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

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-relativistc 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 inifnite 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.
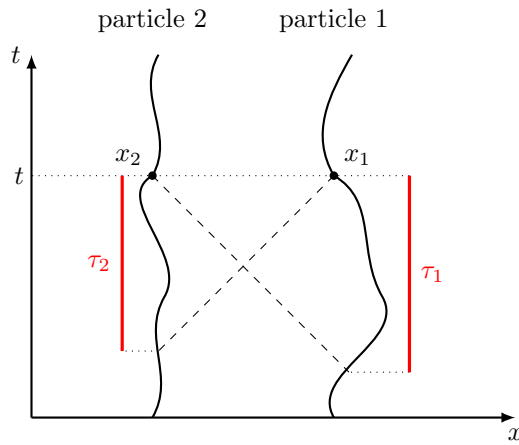
$$F_{12} = F(x_2(t - \tau_2), \dot{x}_2(t - \tau_2), \ldots, x_1(t), \dot{x}_1(t), \ldots) \tag{1.1}$$

and similarly

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

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 uncertainity 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.

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

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

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.

3

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.

$$L = \int \mathcal{L} \, d^3\mathbf{x} \tag{2.1}$$

We consider Lagrangian densities of the form

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

(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

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

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

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

(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

2.1.1 Lagrangin Dynamics

Action — $S = \int \mathcal{L} \, d^4x$

$$\boxed{\text{Nature is lazy}}$$

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

$$\delta S = \int \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi}\delta(\partial_\mu\phi) + \frac{\delta\mathcal{L}}{\delta\phi}\delta(\phi) \, d^4x$$

$$= \int -\delta\phi\,\partial_\mu\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi} + \delta\phi\frac{\delta\mathcal{L}}{\delta\phi} + \partial_\mu\left(\delta\phi\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi}\right) \, d^4x \tag{2.5}$$

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

$$\partial_\mu \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \tag{2.6}$$

2.1.2 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.

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

The Hamiltonian is then

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

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.

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

The equation of motion arising from this lagrangian is simply

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

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

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

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

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

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

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

and

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

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

We can also write this as

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

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

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

and

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

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.

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

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

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

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

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

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

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

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

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 continous index x. Therefore, we impose a similar commutation relation

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

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.

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

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

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

Using this,

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

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

$$[a(\mathbf{p_1}), a^\dagger(\mathbf{p_2})] = \frac{-i}{2}\int[\phi(0, \mathbf{x_1}), \Pi(0, \mathbf{x_2})]e^{-i\mathbf{p_1}\cdot\mathbf{x_1}}e^{i\mathbf{p_2}\cdot\mathbf{x_2}}d^3\mathbf{x_1}d^3\mathbf{x_2} + \frac{i}{2}\int[\Pi(0, \mathbf{x_1}), \phi(0, \mathbf{x_2})]e^{-i\mathbf{p_1}\cdot\mathbf{x_1}}e^{i\mathbf{p_2}\cdot\mathbf{x_2}}d^3\mathbf{x_1}d^3\mathbf{x_2}$$

$$= \int\delta^3(\mathbf{x_1} - \mathbf{x_2})e^{-i\mathbf{p_1}\cdot\mathbf{x_1}}e^{i\mathbf{p_2}\cdot\mathbf{x_2}}d^3\mathbf{x_1}d^3\mathbf{x_2} = (2\pi)^2\delta^3(\mathbf{p_1} - \mathbf{p_2}) \tag{3.2}$$

It also holds that

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

3.1 Hamiltonian in terms of the creation and annihilation operators

The expression for Hamiltonian of the theory is

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

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

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

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

which again simplifies to

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

and similarly

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

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

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

3.2 Momentum of Fields

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

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

Why?

We can see that

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

10

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

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

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)

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

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

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

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

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

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

3.4.1   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

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

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.

11

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 ambiuity 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 $\dfrac{\hat{x}\hat{p} + \hat{p}\hat{x}}{2}$ 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 $:\hat{O}:$.

The normal ordered Hamiltonian is

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

and with this choice of H, we get

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

3.4.2 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

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

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.

3.4.3 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

$$
\begin{aligned}
\langle\mathbf{p_2},\mathbf{q_2}|\mathbf{p_1},\mathbf{q_1}\rangle &= \langle 0|a_{\mathbf{p_2}}a_{\mathbf{q_2}}a_{\mathbf{p_1}}^\dagger a_{\mathbf{q_1}}^\dagger\,|0\rangle \\
&= \langle 0|a_{\mathbf{p_2}}a_{\mathbf{p_1}}^\dagger a_{\mathbf{q_2}}a_{\mathbf{q_1}}^\dagger\,|0\rangle + (2\pi)^3\delta^3(\mathbf{q_2}-\mathbf{p_1})\langle 0|a_{\mathbf{p_2}}a_{\mathbf{q_1}}^\dagger\,|0\rangle \\
&= \langle 0|a_{\mathbf{p_2}}a_{\mathbf{p_1}}^\dagger\,|0\rangle\,(2\pi)^3\delta^3(\mathbf{q_2}-\mathbf{q_1}) + (2\pi)^6\delta^3(\mathbf{q_2}-\mathbf{p_1})\delta^3(\mathbf{p_2}-\mathbf{q_1}) \\
&= (2\pi)^6\delta^3(\mathbf{q_2}-\mathbf{q_1})\delta^3(\mathbf{p_2}-\mathbf{p_1}) + (2\pi)^6\delta^3(\mathbf{q_2}-\mathbf{p_1})\delta^3(\mathbf{p_2}-\mathbf{q_1})
\end{aligned}
$$

13

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.

3.4.4 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

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

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

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

14

This is not the complete picture since,

$$P_2 \left| \mathbf{k_1}, \mathbf{k_2} \right\rangle = 2 \left| \mathbf{k_1}, \mathbf{k_2} \right\rangle$$

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

Therefore, we need a prefactor

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

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

$$P_n = \frac{1}{n!} \int \frac{d^3\mathbf{p_1} \cdots d^3\mathbf{p_n}}{(2\pi)^{3n}} \left| \mathbf{p_1}, \ldots, \mathbf{p_n} \right\rangle \left\langle \mathbf{p_1}, \ldots, \mathbf{p_n} \right|$$

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

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

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

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

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

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

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

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

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

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

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

$$\phi(0, \mathbf{x}) \to \phi(0, \mathbf{x}) + \epsilon\varphi(0, \mathbf{x}) \tag{4.5}$$

Under such a transformation, the wavefunctional transforms as

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

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

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

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. $\phi(0, \mathbf{x}) \to \phi(0, \mathbf{x} + \mathbf{a})$, 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. $\psi[\phi] \to \psi[\phi + \varphi]$.

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

$$\psi[\Pi] = \int \mathfrak{D}\phi \; \psi[\phi]\mathrm{e}^{-i\int \phi\Pi d^3\mathbf{x}} \tag{4.8}$$

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

We can also write a Schrodinger equation given as

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

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

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

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

In terms of these, the commutation relation is

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

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

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

To check that this is true, see that

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

which is indeed the required relation.

We had

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

which now becomes

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

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

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

This is solved by

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

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

$$\langle A \rangle = \frac{\int \mathfrak{D}\phi \ \psi[\phi]^2 A[\phi]}{\int \mathfrak{D}\phi \ \psi[\phi]^2} \tag{4.19}$$

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

$$\int_R O(x)f(x)d^4x \tag{5.1}$$

We can also construct operator like

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

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

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

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

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

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 guarentee that it will be satisfied at some other t as can be seen in the example ($\diamondsuit$))

Is this criteria satisfied in our theory?

We can explicitly calculate this commutator as follows

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

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

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

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

$$\int d^4 p \, \delta^4(p^2 - m^2)\theta(p^0) \tag{5.5}$$

Now this integral is manifestly Lorentz invarient, because it has an integral over d^4p, a δ function in $p^2 - m^2$ and a θ function in p^0. The θ function is also Lorentz invarent 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

$$\int d^3\mathbf{p} \, dp^0 \delta(p_0^2 - \mathbf{p}^2 - m^2)\delta(p^0)$$

$$= \int d\mathbf{p} \, dp^0 \frac{\delta(p_0 - \omega_p)}{2p_0}$$

$$= \int \frac{d^3\mathbf{p}}{2\omega_p}$$

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

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

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 simultaenous, 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

$$f(q) = \int e^{-ip\cdot q} \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}} = e^{-i\omega_p q^0} \int e^{i\mathbf{p}\cdot\mathbf{q}} \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}}$$

$$= e^{-i\omega_p q^0} \int e^{ipq\cos\theta} p^2 dp d\cos\theta \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}} \text{ where now } p \text{ \& } q \text{ stand for } |\mathbf{p}| \text{ \& } |\mathbf{q}|$$

$$= e^{-i\omega_p q^0} \int_0^\infty \frac{1}{iq} \left(e^{ipq} - e^{-ipq}\right) pdp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}}$$

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

$$f(q) = e^{-i\omega_p q^0} \int_0^\infty \frac{1}{iq} e^{ipq} pdp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}} + e^{-i\omega_p q^0} \int_{-\infty}^0 \frac{1}{iq} e^{ipq} pdp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}}$$

$$= e^{-i\omega_p q^0} \int_{-\infty}^\infty \frac{1}{iq} e^{ipq} pdp \frac{1}{(2\pi)^2 2\omega_{\mathbf{p}}}$$

$$= \frac{-i}{4\pi q} \frac{\partial}{\partial q} \int_{-\infty}^\infty e^{-i\omega_p q^0 + ipq} \frac{dp}{2\pi\omega_p}$$

Now, we can make the substitutions

$$\omega_p = m\cosh\phi \ \& \ p = m\sinh\phi \implies \frac{dp}{\omega_p} = d\phi$$

$$q^0 = \alpha\cosh\phi_0 \ \& \ q = \alpha\sinh\phi_0$$

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

21

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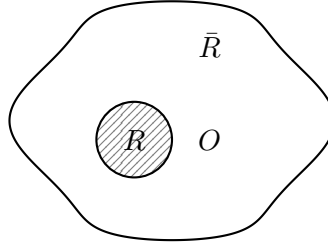
For an excitation $|\psi(R)\rangle$ to be local in R, we require that

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

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

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

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 ϕs 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

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

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 $\phi_i(x) \to \phi_i(x) + \lambda D\phi_i(x)$, (where $D\phi_i$ is just a notation and can denote any function, but is usually a function of original ϕ_is 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

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

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,

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

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

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

$$\phi \to \phi + \lambda D\phi$$

which does not change S.

Now imagine a spacetime dependent variation,

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

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

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

since these are the only terms that are Lorentz invariant and first order in λ.
The action transforms as

$$S \to S + \int d^4x\ \lambda(x)\partial_\mu(K^\mu) + \partial_\mu(\lambda(x))P^\mu$$
$$= S + \int d^4x\ \lambda(x)\partial_\mu(K^\mu - P^\mu)$$

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

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

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

6.2.1 Spacetime translations

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

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

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

where λ is just to keep track of orders.

Since $\mathcal{L}$ is a Lorentz scalar,

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

Therefore, the conserved current is

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

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

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

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 ith 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.

6.2.2   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

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

For this to be a Lorentz transformation

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

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

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

27

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 $\partial_\mu(\epsilon^\mu{}_\nu x^\nu) = \delta^\nu{}_\mu\epsilon^\mu{}_\nu = \eta^{\mu\nu}\epsilon_{\mu\nu} = 0$ ($\because$ *eta* 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

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

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. $\int T^{01} d^3\mathbf{x} = P^i$, meaning

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

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

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

which is just the classical defninition of momentum.

6.2.3   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,

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

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

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

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

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

The corresponding current is

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

Since phi_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

$$\boxed{\text{Prepare a state } |\psi\rangle} \longrightarrow \boxed{\text{Create a disturbance at } x} \longrightarrow \boxed{\text{Measure its effect at } y}$$

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
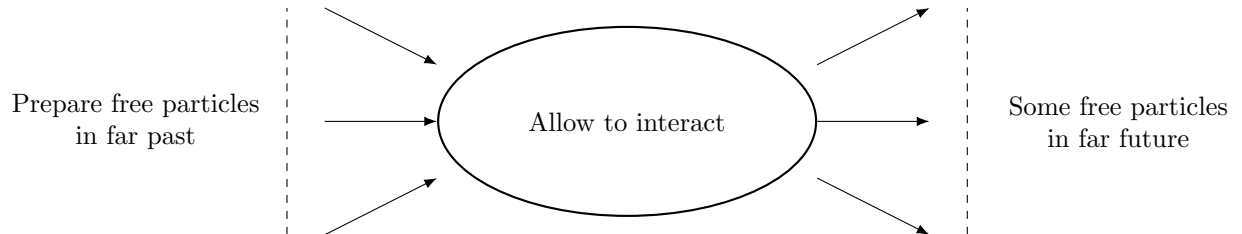
That is, we create free particles in far past, let them evolve and interect, and ask what are the free particles and their properties in far future. This kind of question is answered by something called the S matrix. One might think that this is less general than the other question, but in fact it turns 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 unambigously decided observable. That is if we ask the question like what is the correlator

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

31

this is very ambigous. This is because in interacting theory, there is no unambigous *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 it is not clear how to choose what is the *unique* field that we want to talk about.

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 unambigously 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}

$$\mathbb{H}_{in} = \mathrm{span}\{|0\rangle, |\mathbf{k}\rangle, |\mathbf{k}_1, \mathbf{k}_2\rangle, \cdots\}$$
$$\mathbb{H}_{out} = \mathrm{span}\{|0\rangle, |\mathbf{k}\rangle, |\tilde{\mathbf{k}}_1, \tilde{\mathbf{k}}_2\rangle, \cdots\}$$

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 $(\mathbf{k}_1, \mathbf{k}_2, \cdots)$ can transform into the state $(\tilde{\mathbf{k}}_1, \tilde{\mathbf{k}}_2, \cdots)$ as long as the energy-momentum is conserved.

We can immediately see that

$$_{in}\langle 0|0\rangle_{out} = 1$$

$$_{in}\langle \mathbf{k}|\tilde{\mathbf{k}}\rangle_{out} = (2\pi)^3 \delta(\mathbf{k} - \tilde{\mathbf{k}})$$

and there the non-trivial matrix elements are

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

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. That is, the $S-$matrix is the overlaps between the states in far past with the states in far future. 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

$$|\psi\rangle \to e^{-iHt}|\psi\rangle$$
$$O \to O$$

In QFT, the observables are the fields, and terefore, 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.

$$|\psi\rangle \to |\psi\rangle$$
$$O \to \mathrm{e}^{iHt} O \mathrm{e}^{-iHt}$$

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

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

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

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

(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 $(H_{\mathrm{int}})_I$ that are time dependent).
This is equal to

$$\frac{d}{dt} U_I(t) = -i\mathrm{e}^{iH_0 t} (H - H_0) \mathrm{e}^{-iHt}$$
$$= -i\mathrm{e}^{iH_0 t} (H - H_0) \mathrm{e}^{-iH_0 t} \mathrm{e}^{iH_0 t} \mathrm{e}^{-iHt}$$
$$= -i(H_{\mathrm{int}})_I(t) U_I(t)$$

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,

$$i\frac{d}{dt}U_I(t) = \lambda e^{iH_0 t} \int d^3\mathbf{x} \; \lambda\phi(0, \mathbf{x})^4 e^{-iH_0 t} U_I(t)$$

But

$$e^{iH_0 t} \int d^3\mathbf{x} \; \lambda\phi(0, \mathbf{x})^4 e^{-iH_0 t} = \int d^3\mathbf{x} \; \lambda\phi(t, \mathbf{x})^4$$

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,

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

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

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

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

$$U_I(t) = 1 + (-i)\int_0^t H_I(t')dt' + (-i)^2\int_0^t H_I(t')\int_0^{t'}H_I(t'')dt''dt'$$
$$+ (-i)^3\int_0^t H_I(t')\int_0^{t'}H_I(t'')\int_0^{t''}H_I(t''')dt'''dt''dt'\cdots$$

(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

$$\frac{d}{dt}U_I(t) = -iH_I(t) + (-i)^2H_I(t)\int_0^t H(t'')dt'' + (-i)^3H_I(t)\int_0^t H_I(t'')\int_0^{t''}H_I(t''')dt'''\cdots$$
$$= -iH_I(t)U_I(t)$$

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

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

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

36

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

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

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
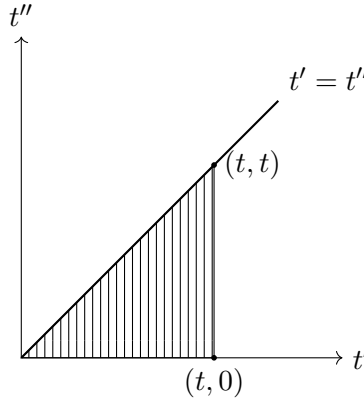
We can also have the t'' integration to go from 0 to t, since we are explicitly enforcing time ordering, we will get the same expression but twice. (Think carefully about this, if there were no time ordering, this manipulation would not be possible, since for the upper triangle, we would be having the 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

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

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

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

37

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 very 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

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

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

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

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

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

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,

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

(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 Heisrnberg picture in the presence of interactions, there is no splitting of the field into creation and annihilation operators. This ordering is only valid for interactin picture fields

We define a contraction as

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

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

$$:\phi_I(x_1)\phi_I(x_2): = \theta(x_1^0 - x_2^0)\left[\phi_I^+(x_1)\phi_I^+(x_2) + \phi_I^-(x_2)\phi_I^+(x_1) + \phi_I^-(x_1)\phi_I^+(x_2) + \phi_I^-(x_1)\phi_I^-(x_2)\right]$$
$$+ \theta(x_2^0 - x_1^0)\left[\phi_I^+(x_1)\phi_I^+(x_2) + \phi_I^-(x_2)\phi_I^+(x_1) + \phi_I^-(x_1)\phi_I^+(x_2) + \phi_I^-(x_1)\phi_I^-(x_2)\right]$$

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

$$\overset{\frown}{\phi_I(x_1)\phi_I(x_2)} = \theta(x_1^0 - x_2^0)\left(\phi_I^+(x_1)\phi_I^-(x_2) - \phi_I^-(x_2)\phi_I^+(x_1)\right) + \theta(x_2^0 - x_1^0)\left(\phi_I^+(x_2)\phi_I^-(x_1) - \phi_I^-(x_1)\phi_I^+(x_2)\right)$$
$$= \theta(x_1^0 - x_2^0)[\phi_I^+(x_1), \phi_I^-(x_2)] + \theta(x_2^0 - x_1^0)[\phi_I^+(x_2), \phi_I^-(x_1)]$$

These commutators are only numbers, and not operators.

8.2 The theorem

Wick's theorem simply tell that

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

As an example,

$$
\begin{aligned}
T[\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)] =\; & :\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4): + :\overline{\phi_I(x_1)\phi}_I(x_2)\phi_I(x_3)\phi_I(x_4): \\
& + :\overline{\phi_I(x_1)\phi_I(x_2)}\phi_I(x_3)\phi_I(x_4): + :\overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)}\phi_I(x_4): \\
& + :\phi_I(x_1)\overline{\phi_I(x_2)\phi}_I(x_3)\phi_I(x_4): + :\phi_I(x_1)\overline{\phi_I(x_2)\phi_I(x_3)}\phi_I(x_4): \\
& + :\phi_I(x_1)\phi_I(x_2)\overline{\phi_I(x_3)\phi}_I(x_4): + :\overline{\phi_I(x_1)\phi}_I(x_2)\overline{\phi_I(x_3)\phi}_I(x_4): \\
& + :\overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)}:
\end{aligned}
$$

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

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

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

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

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

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

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

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

We can write it in a slighly simpler way as

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

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
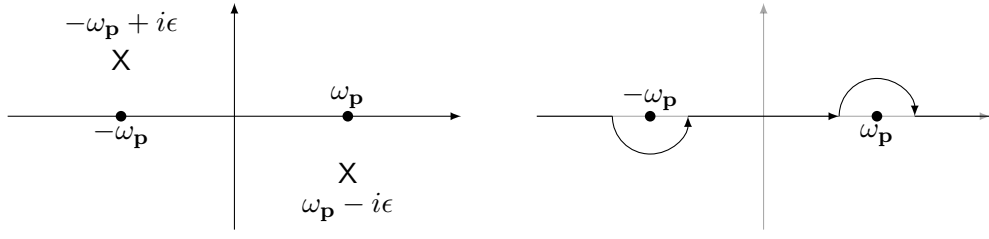
43

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 $|p_0| \to \infty$ 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

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

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

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

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

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

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 progagator also satisfies

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

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

$$\implies (\Box + m^2)\langle 0|T(\phi_I(x)\phi_I(y))|0\rangle = (\Box + m^2)\int \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)} \frac{d^4p}{(2\pi)^4}$$

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

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

$$= -i\delta^4(x-y)$$

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

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

(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 invarent $4-$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 $\mathbf{k} \to -\mathbf{k}$.

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 baheviour, 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

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

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

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

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

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

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

$$\frac{\partial^2}{\partial(x^0)^2}\left(\theta(x^0-y^0)\langle 0|\,[\phi_I(x),\phi_I(y)]\,|0\rangle\right) = \delta(x^0-y^0)\langle 0|\,[\Pi_I(x),\phi_I(y)]\,|0\rangle + \theta(x^0-y^0)\langle 0|\,[\ddot\phi_I(x),\phi_I(y)]\,|0\rangle$$
$$= -i\delta(x^0-y^0)\delta^3(\mathbf{x}-\mathbf{y}) + \theta(x^0-y^0)\langle 0|\,[\ddot\phi_I(x),\phi_I(y)]\,|0\rangle$$

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

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

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

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

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

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

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

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

Claim - the whole retarded propagator can be written as

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

47

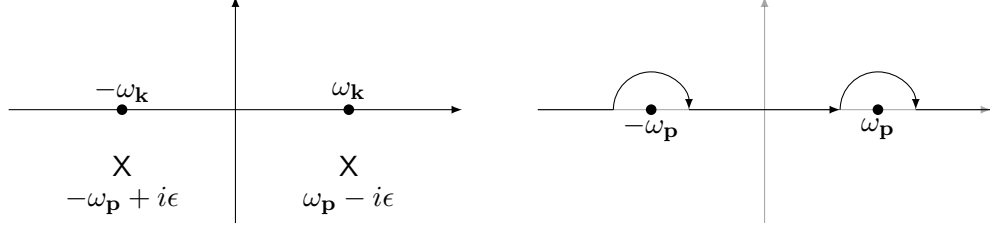In this case, the poles are at $k^0 = \pm\omega_{\mathbf{k}} - i\epsilon$.

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 classical electrodynamics. To understand why this is the case, we need to understand the physical interpretation of this propagator. In QM, we are allowed to turn on sources for fields. These sources are unitary operators. Say we turn on source, which is equivalent to saying that we do

$$|\psi\rangle \to e^{i\int d^4y J(y)\phi(y)}|\psi\rangle$$

that is, we add the term $J(y)\phi(y)$ to the Hamiltonian, with $J(y)$ controlling the strenght and duration of the current. (This is basically saying that we create excitations in a spacetime, whose strength and the region of excitation are defined by J). We can now measure the effect at x, which is basically seeing how much the following expectation value deviates from $\langle\psi|\phi(x)|\psi\rangle$

$$\langle\psi|e^{-i\int d^4y J(y)\phi(y)}\phi(x)e^{i\int d^4y J(y)\phi(y)dy}|\psi\rangle$$

The zeroth order term is $\langle \psi | \phi(x) | \psi \rangle$ itself, and the first order term is

$$\int d^4y\, J(y)\theta(x_0 - y_0)\, \langle \psi | \left[\phi(x), \phi(y) \right] | \psi \rangle$$

Physically, what we are doing is adding a source at y and measuring its effect at x. Causality requires that there cannot be any effect if $y_0 > x_0$, which brings a θ function. Therefore, the effect is, upto first order

$$\int d^4y\, J(y) G_R(x - y)$$

which is the retarded Green's function. (Notice that the above discussion was not precise. To make it precise we need to consider the time ordered exponential as the unitary operator, in which case the θ function naturally emerges. The above discussion is purely for extracting physical insights). Therefore, $G_R(x - y)$ is the linear response at x to a disturbance at y. This is exactly what we do in Classical Mechanics, where we do a disturbance and measure its affect at some different time and different region.

8.5 Wightman Functions

They are simply

$$\langle 0 | \phi_I(x)\phi_I(y) | 0 \rangle = W(x, y)$$

This obeys the equation of motion

$$(\Box + m^2)W(x, y) = 0$$

See that $\phi_I^+(y)$ annihilates the vacuum on the right and $\phi_I^-(x)$ annihilates the vacuum on right, and therefore the only surviving term above is

$$W(x, y) = \langle 0 | \phi_I^+(x)\phi_I^-(y) | 0 \rangle$$

In momentum space, this is

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

which we can commute and obtain

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

As an integral over the four momentum, this can be written as

$$W(x, y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip \cdot (x - y)} 2\pi \delta(p^2 - m^2)\theta(p^0)$$

Consider the object

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

where now the fields are in Heisenberg picture.

We can insert an identity in the form of complete set of momentum eigenstates (the vacuum plus one-particle, multi-particle, etc.) (the actual insertion would involve an integration with some measure and stuff, which is irrelevant here, so we write the insertion as a sum)

$$\langle 0| \phi(x)\phi(y) |0\rangle = \sum_k \langle 0| \phi(x) |k\rangle \langle k| \phi(y) |0\rangle$$

We can always do this, and this evaluates to

$$\sum_k e^{-ik\cdot x} \langle 0| \phi(0) |k\rangle e^{-ik\cdot y} \langle k| \phi(0) |0\rangle$$

This is because of the translational invariance of the Heisenberg-picture field. That is,

$$\phi(x) = e^{iP\cdot x} \phi(0) e^{-iP\cdot x},$$

with $P|k\rangle = k|k\rangle$, $P|0\rangle = |0\rangle$

Therefore the above expression looks like

$$\sum_k e^{-ik\cdot(x-y)} \langle 0| \phi(0) |k\rangle \langle k| \phi(0) |0\rangle$$

Notice that even in the case of interacting fields, we could still insert a complete set of states and got an expression similar to this. Now if we extend $x - y \to x - y - iq$, where q is a future directed (which means $q_0 > 0$) timelike vector ($x - y$ is a $4-$vector, and therefore q should be a $4-$vector). When we do this,

$$W = \sum_k e^{-k\cdot q} e^{-ik\cdot(x-y)} \langle 0| \phi(0) |k\rangle \langle k| \phi(0) |0\rangle$$

The sign of $k \cdot q$ is always positive since k and q are future directed timelike vectors, and their product is always positive ($\because$ $k \cdot q = k_0 q^0 - \mathbf{k} \cdot \mathbf{q}$, and for future directed timelike vectors, this is always positive). Therefore, this addition always improves convergence. So the Wightman function is analytic when $x - y$ is extended to the complex plane with a future directed timelike vector. The only requirement was that k be a timelike vector, i.e. for every momentum eigenstate, the energy should be greater than the total momentum, which is guaranteed by the *spectrum condition in QFT*, and this is true even in interacting theory.

8.6   Normalisation of momentum states

We want to start doing perturbation theory, but before that we need to discuss the normalisation of the states.

So far, we have been working with the states that have the normalisation

$$\langle \mathbf{k}'|\mathbf{k}\rangle = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}')$$

However, there is a problem with this normalisation.

Suppose we Lorentz transform this by Λ. In QFT, the Lorentz transformation acts on the state by some unitary transformation $U(\Lambda)$. With the normalisation we have we see that

$$U(\Lambda)|\mathbf{k}\rangle \neq |\Lambda\mathbf{k}\rangle$$

rather it has to be multiplied by a number. Why?

Suppose the above equation was true. Then we would have

$$\langle \Lambda\mathbf{k}'|\Lambda\mathbf{k}\rangle = \langle \mathbf{k}'| U(\Lambda)^\dagger U(\Lambda)|\mathbf{k}\rangle = \langle \mathbf{k}'|\mathbf{k}\rangle$$

But when we compare the normalisations, we have

$$\langle \Lambda\mathbf{k}'|\Lambda\mathbf{k}\rangle = (2\pi)^3 \delta^3(\Lambda\mathbf{k} - \Lambda\mathbf{k}') \neq (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') = \langle \mathbf{k}'|\mathbf{k}\rangle$$

And we have a contradiction.

The correct transformation would be

$$U(\Lambda)|\mathbf{k}\rangle = \frac{k^0}{(\Lambda k)^0}|\Lambda\mathbf{k}\rangle$$

For our convenience, we use the Lorentz invarient normalisation.

See that

$$\mathbb{I} = \int \frac{d^3\mathbf{k}}{(2\pi)^3}|\mathbf{k}\rangle\langle \mathbf{k}| = \int \frac{d^4 k}{(2\pi)^4}\delta(k^2 - m^2)\theta(k^0)|\mathbf{k}\rangle\langle \mathbf{k}|\, 2\omega_{\mathbf{k}}$$

and therefore it would be convenient if we define

$$|\mathbf{k}\rangle_{\text{new}} = \sqrt{2\omega_{\mathbf{k}}}\,|\mathbf{k}\rangle$$

In this normalisation,

$$\langle \mathbf{k}'|\mathbf{k}\rangle_{\text{new}} = (2\pi)^3 2\omega_{\mathbf{k}}\delta(\mathbf{k} - \mathbf{k}')$$

and this has the property

$$U(\Lambda)|\mathbf{k}\rangle_{\text{new}} = |\Lambda\mathbf{k}\rangle_{\text{new}}$$

since the extra factors now cancel out with the factors these states themselves carry. There was nothing wrong with the previous states, but this normalisation, when doing perturbation theory will give rise to some simpler rules, and therefore are convenient to use. Therefore, from now on, we will use

$$|\mathbf{k}\rangle \equiv |\mathbf{k}\rangle_{\text{new}}$$

With this, we also have to modify the creation and annihilation operators as

$$a_{\mathbf{k}\text{new}} = \sqrt{2\omega_{\mathbf{k}}}a_{\mathbf{k}}, \quad a^{\dagger}_{\mathbf{k}\text{new}} = \sqrt{2\omega_{\mathbf{k}}}a^{\dagger}_{\mathbf{k}}$$

and they satisfy

$$[a_{\mathbf{k}\text{new}}, a^{\dagger}_{\mathbf{k}\text{new}}] = 2\omega_{\mathbf{k}}(2\pi)^3\delta(\mathbf{k} - \mathbf{k}')$$

How we normalise the creation and annihilation operators is upto us and does not modify the content of the theory as long as we keep it consistent. The normalisation of $\phi(x)$ is fixed by the canonical commutation relation, but the definitions of creation and annihilation operators are upto us. What is important, in later calculations, is the objeect $\langle 0|\,\phi(x)\,|k\rangle$, and we require it to satisfy the normalisation

$$\langle 0|\,\phi(x)\,|k\rangle = \mathrm{e}^{-ik\cdot x}$$

With the original normalisation, we had

$$
\begin{aligned}
\langle 0|\,\phi(x)\,|k\rangle &= \langle 0|\int \frac{d^3\mathbf{p}}{\sqrt{2\omega_{\mathbf{p}}}(2\pi)^3}(a_{\mathbf{p}}\mathrm{e}^{-ip\cdot x} + a^{\dagger}_{\mathbf{p}}\mathrm{e}^{ip\cdot x})a^{\dagger}_{\mathbf{k}}\,|0\rangle \\
&= \langle 0|\int \frac{d^3\mathbf{p}}{\sqrt{2\omega_{\mathbf{p}}}(2\pi)^3}a_{\mathbf{p}}\mathrm{e}^{-ip\cdot x}a^{\dagger}_{\mathbf{k}}\,|0\rangle \\
&= \int \frac{d^3\mathbf{p}}{\sqrt{2\omega_{\mathbf{p}}}}\delta^3(\mathbf{p} - \mathbf{k})\mathrm{e}^{-ip\cdot x} \\
&= \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\mathrm{e}^{-ik\cdot x}
\end{aligned}
$$

With the new normalisation, we have

$$
\begin{aligned}
\langle 0|\,\phi(x)\,|k\rangle &= \langle 0|\int \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}(2\pi)^3}(a_{\mathbf{p}\text{new}}\mathrm{e}^{-ip\cdot x} + a^{\dagger}_{\mathbf{p}\text{new}}\mathrm{e}^{ip\cdot x})a^{\dagger}_{\mathbf{k}\text{new}}\,|0\rangle \\
&= \langle 0|\int \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}(2\pi)^3}a_{\mathbf{p}\text{new}}\mathrm{e}^{-ip\cdot x}a^{\dagger}_{\mathbf{k}\text{new}}\,|0\rangle \\
&= \int d^3\mathbf{p}\,\delta^3(\mathbf{p} - \mathbf{k})\mathrm{e}^{-ip\cdot x} \\
&= \mathrm{e}^{-ik\cdot x}
\end{aligned}
$$

From the next section onwards we simply call $a_{\mathbf{k}} \equiv a_{\mathbf{k}\text{new}}$ & $a^{\dagger}_{\mathbf{k}} \equiv a^{\dagger}_{\mathbf{k}\text{new}}$

9   Perturbation Theory

We will consider two models for the perturbation theory,

1. ϕ^4 theory —
$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{m^2}{2}\phi^2 - \frac{\lambda}{4!}\phi^4 \text{ — one field self interacting.}$$

2. Scalar Yukawa theory —

$$\mathcal{L} = |\partial_\mu \psi|^2 - M^2|\psi|^2 + \frac{1}{2}(\partial_\mu \phi)^2 - \frac{m^2}{2}\phi^2 - g|\psi|^2\phi \text{ — three scalar fields } \psi, \psi^\dagger, \phi \text{ interacting with each other.}$$

Free ψ theory —
$$\mathcal{L} = |\partial_\mu \psi(x)|^2 - M^2|\psi(x)|^2 = \partial_\mu \psi(x)\partial^\mu \psi^\dagger(x) - M^2\psi(x)\psi^\dagger(x)$$

The equation of motion for this Lagrangian, when varying with respect to

$$\psi \rightarrow \partial_\mu(\partial^\mu \psi^\dagger(x)) + M^2\psi^\dagger(x) = 0$$
$$\psi^\dagger \rightarrow \partial_\mu(\partial^\mu \psi(x)) + M^2\psi(x) = 0$$

which are again decoupled Klein-Gordon equations. Therefore the solutions can be found, as before using the following Fourier transformation,

$$\psi(t, \mathbf{x}) = \int d^3\mathbf{p} \, \psi(t, \mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}} \implies \psi^\dagger(t, \mathbf{x}) = \int d^3\mathbf{p} \, \psi^\dagger(t, \mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} = \int d^3\mathbf{p} \, \psi^\dagger(t, -\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}}$$

as did in section (2.2) are

$$\psi(t, \mathbf{p}) = b_1(\mathbf{p})\exp(-i\omega_\mathbf{p} t) + b_2(\mathbf{p})\exp(i\omega_\mathbf{p} t)$$
$$\psi^\dagger(t, -\mathbf{p}) = c_1(\mathbf{p})\exp(-i\omega_\mathbf{p} t) + c_2(\mathbf{p})\exp(i\omega_\mathbf{p} t) \implies \psi^\dagger(t, \mathbf{p}) = c_1(-\mathbf{p})\exp(-i\omega_\mathbf{p} t) + c_2(-\mathbf{p})\exp(i\omega_\mathbf{p} t)$$

Comparing the above two equations, we get

$$c_1(-\mathbf{p}) = b_2^*(\mathbf{p}) \ \& \ c_2(-\mathbf{p}) = b_1^*(\mathbf{p})$$

and therefore the fields are (writing everything in terms of $b_1 \equiv b$ and $c_1 \equiv c$)

$$\psi(t, \mathbf{x}) = \int b(\mathbf{p})e^{-i\omega_\mathbf{p} t + i\mathbf{p}\cdot\mathbf{x}} + c^*(-\mathbf{p})e^{i\omega_\mathbf{p} t + i\mathbf{p}\cdot\mathbf{x}} \frac{1}{\sqrt{2\omega_\mathbf{p}}}\frac{d^3\mathbf{p}}{(2\pi)^3}$$
$$\psi^\dagger(t, \mathbf{x}) = \int c(-\mathbf{p})e^{-i\omega_\mathbf{p} t - i\mathbf{p}\cdot\mathbf{x}} + b^*(\mathbf{p})e^{i\omega_\mathbf{p} t - i\mathbf{p}\cdot\mathbf{x}} \frac{1}{\sqrt{2\omega_\mathbf{p}}}\frac{d^3\mathbf{p}}{(2\pi)^3}$$

We can again change the variables in half of both the integrals from $\mathbf{p} \to -\mathbf{p}$ and obtain

$$\psi(t, \mathbf{x}) = \int b(\mathbf{p}) e^{-i\omega_{\mathbf{p}} t + i\mathbf{p} \cdot \mathbf{x}} + c^*(\mathbf{p}) e^{i\omega_{\mathbf{p}} t - i\mathbf{p} \cdot \mathbf{x}} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \frac{d^3\mathbf{p}}{(2\pi)^3}$$

$$\psi^\dagger(t, \mathbf{x}) = \int c(\mathbf{p}) e^{-i\omega_{\mathbf{p}} t + i\mathbf{p} \cdot \mathbf{x}} + b^*(\mathbf{p}) e^{i\omega_{\mathbf{p}} t - i\mathbf{p} \cdot \mathbf{x}} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \frac{d^3\mathbf{p}}{(2\pi)^3}$$

Notice that the theory forces the fields to have creation operator of one kind and annihilation of other kind. Also, we are free to call any of the operators b and the other b^*, but the convention is that the one that comes with a $\exp(+i\omega_{\mathbf{p}} t)$ is $b^\dagger$ since it is the creation operator.

Further, since in the last section we decided on a *redefinition of the momentum eigenstates*, we need to rescale the creation and annihilation operators by multiplying them by $\sqrt{2\omega_{\mathbf{p}}}$, and in terms of the **new** operators, the fields are

$$\psi(t, \mathbf{x}) = \int b(\mathbf{p}) e^{-i\omega_{\mathbf{p}} t + i\mathbf{p} \cdot \mathbf{x}} + c^\dagger(\mathbf{p}) e^{i\omega_{\mathbf{p}} t - i\mathbf{p} \cdot \mathbf{x}} \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}(2\pi)^3}$$

$$\psi^\dagger(t, \mathbf{x}) = \int c(\mathbf{p}) e^{-i\omega_{\mathbf{p}} t + i\mathbf{p} \cdot \mathbf{x}} + b^\dagger(\mathbf{p}) e^{i\omega_{\mathbf{p}} t - i\mathbf{p} \cdot \mathbf{x}} \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}(2\pi)^3}$$

For the complex field, we also have

$$\langle 0| \, T(\psi(x)\psi^\dagger(y)) \, |0\rangle = \int \frac{i}{p^2 - M^2 + i\epsilon} e^{ip \cdot (x-y)} \frac{d^4p}{(2\pi)^4}$$

9.1 ϕ^4 theory

9.1.1 Amplitude for $\phi\phi \to \phi\phi$

We can start by asking the simple question — If we start with "$\mathbf{k}_1, \mathbf{k}_2$" in the past, what is the amplitude for ending up with "$\mathbf{k}_3, \mathbf{k}_4$" in the future. There are a few subtleties, especially regarding the meaning of the states $\mathbf{k}_1, \mathbf{k}_2$ and $\mathbf{k}_3, \mathbf{k}_4$. What we require is that the particles with momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ well separated in far past (i.e. wavepackets well separated in position space, i.e. not a delta function in momentum space, but with a little spread, enough thin to approximated well as a delta function), and similarly in the far future, and then we allow them to interact in between.

But for now we will take a naive approach, assuming that the Hilbert space for the interacting theory is the same as that for a free theory. **This is a wrong assumption, but nevertheless we continue with this assumptionm since it turns out that it doesn't matter at leading order.**

Therefore what we compute is

$$\lim_{t_o \to \infty} \lim_{t_i \to -\infty} \langle \mathbf{k}_3, \mathbf{k}_4 | \, e^{-iHt_o} e^{-iHt_i} \, |\mathbf{k}_1, \mathbf{k}_2\rangle$$

which is nothing but the overlap between the free two-particle states evolved to $-infty$ which are nothing but the in-states, and the free two particle states evolved to $+\infty$ which are the out-states. Therefore, the

above amplitude is the $S-$matrix element for 2 particle in and out states.

Here we are evolving with the full time evolution operator. But since we assume that the Hilbert space is the same as that of free theory, upto a phase, this is the same as

$$\langle \mathbf{k}_3, \mathbf{k}_4 | \, T \left(e^{-i \int_{-\infty}^{\infty} H_I(t)dt} \right) |\mathbf{k}_1, \mathbf{k}_2 \rangle$$

Remember that

$$U_I(t) = e^{iH_0 t} e^{-iHt} \implies e^{-iHt} = e^{-iH_0 t} U_I(t)$$

therefore, in replacing the full time translation operator by the interaction time evolution operator, we only gain a phase $e^{-iE_\psi t}$ if $|\psi\rangle$ is a momentum eigenstate in free theory.

The in and out states are therefore

$$|\mathbf{k}_1, \mathbf{k}_2 \rangle = a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger |0\rangle \,, \quad |\mathbf{k}_3, \mathbf{k}_4 \rangle = a_{\mathbf{k_3}}^\dagger a_{\mathbf{k_4}}^\dagger |0\rangle$$

There is another subtlety about what the vacuum is, since the vacuum of the interacting theory is not the same as that of the free theory. We will assume that in the free theory and the interacting theory the vacuum is the same as it turns out to be at leading order. One way to "arrange" for this is by having the turning on-off function. All of these subtleties, especially assuming the Hilbert spaces and the vacuum to be same in free and interacting theories, they are all important, but they are not important in leading order, and therefore "for now" we will proceed with these assumptions.

In the above equation, we can call

$$T \left(e^{-i \int_{-\infty}^{\infty} H_I(t)dt} \right) = S$$

since this is the matrix that relates the in-states to the out-states as required by the $S-$matrix. The leading order term in this is the identity. This term simply says that the initial two particles go straight through each other without nothing happening. Therefore, the convention is to calculate $S - 1$. (In some cases, in some condensed matter systems etc, it turns out that the 1 also gets corrected when we compute the exact $S-$matrices and we obtain some interesting phases.)

Let us compute the first term. The first term is

$$-i\frac{\lambda}{4!} \langle \mathbf{k}_3, \mathbf{k}_4 | \quad \int_{-\infty}^{\infty} dt \int d^3\mathbf{x}\, \phi_I(t, \mathbf{x})^4 \quad |\mathbf{k}_1, \mathbf{k}_2 \rangle \equiv -i\frac{\lambda}{4!} \langle \mathbf{k}_3, \mathbf{k}_4 | \quad \int d^4x\, \phi_I(x)^4 \quad |\mathbf{k}_1, \mathbf{k}_2 \rangle$$

(where we could convert the $d^3\mathbf{x}$ integral to d^4x only because we were computing the transition from asymptotic past to the asymptotic future. In other cases, this would not have been possible)

At the lowest order, we are calculating

$$\langle \mathbf{k}_3, \mathbf{k}_4 | \quad \int d^4 x \phi_I(x)^4 \quad | \mathbf{k}_1, \mathbf{k}_2 \rangle$$

Let us consider the term

$$\langle \mathbf{k}_3, \mathbf{k}_4 | \, \phi_I(x)\phi_I(x)\phi_I(x)\phi_I(x) \, | \mathbf{k}_1, \mathbf{k}_2 \rangle$$

Since each field operator has one creation and insertion operator, the product will have $2^4 = 16$ terms. In each term there can be only either a creation operator, or an annihilation operator of each field. A term cannot have both creation and annihilation operator both of same field. The fields are inserted at same spacetime point, and therefore they all commute with each other. Only the annihilation contract with $|\mathbf{k}_1, \mathbf{k}_2\rangle$ (since in creating $|\mathbf{k}_1, \mathbf{k}_2\rangle$ we acted on vacuum with the creation operators), while only creation operators act on $\langle \mathbf{k}_3, \mathbf{k}_4 |$. Therefore, one possible term is where the following contractions happen

$$\langle \mathbf{k}_3, \mathbf{k}_4 | \phi_I(x)\phi_I(x)\phi_I(x)\phi_I(x) | \mathbf{k}_1, \mathbf{k}_2 \rangle$$

The two contractions with $\mathbf{k}_3$ and $\mathbf{k}_4$ gives $\mathrm{e}^{i(k_3+k_4)\cdot \mathbf{x}}$ while the other two give $\mathrm{e}^{-i(k_1+k_2)\cdot \mathbf{x}}$. Doing the integral over $d^4 x$, we get one term

$$\frac{-i\lambda}{4!} \int d^4 x \langle \mathbf{k}_3, \mathbf{k}_4 | \phi_I(x)\phi_I(x)\phi_I(x)\phi_I(x) | \mathbf{k}_1, \mathbf{k}_2 \rangle = \frac{-i\lambda}{4!}(2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4)$$

Now there are 4 ways to first contract $\mathbf{k}_1$, 3 ways to then contract $\mathbf{k}_2$, 2 ways to contract $\mathbf{k}_3$ and then only one of the left fields can contract $\mathbf{k}_4$. Therefore, there are 4! such terms, and each term is exactly the same, giving rise to the element

$$\langle \mathbf{k}_3, \mathbf{k}_4 | \int d^4 x \phi_I(x)^4 | \mathbf{k}_1, \mathbf{k}_2 \rangle = -i\lambda(2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4)$$

This is the answer to the question that we asked in the beginning. If we start with "$\mathbf{k}_1, \mathbf{k}_2$" in the past, the amplitude (to the leading order) for ending up with "$\mathbf{k}_3, \mathbf{k}_4$" in the future is $-i\lambda(2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4)$

Let us make the above discussion more precise. In the expansion of ϕ^4, there will be terms of the form $aaaa$, $aaaa^\dagger$, $aaa^\dagger a^\dagger$, $aa^\dagger a^\dagger a^\dagger$ & $a^\dagger a^\dagger a^\dagger a^\dagger$. In our in states there are two creation operators, and our out states has two annihilation operators.
Let us see what happens for the term with three annihilation and one creation operators, say $a^\dagger aaa$. The object that we are computing then is of the form

$$\langle 0|aa(a^\dagger aaa)a^\dagger a^\dagger|0\rangle$$

We can first switch the creation operator on the left with the annhilation operator to its left, picking up a

commutator in the form of delta function

$$= \langle 0|aa^\dagger aaaaa^\dagger a^\dagger|0\rangle + \delta\,\langle 0|aaaaa^\dagger a^\dagger|0\rangle$$

We can keep on doing this as

$$= \langle 0|a^\dagger aaaaaa^\dagger a^\dagger|0\rangle + \delta\,\langle 0|aaaaa^\dagger a^\dagger|0\rangle + \delta\,\langle 0|aaaaa^\dagger a^\dagger|0\rangle$$

$$= 0 + \delta\,\langle 0|aaaa^\dagger aa^\dagger|0\rangle + \delta\delta\,\langle 0|aaaa^\dagger|0\rangle + \delta\,\langle 0|aaaa^\dagger aa^\dagger|0\rangle + \delta\delta\,\langle 0|aaaa^\dagger|0\rangle$$

$$= \delta\delta\,\langle 0|aaaa^\dagger|0\rangle + \delta\delta\delta\,\langle 0|aa|0\rangle + \delta\delta\,\langle 0|aaaa^\dagger|0\rangle + \delta\delta\delta\,\langle 0|aa|0\rangle$$

$$= \delta\delta\delta\,\langle 0|aa|0\rangle + \delta\delta\delta\,\langle 0|aa|0\rangle + \delta\delta\delta\,\langle 0|aa|0\rangle + \delta\delta\delta\,\langle 0|aa|0\rangle = 0$$

(where we are not combining similar looking terms since they are different terms with delta functions, and $a\ a^\dagger$ in different momenta, and from third line onwards directly eliminated terms that are 0 without explicitly showing them.)

If we need a non-zero term, we need two creation operators and two annihilation operators that can each commute with the creation and annihilation operators in the in-out states, giving rise to a constant term. Another way to look at this is that we need two annihilation operators to annihilate the two momenta in incoming state giving the vacuum, and then we need two creation operators to create two momenta, and only this term will have non zero overlap with the out-states. For other combinations of creation and annhilation operators, we will end up with either one or three particle states, which will have zero overlap with the outgoing state.

Let us calculate one such term (not accurately, just to give a gist)

$$\int dp\ dq\ dr\ ds\ \mathrm{e}^{-i(p+q-r-s)\cdot x}\,\langle 0|a_{\mathbf{k_3}}a_{\mathbf{k_4}}a_{\mathbf{p}}^\dagger a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle$$

$$=\int dp\ dq\ dr\ ds\ \mathrm{e}^{-i(p+q-r-s)\cdot x}\,\langle 0|a_{\mathbf{k_3}}a_{\mathbf{p}}^\dagger a_{\mathbf{k_4}}a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle + \delta(\mathbf{k_4}-\mathbf{p})\,\langle 0|a_{\mathbf{k_3}}a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle$$

$$=\int dp\ dq\ dr\ ds\ \mathrm{e}^{-i(p+q-r-s)\cdot x}\delta(\mathbf{k_3}-\mathbf{p})\,\langle 0|a_{\mathbf{k_4}}a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle + \delta(\mathbf{k_4}-\mathbf{p})\,\langle 0|a_{\mathbf{k_3}}a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle$$

$$=\int dp\ dq\ dr\ ds\ \mathrm{e}^{-i(p+q-r-s)\cdot x}\delta(\mathbf{k_3}-\mathbf{p})\,\langle 0|a_{\mathbf{k_4}}a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle + \delta(\mathbf{k_4}-\mathbf{p})\,\langle 0|a_{\mathbf{k_3}}a_{\mathbf{q}}^\dagger a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle$$

$$=\int dp\ dq\ dr\ ds\ \mathrm{e}^{-i(p+q-r-s)\cdot x}\delta(\mathbf{k_3}-\mathbf{p})\delta(\mathbf{k_4}-\mathbf{q})\,\langle 0|a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle + \delta(\mathbf{k_3}-\mathbf{q})\delta(\mathbf{k_4}-\mathbf{p})\,\langle 0|a_{\mathbf{r}}a_{\mathbf{s}}a_{\mathbf{k_1}}^\dagger a_{\mathbf{k_2}}^\dagger|0\rangle$$

$$\vdots$$

$$=\int dp\ dq\ dr\ ds\ \mathrm{e}^{-i(p+q-r-s)}\delta(\mathbf{k_3}-\mathbf{p})\delta(\mathbf{k_4}-\mathbf{q})\delta(\mathbf{k_1}-\mathbf{r})\delta(\mathbf{k_2}-\mathbf{s}) + \text{3 other delta function combinations}$$

$$=4\mathrm{e}^{-i(k_3+k_4-k_2-k_1)\cdot x}$$

Notice that the delta functions will always in all cases end up such that the final answer will have incoming momenta minus the outgoing momenta (or vice versa).

Now, $a^\dagger a^\dagger aa$ was one of the possible terms. There are $^4C_2 = 6$ terms ($a^\dagger aa^\dagger a$, $aaa^\dagger a^\dagger$, etc.) in total that

give rise to the same answer, and therefore the final answer for this case would be

$$\int dx \, 6 \times 4 \times e^{-i(k_3+k_4-k_2-k_1)\cdot x} = 4! \times (2\pi)^4 \delta(k_3 + k_4 - k_2 - k_1)$$

and the 4! from here cancels with the 4! in the coupling constant.

Notice that this entire ordeal is equivalent to saying that I commute the creation(annihilation) operator from one of the fields with a annihilation(creation) operator from one of the out(in) states, to get a factor of $e^{+(-)ikx}$ and that there are 4! ways to do this (4 ways to commute k_1, 3 to commute k_2 and so on), each giving rise to the term $e^{i(\sum_{in} k - \sum_{out} k)}$

9.2 Scalar Yukawa Theory

Following Coleman, let us call the ψ fields nucleons (N) and ϕ fields mesons. This is not accurate since nucleons are fermions, but let us keep this for the time being. (Also keep note that from now, I am dropping the subscript I on the interaction picture fields. Assume that the fields are always in interaction picture)

9.2.1 Amplitude for $\phi \to N\bar{N}$

In this case we can ask a different question - If we started off with a meson ϕ with momentumm $\mathbf{k}_1$ in the far past, what is the amplitude to end up with two nucleons N(created by ψ) & $\bar{N}$(created by $\psi^\dagger$) in the far future with momenta $\mathbf{k}_2$ and $\mathbf{k}_3$.

Even in this case, we follow the same steps as before. We call

$$|\text{in}\rangle = |\mathbf{k}_1\rangle = a^\dagger_{\mathbf{k_1}} |0\rangle$$
$$|\text{out}\rangle = |\mathbf{k}_2, \mathbf{k}_3\rangle = b^\dagger_{\mathbf{k_2}} c^\dagger_{\mathbf{k_3}} |0\rangle$$

We have $H_I(t) = -ig \int d^3\mathbf{x}\phi(x)\psi(x)\psi^\dagger(x)$ and therefore we want the matrix element

$$-ig \langle \mathbf{k}_2, \mathbf{k}_3| \int d^4x \psi(x)\psi^\dagger(x)\phi(x) |\mathbf{k}_1\rangle$$

Now we have fewer choices, ϕ can give only an annihilation operator, and both ψ and $\psi^\dagger$ should give creation operators. Even in the creation operators, there is only one possible contraction; since ψ creates particles of type N and annihilates particles of type $\bar{N}$ and vice versa for $\psi^\dagger$, we need the creation operator from ψ to contract with $\mathbf{k}_3$ and from $\psi^\dagger$ to contract with $\mathbf{k}_2$. Therefore the only possible contraction is the following

$$-ig \int d^4x \langle \mathbf{k}_2, \overset{\frown}{\mathbf{k}_3 | \psi(x)} \psi^\dagger(x) \overset{\frown}{\phi(x) | \mathbf{k}_1}\rangle$$

Now the ϕ contraction should bring in $e^{-ik_1 \cdot x}$, and the other two contractions should bring in $e^{ik_2 \cdot x}$ and

$e^{ik_3 \cdot x}$. Therefore, the above term is

$$-ig \int d^4x e^{-i(k_1-k_2-k_3)\cdot x} = -ig(2\pi)^4 \delta^4(k_1 - k_2 - k_3)$$

In general, for three point amplitudes, it is not easy to find the kinematic configuration where all the delta functions as found above is satisfied. To put things to picture, suppose you work in a frame where k_3 is at rest, i.e.

$$k_3 = M(1, 0, 0, 0)$$

But I can't at the same time, set $k_2 = M(1, 0, 0, 0)$. But we can choose it to be

$$k_2 = M(\cosh\theta, \sinh\theta, 0, 0)$$

which is equivalent to setting the z axis of my coordinate system parallel to $\mathbf{k}_2$.
Now in order for the above delta function to be satisfied, k_1 has to be $k_2 + k_3$, which is

$$k_1 = M(1 + \cosh\theta, \sinh\theta, 0, 0)$$

But for k_1, the mass is m, and therefore it is necessary that $k_1^2 = m^2$, i.e.,

$$M^2((1 + \cosh\theta)^2 - \sinh^2\theta) = m^2 \implies M^2(1 + \cosh^2\theta + 2\cosh\theta - \sinh^2\theta) = m^2$$

which gives

$$2M^2(1 + \cosh\theta) = m^2$$

So we see that unless the particles have a very specific relationship between the masses and this rapidity, we could never had satisfied this delta function. As an example, if the mass m is light compared to M, this delta function can never be satisfied. Therefore, the three point interactions are realised only in very special cases.

9.2.2 Amplitude for $NN \to NN$

We can try asking more questions. For example if we start two N particles in the past, what is the amplitude for getting two N particles in the future.

$$|\text{in}\rangle = |\mathbf{k}_1, \mathbf{k}_2\rangle = c_{\mathbf{k_1}}^\dagger c_{\mathbf{k_2}}^\dagger |0\rangle$$
$$|\text{out}\rangle = |\mathbf{k}_3, \mathbf{k}_4\rangle = c_{\mathbf{k_3}}^\dagger c_{\mathbf{k_4}}^\dagger |0\rangle$$

For this, first order term would be

$$-ig \langle \mathbf{k}_3, \mathbf{k}_4| \int d^4x \; \phi(x)\psi(x)\psi^\dagger(x) |\mathbf{k}_1, \mathbf{k}_2\rangle$$

This is exactly zero since we need two $\psi^\dagger$ and two ψ to contract with the in and out states, and this simply doesn't have the required field insertions.

Let us now try in the second order. In second order, we now have to worry about the time ordering in the time translation operator. Therefore, the term is

$$-\frac{g^2}{2} \langle \mathbf{k}_3, \mathbf{k}_4| \int d^4x d^4y \; T\left[\phi(x)\psi(x)\psi^\dagger(x)\phi(y)\psi(y)\psi^\dagger(y)\right] |\mathbf{k}_1, \mathbf{k}_2\rangle$$

Using Wick's theorem, we can write the time ordered product in terms of different contractions. So we now need to worry about which of the contractions is actually going to contribute to the amplitude. Notice that we need two ψs and two $\psi^\dagger$s to hit the in(out) state to give a non zero contribution. Therefore, the only contraction in which all ψs and $\psi^\dagger$s are left uncontracted while the ϕs are contracted gives a non-zero value. Therefore, the only term we need to worry about is

$$-\frac{g^2}{2} \langle \mathbf{k}_3, \mathbf{k}_4| \int d^4x d^4y \; :\psi(x)\psi^\dagger(x)\psi(y)\psi^\dagger(y): \overbrace{\phi(x)\phi(y)}^{} |\mathbf{k}_1, \mathbf{k}_2\rangle$$

If we did not contract the ϕs we would have ended up with zero, and if we had contracted any of ψ or $\psi^\dagger$s, we would have also obtained zero. See that with Wick's theorem, we do not have to worry about time ordering anymore, we do not have to do horrible integrals where the limits of one integral is the integration variable of the other and so on.

With this, we now have 4 possible terms, which arises from the choices of which ψ to use to annihilate the out states and which $\psi^\dagger$s to annihilate the in states. For one such choice, we get

$$-\frac{g^2}{2} \int d^4x d^4y \underbrace{e^{-ik_1 \cdot y} e^{-ik_2 \cdot x} e^{ik_3 \cdot y} e^{ik_4 \cdot x}}_{\langle \mathbf{k}_3, \mathbf{k}_4|:\psi(x)\psi^\dagger(x)\psi(y)\psi^\dagger(y):|\mathbf{k}_1, \mathbf{k}_2\rangle} \overbrace{\frac{ie^{-ip\cdot(x-y)}}{p^2 - m^2 + i\epsilon}}^{\phi(x)\phi(y)} \frac{d^4p}{(2\pi)^4}$$

Now we can do the x and y integrals to get

$$-\frac{g^2}{2} \int d^4p \; \frac{i}{p^2 - m^2 + i\epsilon}(2\pi)^4\delta^4(k_2 - k_4 + p)\delta^4(k_3 - k_1 + p)$$

We can now do the p integral

$$-\frac{g^2}{2} \frac{i}{(k_1 - k_3)^2 - m^2}(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)$$

The other possible term will be with $x \leftrightarrow y$ but only for the annihilation operator, which will give

$$-\frac{g^2}{2} \int d^4x d^4y e^{-ik_1 \cdot y} e^{-ik_2 \cdot x} e^{ik_4 \cdot y} e^{ik_3 \cdot x} \frac{ie^{-ip\cdot(x-y)}}{p^2 - m^2 + i\epsilon} \frac{d^4p}{(2\pi)^4}$$

60

which will give

$$-\frac{g^2}{2} \int d^4p \, \frac{i}{p^2 - m^2 + i\epsilon} (2\pi)^4 \delta^4(k_2 - k_3 + p) \delta^4(k_4 - k_1 + p)$$

This integral will evaluate to

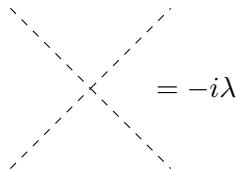$$-\frac{g^2}{2} \frac{i}{(k_1 - k_4)^2 - m^2} (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4)$$

The other two terms are exactly equal to these two terms (since we can obtain them by doing $x \leftrightarrow y$ in both creation and annihilation operators which is simply equivalent to changing the integration variable), and therefore give a factor of 2. Therefore, the final answer at second order is

$$-ig^2 \left(\frac{1}{(k_1 - k_3)^2 - m^2} + \frac{1}{(k_1 - k_4)^2 - m^2} \right) (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4)$$

9.3 Feynman Diagrams

The way we understand Feynman diagrams today is not the way it was developed. In the development phase, it was used to genuinely understand the underlying quantum field theory, and the diagrams had their interpretation. This was a messy business and it was Dyson who cleaned up things by putting everything in order. Today Feynman diagrams is simply a graphical way repackaging of the results of the Dyson series and the Wick's theorem. There is nothing conceptually new in this. It is just a useful simple calculation tool.
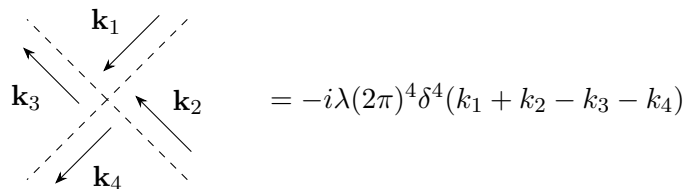
The structure of Feynman diagrams is very simple. There are vertices, and there are propagators connecting these vertices. Vertices are those that come from a single insertion of H_I, and they bring with them a value. In the ϕ^4 theory, the H_I is $\frac{-i\lambda}{4!}\phi^4(x)$, and the corresponding vertex is

$$= -i\lambda$$

where $-i\lambda$ is the value associated with the vertex. Notice that we are not putting the 4!, since the symmetry of the diagram already takes care of it.
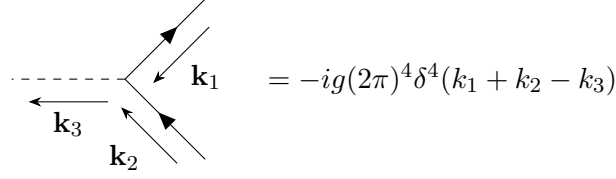
In addition, we are supposed to associate momenta with each leg of a vertex, and impose energy momentum conservation, i.e.

$$= -i\lambda(2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4)$$

What about the $\phi\psi\psi^\dagger$ model? Now there are two kinds of particles, and therefore we need to come up with a notation. Let us call the dotted line ϕ, line with arrow pointing towards the vertex ψ and the line with arrow pointing away $\psi^\dagger$. Therefore, at the vertex, the scalar line meets with a line with arrow pointing towards the vertex and another line with arrow pointing away from the vertex.

$$\text{(diagram)} \quad \mathbf{k}_1 \quad = -ig(2\pi)^4\delta^4(k_1 + k_2 - k_3)$$

Notice that since the in-states enter the calculations as kets and the out-states enter the calculation as bras, the action of ψ and $\psi^\dagger$ on the in-states is not the same as their actions on the out-states. That is, if ψ acts on an in-state it creates an N, while if it acts on an out-state it creates an $\bar{N}$, and vice versa for $\psi^\dagger$. Therefore, the line with arrow towards the vertex, with the momentum also towards the vertex will be an incoming N ($\equiv \psi$) , a line with arrow towards the vertex with momentum arrow away from the vertex will be an outgoing $\bar{N}$ ($\equiv \psi^\dagger$), and vice versa for the line whose arrow points away from the vertex.

Therefore the vertex can be interpreted in multiple ways — it can be considered as having one ingoing ψ and one ingoing $\bar{\psi}$, or it can be one ingoing ψ and one outgoing ψ, or it can be one ingoing $\bar{\psi}$ and one outgoing $\bar{\psi}$. All these are allowed interpretations of this vertex, and they depend on the momentum flow in the vertex.

The propagators are simply lines connecting two vertices, and they are exactly equal to the value (in momentum space) of the time ordered product of two fields.
That is, for the ϕ^4 theory

$$\text{(diagram)} \quad = \mathcal{M}(\langle 0|T(\phi(x)\phi(y))|0\rangle) = \frac{i}{k^2 - m^2}$$

(where I have used $\mathcal{M}$ to denote the momentum space representation) and for the $\psi(\psi^\dagger)$ fields in $\phi\psi\psi^\dagger$ theory,

$$\text{(diagram)} \quad = \mathcal{M}(\langle 0|T(\psi^\dagger(x)\psi(y))|0\rangle) = \frac{i}{k^2 - M^2}$$

where for the vertex on the left, the line has outgoing arrow and therefore is $\psi^\dagger$ and for that on the right, the line has incoming arrow and therefore is ψ. Therefore the relevent propagator is $\langle 0|T(\psi^\dagger(x)\psi(y))|0\rangle$.

The rules for calculations are simple. You join the vertices with propagators to make a diagram that represents the process you are trying to compute. Label the edges with their corresponding momenta (the momenta decide the flow of the diagram), read off the integrals by assigning the corresponding values, and integrate over all undetermined momenta.

Let us try to get the previously obtained results using these diagrams.

The first two are trivial and are directly read from the vertices. The non-trivial result is of the $NN \to NN$. One possible diagram is
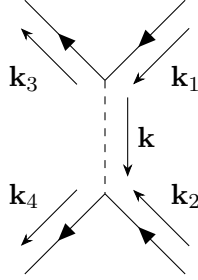
In this diagram the ψ lines are with in-states and therefore are incoming N particles, while $\psi^\dagger$ lines are with out-states and therefore are outgoing N particles. Therefore this diagram indeed represents a $NN \to NN$ scattering.
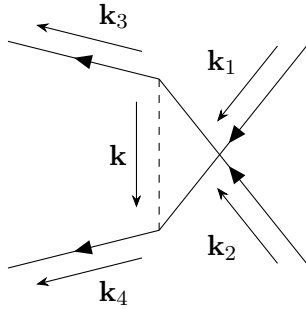
The value of this diagram would be

$$\int \frac{d^k}{(2\pi)^4}(-ig)(2\pi)^4\delta^4(k_1 - k - k_3) \times (-ig)(2\pi)^4\delta^4(k_2 + k - k_4) \times \frac{i}{k^2 - m^2}$$

which is equal to

$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\frac{1}{(k_1 - k_3)^2 - m^2}$$

There is another diagram that we can draw, since the incoming(outgoing) particles are identical, which is

Notice that the location and direction of the incoming and outgoing momenta should be fixed when we draw different diagrams.

For this diagram, the associated value is

$$\int \frac{d^k}{(2\pi)^4}(-ig)(2\pi)^4\delta^4(k_1 - k - k_4) \times (-ig)(2\pi)^4\delta^4(k_2 + k - k_3) \times \frac{i}{k^2 - m^2}$$

which is equal to

$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\frac{1}{(k_1 - k_4)^2 - m^2}$$

And therefore, the amplitude, as before is

$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\left(\frac{1}{(k_1 - k_3)^2 - m^2} + \frac{1}{(k_1 - k_4)^2 - m^2}\right)$$

One can attach a *story* to these diagrams, which is to say that two particles approached each other, exchanged a meson and then flew off. This is not to be taken seriously, these diagrams should only be seen as a calculation tool.

Now that we are equipped with the machinery for computing the scattering amplitudes, we can try to calculate the amplitudes for other processes

9.3.1 Amplitude for $\bar{N}N \to \bar{N}N$

For this process, we can write the following diagrams
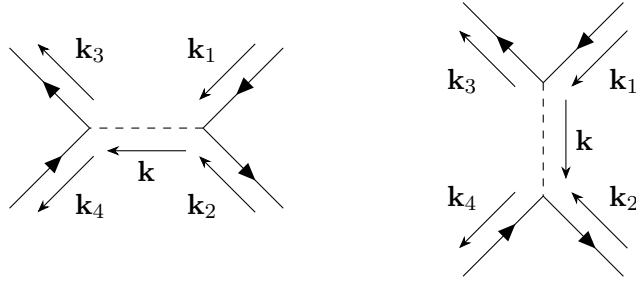
Notice that we have fixed the statement that the incoming N has momentum $\mathbf{k}_1$, and $\bar{N}$ has momentum $\mathbf{k}_2$ and similarly for outgoing nucleons, and since a nucleon is not the same as an antinucleon, there doesnt exist another diagram as seen in the $NN \to NN$ process, where the incoming legs are exchanged.

For the first diagram, the value is

$$\int \frac{d^k}{(2\pi)^4}(-ig)(2\pi)^4\delta^4(k_1 + k_2 - k) \times (-ig)(2\pi)^4\delta^4(k_3 + k_4 - k) \times \frac{i}{k^2 - m^2}$$

which evaluates to

$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\frac{1}{(k_1 + k_2)^2 - m^2}$$

and the second diagram evaluates to

$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\frac{1}{(k_1 - k_3)^2 - m^2}$$

and therefore the total amplitude is
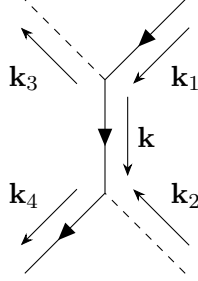
$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\left(\frac{1}{(k_1 + k_2)^2 - m^2} + \frac{1}{(k_1 - k_3)^2 - m^2}\right)$$

Now the story that goes along these two diagrams are different and not the same as was the case in

$NN \to NN$. The first diagram in this corresponds to the N and $\bar{N}$ annihilating each other to give a virtual meson, which then decays back into $N\bar{N}$ pair. Notice that is a virtual meson and not a real meson since it is not on-shell, i.e., it doesnt have momentum $k^2 = m^2$, rather its momenta is variable and determined by the incoming particles. The second one corresponds to two particles approach each other, exchange a meson and keep going.

### 9.3.2 Amplitude for $\phi N \to \phi N$

For this, one possible diagram is The value associated with this is

$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\frac{1}{(k_1 - k_3)^2 - M^2}$$

The story for this diagram is that a nucleon emits a virtual nucleon and turns into a meson, while the meson absorbs the virtual nucleon and becomes a nucleon. See that with interactions, particle identity is not fixed, particles can convert from one to another.

The other diagram one can draw is with the associated value
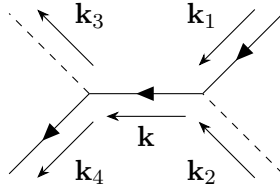
$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\frac{1}{(k_1 + k_2)^2 - M^2}$$

The story again is different, the nucleon and meson meet each other and turn into a virtual nucleon which then again emits a nucleon meson pair. The total amplitude, as usual, is the sum of the values for the two diagrams.

9.4   Detour — Physics of the Yukawa Potential

The Yukawa potential is due to the exchange of massive particles, and is give by

$$V \sim \frac{e^{-mr}}{r}$$

Note that potential is a non-relativistic concept. Relativistic dynamics cannot speak about potential, the reason being the same as discussed int he first section. In case of statics/non-relativistic dynamics, we can speak about potentials, and it is not necessary for all quantum field theories to have deducible potentials. However it is illuminating to see how the things we are calculating in QFT translates to the physics that we can observe. That is, in the non-relativistic limit, the calculations we do in QFT should have some interpretation as some potential.

The objective of this discussion is to consider the process $NN \to NN$, calculate the $S-1$ matrix element, find its non-relativistic limit, and then from that, infer what the non-relativistic potential should be. Remember that for the $NN \to NN$ process, the matrix element

$$\langle \mathbf{k}_3, \mathbf{k}_4 | \, S-1 \, | \mathbf{k}_1, \mathbf{k}_2 \rangle = -ig^2 (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4) \left(\frac{1}{(k_1 - k_3)^2 - m^2} + \frac{1}{(k_1 - k_4)^2 - m^2} \right)$$

First let us discuss some non-relativistc quantum mechanics. The degrees of freedom are described by a wave function which obeys

$$i \frac{\partial}{\partial t} \psi(\mathbf{x}_1, \mathbf{x}_2, t) = \frac{1}{2m} \left(-\nabla_1^2 - \nabla_2^2 \right) \psi(\mathbf{x}_1, \mathbf{x}_2, t) + V(\mathbf{x}_1 - \mathbf{x}_2) \psi(\mathbf{x}_1, \mathbf{x}_2, t)$$

The potential depends on $\mathbf{x}_1 - \mathbf{x}_2$ since we require translation invariance in our theory, which was also inherent in our field theoretic description. Since the potential depends only on $\mathbf{x}_1 - \mathbf{x}_2$, the simplest kind of wavefunction we can write is of the form

$$\psi = \underbrace{\psi_d(\mathbf{x}_1 - \mathbf{x}_2, t)}_{\text{wavefunction of separation DOF}} \times \overbrace{\psi_c(\mathbf{x}_1 + \mathbf{x}_2, t)}^{\text{wavefunction of center of mass DOF}}$$

We can also write the kinetic term in the form

$$\frac{1}{2m} \left(-\nabla_1^2 - \nabla_2^2 \right) = \frac{1}{4m} \left(-\nabla_{\mathbf{x}_1 - \mathbf{x}_2}^2 - \nabla_{\mathbf{x}_1 + \mathbf{x}_2}^2 \right)$$

$$\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2, \quad \mathbf{R} = \mathbf{x}_1 + \mathbf{x}_2$$

$$\nabla_1 = \frac{\partial \mathbf{r}}{\partial \mathbf{x}_1} \nabla_{\mathbf{r}} + \frac{\partial \mathbf{R}}{\partial \mathbf{x}_1} \nabla_{\mathbf{R}} = \nabla_{\mathbf{r}} + \nabla_{\mathbf{R}}$$

$$\nabla_2 = \frac{\partial \mathbf{r}}{\partial \mathbf{x}_2} \nabla_{\mathbf{r}} + \frac{\partial \mathbf{R}}{\partial \mathbf{x}_2} \nabla_{\mathbf{R}} = -\nabla_{\mathbf{r}} + \nabla_{\mathbf{R}}$$

squaring and adding, we get

$$\nabla_1^2 + \nabla_2^2 = 2(\nabla_{\mathbf{r}}^2 + \nabla_{\mathbf{R}}^2)$$

Using this, we can read off the equations of motion for the center of mass and for the separation as

$$i\frac{\partial}{\partial t}\psi_c(\mathbf{x}_1 + \mathbf{x}_2, t) = \frac{1}{4m}\left(-\nabla^2_{\mathbf{x}_1+\mathbf{x}_2}\right)\psi_c(\mathbf{x}_1 + \mathbf{x}_2, t)$$

$$i\frac{\partial}{\partial t}\psi_d(\mathbf{x}_1 - \mathbf{x}_2, t) = \frac{1}{4m}\left(-\nabla^2_{\mathbf{x}_1-\mathbf{x}_2}\right)\psi_d(\mathbf{x}_1 - \mathbf{x}_2, t) + V(\mathbf{x}_1 - \mathbf{x}_2)\psi_d(\mathbf{x}_1 - \mathbf{x}_2, t)$$

We also want the momentum space representation. Notice that

$$\exp(i\mathbf{k}_1 \cdot \mathbf{x}_1 + i\mathbf{k}_2 \cdot \mathbf{x}_2) = \exp\left(i(\mathbf{k}\mathbf{x}_1 + \mathbf{x}_2) \cdot \frac{(\mathbf{k}_1 + \mathbf{k}_2)}{2} + i(\mathbf{x}_1 - \mathbf{x}_2) \cdot \frac{(\mathbf{k}_1 - \mathbf{k}_2)}{2}\right)$$

What this means is that the momentum associated with the ψ_c degree of freedom is $(\mathbf{k}_1 + \mathbf{k}_2)/2$, while ψ_d has the momentum $(\mathbf{k}_1 - \mathbf{k}_2)/2$. However, notice that these are not the good momentum eigenstates to expand our position space eigenstates in. This is because the particles are identical and therefore we require the wavefunciton to be symmetric under $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$. ψ_c already has this property, and on ψ_d we need to impose this property. Therefore the right eigenstates basis for ψ_d will not be the complete expontial, but only the cosine part of it. Field theory was clever, it automatically took care of the Bose statistics. But in non-relativistic QM, we need to impose Bose statistics by hand.

We can now calculate the non-relativistic scattering amplitude for this process. The scattering amplitude for ψ_c and ψ_d are to be considered separately, and the total amplitude will be the product of these two amplitudes.

First let us consider the COM scattering, which is just free particle scattering. Free particle scattering cannot change the momenta, we need the in momenta to be equal to the out momenta, and therefore the free particle scattering amplitude is simply

$$(2\pi)^3\delta^3(\mathbf{k}_3 + \mathbf{k}_4 - (\mathbf{k}_1 + \mathbf{k}_2))$$

The difference scattering can be calculated using the general non-relativistic scattering amplitude formula

$$\langle \mathbf{k}' | S - 1 | \mathbf{k}\rangle = -i(2\pi)\delta(E - E') \langle \mathbf{k}| \frac{1}{1 - VG_\perp} - 1 |\mathbf{k}'\rangle$$

To leading order, this is equal to

$$-i(2\pi)\delta(E - E') \int V(\mathbf{x})e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} d^3\mathbf{x}$$

This is for a one particle scattering from a potential. In our case $\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2$ and $\mathbf{k}' = \mathbf{k}_3 - \mathbf{k}_4$. We also need to impose Bose statistics, which means that $\mathbf{k} \to -\mathbf{k}$ and $\mathbf{k}' \to -\mathbf{k}'$ should also be included, and therefore in our case, the leading order would be

$$-i(2\pi)\delta(E - E') \int V(\mathbf{x}) \left(e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} + e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}}\right) d^3\mathbf{x}$$

(where we have not kept track of (a lot of) factors of 2. The important factor is that there will be terms where both initial and final momenta have same sign, and terms where they have the opposite sign.).

First take a look at the delta function. The delta function in this term can be expressed in terms of the momenta by seeing that

$$\delta(E - E') \sim \frac{1}{m}\delta\left((\mathbf{k}_1 - \mathbf{k}_2)^2 - (\mathbf{k}_3 - \mathbf{k}_4)^2\right)$$

Now since the total amplitude is the product of this expression with the free particle amplitude, this delta function gets multiplied with $\delta^3(\mathbf{k}_3 + \mathbf{k}_4 - (\mathbf{k}_1 + \mathbf{k}_2))$, and therefore we can add $(\mathbf{k}_1 + \mathbf{k}_2)^2$ and subtract $(\mathbf{k}_3 + \mathbf{k}_4)^2$ to get

$$\delta(E - E') \sim \frac{1}{m}\delta\left((\mathbf{k}_1 - \mathbf{k}_2)^2 + (\mathbf{k}_1 + \mathbf{k}_2)^2 - (\mathbf{k}_3 - \mathbf{k}_4)^2 - (\mathbf{k}_3 + \mathbf{k}_4)^2\right) =\sim \frac{1}{m}\delta(k_1^2 + k_2^2 - k_3^2 - k_4^2)$$

This is the non relativistic limit of the delta function

$$\delta(k_1^0 + k_2^0 - k_3^0 - k_4^0)$$

and therefore

$$(2\pi)^4\delta(E - E')\delta^3(\mathbf{k}_3 + \mathbf{k}_4 - (\mathbf{k}_1 + \mathbf{k}_2)) \stackrel{\text{non-relativistic}}{=} (2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)$$

In the amplitude we calculated from the field theory, the delta function was multiplied by terms of the form $((k_1 - k_3)^2 - m^2)^{-1}$, which are four vector terms. We need to match these terms too in our non-relativistic calculations. To do so, we can make use of the fact that the field theoretic amplitude is Lorentz invariant, and therefore we can choose to go to the COM frame. In the COM frame, we have $\mathbf{k}_1 = -\mathbf{k}_2$ and $\mathbf{k}_3 = -\mathbf{k}_4$, which means that the energies are the same, i.e. $k_1^0 = k_2^0 = k_3^0 = k_4^0$. Therefore, the term

$$-ig^2\left(\frac{1}{(k_2 - k_4)^2 - m^2} + \frac{1}{(k_1 - k_4)^2 - m^2}\right) = ig^2\left(\frac{1}{(\mathbf{k}_2 - \mathbf{k}_4)^2 + m^2} + \frac{1}{(\mathbf{k}_1 - \mathbf{k}_4)^2 + m^2}\right)$$

In the COM frame, $\mathbf{k}' = \mathbf{k}_3 - \mathbf{k}_4 = 2\mathbf{k}_3 = -2\mathbf{k}_4$ and $\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2 = 2\mathbf{k}_1 = -2\mathbf{k}_2$. Therefore

$$\mathbf{k} - \mathbf{k}' \sim \mathbf{k}_2 - \mathbf{k}_4$$

and

$$\mathbf{k} + \mathbf{k}' \sim \mathbf{k}_1 - \mathbf{k}_4$$

which means the above term would be

$$ig^2\left(\frac{1}{(\mathbf{k} - \mathbf{k}')^2 + m^2} + \frac{1}{(\mathbf{k} + \mathbf{k}')^2 + m^2}\right) \tag{9.1}$$

in COM frame only.

Therefore, we need a potential $V(r)$ such that

$$\int V(\mathbf{x}) \left(e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} + e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} \right) d^3\mathbf{x} \sim \left(\frac{1}{(\mathbf{k}-\mathbf{k}')^2 + m^2} + \frac{1}{(\mathbf{k}+\mathbf{k}')^2 + m^2} \right)$$

The general integral

$$\int V(\mathbf{x}) e^{i\mathbf{q}\cdot\mathbf{x}} d^3\mathbf{x}$$

can be calculated by going to polar coordinates (assuming the potential to be spherically symmetric) as

$$2\pi \int V(r) e^{iqr\cos\theta} d(\cos\theta) r^2 dr = 4\pi \int V(r) \frac{r}{q} \sin(qr) dr$$

Therefore, for the equation (9.4) to be true, we require

$$\int V(r) \frac{r}{q} \sin(qr) dr = \frac{1}{q^2 + m^2}$$

Now we have

$$\int \frac{e^{-mr}}{r} \frac{r\sin(qr)}{q} dr = \mathrm{Im} \int \frac{e^{-mr+iqr}}{q} dr = \mathrm{Im} \left(\frac{1}{q} \frac{m+iq}{m^2+q^2} \right) = \frac{1}{q^2+m^2}$$

and therefore we get

$$V(r) \sim \frac{e^{-mr}}{r}$$

9.5 Crossing Symmetry in Feynman Diagrams

It refers to the fact that the amplitudes for some processes are related to the amplitudes for some other processes where some of the incoming and outgoing momenta are exchanged. That is, it is the statement that sometimes we can make various transformations to Feynman Diagrams and obtain different processes with same amplitude. The simplest example is the previously studied vertex There are six possible physical
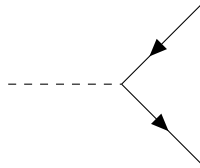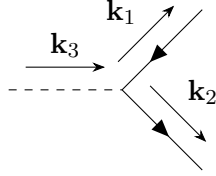

interpretations of this vertex, corresponding to $2^3 - 2$ possible momentum directions (subtract 2 since all momenta inwards and all momenta outwards make no sense physically)

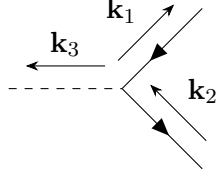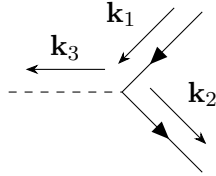1. Nucleon and anti nucleon annihilate to give meson
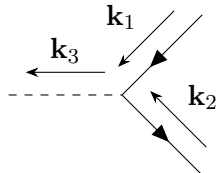
2. Meson creates a nucleon - anti nucleon pair

3. Anti-nucleon emits a meson and continues as an anti-nucleon

4. Nucleon emits a meson and continues as a nucleon

5. Anti-nucleon absorbs a meson and continues as an anti-nucleon
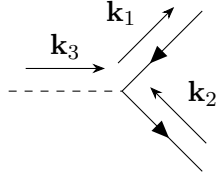
6. Nucleon absorbs a meson and continues as a nucleon
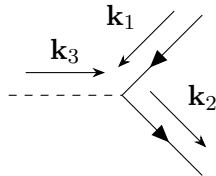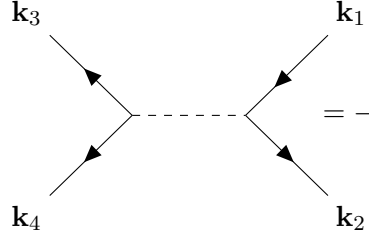
The Feynman diagrams for all these diagrams always assign to these diagrams a $-ig$. The only difference is the signs of the momenta entering the delta function. Its redundent to think of these processes differently, and therefore one can write a compact notation to represent all these values in a single diagram. The notation is to mark all the external momenta as incoming. This is fine since all the external momenta have a particular property, that is, they are on shell $k_i^2 = m^2/M^2$ and they have positive Energy, i.e. $k_i^0 > 0$. Therefore, when we mark all the momenta as incoming, the actual outgoing momenta will have

$k_i^0 < 0$ $\because$ $k_i^\mu = -p^\mu$ for outgoing, and therefore $k_i^0 = -p^0 < 0$. And this will mean that the physical particle is outgoing. Therefore, it is consistent to mark all momenta as incoming since the outgoing particles will automatically get assigned a negative energy, therefore removing confusion. Therefore, we can denote all these scattering amplitudes by
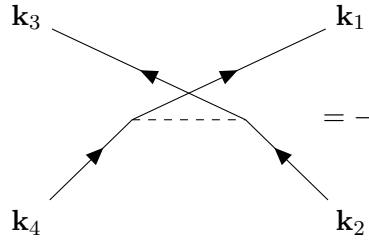
$$-ig(2\pi)^4\delta^4(k_1 + k_2 + k_3)$$

This can be extended to more complicated diagrams too. As an example, the diagram

$$= -ig^2 \frac{1}{(k_1 + k_2)^2 - m^2}(2\pi)^4\delta^4(k_1 + k_2 + k_3 + k_4)$$

There is another different diagram we need to consider, owing to Bose statistics, which is the following

$$= -ig^2 \frac{1}{(k_1 + k_3)^2 - m^2}(2\pi)^4\delta^4(k_1 + k_2 + k_3 + k_4)$$

and the full amplitude is the sum of these two diagrams. What crossing symmetry says is that these are the only possible diagrams, and one can read off all possible processes from this by choosing which k_i^0 to set negative.

To look at it in action, consider the original old diagrams for $N\bar{N} \to N\bar{N}$. which had the amplitude as
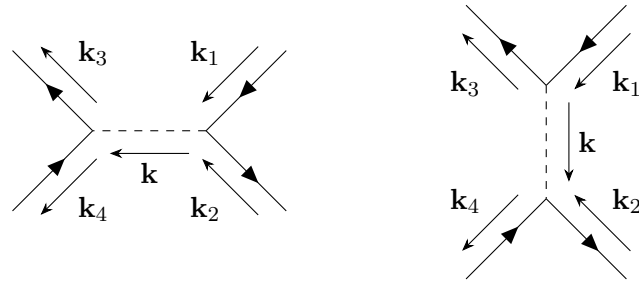
$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 - k_3 - k_4)\left(\frac{1}{(k_1 + k_2)^2 - m^2} + \frac{1}{(k_1 - k_3)^2 - m^2}\right)$$

This can be exactly obtained by the above two digrams in our new conventions, by doing $k_{3/4} = -k_{3/4}$. Notice that in the new convention, we justified the second diagram by invoking Bose statistice, but in the original diagrams, we said that the Bose statistics was not really important. This is due to the fact that

71

the processess $N\bar{N} \to N\bar{N}$, $\bar{N}\bar{N} \to \bar{N}\bar{N}$, $NN \to NN$, etc., are all related to one another by this crossing symmetry, and if Bose statistics is invoked in one, it can and should be invoked in some way in all other.

Let us look at another process $N\bar{N} \to \phi\phi$

In the new convention, with all momenta incoming, these two are the only possible diagrams, and they give the total amplitude

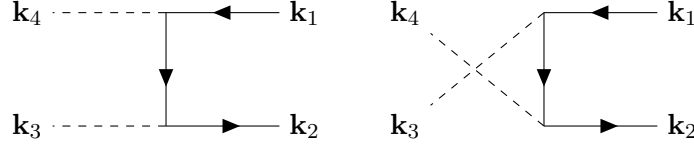$$-ig^2(2\pi)^4\delta^4(k_1 + k_2 + k_3 + k_4)\left(\frac{1}{(k_1 + k_4)^2 - m^2} + \frac{1}{(k_1 + k_3)^2 - m^2}\right)$$

Now using the same diagram, we can represent also the processes $N + \phi \to N + \phi$, $\phi\phi \to NN$, $\phi\phi \to \bar{N}\bar{N}$ etc.

9.6 Mandelstam Variables

Notice that in the 4 point processes that we have calculated so far, we have always encountered terms like $(k_1 + k_3)^2$, $(k_1 + k_4)^2$, etc. We can define a different kinematic representation for momenta that are slightly more convenient in representing the above combinations. These are called Mandelstam variables.

Suppose the four momenta are k_1, k_2, k_3, k_4 (all incoming), then

1. $s = (k_1 + k_2)^2$

2. $t = (k_1 + k_3)^2$

3. $u = (k_1 + k_4)^2$

These are all Lorentz invariant, and they are basically the only unique Lorentz invariant we can construct out of the incoming momenta. All other Lorentz invariant objects (like $k_1 \cdot k_2$ etc) can be written in terms of these. Therefore, all the amplitudes we write should be only functions of s, t, & u. However these are not all independent, they satisfy

$$\begin{aligned} 2(s + t + u) &= (k_1 + k_2)^2 + (k_1 + k_3)^2 + (k_1 + k_4)^2 \\ &= (k_3 + k_4)^2 + (k_2 + k_4)^2 + (k_2 + k_3)^2 \quad (\because k_1 + k_2 + k_3 + k_4 = 0) \\ &= 3(k_1^2 + k_2^2 + k_3^2 + k_4^2) + 2k_1 \cdot k_2 + \text{ other combinations} \end{aligned}$$

Note that $(k_1 + k_2 + k_3 + k_4)^2 = 0 = k_1^2 + \cdot + 2k_1 \cdot k_2 + \cdot \cdot$. Therefore, we get

$$2(s + t + u) = 2(k_1^2 + k_2^2 + k_3^2 + k_4^2) + 0 = 2(\sum_i m_i^2)$$

$$\implies s + t + u = \sum m_i^2$$

Therefore in the above $N\bar{N} \to \phi\phi$, the two diagrams have propagators in t and u, and they are called $t-$channel and $u-$channel diagrams.

The Mandelstam variables are useful even in higher point scatterings, however they become a little complicated sometimes. For higher point correlators, the definitions of Mandelstam variables are also standardized. That is,

$$s_{ij} = k_i \cdot k_j$$

But for these, we get more constraints beyond the obvious constraints we got from $\sum k_i = 0$ and $k_i^2 = m_i^2$. To see how, lets drop s_{nj} (where n is the number of incoming momenta), since $k_n = -\sum_{i=1}^{n-1} k_i$. We are left with

$$^{n-1}C_2 = \frac{(n-1)(n-2)}{2}$$

variables since we can choose i in $n - 1$ ways (leaving n) and j in $n - 2$ ways. That is, we drop all s_{nj}s since we can write the s_{nj} in terms of k_i, $i \neq n$, and $s_{ij} = s_{ji}$. Therefore we need to choose i from $n - 1$ numbers and j from $n - 2$ numbers.

Now in d dimensions, for n point scattering, there are $n(d-1)$ degrees of freedom ($d - 1$ and not d because of mass shell condition). Then there are additional d delta function constraints for the momenta, and therefore the total degrees of freedom is

$$d(n - 1) - d$$

Notice that we have now $\propto n$ degrees of freedom, and $\propto n^2$ free Mandelstam variables, and therefore, there are more constraints, which are not trivial but are some non-trivial non-linear constraints, that should bring down the variables to $\propto n$. This is something we need to keep in mind when working with scattering amplitudes. One way to get around this is to consider scattering in large dimensions, say n, in some processes where the details of the scattering doesn't depend on the dimensionality of the space. Then one can use Mandelstam variables, and calculate the kinematics. But we need to remember that, in finite dimensions, for higher point processes, there are some more constraints that exist, but it is very difficult to handle these, in-fact there is no existing method to obtain the constraints given an $n-$point scattering in d dimensions. Often, it is fine to treat the Mandelstam variables as independent, since the additional non-linear constraint do not have much affect on the kinematics.

9.7 Probabilities and Cross-Sections

What we calculated so far was amplitudes. To get probabilities we should simply square them. The amplitudes are always of the form

$$M_{\text{if}}\ (2\pi)^4 \delta^4(k_i - k_f)$$

and squaring this to get the probability requires us to "square" the delta function, which makes no sense mathematically.

Conceptually speaking, this problem arises because we chose our in and out states as plane wave states, which are not normalisable. If instead we had chosen to work with smeared in and out states normalised to 1, the delta functions would pair with the integration associated with the smearing, and give a finite number for the probability. Notice that the valid physical object is the smeared state, and it is this smeared state that one encounters in experiments. Therefore, we should be studying the wavepackets, which would be a physically right kind of state, and would also circumvent the problem of squaring the delta function.

In practice, we never do this. If we introduce wavepacket in-out states, the results will depend on the form of the wavepackets we consider, and therefore some arbitrariness will creep into our results and we loose the universality of the results we obtain. Further, when we do experiments, the particles we encounter are almost of the form of plane waves with a little spread (i.e. they have a spread very large compared to the region of interaction, and also have sharp momentum distributions), but to an approximation one can always drop this spread and say that the particle has a fixed momentum k. Therefore, if we ask the *physically right kind of questions*, we can continue to work with plane wave approximations.

To be precise, in the following section, we will regulate our calculations by considering wavepackets, and in the end we will consider some limits where the details of the wavepackets do not matter.

The first point is, there are two kinds of plane waves, the incoming plane waves and the outgoing plane waves. At the level of amplitude, there was a symmetry in how the incoming and outgoing states were treated. But when we discuss probabilities, there arises an asymmetry in how we treat them. That is, when discussing probabilities, we average over the initial states and sum over the final states. This is the kind of question we choose to ask. Mathematically, we could have chosen a different question, but this is a physically relevant question. Physically, we are setting up an experiment and there is an uncertainty in the initial conditions due to verious experimental errors, and therefore we average over the initial states, and then there are various possible outcomes and we sum over them. (That is, if we have a few mutually exclusive possible outcomes, the total probability is the sum of the probabilities for the individual processes. As an example consider a process where two electrons are produced. There are different possible outcomes regarding the different possible spins. The probability for the production of two electrons would be the sum of the probabilities of electron production with different spin combinations. This is the right kind of question to ask physically.)

What this means in our case of scattering is that, we consider an initial state of the form $|\text{in}\rangle = \int \psi(p_i)\, |p_1, \ldots, p_n\rangle$, with $\langle \text{in}|\text{in}\rangle = \int \psi(p_i)^2 dp_i = 1$, and ask for the the probability to scatter into $k_f^i + dk_f^i$ for every final state k_f^i. The wavepacket in-state has an integral over the momenta, and therefore absorbs the delta function in the amplitude, giving a finite value. Therefore, we do not need to consider wavepackets in the final states, but just ask for a probability to scatter into an infinitesimal spread in the momenta.

That is, the probability we will consider will be

$$dP = |\langle \text{in} \mid k_1, \ldots, k_m\rangle|^2 \frac{d^3\mathbf{k_1}}{(2\pi)^3 2\omega_{\mathbf{k_1}}} \cdots \frac{d^3\mathbf{k_m}}{(2\pi)^3 2\omega_{\mathbf{k_m}}}$$

Note that $\langle \text{in} \mid k_1, \ldots, k_m\rangle$ is already finite. Also note that if $\psi(p_i)$ was a delta function in the p_is then the delta function in the amplitude would survive and we would not get a finite answer. For the general case when $|\text{in}\rangle$ is not necessarily normalised, the infinitesimal probability would be

$$dP = \frac{|\langle \text{in} \mid k_1, \ldots, k_m\rangle|^2}{\langle \text{in}|\text{in}\rangle} \frac{d^3\mathbf{k_1}}{(2\pi)^3 2\omega_{\mathbf{k_1}}} \cdots \frac{d^3\mathbf{k_m}}{(2\pi)^3 2\omega_{\mathbf{k_m}}}$$

Although the value is finite now, we still have an arbitrariness in the probability since there is a wavepacket in the in-state. We want to get rid of this, by considering in-states that are very close to plane waves. That is, we study the limit where the in-states are plane waves. One way to take this limit is by making the plane wave states normalisable by putting the in-states in a box. That is, we take a box of side L, and consider that all processess happen within this box in a time interval T. In this box, the plane waves are normalisable, and in the limit $L, T \to \infty$ we recover the original plane waves.

When we go back to plane wave incoming states, the delta functions reappear. But see that the problematic term in the amplitude squared was

$$\left((2\pi)^4 \delta^4(\sum k_i + \sum p_i)\right)^2$$

where the delta functions came from

$$\left((2\pi)^4 \delta^4(\sum k_i + \sum p_i)\right)^2 = \left((2\pi)^4 \delta^4(Q)\right)^2 \sim \lim_{L,T\to\infty} \int e^{iQ\cdot x} d^4x \int e^{iQ\cdot y} d^4y$$

We do a bit of handwavery here, you can do the calculations precisely by taking a plane wave limit of the wavepacket and somehow turning off interactions in far past and see that the calculations do match, but here we do not keep that level of precision.

Notice that in doing the first integral above, we get a delta function, which we can use to set $Q = 0$ in the second integrand, giving

$$\left((2\pi)^4 \delta^4(Q)\right)^2 \sim \lim_{L,T\to\infty} (2\pi)^4 \delta^4(Q) \int 1 \, d^4y = \lim_{V,T\to\infty} VT (2\pi)^4 \delta^4(Q)$$

75

(where $V = L^3$) is the volume of the space.

Further, in the plane wave limit, the denominator $\langle \text{in}|\text{in}\rangle$ is no longer 1, but becomes

$$(2\pi)^3 \delta(0) 2\omega_{\mathbf{p}}$$

which happens because of our normalisation

$$\langle \mathbf{p}|\mathbf{p}'\rangle = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') 2\omega_{\mathbf{p}}$$

But this delta function also arises due to an integral over spacetime as

$$\langle \mathbf{p}|\mathbf{p}'\rangle = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') 2\omega_{\mathbf{p}} = 2\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}} \lim_{L,T \to \infty} \int e^{-i(\mathbf{p} - \mathbf{p}') \cdot \mathbf{x}} d^3\mathbf{x}$$

which in $\mathbf{p} = \mathbf{p}'$ case, gives

$$\langle \text{in}|\text{in}\rangle = 2\omega_{\mathbf{p}} \lim_{L,T \to \infty} \int 1 \, d^3\mathbf{x} = 2\omega_{\mathbf{p}} \lim_{V,T \to \infty} V$$

For n incoming particles, this would be

$$\langle \mathbf{p}_1, \ldots, \mathbf{p}_n | \mathbf{p}_1, \ldots, \mathbf{p}_n \rangle = \lim_{V,T \to \infty} V^n \prod_i (2\omega_{\mathbf{p_i}})$$

Putting all these together, we get

$$dP = |M_{\text{if}}|^2 \left(\prod_{j \in \text{out}} \frac{d^3\mathbf{k}_j}{(2\pi)^3 2E_j} \right) \left(\prod_{i \in \text{in}} \frac{1}{2E_i} \right) \times (2\pi)^4 \delta^4 \left(\sum k_{\text{in}} - \sum k_{\text{out}} \right) \lim_{V,T \to \infty} \frac{VT}{V^n}$$

where M_{if} is the part of the amplitude that doesn't contain the delta function. Notice that this experession still has the regulators in the form of volume and time factors.

Now, if we ask the *right kind of physical questions*, the volume and time factors and limits will become irrelevant. Let us consider a few cases. To see what are the questions we should ask, consider first the decay of a particle, which has $n = 1$. In this case, the volume factors cancel off immediately, and only the T factor remains. Therefore the physical question to ask is the **decay probability per unit time**, which is given by

$$\overbrace{\frac{1}{2E_{\text{in}}}}^{\text{not L.I}} |M_{\text{if}}|^2 \underbrace{\left(\prod_{j \in \text{out}} \frac{d^3\mathbf{k}_j}{(2\pi)^3 2E_j} \right) \times (2\pi)^4 \delta^4 \left(\sum k_{\text{in}} - \sum k_{\text{out}} \right)}_{\text{L.I}}$$

where we see that the answer is not Lorentz Invariant (L.I), since the decay probability per unit *time* depends on which frame we are in. The decay probability is minimized in the incoming particle's rest

76

frame, where $E_{\text{in}} = m$. The decay lifetime, which is the inverse of the decay probability per unit time, therefore satisfies

$$\frac{t^{\text{life}}_{\text{moving}}}{t^{\text{life}}_{\text{rest}}} = \frac{E_{\text{moving}}}{m}$$

That is, sometimes a particle that is very short lived can have a longer lifetime by boosting the particle. This is a standard example we consider when we learn special relativity — the muon decay.

Let us consider the case $n = 2$. In this case, the right question to ask would be the *differential cross-section* which is the **probability per unit time per unit of incoming flux**. This is what is is measured in an experiment, and also gives a finite and right answer when calculated.

To discuss this, we first need to do a kinematic computation of flux. Suppose there is a volume $V = L^3$, with one particle having velocity $\mathbf{v}$ at an instant. What we want is the flux of this particle as it goes by. Flux is given by number of particles per unit area per time. In this case, the number of particles is 1, the area of a side is L^2 and the time taken for the particle to cross the volume is $\frac{L_1}{v}$. Therefore, the flux is

$$\text{flux} = \frac{|\mathbf{v}|}{L} \frac{1}{L^2} = \frac{|\mathbf{v}|}{L^3} = \frac{|\mathbf{v}|}{V}$$

What we have is two particles approaching each other, in which case the flux would be

$$\text{flux} = \frac{|\mathbf{v}_1 - \mathbf{v}_2|}{V}$$

The differential cross section is given by

$$d\sigma = \lim_{V,T\to\infty} |M_{\text{if}}|^2 \left(\frac{1}{2E_1} \frac{1}{2E_2} \right) \left(\prod_{j\in\text{out}} \frac{d^3\mathbf{k}_j}{(2\pi)^3 2E_j} \right) \times (2\pi)^4 \delta^4 \left(\sum k_{\text{in}} - \sum k_{\text{out}} \right) \frac{VT}{V^2} \frac{V}{|\mathbf{v}_1 - \mathbf{v}_2|} \frac{1}{T}$$

where now all the volume and time factors cancel out, givin a finite result

$$d\sigma = \underbrace{|M_{\text{if}}|^2}_{\text{amplitude squared}} \overbrace{\left(\frac{1}{2E_1} \frac{1}{2E_2} \right) \left(\prod_{j\in\text{out}} \frac{d^3\mathbf{k}_j}{(2\pi)^3 2E_j} \right) \times (2\pi)^4 \delta^4 \left(\sum k_{\text{in}} - \sum k_{\text{out}} \right) \frac{1}{|\mathbf{v}_1 - \mathbf{v}_2|}}^{\text{phase space factor}}$$

Again, we see that if we ask the right physical question, the details of the regulators are no more relevant. This is a very important formula in particle physics, and we see that a lot of the details of the cross section comes from the kinematical factors and not the amplitude. These are called the phase space factors.

For an important and special case of 2 to 2 scattering, let us work out the phase space factor in the COM frame. Let us consider the center of mass frame, where the incoming particles have $\mathbf{k}_1 + \mathbf{k}_2 = 0$ and

$E_1 + E_2 = E_T$. The phase space factor for this case would be

$$\left(\frac{1}{2E_1}\frac{1}{2E_2}\right)\left(\frac{d^3\mathbf{k}_3}{(2\pi)^3 2E_3}\frac{d^3\mathbf{k}_4}{(2\pi)^3 2E_4}\right)(2\pi)^4\delta(E_T - E_3 - E_4)\delta^3(\mathbf{k}_3 + \mathbf{k}_4)\frac{1}{|\mathbf{v}_1 - \mathbf{v}_2|}$$

We can use the delta function in the outgoing momenta to eliminate one integral, giving the factor to be

$$\frac{1}{2E_1}\frac{1}{2E_2}\frac{1}{2E_3}\frac{1}{2E_4}\left(\frac{d^3\mathbf{k}_3}{(2\pi)^3}\right)(2\pi)\delta(E_T - E_3 - E_4)\frac{1}{|\mathbf{v}_1 - \mathbf{v}_2|}$$

Now in the measure $d^3\mathbf{k}_3 \equiv |\mathbf{k}_3|^2 d|\mathbf{k}_3|d\Omega_3$, there is an integral over the magnitude of the momentum and direction. But the magnitude is fixed by the delta function

$$\delta(E_T - E_3 - E_4) = \delta\left(E_T - \sqrt{|\mathbf{k}_3|^2 + m_3^2} - \sqrt{|\mathbf{k}_3|^2 + m_4^2}\right)$$

This delta function can be written as

$$\frac{\delta\left(|\mathbf{k}_3| - |\mathbf{k}_3|^0\right)}{\left|\frac{\partial E_3}{\partial|\mathbf{k}_3|} + \frac{\partial E_4}{\partial|\mathbf{k}_3|}\right|}$$

Using the one-dimensional delta-function identity

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|},$$

with

$$f(|\mathbf{k}_3|) = E_T - \sqrt{|\mathbf{k}_3|^2 + m_3^2} - \sqrt{|\mathbf{k}_3|^2 + m_4^2}$$

and denoting by $|\mathbf{k}_3|^0$ the (positive) solution of $f(|\mathbf{k}_3|^0) = 0$, we get

$$\delta(E_T - E_3 - E_4) = \delta(f(|\mathbf{k}_3|)) = \frac{\delta\left(|\mathbf{k}_3| - |\mathbf{k}_3|^0\right)}{|f'(|\mathbf{k}_3|^0)|}.$$

Since

$$f'(|\mathbf{k}_3|) = -\frac{\partial}{\partial|\mathbf{k}_3|}\left(\sqrt{|\mathbf{k}_3|^2 + m_3^2}\right) - \frac{\partial}{\partial|\mathbf{k}_3|}\left(\sqrt{|\mathbf{k}_3|^2 + m_4^2}\right) = -\left(\frac{\partial E_3}{\partial|\mathbf{k}_3|} + \frac{\partial E_4}{\partial|\mathbf{k}_3|}\right),$$

we have

$$|f'(|\mathbf{k}_3|^0)| = \frac{\partial E_3}{\partial|\mathbf{k}_3|}\bigg|_{|\mathbf{k}_3|^0} + \frac{\partial E_4}{\partial|\mathbf{k}_3|}\bigg|_{|\mathbf{k}_3|^0}.$$

obtaining

$$\delta(E_T - E_3 - E_4) = \frac{\delta\left(|\mathbf{k}_3| - |\mathbf{k}_3|^0\right)}{\frac{\partial E_3}{\partial|\mathbf{k}_3|}\bigg|_{|\mathbf{k}_3|^0} + \frac{\partial E_4}{\partial|\mathbf{k}_3|}\bigg|_{|\mathbf{k}_3|^0}}.$$

Now since

$$\frac{\partial E_{3/4}}{\partial|\mathbf{k}_3|} = \frac{\mathbf{k}_3}{E_{3/4}}$$

and therefore, we get

$$\delta(E_T - E_3 - E_4) = \frac{\delta\left(|\mathbf{k}_3| - |\mathbf{k}_3|^0\right)}{\dfrac{|\mathbf{k}_3|^0}{E_3} + \dfrac{|\mathbf{k}_3|^0}{E_4}} = \delta\left(|\mathbf{k}_3| - |\mathbf{k}_3|^0\right)\frac{E_3 E_4}{|\mathbf{k}_3|^0(E_3 + E_4)} = \delta\left(|\mathbf{k}_3| - |\mathbf{k}_3|^0\right)\frac{E_3 E_4}{|\mathbf{k}_3|^0 E_T}$$

Calling $|\mathbf{k}_3|^0 \equiv k_3$ from here on, we get the phase space factor as

$$\frac{1}{2E_1}\frac{1}{2E_2}\frac{k_3}{4E_T}\left(\frac{d\Omega_3}{(2\pi)^3}\right)(2\pi)\frac{1}{|\mathbf{v}_1 - \mathbf{v}_2|}$$

We still need to look at the $|\mathbf{v}_1 - \mathbf{v}_2|$ factor. We can use $(m^{rel} = \gamma_v m)$

$$\mathbf{k}_{1/2} = m_{1/2}^{rel}\mathbf{v}_{1/2}, \quad E_{1/2} = m_{1/2}^{rel} \implies \mathbf{v}_{1/2} = \frac{\mathbf{k}_{1/2}}{E_{1/2}}$$

and therefore

$$\frac{1}{2E_1}\frac{1}{2E_2}\frac{k_3}{4E_T}\left(\frac{d\Omega_3}{(2\pi)^3}\right)(2\pi)\frac{1}{\left|\dfrac{\mathbf{k}_1}{E_1} - \dfrac{\mathbf{k}_2}{E_2}\right|}$$

In the COM frame, $\mathbf{k}_2 = -\mathbf{k}_1$, and therefore, we get

$$\frac{1}{2E_1}\frac{1}{2E_2}\frac{k_3}{4E_T}\left(\frac{d\Omega_3}{(2\pi)^3}\right)(2\pi)\frac{E_1 E_2}{k_1 E_T}$$

where we are now calling $|\mathbf{k}_1| \equiv k_1$.

This can be simplified to give

$$\frac{1}{64\pi^2}\frac{1}{E_T^2}\frac{k_3}{k_1}d\Omega_3$$

Therefore, the differential cross section is

$$d\sigma = |M|^2 \frac{1}{64\pi^2}\frac{1}{E_T^2}\frac{k_3}{k_1}d\Omega_3$$

In the COM frame, there is only one degree of freedom which is unspecified, which is the angle of emergance of one of the two outgoing particles. This gives the cross section per unit solid angle, which can be then integrated over some solid angle, or even the entire sphere to get physical results.

Notice that the cross section is inversely proportional to the magnitude of incoming momentum k_1. Therefore, larger the incoming momentum, the smalller is the scattering probability. In the limit, $k_1 \to 0$, the kinematical factor is the maximum, but of course this will be accompanied by changes in M line $M|_{k_1 \to 0} = 0$, and there will be a sweet spot somewhere in between where the cross section is maximised.

# 10 Systematic Formulation of Perturbation Theory

So far we have taken a simple minded approach, not being careful about how we define the in and out states, and also the vacuum. This was not a problem at leading order, but will come to bite us when we discuss perturbation theory in higher order.

To be precise, if we have the free vacuum $H_0 |0\rangle = 0$, and the vacuum of the full theory $H |\Omega\rangle = 0$, the states $|0\rangle$ and $|\Omega\rangle$ are definitely not the same, since H and H_0 do not at all commute. Secondly, so far we had our in and out states as $|k_1, k_2, \ldots\rangle$, but these were created by the action of the creation(annihilation) operators on the free vacuum and that cannot be the right way to think about these states. At lowest order these were fine, but at higher order, these will create problems.

As an example, consider in the ϕ^4 theory the amplitude for 2 to 2 scattering in 2nd order. In 2nd order, one of the diagrams is
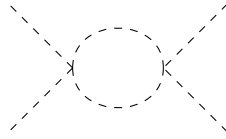
But there can be other disconnected diagrams of the form
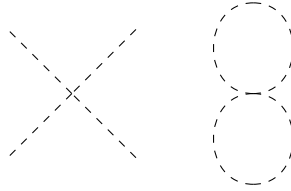
which are not forbidden by Wick's theorem, and therefore MUST be included in our calculations.

This diagram is troublesome since in the loop, the momenta is not constrained at all unlike the first diagram, and we get an infinity that we cannot regulate. That is, the loop in this diagram contributes a diverging value (which cannot be regulated)

$$\left[\int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 - m^2 + i\epsilon} \right]^2 = \infty^2$$

The presence of such diagrams is an indication that we have not been careful with our constructions.

To fix these problems, we need to study the correlation functions (Greens Functions) of the form

$$\langle \Omega | \, T \left(\phi_H(x_1), \ldots, \phi_H(x_n) \right) \, |\Omega\rangle$$

These objects are very important to understand the theory. First of all, we will relate this to the $S-$matrix. But there is also an intrinsic interest in these objects because this forms a general class of observables that

can be studied in a QFT. The formal properties of QFT often will involve these correlators. Here, we will only focus on relating these correlation functions to the $S-$matrix.

10.1 Calculating the Green's Functions

In perturbation theory, the correlation function (Green's function) above discussed can be obtained as

$$\langle \Omega | \ T \left(\phi_H(x_1), \ldots, \phi_H(x_n) \right) \ |\Omega\rangle = \frac{\langle 0 | \ T \left(\phi_I(x_1), \ldots, \phi_I(x_n) e^{-i \int_{-\infty}^{\infty} H_I(t)dt} \right) \ |0\rangle}{\langle 0 | \ T \left(e^{-i \int_{-\infty}^{\infty} H_I(t)dt} \right) \ |0\rangle} \qquad (10.1)$$

10.1.1 Diversion ...

Before proving the above equation, we want to quickly discuss another relation. Suppose we made the split $H = H_0 + H_{\text{int}}$ at some time t_0, then the evolution operator from t_1 to t_2 is given by

$$U_I(t_2, t_1) = T \left(e^{-i \int_{t_1}^{t_2} H_I(t)dt} \right) = e^{iH_0(t_2-t_0)} e^{-iH(t_2-t_1)} e^{-iH_0(t_1-t_0)}$$

which can be written as

$$U_I(t_2, t_1) = \overbrace{e^{iH_0(t_2-t_0)} e^{-iH(t_2-t_0)}}^{U_I(t_2)} \overbrace{e^{iH(t_1-t_0)} e^{-iH_0(t_1-t_0)}}^{U_I^\dagger(t_1)}$$

To see that this equation is indeed correct, check that the differential equation the LHS and RHS satisfies are the same. Suppose we differentiate RHS with respect to t_2, we get

$$i\frac{\partial U_I(t_2, t_1)}{\partial t_2} = \frac{\partial U_I(t_2)}{\partial t_2} U_I(t_1) = H_I(t_2)U_I(t_2)U_I^\dagger(t_1) = H_I(t_2)U_I(t_2, t_1)$$

while LHS gives

$$i\frac{\partial U_I(t_2, t_1)}{\partial t_2} = i\frac{\partial}{\partial t_2} T \left(e^{-i \int_{t_1}^{t_2} H_I(t)dt} \right) = H_I(t_2)U_I(t_2, t_1)$$

and similarly

$$i\frac{\partial U_I(t_2, t_1)}{\partial t_1} = -U_I(t_2, t_1)H_I^\dagger(t_1)$$

for LHS and RHS. Therefore,

$$T \left(e^{-i \int_{t_1}^{t_2} H_I(t)dt} \right)$$

and

$$U_I(t_2)U_I^\dagger(t_1)$$

satisfy the same differential equation, and are hence equal. We can now extend this to have the following

$$U_I(t_2, t_m)U_I(t_m, t_1) = U_I(t_2)U_I^\dagger(t_m) \cdot U_I(t_m)U_I^\dagger(t_1) = U_I(t_2, t_1), \quad t_2 > t_m > t_1$$

(since U_I is unitary).

Armed with these results, we can now proceed to prove the equation (10.1).

10.1.2   Continued . . .

Without loss of generality, let us suppose $x_1^0 > x_2^0 > \cdots > x_n^0$. The numerator of RHS in eq (10.1) becomes

$$\langle 0| \ T\left(\phi_I(x_1), \ldots, \phi_I(x_n)e^{-i\int_{-\infty}^{\infty} H_I(t)dt}\right) \ |0\rangle$$
$$= \langle 0| \ T\left(e^{-i\int_{t_1}^{\infty} H_I(t)dt}\right) \phi_I(x_1)T\left(e^{-i\int_{t_2}^{t_1} H_I(t)dt}\right) \phi_I(t_2)T\left(e^{-i\int_{t_3}^{x_2} H_I(t)dt}\right) \phi_I(x_3)\ldots\phi(x_n)T\left(e^{-i\int_{-\infty}^{t_n} H_I(t)dt}\right) \ |0\rangle$$

$(t_n \equiv x_n^0)$, where we were able to split the $T\left(e^{-i\int_{-\infty}^{\infty} H_I(t)dt}\right)$ owing to the relation we derived above, and since each of the part in the splitting has operators at time $t \in [x_{i-1}, \ x_i]$, the time ordering simply places them as above.

Therefore the above term can be written as

$$\langle 0| \ U_I(\infty, t_1)\phi_I(x_1)U_I(t_1, t_2)\phi_I(t_2)U_I(t_2, t_3)\phi_I(x_3)\ldots\phi(x_n)U_I(t_n, -\infty) \ |0\rangle$$
$$= \langle 0| \ U_I(\infty)\underbrace{U_I^\dagger(t_1)\phi_I(x_1)U_I(t_1)}_{\phi_H(x_1)}\underbrace{U_I^\dagger(t_2)\phi_I(t_2)U_I(t_2)}_{\phi_H(x_2)}U_I^\dagger(t_3)\phi_I(x_3)\ldots\phi(x_n)U_I(t_n)U_I^\dagger(-\infty) \ |0\rangle$$
$$= \langle 0| U_I(\infty)\phi_H(x_1)\phi_H(x_2)\ldots\phi_H(x_n)U^\dagger(-\infty) |0\rangle$$

By definition,

$$U_I^\dagger(t) \equiv U_I^\dagger(t, t_0) = e^{iH(t-t_0)}e^{-iH_0(t-t_0)}$$

and therefore

$$U_I^\dagger(-\infty) |0\rangle = \lim_{T\to-\infty} e^{iH(T-t_0)}e^{-iH_0(T-t_0)} |0\rangle$$

But $H_0 |0\rangle = 0$, and therefore,

$$U_I^\dagger(-\infty) |0\rangle = \lim_{T\to-\infty} e^{iH(T-t_0)} |0\rangle$$

Now since the eigenstates of the full Hamiltonian forms a complete set of basis, we can perform the following expansion

$$|0\rangle = c |\Omega\rangle + \sum c_n |E_n\rangle$$

and under the evolution

$$\lim_{T\to-\infty} e^{iH(T-t_0)}\left(c |\Omega\rangle + \sum c_n |E_n\rangle\right) = c |\Omega\rangle + \lim_{T\to\infty} \sum c_n e^{-iE_n(T+t_0)} |E_n\rangle$$

only the $c |\Omega\rangle$ term is not multiplied with a phase, while all other terms are multiplied a wildly oscillating phase. In particular, if you take any other state

$$|\psi\rangle = k |0\rangle + \sum k_m |E_m\rangle$$

and if we look at $\lim_{T \to -\infty} \langle \psi | e^{iH(T-t_0)} | 0 \rangle$, then we get

$$\lim_{T \to -\infty} \langle \psi | e^{iH(T-t_0)} | 0 \rangle = ck + \lim_{T \to -\infty} \sum_{n,m} c_n k_m e^{-iE_n(T+t_0)}$$

Now in QFT, the spectrum of energy eigenstates is always continuous, and if we start with some smooth distribution of k_n, what we end up with is the same distrubution k_n but weighted with wildly oscillating phases, and when that happens, the sum (integral) is always zero. This follows from the Riemann-Lebesgue lemma which says that for any well-behaved function $f(x)$

$$\lim_{\mu \to \infty} \int_a^b dx \ f(x) e^{i\mu x} = 0$$

(Notice that in the above sum over n, we are actualy integrating over dn since the spectrum is continous). Therefore, we can ignore these rapidly oscillating phases since they do not contribute to the matrix elements (we are making a slight assumption here that we will take inner products with states which are slightly smeared in energy, i.e. having smooth distributions of k_n. If we take *a particular state*, then the inner product may or may not be zero).

Therefore, for all effects and purposes

$$U_I^\dagger(-\infty) | 0 \rangle = \lim_{T \to -\infty} e^{iH(T-t_0)} | 0 \rangle = |\Omega\rangle \langle \Omega | 0 \rangle$$

By the same argument,

$$\langle 0 | U_I(\infty) = \lim_{T \to \infty} \langle 0 | e^{-iH(T-t_0)} = \langle 0 | \Omega \rangle \langle \Omega |$$

Therefore, we get

$$\langle 0 | T \left(\phi_I(x_1), \ldots, \phi_I(x_n) e^{-i \int_{-\infty}^{\infty} H_I(t) dt} \right) | 0 \rangle = \langle 0 | \Omega \rangle \langle \Omega | T \left(\phi_H(x_1) \phi_H(x_2) \ldots \phi_H(x_n) \right) | \Omega \rangle \langle \Omega | 0 \rangle$$

Now the denominator of RHS in eq (10.1) is simple, it is just the above derivation, but without the $\phi(x_i)$s, and therefore simply gives

$$\langle 0 | \ T \left(e^{-i \int_{-\infty}^{\infty} H_I(t) dt} \right) \ | 0 \rangle = \langle 0 | \Omega \rangle \langle \Omega | 0 \rangle$$

which cancels out with the factors in the numerator, and hence proved.

We know exactly how to calculate the objects in the RHS, which is simply using the Wick's theorem. As an example, if we want to compute the four point function in ϕ^4 theory,

$$\langle \Omega | T \left(\phi_H(x_1) \phi_H(x_2) \phi_H(x_3) \phi_H(x_4) \right) | \Omega \rangle = \langle 0 | T \left(\phi_I(x_1) \phi_I(x_2) \phi_I(x_3) \phi_I(x_4) e^{-i \int_{-\infty}^{\infty} H_I(t) dt} \right) | 0 \rangle$$

In first order in perturbation theory the above is

$$\langle 0 | T \left(\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4) \left(-i \int_{-\infty}^{\infty} H_I(t)dt \right) \right) | 0 \rangle$$

when Wick contracting, the only term that is non-zero is

$$-i\lambda \int d^4x \, G(x_1,x)G(x_2,x)G(x_3,x)G(x_4,x)$$

where G is the propagator. (We do not pick up the disconnected pieces where two x_is contract with each other leaving a disconnected leg). This is calculated using a feynman diagram where now we also add a propagator to the external legs (legs which connect a 4-valent vertex (black filled dots in following diagrams) to an external field insertion (white filled dots)).
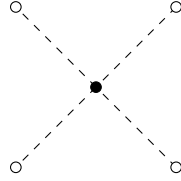
In momentum space, this looks like

$$-i\lambda \left(\frac{i}{k_1^2 - m^2 + i\epsilon} \right) \left(\frac{i}{k_2^2 - m^2 + i\epsilon} \right) \left(\frac{i}{k_3^2 - m^2 + i\epsilon} \right) \left(\frac{i}{k_4^2 - m^2 + i\epsilon} \right) (2\pi)^4 \delta^4 \left(\sum_i k_i \right)$$

The Feynman rules is similar to the one we used to calculate scattering amplitudes where we are now also adding a propagator for the external legs. LSZ formula simply gives a way to calculate the $S-$matrix elements from these correlation functions by stripping off the external propagators.

Let us see how this "corrects" the issue we had with the disconnected diagrams. Suppose we wanted to calculate the four point function in the ϕ^4 theory as before. The numerator on the RHS of the equation (10.1) can be calculated via Feynman diagrams as discussed above. The relevant Feynman diagrams are
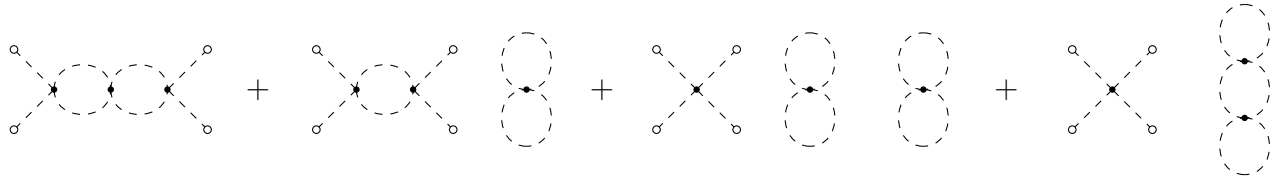
First order:

Second order:

Third order:

and so on.

If we collect together the subdiagrams which have all the lines connected to external legs and call them connected diagrams,

and the subdiagrams where there are no external legs at all the disconnected diagrams

then we see that we can factor out the disconnected diagrams from the sum over all the diagrams to give

$$\langle 0 | \, T \left(\phi_I(x_1) \dots \phi_I(x_n) \left(e^{-i \int_{-\infty}^{\infty} H_I(t)dt} \right) \right) | 0 \rangle = \left(\sum \text{connected pieces} \right) \times \left(\sum \text{disconnected pieces} \right)$$

The denominator

$$\langle 0 | \, T \left(e^{-i \int_{-\infty}^{\infty} H_I(t)dt} \right) | 0 \rangle$$

is nothing but exactly the sum over all disconnected pieces, and therefore we get

$$\langle \Omega | \, T \left(\phi_H(x_1), \dots, \phi_H(x_n) \right) | \Omega \rangle = \sum \text{connected Feynman diagrams}$$

(Notice that there are also diagrams of the form
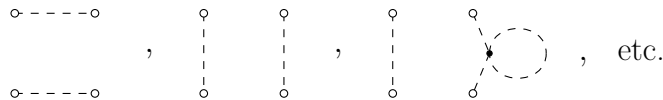
at all orders, which might on the surface seem like disconnected diagram, but in-fact are also connected diagrams, since every line is connected to some external line. These diagrams should also be included in the calculation of Green's functions.)

The numerator and denominator in the RHS are not something intrinsically meaningful for the theory. It depended on how we chose our vacuum. The fact that this vacuum was not the actual vacuum of the theory

is the reason why the disconnected pieces arose. In the Green's function, which is physically meaningful object, all these disconnected diagrams vanish, and only the connected diagrams remain.

10.2 Another way to interpret the Green's functions

Consider deforming the Hamiltonian by adding a source to the Hamiltonian

$$H \to H + \int \rho(x)\phi_H(x)d^3\mathbf{x}$$

which is equivalent to modifying the Hamiltonian density as

$$\mathcal{H} \to \mathcal{H} + \rho(x)\phi_H(x)$$

What this source does is to add some energy proportional to the field, with the proportionality constant depending on spacetime. We can ask now, in the presence of this source, what is the amplitude for the vacuum $|\Omega\rangle$ back to the original vacuum. To calculate this, we can treat the $\int \rho(x)\phi_H(x)d^3\mathbf{x}$ as the perturbation term, and obtain

$$\langle\Omega| S_\rho |\Omega\rangle = \langle\Omega| T \left(e^{i \int \rho(x)\phi_H(x)d^4x}\right) |\Omega\rangle = Z[\rho]$$

This is called the vacuum persistance amplitude in the presence of the source $\rho(x)$, and is sometimes also called the generating function for correlation functions.

We can expand the above as

$$Z[\rho] = \sum_n \int d^4x_1 d^4x_2 \ldots d^4x_n \frac{i^n}{n!} \rho(x_1)\rho(x_2)\ldots\rho(x_n) \langle\Omega| T \left(\phi_H(x_1)\phi_H(x_2)\cdots\phi_H(x_n)\right) |\Omega\rangle$$

(the time ordering doesn't care about the ρs since they are simply functions and not operators.)

Now if we take the functional derivative of this generating functional w.r.to the $\rho(x_i)$s,

$$(-i)^n \frac{\delta^n Z[\rho]}{\delta\rho(x_1)\delta\rho(x_2)\ldots\delta\rho(x_n)}\bigg|_{\rho(x)=0} = \langle\Omega| T \left(\phi_H(x_1)\phi_H(x_2)\cdots\phi_H(x_n)\right) |\Omega\rangle$$

(where the $n!$ cancelled out since there are $n!$ ways to act the functional derivatives on the n terms, and $\rho(x) = 0$ is important to kill off all the terms with power higher than n, which are not killed by the derivative)

Therefore, another interpretation of the Green's function is that it tells us how the theory responds to the insertion of a source. The $Z[\rho]$ is called the "Generating Function for Correlation Functions", since it generates the correlation functions for us.

10.3 LSZ Formula

Remember that in the full interacting theory, there is no notion of *one field*, but one is free to redefine the field operators while introducing interaction vertices, which keeps the theory unchanged. This is why we said that the above discussed Green's function was not a good physical question to ask, but rather we should discuss the $S-$matrix. The LSZ formula relates the Greens functions which we have been discussing so far to $S-$matrix elements. In Layman language, this formula simply tells to normalise the $\phi(x)$ correctly, strip off the external propagators, and then put the external momenta on-shell. In the following, we will discuss the same in precise language.

First we will have to *define carefully* what an $S-$matrix is. So far what we had was some matrix connecting some in-states to some out-states, but we want to make that notion more precise. First of all, let us impose that the fields we want to discuss about must satisfy

$$\langle \Omega | \phi_H(x) | \Omega \rangle = 0 \quad \leftarrow \quad \text{subtracting off the vacuum expectation value}$$

and

$$\langle k | \phi_H(x) | \Omega \rangle = e^{ik \cdot x} \quad \leftarrow \quad \text{normalising the field}$$

That is, if our original field operators did not satisfy this, then make a field redefinition such that the new fields $\phi_H(x)$ satisfy the above. We consider these two conditions since the vacuum and the one particle states are special states. The vacuum is the lowest energy state, and the one particle states are also uniquely specified by the conserved energy-momentum charges, and this is true even in the interacting theory. Notice that we have not yet specified what the states $|k\rangle$ are. The only requirement is that they satisfy $k^2 = m^2$.

To calculate the $S-$matrix elements, we need some in and out states, to create which, we want to now *extract* something that acts like a "creation" operator from the field $\phi_H(x)$. (This "creation" operator will not satisfy the free field algebra, but it will be some operator that will create one particle states). We can do this by using the so-called Klein-Gordon inner product. To do this, we start with $f(t, \mathbf{x})$ that satisfies the Klein-Gordon equation, i.e.

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2 \right) f(t, \mathbf{x}) = m^2 f(t, \mathbf{x})$$

We can now specify if this $f(t, \mathbf{x})$ has only positive or only negative frequencies, and we choose for it to have only negative frequencies. That is

$$f(t, \mathbf{x}) = \int \frac{d^3 \mathbf{k}'}{(2\omega'_\mathbf{k})(2\pi)^3} f(k') e^{-ik' \cdot x}$$

Now consider a smearing of the fields,

$$\phi_f(t) = i \int \left[\phi_H(t, \mathbf{x}) \partial_0 f(t, \mathbf{x}) - (\partial_0 \phi_H(t, \mathbf{x})) f(t, \mathbf{x}) \right] d^3 \mathbf{x}$$

This is called a Klein-Gordon inner product.

Now see that

$$\langle k| \phi_f(t) |\Omega\rangle = i \int \left[e^{ik\cdot x} \partial_0 f(t,\mathbf{x}) - i\omega_\mathbf{k} e^{ik\cdot x} f(t,\mathbf{x}) \right] d^3\mathbf{x}$$

substituting the expansion of $f(t,\mathbf{x})$, we get

$$\langle k| \phi_f(t) |\Omega\rangle = i \int \left[e^{ik\cdot x}(-i\omega_{\mathbf{k}'})e^{-ik'\cdot x} + e^{ik\cdot x}(-i\omega_\mathbf{k})e^{-ik'\cdot x} \right] f(k')d^3\mathbf{x} \frac{d^3\mathbf{k}'}{(2\pi)^3(2\omega_\mathbf{k}')}$$

The integral over $\mathbf{x}$ gives a delta function which sets $\mathbf{k} = \mathbf{k}'$. What remains in the exponents is $e^{i\omega_\mathbf{k} t - i\omega_\mathbf{k} t}$ which again cancel out since both $\mathbf{k}'$ and $\mathbf{k}$ are on-shell, and therefore we are left with

$$\langle k| \phi_f(t) |\Omega\rangle = f(k)$$

and this is independent of time.

A similar calcuation for

$$\langle\Omega| \phi_f(t) |k\rangle = i \int \left[e^{-ik\cdot x}(-i\omega_{\mathbf{k}'})e^{-ik'\cdot x} + e^{-ik\cdot x}(i\omega_\mathbf{k})e^{-ik'\cdot x} \right] f(k')d^3\mathbf{x} \frac{d^3\mathbf{k}'}{(2\pi)^3(2\omega_\mathbf{k}')} = 0$$

In a sense, the ϕ_f acts like a "creation" operator, which creates the wavepacket given by $f(k)$. However this is not a creation operator in the sense of a free field creation operator. That is, $\phi_f(t)$ can also create multi-particle states since the theory is an interacting theory. This is because, for an interacting field

$$\langle q| \phi_H(x) |\Omega\rangle \neq 0$$

where $|q\rangle$ is some multiparticle state satisfying $P_\mu |q\rangle = q_\mu |q\rangle$, since $\phi_H(x)$ doesn't only contain creation and annihilation operators, but will also have other operators that create multiple particles. Therefore simply smearing this field with some $f(t,\mathbf{x})$ will not end up creating only single particle states, but will also create the multiparticle states.

To be precise, let us compute

$$\langle q| \phi_f(t) |\Omega\rangle$$

To calculate this, notice that

$$\langle q| \phi_H(x) |\Omega\rangle = \langle q| e^{iP\cdot x}\phi_H(0)e^{-iP\cdot x} |\Omega\rangle = e^{iq\cdot x} \langle q| \phi_H(0) |\Omega\rangle$$

which is given from Lorentz invariance. Therefore

$$\langle q| \phi_f(t) |\Omega\rangle = i \int \left[e^{iq\cdot x}\partial_0 f(t,\mathbf{x}) - iq^0 e^{ik\cdot x} f(t,\mathbf{x}) \right] d^3\mathbf{x} \ \langle q| \phi_H(0) |\Omega\rangle$$

which gives

$$\langle q| \, \phi_f(t) \, |\Omega\rangle = i \int \left[e^{iq \cdot x}(-i\omega_{\mathbf{k}'})e^{-ik' \cdot x} + e^{iq \cdot x}(-iq^0)e^{-ik' \cdot x} \right] f(k') d^3\mathbf{x} \frac{d^3\mathbf{k}'}{(2\pi)^3(2\omega'_{\mathbf{k}})} \, \langle q| \, \phi_H(0) \, |\Omega\rangle$$

Once again we can do the integral over $\mathbf{x}$ to get a delta function $\mathbf{q} = \mathbf{k}$. However, now $q^0 \neq \omega_{\mathbf{q}}$ and therefore we will be left with

$$\langle q| \, \phi_f(t) \, |\Omega\rangle = \frac{\omega_{\mathbf{q}} + q^0}{2\omega_{\mathbf{q}}} e^{i(q^0 - \omega_{\mathbf{q}})t} f(q) \, \langle q| \, \phi_H(0) \, |\Omega\rangle$$

Now see that this is time dependent and not time independent. Notice that $q^0 > \omega_{\mathbf{q}}$, that is, for a given momentum, a multiparticle state has always more energy than a single particle state

$$q^0 = \sqrt{\mathbf{q}^2 + \sum m_i^2}, \quad \omega_{\mathbf{q}} = \sqrt{\mathbf{q}^2 + m^2}$$

and therefore, we get an oscillating phase factor which depends on time.

Physically speaking, $\phi_f(t)$ creates a single particle wavepacket specified by $f(k)$, and along with that it also creates multiparticle garbage. But we can now see what happens in the limit $t \to \pm\infty$. Suppose we take a smear the momenta states by considering a superposition of multiparticle states with different momenta and energy,

$$|\psi\rangle \approx \int d^3\mathbf{q} g(q) \, |q\rangle$$

then

$$\lim_{t \to \pm\infty} \langle\psi| \, \phi_f(t) \, |\Omega\rangle \propto \int d^3\mathbf{q} \, g(q) e^{i(q - \omega_{\mathbf{q}})t} = 0$$

Therefore, in the far past and far future, $\phi_f(t)$ does not create multiparticle states, and the only remaining state is the single particle wavepacket. Physically speaking, $\phi_f(t)$ creates a wavepacket of single particle states along with some multiparticle garbage states, but if we look at the state $\phi_f(t) \, |\Omega\rangle$ in some definite region in space time (i.e.,take its inner product with some definite state $|\psi\rangle$), and send the time of reaction to $-\infty$, all the multiparticle states will run away. Of course the one particle state may run away too. We prevent this by modifying the state we create in such a way that the single particle wavepacket always has the same relationship to the observer, and that's done by the funny combination $\phi_f(t)$. No multiparticle state has the right dispersion relation to keep the same relationship to the observer under this modification.

To create multiparticle states, we take two different functions f_1 and f_2, and we define a two particle state to be

$$|f_1, f_2\rangle_{\text{in}} = \lim_{t \to -\infty} \phi_{f_1}(t)\phi_{f_2}(t) \, |\Omega\rangle \quad _{\text{out}}\langle f_3, f_4| = \lim_{t \to \infty} \langle\Omega| \, \phi^\dagger_{f_3}(t)\phi^\dagger_{f_4}(t)$$

For most choices of f_1 and f_2, we find that the two wavepackets separate at far past and far future. This is because since the wavepackets are headed in different directions in space, in the far past and far future, they are separated well enough that acting $\phi_{f_2}(t)$ on $\phi_{f_1}(t) \, |\Omega\rangle$ will be the same as acting $\phi_{f_2}(t)$ on the

vacuum for a local observer.

Notice that even if $f_1 = f_3$ and $f_2 = f_4$, the in-states are not the same as out states, since the limits $t \to +\infty$ and $t \to -\infty$ of the states are not the same. It does indeed create two particle states, but it is a different two particle states in the in-states and the out-states for the same functions are different. These are the precise in and out states in the full quantum field theory, and the $S-$matrix will be the overlap between the two.

Therefore, the $S-$matrix elements are of the form

$$_{\text{out}} \langle f_3, f_4 | f_1, f_2 \rangle_{\text{in}} = \lim_{t_1, t_2 \to -\infty} \lim_{t_3, t_4 \to \infty} \langle \Omega | \phi_{f_3}^\dagger(t_3) \phi_{f_4}^\dagger(t_4) \phi_{f_1}(t_1) \phi_{f_2}(t_2) | \Omega \rangle$$

Note that the time ordering is already put in the formula. There is a subtlety regarding the individual time ordering of t_1 & t_2, that is, which limit to take first, $t_1 \to -\infty$ or $t_2 \to -\infty$ (and similarly for t_3, t_4), but this doesn't matter since the wavepackets are largely separated and therefore they commute.

One way to do this would be to explicitly solve for this in t and take the limit $t \to \pm\infty$. The LSZ reduction formula gives a convenient way to calculate this without having to go through all that hassle. The formula simply states (for the 4 particle state, this can be generalised to n-particle states too)

$_{\text{out}} \langle f_3, f_4 | f_1, f_2 \rangle_{\text{in}}$

$$= \int d^4 x_1 \ldots d^4 x_4 \ f_3^*(x_3) f_4^*(x_4) f_1(x_1) f_2(x_2) \prod_{j=1}^{4} i \left(\Box_j + m^2 \right) \langle \Omega | T \{ \phi_H(x_1) \phi_H(x_2) \phi_H(x_3) \phi_H(x_4) \} | \Omega \rangle$$

In momentum space

$$\Box_j + m^2 \equiv k_j^2 + m^2$$

which would cancel out the external propagators in the Greens function, iff it were a free theory. In the interacting theory it doesn't exactly cancel out the propagator, but the integration over the fs will put the external momenta on-shell since they are the solutions of the KG equation and they have exactly the momenta that are on-shell.

See that the Greens function has poles at all $k_i^2 = m^2$ if the k_i are on-shell, and therefore the above procedure simply tells to pick up the residues at these poles. Therefore, $S-$matrix is the residue at the poles of the Greens function when the momenta are put on-shell.

Proof—

Before proceeding to the proof, we will prove some intermediate results. First, we want to simplify the

following

$$i \int f(x)(\Box + m^2)\phi_H(x)d^4x = i \int f(x)\left(\partial_0^2 - \nabla^2 + m^2\right)\phi_H(x)d^4x$$
$$= i \int f(x)\partial_0^2\phi_H(x) - \left(\nabla^2 f(x)\right)\phi_H(x) + m^2 f(x)\phi_H(x)d^4x$$

We are allowed to do this since we assume that $f(x)$ is a wavepacket of finite extent, and therefore has no support at infinity, therefore the boundary terms are zero.

But $f(x)$ satisfies

$$\left(\partial_0^2 - \nabla^2 + m^2\right)f(x) = 0 \implies \left(\nabla^2\right)f(x) = (\partial_0^2 + m^2)f(x)$$

Therefore,

$$i \int f(x)(\Box + m^2)\phi_H(x)d^4x = i \int f(x)\partial_0^2\phi_H(x) - \left(\partial_0^2 f(x)\right)\phi_H(x)d^4x$$
$$= i \int \partial_0\left(f(x)\partial_0\phi_H(x)\right) - \partial_0\left(\phi_H(x)\partial_0 f(x)\right)d^4x$$
$$= \int_{-\infty}^{\infty}\partial_0\left(i \int \left(f(x)\partial_0\phi_H(x) - \phi_H(x)\partial_0 f(x)\right)d^3\mathbf{x}\right)dt$$
$$= \int_{-\infty}^{\infty}\partial_0\left(-\phi_f(t)\right)dt$$
$$= \phi_f(-\infty) - \phi_f(\infty)$$

We have a similar result for $f^*(x)$,

$$i \int f^*(x)(\Box + m^2)\phi_H(x)d^4x = \phi_f^\dagger(\infty) - \phi_f^\dagger(-\infty)$$

Notice that the different sign is due to the $i \to -i$ when $\phi_f \to \phi_f^\dagger$.

With these results, we now proceed to prove the LSZ formula.

$_{\text{out}}\langle f_3, f_4 | f_1, f_2 \rangle_{\text{in}}$

$$= \int d^4x_1 \ldots d^4x_4\ f_3^*(x_3)f_4^*(x_4)f_1(x_1)f_2(x_2)\prod_{j=1}^{4} i\left(\Box_j + m^2\right)\langle\Omega|\,T\left\{\phi_H(x_1)\phi_H(x_2)\phi_H(x_3)\phi_H(x_4)\right\}|\Omega\rangle$$

First, we want to move the $\Box + m^2$ in the LHS into the time ordering. But this is a subtle step since the time ordering has theta functions in time, and the box operator has derivatives w.r.to time.

Remember that the difference between one time ordering and another time ordering (i.e. $\phi(x)\phi(y)$ and $\phi(y)\phi(x)$) is just the commutator, and it vanishes as long as the points are spacelike separated, and gives something that looks like a delta function in $x - y$. Therefore, when we take the box operator into the time ordering, we will get some terms where the box operator is acting on the fields, and the other terms

will form commutators of the above kind, which are non-zero only in the region where x_i & x_j are close to each other. In the limit where the fs are orthogonal to each other (i.e. they do not have overlapping support), these terms will identically give 0. And therefore, in this limit, it is fine to move the derivatives into the time ordering.

A rough handwaving argument would be

$$(\partial_0)_x \left(\theta(x-y)\phi(x)\phi(y) + \theta(y-x)\phi(y)\phi(x) \right) = \delta(x-y)\phi(x)\phi(y) - \delta(x-y)\phi(y)\phi(x) + \text{derivative inside T.O.}$$
$$= \delta(x-y)\left[\phi(x), \phi(y)\right] + \text{derivative inside T.O.}$$

which implies

$$\left(\partial_0^2\right)_x \left(\theta(x-y)\phi(x)\phi(y) + \theta(y-x)\phi(y)\phi(x) \right) = \partial_0 \delta(x-y)[\phi(x), \phi(y)] + \delta(x-y)[\Pi(x), \phi(y)]$$
$$+\text{derivatives inside T.O.}$$

The extra terms give zero when the fs do not have overlapping support.

After moving the box operators inside and using the above derived relations, we get the following

$$_{\text{out}} \langle f_3, f_4 | f_1, f_2 \rangle_{\text{in}}$$
$$= \langle \Omega | T \left\{ (\phi_{f_1}(-\infty) - \phi_{f_1}(\infty))(\phi_{f_2}(-\infty) - \phi_{f_2}(\infty)) \left(\phi_{f_3}^\dagger(\infty) - \phi_{f_3}^\dagger(-\infty) \right) \left(\phi_{f_4}^\dagger(\infty) - \phi_{f_4}^\dagger(-\infty) \right) \right\} | \Omega \rangle$$

The product will expand to give $2^4 = 16$ terms, and for each term the time ordering pushes all the $+\infty$s to the left and $-\infty$s to the right. Let us consider one term of our interest,

$$\langle \Omega | \phi_{f_3}^\dagger(\infty)\phi_{f_4}^\dagger(\infty)\phi_{f_1}(-\infty)\phi_{f_2}(-\infty) | \Omega \rangle$$

which is exactly equal to

$$_{\text{out}} \langle f_3, f_4 | f_1, f_2 \rangle_{\text{in}}$$

Therefore, it should be true that all other 15 terms are zero. To see why, let us look at some other term

$$\langle \Omega | \phi_{f_1}(\infty)\phi_{f_2}(\infty)\phi_{f_3}^\dagger(-\infty)\phi_{f_4}^\dagger(-\infty) | \Omega \rangle$$

This term is zero, since $\phi_f^\dagger$ is an annihilation operator, and therefore gives zero. Another term would be of the form

$$\langle \Omega | \phi_{f_3}^\dagger(\infty)\phi_{f_2}(\infty)\phi_{f_1}(-\infty)\phi_{f_4}^\dagger(-\infty)\phi_{f_2}(-\infty) | \Omega \rangle$$

There is an subtlety in this, a question if one can commute the $\phi_{f_i}^\dagger$s with the ϕ_{f_j}s to annihilate the vacuum, and the answer is that we can indeed commute them since at $\pm\infty$ the wavepackets are well separated, therefore giving the above term also to be 0. Similarly all other 13 terms give 0, and we have proved the

92

LSZ formula.

There are other ways to derive this formula, which would roughly be to start with this definition

$$_{\text{out}} \langle f_3, f_4 | f_1, f_2 \rangle_{\text{in}} = \langle \Omega | \phi_{f_3}^\dagger(\infty) \phi_{f_4}^\dagger(\infty) \phi_{f_1}(-\infty) \phi_{f_2}(-\infty) | \Omega \rangle$$

and then argue that by taking the times to infinty, we pick the poles of the Green's function, and when we strip off the poles, we get the $S-$matrix. This approach can be found in Peskin and Schroeder, and we do not discuss this here.

A very important point — in the LSZ reduction formula, what is the ϕ_H that goes into its definition? There is no consensus on what should be the unique field in an interacting theory, and different field redefinitions give different interactions. In principle any redefinition is also a good definition for the theory, and in deriving the LSZ reduction formula, we never required that the ϕ_H was an elementary field, i.e. the field that enters the free Lagrangian. The LSZ formula works with any operators $Q(x)$, as long as they satisfy the normalisation condition

$$\langle k | Q(x) | \Omega \rangle = \text{e}^{ik \cdot x}$$

For instance, lets say we want to compute the $S-$matrix for some bound state. We do not have "fields" which create these bound states, but we can still compute the $S-$matrix for those bound states by the virtue that if the bound state is a stable state, then there should be some operator that has the above discussed overlap with the one-particle bound state, and we can calculate the $S-$matrix from that.

11 Loop Amplitudes and Renormalisation

Consider a simple second order diagram in the calculation of $4-$point function in the ϕ^4 theory.
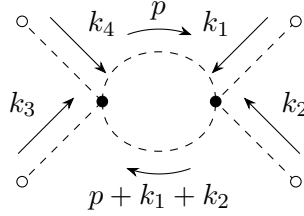
This will have an overall energy-momentum conserving delta function, but there is no restriction on p, that is, any value of p will satisfy the energy-momentum conservation. Therefore, the above diagram will have propagators

$$\int d^4 p \frac{1}{p^2 - m^2 + i\epsilon} \times \frac{1}{(p + k_1 + k_2)^2 - m^2 + i\epsilon}$$

and since there is no restrictions on p in the form of some delta functions, the above integral is divergant.

It is common in field theories that as we go to diagrams with loops, they diverge, and this divergence is not because we were wrong in our definitions, but rather it has to attribute to the fact that we are not being precise with what are the limits of our theory and what questions we can really ask. The quantum field theory we are discussing is, in a sense an *effective field theory* of an underlying "ultimate" description of the microscopic scales. (This statement is true in almost every physical theories, we do not model the flow of a liquid based on the movements of the individual particles in it, and similarly in the case of field theories). At long distances, the microscopic parameters combine in some way to give us some "physical" observables, and the effective field theory relates one set of physical observables to another set of physical observables. If we try to directly relate the physical observables we have, to the microscopic parameters, then the relationship will be complicated, giving all sorts of divergences. The field theory we have is a good description of physics at scales only larger than some length scale, and the above infinity arises because we wrongly assume that our effective field theory is valid also in scales smaller than this. To be precise, the above divergence arose because we allowed arbitrarily large momenta to circulate in our loop (remember that higher the energy of a process, the smaller the scale probed).

An important question one would therefore ask is how would the various parameters in the Lagrangian grow and decay with energy. As said, there might be some parameters in the complete theory, and what we see in the experiments will be some low energy effective parameters. Some basic dimensional analysis will allow us to determine what are the possible low energy effective parameters that we can consider in our field theories.

11.1 Dimensional Analysis

In the units we have chosen, $\hbar = c = 1$, since action has the same dimensions as $\hbar$, we have

$$[S] = 1 \implies [L] = M^{-1}$$

Therefore, for an arbitray interacting scalar field theory

$$S = \int d^4x \frac{1}{2}(\partial_\mu \phi(x))^2 - \frac{m^2}{2}\phi(x)^2 + \sum_{n \geq 3} \frac{\lambda_n}{n!}\phi(x)^n$$

should be dimensionless, and the first term gives

$$[x]^4[x]^{-2}[\phi]^2 = 1 \implies [\phi] = L^{-1} = M$$

Given this, we can now ask what are the dimensions of the various parameters that arise in the Lagrangian density. See that

$$[m] = M$$

and

$$[\lambda_n] = \frac{L^n}{L^4} = L^{n-4} = M^{4-n}$$

We now ask what happens to the different parameters at different energy scales. To answer the question of scales, we always need to consider dimensionless quantities, since the notion of large/small does not make sense for dimensionful quantities. Therefore, for a given parameter, at an energy scale $E \sim \Lambda$ we need to form a dimensionless quantities.

For m the dimensionless quantity would be

$$\frac{m}{\Lambda}$$

and this becomes small at high energies and large at low energies. This kind of parameter is called "relevant parameter".

For the other couplings, the dimensionless quantities are

$$\frac{\lambda_n}{\Lambda^{4-n}}$$

and we see that for $n = 3$ this is again a relevant parameter.

For $n > 4$, we see that the dimensionless parameter is

$$\lambda_n \Lambda^{n-4}$$

which becomes large at large energies and small at low energies, and this kind of parameter is called "irrelevant" since it does not have any effect on our effective field theoretical descriptions at low energies.

An example of irrelevant coupling in nature would be gravity, since

$$G_N \propto \frac{1}{L_{\text{Planck}}^2} \implies G_N \Lambda^2$$

is the dimensionless parameter.

In the case of $n = 4$, the parameter is dimensionless, and therefore this does not scale with energy at all, it is relevant at all energy scales. This is called "marginal parameter". Now when we add quantum effects, this properties can change, and the coupling can become marginally relevant or marginally irrelevant. We will see later that in the ϕ^4 theory, the coupling becomes marginally irrelevant.

11.2 Renormalisation

Theories that have marginal and relevant parameters are called *renormalisable theories*, while theories with irrelevant parameters are called *non-renormalisable theories*. Physically, this is because the parameters that grow large with energy are sensitive to what happens at high energies, and we therefore cannot define a *UV completion* of the theory by introducing some mathematical guesswork. But in the case for theories with only relevant and marginal parameters, we can define some UV completion for the theory, without affecting the details of the low energy theory at all. What renormalisation does is completes the UV physics by making some guess, which might not be the way the physical things works, while being rest assured that the low energy physics is unaltered. Operationally what this tells us is that the $\mathcal{L}$ is the bare Lagrangian, and we should adjust its parameters in the correct way to get the low-energy physics. That is, in

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

the parameters m and λ are not the parameters we can obtain from the low energy experiments. Therefore, we should write a counterterm lagrangian which might have different parameters, so that the above m and λ are now the parameters we measure in low-energy experiments.
For the ϕ^4 theory, the counterterm lagrangian would be

$$\mathcal{L}_{\text{CT}} = A\phi + \frac{B}{2}(\partial_\mu \phi)^2 - \frac{C}{2}\phi^2 + \frac{D}{4!}\phi^4$$

and we fix the parameters A, B, C, D by the low energy physical data. What are the possible physical data? For example, we could have (Note:- from now on $|0\rangle$ will unambigously mean the vacuum of the full Hamiltonian and not the free vacuum, since we will not invoke the free theory vacuum any more in our discussions.)

$$\langle 0| \phi(x) |0\rangle = 0, \quad \langle k| \phi(x) |0\rangle = e^{ik \cdot x}$$

and we can use these to fix the parameters A and B of the counterterm lagrangian. Further, we can require the mass of the excitation to be m which will fix C, and then we can also get the amplitude for some low energy scattering which would also fix D.

For the scalar Yukawa theory with

$$\mathcal{L} = |\partial_\mu \psi|^2 + \frac{1}{2}(\partial_\mu \phi)^2 - \frac{m^2}{2}\phi^2 - M^2|\psi|^2 - g\phi|\psi|^2$$

the counterterm Lagrangian would be

$$\mathcal{L}_{\mathrm{CT}} = A\phi + \frac{B}{2}(\partial_\mu \phi)^2 + \frac{C}{2}\phi^2 + D|\partial_\mu \psi|^2 + E|\psi|^2 - F\phi|\psi|^2$$

Notice that there is no term proportional to ψ, since a) the Lagrangian density should be a real function, and b) the theory has $SO(2)$ symmetry in $\psi - \psi^\dagger$, and the counterterm Lagrangian should follow the symmetry too.

The physical requirements area

- $\langle 0| \phi |0\rangle \, 0 \to \ A$

- $\langle q| \phi(x) |0\rangle = \mathrm{e}^{iq\cdot x} \to B$

- $\langle p| \psi(x) |0\rangle = \mathrm{e}^{ip\cdot x} \to D$

- Mass of $\phi = m \to C$

- Mass of $\psi = M \to E$

- g matches the scattering amplitude of some low energy scattering $\to F$

At the heart of this procedure is the idea that you split the bare parameters as

$$x_{\mathrm{bare}} = x_{\mathrm{ren}} + x_{\mathrm{C.T}}$$

and we impose physical conditions on x_{ren} which will give us also $x_{\mathrm{C.T.}}$. When we compute loops and include both the renormalised and C.T vertices, the divergent parts of the loop integrals are exactly canceled by the divergences in the CT insertions, leaving a finite result.

In the following text, unless explicitly specified, λ, m, g etc. are renormalised parameters.

Let us make some of these more formal and precise and then work out what the counter terms are. We will oscillate between the two models ϕ^4 and $\phi|\psi|^2$, using whichever is convenient to discuss the technicalities.

11.2.1   Condition on v.e.v. of $\phi(x)$

This condition lets us set the counterterm A. To do this, see that the vertex for it would be simply a linear vertex (we are using $\times$ to denote the A vertex)

$$- - - - - - \times$$

The condition that $\langle 0|\phi(x)|0\rangle$ simply implies that

All other graphs with
only one external leg

What this means is that we never have to calculate what A actually is, all A does is sets such graphs with only one external leg to zero.

A similar argument holds in the scalar Yukawa theory too, and in the Yukawa theory, the term sets the diagrams of the following kind to zero. (we did not give explicit example in ϕ^4 theory because there are no simple such diagrams in this theory)
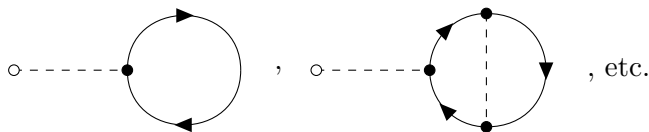

Such diagrams, which have only one external point and have multiple internal vertex are called tadpoles and our renormalisation procedure tells us that they can be neglected.

In the bare theory there also arises uncontrollable infinities in the form of the tadpole diagrams. As an example, consider the 3-point function in the scalar Yukawa theory, with one internal vertex. So far we have only seen the following diagram
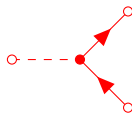
which was finite. But there are also divergant diagrams of the following form
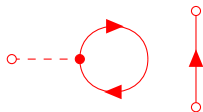
which contribute to the 3-point function.

Without the term $A\phi(x)$, the quantity $\langle 0|\phi(x)|0\rangle$ would be some infinity which is given by the sum of all the tadpole diagrams. But introducing the extra term linear in ϕ allows us to set the value of A such that the extra vertex that we get from $A\phi(x)$ exactly cancels out all these infinities, allowing us to set $\langle 0|\phi(x)|0\rangle = 0$. What this means for the perturbation expansion for the $S-$matrix is that one can conveniently ignore all the terms which have the tadpole diagrams knowing that the extra vertex due to A should also be included, and it will inevitable cancel these infinities.

11.2.2   Normalisation and Mass Condition - I

Lets now consider the other two conditions $\langle q|\phi(x)|0\rangle = e^{iq\cdot x}$ and that the mass m should be the physical mass. Before discussing this, we want to reframe these conditions in terms of the propagator

$$G(x,y) = \langle 0|\, T\left\{\phi(x)\phi(y)\right\}|0\rangle$$

We will prove two results about this propagator, the first one being the Källén-Lehmann spectral representation theorem, and the other being a concept of self-energy, and in terms of these we will discuss these renormalisation conditions.

<u>Källén-Lehmann spectral representation theorem</u>

First start with the above propagator and insert identity in the form of a complete set of eigenstates,

$$G(x,y) = \theta(x_0 - y_0)\sum_E \langle 0|\,\phi(x)\,|E\rangle\,\langle E|\,\phi(y)\,|0\rangle + \ y_0 > x_0 \text{ term}$$

There will be a vacuum state in the complete set, which will give zero owing to our above condition. The first contribution we get will come from single particle states, which would be

$$\sum_{\mathbf{k}} \langle 0|\,\phi(x)\,|k\rangle\,\langle k|\,\phi(y)\,|0\rangle + \ \text{ the other term}$$

We know what the measure for the momentum intergration is, and we can indeed perform this summation (integration). If the renormalisation condition we required above is to be true, then we know that this is simply

$$\int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}}\left(\theta(x_0 - y_0)e^{-ik(x-y)} + \theta(y_0 - x_0)e^{-ik(y-x)}\right)$$

which is just the usual free field propagator.

The other terms, which would be the contribution from the multi-particle states, would look like

$$\sum_{n\notin\text{s.p. states}}\left(\theta(x_0 > y_0)e^{-iQ_n(x-y)}|\,\langle 0|\phi(0)|n\rangle\,|^2 + \ y_0 > x_0 \text{ term}\cdots\right)$$

In the free field theory such terms never arose since the field ϕ had no overlap between the vacuum and multiparticle states. ϕ always created only single particle states in the free theory. But this is not the case in the interacting theory. We do not know how to write this since we do not know what the measure is, and we also do not know any conditions what the elements $\langle 0|\phi(0)|n\rangle$ should be.

We can put this in a nicer form by looking at the Fourier transform (calling $z = y - x$)

$$\sum_{n\notin\text{s.p. states}}\int d^4z\ e^{-ipz}\ e^{iQ_n z}|\,\langle 0|\phi(0)|n\rangle\,|^2 = \sum_{n\notin\text{s.p. states}}\delta^4(p - Q_n)\,|\,\langle 0|\phi(0)|n\rangle\,|^2$$

This only depends on p since everythin else is getting summed over. But this is a Lorentz invariant object, and therefore whatever p dependence exists should be of the form $\sigma(p^2)$, i.e.

$$\sum_{n \notin \text{s.p. states}} \delta^4(p - Q_n) \, |\, \langle 0|\phi(0)|n\rangle \,|^2 = \sigma(p^2)$$

Notice that since $Q_n^2 > m^2$, and therefore $\sigma(p^2)$ has no support for $p^2 < m^2$, i.e. $p^2 < m^2 \implies \sigma(p^2) = 0$. Further $\sigma(p^2)$ should be positive, since we are summing over the square of absolute values, which are all positive.

Therefore we get

$$\sum_{n \notin \text{s.p. states}} e^{-iQ_n(x-y)} |\, \langle 0|\phi(0)|n\rangle \,|^2 = \int e^{ipz} \sigma(p^2) \frac{d^4p}{(2\pi)^4}, \qquad \sigma(p^2) \geq 0, \ \sigma(p^2)|_{p^2 < m^2} = 0$$

We can now insert a 1 in the above term as

$$\int e^{ipz} \sigma(p^2) \, (2\pi)\delta(p^2 - a^2) \frac{da^2}{(2\pi)} \frac{d^4p}{(2\pi)^4}$$

and we can switch the order of integration, giving

$$\int \left(\int e^{ipz} \sigma(p^2) \, (2\pi)\delta(p^2 - a^2) \, \frac{d^4p}{(2\pi)^4} \right) \frac{da^2}{(2\pi)}$$

Since the $\sigma(p^2)$ depends only on p^2 and the delta function explicitly sets $p^2 = a^2$, we can pull it out of the integral to get

$$\int \sigma(a^2) \left(\int e^{ipz} \, (2\pi)\delta(p^2 - a^2) \, \frac{d^4p}{(2\pi)^4} \right) \frac{da^2}{(2\pi)}$$

Now notice that (eq 5.5) the inner integral is simply an integral over $d^3\mathbf{p}$ with the mass set to a, and we get

$$\int \sigma(a^2) \left(\int e^{ipz} \, \frac{d^3\mathbf{p}}{(2\pi)^3 2\sqrt{\mathbf{p}^2 + a^2}} \right) da^2$$

Therefore, combining this with the term with the other time ordering, we get a propagator but with mass a,

$$\int \sigma(a^2) \left(\int e^{ipz} \frac{i}{p^2 - a^2 + i\epsilon} \frac{d^4p}{(2\pi)^4} \right) da^2$$

and therefore we get the final result as

$$G(x, y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} \left[\frac{i}{p^2 - m^2 + i\epsilon} + \left(\int \frac{i\sigma(a^2)}{p^2 - a^2 + i\epsilon} da^2 \right) \right]$$

Sometimes, this is also written in a shorthand

$$G(x, y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} \int \frac{i\tilde{\sigma}(a^2)}{p^2 - a^2 + i\epsilon} da^2$$

100

where $\tilde{\sigma}(a^2) = \delta(a^2 - m^) + \sigma(a^2)$. Notice that there is only one delta function in $\tilde{\sigma}(a^2)$.

This is called the Källén-Lehmann spectral representation theorem. Our renormalisation condition fixes the coefficient and location of the delta function in $\tilde{\sigma}(a^2)$; it fixes that the position of the delta function to be at $a^2 = m^2$, and the coefficient to be 1.

Conversely $G(p^2)$ which is propagator in momentum space has pole at $p^2 = m^2$ with coefficient 1, and we can reformulate our renormalisation conditions in terms of this too.

Let us compute $G(p^2) = $ mom. rep. of $\langle 0|T(\phi_H(x)\phi_H(y)|0\rangle$ in perturbation theory in $\phi|\psi|^2$ theory. This is equal to

$$G(p^2) = $$

which is just as we had shown above, a free propagator plus some other terms. These other terms, in fact, have a simple structure as follows. We define the notion of a 1PI diagrams which stands for 1 Particle Irreducible diagrams. These are the diagrams which cannot be cut into two diagrams by cutting one line.
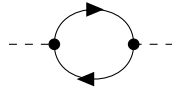
This is a 1PI diagram, but
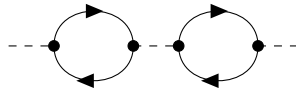
is not a 1PI diagram, since we can break the internal scalar propagator to get two diagrams. (Notice that we are not including the external propagators in the 1PI diagrams). We define $-i\Sigma(p^2)$ to be the sum of all 1PI diagrams

$$= -i\Sigma(p^2)$$

The propagator therefore looks like

Suppose the 1PI diagrams have values $a_1, a_2, \cdots$ such that $\sum_i a_i = S$, and let us call the pfree ropagator p. Then a general diagram in the full propagator would look like $p \cdot a_k \cdot p \cdot a_l \cdot p \cdot a_m \cdots p$. We can divide these into diagrams with one 1PI, diagrams with two 1PI, etc. The diagrams with one 1PI will all sum to

give pSp. The diagrams with two 2PI diagrams will sum to $p\sum_{ij}a_ipa_jp = pS^2p^2$, the diagrams with three 2PI diagrams will sum to $p\sum ijka_ipa_jpa_kp = pS^3p^3$ and so on.

and this can be summed to get

$$\frac{-i}{p^2 - m^2 + i\epsilon} + \left(\frac{-i}{p^2 - m^2 + i\epsilon}(-i\Sigma(p^2))\frac{-i}{p^2 - m^2 + i\epsilon}\right) + \left(\frac{-i}{p^2 - m^2 + i\epsilon}\left((-i\Sigma(p^2))\frac{-i}{p^2 - m^2 + i\epsilon}\right)^2\right) + \cdots$$

which forms a geometric series, whose sum looks like

$$\frac{\dfrac{-i}{p^2 - m^2 + i\epsilon}}{1 - \dfrac{\Sigma(p^2)}{p^2 - m^2 + i\epsilon}} = \frac{-i}{p^2 - m^2 + i\epsilon} \times \frac{p^2 - m^2 + i\epsilon}{p^2 - m^2 - \Sigma(p^2) + i\epsilon} = \frac{1}{p^2 - m^2 - \Sigma(p^2) + i\epsilon}$$

$\Sigma(p^2)$ is called the self energy, and this is sensitive to the loop diagrams, and our renormalisation requirement is to set the counterterm such that the propagator has a pole at $p^2 = m^2$ with residue 1. For $\Sigma(p^2)$, the pole condition simply reads $\Sigma(m^2) = 0$. We also need to see what the condition about the residue looks like for Σ. To do that, we can expand Σ around $p^2 = m^2$ to get

$$\Sigma(p^2) = \Sigma(m^2) + (p^2 - m^2)\frac{d}{dp^2}\Sigma(m^2) + \text{hots (higher order terms)}$$

Substituting this, we get the propagator to be

$$\frac{1}{(p^2 - m^2) - (p^2 - m^2)\dfrac{d}{dp^2}\Sigma(m^2) + \text{hots} + i\epsilon} = \frac{1}{p^2 - m^2}\frac{1}{1 - \dfrac{d}{dp^2}\Sigma(m^2) + \text{hots}}$$

giving the residue at $p^2 = m^2$ to be $(1 - \frac{d}{dp^2}\Sigma(m^2))$. Therefore, for the residue of the pole to be unchanged, we require

$$\frac{d}{dp^2}\Sigma(p^2)\bigg|_{p^2=m^2} = 0$$

That is, the counter terms should be chosen such that

$$\Sigma(m^2) = 0 \quad \& \quad \frac{d}{dp^2}\Sigma(p^2)\bigg|_{p^2=m^2} = 0$$

These conditions will allow us to set the counterterms $\dfrac{B}{2}(\partial_\mu\phi)^2$ and $\dfrac{C}{2}\phi^2$.

We will not compute the counterterms in the ϕ^4 theory since these terms do not appear at the first non-trivial order, since in ϕ^4, the first order loop correction to the propagator would be
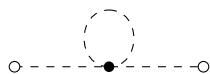
which is trivial since it has no dependence on the external kinematics. To get a non-trivial dependance, we need to go to the two loop diagrams of the form
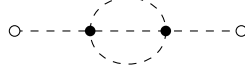
We can indeed calculate these counterterms in the $\phi|\psi|^2$ theory, where it has non-trivial one-loop contributions, and we will do it later in section 11.2.5

11.2.3 Condition on Interaction Coupling

This lets us fix the counterterm for the interaction $\dfrac{D}{4!}\phi^4$. What this requires is that the scattering amplitude for some physical values is equal to λ, i.e.

$= -i\lambda$ at some $s = s_0$, $t = t_0$ and $u = u_0$

(this diagram represents the full $4-$point Greens function summed to all orders in the perturbation series.) where in the above diagram we have considered the amputated Greens function, which is the full four point function with the external propagators stripped off. We could have considered the $S-$matrix, but $S-$matrix forces the external momen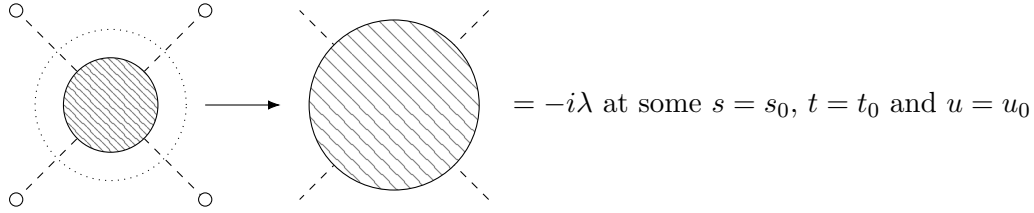ta to be on-shell. But by amputating, we are keeping the external momenta arbitrary and off-shell for our convenience. We will use this requirement to calculate the value of D in a perturbation theory in λ, and show how this will allow us to get finite result for one loop calculations.

At tree level, the above requirement is already met, and therefore D is zero at the leading order in lambda. Let us now consider the one loop diagrams in $4-$point functions for the ϕ^4 theory. There are three diagrams, which are
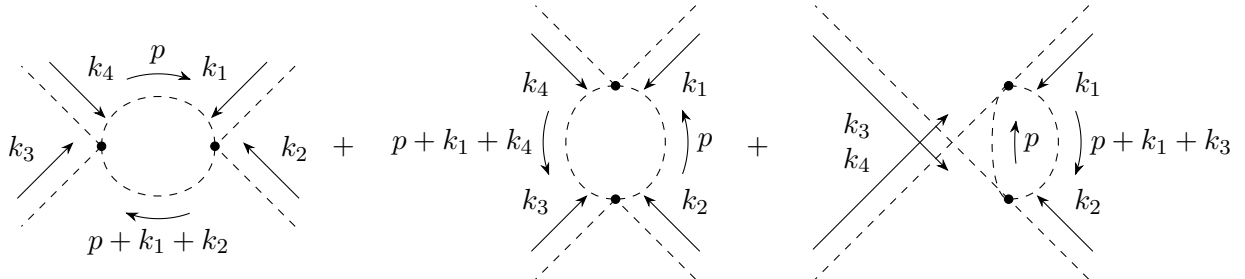
which are the s, t, and u channel diagrams respectively. Let us first look at the $s-$channel diagram. The other two answers will be very similar to this. From the Feynman rules ths $s-$channel diagram is equal to

$$\frac{1}{2}(-i\lambda)^2 \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} \frac{i}{(p + k_1 + k_2)^2 - m^2 + i\epsilon}$$

where the $1/2$ comes due to the symmetry factor in the exchange of $p \leftrightarrow p + k_1 + k_2$.

To evaluate this, we will use a *trick* which is one of the standard ways to evaluate this.

Since the denominators are different we use the *Feynman's parameter trick* which is based on

$$\frac{1}{ab} = \int_0^1 \frac{dx}{(ax + b(1-x))^2} =$$

we can check that this equation is true by explicitly doing the integral to get

$$\int_0^1 \frac{dx}{((a-b)x + b)^2} = \frac{1}{a-b} \left(\frac{-1}{ax + b(1-x)} \Big|_0^1 \right) = \frac{1}{a-b} \left(\frac{-1}{a} - \frac{-1}{b} \right) = \frac{1}{ab}$$

This is useful since we can write the above integral in p as

$$-\frac{1}{2}(-i\lambda)^2 \int dx \int \frac{d^4p}{(2\pi)^4} \frac{1}{((p^2 - m^2 + i\epsilon)x + ((p + k_1 + k_2)^2 - m^2 + i\epsilon)(1-x))^2}$$

The denominator can be simplified to get

$$p^2(x + (1-x)) - m^2(x + (1-x)) + i\epsilon(x + (1-x)) + (k_1 + k_2)^2(1-x) + 2p(k_1 + k_2)(1-x)$$
$$= p^2 - m^2 + i\epsilon + ((k_1 + k_2)^2 + 2p(k_1 + k_2))(1-x)$$

Since $(k_1 + k_2)^2 = s$, the above integral becomes (ignoring the dx integral for the time being)

$$-\frac{1}{2}(-i\lambda)^2 \int \frac{d^4p}{(2\pi)^4} \frac{1}{\{p^2 - m^2 + i\epsilon + (s + 2(k_1 + k_2))(1-x)\}^2}$$
$$= -\frac{1}{2}(-i\lambda)^2 \int \frac{d^4p}{(2\pi)^4} \frac{1}{\{(p + (k_1 + k_2)(1-x))^2 - m^2 + i\epsilon + s(1-x) - s(1-x)^2\}^2}$$

We can now do the variable change as $p + (k_1 + k_2)(1-x) \to p$, giving

$$-\frac{1}{2}(-i\lambda)^2 \int \frac{d^4p}{(2\pi)^4} \frac{1}{\{p^2 - m^2 + i\epsilon + s(x - x^2)\}^2} = -\frac{1}{2}(-i\lambda)^2 \int \frac{d^4p}{(2\pi)^4} \frac{1}{\{p^2 + a(x) + i\epsilon\}^2}$$

where $a(x) = -m^2 + s(x - x^2)$. Now, the integral over p^0 in this would be

$$-\frac{1}{2}(-i\lambda)^2 \int \frac{dp_0 d^3\mathbf{p}}{(2\pi)^4} \frac{1}{\{p_0^2 - \mathbf{p}^2 + a(x) + i\epsilon\}^2}$$

and we need to see where the poles are.

There are two possibilities.

1. If $-\mathbf{p}^2 + a(x) < 0$, then the poles of this integral are at $p_0 = \sqrt{\mathbf{p}^2 - a(x)} - i\epsilon$, $p_0 = -\sqrt{\mathbf{p}^2 - a(x)} + i\epsilon$.

2. If $-\mathbf{p}^2 + a(x) > 0$, then the poles are at $p_0 = i\sqrt{a(x) - \mathbf{p}^2} - \epsilon$, $p_0 = -i\sqrt{a(x) - \mathbf{p}^2} + \epsilon$. These are on the imaginary axis, but shifted by ϵ.
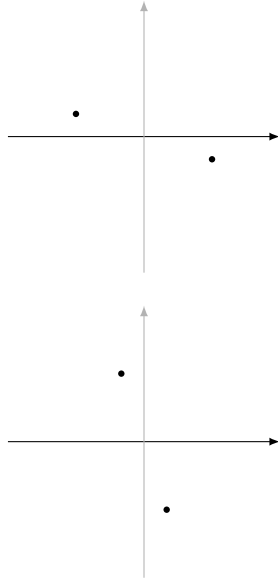
Since we are parforming the integral over x too, in the above integral we will get both these poles. Notice that, due to the way the poles are placed, in both the above cases, we are free to rotate the contour of integration (which is the real axis with some closure in either of the half plane) to now sit on the imaginary axis, since we can do this without crossing any poles. Therefore, p_0 now runs from $-i\infty$ to $i\infty$, and it is convenient to define $p_4 = -ip_0$, such that when p_0 goes from $-i\infty$ to $i\infty$, p_4 runs from $-\infty$ to ∞.
$p_0^2 = -p_4^2$, and $p^2 = p_0^2 - \mathbf{p}^2 = -p_4^2 - \mathbf{p}^2 = -p_{\text{Euclidean}}^2$
All the rotation manipulation we did above is called the *Wick rotation* which stands for the rotation from doing the Lorentzian integral to the Euclidean integral.

Notice that if the pole placement was different we would have gotten a different overall sign, since the allowed rotation in the case of a different pole placement would be the opposite of this case.

The full integral now looks like

$$-\frac{1}{2}(-i\lambda)^2 \int i\frac{d^4p_E}{(2\pi)^4}\frac{1}{\{-p_E^2 + a(x) + i\epsilon\}^2}$$

The Euclidean integral has spherical symmetry, and therefore we can go to polar coordinates in $4D$, but with a different volume of the unit sphere pulled out (in 3D case, the volume was 4π, and in $4D$, the volume is $2\pi^2$). Therefore, we get (where now $p_E \equiv |p_E|$)

$$\frac{i}{2}\lambda^2 \int 2p_E^3\frac{dp_E}{16\pi^2}\frac{1}{\{-p_E^2 + a(x) + i\epsilon\}^2}$$

Making a change of variables $p_E^2 = z \implies 2p_E dp_E = dz$, we get

$$\frac{i}{2}\lambda^2\frac{1}{16\pi^2}\int \frac{zdz}{\{-z + a(x) + i\epsilon\}^2}$$

105

So far what we did was exact manipulation, and we did not do anything to remove the divergance. This integral is still log divergant, at large z. But this form is very convenient for us to fix the divergance. Since we have a Euclidean integral, this integral can be thought of as put on some discrete Euclidean lattice, and the Euclidean momentum will be dual to the lattice spacing. What we want to do is to say that we are working in some finite lattice spacing, which in the momentum space translates to a cutoff in the integral at $z = \Lambda^2$. The motivation for the cutoff is all that we have been saying so far, which is that we don't know what the UV theory is, and that our assumption that the integral is valid all the way upto some arbitrarily high momenta is wrong since we do not know what happens at these arbitrarily high momenta. Whatever that happens at low energy should not care about what happens at high momenta, and therefore such a cutoff is completely valid in our *effective calculations*. Later we see that the cutoffs will drop out of all the physical results.

It is of importance that the cutoff is put after doing the Wick rotation. This is because in the Lorentzian integral the divergenes were not coming from some arbitrarily high momenta, but from the momenta that are very close to the light-cone.

With the cutoff, we can do the computation of the integral
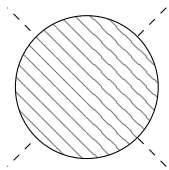
$$
\frac{i}{2}\lambda^2 \frac{1}{16\pi^2} \int \frac{(z - a(x) + i\epsilon + a(x) - i\epsilon)dz}{\{-z + a(x) + i\epsilon\}^2} = \frac{i}{2}\lambda^2 \frac{1}{16\pi^2} \int \frac{(z - a(x) - i\epsilon)}{\{-z + a(x) + i\epsilon\}^2} + \frac{a(x) + i\epsilon}{\{-z + a(x) + i\epsilon\}^2}dz
$$

$$
= \frac{i}{2}\lambda^2 \frac{1}{16\pi^2} \left(\ln(z - a(x) - i\epsilon) - \frac{a(x)}{z - a(x) - i\epsilon} \right)\Bigg|_0^{\Lambda^2}
$$

$$
= \frac{i}{2}\lambda^2 \frac{1}{16\pi^2} \left(\ln\left(\Lambda^2 - a(x) - i\epsilon\right) - \ln(-a(x) - i\epsilon) - \frac{a(x)}{\Lambda^2 - a(x) - i\epsilon} - 1 \right)
$$

(notice that in the numerator of a fraction, the $i\epsilon$ doesn't matter)

Therefore, the integral for the $s-$channel diagram is simply

$$
\int_0^1 dx \frac{i\lambda^2}{32\pi^2} \left(\ln\left(\Lambda^2 - a(x) - i\epsilon\right) - \ln(-a(x) - i\epsilon) - \frac{a(x)}{\Lambda^2 - a(x) - i\epsilon} + 1 \right)
$$

There are also integrals for the t and u channels. What we did so far did not care about what we called k_1 and k_2. Therefore, for the t and u channels, we get exactly the same answer, but with some different $a(x)$s. The full answer upto second order in perturbation theory therefore is

 $= -i\lambda + \sum_{c \in \{s,t,u\}} \int_0^1 dx \frac{i\lambda^2}{32\pi^2} \left(\ln\left(\Lambda^2 - a_c(x)\right) - \ln(-a_c(x)) - \frac{a_c(x)}{\Lambda^2 - a_c(x)} - 1 \right) + \mathrm{i}D$

where we have slightly redefined $a_c(x)$ to be

$$a_c(x) = c(x - x^2) - m^2 + i\epsilon \qquad \text{for } c \in \{s, t, u\}$$

and the iD contribution is from the counterterm.

We only included tree level contribution from the D vertex, since we calculate D order by order in λ. We already said that $D = \sum_{i=2}^{\infty} D^{(i)}$, $D^{(i)} \sim \lambda^i$, and the D we calculated above was just $D^{(2)}$. In regulating the diverging 3-vertex integrals, we will use diagrams with single $D^{(2)}$ insertion and tree level diagrams in $D^{(3)}$, and in regulating 4-vertex integrals in λ, we will use one-loop diagrams in $D^{(2)}$ and tree level diagram in $D^{(4)}$, and so on.

Notice that, importantly, $\Lambda \gg a_c(x)$, since the momentum scaling of a_c is s, t & u, and we required that our description to be of low energy physics. Therefore, we can ignore the terms of the form $a_c(x)\Lambda^{-2}$, and we can simply write the above result as

$$-i\lambda + \frac{i\lambda^2}{32\pi^2} 3\ln(\Lambda^2) + \sum_{c\in\{s,t,u\}} \int_0^1 \frac{i\lambda^2}{32\pi^2} \left(-\ln(-a_c(x))\right) dx - 3\frac{i\lambda^2}{32\pi^2} + iD + O\left(\frac{1}{\Lambda^2}\right)$$

Now our renormalisation condition requires that we should set D such that at some $s = s_0$, $t = t_0$, $u = u_0$,

$$\frac{i\lambda^2}{32\pi^2} 3\ln(\Lambda^2) + \sum_{c\in\{s_0,t_0,u_0\}} \int_0^1 \frac{i\lambda^2}{32\pi^2} \left(-\ln(-a_c(x))\right) dx - 3\frac{i\lambda^2}{32\pi^2} + iD = 0$$

This does not mean that the one loop contribution is zero. At some different values of s, t, and u, we will still get one loop contributions. What D does is it will cancel out the divergant $\ln\Lambda^2$ term in the result, but there will be a contribution from the $a_c(x)$ which would be different for different values of s, t, and u.

Therefore, for some other s, t, u, the full answer will be

$$-i\lambda - \sum_{c\in\{s_0,t_0,u_0\}} \int_0^1 \frac{i\lambda^2}{32\pi^2} \ln\left(\frac{a_c(x)}{a_c^0(x)}\right) dx$$

We defined the physical coupling constant at some scale, and we can calculate the amplitudes at some other scale using these physical coupling. This is in line with our philosophy that we should relate some physical observables to some other and not to the bare microscopic quantitites. That is, we do not know what are the parameters that enter the Lagrangian at the microscopic scale, and we can only get the effective parameters from low energy experiments, and we always need to compare values for different low energy experiments and never to the bare microscopic parameter.

11.2.4 Interlude — Regularisation methods

In the above calculation, we regularised our diverging loop integral by introducing a momentum cutoff. Introducing a momentum cutoff is the most simplest and the most physical way of regularising. But this is not the only way one can regularise divergances, and here we will discuss some other methods and schemes for regularising divergances.

One method of regularising is called the Pauli-Villars regularisation, in which we simply replace the free propagator by

$$\frac{1}{p^2 - m^2 + i\epsilon} \longrightarrow \frac{1}{p^2 - m^2 + i\epsilon} - \frac{1}{p^2 - \Lambda^2 + i\epsilon}$$

We are introducing another (ficticious) particle with mass Λ that propagates with a negative sign, and this has the same effect of cutting off the divergences at large p^2. Physically what this does is that if we go to very large value of p^2, the two terms will cancel out, and the propagator is basically zero, and therefore this regularisation does not allow the high momenta modes to propagate at all. When $p^2 \ll \Lambda^2$, the propagator is the usual propagator.

Another method is called dimensional regularisation, and this also corresponds to a cutoff. Suppose we have a log divergent integral of the form as we had before

$$\sim \int d^4p \frac{1}{(p^2 - \ldots)(p^2 - \ldots)}$$

we replace d^4p with $d^{4-\epsilon}p$. This is a trick that says that instead of working with 4 dimensions, let us work with some fractional $4 - \epsilon$ dimensions. Practically what this means is that if we Wick rotate the above integral and write in the spherical coordinates in $4D$, we get

$$\sim \int p^{3-\epsilon} dp \frac{1}{(p^2 - \ldots)(p^2 - \ldots)}$$

We see that at some fininte value of ϵ this is not divergant, since the integrand will now look like

$$\sim \frac{1}{p^{1+\epsilon}}$$

and this integral has a divergance at $\epsilon = 0$, but there is no divergance if $\epsilon > 0$

$$\int_{p_0}^{\infty} \frac{1}{p^{1+\epsilon}} dp = \frac{p^{-\epsilon}}{-\epsilon}\bigg|_{p_0}^{\infty} = \frac{p_0^{-\epsilon} - \infty^{-\epsilon}}{\epsilon}$$

and this is finite since $\infty^{-\epsilon} = 0$ for $\epsilon > 0$. If $\epsilon = 0$, then the above integral is rather

$$\int_{p_0}^{\infty} \frac{dp}{p} = \ln \infty - \ln p_0$$

which is diverging.

This looks like some mathematical trick, but this is also equivalent to introducing a momentum cutoff. When we replace $d^4p \to d^{4-\epsilon}p$, we need to ensure that we are dimensionally consistent, and therefore we need to multiply by a scale μ^ϵ

$$\int \mu^\epsilon p^{3-\epsilon} dp \frac{1}{(p^2 - \ldots)(p^2 - \ldots)} \sim \left(\frac{1}{\epsilon} + \text{ finite terms }\right) \mu^\epsilon$$

(open Peskin and Schroeder for a more detailed calculation).

Now, the divergant part looks like

$$\frac{\mu^\epsilon}{\epsilon} = \frac{e^{\epsilon \ln \mu}}{\epsilon} \simeq \frac{1}{\epsilon} + \ln \mu = \ln\left(\mu e^{\frac{1}{\epsilon}}\right) = \ln \Lambda^2$$

Which is exactly a cutoff $\Lambda^2 = \mu e^{\frac{1}{\epsilon}}$.

11.2.5 Normalisation and Mass Condition - II

In the previous discussion, we fixed $A\phi$ and $\frac{D}{4!}\phi^4$ terms, but we did not discuss $\frac{B}{2}(\partial_\mu \phi)^2$ and $\frac{C}{2}\phi^2$ terms. This is because in the ϕ^4, these terms do not have any contribution at one loop order, so we will look at the $\phi|\psi|^2$ theory to see how to fix these counterterms.

$$\mathcal{L}_{\text{CT}} = A\phi + \frac{B}{2}(\partial_\mu \phi)^2 + \frac{C}{2}\phi^2 + D|\partial_\mu \psi|^2 + E|\psi|^2 - F\phi|\psi|^2$$

In this model, the lowest order correction to the amputated two-point function of the ϕ field is at $O(g^2)$ and is the following,
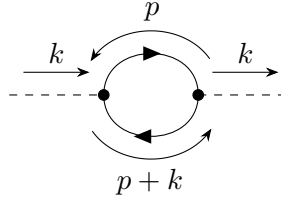
This is a 1PI diagram, and it evaluates to

$$-i\Sigma(k^2) = (-ig)(-ig) \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - M^2 + i\epsilon} \frac{i}{(p+k)^2 - M^2 + i\epsilon}$$

This integral is the same as the one we did before.

The counterterms B and C give the following vertices

.

We get an extra k^2 in the B vertex since there is a square of derivative which pulls down a factor of k^2.

Therefore the total corrections to the propagator at one loop would be

$$-i\Sigma(k^2) = (-ig)(-ig) \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - M^2 + i\epsilon} \frac{i}{(p+k)^2 - M^2 + i\epsilon} + iC + iBk^2$$

To fix B and C, we demand the requirements we derived above, i.e.

$$\Sigma(m^2) = 0, \quad \left. \frac{d}{dk^2}\Sigma(k^2)\right|_{k^2=m^2} = 0$$

A slight aside — The first condition came from the requirement that the physical propagator has pole at $p^2 = m^2$, and is stringent and cannot be traded for. But the second condition can be rephrased slightly. We could rather say that we want to work with a propagator of the form

$$\frac{Z}{p^2 - m^2 + i\epsilon}$$

Now this is equivalent to demanding that the field ϕ be such that $\langle k|\phi(0)|0\rangle = \sqrt{Z}$. This is called wave-function renormalisation, and this kind of modification will change the LSZ formula by some factors too. We will not work with this, but we will rather demand that the field be normalised correctly, but that will require the above condition in derivative to be true.

The above integral, after introducing a cutoff Λ is (where we have removed the terms of form $a(x)\Lambda^{-2}$, and have absorbed the constant term into C)

$$-i\Sigma(k^2) = \frac{ig^2}{16\pi^2} \int_0^1 dx \ln\left(\frac{\Lambda^2}{-a(x)}\right) + ik^2B + iC \tag{11.1}$$

where $a(x) = k^2(x - x^2) - M^2 + i\epsilon$

What follows is some algebra to set the values of B and C.
The first condition is

$$m^2 B + C + \frac{g^2}{16\pi^2} \int_0^1 dx \ln\left(\frac{\Lambda^2}{-m^2(x - x^2) + M^2 - i\epsilon}\right) = 0$$

$$\implies C = -m^2 B - \frac{g^2}{16\pi^2} \int_0^1 dx \ln\left(\frac{\Lambda^2}{-m^2(x - x^2) + M^2 - i\epsilon}\right)$$

$$\implies k^2 B + C = (k^2 - m^2)B - \frac{g^2}{16\pi^2} \int_0^1 dx \ln\left(\frac{\Lambda^2}{-m^2(x - x^2) + M^2 - i\epsilon}\right)$$

The second condition directly gives us B

$$B + \frac{g^2}{16\pi^2} \int_0^1 dx \frac{-k^2(x-x^2) + M^2 - i\epsilon}{\Lambda^2} \frac{\Lambda^2}{\{-k^2(x-x^2)+M^2-i\epsilon\}^2}(x-x^2)\bigg|_{k^2=m^2} = 0$$

$$\implies B = -\frac{g^2}{16\pi^2} \int_0^1 dx \frac{(x-x^2)}{-m^2(x-x^2)+M^2-i\epsilon}$$

We get

$$k^2 B + C = -\frac{g^2}{16\pi^2} \left((k^2 - m^2) \int_0^1 dx \frac{(x-x^2)}{-m^2(x-x^2)+M^2-i\epsilon} + \int_0^1 dx \ln\left(\frac{\Lambda^2}{-m^2(x-x^2)+M^2-i\epsilon}\right) \right)$$

If we substitute this back in eq 11.1, we see tha the $\ln \Lambda^2$s cancel, and what we are left with is some finite terms. We can write down the final answer for the self energy as

$$i\Sigma(k^2) = i\frac{g^2}{16\pi^2} \int_0^1 dx \ln\left(\frac{-m^2(x-x^2)+M^2-i\epsilon}{-k^2(x-x^2)+M^2-i\epsilon}\right) - i\frac{g^2}{16\pi^2}(k^2-m^2) \int_0^1 dx \frac{(x-x^2)}{-m^2(x-x^2)+M^2-i\epsilon}$$

which is a perfectly finite quantity. Notice that the momentum cutoff that we imposed drop out of the final answer.

When we have external legs, we never have to worry about the $\Sigma(k^2)$, since for the external legs, k^2 is on shell, i.e. $k^2 = m^2$, and therefore Σ is guaranteed to be zero. We only have to worry about the self-energy when we are considering internal legs. Even for the internal line, at leading order the propagator is

$$\frac{1}{p^2 - m^2 + i\epsilon}$$

but if we go to higher orders, we will have to consider

$$\frac{1}{p^2 - m^2 - \Sigma(p^2) + i\epsilon}$$

The Σ we calculated above is already pretty ugly even with only one-loop correction, and this is the reason why perturbation theory is hard, even with computers, and therefore is a reason why a lot of research is done on how to efficiently calculate them.

11.3 Loss of Scale Invariance

What we have done so far is defined a renormalisation perturbation theory using a scale. In the ϕ^4 theory, classically, we saw that the coupling constant was of the order $[\lambda] = M^0$ (see section 11.1). We might have concluded that the coupling constant has the same strength at high and low energies. But this is not true, since to define the quantum theory we had to introduce a cutoff. In fact, the strength of the coupling does vary as a function of the cutoff, and lets see how the variation happens, and to this extent lets define the beta function.

Lets work, for a moment, with the bare parameters λ. This λ is not the coupling constant of the *UV complete microscopic theory*, but rather the non-renormalised coupling constant in some theory with a cutoff Λ. If it were the bare parameter of a UV complete microscopic theory, then it would not depend on scale.

Say we define our theory at some lattice scale, i.e. lets say we have a lattice, we put our theory on the lattice, and then with the help of a computer solve the integrals numerically on the lattice, and got some long-distance physics. Now if I want to work on a finer or a coarser lattice, the question is how do I scale my coupling constant such that the new lattice with the new scale gives the same answers as the old lattice. In terms of the bare parameters, we can see this as following —

When we computed the $2-2$ scattering in ϕ^4, we found the answer to be

$$G_4 \sim -i\lambda + 3\frac{i\lambda^2}{32\pi^2}\ln\Lambda^2 - \frac{i\lambda^2}{32\pi^2}\sum_c \int_0^1 \ln(-a_c(x))dx$$

As we change Λ, how should we change λ to get the same predictions. To do this, see that G_4 is a physical thing, and therefore it is necessary that

$$\frac{dG_4}{d(\ln\Lambda)} = 0$$

and what this means is that

$$-i\frac{d\lambda}{d\ln\Lambda} + 6\frac{i\lambda^2}{32\pi^2} + 6\frac{i2\lambda}{32\pi^2}\ln\Lambda^2\frac{d\lambda}{d\ln\Lambda} = 0$$

see that when we solve this perturbatively, at leading order, we don't have to consider the third term, i.e. first we will solve the first two terms, and then we substitute it in the third term to get a better approximation and so on. Therefore, at leading order, we get

$$-i\frac{d\lambda}{d\ln\Lambda} + 6\frac{i\lambda^2}{32\pi^2} = 0$$

which we can solve as

$$\frac{d\lambda}{\lambda^2} = \frac{3}{16\pi^2}d\ln\Lambda$$
$$\implies \frac{1}{\lambda_i} - \frac{1}{\lambda_f} = \frac{3}{16\pi^2}\left(\ln\Lambda_f - \ln\Lambda_i\right)$$
$$\implies \frac{1}{\lambda_i} - \frac{1}{\lambda_f} = \frac{3}{16\pi^2}\ln\left(\frac{\Lambda_f}{\Lambda_i}\right)$$
$$\implies \lambda_f = \frac{\lambda_i}{1 - \lambda_i\frac{3}{16\pi^2}\ln\left(\frac{\Lambda_f}{\Lambda_i}\right)}$$

Now notice that at some point, as we move to higher cutoff, the denominator becomes zero. That is, the coupling constant behaves like

$$\ln \Lambda$$
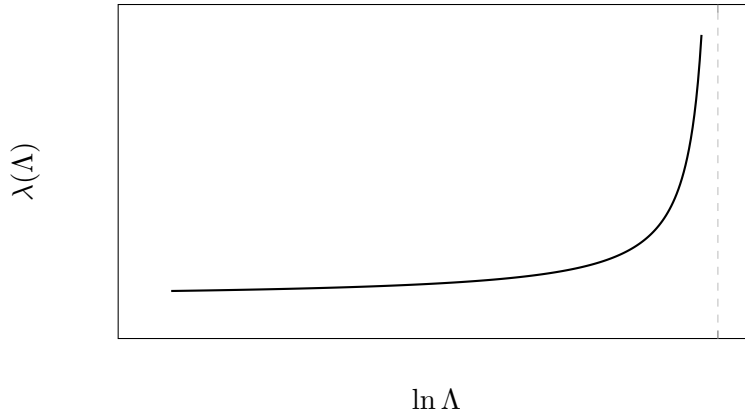
This kind of pole is called a Landau pole, and it tells us that for this ϕ^4 theory, as we go to larger scales, the theory becomes strongly coupled, and there is no good continuum limit (unless something very subtle happens at large Λ). We can still define an *effective field theory* that we did so far, but what this tells is that the theory does not have a UV completion, i.e. it is not valid even at UV scale.

The equation that defines the rate of change of λ with $\ln \Lambda$

$$\frac{d\lambda}{d \ln \Lambda} = \frac{3\lambda^2}{16\pi^2}$$

is called the beta function.

The beta functions of many other theories had Landau poles, and this is the reason why people initially thought that QFTs were not good theories. Famously, in the case of QCD, the beta function has the opposite sign, and therefore they have a good continuum limit, where as we go to larger scales, the coupling constant gets weaker. There are some theories, the beta function vanishes, and the coupling constant does not change as a function of scale at all. These are called conformal field theories.

Even when doing counterterm renormalisation, the renormalised parameters do change when chaning the cutoff. To see why, consider the following —
The UV complete theory has the Lagrangian $\mathcal{L} = \sum_i g_i O_i$ where g_i are the bare parameters, and O_i are the local operators. In doing counterterm renormalisation, we split the g_i as $g_i = g_i^{ren} + g_i^{ct}$, and fine tuned the g_i^{ct} such that they cancel out the infinities that arise from g_i^{ren}. As we saw in the previous discussion, the counterterms depended on the cutoff Λ. But since the bare parameters should be independent of this artificial cutoff we introduced, the g_i^{ren} should also be a function of Λ such that $g_i^{ren}(\Lambda) + g_i^{ct}(\Lambda) = g_i^{bare}$ is not a function of Λ. Therefore, we see that the renormalisable parameters also vary with the cutoff. But it is necessary that low energy observables that do not depend on high momentum modes should not change when we shift the cutoff Λ, and it is this exact requirement that lead to the β function.

113

11.4 The Wilsonian Perspective of Renormalisation

(A good reference — arXiv:9210046)

To understand the physics behind the procedure of renormalisation, and why the renormalisibility of an interaction depended on the mass dimension of the coupling constant, it is instructive to look at the Wilsonian perspective.

The Wilsonian perspective is to consider a Lagrangian

$$\mathcal{L} = \sum_i g_i O_i$$

and take seriously the idea that we have a cutoff Λ. When there is cutoff, by construction, there are no infinities, however we can now ask how the parameters g_i change as we *flow* the cutoff. To answer this in a more formal way, we need to consider the path integral approach (which will be discussed in the upcoming sections). The path integral approach to calculating expectation value of some operator A is to calculate the integral (this is in Euclidean space)

$$\int e^{-S_\Lambda} A \, \mathfrak{D}\phi_{<\Lambda}$$

The integral measure $\mathfrak{D}\phi$ can be written in momentum space too, in which it forms an integral over fields of different momentum modes, and we impose a cutoff Λ by simply integrating over modes that have momentum less than Λ. Now let us say we do a part of the integral over the modes that lie between Λ and $\Lambda' < \Lambda$. I.e. we can decompose our fields as $\phi_{<\Lambda'} + \phi_{>\Lambda',<\Lambda}$, in which case the measure factors too, giving

$$\int \int e^{-S_\Lambda} A \, \mathfrak{D}\phi_{>\Lambda',<\Lambda} \mathfrak{D}\phi_{<\Lambda'}$$

If the observale A does not depend on modes between Λ' and Λ, then one can write the above integral equivalently as

$$\int e^{-S_{\Lambda'}} A \mathfrak{D}\phi_{<\Lambda'}$$

where

$$e^{-S_{\Lambda'}} = \int e^{-S_\Lambda} \, \mathfrak{D}\phi_{>\Lambda',<\Lambda}$$

(Notice that we can only flow the cutoff inwards, we can never flow it outwards.)

The action at cutoff Λ' is not necessarily the same as the action at cutoff Λ, and therefore, at a lower cutoff, we need different parameters in the action so as to obtain the same physical answers. This process defines a flow in the space of parameters as a function of Λ, and this is the renormalisation group flow.

For a scalar field theory which has $[\phi] = M = L^{-1}$, some operator O_i that has M ϕ's and N derivatives, has dimensions $[O_i] = M + N = \delta_i$, and therefore the g_i should have $[g_i] = 4 - \delta_i$. For this, we write the

flow equation in the form

$$\Lambda \frac{dg_i}{d\Lambda} = \Lambda^{4-\delta_i} \beta_i(g_i \Lambda^{\delta_i-4})$$

where the factors of Λ are set to make everything dimensionally consistent. Now suppose we consider a dimensionless parameter,

$$\lambda_i = g_i \Lambda^{\delta_i - 4}$$

this will satisfy the equation

$$\Lambda \frac{d\lambda_i}{d\Lambda} = (\delta_i - 4)\lambda_i + \beta_i(\lambda_i) \tag{11.2}$$

β will be a perturbative function of, in principle, all λ_is, and therefore the leading term would be the first term. In the case O_i has dimension $\delta_i > 4$, we get a leading term term that damps the flow of λ_i as we lower the cutoff, and vice versa for $\delta_i < 4$. That is, the flow of relevant parameters grow in size as we lower the cutoff and vice versa for irrelevant. Therefore, what remains at low energies are only marginal and relevant parameters, the irrelevant parameters flow to zero as we go to lower energies. At low cutoffs, we always see effective field theories with only marginal and relevant parameters. The surprising fact that the standard model is renormalisable is a consequence of this statement.

One might think that if we start with irrelevant parameters at very high energies, and then flow to lower energies, these irrelevant parameters do not matter at all. But this is not the case, since the irrelevant parameters can still alter the flow of the other parameters. As an example, consider a theory with λ_4 and λ_6 interactions. At low energies, the λ_6 goes to zero, and one might assume that a theory where we started off with only λ_4 will match the theory we have. But this is not the case since at high energies λ_6 is large and will therefore affect the flow of λ_4 at high energies, and therefore the theory that we get from flowing the complete theory to some low energy will be manifestly different from the theory that we get if we started off with only λ_4.

For a marginal coupling, the leading order term in eq (11.2) is already zero, and the behaviour as we increase the cutoff or decrease the cutoff wholly depends on the β function. To calculate the β function for the marginal coupling, it is *easy* to calculate as we did in the previous section, using the non-renormalised Lagrangian as opposed to using the counter-term renormalised Lagrangian. We could of course use the renormalised Lagrangian, but it is easier without it.

12   An Aside — Dualities in Quantum Field Theories

(The following is the notes for two lectures by Prof. Subhro Bhattacharjee which are complementary to the other lectures given by Prof. Suvrat Raju)

To discuss dualities in quantum field theories, let us first re-discuss how a quantum field theory is defined. Quantum field theory is about trying to understand how laws of QM apply for *many degrees of freedom*, which might be interacting with each other. These many degrees of freedom could be, in-principle, anything, they could be particle fields, or electric and magentic fields, or even spacetime itself, OR it could be some bunch of electrons and ions which condensed matter physicists study. And here, we will try to get a point of view from the condensed matter side, which is often called "Many-Body Physics". This is essentially a different name for quantum field theory, where now the degrees of freedom are particular objects like electrons, etc.

To define a QFT, we need to specify the degrees of freedom. The question, how do we go on to define degrees of freedom, is a very intricate question and is at the heart of QFT. Once we know the degrees of freedom, we can put some quantum mechanics in these degrees of freedom, which we already know how to do, and therefore, the fundamental question is how to define these degrees of freedom. One way to define many-body degrees of freedom is to take the space that contains these many-body degrees of freedom (which might be a box of gas, or a box of electromagnetic radiation, etc.) and divide it into small parts a_i at length scale a which defines the smallest length scale for the theory (this is equivalent to introducing a largest momentum scale, and as we discussed in the previous sections, one requires to specify a cutoff to define a QFT)
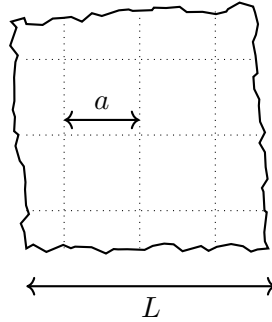
Once we define these parts, we can define Hilbert spaces associated with each part

$$\mathbf{H}_i = \{|\alpha_j^i\rangle \mid j = 1, \cdots, n_i\}$$

where n is the dimension of the Hilbert space associated with i^{th} part, and we can have a quantum mechanics in each part. We can get the complete Hilbert space as

$$\mathbf{H} = \otimes \mathbf{H}_i$$

The length scale a (also called the UV scale) can be ad-hoc and be dictated by experiments what we want to work with, or in some theories like spin lattice systems they can also be fundamental to the theory itself. There is another length scale associated with the system which is L that denotes the system size, and this is called the IR length scale.

Once we define the length scale a and L and do the above procedure, we have defined the degrees of freedom *at the scale a and L*. Given two length scales $a \ll L$, there is a length scale $a \ll l \ll L$ at which we are probing the system. The whole richness of QFT comes due to the fact that the description of the physical system can change depending on the length scale l. This is the non-trivial part that never occurs in Quantum Mechanics. Most Quantum Field Theories are best described in the regime

$$\frac{l}{a} \gg 1, \qquad \frac{l}{L} \ll 1$$

which is called the scaling limit. There are two ways of taking this limit for a given theory. In the case of field theories or electromagnetic systems, we take $a \to 0$ while keeping L fixed to get the entire limit, and this is called the continuum limit. In the other case, we take $L \to \infty$ while keeping a fixed, and this is called Thermodynamic limit, and this is done in statistical mechanics and condensed matter studies.

A very interesting fact is that even at the same length scale, QFTs may have two different representation, i.e. there may be two different actions with completely different set of variables describing the same quantum field theory. Suppose I have an action

$$S_{[a,L,l]}[\phi]$$

It turns out that there can be a completely different action in completely different fields

$$\tilde{S}_{[a,L,l]}[\tilde{\phi}]$$

such that for every correlation function I can calculate in S, there is a corresponding correlation function in $\tilde{S}$ with a one-to-one mapping, although these QFTs are quite different. These mappings are extremely useful especially in the cases where calculating correlation functions in one theory might be difficult, but will be extremely easy in the other. Such mappings are called dualities, and it is these that we will be discussing here.

12.1 Ising Model - Ising Electromagnetism Duality

This model consists of a lattice of spacing a, with a spin-1/2 particle placed at each lattice point. Therefore, at each lattice point, there are basis states $|\uparrow\rangle$ & $|\downarrow\rangle$, with the operators S^z, S^x, S^y. The Hamiltonian is given by

$$\mathcal{H} = -\sum_{i,j \in \text{neighbours}} S_i^z S_j^z - \Gamma \sum_i S_i^x$$

117

where $i \equiv \vec{i} = (i_x, i_y)$.

The model with $\Gamma = 0$ has a phase transition at some temperature T_c such that $T < T_c \implies \langle S^z \rangle \neq 0$ and $T > T_c \implies \langle S^z \rangle = 0$. At $\Gamma = 0$ & $T = 0$, the state of the system is $|\uparrow, \uparrow, \ldots, \uparrow\rangle$ or $|\downarrow, \downarrow, \ldots, \downarrow\rangle$ while at higher temperature there will be some spins flipped, and beyond the critical temperature the spins are 50-50 up/down.

We are interested in what happens if we increase Γ with $T = 0$. At $\Gamma \to \infty$, we can neglect the first term in Hamiltonian, and therefore the state at $T = 0$ is $|\uparrow_x, \uparrow_x, \ldots, \uparrow_x\rangle$. Therefore, as we increase Γ, there is another phase transition.

One can study the phase transition, etc. for this model, but we won't do that here. Rather, we will write down an action, and we show that there is a completely different action that will describe the same system. In this case, it turns out that the dual theory is *Ising Gauge Theory/Ising Electromagnetism*. There aren't many examples of such dualities, and here we will work out in-detail the transformations that will take us from one theory to the other.

12.1.1 Deriving the Action

We can derive the action by making use of the fact that the partition function $\mathcal{Z} = \mathrm{Tr}\left(e^{-\beta H}\right)$ is equivalent to a path integral in imaginary time.

$$\mathcal{Z} = \mathrm{Tr}\left(e^{-\beta H}\right) = \sum_{\{s^z\}} \langle \{s^z\}| e^{-\beta H} |\{s^z\}\rangle$$

where $\{s^z\}$ is the basis vector for the entire lattice in $z-$basis. Writing $\beta = \epsilon N$, where $N \to \infty, \epsilon \to 0$ gives the continuum limit,

$$\mathcal{Z} = \sum_{\{s^z\}} \langle \{s^z\}| \prod_1^N e^{-\epsilon H} |\{s^z\}\rangle = \sum_{\{s^z\}} \langle \{s^z\}| e^{-\epsilon H} \ e^{-\epsilon H} \cdots e^{-\epsilon H} |\{s^z\}\rangle$$

Inserting identity in the basis between every $e^{-\epsilon H}$, we get

$$\mathcal{Z} = \sum_{\{s_1^z\}} \sum_{\{s_2^z\}} \cdots \sum_{\{s_n^z\}} \langle \{s_1^z\}| e^{-\epsilon H} |\{s_2^z\}\rangle \langle \{s_2^z\}| \ e^{-\epsilon H} |\{s_3^z\}\rangle \cdots \langle \{s_n^z\}| e^{-\epsilon H} |\{s_1^z\}\rangle$$

One of the terms in the above equation looks like

$$\langle \{s_\tau^z\}| e^{-\epsilon H} |\{s_{\tau+1}^z\}\rangle$$

If Γ was zero, then the above term would be easy to calculate, since the identity insertion would be eigenbasis of $\mathcal{H}$. It is the Γ term that makes the problem quantum, since it is this term that brings a

non-trivial commutation relation.

$$\langle\{s_\tau^z\}|\,\mathrm{e}^{-\epsilon H}\,|\{s_{\tau+1}^z\}\rangle = \mathrm{e}^{\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z}\,\langle\{s_\tau^z\}|\,\mathrm{e}^{-\epsilon\Gamma\sum_i s_i^x}\,|\{s_{\tau+1}^z\}\rangle$$

where we could ignore the commutator terms in splitting the exponential since they are of the order ϵ^2. Now, we can introduce an identity in s^x basis as

$$\sum_{\{s_\tau^x\}}\mathrm{e}^{\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z}\,\langle\{s_\tau^z\}|\,|\{s_\tau^x\}\rangle\,\langle\{s_\tau^x\}|\,\mathrm{e}^{-\epsilon\Gamma\sum_i s_i^x}\,|\{s_{\tau+1}^z\}\rangle$$

$$=\sum_{\{s_\tau^x\}}\mathrm{e}^{\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z}\,\langle\{s_\tau^z\}|\{s_\tau^x\}\rangle\,\mathrm{e}^{-\epsilon\Gamma\sum_i s_i^x}\,\langle\{s_\tau^x\}|\{s_{\tau+1}^z\}\rangle$$

To calculate the inner products that are left, see that

$$\langle\uparrow_x | \uparrow\rangle = \tfrac{1}{\sqrt{2}} \quad \langle\downarrow_x | \uparrow\rangle = \tfrac{1}{\sqrt{2}}$$
$$\langle\uparrow_x | \downarrow\rangle = \tfrac{1}{\sqrt{2}} \quad \langle\downarrow_x | \downarrow\rangle = -\tfrac{1}{\sqrt{2}}$$

which can be written in a compact form as

$$\langle s^x|s^z\rangle = \mathrm{e}^{i\frac{\pi}{4}(1-s^x)(1-s^z)}$$

(see that the exponent is 1 when either s^x or s^z is $up(+1)$, and is -1 only when both are *down*).

Therefore we have

$$\langle\{s_\tau^z\}|\,\mathrm{e}^{-\epsilon H}\,|\{s_{\tau+1}^z\}\rangle = \sum_{\{s_\tau^x\}}\mathrm{e}^{\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z}\mathrm{e}^{-\epsilon\Gamma\sum_i s_i^x}\mathrm{e}^{i\frac{\pi}{4}\sum_i(1-s_{i\tau}^x)(1-s_{i\tau}^z)}\mathrm{e}^{-i\frac{\pi}{4}\sum_i(1-s_{i\tau}^x)(1-s_{i(\tau+1)}^z)}$$

$$=\sum_{\{s_\tau^x\}}\exp\left(\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z - \epsilon\Gamma\sum_i s_i^x + i\frac{\pi}{4}\sum_i(1-s_{i\tau}^x)(s_{i(\tau+1)}^z - s_{i\tau}^z)\right)$$

where now everything in the exponent is simply a number, and there are no operators.

We can add and subtract 1 to the Γ term to make it match the last term as (not worrying about any constants we add/subtract since it doesn't change the physics)

$$=\sum_{\{s_\tau^x\}}\exp\left(\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z + \epsilon\Gamma\sum_i(1-s_i^x) + i\frac{\pi}{4}\sum_i(1-s_{i\tau}^x)(s_{i(\tau+1)}^z - s_{i\tau}^z)\right)$$

$$=\sum_{\{s_\tau^x\}}\exp\left(\epsilon\sum_{ij}s_{i\tau}^z s_{j\tau}^z + \sum_i\frac{(1-s_i^x)}{2}\left(2\epsilon\Gamma + i\frac{\pi}{2}(s_{i(\tau+1)}^z - s_{i\tau}^z)\right)\right)$$

The second term is diagonal in i and therefore we can write it as

$$\prod_i \sum_{\{s_\tau^x\}} \exp\left(\frac{(1-s_i^x)}{2}\left(2\epsilon\Gamma + i\frac{\pi}{2}(s_{i(\tau+1)}^z - s_{i\tau}^z)\right)\right)$$

and this sum over $s^x = \{1, -1\}$, we can do explicitly. All these manipulations were done with a sole purpose of reaching a form where we can perform the sum explicitly.

Notice that when $s_i^x = +1$, the exponential is simply 1, and therefore the above sum is simply

$$\prod_i \left(1 + \exp\left(2\epsilon\Gamma + i\frac{\pi}{2}(s_{i(\tau+1)}^z - s_{i\tau}^z)\right)\right)$$

Now

$$\exp\left(i\frac{\pi}{2}(s_{i(\tau+1)}^z - s_{i\tau}^z)\right) = \begin{cases} +1 & \text{if } s_{i(\tau+1)}^z = \pm 1 \text{ and } s_{i\tau}^z = \pm 1 \\ -1 & \text{if } s_{i(\tau+1)}^z = \pm 1 \text{ and } s_{i\tau}^z = \mp 1 \end{cases} = s_{i(\tau+1)}^z s_{i\tau}^z$$

Rewriting

$$e^{2\epsilon\Gamma} = \tanh J_\tau = \frac{\sinh J_\tau}{\cosh J_\tau}$$

we get the above term as

$$\prod_i \left(1 + \exp\left(2\epsilon\Gamma + i\frac{\pi}{2}(s_{i(\tau+1)}^z - s_{i\tau}^z)\right)\right) = \frac{1}{\cosh J_\tau}\prod_i \left(\cosh J_\tau + \sinh J_\tau s_{i(\tau+1)}^z s_{i\tau}^z\right)$$

Again

$$\cosh J_\tau + \sinh J_\tau s_{i(\tau+1)}^z s_{i\tau}^z = \begin{cases} e^{J_\tau} & \text{if } s_{i(\tau+1)}^z s_{i\tau}^z = 1 \\ e^{-J_\tau} & \text{if } s_{i(\tau+1)}^z s_{i\tau}^z = -1 \end{cases} = e^{J_\tau s_{i(\tau+1)}^z s_{i\tau}^z}$$

Putting all this together, for a single time slice we get

$$\langle\{s_\tau^z\}|\, e^{-\epsilon H}\,|\{s_{\tau+1}^z\}\rangle = \exp\left(\epsilon \sum_{ij} s_{i\tau}^z s_{j\tau}^z + \sum_i J_\tau s_{i(\tau+1)}^z s_{i\tau}^z\right)$$

where in the last term we conveted the product of exponentials to exponential of sum.

Now using this, we get the partition function as

$$\mathcal{Z} = \sum_{\{s_1^z\}}\sum_{\{s_2^z\}}\cdots\sum_{\{s_n^z\}} \exp\left(\epsilon \sum_{\tau,ij} s_{i\tau}^z s_{j\tau}^z + J_\tau \sum_{\tau,i} s_{i(\tau+1)}^z s_{i\tau}^z\right)$$

Notice that the sum in the above equation is equivalent to a sum over all possible configurations in a 3D lattice with the third dimension as τ, this is a path integral. Further, see that the first term in the exponential has the coupling between sites on a same time slice, while the second term has the coupling

between sites on neighbouring time slice on a same spatial point. Therefore, we can write the above equation compactly as

$$\mathcal{Z} = \sum_{\{\{s^z\}\}} \exp\left(\sum_{IJ \in \{\text{nearest neighbours}\}} J_{IJ} s_I^z s_J^z\right) = \sum_{\{\{s^z\}\}} \exp\left(-\mathcal{S}\right)$$

where now $I \equiv (i_x, i_y, \tau)$. We have made the coupling index dependent since depending on temporal or spatial coupling, we have J_τ or ϵ as the coupling. Therefore, the action for the 2D Ising model is simply

$$\mathcal{S} = \sum_{IJ \in \{\text{nearest neighbours}\}} -J_{IJ} s_I^z s_J^z$$

What we will do now, is to rewrite the same partition function as some

$$\mathcal{Z} = \sum_{\{\{\}\}} e^{-\tilde{\mathcal{S}}}$$

where the degrees of freedom and the action are different from what we were doing so far.

12.1.2 The Dual Action

Starting with the above derived action

$$\mathcal{S} = \sum_{IJ \in \{\text{n. n}\}} -J_{IJ} s_I^z s_J^z$$

First we will do one simplification, i.e. we assume an isotropic spacetime lattice, which means that $J_{IJ} = J$ independent of which direction we are coupling with. With this,

$$\mathcal{S} = -J \sum_{IJ \in \{\text{n. n}\}} s_I^z s_J^z$$

which implies

$$\mathcal{Z} = \sum_{\{\{s^z\}\}} e^{J \sum_{IJ} s_I s_J}$$

$$= \sum_{\{\{s^z\}\}} \prod_{IJ} [\cosh J + s_I s_J \sinh J] = (\cosh J)^\# \sum_{\{\{s^z\}\}} \prod_{IJ} [1 + s_I s_J \tanh J]$$

Suppose a new coupling constant K such that

$$e^{-2K} = \tanh J$$

then

$$\mathcal{Z} = (\cosh J)^{\#} \sum_{\{\{s^z\}\}} \prod_{IJ \in \text{n.n}} \left[1 + e^{-2K} e^{i\frac{\pi}{2}(1-s_I)} e^{i\frac{\pi}{2}(1-s_J)} \right]$$

where we have used

$$s_I = \exp\left(\pm i \frac{\pi}{2}(1 \mp s_I) \right) \quad \text{if } s_I \in \{-1, 1\}$$

and therefore

$$\mathcal{Z} = \sum_{\{\{s^z\}\}} \prod_{IJ \in \text{n.n}} \left[1 + \exp\left(-2K + i\frac{\pi}{2}(s_J - s_I) \right) \right]$$

With this modification, we will now proceed to introduce the new degrees of freedom in which we will formulate the dual QFT. (notice that there are a lot of prefactors and constants dropped as they do not change the physics)

The above can be written also as

$$\mathcal{Z} = \sum_{\{\{s^z\}\}} \prod_{IJ \in \text{n.n}} \sum_{n_{IJ}=0,1} \exp\left(n_{IJ}\left(-2K + i\frac{\pi}{2}(s_J - s_I) \right) \right)$$

where we have now introduced a new bond variable n_{IJ} that connects every neighbouring sites I and J, which can take values 0 or 1. Notice that in the above, the product over the neighbouring sites (i.e. over the bonds) can be done first and then one can perform a sum over all possible n_{IJ} configurations

$$\prod_{b=1}^{B}\left(\sum_{n_b=0}^{1} f_b(n_b) \right) = \sum_{n_1=0}^{1} \sum_{n_2=0}^{1} \cdots \sum_{n_B=0}^{1} \prod_{b=1}^{B} f_b(n_b) = \sum_{\{\{n\}\}} \prod_{b=1}^{B} f_b(n_b)$$

to get the same result as

$$\mathcal{Z} = \sum_{\{\{s^z\}\}} \sum_{\{\{n\}\}} \exp\left(\sum_{IJ \in \text{n.n}} n_{IJ}\left(-2K + i\frac{\pi}{2}(s_J - s_I) \right) \right)$$

The second term in the above exponential is a sum over all bonds of $n_{IJ}s_J - n_{IJ}s_J$. Since we are also summing over all values of s_J, and since $s_J \in \{-1, 1\}$, we can also write this as sum over all bonds of $n_{IJ}s_J + n_{IJ}s_J$. That is, you pick a bond, multiply it with the values at its end points, and then add them. We can also write the entire sum equivalently as, pick a lattice site, take all the bonds attached to it, multiply the site value to each bond value and add them all together, i.e. the above equation is equivalent to

$$\mathcal{Z} = \sum_{\{\{s^z\}\}} \sum_{\{\{n\}\}} \exp\left(-2K \sum_{IJ \in \text{n.n}} n_{IJ} \right) \times \exp\left(i\frac{\pi}{2} \sum_I s_I \sum_{J \in \{\text{n.n of } I\}} n_{IJ} \right)$$

Now, we can perform the sum over $\{\{s^z\}\}$. The first term is independent of s_I, and therefore we can pull

122

it out of the sum.

$$\mathcal{Z} = \sum_{\{\{n\}\}} \exp\left(-2K \sum_{IJ \in \text{n.n}} n_{IJ}\right) \sum_{\{\{s^z\}\}} \prod_I \exp\left(i\frac{\pi}{2} s_I \sum_{J \in \{\text{n.n of } I\}} n_{IJ}\right)$$

Further note that

$$\sum_{\{\{s^z\}\}} \equiv \sum_{s_1 \in \{-1,1\}} \sum_{s_2 \in \{-1,1\}} \cdots \sum_{s_I \in \{-1,1\}} \cdots \sum_{s_N \in \{-1,1\}}$$

and therefore

$$\mathcal{Z} = \sum_{\{\{n\}\}} \exp\left(-2K \sum_{IJ \in \text{n.n}} n_{IJ}\right) \prod_I \sum_{s_I \in \{-1,1\}} \exp\left(i\frac{\pi}{2} s_I \sum_{J \in \{\text{n.n of } I\}} n_{IJ}\right)$$

$$= \sum_{\{\{n\}\}} \exp\left(-2K \sum_{IJ \in \text{n.n}} n_{IJ}\right) \prod_I \cos\left(\frac{\pi}{2} \sum_{J \in \{\text{n.n of } I\}} n_{IJ}\right)$$

The cosine term is a constraint since it vanishes for the configurations where for any site I, $\sum_{J \in \{\text{n.n of } I\}} n_{IJ} =$ odd integer and it is 1 for even integer sums, therefore selecting only certain configurations for the path integral. The action is only the first term.

To see what the cosine condition actually means, consider one lattice site I. The sum over n_{IJ} for the site I is a sum over all bonds that are starting from I. If we consider a volume around the site I, then this sum is an object similar to the flux of bond lines, and is therefore just a lattice version of the divergence. The cosine condition is simply cosine of $\pi/2$ times the lattice divergance, and the condition simply translates as

$$\forall I \quad \sum_{J \in \{\text{n.n of } I\}} n_{IJ} = 0 (\text{mod } 2)$$

This has a form of Gauss' law with some subtleties because we working with weird variables.

THE PROCEEDING PART OF THE VIDEO HAS NO AUDIO AND THEREFORE I HAVE SKIPPED THE LECTURE.

13 Path Integrals

13.1 Path Integrals in Quantum Mechanics

In quantum mechanics, the time evolution operator

$$\langle q_f | e^{-iHT} | q_i \rangle$$

can be written as a path integral.

To derive the path integral representation, we can decompose the time interval in the above equation into N smaller intervals.

$$\langle q_f | \underbrace{e^{-iH\epsilon} e^{-iH\epsilon} \cdots e^{-iH\epsilon}}_{N+1 \text{ intervals}} | q_i \rangle$$

where $(N+1)\epsilon = T$. As we did in the previous section, we can now insert identities in complete basis between each of these intervals as

$$\int dq_1 dq_2 \ldots dq_N \ \langle q_f | e^{-iH\epsilon} | q_N \rangle \langle q_N | e^{-iH\epsilon} | q_{N-1} \rangle \langle q_{N-1} | \cdots | q_1 \rangle \langle q_1 | e^{-iH\epsilon} | q_i \rangle$$

First let us suppose the Hamiltonian is a function of only q, of the form $g(q)$. Then, one of the terms in the above integral would be (notice that the operators and classical variables are not distinguished anywhere here. Please proceed carefully regarding the correct ones as operators)

$$\langle q_t | e^{-ig(q)\epsilon} | q_{t-1} \rangle = e^{-ig(q_{t-1})\epsilon} \langle q_t | q_{t-1} \rangle$$
$$= e^{-ig(q_{t-1})\epsilon} \delta(q_t - q_{t-1})$$
$$= e^{-ig\left(\frac{q_t + q_{t-1}}{2}\right)\epsilon} \delta(q_t - q_{t-1})$$

where in the last line we used the fact that tere is a delta function that sets $q_t = q_{t-1}$ to write $g(q_{t-1}) = g\left(\dfrac{q_t + q_{t-1}}{2}\right)$.

It is also useful to write the delta function in the momentum space as

$$\int e^{-ig\left(\frac{q_t + q_{t-1}}{2}\right)\epsilon} e^{ip(q_t - q_{t-1})} \frac{dp}{2\pi}$$

Now let us suppose the Hamiltonian is a function only of momentum p, of the form $f(p)$. One of the terms in the above integral can be simplified by introducing a complete basis in momentum space as

$$\langle q_t | e^{-if(p)\epsilon} | q_{t-1} \rangle = \int \frac{dp}{2\pi} \langle q_t | e^{-if(p)\epsilon} | p \rangle \langle p | q_{t-1} \rangle$$
$$= \int \frac{dp}{2\pi} e^{-if(p)\epsilon} \langle q_t | p \rangle \langle p | q_{t-1} \rangle$$
$$= \int \frac{dp}{2\pi} e^{-if(p)\epsilon} e^{ip(q_t - q_{t-1})}$$

From the above two, for $H = h(p, q)$ one might assume that

$$\langle q_t | e^{-ih(p,q)\epsilon} | q_{t-1} \rangle \overset{?}{=} \int e^{-ih\left(p, \frac{q_t + q_{t-1}}{2}\right)\epsilon} e^{ip(q_t - q_{t-1})} \frac{dp}{2\pi}$$

that is, the matrix element is a simple expression involving the classical Hamiltonian evaluated at the mid point of the two qs.

In general this is not true since p and q do not commute. To see why, see one possible form of Hamiltonian $H = q^r p^s$. Then, at first order,

$$\langle q_t | e^{-iq^r p^s \epsilon} - 1 | q_{t-1} \rangle = i\epsilon \langle q_t | q^r p^s | q_{t-1} \rangle = i\epsilon q_t^r \langle q_t | p^s | q_{t-1} \rangle = i\epsilon q_t^r \int p^s e^{ip(q_t - q_{t-1})} \frac{dp}{2\pi}$$

which is asymmetric in q_t & q_{t-1}, and therefore not the same as what we would get if we simply used the mid-point prescription as assumed. One might also assume that starting with a more symmetric Hamiltonian of the form

$$H = \frac{q^r p^s + p^s q^r}{2}$$

might help, but see that this evaluates at first order to

$$i\epsilon \frac{(q_t^r + q_{t-1}^r)}{2} \int p^s e^{ip(q_t - q_{t-1})} \frac{dp}{2\pi}$$

which is again not the same as the expected result which would have $\dfrac{(q_t + q_{t-1})^r}{2}$.

This is a fundamental problem, mainly due to the operator ordering issue when going from classical theory to quantum theory that we have discussed long ago.

There is an *ordering* that we can choose for H that makes the above formula true. This is called Weyl ordering, and is defined as

$$W(q^m p^r) = \frac{1}{2^m} \sum_{s=0}^{m} \binom{m}{s} q^{m-s} p^r q^s$$

(this is an operator equation).
Suppose we have a Hamiltonian

$$H = \sum_{r,s} q^r p^s$$

the Weyl ordered H_W will be of the form

$$H_W = \sum_{r,s} W(q^r p^s)$$

Then, at first order

$$
\begin{aligned}
\langle q_t | H_W | q_{t-1} \rangle &= \sum_{r,s} \langle q_t | W(q^r p^s) | q_{t-1} \rangle \\
&= \sum_{r,s} \frac{1}{2^r} \sum_{m=0}^{r} \binom{m}{s} \langle q_t | q^{r-m} p^s q^m | q_{t-1} \rangle \\
&= \sum_{r,s} \frac{1}{2^r} \sum_{m=0}^{r} \binom{m}{s} \langle q_t | q^{r-m} p^s q^m | q_{t-1} \rangle \\
&= \sum_{r,s} \frac{1}{2^r} \sum_{m=0}^{r} \binom{m}{s} \int \frac{dp}{2\pi} q_t^{r-m} p^s q_{t-1}^m e^{ip(q_t - q_{t-1})}
\end{aligned}
$$

which is exactly equal to

$$
\sum_{r,s} \int \frac{dp}{2\pi} \left(\frac{q_t + q_{t-1}}{2} \right)^r p^s e^{ip(q_t - q_{t-1})} = \int \frac{dp}{2\pi} H \left(p, \frac{q_t + q_{t-1}}{2} \right) e^{ip(q_t - q_{t-1})}
$$

That is, if I have a classical Hamiltonian H, then if I calculate the above matrix element for a *Weyl ordered* quantum Hamiltonian H_W, I recover some expression of the classical Hamiltonian evaluated at the mid point. That is, when going to a path integral we make a choice of ordering in our quantum Hamiltonian.

Therefore, (assuming from now on that all the Hamiltonian operators are Weyl ordered),

$$
\begin{aligned}
&\langle q_f | e^{-iHT} | q_i \rangle \\
&= \int dq_1 \ldots dq_N \frac{dp_0}{2\pi} \cdots \frac{dp_N}{2\pi} \ e^{-iH\left(\frac{q_f + q_N}{2}, p_N \right)\epsilon} e^{ip_N(q_f - q_N)} e^{-iH\left(\frac{q_N + q_{N-1}}{2}, p_{N-1} \right)\epsilon} e^{ip_{N-1}(q_N - q_{N-1})} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdots e^{-iH\left(\frac{q_1 + q_i}{2}, p_0 \right)\epsilon} e^{ip_0(q_1 - q_i)}
\end{aligned}
$$

Notice that there are only N integrals over dq and $N+1$ integrals over dp since there are $N+1$ intervals into which we have divided the integral by inserting N identities in position basis.

We can write the above expression as

$$
\begin{aligned}
\int dq_1 \ldots dq_N \frac{dp_0}{2\pi} \cdots \frac{dp_N}{2\pi} \ \exp \Bigg\{ i\epsilon \Bigg(&-H \left(\frac{q_f + q_N}{2}, p_N \right) - H \left(\frac{q_N + q_{N-1}}{2}, p_{N-1} \right) \cdots - H \left(\frac{q_1 + q_i}{2}, p_0 \right) \\
&+ p_N \frac{(q_f - q_N)}{\epsilon} + p_{N-1} \frac{(q_N - q_{N-1})}{\epsilon} \cdots + p_0 \frac{(q_1 - q_i)}{\epsilon} \Bigg) \Bigg\}
\end{aligned}
$$

The steps we did so far was correct upto first order in ϵ. At higher orders in ϵ, there would be exponentials of commutators coming from BCH formula. To make the above expression exact, we should take the limit $\epsilon \to 0$ along with $N \to \infty$, such that $T = N\epsilon$ is unchanged. Under this limit, the Hamiltonians terms sum to give an integration over the Hamiltonian, while the momentum terms sum to give an integral over $p\dot{q}$,

giving

$$\langle q_f | e^{-iHT} | q_i \rangle = \int \mathfrak{D}q \mathfrak{D}p \; e^{i \int p\dot{q} - H(q,p)dt}$$

where $\mathfrak{D}q$ is a shorthand for $\lim_{N \to \infty} dq_1 dq_2 \ldots dq_N$ and similarly for $\mathfrak{D}p$. This above expression is an integral over all paths in *phase space.*

To go from this expression to the usual form where we write it in terms of the action, we need to make some assumptions about the Hamiltonian. That is, we assume that the Hamiltonian is of the form

$$H = \frac{p^2}{2m} + V(q)$$

In this case, we can do the p integrals which are all Gaussian integrals. The basic formula for Gaussian integral is

$$\int e^{-ax^2 + bx} dx = e^{-\frac{b^2}{4a}} \sqrt{\frac{\pi}{a}}$$

then in our expression, $a = \dfrac{i\epsilon}{2m}$ and $b = i(q_t - q_{t-1})$. Notice that this substitution is not strictly convergant because the formula for the Gaussian integral assumes that $\mathfrak{Re}(a) > 0$, which is not the case here. Therefore, to perform the integral, we need to rotate time slightly such that $\mathfrak{Im}(\epsilon) < 0$. Amother way one could evaluate this would be to perform a Wick rotation, converting this into a Euclidean path integral, and then rotate back to Minkowski space. With this, we get each integral as

$$\frac{1}{2\pi} \sqrt{\frac{2\pi m}{i\epsilon}} \exp\left(i \frac{m}{2\epsilon}(q_t - q_{t-1})^2 \right)$$

and the path integral becomes

$$\int \sqrt{\frac{m}{2\pi i \epsilon}} dq_1 \cdots \sqrt{\frac{m}{2\pi i \epsilon}} dq_N \; \exp\left\{ i\epsilon \left(-V\left(\frac{q_f + q_N}{2}\right) - V\left(\frac{q_N + q_{N-1}}{2}\right) \cdots - V\left(\frac{q_1 + q_i}{2}\right) \right. \right.$$
$$\left. \left. + \frac{m}{2} \frac{(q_f - q_N)^2}{\epsilon^2} + \frac{m}{2} \frac{(q_N - q_{N-1})^2}{\epsilon^2} \cdots + \frac{m}{2} \frac{(q_1 - q_i)^2}{\epsilon^2} \right) \right\}$$

See now that the measure itself has ϵ in it. In the limit $\epsilon \to 0$, the term in the exponential becomes S, and the path integral simply is

$$\int_{q(0)=q_i}^{q(t)=q_f} \mathfrak{D}q \; e^{iS}$$

where now $\mathfrak{D}q \equiv \lim_{N \to \infty, \epsilon \to 0} \sqrt{\frac{m}{2\pi i \epsilon}} dq_1 \cdots \sqrt{\frac{m}{2\pi i \epsilon}} dq_N$.

Now one can ask what happens to the expressions of the form

$$\int_{q(0)=q_i}^{q(t)=q_f} \mathfrak{D}q \; q(t_2) q(t_1) e^{iS}$$

We interpret this in the following way. Imagine $0 < t_1 < t_2 < t$, then we can write the above path integral

as

$$\int dq_1 dq_2 \int_{q(t_2)=q_2}^{q(t)=q_f} \mathfrak{D}q e^{iS} q_2 \int_{q(t_1)=q_1}^{q(t_2)=q_2} \mathfrak{D}q e^{iS} q_1 \int_{q(0)=q_i}^{a(t_1)=q_1} \mathfrak{D}q e^{iS}$$

That is, we first integrate over all paths which start at q_i and end at some q_1 at t_1, then all paths that start at q_1 at t_1 and end at some q_2 at time t_2, and then all the paths that start at q_2 and end at q_f at t. Further, we multiply by q_1 and q_2, and integrate over all q_1, q_2 to recover the complete path integral.

Now we know exactly how to interpret the individual pieces in the above equation, which is

$$\int dq_1 dq_2 \; q_1 q_2 \, \langle q_f | e^{-iH(t-t_2)} | q_2 \rangle \, \langle q_2 | e^{-iH(t_2-t_1)} | q_1 \rangle \, \langle q_1 | e^{-iH(t_1-0)} | q_i \rangle$$

Now

$$\int dq_1 \; q_1 \, |q_1\rangle \, \langle q_1| = \hat{q}_{1S}$$

(where subscript S denotes that the operator is in Schrodinger picture)
and therefore, the above path integral is

$$\langle q_f | e^{-iH(t-t_2)} q(t_2) e^{-iH(t_2-t_1)} q(t_1) e^{-iH(t_1-0)} | q_i \rangle$$

Now, we can write the above in terms of operators in Heisenberg representation as

$$\langle q_f | e^{-iHt} q_H(t_2) q_H(t_1) | q_i \rangle$$

If we had started at some time t_i instead of 0, the above expression would be

$$\langle q_f | e^{-iHt_f} q_H(t_2) q_H(t_1) e^{iHt_i} | q_i \rangle$$

Notice that the expression is automatically time ordered.

13.2   Path Integrals for Field Theory

Field theories are naturally Quantum mechanics of variables that live in *spacetime*. Therefore, path integrals for field theories are exactly the same as the path integrals for the quantum mechanics that we derived above, but now the paths are not only in time, but also in space, i.e. we integrate over all configurations in spacetime. That is

$$\langle \phi_f | e^{-iHt} | \phi_i \rangle = \int_{\phi(0,\mathbf{x})=\phi_i}^{\phi(t,\mathbf{x})=\phi_2} \mathfrak{D}\phi \; e^{iS}$$

Now in perturbative QFT, we were interested in calculating correlation functions of field insertions in the true vacuum. There is a way to calculate those in the path integral, and the idea is as follows. Consider

$$\int_{\phi(-\infty,\mathbf{x})=\phi_i}^{\phi(\infty,\mathbf{x})=\phi_f} \mathfrak{D}\phi \; e^{iS} \phi(t_1, \mathbf{x}_1) \dots \phi(t_n, \mathbf{x}_n)$$

We can evaluate this the same way we did in the case of quantum mechanics as

$$= \langle \phi_f | \, \mathrm{e}^{-iH(\infty)} T \left\{ \phi_H(t_n) \dots \phi_H(t_1) \right\} \mathrm{e}^{iH\infty} \, | \phi_i \rangle$$

Notice that the infinite time evolution will always project any state to the true vacuum

$$= \langle \phi_f | \Omega \rangle \, \langle \Omega | \, T \left\{ \phi_H(t_n) \dots \phi_H(t_1) \right\} | \Omega \rangle \, \langle \Omega | \phi_i \rangle$$

Also notice that

$$\int_{\phi(-\infty,\mathbf{x})=\phi_i}^{\phi(\infty,\mathbf{x})=\phi_f} \mathfrak{D}\phi \; \mathrm{e}^{iS} = \langle \phi_f | \Omega \rangle \, \langle \Omega | \phi_i \rangle$$

and therefore we get

$$\langle \Omega | \, T \left\{ \phi_H(t_n) \dots \phi_H(t_1) \right\} | \Omega \rangle = \frac{\int_{\phi(-\infty,\mathbf{x})=\phi_i}^{\phi(\infty,\mathbf{x})=\phi_f} \mathfrak{D}\phi \; \mathrm{e}^{iS} \phi(t_1, \mathbf{x}_1) \dots \phi(t_n, \mathbf{x}_n)}{\int_{\phi(-\infty,\mathbf{x})=\phi_i}^{\phi(\infty,\mathbf{x})=\phi_f} \mathfrak{D}\phi \; \mathrm{e}^{iS}}$$

Notice that the boundary condition is totally irrelevant, and therefore we can write the above generally as

$$\langle \Omega | \, T \left\{ \phi_H(t_n) \dots \phi_H(t_1) \right\} | \Omega \rangle = \frac{\int \mathfrak{D}\phi \; \mathrm{e}^{iS} \phi(t_1, \mathbf{x}_1) \dots \phi(t_n, \mathbf{x}_n)}{\int \mathfrak{D}\phi \; \mathrm{e}^{iS}}$$

There is an important caveat here.

$$\lim_{T \to \infty} \mathrm{e}^{iHt} \, | \phi \rangle = | \Omega \rangle \, \langle \Omega | \phi \rangle$$

is not a strictly true equation. It is true only when integrated over some smooth function. Therefore, if we do the path integrals with time on the real axis, they are not well defined. We need to wither go slightly off the real axis ($\tau = \mathrm{e}^{-i\theta}t$), or we need to work with the Euclidean spacetime which one can obtain by taking $\theta = \frac{\pi}{2} \equiv \tau = -it$. With this,

$$\mathrm{e}^{iS} \to \mathrm{e}^{-S_E}$$

and what we actually should calculate is objects of the form

$$\langle \Omega | \, T \left\{ \phi_H(\tau_n) \dots \phi_H(\tau_1) \right\} | \Omega \rangle = \frac{\int \mathfrak{D}\phi \; \mathrm{e}^{-S_E} \phi(\tau_1, \mathbf{x}_1) \dots \phi(\tau_n, \mathbf{x}_n)}{\int \mathfrak{D}\phi \; \mathrm{e}^{iS}}$$

But we are interested in the correlations in Minkowski time and not in Euclidean time, and to go from the Euclidean correlators to the Minkowski correlators, it is very important to know the analyticity properties of the correlators.

First let us take a look at the correlation function without time ordering

$$\langle \Omega | \, \phi(t_1, \mathbf{x}_1) \dots \phi(t_n, \mathbf{x}_n) \, | \Omega \rangle$$

The way to see its analyticity properties is to insert a complete set of energy eigenstates as

$$\sum_{E_1} \cdots \sum_{E_{n-1}} \langle\Omega| \phi(t_1, \mathbf{x}_1) |E_1\rangle \langle E_1| \ldots |E_{n-1}\rangle \langle E_{n-1}| \phi(t_n, \mathbf{x}_n) |\Omega\rangle$$

$$= \sum_{E_1} \cdots \sum_{E_{n-1}} \langle\Omega| \phi(0, \mathbf{x}_1) |E_1\rangle \, \mathrm{e}^{-iE_1 t_1} \mathrm{e}^{iE_1 t_2} \langle E_1| \phi(0, \mathbf{x}_2) \ldots \phi(0, \mathbf{x}_{n-1}) |E_{n-1}\rangle \, \mathrm{e}^{-iE_{n-1} t_{n-1}} \mathrm{e}^{iE_{n-1} t_n} \langle E_{n-1}| \phi(0, \mathbf{x}_n) |\Omega\rangle$$

$$= \sum_{E_1} \cdots \sum_{E_{n-1}} \langle\Omega| \phi(0, \mathbf{x}_1) |E_1\rangle \, \mathrm{e}^{iE_1(t_2 - t_1)} \langle E_1| \phi(0, \mathbf{x}_2) \ldots \phi(0, \mathbf{x}_{n-1}) |E_{n-1}\rangle \, \mathrm{e}^{iE_{n-1}(t_n - t_{n-1})} \langle E_{n-1}| \phi(0, \mathbf{x}_n) |\Omega\rangle$$

The singularities in the above function (we are only interested in the singularities along the time axis) are present only in the limit $E_i \to \infty$. If this correlator is well defined for some value of $t_j - t_{j-1}$, since all E_i are always positive, the correlator is always analytic in the upper half plane of $t_j - t_{j-1}$, since giving a positive imaginary part will lead to a negative real exponential which will only improve the convergance. These correlators are called Wightman functions.

Notice that the Wightman function in Minkowski space has no singularities on the upper half plane for each $t_j - t_{j-1}$, while on the lower half plane there are singularities. The Wick rotation described above is a rotation by $\frac{\pi}{2}$ in the clockwise direction, therefore one can perform a rotation of the integration contour from the real axis to imaginary axis only for the negative part of the real axis since the positive part crosses some singularities when rotated since it rotates in the lower half of the plane. Therefore, if we start with Wightman functions in Minkowski space, and want to Wick rotate it to Euclidean space, it is only possible if the Wightman function is already time ordered, i.e. only if $t_j - t_{j-1} < 0$. Conversely, if we start with a Wightman function in Euclidean space (we do not worry about how it is defined), and Wick rotate all the time differences to Minkowski space at the same time, we recover the time ordered correlation functions automatically.

13.3   Gaussian Integrals in $n-$dimensions

When doing path integrals, we usually do only Gaussian integrals since that is the only thing which we know how to evaluate. The general formula for Gaussian integrals in $n-$dimensions is

$$\int \exp\left\{-x_i A_{ij} x_j + B_i x_i\right\} \prod_i dx_i = \int \exp\left\{-x^T A x + B x\right\} \prod_i dx_i = \exp\left\{-\frac{B^T A^{-1} B}{4}\right\} \sqrt{\frac{1}{\det\left(\frac{A}{\pi}\right)}}$$

You can derive this by just making an rotation of basis $x \to \tilde{x} = O^T x$ where the A is diagonal in $\tilde{x}$ (i.e. $x^T A x = \tilde{x}^T (O^T A O) \tilde{x}$ and $O^T A O$ is diagonal), giving

$$\int \exp\left\{-\tilde{x}_i \tilde{A}_{ii} \tilde{x}_i + \tilde{B}_i \tilde{x}_i\right\} \prod_i d\tilde{x}_i = \prod_i \sqrt{\frac{\pi}{\tilde{A}_{ii}}} \exp\left\{-\frac{\tilde{B}_i^2}{4\tilde{A}_{ii}}\right\}$$

which when we transform back to the original basis is nothing but the above equation.

One more important formula that follows by differentiating the one dimensional gaussian integral once with respect to a is

$$\int e^{-ax^2} x^2 dx = \sqrt{\frac{\pi}{a}} \frac{1}{2a}$$

The vector generalisation of this is

$$\int \left(\prod_i dx_i \right) x_m x_n e^{-x_i A_{ij} x_j} = \sqrt{\frac{1}{\det\left(\frac{A}{\pi}\right)}} \frac{1}{2} A_{mn}^{-1} \tag{13.1}$$

14 Path Integrals for Free Scalar Fields

For a free scalar field, we have

$$S = \frac{1}{2} \int d^4x \ (\partial_\mu \phi)^2 - m^2 \phi^2$$

However, we can also perform integration by parts on the first term, assuming the fields at infinity die off to zero, to get

$$S = \frac{1}{2} \int d^4x \ \phi(-\Box - m^2)\phi$$

The Euclidean action looks like

$$S_E = \frac{1}{2} \int d^4x_E \phi(-\Box_E + m^2)\phi$$

The signs are due to our sign convention. Notice that upon Wick rotation, i.e. $\tau = -it \implies \partial_\tau = i\partial_t \implies \partial_t^2 = -\partial_\tau^2$, and the measure also transforms as