subtracting it to get the commutator, the term proportional to N_1 cancels. It should not be surprising that the answer is proportional to i since the commutator is an antihermitian quantity and therefore should be purely imaginary.

The commutator is finite in timelike separations and exactly zero for spacelike separations.

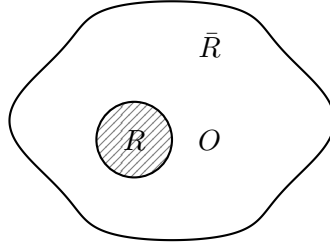
Therefore, causality is preserved in QFT.

5.3 Localized Excitations in QFT

Before discussing localised excitations, let us first discuss what the action of $\phi(x)$ on the vacuum is. We called $a_{\mathbf{p}}^\dagger |0\rangle$ s as the one particle states of our system. The $\phi(x)$ is a linear combination of $a_{\mathbf{p}}^\dagger$ s, with different p s given different weights. Therefore, $\phi(x) |0\rangle$ is a superposition of infinite number of one particle states with different momenta, and therefore is itself a one particle state. To create higher particle states, one needs to operate by more $\phi(x)$ s

Now let us discuss how can one define localised excitations in some region R ?

Suppose we have a region R , and an operator O in \bar{R} which is spacelike to R



For an excitation $|\psi(R)\rangle$ to be local in R , we require that

$$\langle \psi(R) | O(\bar{R}) | \psi(R) \rangle = \langle 0 | O(\bar{R}) | 0 \rangle \quad \forall O \text{ in region spacelike to } R$$

That is, the expectation values of operators in regions spacelike separated from R should be the same as their vacuum expectation values.

How can we create such a state?

Our first guess would be

$$|\psi(R)\rangle = \int_R d^4x f(x) \phi(x) |0\rangle$$

We can check if this is indeed a localised expectation,

$$\langle \psi(R) | O(\bar{R}) | \psi(R) \rangle = \int f(x) f(y) \langle 0 | \phi(x) O \phi(y) | 0 \rangle$$

since ϕ in R and O in \bar{R} must commute, the above should be equal to

$$\langle \psi(R) | O(\bar{R}) | \psi(R) \rangle = \int f(x) f(y) \langle 0 | O \phi(x) \phi(y) | 0 \rangle$$

which is not equal to the vacuum expectation value of O . See that even $\phi(z)$ does not produce a local excitation **at** z , since $\phi(z)$ is obtained by simply setting $f(x) = \delta(x - z)$, $f(y) = \delta(z - y)$ in the above integral.

What do we need to create a localized excitation?

In order for $|\psi(R)\rangle = U(R) |0\rangle$ to be a localised excitation, we need

$$\langle \psi(R) | O(\bar{R}) | \psi(R) \rangle = \langle 0 | U^\dagger(R) O(\bar{R}) U(R) | 0 \rangle = \langle 0 | O(\bar{R}) U^\dagger(R) U(R) | 0 \rangle = \langle 0 | O(\bar{R}) | 0 \rangle$$

What this implies is that the operator U should satisfy $U^\dagger U = I$, i.e. it should be unitary. Therefore

$$U(R) = e^{i \int_R f(x) \phi(x) dx} |0\rangle$$

is a localised excitation in R . There is no unique localised excitation, there can be many unitary operators in R that can produce localised excitations. One can also see that since ϕ are Hermitian, the only way to create unitary operators is by exponentiating the ϕ s with a factor of i . If you produce an excitation in a region of space \mathbf{R} , it propagates outwards at the speed of light, such that all the region which is timelike to \mathbf{R} is affected, while regions spacelike to \mathbf{R} are not at all affected.

Now these localised excitations can never be single particle states. Since the exponential can be written as

$$e^{i \int_R f(x) \phi(x) dx} |0\rangle = \left(1 + i \int_R f(x) \phi(x) dx + \frac{1}{2} i \iint_R f(x) f(y) \phi(x) \phi(y) dx dy + \dots \right) |0\rangle$$

the localised excitation is always a superposition of states of all possible particle numbers.

6 Interlude — Symmetries

What we mean when we say that the action has symmetry is that when we make a change in the fields $\phi_i(x) \rightarrow \phi_i(x) + \lambda D\phi_i(x)$, (where $D\phi_i$ is just a notation and can denote any function, but is usually a function of original ϕ_i s and λ is a parameter to keep track of order of terms), the action remains invariant. For this to be true, the Lagrangian density should transform as

$$\mathcal{L} \rightarrow \mathcal{L} + \partial_\mu K^\mu$$

Now this invariance should hold for all the field configurations, and not just some special ones, i.e. the configurations which satisfy the equations of motion. For the configurations which satisfy the equations of motion, the requirement is identically satisfied, since we define the equations of motion as those configurations for which the small variations of fields do not change the action.

6.1 Noether's Theorem

When \mathcal{L} has a symmetry, there is a conserved current of the form

$$\partial_\mu j^\mu = 0$$

only for the solutions of the equations of motion.

Proof —

Under the above transformation,

$$\begin{aligned} \delta\mathcal{L} &= \left(\frac{\delta\mathcal{L}}{\delta\phi_i} D\phi_i + \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i} \partial_\mu D\phi_i \right) = \partial_\mu K^\mu \\ \implies \left(\frac{\delta\mathcal{L}}{\delta\phi_i} D\phi_i - \partial_\mu \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i} D\phi_i + \partial_\mu \left(\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i} D\phi_i \right) \right) &= \partial_\mu K^\mu \end{aligned}$$

When we are on-shell (i.e. when ϕ satisfies the equation of motion), the first two term sum to 0, and we are left with

$$\partial_\mu \left(\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i} D\phi_i \right) = \partial_\mu K^\mu$$

which implies that

$$\partial_\mu j^\mu(x) = 0, \quad \text{where } j^\mu = \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i} D\phi_i - K^\mu$$

Noether's current is not something new, it is just re-expressing the equations of motion. Noether's current has inherent ambiguity in definition, because we can always add a divergence free 4-vector to it, i.e. we can add terms of the form $\epsilon^{\mu\nu} \partial_\nu \phi$ where ϵ is antisymmetric tensor, since $\partial_\mu \epsilon^{\mu\nu} \partial_\nu \phi = 0$. But nevertheless this is a powerful statement and allows us to derive conclusions without having to do tedious calculations.

When we have a conserved current, we have a conserved charge.

$$\partial_0 \int j^0 d^3 \mathbf{x} + \int d^3 \mathbf{x} \partial_i j^i = 0 \implies \partial_t \int j^0 d^3 \mathbf{x} = 0$$

since the second term is a total divergence. What this tells is that the quantity

$$Q = \int j^0 d^3 \mathbf{x}$$

is conserved. All conservation of laws that we have known so far arise in this very fashion, and are consequences of the Noether's theorem.

There is another derivation of the Noether's theorem, which is at the level of action. Suppose we have a symmetry under

$$\phi \rightarrow \phi + \lambda D\phi$$

which does not change S .

Now imagine a spacetime dependent variation,

$$\phi \rightarrow \phi + \lambda(x) D\phi$$

The action is not invariant under this transformation. The Lagrangian, upto first order in λ changes as

$$\mathcal{L} \rightarrow \mathcal{L} + \lambda(x) \partial_\mu (K^\mu) + \partial_\mu (\lambda(x)) P^\mu$$

since these are the only terms that are Lorentz invariant and first order in λ .

The action transforms as

$$\begin{aligned} S &\rightarrow S + \int d^4 x \lambda(x) \partial_\mu (K^\mu) + \partial_\mu (\lambda(x)) P^\mu \\ &= S + \int d^4 x \lambda(x) \partial_\mu (K^\mu - P^\mu) \end{aligned}$$

On solutions to equations of motions, the action should be invariant under these transformations too. Therefore

$$\partial_\mu (K^\mu - P^\mu) = 0$$

Now what is P^μ ?

Under the given transformation, the Lagrangian changes as

$$\mathcal{L} \rightarrow \mathcal{L} + \left(\frac{\delta \mathcal{L}}{\delta \phi} \lambda(x) D\phi + \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} \partial_\mu (\lambda(x) D\phi) \right)$$

from which we can see that P^μ is simply the part that multiplies to $\partial_\mu \lambda(x)$, which is

$$P^\mu = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} D\phi$$

6.2 Applications

Let us look at some applications of the Noether's theorem

Spacetime translations

Consider $x^\mu \rightarrow x^\mu + a^\mu$.

When \mathcal{L} doesn't explicitly depend on position, this is a symmetry. In the language of fields, this transformation is

$$\phi \rightarrow \phi + \lambda \partial_\mu \phi^\mu$$

where λ is just to keep track of orders.

Since \mathcal{L} is a Lorentz scalar,

$$\mathcal{L} \rightarrow \mathcal{L} + \lambda a^\mu \partial_\mu \mathcal{L} \implies K^\mu = a^\mu \mathcal{L}$$

Therefore, the conserved current is

$$j^\mu = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} a^\nu \partial_\nu \phi - a^\nu \delta_\nu^\mu \mathcal{L}$$

Now this current is conserved for all possible a^μ which means that

$$T_\nu^\mu = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} \partial_\nu \phi - \delta_\nu^\mu \mathcal{L} \implies T^{\mu\nu} = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L}$$

must be conserved. This is called stress-energy tensor and this is conserved. Let us check what this is for few cases

$$T^{00} = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} \dot{\phi} - \mathcal{L} = \mathcal{H}$$

which is the Hamiltonian density. Integrating this over all space gives the Hamiltonian which is the energy, and this shows that energy is conserved.

$$T^{0i} = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} \partial^i \phi$$

the corresponding conserved charge is

$$P^i = \int d^3 \mathbf{x} \frac{\delta \mathcal{L}}{\delta \dot{\phi}} \partial^i \phi$$

which is nothing but the spatial momentum in i th direction as we had defined previously. The stress tensor therefore gives 4 conserved charges (we need to keep $\mu = 0$ and integrate, and there are 4 choice of ν) which are energy and the three components of momentum.

The formula for stress-tensor is not manifestly symmetric in μ and ν , but we can use the ambiguity that

we discussed to subtract the antisymmetric part and make it symmetric. But in most cases of physical interest, stress-tensor is explicitly symmetric.

Lorentz transformation

Lorentz transformation

$$\tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu$$

The requirement is that the transformation preserves vector inner products

$$\tilde{x}^\mu \tilde{y}^\nu \eta_{\mu\nu} = x^\mu y^\nu \eta_{\mu\nu} \implies \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \eta_{\rho\sigma} = \eta_{\mu\nu}$$

An infinitesimal transformation of x would be

$$x^\mu \rightarrow x^\mu + \lambda \epsilon^\mu{}_\nu x^\nu$$

For this to be a Lorentz transformation

$$\begin{aligned} (x^\mu + \lambda \epsilon^\mu{}_\nu x^\nu)(y_\mu + \lambda \epsilon_{\mu\nu} y^\nu) &= x^\mu y_\mu \\ \implies x^\mu y_\mu + \lambda \epsilon^\mu{}_\nu x^\nu y_\mu + \lambda \epsilon_{\mu\nu} x^\mu y^\nu + O(\lambda^2) &= x^\mu y_\mu \\ \implies \epsilon_{\mu\nu} x^\nu y^\mu + \epsilon_{\mu\nu} x^\mu y^\nu &= 0 \end{aligned}$$

Now since in the first term, μ and ν are dummy indices, we can interchange them to get

$$(\epsilon_{\nu\mu} + \epsilon_{\mu\nu}) x^\mu y^\nu = 0 \implies \epsilon_{\nu\mu} = -\epsilon_{\mu\nu}$$

Which implies that the $\epsilon_{\mu\nu}$ is an antisymmetric tensor. (Notice that this would not work if we compare tensors with one upper and one upper index). There are 6 components to the antisymmetric tensor, and therefore 6 conserved currents — three for boosts and three for rotations. As an example consider $\epsilon_{01} = 1$, $\epsilon_{10} = -1$ rest all 0. Under this

$$\delta x^0 = \lambda x^1, \quad \delta x^1 = \lambda x^0$$

which are infinitesimal boosts. If $\epsilon_{12} = 1$, $\epsilon_{21} = -1$, then

$$\delta x^1 = -\lambda x^2, \quad \delta x^2 = \lambda x^1$$

This is an infinitesimal rotation.

Under Lorentz transformation, the fields transform as

$$\delta x^\mu = \lambda \epsilon^\mu{}_\nu x^\nu \implies \delta \phi = \lambda \partial_\mu \phi \epsilon^\mu{}_\nu x^\nu$$

Under this kind of transformation, the lagrangian density also transforms as

$$\delta\mathcal{L} = \lambda(\partial_\mu\mathcal{L})\epsilon^\mu{}_\nu x^\nu$$

Now since $\partial_\mu(\epsilon^\mu{}_\nu x^\nu) = \delta^\nu{}_\mu \epsilon^\mu{}_\nu = \eta^{\mu\nu}\epsilon_{\mu\nu} = 0$ (\because η is symmetric and ϵ is antisymmetric) (another way to look at this is that $\delta^\nu{}_\mu \epsilon^\mu{}_\nu = \text{Tr}\epsilon = 0$) we can add this term to make the change in lagrangian density a total derivative

$$\delta\mathcal{L} = \lambda\partial_\mu(\mathcal{L}\epsilon^\mu{}_\nu x^\nu)$$

To proceed, let us consider an explicit example of a theory, the lagrangian density we have already considered. For this theory,

$$\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi} = \partial^\mu\phi$$

and therefore,

$$j^\mu = \partial^\mu\phi\partial_\rho\phi\epsilon^\rho{}_\nu x^\nu - \mathcal{L}\epsilon^\mu{}_\nu x^\nu$$

Since this should be conserved for all ϵ , we need to pull it out of the equation as

$$j^\mu = (\partial^\mu\phi\partial^\rho\phi x^\nu - \mathcal{L}\eta^{\rho\mu}x^\nu)\epsilon_{\rho\nu}$$

The conserved current is therefore the antisymmetric part of

$$\partial^\mu\phi\partial^\rho\phi x^\nu - \mathcal{L}\eta^{\rho\mu}x^\nu$$

since the derivative of the symmetric part is identically zero, owing to the presence of ϵ . Therefore the conserved current is

$$M^{\mu\nu\rho} = (\partial^\mu\phi\partial^\rho\phi x^\nu - \mathcal{L}\eta^{\rho\mu}x^\nu) - (\rho \leftrightarrow \nu)$$

There are 6 possible choices of ρ and ν for the antisymmetric part, and therefore 6 conserved currents.

The above quantity is exactly equal to

$$M^{\mu\nu\rho} = x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu}$$

where T is the stress tensor. The six conserved charges are

$$Q^{\nu\rho} = \int x^\nu T^{0\rho} - x^\rho T^{0\nu} d^3\mathbf{x}$$

The different cases —

$$Q^{ij} = \int x^i T^{0j} - x^j T^{0i} d^3\mathbf{x}$$

This is angular momentum associated with the plane of rotation ij .

$$Q^{0i} = \int x^0 T^{0i} - x^i T^{00} d^3 \mathbf{x}$$

Let us write it in a slightly different way. $\int T^{0i} d^3 \mathbf{x} = P^i$, meaning

$$Q^{0i} = t P^i - \int x^i T^{00} d^3 \mathbf{x}$$

Differentiating with respect to time (and using momentum is conserved),

$$P^i - \frac{d}{dt} \int x^i T^{tt} d^3 \mathbf{x} = 0 \implies P^i - \left(\int T^{tt} d^3 \mathbf{x} \right) \times \frac{d}{dt} \frac{\int x^i T^{tt} d^3 \mathbf{x}}{\int T^{tt} d^3 \mathbf{x}}$$

Since T^{tt} is simply the mass density, the integral in the brackets is simply the total mass, and the object getting differentiated is the position of the center of mass. That is

$$P^i = M \frac{d}{dt} x_{COM}^i$$

which is just the classical definition of momentum.

Internal Symmetries

So far we discussed transformations in ϕ due to changes to spacetime. Now we will discuss something different. Consider a theory with two fields,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1)^2 + \frac{1}{2}(\partial_\mu \phi_2)^2 - \frac{m^2}{2}\phi_1^2 - \frac{m^2}{2}\phi_2^2$$

This theory has a continuous symmetry under rotations of ϕ_1 into ϕ_2 , i.e.

$$\phi_1 \rightarrow \cos \theta \phi_1 - \sin \theta \phi_2, \quad \phi_2 \rightarrow \sin \theta \phi_1 + \cos \theta \phi_2$$

This transformation leaves the lagrangian density unchanged.

infinitesimally, the above transformation looks like

$$\delta \phi_1 = -\phi_2, \quad \delta \phi_2 = \phi_1$$

The associated current is

$$j^\mu = \phi_1 \partial^\mu \phi_2 - \phi_2 \partial^\mu \phi_1$$

The associated charge is

$$Q = \int \phi_1 \partial^0 \phi_2 - \phi_2 \partial^0 \phi_1 d^3 \mathbf{x}$$

Charges like electric charge etc. are of this form.

These fields are real fields. However we can combine these two into one complex object as

$$\chi = (\phi_1 + i\phi_2)\frac{1}{\sqrt{2}}, \quad \chi^* = (\phi_1 - i\phi_2)\frac{1}{\sqrt{2}}$$

With this,

$$\frac{1}{2}(\phi_1^2 + \phi_2^2) = |\chi|^2$$

and

$$\frac{1}{2}((\partial_\mu \phi_1)^2 + (\partial_\mu \phi_2)^2) = |\partial_\mu \chi|^2$$

Therefore, the Lagrangian density becomes

$$\mathcal{L} = |\partial_\mu \chi|^2 - m^2 |\chi|^2$$

In this language, the symmetry is

$$\chi \rightarrow e^{i\alpha} \chi, \quad \chi^* \rightarrow e^{-i\alpha} \chi^*$$

The corresponding current is

$$j^\mu = i((\partial_\mu \chi^*)\chi - (\partial_\mu \chi)\chi^*)$$

Since ϕ_1 and ϕ_2 transformed into each other, they are not the fields corresponding to particles of positive and negative charges. χ and χ^* are the fields that correspond to particles of positive and negative charges. This is an example of $U(1)$ internal symmetry. We can have theories with more elaborate symmetries like $SU(N)$, $Sp(N)$ etc.

7 Interactions

Free field theory — theory whose Lagrangian density has only terms that are quadratic in fields. These are exactly solvable, as we did previously. Interacting theory — introducing any other in the Lagrangian will lead to interactions between the different fields or with itself.

So far, we discussed free field theories, and we want to extend our discussions interacting theories. Before that, let us discuss what kind of questions can we ask in the presence of interactions.

7.1 The S -matrix

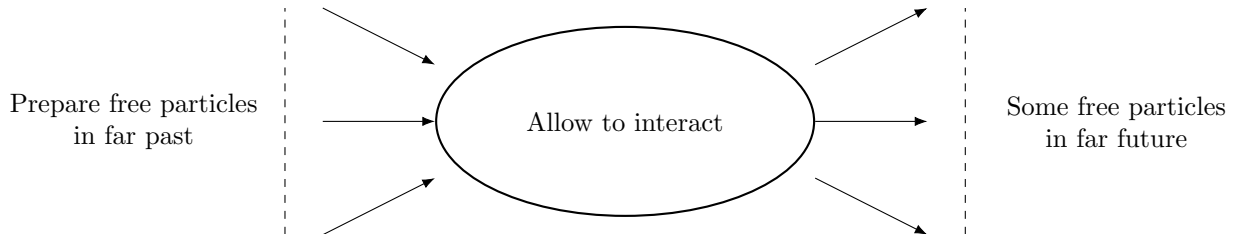
The most general kind of question would be



Such questions are answered by computing the correlators of the form — $\langle 0|\phi(x_1)\cdots\phi(x_n)|0\rangle$.

However, it turns out that in QFT, although we can ask these questions, a lot of focus is on slightly *simpler* questions.

The simpler question is



That is, we create free particles in far past, let them evolve and interact, and ask what are the free particles and their properties in far future. This kind of question is answered by something called the S matrix. One might think that this is less general than the other question, but in fact it turns out that it has all the information we want to have about the theory.

Why we talk about S matrix?

1. Experiments — most physics we learn from experiments is from this kind of experiments.
2. Theoretical reasons — two theoretical reasons
 - S matrix is an unambiguously decided observable. That is if we ask the question like what is the correlator

$$\langle\psi|\phi(x_1)\cdots\phi(x_n)|\psi\rangle$$

this is very ambiguous. This is because in interacting theory, there is no unambiguous *definition* of field. One can always redefine the fields at the cost of adding interaction terms, without altering the physical content of the theory. When the fields are free and don't interact, there is 'a' free field, and field redefinitions do not obey the equations of motions whereas the original field obeys. But in case of interacting theories, no fields obey the equation of motion, and therefore there is no more unique field.

There is a class of field theories called conformal field theories, where there are other principles that allows one to uniquely identify field variables even in the case of interactions, and there we can talk about the first type of questions. But in general S matrix is unambiguously defined since all we talk about is free theories in far past and future.

- In Quantum Gravity, maybe only the S matrix makes sense. The reason for that is that the first type of questions involve some local questions about what happens in some local regions of spacetime, and when the metric itself is fluctuating and spacetime itself is fluctuating, such questions are not well defined. But when we ask the second type of questions, in far past and future, where such fluctuations are not present, it is perfectly well defined.

Therefore there are good reasons to consider the S matrix. Now we develop the machinery needed to calculate these S matrices.

7.2 The Interaction Hamiltonian

For an interacting theory, the Hamiltonian looks like

$$H = H_0 + H_{\text{int}}$$

where H_0 is the free part, and H_{int} is the interacting part.

To make the intuition more precise, one can also consider adding a *turning on-off function* $f(t)$ such that

$$H = H_0 + f(t)H_{\text{int}}$$

and $f(t)$ is such that it starts at 0 at $-\infty$ and it rapidly rises to 1, stays 1 for a long time, and then dies off rapidly to 0 again at ∞ . The purpose of this function is to make the Hamiltonian free in far past and far future, while keeping the interactions turned on in between. It is often not necessary to think of this turning on-off function if we talk about the right initial and final states, i.e. if we discuss about particles that separate sufficiently as we go in the far future/past, the *turning off* of interactions automatically happens.

If we call fields in the far past ϕ_{in} and fields in far future ϕ_{out} , we can take these fields and act them on the vacuum, we can create free particles in far past and far future. We get two Fock spaces, corresponding

to ϕ_{in} and ϕ_{out}

$$\begin{aligned}\mathbb{H}_{in} &= \text{span}\{|0\rangle, |\mathbf{k}\rangle, |\mathbf{k}_1, \mathbf{k}_2\rangle, \dots\} \\ \mathbb{H}_{out} &= \text{span}\{|0\rangle, |\mathbf{k}\rangle, |\tilde{\mathbf{k}}_1, \tilde{\mathbf{k}}_2\rangle, \dots\}\end{aligned}$$

The vacuum in both the spaces are the same as they both are the vacuum of the same H_0 . The one particle states are also the same, as required by the conservation of momentum. However, the higher particle states are not necessarily the same. That is, if we started with a single particle in the far past with momentum \mathbf{k} , it is necessary that in the far future the momentum is still \mathbf{k} . But this is not the case in case of multi-particle states, where a state with momentum $(\mathbf{k}_1, \mathbf{k}_2, \dots)$ can transform into the state $(\tilde{\mathbf{k}}_1, \tilde{\mathbf{k}}_2, \dots)$ as long as the energy-momentum is conserved.

We can immediately see that

$$\begin{aligned}_{in} \langle 0|0\rangle_{out} &= 1 \\ _{in} \langle \mathbf{k}|\tilde{\mathbf{k}}\rangle_{out} &= (2\pi)^3 \delta(\mathbf{k} - \tilde{\mathbf{k}})\end{aligned}$$

and there the non-trivial matrix elements are

$$_{in} \langle \mathbf{k}_1, \mathbf{k}_2, \dots |\tilde{\mathbf{k}}_1, \tilde{\mathbf{k}}_2, \dots\rangle_{out}$$

where the number of momenta in the in states need not be equal to the number of momenta in the out space. That is, we can have a one particle state going to a two particle state and so on too.

These two \mathbb{H} spaces should be related by unitary transformation, since we have one description that describes all degrees of freedom in far past and another description for the far future, it is necessary, for the probabilities to be conserved, that they both are related by a unitary transformation. If we tell all the matrix elements we discussed above, we know what the unitary transformation is. And it is this matrix that is called the S -matrix. The S -matrix is a particular limit of the correlation functions $\langle 0|\phi(x_1) \cdots \phi(x_n)|0\rangle$, with some t taken to $-\infty$ and some to $+\infty$, and this limit is given by the LSZ formula.

To compute the S -matrix, we need the time evolution operator, $U(-\infty, \infty)$. Obtaining this operator exactly is not possible in most interacting field theories, and therefore we need to resort to a perturbation theory. One tool we use to calculate this is called the interaction picture.

7.3 The Interaction Picture

In the Schrodinger picture, the observables do not evolve, but the states evolve. That is

$$\begin{aligned}|\psi\rangle &\rightarrow e^{-iHt} |\psi\rangle \\ O &\rightarrow O\end{aligned}$$

In QFT, the observables are the fields, and therefore, we fix a time slice, lets say $t = 0$ and call all the fields $\phi(0, \mathbf{x})$ in this time slice as observables.

In the Heisenberg picture, which is little more natural in field theory, and therefore is the picture in which we had our discussions so far, the states do not evolve, but the operators do.

$$\begin{aligned} |\psi\rangle &\rightarrow |\psi\rangle \\ O &\rightarrow e^{iHt} O e^{-iHt} \end{aligned}$$

In this description, as we have already seen, the operators are $\phi(t, \mathbf{x}) \equiv \phi(x)$. This is more natural because we are concerned with observations in different times, and for someone at a later time t , it doesn't make sense to discuss observations in terms of the operators $\phi(0, \mathbf{x})$.

In the interaction picture, we solve exactly for the free part of the Hamiltonian and perturbatively for the interacting part. In this picture, we introduce a new operator

$$O_I(t) = e^{iH_0 t} O(0) e^{-iH_0 t}$$

which evolves according to the free Hamiltonian.

To find its relation to the Heisenberg picture operator, we first undo the evolution by free Hamiltonian and then evolve it according to the full Hamiltonian, i.e.

$$O_H(t) = \underbrace{e^{iHt} e^{-iH_0 t}}_{U_I^\dagger(t)} \overbrace{e^{iH_0 t} O(0) e^{-iH_0 t}}^{O_I(t)} \underbrace{e^{iH_0 t} e^{-iHt}}_{U_I(t)}$$

$U_I(t)$ is a unitary operator, but is different from the regular time evolution operator. To see why this is useful, see that the equation of motion it follows is

$$\frac{d}{dt} U_I(t) = e^{iH_0 t} (iH_0) e^{-iHt} + e^{iH_0 t} (-iH) e^{-iHt}$$

(notice that H and H_0 do not necessarily commute, also H_0 and H_{int} are time independent. It is $(H_0)_I$ and $(H_{\text{int}})_I$ that are time dependent).

This is equal to

$$\begin{aligned} \frac{d}{dt} U_I(t) &= -ie^{iH_0 t} (H - H_0) e^{-iHt} \\ &= -ie^{iH_0 t} (H - H_0) e^{-iH_0 t} e^{iH_0 t} e^{-iHt} \\ &= -i(H_{\text{int}})_I(t) U_I(t) \end{aligned}$$

This equation is similar to the one satisfied by the regular time evolution operator, but with H_{int} instead of the full Hamiltonian.

To see why this is a very important picture, let us consider adding a term $\lambda\phi^4(x)$. When we make the split between H_0 and H_{int} , we need to choose a specific time, since these terms can mix with time when evolved under the full Hamiltonian. Let us therefore consider the splitting at time $t = 0$

$$H_{\text{int}} = \int d^3\mathbf{x} \lambda\phi(0, \mathbf{x})^4$$

For this interaction,

$$i\frac{d}{dt}U_I(t) = \lambda e^{iH_0t} \int d^3\mathbf{x} \lambda\phi(0, \mathbf{x})^4 e^{-iH_0t} U_I(t)$$

But

$$e^{iH_0t} \int d^3\mathbf{x} \lambda\phi(0, \mathbf{x})^4 e^{-iH_0t} = \int d^3\mathbf{x} \lambda\phi(t, \mathbf{x})^4$$

where now the fields evolve according to the free Hamiltonian, and we already know the explicit expression for the interaction picture fields at time t . That is, this equation converts everything in terms of interaction picture at time t .

But there is a problem, the equation of motion for the actual time evolution operator depends on H which is the full Hamiltonian, and is therefore a time independent quantity. But in the above equation $(H_{\text{int}})_I(t)$ is also a time dependent quantity, and therefore we need to treat it carefully.

7.4 The Time Evolution Operator

In the case of the complete Hamiltonian, the actual time evolution operator also follows

$$\frac{d}{dt}U(t) = -iHU(t)$$

In this case, the H is time independent, and the equation has a solution

$$U(t) = e^{-i\int_0^t H dt'} = e^{-iHt} \because H \text{ is time independent.}$$

Let us try to solve the previously discussed equation by (wrongly) assuming that the solution in this case too is an exponential, and seeing where it goes wrong. Let us guess (from now on we simply write $H_I(t)$ for $(H_{\text{int}})_I(t)$)

$$U^g(t) = e^{-i\int_0^t H_I(t') dt'}$$

where the superscript g stands for guess. Expanding this in Taylor series,

$$U^g(t) = 1 - i \int_0^t H_I(t') dt' - \frac{1}{2} \int_0^t H_I(t') dt' \int_0^t H_I(t'') dt'' + \dots$$

We can now check if this satisfies the above equation by taking the time derivative

$$\frac{d}{dt}U^g(t) = -iH_I(t) - \frac{1}{2}H_I(t) \left(\int_0^t H_I(t')dt' \right) - \frac{1}{2} \left(\int_0^t H_I(t')dt' \right) H_I(t) + \dots$$

Now if $H_I(t)$ commuted with $H_I(t')$, we could have moved the $H_I(t)$ to the left in the third term and add it to the second term to form $H_I(t) \int_0^t H_I(t')dt'$, and this expression would be the right expression for U_I . But since they do not commute, we cannot do such a manipulation. We face similar problems for all higher order powers, and therefore our guess for the time evolution operator is not the right one.

We can fix this by considering the following expression

$$\begin{aligned} U_I(t) = 1 &+ (-i) \int_0^t H_I(t')dt' + (-i)^2 \int_0^t H_I(t') \int_0^{t'} H_I(t'')dt''dt' \\ &+ (-i)^3 \int_0^t H_I(t') \int_0^{t'} H_I(t'') \int_0^{t''} H_I(t''')dt'''dt''dt' \dots \end{aligned}$$

(The individual integral limits are now changed, they are not all t anymore. Also there are no $\frac{1}{n!}$ accompanying each terms).

When differentiating this, we get

$$\begin{aligned} \frac{d}{dt}U_I(t) &= -iH_I(t) + (-i)^2 H_I(t) \int_0^t H(t'')dt'' + (-i)^3 H_I(t) \int_0^t H_I(t'') \int_0^{t''} H_I(t''')dt''' \dots \\ &= -iH_I(t)U_I(t) \end{aligned}$$

which is the required equation.

We can rewrite this in a slightly compact way by introducing *time ordering*. The above equation has the property that operators later in time are on the left, i.e. if $t' > t''$, then in the expansion $H(t')$ will always be to the left of $H(t'')$.

To see this explicitly, consider a discretized version of the double integral of the above form

$$\sum_0^t H(a) \sum_0^a H(b)$$

Upon expanding the series we get

$$H(0)H(0) + H(1)[H(0) + H(1)] + H(2)[H(0) + H(1) + H(2)] + \dots + H(t)[H(0) + \dots + H(t)]$$

See that at no point in the entire expansion we had a term of the form $H(a)H(b)$ with $a < b$.

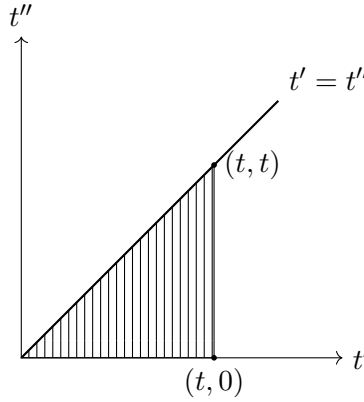
It is convenient to introduce the notion of time ordering which is defined as

$$T(A(t_1, \mathbf{x})B(t_2, \mathbf{y})) = \begin{cases} A(t_1, \mathbf{x})B(t_2, \mathbf{y}) & : t_1 > t_2 \\ B(t_2, \mathbf{y})A(t_1, \mathbf{x}) & : t_2 > t_1 \end{cases}$$

Using this, we can write the above expression as

$$U_I(t) = 1 - i \int_0^t T(H(t'))dt' + (-i)^2 \int_0^t \int_0^{t'} T(H(t')H(t''))dt''dt' + \dots$$

This is the expression we already had above, it was already time ordered, we just made it explicit here, without changing anything. We can see that (considering specifically the double integral), the integral runs dt' from 0 to t , and t'' from 0 to t' . That is, we are integrating over the region



We can also have the t'' integration to go from 0 to t , since we are explicitly enforcing time ordering, we will get the same expression but twice. (Think carefully about this, if there were no time ordering, this manipulation would not be possible, since for the upper triangle, we would be having the products wrongly ordered. The presence of time ordering allows one to extend the integral without altering its value.) Similarly, we can extend all the integrals to t , while being reassured by the time ordering that we are getting consistent results, while dividing by $n!$ which removes the factor of adding over all the possible arrangements of the n dummy variables. Therefore, we can write the above also as

$$U_I(t) = 1 - i \int_0^t T(H(t'))dt' + (-i)^2 \frac{1}{2!} \int_0^t \int_0^t T(H(t')H(t''))dt''dt' + \dots$$

Let us again explicitly see this in the case of the discrete sum which now would be rewritten as

$$\sum_0^t \sum_0^t T(H(a)H(b))$$

This would be equal to

$$\sum_{a>b} H(a)H(b) + \sum_{a<b} H(b)H(a) + \sum_{a=b} H(a)^2$$

Since the first two sums are equal, we get

$$2 \sum_{a>b} H(a)H(b) + \sum_{a=b} H(a)^2$$

In continuum, the diagonal term has zero measure and drops out, and therefore, we will be left with twice the intended integral as expected.

Therefore

$$U_I(t) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \left(\int_0^t \right)^n T(H(t_1)H(t_2) \cdots H(t_n)) dt_1 dt_2 \cdots dt_n$$

and there is a formal way to write this, called the Dyson's formula

$$U_I(t) = T \left\{ e^{-i \int_0^t H_I(t') dt'} \right\}$$

What this means is that we expand the exponential in its Taylor series and then apply the time ordering to every term.

Given an operator in terms of the free fields $\phi_I(t, \mathbf{x})$, the above derived $U_I(t)$ gives us a way to obtain the full Heisenberg picture operators. An inherent advantage is that this formula is already in the form of a power series in H_I , and therefore naturally provides a perturbative series.

8 Wick's Theorem

In general, what we want to compute is what the matrix elements of the above derived $U_I(t)$ are, since this gives us the amplitudes for some in states going to some out states. In computing these matrix elements, we have to compute the matrix elements of the kind

$$\langle \psi_1 | T[H_I(t_1), \dots, H_I(t_n)] | \psi_2 \rangle$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are the in and out states (we will later formulate these only in terms of vacuum expectation values, but it is more physically relevant to talk about these matrix elements). Once we know all these elements, we can sum them according to the Dyson's formula to get the matrix element of U_I which gives us the S -matrix.

The $H_I(t)$ is some polynomial in $\phi_I(t, \mathbf{x})$, and therefore, the above expectation values reduce to calculating elements of the form

$$\langle \psi_1 | T[\phi_I(x_1), \dots, \phi_I(x_n)] | \psi_2 \rangle$$

Remember that the interaction picture fields are simply

$$\phi_I(t, \mathbf{x}) = \int a_{\mathbf{p}} e^{-ip \cdot x} + a_{\mathbf{p}}^\dagger e^{ip \cdot x} \frac{d^3 \mathbf{p}}{(2\pi)^3 \sqrt{2\omega_{\mathbf{p}}}}$$

and the above matrix values are simply matrix elements of the time ordered products of the fields of this kind. But see that the time ordered product is going to be a mess, to see why consider a simple case of three field insertions

$$\begin{aligned} T[\phi(x_1)\phi(x_2)\phi(x_3)] &= \theta(x_1^0 - x_2^0)\theta(x_1^0 - x_3^0)\theta(x_2^0 - x_3^0)\phi(x_1)\phi(x_2)\phi(x_3) \\ &\quad + \theta(x_2^0 - x_1^0)\theta(x_2^0 - x_3^0)\theta(x_1^0 - x_3^0)\phi(x_2)\phi(x_1)\phi(x_3) \\ &\quad + 4 \text{ other terms} \end{aligned}$$

Each of the term above has further 2^3 terms since each field insertion has a creation and annihilation operator, and therefore the product of n such field insertions has 2^n terms. This expression only gets more complicated for polynomials with more field insertions, which is natural in such interaction Hamiltonians, and therefore the entire matrix element is a mess. The tool that is used to simplify this mess is called the Wick's theorem.

8.1 Contractions

Earlier, we had encountered the normal ordering, which was simply an instruction to place all the annihilation operators on the right. We can extend this definition to the interaction picture field operators by

expanding them in terms of creation and annihilation operators and normal order it. That is,

$$\phi_I(x) = \phi_I^+(x) + \phi_I^-(x)$$

where

$$\phi_I^+(x) = \int a_{\mathbf{p}} e^{-ip \cdot x} \frac{d^3 \mathbf{p}}{(2\pi)^3 \sqrt{2\omega_{\mathbf{p}}}} \quad \& \quad \phi_I^-(x) = \int a_{\mathbf{p}}^\dagger e^{ip \cdot x} \frac{d^3 \mathbf{p}}{(2\pi)^3 \sqrt{2\omega_{\mathbf{p}}}}$$

(which is taken by convention — positive frequency contains the annihilation operators and negative frequency contains the creation operators) then,

$$\begin{aligned} :\phi_I(x_1)\phi_I(x_2): &= :\phi_I^+(x_1)\phi_I^+(x_2): + :\phi_I^+(x_1)\phi_I^-(x_2): + :\phi_I^-(x_1)\phi_I^+(x_2): + :\phi_I^-(x_1)\phi_I^-(x_2): \\ &= \phi_I^+(x_1)\phi_I^+(x_2) + \phi_I^-(x_2)\phi_I^+(x_1) + \phi_I^-(x_1)\phi_I^+(x_2) + \phi_I^-(x_1)\phi_I^-(x_2) \end{aligned}$$

(where only the second term is different from the regular product) Notice that normal ordered product is NOT defined for operators in Heisenberg picture since in Heisenberg picture in the presence of interactions, there is no splitting of the field into creation and annihilation operators. This ordering is only valid for interaction picture fields

We define a contraction as

$$\overline{\phi_I(x_1)\phi_I(x_2)} = T[\phi_I(x_1)\phi_I(x_2)] - :\phi_I(x_1)\phi_I(x_2):$$

The normal ordered product and time ordered product are operators, but contraction is just a number. We know that

$$T[\phi_I(x_1)\phi_I(x_2)] = \theta(x_1^0 - x_2^0)\phi_I(x_1)\phi_I(x_2) + \theta(x_2^0 - x_1^0)\phi_I(x_2)\phi_I(x_1)$$

and the normal ordered product is

$$:\phi_I(x_1)\phi_I(x_2): = \phi_I^+(x_1)\phi_I^+(x_2) + \phi_I^-(x_2)\phi_I^+(x_1) + \phi_I^-(x_1)\phi_I^+(x_2) + \phi_I^-(x_1)\phi_I^-(x_2)$$

we can simply insert a θ function in the normal ordered product without altering it as

$$\begin{aligned} :\phi_I(x_1)\phi_I(x_2): &= \theta(x_1^0 - x_2^0) [\phi_I^+(x_1)\phi_I^+(x_2) + \phi_I^-(x_2)\phi_I^+(x_1) + \phi_I^-(x_1)\phi_I^+(x_2) + \phi_I^-(x_1)\phi_I^-(x_2)] \\ &\quad + \theta(x_2^0 - x_1^0) [\phi_I^+(x_1)\phi_I^+(x_2) + \phi_I^-(x_2)\phi_I^+(x_1) + \phi_I^-(x_1)\phi_I^+(x_2) + \phi_I^-(x_1)\phi_I^-(x_2)] \end{aligned}$$

Then the difference in the two products would be (noticing that the only different term is the second term in the normal ordering)

$$\begin{aligned} \overline{\phi_I(x_1)\phi_I(x_2)} &= \theta(x_1^0 - x_2^0) (\phi_I^+(x_1)\phi_I^-(x_2) - \phi_I^-(x_2)\phi_I^+(x_1)) + \theta(x_2^0 - x_1^0) (\phi_I^+(x_2)\phi_I^-(x_1) - \phi_I^-(x_1)\phi_I^+(x_2)) \\ &= \theta(x_1^0 - x_2^0) [\phi_I^+(x_1), \phi_I^-(x_2)] + \theta(x_2^0 - x_1^0) [\phi_I^+(x_2), \phi_I^-(x_1)] \end{aligned}$$

These commutators are only numbers, and not operators.

8.2 The theorem

Wick's theorem simply tell that

$$\begin{aligned}
T[\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)] = & \phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n) + \overbrace{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)}^{1 \text{ contraction}} \\
& + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\cdots\phi_I(x_n)}^{2 \text{ contractions}} + \text{other terms with 1 contraction} \\
& + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)\cdots\phi_I(x_n)}^{3 \text{ contractions}} + \text{other terms with 2 contractions} \\
& + \cdots \text{ terms with 3 and more contractions} \\
& + \begin{cases} \overbrace{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_{n-2})\phi_I(x_{n-1})\phi_I(x_n)}^{1 \text{ uncontracted field}} + \text{terms with only one uncontracted field} & : n \text{ odd} \\ \overbrace{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_{n-1})\phi_I(x_n)}^{0 \text{ uncontracted field}} + \text{terms with no uncontracted field} & : n \text{ even} \end{cases}
\end{aligned} \tag{8.1}$$

As an example,

$$\begin{aligned}
T[\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)] = & \phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4) + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{1 \text{ contraction}} \\
& + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{2 \text{ contractions}} + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{3 \text{ contractions}} \\
& + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{4 \text{ contractions}} + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{5 \text{ contractions}} \\
& + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{6 \text{ contractions}} + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{7 \text{ contractions}} \\
& + \overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}^{8 \text{ contractions}}
\end{aligned}$$

Proof

We prove the theorem by induction.

For $n = 2$, the theorem is definitely true, by the definition of contraction.

Suppose the theorem is true for some n . That is

$$T[\phi_I(x_1)\cdots\phi_I(x_n)] = W(x_1, \cdots, x_n)$$

where we used W as the shorthand for the RHS in equation (8.1)

For $n + 1$, let us assume without loss of generality that x_1^0 is the latest time. We can do this because inside the T all operators commute, since at the end T rearranges them according to their time ordering, and therefore, we can bring the field with the latest time to the first and call it $\phi_I(x_1)$. Then

$$T(\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_{n+1})) = \phi_I(x_1)T(\phi_I(x_2)\cdots\phi_I(x_{n+1}))$$

From our assumption that the theorem is true for n , $T(\phi_I(x_2) \cdots \phi_I(x_{n+1})) = W(x_2, \cdots x_{n+1})$. Now

$$\phi_I(x_1) = \phi_I^+(x_1) + \phi_I^-(x_1)$$

which means

$$T(\phi_I(x_1)\phi_I(x_2) \cdots \phi_I(x_{n+1})) = (\phi_I^-(x_1) + \phi_I^+(x_1))W(x_2, \cdots x_{n+1})$$

The first term in the RHS is already has ϕ^- on the left, but the second term is not normal ordered. Therefore, we can commute the ϕ^+ to normal order the second term too as

$$T(\phi_I(x_1)\phi_I(x_2) \cdots \phi_I(x_{n+1})) = \phi_I^-(x_1)W(x_2, \cdots x_{n+1}) + W(x_2, \cdots x_{n+1})\phi_I^+(x_1) + [\phi_I^+(x_1), W(x_2, \cdots x_{n+1})]$$

The first two terms when added together simply gives

$$:\phi_I(x_1)W(x_2, \cdots x_{n+1}):$$

which is the set of $:\phi_I(x_1)\phi_I(x_2) \cdots \phi_I(x_{n+1}):$ and all terms with $\phi_I(x_1)$ uncontracted.

The next part, which is the commutator, will have terms of the form

$$[\phi_I^+(x_1), \phi_I(x_m)] = [\phi_I^+(x_1), \phi_I^-(x_m)] \quad \because \phi^+ \text{ at different spacetime points commute}$$

But since $x_1^0 > x_m^0 \forall m$, we can also rewrite this as

$$\theta(x_1^0 - x_m^0)[\phi_I^+(x_1), \phi_I^-(x_m)] + \theta(x_m^0 - x_1^0)[\phi_I^+(x_m), \phi_I^-(x_1)] = \overline{\phi_I^+(x_1)\phi_I^-(x_m)}$$

Therefore, the second part will have contractions of $\phi_I(x_1)$ with all $\phi_I(x_m)$.

The sum of the first and second part will have everything that is required in $W(x_1, \cdots, x_{n+1})$, i.e. there is a normal ordered product, plus all possible contractions of all fields. Therefore, for $n + 1$

$$T(\phi_I(x_1)\phi_I(x_2) \cdots \phi_I(x_{n+1})) = W(x_1, \cdots, x_{n+1})$$

Therefore, Wick's theorem is proved by mathematical induction.

What makes Wick's theorem useful, and therefore significant, is that it makes evaluation of these time ordered products simple. Suppose we were calculating the vacuum expectation values of a given $T(\phi_I(x_1) \cdots \phi_I(x_n))$. The only contribution would be from the terms where all the fields are contracted, since other terms are normal ordered and normal ordered products annihilate vacuum and therefore have zero vacuum expectation values.

8.3 The Feynman Propagator

See that

$$\langle 0|T(\phi_I(x)\phi_I(y))|0\rangle = \langle 0|\overline{\phi_I(x)\phi_I(y)}|0\rangle = \overline{\phi_I(x)\phi_I(y)}$$

Since the normal ordered product annihilates the vacuum, contraction is a number, and $\langle 0|0\rangle = 1$.

For $x^0 > y^0$, this evaluates to

$$\int \langle 0|a_{\mathbf{q}}e^{-iq\cdot x}a_{\mathbf{p}}^\dagger e^{ip\cdot y}|0\rangle \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6 2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}}$$

since all other terms give zero. We can now commute the a and a^\dagger to get a delta function, which removes one integral in \mathbf{q} , and obtain

$$\int e^{-ip\cdot(x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3\mathbf{p}}{(2\pi)^3}$$

When $y^0 > x^0$, we similarly get

$$\int e^{ip\cdot(x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3\mathbf{p}}{(2\pi)^3}$$

and therefore the contraction looks like

$$\overline{\phi_I(x)\phi_I(y)} = \theta(x^0 - y^0) \int e^{-ip\cdot(x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3\mathbf{p}}{(2\pi)^3} + \theta(y^0 - x^0) \int e^{ip\cdot(x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3\mathbf{p}}{(2\pi)^3} \quad (8.2)$$

We can write it in a slightly simpler way as

$$\overline{\phi_I(x)\phi_I(y)} = \int \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)} \frac{d^4p}{(2\pi)^4} = \int e^{-ip_0(x^0 - y^0)} e^{i\mathbf{p}\cdot(\mathbf{x} - \mathbf{y})} \frac{i}{p_0^2 - \mathbf{p}^2 - m^2 + i\epsilon} \frac{dp_0 d^3\mathbf{p}}{(2\pi)^4} \quad (8.3)$$

The $i\epsilon$ is just a prescription to indicate in which direction the contour should be closed. See that when we do the integral over p_0 without the $i\epsilon$ shifting, it has two poles,

$$p_0^2 - \mathbf{p}^2 - m^2 = 0 \implies p_0 = \pm\omega_{\mathbf{p}}$$

on the real axis and therefore the integral is ill defined. The $i\epsilon$ shifts the poles in such a way that taking a contour along the real axis and closing the contour along one of the two half planes will give the correct result.

With the introduction of $i\epsilon$, the poles now move to

$$p_0 = \pm\sqrt{\omega_{\mathbf{p}}^2 - i\epsilon} \approx \pm\left(\omega_{\mathbf{p}} - \frac{i\epsilon}{2\omega_{\mathbf{p}}}\right) \approx \pm(\omega_{\mathbf{p}} - i\epsilon)$$

(notice that the $i\epsilon$ is simply an integral prescription and its value has no meaning). Therefore, for the p_0 integral the location of poles is shown in figure (1)

To check that this expression with this pole prescription gives the correct result, consider the two cases

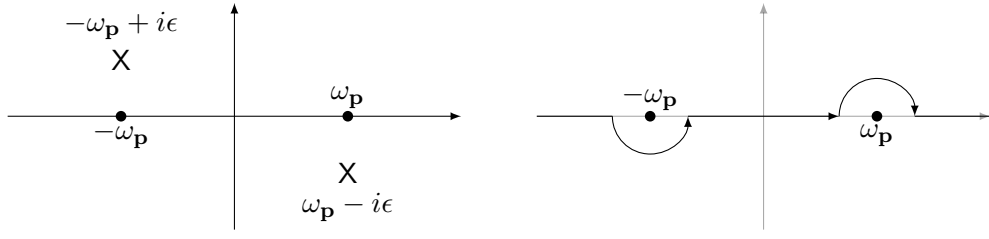


Figure 1: The position of poles in case of $i\epsilon$ prescription (left), and the equivalent contour we need to consider when working without the shifting (right)

1. $x_0 > y_0$

The integrand is

$$e^{-ip_0(x^0-y^0)} = e^{-i\text{Re}(p_0)(x^0-y^0)} e^{\text{Im}(p_0)(x^0-y^0)}$$

In the upper half of the complex plane, this blows at $|p_0| \rightarrow \infty$ and therefore we close the contour in lower half of the complex plane, picking up the pole at $\omega_{\mathbf{p}} - i\epsilon$. The contour is closed in clockwise sense and the value of the integral is

$$-2\pi i \frac{1}{2\omega_{\mathbf{p}}} e^{-i\omega_{\mathbf{p}}(x^0-y^0)}$$

When plugging the value of this integral into the equation (8.3) where one factor of 2π cancels with the $(2\pi)^4$ in denominator, and $-i$ multiplies with i to give 1, we get

$$\int e^{-ip \cdot (x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3 \mathbf{p}}{(2\pi)^3}$$

2. $x_0 < y_0$

In this case the integrand is zero at infinity in the upper half plane and therefore we complete the contour in the upper half plane, picking up the pole at $p_0 = -\omega_{\mathbf{p}}$. The contour is closed counterclockwise, and therefore the integral is

$$(2\pi i)(-1) \frac{1}{2\omega_{\mathbf{p}}} e^{i\omega_{\mathbf{p}}(x^0-y^0)}$$

Plugging this back, we get

$$\int e^{i\omega_{\mathbf{p}}(x_0-y_0)+i\mathbf{p} \cdot (\mathbf{x}-\mathbf{y})} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3 \mathbf{p}}{(2\pi)^3}$$

We can convert the integral from \mathbf{p} to $-\mathbf{p}$ to get

$$\int e^{ip \cdot (x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3 \mathbf{p}}{(2\pi)^3}$$

Therefore we recover

$$\overline{\phi_I(x)\phi_I(y)} = G_F(x-y) = \begin{cases} \int e^{-ip \cdot (x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3 \mathbf{p}}{(2\pi)^3} & : x_0 > y_0 \\ \int e^{ip \cdot (x-y)} \frac{1}{2\omega_{\mathbf{p}}} \frac{d^3 \mathbf{p}}{(2\pi)^3} & : x_0 < y_0 \end{cases}$$

as required. This is called the Feynman Green's function (Feynman Propagator). (Notice that it was very important to consider $+i\epsilon$ in the denominator. Any other addition would not give this result.)

There is another way to understand this $i\epsilon$ prescription as

$$\frac{1}{x-a-i\epsilon} = P\left(\frac{1}{x-a}\right) + i\pi\delta(x-a)$$

where $P()$ stands for the principle value. The above equation is, in essence, a distribution valued equation, and it makes sense only under an integral

$$\int \frac{f(x)}{x-a-i\epsilon} dx = P \int \frac{f(x)}{x-a} dx + i\pi f(a) = \lim_{\delta \rightarrow 0} \left(\int_{-\infty}^{a-\delta} \frac{f(x)}{x-a} dx + \int_{a+\delta}^{\infty} \frac{f(x)}{x-a} dx \right) + i\pi f(a)$$

To prove this, let us consider the equation

$$\frac{1}{x-a-i\epsilon} - \frac{1}{x-a+i\epsilon} = 2i\pi\delta(x-a)$$

which we obtained from adding the above equation to its complex conjugate.

This is equivalent to stating that

$$\int dx \frac{f(x)}{x-a-i\epsilon} - \int dx \frac{f(x)}{x-a+i\epsilon} = 2i\pi f(a)$$

The first term in the above equation simply implies the choice of the contour



and the second term implies the choice of the contour



Subtracting the above two is equivalent to reversing the direction of the second contour and adding, in which case the only piece that remains is



which is equal to $2\pi i$ times the residue of the function at $x = a$, which is exactly equal to $2\pi i f(a)$.

Notice that the Feynman propagator also satisfies

$$(\square + m^2) \langle 0 | T(\phi_I(x) \phi_I(y)) | 0 \rangle = -i\delta^4(x-y)$$

$$\begin{aligned}
\langle 0|T(\phi_I(x)\phi_I(y))|0\rangle &= \int \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip \cdot (x-y)} \frac{d^4 p}{(2\pi)^4} \\
\Rightarrow (\square + m^2) \langle 0|T(\phi_I(x)\phi_I(y))|0\rangle &= (\square + m^2) \int \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip \cdot (x-y)} \frac{d^4 p}{(2\pi)^4} \\
&= \int \frac{i}{p^2 - m^2 + i\epsilon} ((-ip)^2 + m^2) e^{-ip \cdot (x-y)} \frac{d^4 p}{(2\pi)^4} \\
&= -i \int e^{-ip \cdot (x-y)} \frac{d^4 p}{(2\pi)^4} \\
&= -i\delta^4(x - y)
\end{aligned}$$

The Feynman propagator has the property that it propagates positive frequencies forward in time, and negative frequencies backwards in time. To see what this means, see that from equation (8.2), we see that the Feynman propagator is of the form

$$\int \theta(x^0 - y^0) e^{-ik \cdot (x-y)} \theta(k^0) f(k) \frac{d^4 k}{(2\pi)^4} + \int \theta(y^0 - x^0) e^{-ik \cdot (x-y)} \theta(-k^0) f(k) \frac{d^4 k}{(2\pi)^4}$$

(where $f(k)$ is the $\delta(k^2 - m^2)$, and in the above equation we have simply taken equation (8.2) and converted it into a manifestly Lorentz invariant 4-integral).

In the second integral, rather than writing the conversion as $e^{ik \cdot (x-y)} \theta(k^0) f(k)$, we wrote it as $e^{-ik \cdot (x-y)} \theta(-k^0) f(k)$ and used the invariance of the 3-integral under $\mathbf{k} \rightarrow -\mathbf{k}$.

The propagator in this form has the combinations $\theta(x^0 - y^0) \theta(k^0)$ and $\theta(y^0 - x^0) \theta(-k^0)$ which imply that the Feynman propagator propagates positive frequencies forward in time and negative frequencies backward in time.

Feynman propagator is one kind of propagator. We could have considered an anti-time ordered propagator, which would come with a $-i\epsilon$, and would have the opposite behaviour, propagating positive frequencies backward in time and negative frequencies forward in time. Then there are other propagators which send both the frequencies either backwards or forwards.

8.4 The retarded propagator

The retarded propagator is simply

$$\theta(x^0 - y^0) \langle 0|[\phi_I(x), \phi_I(y)]|0\rangle = G_R(x - y)$$

Notice that this propagator also satisfies

$$(\square + m^2)G_R(x - y) = -i\delta^4(x - y)$$

This implies that G_F and G_R should differ by only a solution of the equation of motion. To check that this is true, notice that

$$\frac{\partial}{\partial x^0} (\theta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle) = \delta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle + \theta(x^0 - y^0) \frac{\partial}{\partial x^0} \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle$$

The first term is zero, since at equal times, the fields commute and the delta function enforces equal time. Therefore we get

$$\frac{\partial}{\partial x^0} (\theta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle) = \theta(x^0 - y^0) \langle 0 | [\dot{\phi}_I(x), \phi_I(y)] | 0 \rangle$$

The second derivative w.r.to x^0 would therefore be

$$\begin{aligned} \frac{\partial^2}{\partial (x^0)^2} (\theta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle) &= \delta(x^0 - y^0) \langle 0 | [\Pi_I(x), \phi_I(y)] | 0 \rangle + \theta(x^0 - y^0) \langle 0 | [\ddot{\phi}_I(x), \phi_I(y)] | 0 \rangle \\ &= -i\delta(x^0 - y^0)\delta^3(\mathbf{x} - \mathbf{y}) + \theta(x^0 - y^0) \langle 0 | [\ddot{\phi}_I(x), \phi_I(y)] | 0 \rangle \end{aligned}$$

where in the first term the delta function imposed equal times, and therefore the commutator becomes simply an equal time commutation relation.

There are other terms which are $(-\nabla_x^2 + m^2)G_R$. The ∇_x doesn't care about the θ function and therefore we would simply get

$$(-\nabla_x^2 + m^2) (\theta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle) = \theta(x^0 - y^0) \langle 0 | [(-\nabla_x^2 + m^2)\phi_I(x), \phi_I(y)] | 0 \rangle$$

Adding this and the above equation, we see that we get

$$(\square + m^2) (\theta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle) = -i\delta(x^0 - y^0)\delta^3(\mathbf{x} - \mathbf{y}) + \theta(x^0 - y^0) \langle 0 | [(\square + m^2)\phi_I(x), \phi_I(y)] | 0 \rangle$$

But since $(\square + m^2)\phi_I(x) = 0$, we get

$$(\square + m^2) (\theta(x^0 - y^0) \langle 0 | [\phi_I(x), \phi_I(y)] | 0 \rangle) = -i\delta(x^0 - y^0)\delta^3(\mathbf{x} - \mathbf{y}) = -i\delta^4(x - y)$$

Let us try evaluating this expression in momentum space, the propagator becomes

$$\begin{aligned} \theta(x^0 - y^0) [\phi_I(x), \phi_I(y)] &= \int \left([a_{\mathbf{q}}, a_{\mathbf{p}}^\dagger] e^{-iq \cdot x + ip \cdot y} + [a_{\mathbf{q}}^\dagger, a_{\mathbf{p}}] e^{iq \cdot x - ip \cdot y} \right) \frac{d^3 \mathbf{p} d^3 \mathbf{q}}{(2\pi)^6 2\sqrt{\omega_{\mathbf{p}} \omega_{\mathbf{q}}}} \\ &= \int e^{-iq \cdot (x-y)} - e^{iq \cdot (x-y)} \frac{d^3 \mathbf{q}}{(2\pi)^3 2\omega_{\mathbf{q}}} \theta(x^0 - y^0) \end{aligned}$$

Claim - the whole retarded propagator can be written as

$$G_R(x - y) = \int \frac{i}{(k_0 + i\epsilon)^2 - \mathbf{k}^2 - m^2} e^{-ik \cdot (x-y)} \frac{d^4 k}{(2\pi)^4}$$

In this case, the poles are at $k^0 = \pm\omega_{\mathbf{k}} - i\epsilon$.

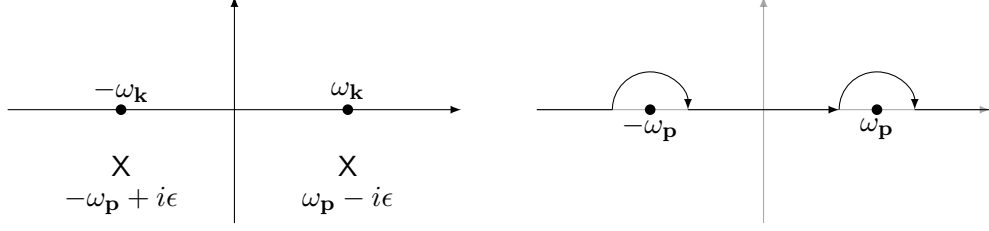


Figure 2: The position of the shifted poles (left), and the equivalent contour (right).

For $x^0 < y^0$, we need to close the contour in the upper half plane, where it does not enclose any poles. Therefore the integral is zero.

For $x^0 > y^0$, we close the contour in the lower half plane, picking up residues of both the poles. The residues are the same as we had calculated for the case of Feynman propagator, and therefore the above integral for $x^0 > y^0$ becomes

$$(-2\pi i) \frac{i}{2\pi} \left(\int e^{-ik \cdot (x-y)} \frac{1}{2\omega_{\mathbf{k}}} \frac{d^3\mathbf{k}}{(2\pi)^3} + \int e^{ik \cdot (x-y)} \frac{1}{-2\omega_{\mathbf{k}}} \frac{d^3\mathbf{k}}{(2\pi)^3} \right)$$

therefore giving

$$G_R(x-y) = \int \frac{i}{(k_0 + i\epsilon)^2 - \mathbf{k}^2 - m^2} e^{-ik \cdot (x-y)} \frac{d^4k}{(2\pi)^4} = \theta(x^0 - y^0) \int e^{-ik \cdot (x-y)} \frac{1}{2\omega_{\mathbf{k}}} (2\pi)^3 - e^{ik \cdot (x-y)} \frac{1}{2\omega_{\mathbf{k}}} \frac{d^3\mathbf{k}}{(2\pi)^3}$$

as required.

There is also an advanced propagator, with $\theta(y^0 - x^0)$ and for this case the pole prescription would be $(k_0 + i\epsilon)^2 - \mathbf{k}^2 - m^2$.

The retarded propagator is the propagator we study in electrodynamics. To understand why this is the case, we need to understand the physical interpretation of this propagator.