QFT

Table of contents

1 The Klein-Gordon equation

We start out with the aim of generalizing QM for the special relativistic case. Let us take the Schrödinger equation for a single free mass point with no potential

$$-\frac{\hbar^2}{2m}\nabla^2\phi = i\hbar\frac{\partial}{\partial t}\phi$$

We use natural units $\hbar = c = 1$ and relativistic notation

$$t = x^0 \qquad x^\mu = (t, \vec{x}) \qquad \partial_\mu = \frac{\partial}{\partial x^\mu}$$

and a metric

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

The free Schrödinger equation thus reads

$$-\frac{1}{2m}\partial_i\partial_i\phi = i\partial_0\phi$$

Solutions are given by plane waves

$$\phi = e^{i(\vec{k}\cdot\vec{x}-\omega t)}$$

and we find

$$\frac{\vec{k}^2}{2m}\phi = \omega\phi$$

For nontrivial wave functions ϕ this implies

$$\omega = \frac{\vec{k}^2}{2m}$$

which, using the de Broglie relations $E=\omega$, $p^i=k^i$ exposes the physical content of the Schrödinger equation as the nonrelativistic relation between energy and momentum of a free patricle

$$E=\frac{\vec{p}^2}{2m}$$

For the relativistic generalization, we simply try to replace this relation with the relativistic one

$$E^2=\vec{p}^2+m^2$$

Acting on a plane wave we find

$$(-\omega^2+\vec{k}^2+m^2)\phi=(\partial_0\partial_0-\partial_i\partial_i+m^2)\phi=(\partial_0\partial^0+\partial_i\partial^i+m^2)\phi=(\partial_\mu\partial^\mu+m^2)\phi=0$$

We thus find the Klein-Gordon(KG) equation

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

as a generalizarion of the free Schrödinger equation for the free, relativistic particle.

2 Probability interpretation

2.1 The Schrödinger equation

In quantum mechanics, the wave function is interpreted as a probability amplitude. A probability density ρ can be formed as

$$\rho=\phi^*\phi$$

There are two essential properties for a probability density: It needs to be positive and its integral needs to be conserved (and normalized to 1). Positivity of ρ is obviously fulfilled. To check for the conservation of the total probability, we look at the time derivative of the probability density

$$\begin{aligned}
\partial_0\rho &= \phi\partial_0\phi^*+\phi^*\partial_0\phi \\
&= \frac{1}{2m}(-i\phi\partial_i\partial_i\phi^*+\phi^*i\partial_i\partial_i\phi) \\
&= -\frac{i}{2m}(\phi\partial_i\partial_i\phi^*-\phi^*\partial_i\partial_i\phi) \\
&= -\frac{i}{2m}(\partial_i(\phi\partial_i\phi^*)-(\partial_i\phi)(\partial_i\phi^*)-\partial_i(\phi^*\partial_i\phi)+(\partial_i\phi^*)(\partial_i\phi)) \\
&= \partial_i\bigg(-\frac{i}{2m}(\phi\partial_i\phi^*+\phi^*\partial_i\phi)\bigg)
\end{aligned}$$

We thus have a continuity equation

$$\partial_0\rho-\partial_i j_i=0$$

with a probability current

$$j_i=-\frac{i}{2m}(\phi\partial_i\phi^*-\phi^*\partial_i\phi)$$

which allows us to define a conserved quantity

$$Q=\int_{\mathbb{R}^3}d^3x\rho \qquad \partial_0 Q=\int_{\mathbb{R}^3}d^3x\partial_0\rho=\int_{\mathbb{R}^3}d^3x\nabla\vec{j}=\int_{\partial\mathbb{R}^3}d\vec{A}\cdot\vec{j}$$

as long as the probability current vanishes at spatial infinity.

2.2 The Klein-Gordon equation

In the relativistic case, ρ does not satisfy any continuity equation. In fact, it can not be consistently interpreted as a probability density as it is a Lorentz scalar. Assume e.g. that in a certain reference frame the probability density is given by (1D for simplicity)

$$\rho(x) = \begin{cases} \frac{1}{2L} & |x| < L \\ 0 & |x| \geqslant L \end{cases}$$

so $Q = 1$. In a boosted system $(\bar{t}, \bar{x})$ with relative velocity v, one would then have

$$\rho(\bar{x}) = \begin{cases} \frac{1}{2L} & |\bar{x} - v\bar{t}| < L\sqrt{1-v^2} \\ 0 & |\bar{x} - v\bar{t}| \geqslant L\sqrt{1-v^2} \end{cases}$$

and thus $\bar{Q} = \sqrt{1-v^2}$.

A candidate quantity that has the correct behaviour under Lorentz transformation (i.e. that is a time component of a 4-vector) is $j_0 = -\frac{i}{2m}(\phi\partial_0\phi^* - \phi^*\partial_0\phi)$. We can define a 4-vector

$$j_\mu = -\frac{i}{2m}(\phi\partial_\mu\phi^* - \phi^*\partial_\mu\phi) \tag{1}$$

and, using the Klein-Gordon equation, we find

$$\begin{aligned} \partial_\mu j^\mu &= -\frac{i}{2m}(\partial_\mu(\phi\partial^\mu\phi^*) - \partial_\mu(\phi^*\partial^\mu\phi)) \\ &= -\frac{i}{2m}((\partial_\mu\phi)(\partial^\mu\phi^*) + \phi\partial_\mu\partial^\mu\phi^* - (\partial_\mu\phi^*)(\partial^\mu\phi) - \phi^*\partial_\mu\partial^\mu\phi) \\ &= -\frac{i}{2m}(\phi\partial_\mu\partial^\mu\phi^* - \phi^*\partial_\mu\partial^\mu\phi) = 0 \end{aligned}$$

so that in fact $Q = \int_{\mathbb{R}^3} d^3x j^0$ is a conserved quantity. However, j^0 is not positive definite. We can explicitly write it as

$$j^0 = \frac{i}{2m}(\phi^*\partial_0\phi - \phi\partial_0\phi^*) = \frac{1}{m}(I\partial_0 R - R\partial_0 I)$$

with $R = \Re(\phi)$ and $I = \Im(\phi)$. We thus have a conserved current, but we can not interpret it as a probability current.

3 Negative energy

The Klein-Gordon equation has one more problem. Per construction, it has plane wave solutions that fulfill

$$E^2 = \vec{p}^2 + m^2$$

We thus get

$$E = \pm\sqrt{\vec{p}^2 + m^2}$$

so there are negative energy solutions and their energies decrease with momentum. In a free theory this is not a problem as there is no way to gain/loose energy. But as we couple such a system to anything else, it may accellerate, producing energy in the process. It is not concievable how such a system should have a stable minimum energy state.

4 Field theory

In order to proceed, we could abandon the KG equation, or we could try to replace the probabilistic interpretation. We try the later.

At face value, both Schrödinger and KG are classical field equations. Let us therefore try to quantize the fields themselves. We will try this in the canonical formalism known from QM. This is based on the classical Hamiltonian. So let us try to construct the classical Hamiltonian for our field theory. The first step is to identify the Lagrangean.

4.1 Lagrangean for the classical KG field

We obtain the equations of motion (Euler-Lagrange equations) of an arbitrary system by extremizing its action

$$\delta S = 0$$

Generically, the action S is defined from the Lagrangean $\mathscr{L}$ as

$$S = \int dt L(q, \dot{q})$$

for a local field theory that is Lorentz covariant, we have (more on this later)

$$S = \int d^4x \mathscr{L}(\phi, \partial_\mu \phi)$$

with the Lagrange density $\mathscr{L}$ and ϕ collectively denoting all fields ϕ_i. Extremizing is thus achieved by

$$\begin{aligned}
\delta S &= \int_V d^4x \left(\frac{\partial \mathscr{L}}{\partial \phi} \delta\phi_i + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \delta \partial_\mu \phi_i \right) \\
&= \int_V d^4x \left(\frac{\partial \mathscr{L}}{\partial \phi_i} \delta\phi_i + \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \delta\phi_i \right) - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \delta\phi_i \right) \\
&= \int_V d^4x \left(\frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right) \delta\phi_i + \int_{\partial V} d^3x^\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \delta\phi_i
\end{aligned}$$

Integrating over a region larger than the support of the variation $\delta\phi$ we obtain

$$\delta S = \int_V d^4x \left(\frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right) \delta\phi_i$$

recognizing that otherwise $\delta\phi_i$ is arbitrary, we obtain the Euler-Lagrange equations for a local field theory

$$\frac{\partial \mathscr{L}}{\partial \phi_i} = \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)}$$

A Lagrangean that leads to the KG equation as an equation of motion is easy to find. We take

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) \tag{2}$$

which describes a real-valued scalar field ϕ. In order to arrive at the Lagrangean for a complex scalar field, we have to assume independent real and imaginary components and derive separate equations of motion for both. Alternatively, we may view ϕ and ϕ^* as the two independent components. We can thus simply write

$$\mathscr{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi$$

which results in the equations of motion

$$\begin{aligned} -m^2\phi^* &= \partial^\mu\partial_\mu\phi^* \\ -m^2\phi &= \partial^\mu\partial_\mu\phi \end{aligned}$$

4.2 Hamiltonian

In order to define the Hamiltonian, we need conjugate momenta. Keeping in mind that

$$L = \int d^3x \mathscr{L}(\vec{x})$$

we obtain them in the standard fashion

$$\Pi_i(\vec{x}) = \frac{\delta L}{\delta\dot{\phi}_i(\vec{x})} = \frac{\partial\mathscr{L}(\vec{x})}{\partial(\partial_0\phi_i)}$$

The Hamiltonian is thus

$$H = \int d^3x \sum_i \Pi_i(\vec{x})\dot{\phi}_i(\vec{x}) - L = \int d^3x \mathscr{H}(\vec{x})$$

where we have defined the Hamiltonian density as

$$\mathscr{H}(\vec{x}) = \sum_i \Pi_i(\vec{x})\partial_0\phi_i(\vec{x}) - \mathscr{L}(\vec{x})$$

For the real KG theory we have

$$\Pi = \frac{\partial\mathscr{L}}{\partial(\partial_0\phi)} = \partial^0\phi$$

and for the complex one we get

$$\begin{aligned} \Pi &= \frac{\partial\mathscr{L}}{\partial(\partial_0\phi)} = \partial^0\phi^* \\ \Pi^* &= \frac{\partial\mathscr{L}}{\partial(\partial_0\phi^*)} = \partial^0\phi \end{aligned}$$

For the Hamiltonian density of the real KG theory we thus obtain

$$\mathscr{H} = \partial^0\phi\partial_0\phi - \mathscr{L} = \frac{1}{2}(\partial_0\phi\partial_0\phi + \partial_i\phi\partial_i\phi + m^2\phi^2) = \frac{1}{2}(\Pi^2 + \partial_i\phi\partial_i\phi + m^2\phi^2) \tag{3}$$

whereas for the complex theory we get

$$\mathscr{H} = \partial^0\phi^*\partial_0\phi + \partial^0\phi\partial_0\phi^* - \mathscr{L} = \partial_0\phi^*\partial_0\phi + \partial_i\phi^*\partial_i\phi + m^2\phi^*\phi = \Pi^*\Pi + \partial_i\phi^*\partial_i\phi + m^2\phi^*\phi$$

We see that in both cases the Hamiltonian is nonnegative. If this property is not violated upon quantization, the problem of the relativistic energy being unbounded from below would be solved.

4.3 Nöthers first theorem

Let us go back to the Lagrangean formulation and assume that the equations of motion exhibit an invariance under a continuous transformation. Such a transformation is called a continuous symmetry transformation. An infinitesimal form of this transformation can be written as

$$\phi \to \phi' = \phi + \varepsilon\Delta\phi$$

with an infinitesimal parameter $\Delta\phi$. This transformation will generically change the Lagrangean

$$\mathscr{L} \to \mathscr{L}' = \mathscr{L} + \varepsilon\Delta\mathscr{L}$$

The action thus transforms as

$$S' = \int d^4x \mathscr{L}' = S + \varepsilon \int d^4x \Delta\mathscr{L}$$

Since the transformation is a symmetry, i.e. it leaves the equations of motion invariant, we have

$$\frac{\partial\mathscr{L}'}{\partial\phi_i'} = \partial_\mu \frac{\partial\mathscr{L}'}{\partial(\partial_\mu\phi_i')}$$

and thus the variation of the action S', which can generically be written as

$$\begin{aligned}
\delta S' &= \int_V d^4x \left(\frac{\partial\mathscr{L}'}{\partial\phi_i'}\delta\phi_i' + \frac{\partial\mathscr{L}'}{\partial(\partial_\mu\phi_i')}\delta\partial_\mu\phi_i' \right) \\
&= \int_V d^4x \left(\frac{\partial\mathscr{L}'}{\partial\phi_i'}\delta\phi_i' + \partial_\mu\left(\frac{\partial\mathscr{L}'}{\partial(\partial_\mu\phi_i')}\delta\phi_i' \right) - \partial_\mu\frac{\partial\mathscr{L}'}{\partial(\partial_\mu\phi_i')}\delta\phi_i' \right) \\
&= \int_V d^4x \left(\frac{\partial\mathscr{L}'}{\partial\phi_i'} - \partial_\mu\frac{\partial\mathscr{L}'}{\partial(\partial_\mu\phi_i')} \right)\delta\phi_i' + \int_{\partial V} d^3x^\mu \frac{\partial\mathscr{L}'}{\partial(\partial_\mu\phi_i')}\delta\phi_i' \\
&= 0
\end{aligned}$$

does also vanish. Varying S' in terms of the original fields we obtain

$$\delta S' = \delta S + \delta\left(\varepsilon \int_V d^4x \Delta\mathscr{L} \right)$$

and thus

$$\delta\left(\int_V d^4x \Delta\mathscr{L} \right) \stackrel{!}{=} 0$$

which can only be fulfilled if $\int_V d^4x \Delta\mathscr{L}$ is a constant and thus

$$\Delta\mathscr{L} = \partial_\mu J^\mu$$

can be written as the divergence of a current J^μ. In terms of the original fields, we may now write

$$\begin{aligned}
\varepsilon\partial_\mu J^\mu &= \varepsilon\Delta\mathscr{L} = \varepsilon\left(\frac{\partial\mathscr{L}}{\partial\phi_i}\Delta\phi_i + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}\Delta\partial_\mu\phi_i \right) \\
&= \varepsilon\left(\left(\frac{\partial\mathscr{L}}{\partial\phi_i} - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)} \right)\Delta\phi_i + \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}\Delta\phi_i \right) \right)
\end{aligned}$$

The term $\frac{\partial\mathscr{L}}{\partial\phi_i} - \partial_\mu\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}$ vanishes because of the equations of motion, so we are left with

$$\partial_\mu J^\mu = \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}\Delta\phi_i \right)$$

which is a continuity equation

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

for the Nöther current

$$j^\mu := \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_i)}\Delta\phi_i - J^\mu$$

We have thus proven the first Noether theorem, which states that a continuous symmetry of the equations of motion implies a conserved current.

We can now define a conserved charge

$$Q = \int_{\mathbb{R}^3} d^3x j^0 \qquad \partial_0 Q = \int_{\mathbb{R}^3} d^3x \partial_0 j^0 = \int_{\mathbb{R}^3} d^3x \nabla \vec{j} = \int_{\partial\mathbb{R}^3} d\vec{A}\cdot\vec{j} = 0$$

Example 1. For a free, real massless scalar with

$$\mathscr{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi$$

there is a continuous symmetry

$$\phi \to \phi + \varepsilon$$

under which $\Delta\mathscr{L} = 0$. We have $\Delta\phi = 1$ and thus the Nöther current

$$j^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}\Delta\phi = \partial^\mu\phi$$

and the conserved charge

$$Q = \int_{\mathbb{R}^3} d^3x j^0 = \int_{\mathbb{R}^3} d^3x \dot{\phi}$$

Example 2. For the complex KG theory

$$\mathscr{L} = \partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi$$

there is a continuous symmetry

$$\begin{aligned} \phi &\to e^{i\varepsilon}\phi \\ \phi^* &\to \phi^* e^{-i\varepsilon} \end{aligned}$$

We have $\Delta\phi = i\phi$, $\Delta\phi^* = -i\phi^*$ and $\Delta\mathscr{L} = 0$. The Noether current is thus

$$j_\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}\Delta\phi + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi^*)}\Delta\phi^* = i(\phi\partial_\mu\phi^* - \phi^*\partial_\mu\phi)$$

which, up to a trivial multiplicative factor, is our failed probability current (1)

Example 3. As a final example, we look at a spacetime translation

$$x^\mu \to x^\mu - \varepsilon a^\mu$$

with a constant vector a^μ. Under this transformation the fields transform as

$$\phi(x) \to \phi(x + \varepsilon a) = \phi(x) + \varepsilon a^\mu\partial_\mu\phi(x)$$

while the Lagrangean density

$$\mathscr{L}(x) \to \mathscr{L}(x + \varepsilon a) = \mathscr{L}(x) + \varepsilon a^\mu\partial_\mu\mathscr{L}(x)$$

We can thus identify $\Delta\phi = a^\mu\partial_\mu\phi$ and $\Delta\mathscr{L} = a^\mu\partial_\mu\mathscr{L}$ and consequently $J^\mu = a^\mu\mathscr{L}$. Constructing the Noether current we find

$$j^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}\Delta\phi - J^\mu = a^\alpha\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}\partial_\alpha\phi - a^\mu\mathscr{L}$$

If we take for the a^ν the unit vectors $e^\alpha_{(\nu)} = \delta^\alpha_\nu$, we find

$$j^\mu_{(\nu)} = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \delta^\mu_\nu \mathscr{L}$$

and

$$Q_{(\nu)} = \int_{\mathbb{R}^3} d^3x \left(\frac{\partial \mathscr{L}}{\partial(\partial_0 \phi)} \partial_\nu \phi - \delta^0_\nu \mathscr{L} \right) = \int_{\mathbb{R}^3} d^3x (\Pi \partial_\nu \phi - \delta^0_\nu \mathscr{L})$$

For time translations, $\nu = 0$ and we find

$$Q_{(0)} = \int_{\mathbb{R}^3} d^3x (\Pi \partial_0 \phi - \mathscr{L})$$

which is just the Hamiltonian.

5 Quantization of the real KG field

5.1 The real KG field in momentum space

The Lagrangean density of a free scalar field is given by (2)

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) = \frac{1}{2}(\partial_0 \phi \partial_0 \phi - \partial_i \phi \partial_i \phi - m^2 \phi^2)$$

We write the field in terms of its 3D Fourier components

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

which implies the convention

$$\hat{\phi}(t, \vec{p}) = \int d^3x e^{-i\vec{p}\cdot\vec{x}} \phi(t, \vec{x})$$

The reality of the field $\phi^*(t, \vec{x}) = \phi(t, \vec{x})$ implies

$$\begin{aligned}
\hat{\phi}(t, \vec{p}) &= \int d^3x e^{-i\vec{p}\cdot\vec{x}} \phi(t, \vec{x}) \\
&= \int d^3x e^{-i\vec{p}\cdot\vec{x}} \phi^*(t, \vec{x}) \\
&= \left(\int d^3x e^{i\vec{p}\cdot\vec{x}} \phi(t, \vec{x}) \right)^* \\
&= \hat{\phi}^*(t, -\vec{p})
\end{aligned}$$

Plugging the Fourier decomposition of the field into the Lagrangean gives

$$\begin{aligned}
L &= \frac{1}{2} \int d^3x (\partial_0 \phi \partial_0 \phi - \partial_i \phi \partial_i \phi - m^2 \phi^2) \\
&= \frac{1}{2} \int d^3x \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \int \frac{d^3q}{(2\pi)^3} e^{i\vec{q}\cdot\vec{x}} (\partial_0 \hat{\phi}(t, \vec{p}) \partial_0 \hat{\phi}(t, \vec{q}) - (-p_i q_i + m^2) \hat{\phi}(t, \vec{p}) \hat{\phi}(t, \vec{q})) \\
&= \frac{1}{2} \int d^3x \int \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} e^{i(\vec{p}+\vec{q})\cdot\vec{x}} (\partial_0 \hat{\phi}(t, \vec{p}) \partial_0 \hat{\phi}(t, \vec{q}) - (-p_i q_i + m^2) \hat{\phi}(t, \vec{p}) \hat{\phi}(t, \vec{q})) \\
&= \frac{1}{2} \int \frac{d^3p d^3q}{(2\pi)^3} \delta^3(\vec{p} + \vec{q}) (\partial_0 \hat{\phi}(t, \vec{p}) \partial_0 \hat{\phi}(t, \vec{q}) - (-p_i q_i + m^2) \hat{\phi}(t, \vec{p}) \hat{\phi}(t, \vec{q})) \\
&= \frac{1}{2} \int \frac{d^3p}{(2\pi)^3} (\partial_0 \hat{\phi}(t, \vec{p}) \partial_0 \hat{\phi}(t, -\vec{p}) - (p_i p_i + m^2) \hat{\phi}(t, \vec{p}) \hat{\phi}(t, -\vec{p})) \\
&= \frac{1}{2} \int \frac{d^3p}{(2\pi)^3} (\partial_0 \hat{\phi}(t, \vec{p}) \partial_0 \hat{\phi}^*(t, \vec{p}) - (\vec{p}^2 + m^2) \hat{\phi}(t, \vec{p}) \hat{\phi}^*(t, \vec{p}))
\end{aligned}$$

We can see that field components at momenta $\vec{p}$ only couple to components at $-\vec{p}$. Treating the momentum components of the field as our generalized coordinates, we construct the canonically conjugate momenta to the coordinates $(dp/2\pi)^{\frac{3}{2}}\hat{\phi}(t,\vec{p})$ as

$$(dp/2\pi)^{\frac{3}{2}}\hat{\Pi}(t,\vec{p}) := (dp/2\pi)^{\frac{3}{2}}\frac{\delta L}{\delta(\partial_0\hat{\phi}(t,\vec{p}))} = (dp/2\pi)^{\frac{3}{2}}\partial_0\hat{\phi}(t,-\vec{p})$$

and thus write the Hamiltonian

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3}\hat{\Pi}(t,\vec{p})\partial_0\hat{\phi}(t,\vec{p}) - L \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2}(\hat{\Pi}(t,\vec{p})\hat{\Pi}(t,-\vec{p}) + (\vec{p}^2+m^2)\hat{\phi}(t,\vec{p})\hat{\phi}(t,-\vec{p})) \\ &= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}(\hat{\Pi}(t,\vec{p})\hat{\Pi}^*(t,\vec{p}) + (\vec{p}^2+m^2)\hat{\phi}(t,\vec{p})\hat{\phi}^*(t,\vec{p})) \end{aligned} \tag{4}$$

From the canonical equations of motion

$$\frac{\partial H}{\partial p} = \partial_0 q \qquad\qquad \frac{\partial H}{\partial q} = -\partial_0 p$$

we thus obtain (using $q = (dp/2\pi)^{3/2}\hat{\phi}(t,\vec{p})$ and $p = (dp/2\pi)^{3/2}\hat{\Pi}(t,\vec{p})$)

$$\begin{aligned} \frac{\partial H}{\partial\Big((dp/2\pi)^{\frac{3}{2}}\hat{\Pi}(t,\vec{p})\Big)} &= (dp/2\pi)^{\frac{3}{2}}\hat{\Pi}(t,-\vec{p}) &&= (dp/2\pi)^{\frac{3}{2}}\partial_0\hat{\phi}(t,\vec{p}) \\ \frac{\partial H}{\partial\Big((dp/2\pi)^{\frac{3}{2}}\hat{\phi}(t,\vec{p})\Big)} &= (dp/2\pi)^{\frac{3}{2}}(\vec{p}^2+m^2)\hat{\phi}(t,-\vec{p}) &&= -(dp/2\pi)^{\frac{3}{2}}\partial_0\hat{\Pi}(t,\vec{p}) \end{aligned}$$

so ultimately

$$\partial_0\partial_0\hat{\phi}(t,\vec{p}) = \partial_0\hat{\Pi}(t,-\vec{p}) = -(\vec{p}^2+m^2)\hat{\phi}(t,\vec{p})$$

and we see that every momentum component of the field behaves as a harmonic oscillator

$$\ddot{\varphi} = -\omega^2\varphi$$

where we have introduced the shorthand notation $\varphi = (dp/2\pi)^{3/2}\hat{\phi}(t,\vec{p})$ and $\omega^2 = \vec{p}^2 + m^2$. For later convenience, we also introduce the shorthand notation $\pi = (dp/2\pi)^{3/2}\hat{\Pi}(t,-\vec{p})$ for the corresponding canonically conjugate momentum.

5.2 Quantization of the harmonic oscillator

Quantizing the free, real KG field in momentum space thus amouts to quantizing a harmonic oscillator with angular frequency $\omega = \sqrt{\vec{p}^2+m^2}$ for every momentum mode. According to (4), the contribution to the total Hamiltonian of the momentum $\vec{p}$ and $-\vec{p}$ modes is given by

$$H_p = \frac{1}{2}(\pi^*\pi + \omega^2\varphi\varphi^* + \pi\pi^* + \omega^2\varphi^*\varphi)$$

We now quantize this system by promoting φ and π to operators and demanding canonical commutation relations

$$[\varphi,\pi] = [\varphi^\dagger,\pi^\dagger] = i \qquad [\varphi,\pi^\dagger] = [\pi,\varphi^\dagger] = [\varphi,\varphi^\dagger] = [\pi,\pi^\dagger] = 0 \tag{5}$$

The hamiltonian contribution thus reads

$$H_p = \frac{1}{2}(\pi^\dagger\pi + \omega^2\varphi\varphi^\dagger + \pi\pi^\dagger + \omega^2\varphi^\dagger\varphi)$$

By substituting

$$\begin{aligned}\varphi &= \frac{1}{\sqrt{2\omega}}(a + b^\dagger)\\ \pi &= i\sqrt{\frac{\omega}{2}}(a^\dagger - b)\end{aligned}$$

we can rewrite

$$\begin{aligned}H_p &= \frac{1}{2}(\pi^\dagger\pi + \omega^2\varphi\varphi^\dagger + \pi\pi^\dagger + \omega^2\varphi^\dagger\varphi)\\ &= \frac{1}{4}(\omega(a - b^\dagger)(a^\dagger - b) + \omega(a + b^\dagger)(a^\dagger + b) + \omega(a^\dagger - b)(a - b^\dagger) + \omega(a^\dagger + b)(a + b^\dagger))\\ &= \frac{\omega}{2}(aa^\dagger + b^\dagger b + a^\dagger a + bb^\dagger)\end{aligned}$$

In terms of the original fields and momenta, we can write

$$\begin{aligned}a &= \sqrt{\frac{\omega}{2}}\left(\varphi + \frac{i}{\omega}\pi^\dagger\right)\\ b &= \sqrt{\frac{\omega}{2}}\left(\varphi^\dagger + \frac{i}{\omega}\pi\right)\end{aligned}$$

and from the commutation realtions (5) we find that

$$\begin{aligned}[a, a^\dagger] &= \frac{\omega}{2}\left([\varphi, \varphi^\dagger] + \frac{i}{\omega}[\pi^\dagger, \varphi^\dagger] - \frac{i}{\omega}[\varphi, \pi] + \frac{1}{\omega^2}[\pi^\dagger, \pi]\right) = 1\\ [b, b^\dagger] &= \frac{\omega}{2}\left([\varphi^\dagger, \varphi] + \frac{i}{\omega}[\pi, \varphi] - \frac{i}{\omega}[\varphi^\dagger, \pi^\dagger] + \frac{1}{\omega^2}[\pi, \pi^\dagger]\right) = 1\\ [a, b^\dagger] &= \frac{\omega}{2}\left([\varphi, \varphi] + \frac{i}{\omega}[\pi^\dagger, \varphi] - \frac{i}{\omega}[\varphi, \pi^\dagger] + \frac{1}{\omega^2}[\pi^\dagger, \pi^\dagger]\right) = 0\\ [a, b] &= \frac{\omega}{2}\left([\varphi, \varphi^\dagger] + \frac{i}{\omega}[\pi^\dagger, \varphi^\dagger] + \frac{i}{\omega}[\varphi, \pi] + \frac{1}{\omega^2}[\pi^\dagger, \pi]\right) = 0\end{aligned} \tag{6}$$

so the contributions from a and b decouple and H_p can be written as the sum of two independent harmonic oscillators

$$H_p = H_a + H_b$$

with

$$\begin{aligned}H_a &= \frac{\omega}{2}(aa^\dagger + a^\dagger a) = \omega\left(a^\dagger a + \frac{1}{2}\right)\\ H_b &= \frac{\omega}{2}(bb^\dagger + b^\dagger b) = \omega\left(b^\dagger b + \frac{1}{2}\right)\end{aligned}$$

5.3 Properties of a single momentum mode of the KG field

We concentrate on one of the two harmonic oscillators H_a describing the two momentum modes $\vec{p}$ and $-\vec{p}$ of the real KG field. We label the harmonic oscillator energy eigenstates by $|n\rangle$, starting from the ground state $|0\rangle$. Evidently, the lowest possible eigenvalue of H_a is $\omega/2$ and the corresponding eigenstate fulfills

$$H_a|0\rangle = \omega\left(a^\dagger a + \frac{1}{2}\right)|0\rangle = \frac{\omega}{2}|0\rangle \quad \Rightarrow \quad a^\dagger a|0\rangle = 0 \tag{7}$$

Let us next examine what state we obtain when acting with $a^\dagger$ on an energy eigenstate $|n\rangle$ with eigenvalue E_n. The Hamiltonian acting on such a state gives

$$\begin{aligned} H_a a^\dagger|n\rangle &= \omega\left(a^\dagger a+\frac{1}{2}\right)a^\dagger|n\rangle \\ &= \omega\left(a^\dagger a a^\dagger+a^\dagger\frac{1}{2}\right)|n\rangle \\ &= a^\dagger\omega\left(aa^\dagger+\frac{1}{2}\right)|n\rangle \\ &= a^\dagger\omega\left(a^\dagger a+1+\frac{1}{2}\right)|n\rangle \\ &= a^\dagger(H_a+\omega)|n\rangle \\ &= (E_n+\omega)a^\dagger|n\rangle \end{aligned}$$

so it again is an eigenstate with energy $E_n+\omega$. Acting with a on an eigenstate $|n\rangle$ on the other hand results in a state on which the Hamiltonian gives

$$\begin{aligned} H_a a|n\rangle &= \omega\left(a^\dagger a+\frac{1}{2}\right)a|n\rangle \\ &= \omega\left(a^\dagger a a+a\frac{1}{2}\right)|n\rangle \\ &= \omega\left(aa^\dagger a-1+a\frac{1}{2}\right)|n\rangle \\ &= a\omega\left(a^\dagger a-\frac{1}{2}\right)|n\rangle \\ &= a(H_a-\omega)|n\rangle \\ &= (E_n-\omega)a|n\rangle \end{aligned}$$

so it is also an energy eigenstate with energy $E_n-\omega$. On the ground state $|0\rangle$, this property together with (7) implies that $a|0\rangle=0$. We thus have a method to construct all energy eigenstates, $|n\rangle$ with $n\in\mathbb{N}_0$ from the ground state $|0\rangle$ by repeated application of $a^\dagger$ on it. We define

$$|n\rangle := \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle$$

and find $E_n=\omega(n+1/2)$.

Exercise 1. Show that $\langle n|n\rangle=\langle 0|0\rangle$

It is also easy to show that there are no further energy eigenstates: Let us suppose that $|r\rangle$, $r\notin\mathbb{N}_0$ is an eigenstate of the Hamiltonian with energy $E_r=\omega(r+1/2)$. Because of the positivity of $a^\dagger a$, we know that $r>0$. Let us label with i the largest integer smaller than r. Starting from $|r\rangle$, we can thus construct another energy eigenstate $|r-i\rangle\propto a^i|r\rangle$ with $E_{r-i}=\omega(r-i+1/2)$ and $r-i<1$. Acting on such a state with a can not produce an energy eigenstate $|r-i-1\rangle$ since its energy $E_{r-i}=\omega(r-i-1/2)$ would violate the positivity of $a^\dagger a=H/\omega-1/2$. The only way out is that $a|r-i\rangle=0$. But $a|r-i\rangle=0$ implies $H|r-i\rangle=\omega/2$, so $|r-i\rangle=|0\rangle$ which is in contradiction with our assumption.

The operators $a^\dagger$ and a move us up and down the tower of energy eigenstates and are therefore called ladder operators. The operator

$$N=a^\dagger a$$

evidently has the same eigenstates as H. Its eigenvalues are

$$N|n\rangle=n|n\rangle$$

5.4 The quantum KG field in coordinate space

Remembering the definitions $\varphi = (dp/2\pi)^{3/2}\hat{\phi}(t, \vec{p})$ and $\pi = (dp/2\pi)^{3/2}\hat{\Pi}(t, -\vec{p})$, we define ladder operators per momentum as

$$\begin{aligned} a(t, \vec{p}) &= \left(\frac{2\pi}{dp}\right)^{3/2} a \\ a(t, -\vec{p}) &= \left(\frac{2\pi}{dp}\right)^{3/2} b \end{aligned}$$

so that

$$\begin{aligned} \hat{\phi}(t, \vec{p}) &= \frac{1}{\sqrt{2\omega(\vec{p})}}(a(t, \vec{p}) + a^\dagger(t, -\vec{p})) \\ \hat{\Pi}(t, -\vec{p}) &= i\sqrt{\frac{\omega(\vec{p})}{2}}(a^\dagger(t, \vec{p}) - a(t, -\vec{p})) \end{aligned} \tag{8}$$

We can rewrite the field Hamiltonian (4) in terms of the ladder operators as

$$H = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega(\vec{p})(a^\dagger(t, \vec{p})a(t, \vec{p}) + a(t, \vec{p})a^\dagger(t, \vec{p})) \tag{9}$$

with

$$\omega(\vec{p}) = \sqrt{\vec{p}^2 + m^2}$$

From the commutation relations (6) it follows that

$$[a(t, \vec{p}), a^\dagger(t, \vec{q})] = \frac{(2\pi)^3}{d^3p}\delta_{\vec{p},\vec{q}}$$

and

$$[a(t, \vec{p}), a(t, \vec{q})] = [a^\dagger(t, \vec{p}), a^\dagger(t, \vec{q})] = 0$$

To make sense of the normalization factor $(2\pi)^3/d^3p$, we construct the integral $\int d^3p$ as the limit of a sum over finite momentum space volumes d^3p

$$\int d^3p = \lim_{d^3p\to 0}\sum_{\vec{p}} d^3p$$

Pluging the commutation relation into such an integral we obtain

$$\begin{aligned} \int d^3p f(\vec{p})[a(t, \vec{p}), a^\dagger(t, \vec{q})] &= \lim_{d^3p\to 0}\sum_{\vec{p}} d^3p\frac{(2\pi)^3}{d^3p}\delta_{\vec{p},\vec{q}}f(\vec{p}) \\ &= (2\pi)^3 \lim_{d^3p\to 0}\sum_{\vec{p}} \delta_{\vec{p},\vec{q}}f(\vec{p}) \\ &= (2\pi)^3 f(\vec{q}) \end{aligned}$$

so the commutation relation

$$[a(t, \vec{p}), a^\dagger(t, \vec{q})] = (2\pi)^3\delta^3(\vec{p} - \vec{q})$$

contains the proper Dirac δ. The commutation relations of the fields themselves thus read (8)

$$\begin{aligned}[\hat{\phi}(t,\vec{p}),\hat{\Pi}(t,\vec{q})] &= \frac{i}{2}([a(t,\vec{p}),a^\dagger(t,-\vec{q})]-[a^\dagger(t,-\vec{p}),a(t,\vec{q})]) \\ &= i(2\pi)^3\delta^3(\vec{p}+\vec{q}) \\ [\hat{\phi}(t,\vec{p}),\hat{\phi}(t,\vec{q})] = [\hat{\Pi}(t,\vec{p}),\hat{\Pi}(t,\vec{q})] &= 0\end{aligned}$$

Finally, transforming back to coordinate space, we obtain

$$\begin{aligned}\phi(t,\vec{x}) &= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\hat{\phi}(t,\vec{p}) = \int\frac{d^3p}{(2\pi)^3}\frac{e^{i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega(\vec{p})}}(a(t,\vec{p})+a^\dagger(t,-\vec{p})) \\ \Pi(t,\vec{x}) &= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\hat{\Pi}(t,\vec{p}) = i\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\sqrt{\frac{\omega(\vec{p})}{2}}(a^\dagger(t,-\vec{p})-a(t,\vec{p}))\end{aligned} \tag{10}$$

with the commutation relations

$$\begin{aligned}[\phi(t,\vec{x}),\Pi(t,\vec{y})] &= \int\frac{d^3p}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}e^{i\vec{q}\cdot\vec{y}}[\hat{\phi}(t,\vec{p}),\hat{\Pi}(t,\vec{q})] \\ &= i\int\frac{d^3p}{(2\pi)^3}\int d^3q\,e^{i\vec{p}\cdot\vec{x}}e^{i\vec{q}\cdot\vec{y}}\delta^3(\vec{p}+\vec{q}) \\ &= i\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \\ &= i\delta^3(\vec{x}-\vec{y}) \\ [\phi(t,\vec{x}),\phi(t,\vec{y})] = [\Pi(t,\vec{x}),\Pi(t,\vec{y})] &= 0\end{aligned}$$

5.5 The first divergence - cosmological constant renormalization

Quantizing the real KG field in momentum space revealed a harmonic oscillator per momentum $\vec{p}$ and we have now found the expression for the fields in the ladder operators of these oscillators. But what does this mean physically?

To answer this question, we first look at the full Hamiltonian in terms of the ladder operators (9). Utilizing our commutation relations we find

$$\begin{aligned}H &= \frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\omega(\vec{p})(a^\dagger(t,\vec{p})a(t,\vec{p})+a(t,\vec{p})a^\dagger(t,\vec{p})) \\ &= \int\frac{d^3p}{(2\pi)^3}\omega(\vec{p})a^\dagger(t,\vec{p})a(t,\vec{p})+\frac{1}{2}\int d^3p\omega(\vec{p})\,\delta^3(0)\end{aligned}$$

This expression is ill defined as the second term diverges. It originates from the ground state energies of an infinite number of harmonic oscillators. The problem can be overcome by noticing that - with the exception of gravity - we always are interested in energy differences only. Outside of gravity absolute energy is not an observable and we may redefine the energy of each harmonic oscillator by subtracting the ground state energy before taking the limit of an infinite number of harmonic oscillators (i.e. momentum modes). For gravity the subtraction of a constant energy amounts to a redefinition of the cosmological constant Λ, which would have to absorb the pertaining divergence. The procedure is thus called cosmological constant renormalization.

Technically this elimination can be achieved by a procedure called normal ordering. It consists of commuting a operators to the right and $a^\dagger$ operators to the left and discarding the commutation terms and is denoted as

$$:aa^\dagger:=a^\dagger a$$

Redefining the Hamiltonian as the normal ordered product we obtain

$$\begin{aligned} H &= :\frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega(\vec{p})(a^\dagger(t,\vec{p})a(t,\vec{p}) + a(t,\vec{p})a^\dagger(t,\vec{p})): \\ &= \int \frac{d^3p}{(2\pi)^3}\omega(\vec{p})a^\dagger(t,\vec{p})a(t,\vec{p}) \end{aligned}$$

5.6 Physical interpretation

After normal ordering, the Hamiltonian has a very simple interpretation. If all oscillators are in the ground state, the total energy vanishes. This is the lowest energy state and we call it the vacuum $|0\rangle$. Technically, this vacuum is a direct product over the ground states of each momentum mode

$$|0\rangle = \otimes_{\vec{p}}|0\rangle_{\vec{p}}$$

We may act on the vacuum with a ladder operator $a^\dagger(t,\vec{p})$ for any momentum $\vec{p}$ an arbitrary number of times. The entirety of resulting states spans the Hilbert space of the theory, which is called Fock space.

Each such application of $a^\dagger(t,\vec{p})$ raises the energy of the state by a fixed amount $\omega(\vec{p}) = \sqrt{\vec{p}^2 + m^2}$, which happens to be exactly the energy of a relativistic point mass of momentum $\vec{p}$. We call these field quanta that are created by $a^\dagger(t,\vec{p})$ and annihilated by $a(t,\vec{p})$ particles and the the operators creation resp. annihilation operators. Note that the Hamiltonian literally counts the number of particles in each mode with the number operator $a^\dagger(t,\vec{p})a(t,\vec{p})$ and multiplies it with the relativistic energy of a single particle at that momentum.

Exercise 2. Show that the field momentum, i.e. the normal ordered Noether charge pertaining to spatial translations, can be written as $P^i = \int \frac{d^3p}{(2\pi)^3}p^i a^\dagger(t,\vec{p})a(t,\vec{p})$.

The momentum of the quantum field is given by

$$\vec{P} = \int \frac{d^3p}{(2\pi)^3}\vec{p}a^\dagger(t,\vec{p})a(t,\vec{p})$$

which also counts the number of particles per momentum mode and multiplies it with the momentum of that mode.

Note that the operators for creating two particles at equal times always commute

$$a^\dagger(t,\vec{p})a^\dagger(t,\vec{q})|0\rangle = a^\dagger(t,\vec{q})a^\dagger(t,\vec{p})|0\rangle$$

and thus a multiple particle state is unique.

Finally, we want to clarify the meaning of the operator valued quantum fields themselves. Acting on the vacuum, we obtain a state we call $|t,\vec{x}\rangle$

$$\begin{aligned} |t,\vec{x}\rangle := \phi(t,\vec{x})|0\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{e^{i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega(\vec{p})}}(a(t,\vec{p}) + a^\dagger(t,-\vec{p}))|0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{e^{-i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega(\vec{p})}}a^\dagger(t,\vec{p})|0\rangle \end{aligned} \tag{11}$$

We defne a single particle state of momentum $\vec{p}$ in Fock space as

$$|t,\vec{p}\rangle = N_{\vec{p}}a^\dagger(t,\vec{p})|0\rangle$$

with a normalization factor $N_{\vec{p}}$ still to be chosen. We now have

$$\begin{aligned}\langle t,\vec{x}|t,\vec{p}\rangle &= N_{\vec{p}}\int\frac{d^3q}{(2\pi)^3}\frac{e^{-i\vec{q}\cdot\vec{x}}}{\sqrt{2\omega(\vec{q})}}\langle 0|a(t,\vec{q})a^\dagger(t,\vec{p})|0\rangle\\ &= N_{\vec{p}}\int\frac{d^3q}{(2\pi)^3}\frac{e^{-i\vec{q}\cdot\vec{x}}}{\sqrt{2\omega(\vec{q})}}\langle 0|[a(t,\vec{q}),a^\dagger(t,\vec{p})]|0\rangle\\ &= N_{\vec{p}}\int d^3q\frac{e^{-i\vec{q}\cdot\vec{x}}}{\sqrt{2\omega(\vec{q})}}\delta^3(\vec{p}-\vec{q})\langle 0|0\rangle\\ &= e^{-i\vec{p}\cdot\vec{x}}\frac{N_{\vec{p}}}{\sqrt{2\omega(\vec{p})}}\end{aligned}$$

so $|\vec{x}\rangle$ can be identified with a single particle state at position $\vec{x}$ in complete analogy to the quantum mechanical case where an eigenstate of momentum $\vec{p}$ is given by $\langle t,\vec{x}|t,\vec{p}\rangle=\psi(\vec{x})=e^{-i\vec{p}\cdot\vec{x}}$. For future convenience we choose $N_{\vec{p}}=\sqrt{2\omega(\vec{p})}$ so that

$$|t,\vec{p}\rangle=\sqrt{2\omega(\vec{p})}a^\dagger(t,\vec{p})|0\rangle \qquad \langle t,\vec{x}|t,\vec{p}\rangle=e^{-i\vec{p}\cdot\vec{x}} \qquad \langle t,\vec{p}|t,\vec{q}\rangle=2\omega(\vec{p})(2\pi)^3\delta^3(\vec{p}-\vec{q})$$

6 Spacetime behaviour of the free scalar quantum field

6.1 The dynamical equations

After quantizing the real KG field, we are left with time dependent operators acting on time independent state vectors, i.e. we are in the Heisenberg picture. Following general quantum mechanical principles, the time translation of any operator is generated by the Hamiltonian as

$$O(t+dt)=U(t,dt)O(t)U^\dagger(t,dt) \qquad U(t,dt)=e^{iH(t)dt}$$

which implies

$$\begin{aligned}O(t+dt) &= U(t,dt)O(t)U^\dagger(t,dt)\\ &= (1+iH(t)dt)O(t)(1-iH(t)dt)\\ &= O(t)+i[H(t),O(t)]dt\end{aligned}$$

and thus the operators fulfill the Heisenberg equation

$$\frac{\partial O}{\partial t}=-i[O,H]$$

We could shift the time dependence into the states, i.e. go to the Schrödinger picture, by noting that a generic matrix element, and thus the physics content

$$\begin{aligned}\langle s'(t)|O(t+dt)|s(t)\rangle &= \langle s'(t)|e^{iH(t)dt}O(t)e^{-iH(t)dt}|s(t)\rangle\\ &= \langle s'(t+dt)|O(t)|s(t+dt)\rangle\end{aligned}$$

is invariant when applying the time evolution to the state rather than to the matrix element

$$|s(t+dt)\rangle=e^{-iH(t)dt}|s(t)\rangle=|s(t)\rangle-iH(t)dt|s(t)\rangle$$

Obviously, this implies the Schrödinger equation

$$\frac{\partial}{\partial t}|s\rangle=-iH|s\rangle$$

Notice that there is no need now to adapt the Schrödinger equation to the relativistic case. The Schrödinger equation does not encode the nonrelativistic dispersion relation any more but simply the time evolution of the system. The relativistic physics content is in the field Hamiltonean H.

6.2 Time evolution

We start by computing the time evolution of the creation and annihilation operators

$$\begin{aligned}
\frac{\partial a(\vec{p})}{\partial t} &= -i[a(\vec{p}), H] \\
&= -i\int \frac{d^3q}{(2\pi)^3}\omega(\vec{q})[a(\vec{p}), a^\dagger(\vec{q})]a(\vec{q}) \\
&= -i\int d^3q\omega(\vec{q})\delta^3(\vec{p}-\vec{q})a(\vec{q}) \\
&= -i\omega(\vec{p})a(\vec{p}) \\
\frac{\partial a^\dagger(\vec{p})}{\partial t} &= -i[a^\dagger(\vec{p}), H] \\
&= -i\int \frac{d^3q}{(2\pi)^3}\omega(\vec{q})a^\dagger(\vec{q})[a^\dagger(\vec{p}), a(\vec{q})] \\
&= i\int d^3q\omega(\vec{q})\delta^3(\vec{p}-\vec{q})a^\dagger(\vec{q}) \\
&= i\omega(\vec{p})a^\dagger(\vec{p})
\end{aligned} \tag{12}$$

The time evolution of momentum components of the field and the field momenta are thus

$$\begin{aligned}
\frac{\partial\hat{\phi}(\vec{p})}{\partial t} &= \frac{1}{\sqrt{2\omega(\vec{p})}}\left(\frac{\partial a(\vec{p})}{\partial t}+\frac{\partial a^\dagger(-\vec{p})}{\partial t}\right) \\
&= \frac{1}{\sqrt{2\omega(\vec{p})}}(-i\omega(\vec{p})a(\vec{p})+i\omega(\vec{p})a^\dagger(-\vec{p})]) \\
&= i\sqrt{\frac{\omega(\vec{p})}{2}}(-a(\vec{p})+a^\dagger(-\vec{p})]) \\
&= \hat{\Pi}(\vec{p}) \\
\frac{\partial\hat{\Pi}(-\vec{p})}{\partial t} &= i\sqrt{\frac{\omega(\vec{p})}{2}}\left(\frac{\partial a^\dagger(\vec{p})}{\partial t}-\frac{\partial a(-\vec{p})}{\partial t}\right) \\
&= i\sqrt{\frac{\omega(\vec{p})}{2}}(i\omega(\vec{p})a^\dagger(\vec{p})+i\omega(\vec{p})a(-\vec{p})) \\
&= -\frac{\omega^2(\vec{p})}{\sqrt{2\omega(\vec{p})}}(a^\dagger(\vec{p})+a(-\vec{p})) \\
&= -\omega^2(\vec{p})\hat{\phi}(-\vec{p}) \\
&= -(\vec{p}^2+m^2)\hat{\phi}(-\vec{p})
\end{aligned}$$

and finally for the fields and field momenta in coordinate space we have

$$\begin{aligned}
\frac{\partial\phi(\vec{x})}{\partial t} &= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\frac{\partial\hat{\phi}(\vec{p})}{\partial t} \\
&= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\hat{\Pi}(\vec{p}) \\
&= \Pi(\vec{x})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\Pi(\vec{x})}{\partial t} &= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\frac{\partial\hat{\Pi}(\vec{p})}{\partial t} \\
&= -\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}(\vec{p}^2+m^2)\hat{\phi}(\vec{p}) \\
&= (\vec{\nabla}^2-m^2)\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\hat{\phi}(\vec{p}) \\
&= (\vec{\nabla}^2-m^2)\phi(\vec{x})
\end{aligned}$$

Putting these together we find

$$\frac{\partial^2\phi(\vec{x})}{\partial t^2} = \frac{\partial\Pi(\vec{x})}{\partial t} = (\vec{\nabla}^2 - m^2)\phi(\vec{x})$$

which we can rewrite as

$$(\partial_0\partial_0 - \partial_i\partial_i + m^2)\phi(\vec{x}) = 0$$

or, in relativistic notation

$$(\partial_\mu\partial^\mu + m^2)\phi(\vec{x}) = 0$$

which is the Klein-Gordon equation. Note that this result is not trivial: It tells us that the KG equation is fulfilled exactly by the quantum fields, so it survived quantization.

6.3 Explicit expression for the time evolution of the quantum KG field

Solving the Heisenberg equations for creation and annihilation operators (12), we find

$$a(t,\vec{p}) = e^{-i\omega(\vec{p})t}a(\vec{p}) \qquad a^\dagger(t,\vec{p}) = e^{i\omega(\vec{p})t}a^\dagger(\vec{p})$$

where we have introduced the shorthand notation

$$a(\vec{p}) := a(0,\vec{p}) \qquad a^\dagger(\vec{p}) := a^\dagger(0,\vec{p})$$

for creation and annihilation operators at the reference time $t=0$. Plugging this result into our expression for the momentum state fields (8) we obtain

$$\begin{aligned}
\hat{\phi}(t,\vec{p}) &= \frac{1}{\sqrt{2\omega(\vec{p})}}(e^{-i\omega(\vec{p})t}a(\vec{p}) + e^{i\omega(\vec{p})t}a^\dagger(-\vec{p})) \\
\hat{\Pi}(t,-\vec{p}) &= i\sqrt{\frac{\omega(\vec{p})}{2}}(e^{i\omega(\vec{p})t}a^\dagger(\vec{p}) - e^{-i\omega(\vec{p})t}a(-\vec{p}))
\end{aligned}$$

and after Fourier transformation into coordinate space

$$\begin{aligned}
\phi(t,\vec{x}) &= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\hat{\phi}(t,\vec{p}) = \int\frac{d^3p}{(2\pi)^3}\frac{e^{i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega(\vec{p})}}(e^{-i\omega t(\vec{p})}a(\vec{p}) + e^{i\omega(\vec{p})t}a^\dagger(-\vec{p})) \\
\Pi(t,\vec{x}) &= \int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\hat{\Pi}(t,\vec{p}) = i\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\sqrt{\frac{\omega(\vec{p})}{2}}(e^{i\omega(\vec{p})t}a^\dagger(-\vec{p}) - e^{-i\omega(\vec{p})t}a(\vec{p}))
\end{aligned}$$

We can rewrite these expressions in a slightly more compact way by first defining the time component of the momentum vector as the positive frequency

$$p^0 := \omega(\vec{p}) = +\sqrt{\vec{p}^2 + m^2}$$

and then changing the definition of the dummy momentum integration variable $\vec{p}\to-\vec{p}$ for the creation operators. We obtain

$$\begin{aligned}
\phi(x) &= \int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}(e^{-ip\cdot x}a(\vec{p}) + e^{ip\cdot x}a^\dagger(\vec{p})) \\
\Pi(x) &= i\int\frac{d^3p}{(2\pi)^3}\sqrt{\frac{p^0}{2}}(e^{ip\cdot x}a^\dagger(\vec{p}) - e^{-ip\cdot x}a(\vec{p}))
\end{aligned} \tag{13}$$

where we used the shorthand notation

$$p\cdot x := p_\mu x^\mu$$

6.4 Propagators

We now investigate the propagation of a single particle state, i.e. we compute the probability amplitude for a particle that is created at a space-time point y to reach a spacetime point x. Accorfing to (11), a single particle state at space-time point y is given by

$$|y\rangle = \phi(y)|0\rangle$$

Consequently, the probability amplitude for a particle that is created at a space-time point y to reach a spacetime point x is given by

$$\begin{aligned}
\langle x|y\rangle &= \langle 0|\phi(x)\phi(y)|0\rangle \\
&= \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}\frac{1}{\sqrt{2q^0}}e^{-ip\cdot x}e^{iq\cdot y}\langle 0|a(\vec{p})a^\dagger(\vec{q})|0\rangle \\
&= \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}\frac{1}{\sqrt{2q^0}}e^{-ip\cdot x}e^{iq\cdot y}(2\pi)^3\delta^3(\vec{p}-\vec{q}) \\
&= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2p^0}e^{-ip\cdot(x-y)}
\end{aligned} \tag{14}$$

which is only dependent on the differecne in positions $x-y$. We call this object the propagator

$$D(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2p^0}e^{-ip\cdot x} \tag{15}$$

There is an interesting connection between the propagator and the Green's function of the Klein-Gordon equation. We recall that the Greens function is the solution of an inhomogenous differential equation with a δ source term, so in the case of the KG equation

$$(\partial_\mu\partial^\mu + m^2)G(x) = -i\delta^4(x)$$

We take the four dimensional Fourier transform which is generically defined as

$$\hat{\phi}(p) = \int d^4x e^{ip\cdot x}\phi(x)$$

with the inverse

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

and obtain

$$(-p^\mu p_\mu + m^2)\hat{G}(p) = -i$$

so the solution is

$$\hat{G}(p) = \frac{i}{p^\mu p_\mu - m^2}$$

In position space, the propagator is thus given via the inverse Fourier transform as

$$\begin{aligned}
G(x) &= \int \frac{d^4p}{(2\pi)^4}e^{-ip\cdot x}\hat{G}(p) \\
&= \int \frac{d^4p}{(2\pi)^4}e^{-ip\cdot x}\frac{i}{p^\mu p_\mu - m^2} \\
&= \int \frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\int \frac{dp^0}{2\pi}\frac{ie^{-ip^0x^0}}{p^0p^0 - (\vec{p}^2+m^2)}
\end{aligned}$$

Note that in this expression p^0 is not fixed but rather is an integration variable. We now proceed to compute explicitly the innermost integral over p^0. With the shorthand notation $E = p^0$ and $\omega = +\sqrt{\vec{p}^2 + m^2}$ we need to compute the integral

$$I = \int_{-\infty}^{\infty} \frac{dE}{2\pi} \frac{ie^{-iEx^0}}{E^2 - \omega^2} \tag{16}$$

The integrand has two poles at

$$E = \pm\omega$$

which lie on the real axis, i.e. on our integration contour. We will therefore need to be very careful with our choice of integration contour, but let us first compute the residues. To find them, we choose an integration contour describing an infinitesimal circle around the poles. We substitute

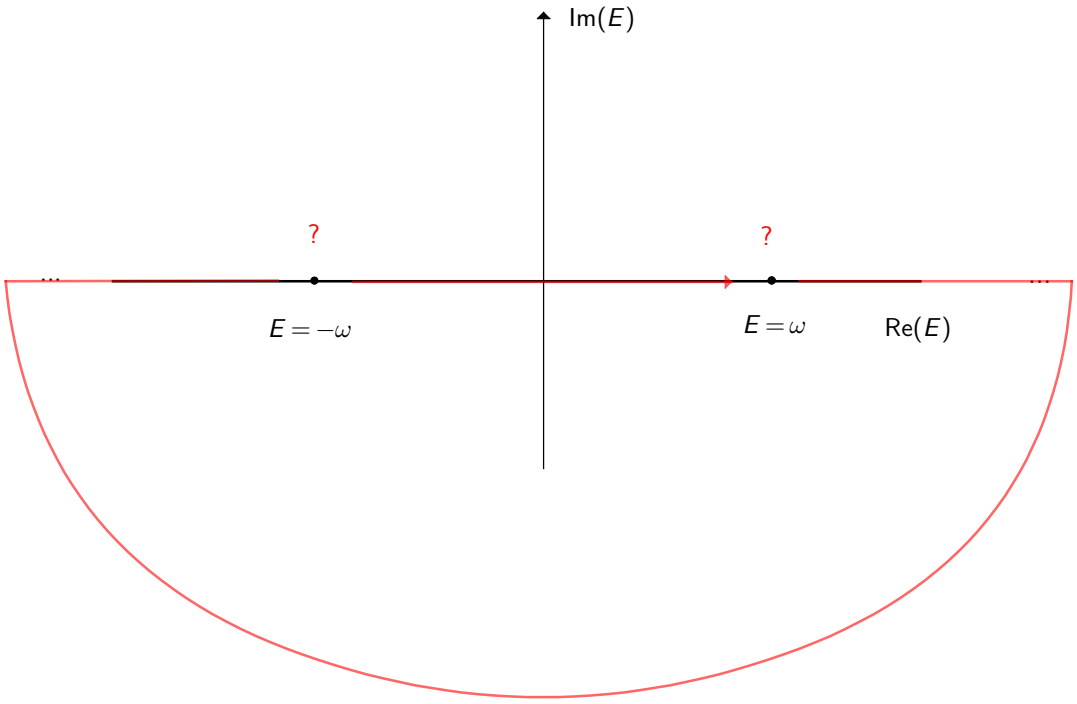
$$E = \pm\omega + \varepsilon e^{i\alpha} \qquad dE = i\varepsilon e^{i\alpha} d\alpha$$

$$\begin{aligned} \lim_{\varepsilon\to 0} \oint \frac{dE}{2\pi} \frac{ie^{-iEx^0}}{E^2 - \omega^2} &= \lim_{\varepsilon\to 0} i\varepsilon \int_0^{2\pi} \frac{e^{i\alpha} d\alpha}{2\pi} \frac{ie^{\mp i\omega x^0}}{\pm 2\omega\varepsilon e^{i\alpha}} \\ &= \mp \frac{1}{2\pi} \int_0^{2\pi} \frac{d\alpha}{2\omega} e^{\mp i\omega x^0} \\ &= \mp \frac{e^{\mp i\omega x^0}}{2\omega} \end{aligned}$$

To compute our integral I, we first note that there are no other poles in the complex plane than those at $\pm\omega$. In addition, if we suppose that our time component x^0 is real and positive,

$$\mathrm{Re}(-iEx^0) < 0 \quad \text{for} \quad \mathrm{Im}(E) < 0$$

so the exponential function in the integrand of (16) vanishes for $\mathrm{Im}(E) \to -\infty$, i.e. if we close the integration path of our integral I from below.



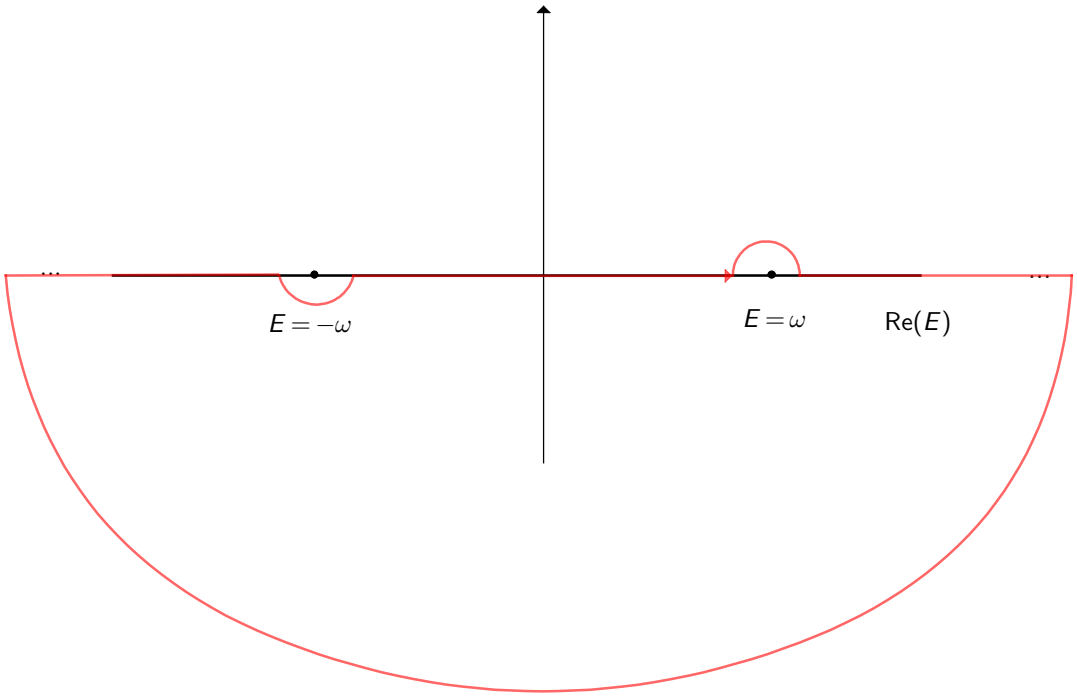
We now have to decide what to do about the poles. Including none of them would render our integral zero according to the residue theorem. Including the pole at $E=\omega$ (thereby encircling it clockwise, so in mathematically negative sense) we would obtain (see (15))

$$G(x)=\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\frac{e^{-i\omega x^0}}{2\omega(\vec{p})}=D(x)$$

which can be interpreted as a momentum $\vec{p}$ mode propagating with positive wave number ω, i.e. energy. Conversely, if we include the pole at $E=-\omega$ into the integration contour, we would obtain

$$G(x)=\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\frac{e^{i\omega x^0}}{-2\omega(\vec{p})}$$

which would correspond to a negative wave number. Since we would like to eliminate negative energy modes, we chose to include the $E=\omega$ pole only, which corresponds to choosing the integration contour depicted in the following figure:



We now investigate the case of a real, negative time component x^0. There we have

$$\mathrm{Re}(-iEx^0)<0 \quad \text{for} \quad \mathrm{Im}(E)>0$$

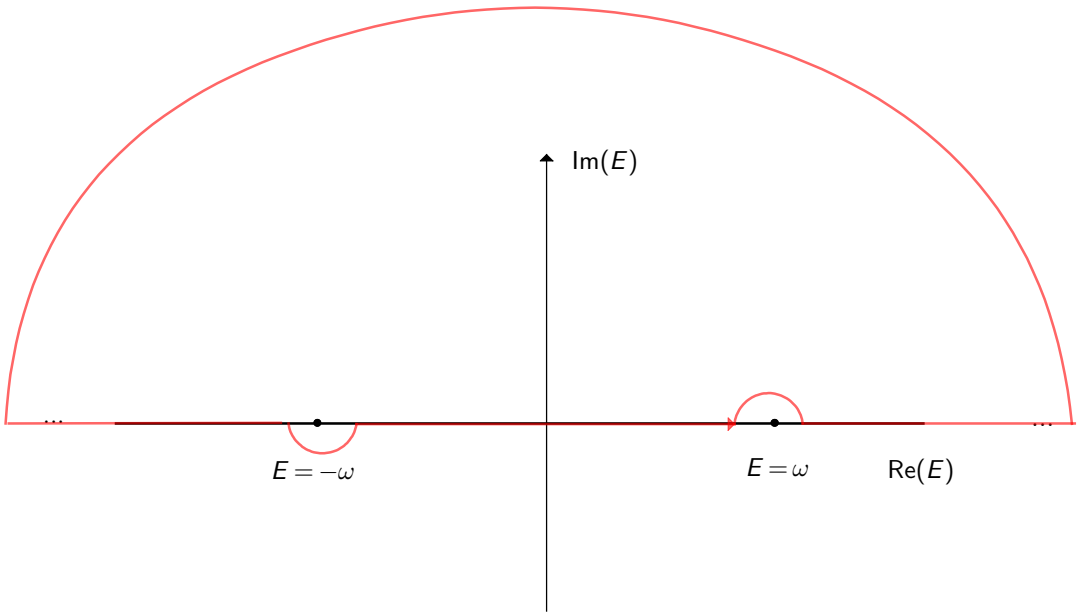
so the exponential function in the integrand of (16) vanishes for $\mathrm{Im}(E)\to+\infty$ and we consequently close our integration contour above. Including a pole does now imply encircling it counterclockwise. From the $E=\omega$ pole we would thus get (15)

$$G(x)=\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\frac{e^{i\omega|x^0|}}{-2\omega(\vec{p})}$$

corresponding to a negative wave number for a causal, forward in time $|x^0|$ propagation. At the $E=-\omega$ pole we would get a contribution

$$G(x)=\int\frac{d^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{x}}\frac{e^{-i\omega|x^0|}}{2\omega(\vec{p})}=D(-x)$$

corresponding to a positive wave number for forward in time propagation. We again chose the positive wave number propagation only and the corresponding contour looks like this:



Thus both of our choices for going around the poles in the $x^0 > 0$ and $x^0 < 0$ cases are identical. We can thus define a Green's function containing positive energy modes only and via (15) relate it to the propagator

$$G(x) = \begin{cases} D(x) & x^0 \geqslant 0 \\ D(-x) & x^0 \leqslant 0 \end{cases}$$

We call this Green's function with the specific contour prescription the Feynman propagator

$$D_F(x) \;:=\; \theta(x^0)D(x) + \theta(-x^0)D(-x)$$

Expressing the propagator in terms of the fields via (14) we obtain

$$D_F(x-y) = \langle 0|\, T(\phi(x)\phi(y))|0\rangle \tag{17}$$

where we have defined the time ordered product

$$T(\phi(x)\phi(y)) := \theta(x^0 - y^0)\phi(x)\phi(y) + \theta(y^0 - x^0)\phi(y)\phi(x)$$

which sorts a product of operators according to their time coordinate with the latest one appearing at the leftmost place. Note that the time ordering of fields is a relativistically consistent prescription since field operators at equal time do commute.

Although we have succeeded now in defining a consistent probability amplitude for single particle propagation that only includes positive energy modes propagating forward in time, the choice of integration contour seems somewhat arbitrary and ad-hoc. This may be mitigated by noticing that encircling the poles as described above can also be achieved by infinitessimally tilting the integration line off the real axis, i.e. integrating along

$$p^0 = \frac{r}{1 - i\delta} \qquad r \in \mathbb{R}$$

instead of simply $p^0 \in \mathbb{R}$. This line will go around the poles in the desired manner, but when closing the contour we will have problems as e.g. in the case $x^0 > 0$ we have

$$\mathrm{Re}(-ip^0x^0) < 0 \quad \text{for} \quad \mathrm{Im}(p^0) < 0$$

but the contour needs to be closed at the $r \to \infty$ side with a positive imaginary part. This can be rectified by also tilting the time axis infinitesimally off the real axis. Using

$$x^0 = t(1 - i\delta) \qquad \delta \in \mathbb{R}^+ \tag{18}$$

we now have

$$\mathrm{Re}(-ip^0x^0) = \mathrm{Re}(-irt) < 0 \quad \text{for} \quad \mathrm{Im}(r) < 0$$

and we thus can properly close the contour in the halfplane of a negative imaginary part of r.

Giving the time an imaginary component might seem even worse than choosing an ad-hoc contour at the moment, but we will soon see that in fact it is not only allowed but even mandatory when isolating the ground state of an interacting theory.

As a notational tool to remember how to go around the poles, it is customary to write in a slightly sloppy fashion

$$D_F(x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot x} \frac{i}{p^2 - m^2 + i\varepsilon} \tag{19}$$

7 Interacting fields

The Klein-Gordon equation is linear and thus its dynamics is trivial. We have explicitly seen that the different momentum modes decouple and thus there is no interaction. There is not even a measurement possible in this theory and thus it is not a good candidate to describe any naturalphenomena.

In order to introduce some nontrivial dynamics, we need to add some interaction. In quantum mechanics, we introduced a potential into the Schrödinger equation. This is not appropriate in a relativistic field theory as an action at a distance violates causality. We will therefore restrict the interaction terms to additionalterms in the Lagrange density.

7.1 Constructing an interaction term

In principle, we can add any local combination of constants, fields and their derivatives to the Lagrange density $\mathscr{L}$ that are Lorentz scalars and of mass dimension $[m]^4$, the latter requirement comming from the fact that the action $S = \int d^4x \mathscr{L}$ needs to be dimensionless. From the free Lagrangean

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

we can deduce that the mass dimension of the field is

$$[m^2\phi^2] = [m]^4 \qquad \Rightarrow \qquad [\phi] = [m]$$

The simplest thing we could do is adding a constant term to $\mathscr{L}$. It is easy to see however that this term does not affect the equations of motion and so is irrelevant.

Similarly, one could add a term $c\phi$ to $\mathscr{L}$. We could however write

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2) + c\phi = \frac{1}{2}\Big(\partial_\mu\phi\partial^\mu\phi - m^2\Big(\phi - \frac{c}{m^2}\Big)^2\Big) + \frac{c^2}{2m^2}$$

and thus simply absorb it into a field redefinition, which will again have no effect on the equations of motion. A quadratic term is already present in $\mathscr{L}$ and so is a quadratic term in the field derivatives. The first nontrivial term we can add thus is a term of the form ϕ^3. This term does have a problem though as it results in a Hamiltonean that is not bounded from below, which can be anticipated from the fact that ϕ^3 is not bounded from below.

Finally, we can add a ϕ^4 term to the Lagrangean density

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

which is nontrivial and results in a Hamiltonean that is bounded from below. We will thus proceed studying this ϕ^4 theory as an example of a nontrivial quantum field theory.

Exercise 3. Show that for a Lagrangean of the form $\mathscr{L} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2\phi^2) - \frac{\lambda}{n!}\phi^n$, $n \in \mathbb{N}$ the canonical equation of motion reads $(\partial_\mu \partial^\mu + m^2)\phi = -\frac{1}{(n-1)!}\phi^{n-1}$.

Exercise 4. Show that the Hamiltonean density of ϕ^4 theory is given by $\mathscr{H} = \frac{1}{2}(\Pi^2 + \partial_i\phi\partial_i\phi + m^2\phi^2) + \frac{\lambda}{4!}\phi^4$ and is thus bounded from below.

7.2 ϕ^4 theory

The Hamiltonean of ϕ^4 theory is given by

$$H(t) = \int d^3x \left(\frac{1}{2}(\Pi^2(t,\vec{x}) + \partial_i\phi(t,\vec{x})\partial_i\phi(t,\vec{x}) + m^2\phi^2(t,\vec{x})) + \frac{\lambda}{4!}\phi^4(t,\vec{x}) \right)$$

When going to Fourier space, the first three terms are diagonalized as for the KG theory, but the last term reads

$$\begin{aligned}
\int d^3x \phi^4(\vec{x}) &= \int d^3x \int \frac{d^3p_1}{(2\pi)^3} \int \frac{d^3p_2}{(2\pi)^3} \int \frac{d^3p_3}{(2\pi)^3} \int \frac{d^3p_4}{(2\pi)^3} e^{i(\vec{p}_1+\vec{p}_2+\vec{p}_3+\vec{p}_4)\cdot\vec{x}} \hat{\phi}(\vec{p}_1)\hat{\phi}(\vec{p}_2)\hat{\phi}(\vec{p}_3)\hat{\phi}(\vec{p}_4) \\
&= \int \frac{d^3p_1}{(2\pi)^3} \int \frac{d^3p_2}{(2\pi)^3} \int \frac{d^3p_3}{(2\pi)^3} \int d^3p_4 \delta^3(\vec{p}_1+\vec{p}_2+\vec{p}_3+\vec{p}_4)\hat{\phi}(\vec{p}_1)\hat{\phi}(\vec{p}_2)\hat{\phi}(\vec{p}_3)\hat{\phi}(\vec{p}_4) \\
&= \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p_2}{(2\pi)^3} \int \frac{d^3p_3}{(2\pi)^3} \hat{\phi}(\vec{p})\hat{\phi}(\vec{p}_2)\hat{\phi}(\vec{p}_3)\hat{\phi}(-\vec{p}-\vec{p}_2-\vec{p}_3)
\end{aligned}$$

making it explicit that momentum modes do not decouple. On the one hand this implies that the theory is not as simply solvable by Fourier Transformation as the KG theory, but on the other hand it implies that the theory has some nontrivial dynamics.

We can determine the dimensionality of λ by power counting. Comparing the $m^2\phi^2$ and ϕ^4 terms, we see that

$$[\phi^4] = [m^2\phi^2] = [\lambda\phi^4] \quad \Rightarrow \quad [\lambda] = 1$$

and thus the coupling constant λ is dimensionless.

8 Perturbation theory

In order to extract predictions from our theory, we want to compute the propagator

$$\langle\Omega|\, T(\phi(x)\phi(y))|\Omega\rangle$$

where $|\Omega\rangle$ denotes the ground state or vacuum of ϕ^4 theory. We introduce this notation to delimit $|\Omega\rangle$ from the vacuum of the KG theory, which we will continue to call $|0\rangle$. Since a closed form solution for this propagator is not easily obtainable, we will make use of perturbation theory. We assume $\lambda \ll 1$ and make an expansion around $\lambda = 0$. For $\lambda = 0$, the theory reduces to the well-known KG case, which we have solved exactly.

Formally, we decompose the Hamiltonean $H = H_0 + H_{\text{int}}$ into a trivial part

$$H_0(t) = \frac{1}{2}\int d^3x (\Pi^2(t,\vec{x}) + \partial_i\phi(t,\vec{x})\partial_i\phi(t,\vec{x}) + m^2\phi^2(t,\vec{x})) \tag{20}$$

and an interaction part

$$H_{\text{int}}(t) = \frac{\lambda}{4!}\int d^3x \phi^4(t,\vec{x})$$

8.1 The interaction picture

We can now decompose the time evolution of our fields ϕ into a trivial part from H_0 and a small correction due to H_{int}, which is proportional to λ. For this purpose we introduce the "interaction picture" field ϕ_I. This field coincides with the field ϕ at a reference time that we call t_0

$$\phi_I(t_0) = \phi(t_0)$$

but from this point on its time evolution is not given by the full Hamiltonean H but rather by the trivial piece H_0 only. This field is called interaction picture field because the time dependence is neither fully in the operator valued field (Heisenberg picture) nor is it fully in the states (Schrödinger picture). In the interaction picture, the trivial time dependence (sloppily speaking the part according to H_0) is absorbed into the operators (fields) whereas the time dependence due to the interaction Hamiltonean has to be accounted for by the time dependence of the states.

At the reference time t_0, the Heisenberg equation for the interaction picture observables O_I generically reads

$$\left.\frac{\partial O_I(t)}{\partial t}\right|_{t_0} = -i[O_I(t_0), H_0(t_0)]$$

Moving away from the reference time, the time evolutions for the free (H_0) and interacting (H) theory diverge. The Hamiltonean of the free theory will have the functional form (20), but with interaction picture fields (i.e. the fields following the free time evolution) instead of the ϕ

$$H_{0I}(t) = \frac{1}{2}\int d^3x (\Pi_I^2(t,\vec{x}) + \partial_i\phi_I(t,\vec{x})\partial_i\phi_I(t,\vec{x}) + m^2\phi_I^2(t,\vec{x}))$$

Similarly, we define the interaction part of the Hamiltonian in terms of the interaction picture fields

$$H_I(t) := \frac{\lambda}{4!}\int d^3x \phi_I^4(t,\vec{x})$$

The advantage of this definition is that the time evolution of ϕ_I is identical to that of a KG field (13), thus the interaction picture fields

$$\phi_I(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}(e^{-ip\cdot x}a(\vec{p}) + e^{ip\cdot x}a^\dagger(\vec{p})) \tag{21}$$

can be written in terms of creation and annihilation operators (again with the implied $p^0 = \sqrt{\vec{p}^2 + m^2}$).

The Heisenberg equation for interaction picture quantities O_I at arbitrary times thus reads

$$\frac{\partial O_I(t)}{\partial t} = -i[O_I(t), H_{0I}(t)]$$

and thus an infinitesmial time evolution has the form

$$\begin{aligned} O_I(t+dt) &= O_I(t) - i[O_I(t), H_{0I}(t)] \\ &= (1 + iH_{0I}(t)dt)O_I(t)(1 - iH_{0I}(t)dt) \\ &= e^{iH_{0I}(t)dt}O_I(t)e^{-iH_{0I}(t)dt} \end{aligned}$$

Defining a unitary time evolution operator as the time ordered product

$$V(t) := T\left(e^{i\int_{t_0}^t d\tau H_{0I}(\tau)}\right)$$

and assuming $dt > 0$, we can write

$$O_I(t) = V(t) O_I(t_0) V^\dagger(t)$$

or

$$\begin{aligned} V^\dagger(t) O_I(t) &= O_I(t_0) V^\dagger(t) \\ O_I(t) V(t) &= V(t) O_I(t_0) \end{aligned}$$

In case of $dt < 0$ we have reverse time ordering, i.e.

$$V(t) := T^\dagger\left(e^{i \int_{t_0}^t d\tau H_{0I}(\tau)} \right)$$

The Heisenberg equation of the full field ϕ reads

$$\frac{\partial O(t)}{\partial t} = -i[O(t), H(t)]$$

and thus

$$O(t + dt) = e^{iH(t)dt} O(t) e^{-iH(t)dt}$$

Similar to the case above, we define a time evolution operator

$$W(t) := \begin{cases} T\left(e^{i \int_{t_0}^t d\tau H(\tau)} \right) & t > t_0 \\ T^\dagger\left(e^{i \int_{t_0}^t d\tau H(\tau)} \right) & t < t_0 \end{cases}$$

we can write

$$O(t) = W(t) O(t_0) W^\dagger(t)$$

or

$$\begin{aligned} W^\dagger(t) O(t) &= O(t_0) W^\dagger(t) \\ O(t) W(t) &= W(t) O(t_0) \end{aligned}$$

Remembering that the fields are equal at the reference time t_0, we have

$$\begin{aligned} \phi(t) &= W(t) \phi(t_0) W^\dagger(t) \\ &= W(t) V^\dagger(t) \phi_I(t) V(t) W^\dagger(t) \\ &= U^\dagger(t) \phi_I(t) U(t) \end{aligned} \tag{22}$$

where we have defined the unitary operator that transforms the field to the interaction picture

$$U(t) := V(t) W^\dagger(t) \tag{23}$$

Let us look at this operator in a bit more detail. Going an infinitesimal time step further, we have

$$\begin{aligned} U(t + dt) &= V(t + dt) W^\dagger(t + dt) \\ &= (1 + i H_{0I}(t) dt) V(t) W^\dagger(t) (1 - i H(t) dt) \\ &= U(t) + i \left(H_{0I}(t) V(t) W^\dagger(t) - V(t) W^\dagger(t) H(t) \right) dt \\ &= U(t) + i \left(V(t) H_{0I}(t_0) W^\dagger(t) - V(t) H(t) W^\dagger(t) \right) dt \\ &= U(t) + i \left(V(t) H_{0I}(t_0) W^\dagger(t) - V(t) H(t_0) W^\dagger(t) \right) dt \\ &= U(t) - i V(t) H_I(t_0) W^\dagger(t) dt \\ &= U(t) - i V(t) H_I(t_0) V^\dagger(t) U(t) dt \\ &= (1 - i H_I(t) dt) U(t) \end{aligned}$$

and thus we can construct the transformation operator from the interaction picture to the full theory as a reverse time ordered product of exponentials of the interaction Hamiltonean in the interaction picture

$$U(t) = \begin{cases} T\Big(e^{-i\int_{t_0}^{t} d\tau H_I(\tau)}\Big) & t > t_0 \\ T^\dagger\Big(e^{-i\int_{t_0}^{t} d\tau H_I(\tau)}\Big) & t < t_0 \end{cases}$$

8.2 The full propagator in terms of the interaction picture fields

We now return to computing the propagator of the full, interacting theory

$$\langle\Omega|\, T(\phi(x)\phi(y))|\Omega\rangle$$

One immediate problem we face is that we do not know the true vacuum $|\Omega\rangle$ of the theory. Assuming that the vacuum state is unique (which will not be true for all theories) and further assuming that the theory has a mass gap, i.e. that the first excited state above the vacuum has a finite energy difference to the vacuum, we can however isolate the vacuum of the theory $|\Omega\rangle$. For this purpose, we first take a state we know from the free KG theory, e.g. the vacuum of the free KG theory $|0\rangle$ at a certain time t. In the interaction picture, this state is time invariant (since there we are in the Heisenberg picture of the KG theory). To transform it into a full Heisenberg picture state, we apply the transformation operator (23) on it

$$U^\dagger(t)|0\rangle$$

Assuming that we are interested in times before our reference time, i.e. $t < t_0$, we have

$$\begin{aligned} U^\dagger(t)|0\rangle &= W(t)V^\dagger(t)|0\rangle \\ &= W(t)T\Big(e^{-i\int_{t_0}^{t} d\tau H_{0I}(\tau)}\Big)|0\rangle \\ &= W(t)e^{-iE_0(t-t_0)}|0\rangle \\ &= W(t)|0\rangle \\ &= T^\dagger\Big(e^{i\int_{t_0}^{t} d\tau H(\tau)}\Big)|0\rangle \end{aligned}$$

where in the second line we have used the fact that $|0\rangle$ is an eigenstate of H_{0I} independent of time and in the third line we have used that its energy $E_0 = 0$ vanishes. To proceed further, we will insert a complete set of eigenstates $|\hat{i}\rangle$ with energies $E_{\hat{i}}$ of the full Hamiltonian H

$$\begin{aligned} U^\dagger(t)|0\rangle &= \sum_i T^\dagger\Big(e^{i\int_{t_0}^{t} d\tau H(\tau)}\Big)|\hat{i}\rangle\langle\hat{i}|0\rangle_t \\ &= \sum_i e^{iE_{\hat{i}}(t-t_0)}|\hat{i}\rangle\langle\hat{i}|0\rangle_t \end{aligned}$$

where the sum over i denotes a sum/integral over all energy eigenstates of the full theory and we have used the fact that $|\hat{i}\rangle$ is an eigenstate of H independent of time. The subscript t on the matrix element $\langle\hat{i}|0\rangle_t$ denotes that it has to be taken at time t, which is important since $|0\rangle$ and $|\hat{i}\rangle$ are time independent eigenstates of different Hamiltoneans.

For real t, different energy components thus obtain different phases. We now remember however, that we had to give time a small imaginary component in order to obtain an unambiguous expression for the propagator in the KG theory (18). Plugging this in and shifting our the time t to the infinite past, we obtain

$$\begin{aligned} \lim_{(t_0-t)\to\infty(1-i\delta)} U^\dagger(t)|0\rangle &= \lim_{(t_0-t)\to\infty(1-i\delta)} \sum_i e^{iE_{\hat{i}}(t-t_0)}|\hat{i}\rangle\langle\hat{i}|0\rangle_t \\ &= \lim_{(t_0-t)\to\infty} \sum_i e^{iE_{\hat{i}}(t-t_0)(1-i\delta)}|\hat{i}\rangle\langle\hat{i}|0\rangle_{t(1-i\delta)} \\ &= \lim_{(t_0-t)\to\infty} \sum_i e^{-E_{\hat{i}}(t_0-t)\delta}e^{iE_{\hat{i}}(t-t_0)}|\hat{i}\rangle\langle\hat{i}|0\rangle_{t(1-i\delta)} \end{aligned}$$

When shifting the time t to the infinite past, all modes will thus be infinitely suppressed with respect to the lowest lying one. Since this lowest energy mode is our vacuum $|\Omega\rangle = |\widehat{0}\rangle$ and under the additional assumption that $\langle\Omega|0\rangle_{t(1-i\delta)} \neq 0$ (which is a very reasonable assumption since we these are vacua of theories that only differ by a small perturbation) we thus obtain

$$U^\dagger(t)|0\rangle \xrightarrow[t-t_0\to-\infty(1-i\delta)]{} e^{iE_\Omega(t-t_0)}|\Omega\rangle\langle\Omega|0\rangle_t$$

leading to an explicit expression for the vacuum of the full theory

$$\begin{aligned}
|\Omega\rangle &= \lim_{(t_0-t)\to\infty(1-i\delta)} \langle\Omega|0\rangle_t^{-1} e^{-iE_\Omega(t-t_0)} U^\dagger(t)|0\rangle \\
&= \lim_{(t_0-t)\to\infty(1-i\delta)} c_t T\Big(e^{i\int_{t_0}^t d\tau H_I(\tau)}\Big)|0\rangle \\
&= \lim_{(t_0-t)\to\infty(1-i\delta)} c_t T\big(e^{-i\int_t^{t_0} d\tau H_I(\tau)}\big)|0\rangle
\end{aligned}$$

where we have defined the constant

$$c_t := \langle\Omega|0\rangle_t^{-1} e^{-iE_\Omega(t-t_0)}$$

In complete analogy, we may express $\langle\Omega|$ as

$$\langle\Omega| = \lim_{(t-t_0)\to\infty(1-i\delta)} c_t^\dagger \langle 0| T\Big(e^{-i\int_{t_0}^t d\tau H_I(\tau)}\Big)$$

and thus we obtain for the propagator

$$\begin{aligned}
\langle\Omega|T(\phi(x)\phi(y))|\Omega\rangle &= \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| T\Big(e^{-i\int_{t_0}^t d\tau H_I(\tau)}\Big) T(\phi(x)\phi(y)) T\big(e^{-i\int_{-t}^{t_0} d\tau H_I(\tau)}\big)|0\rangle \\
&= \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| U(t) T(\phi(x)\phi(y)) U^\dagger(-t)|0\rangle
\end{aligned}$$

where we have implied the small imaginary part on all time variables for notational convenience. This expression may be further simplified by expressing the Heisenberg picture fields ϕ in terms of the interaction picture fields ϕ_I via (22).

$$\begin{aligned}
\langle\Omega|T(\phi(x)\phi(y))|\Omega\rangle &= \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| U(t) T(\phi(x)\phi(y)) U^\dagger(-t)|0\rangle \\
&= \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| U(t) T(U^\dagger(x^0)\phi_I(x)U(x^0)U^\dagger(y^0)\phi_I(y)U(y^0)) U^\dagger(-t)|0\rangle \\
&= \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| T(U(t)\phi_I(x)\phi_I(y)U^\dagger(-t))|0\rangle \\
&= \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| T\Big(\phi_I(x)\phi_I(y) e^{-i\int_{-t}^t d\tau H_I(\tau)}\Big)|0\rangle
\end{aligned}$$

Obviously, this expression generalizes to generic time ordered products of any number of fields. In particular, we can thus compute

$$1 = \langle\Omega|\Omega\rangle = \lim_{t\to\infty} c_t^\dagger c_{-t} \langle 0| T\big(e^{-i\int_{-t}^t d\tau H_I(\tau)}\big)|0\rangle$$

to obtain our final expression for the propagator

$$\langle\Omega|T(\phi(x)\phi(y))|\Omega\rangle = \frac{\lim_{t\to\infty}\langle 0| T\Big(\phi_I(x)\phi_I(y) e^{-i\int_{-t}^t d\tau H_I(\tau(1-i\delta))}\Big)|0\rangle}{\lim_{t\to\infty}\langle 0| T\big(e^{-i\int_{-t}^t d\tau H_I(\tau(1-i\delta))}\big)|0\rangle} \tag{24}$$

where we have again explicitly written the small imaginary tilt in the time axis.

9 Wicks theorem

In order to compute propagators in our interacting theory, we could now simply expand the r.h.s. of (24) to any desired order λ, leaving us with generic vacuum expectation values of time ordered products of interaction picture fields

$$\langle 0|\, T(\phi_I(x_1)\phi_I(x_2)...\phi(x_n))|0\rangle$$

To evaluate these expressions, we remember that interaction picture fields simply evolve according to the free KG equation and can be expressed in terms of creation and annihilation operators as (21)

$$\begin{aligned}\phi_I(x) &= \phi_I^+(x)+\phi_I^-(x)\\ \phi_I^+(x) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}e^{-ip\cdot x}a(\vec{p})\\ \phi_I^-(x) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}e^{ip\cdot x}a^\dagger(\vec{p})\end{aligned}$$

where we have explicitly separated the components with creation (ϕ_I^+) and annihilation (ϕ_I^-) operators. This is useful because these components annihilate the vacuum on one side, i.e.

$$\phi_I^+(x)|0\rangle = 0 \qquad \langle 0|\phi_I^-(x) = 0$$

Let us write explicitly the time ordered product of two interaction picture fields

$$T(\phi_I(x)\phi_I(y)) = \begin{cases} \phi_I^+(x)\phi_I^+(y)+\phi_I^-(x)\phi_I^+(y)+\phi_I^+(x)\phi_I^-(y)+\phi_I^-(x)\phi_I^-(y) & x^0>y^0\\ \phi_I^+(y)\phi_I^+(x)+\phi_I^-(y)\phi_I^+(x)+\phi_I^+(y)\phi_I^-(x)+\phi_I^-(y)\phi_I^-(x) & y^0>x^0\end{cases}$$

Because all creation and annihilation operators commute with each other, $\phi_I^+(x)\phi_I^+(y) = \phi_I^+(y)\phi_I^+(x)$ and $\phi_I^-(x)\phi_I^-(y) = \phi_I^-(y)\phi_I^-(x)$. Thus, for $x^0>y^0$, all terms in the above expression are in normal order except $\phi_I^+(x)\phi_I^-(y)$ and for $y^0>x^0$, all terms are in normal order except $\phi_I^+(y)\phi_I^-(x)$. We can thus write

$$T(\phi_I(x)\phi_I(y)) = :\phi_I(x)\phi_I(y): + \begin{cases} [\phi_I^+(x),\phi_I^-(y)] & x^0>y^0\\ [\phi_I^+(y),\phi_I^-(x)] & y^0>x^0\end{cases}$$

The second term in the above expression is the "Wick contraction", which we denote by[1]

$$\overset{1}{\varphi}_I(x)\overset{1}{\varphi}_I(y) := \begin{cases} [\phi_I^+(x),\phi_I^-(y)] & x^0>y^0\\ [\phi_I^+(y),\phi_I^-(x)] & y^0>x^0\end{cases}$$

thus we have

$$T(\phi_I(x)\phi_I(y)) = :\phi_I(x)\phi_I(y): + \overset{1}{\varphi}_I(x)\overset{1}{\varphi}_I(y) \tag{25}$$

Taking the vacuum expectation value of the above expression, we see that

$$\langle 0|\, T(\phi_I(x)\phi_I(y))|0\rangle = \langle 0|\overset{1}{\varphi}_I(x)\overset{1}{\varphi}_I(y)|0\rangle$$

1. The usual notation for wick contractions is an over- or underbrace. Because of typsetting restrictions, we will use above or below script numbers with the implication that equal numbers are contracted.

and thus, according to (17) is equal the feynman propagator. In addition, the commutator of two fields and thus the Wick contraction is proportional to the unit operator and thus the Wick contraction itself is equal to the Feynman propagator

$$D_F(x-y) = \overset{1}{\varphi}_I(x)\overset{1}{\varphi}_I(y)$$

The result (25) does in fact generalize to arbitrary time ordered products of interaction picture fields. This is the Wick theorem which, sloppily stated, is the equivalence of the time ordered product of interaction picture operators to the sum of the normal ordered product plus all possible contractions. More formally, the Wick theorem states that given a number of interaction picture fields at points x_i for which we will use the shorthand notation

$$\phi_i := \phi_I(x_i)$$

we have

$$\begin{aligned} T(\phi_1...\phi_N) &= :\phi_1...\phi_N: + \\ &\overset{1}{\varphi}_1\overset{1}{\varphi}_2:\phi_3...\phi_N: + \overset{1}{\varphi}_1\overset{1}{\varphi}_3:\phi_2\phi_4...\phi_N: + ... + \overset{1}{\varphi}_{N-1}\overset{1}{\varphi}_N:\phi_1...\phi_{N-2}: + \\ &\overset{1}{\varphi}_1\overset{1}{\varphi}_2\overset{2}{\varphi}_3\overset{2}{\varphi}_4:\phi_5...\phi_N: + ... + \overset{1}{\varphi}_{N-3}\overset{1}{\varphi}_{N-2}\overset{2}{\varphi}_{N-1}\overset{2}{\varphi}_N:\phi_1...\phi_{N-4}: + \\ &... + \\ &\overset{1}{\varphi}_1\overset{1}{\varphi}_2\overset{2}{\varphi}_3\overset{2}{\varphi}_4...\overset{N/2}{\varphi}_{N-1}\overset{N/2}{\varphi}_N + ... + \overset{1}{\varphi}_1\overset{2}{\varphi}_2\overset{3}{\varphi}_3\overset{4}{\varphi}_4...\overset{2}{\varphi}_{N-1}\overset{1}{\varphi}_N + \end{aligned}$$

The most important feature of this identity is exposed when taking the vacuum expectation value. Since contractions are simply numbers and any normal ordered product has zero vacuum expectation value, only the last line is remaining on the r.h.s. of the above equation and we obtain

$$\langle 0|T(\phi_1...\phi_N)|0\rangle = \langle 0|\overset{1}{\varphi}_1\overset{1}{\varphi}_2\overset{2}{\varphi}_3\overset{2}{\varphi}_4...\overset{N/2}{\varphi}_{N-1}\overset{N/2}{\varphi}_N|0\rangle + ... + \langle 0|\overset{1}{\varphi}_1\overset{2}{\varphi}_2\overset{3}{\varphi}_3\overset{4}{\varphi}_4...\overset{2}{\varphi}_{N-1}\overset{1}{\varphi}_N|0\rangle$$

or, in words, the vacuum expectation value of the time ordered product of interaction picture fields is equal to the sum of all their Wick contractions.

We can prove the Wick theorem by complete incuction. We have explicitly seen that the theorem holds for two fields. Let us now assume that it works for $N-1$ fields $\phi_2...\phi_N$ and that, without loss of generality, we have labeled our coordinnates in a time ordered way such that $x_1^0 > x_2^0 > ... > x_N^0$. Then

$$T(\phi_1\phi_2...\phi_N) = \phi_1\phi_2...\phi_N = (\phi_1^+ + \phi_1^-)\phi_2...\phi_N$$

and we can use the Wick theorem for the product $\phi_2...\phi_N$. We obtain

$$\begin{aligned} T(\phi_1...\phi_N) &= (\phi_1^+ + \phi_1^-):\phi_2...\phi_N: + \\ &(\phi_1^+ + \phi_1^-)\overset{1}{\varphi}_2\overset{1}{\varphi}_3:\phi_4...\phi_N: + ... + (\phi_1^+ + \phi_1^-)\overset{1}{\varphi}_{N-1}\overset{1}{\varphi}_N:\phi_2...\phi_{N-2}: + \\ &... + \\ &(\phi_1^+ + \phi_1^-)\overset{1}{\varphi}_2\overset{1}{\varphi}_3\overset{2}{\varphi}_4\overset{2}{\varphi}_5...\overset{N/2}{\varphi}_{N-1}\overset{N/2}{\varphi}_N + ... + (\phi_1^+ + \phi_1^-)\overset{1}{\varphi}_2\overset{2}{\varphi}_3\overset{3}{\varphi}_4\overset{4}{\varphi}_5...\overset{2}{\varphi}_{N-1}\overset{1}{\varphi}_N \end{aligned}$$

The field ϕ_1^- is already at the leftmost position, i.e. it is in the correct position for any normal ordered product. It can thus simply be absorbed into any normal ordered products. In addition, contractions are scalar and thus commute with the fields, so we obtain

$$\begin{aligned} T(\phi_1...\phi_N) &= :\phi_1^-\phi_2...\phi_N: + \phi_1^+:\phi_2...\phi_N: + \\ &\overset{1}{\varphi}_2\overset{1}{\varphi}_3:\phi_1^-\phi_4...\phi_N: + \overset{1}{\varphi}_2\overset{1}{\varphi}_3\,\phi_1^+:\phi_4...\phi_N: + ... + \\ &... + \\ &\phi_1^-\overset{1}{\varphi}_2\overset{1}{\varphi}_3\overset{2}{\varphi}_4\overset{2}{\varphi}_5...\overset{N/2}{\varphi}_{N-1}\overset{N/2}{\varphi}_N + \phi_1^+\overset{1}{\varphi}_2\overset{1}{\varphi}_3\overset{2}{\varphi}_4\overset{2}{\varphi}_5...\overset{N/2}{\varphi}_{N-1}\overset{N/2}{\varphi}_N + ... + \end{aligned}$$

It now remains to commute the ϕ_1^+ all the way to the rightmost place in all normal ordered products. Picking the first (and longest) term we obtain

$$\begin{aligned}\phi_1^+ : \phi_2 ... \phi_N : &= : \phi_2 ... \phi_N : \phi_1^+ + [\phi_1^+, : \phi_2 ... \phi_N :] \\ &= : \phi_1^+ \phi_2 ... \phi_N : + [\phi_1^+, : \phi_2 ... \phi_N :]\end{aligned}$$

For computing the commutator term in the expression above, we can again use the fact that the commutator between fields is proporeional to the unit operator and thus commutes with all other fields. We obtain

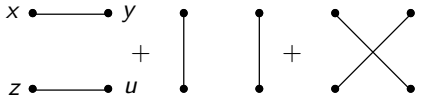
$$\begin{aligned}[\phi_1^+, : \phi_2 ... \phi_N :] &= [\phi_1^+, \phi_2] : \phi_3 ... \phi_N : + [\phi_1^+, \phi_3] : \phi_2 \phi_4 ... \phi_N : + ... + [\phi_1^+, \phi_N] : \phi_2 ... \phi_{N-1} : \\ &= \overset{1}{\varphi}_1 \overset{1}{\varphi}_2 : \phi_3 ... \phi_N : + \overset{1}{\varphi}_1 \overset{1}{\varphi}_3 : \phi_2 \phi_4 ... \phi_N : + ... + \overset{1}{\varphi}_1 \overset{1}{\varphi}_N : \phi_2 ... \phi_{N-1} :\end{aligned}$$

i.e., in addition to the normal ordering we obtain all the contraction terms including ϕ_1, thus proving that the Wick theorem is fulfilled for the product of N fields $\phi_2 ... \phi_N$.

Example 4. As a simple example, let us use the Wick theorem to compute the vacuum expectation value of a time ordered product of four interaction picture fields:

$$\begin{aligned}\langle 0 | T(\phi_I(x)\phi_I(y)\phi_I(z)\phi_I(u)) | 0 \rangle &= \overset{1}{\varphi}_I(x)\overset{1}{\varphi}_I(y)\overset{2}{\varphi}_I(z)\overset{2}{\varphi}_I(u) + \\ &\quad \overset{1}{\varphi}_I(x)\overset{2}{\varphi}_I(y)\overset{1}{\varphi}_I(z)\overset{2}{\varphi}_I(u) + \\ &\quad \overset{1}{\varphi}_I(x)\overset{2}{\varphi}_I(y)\overset{2}{\varphi}_I(z)\overset{1}{\varphi}_I(u) \\ &= D_F(x-y)D_F(z-u) + \\ &\quad D_F(x-z)D_F(y-u) + \\ &\quad D_F(x-u)D_F(y-z)\end{aligned}$$

We can represent this graphically with lines representing the Feynman propagator between two points as follows:

$\langle 0 | T(\phi_I(x)\phi_I(y)\phi_I(z)\phi_I(u)) | 0 \rangle =$ 

10 Perturbative expansion of the full propagator

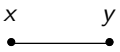
The wick theorem now allows us to write the perturbative expansion of the full propagator in terms of the Feynman propagator for the KG field. We will from now on drop the subscript I on interaction picture fields - all fields are meant to be in the interaction picture unless otherwise noted. Let us first expand the numerator of (24) to first order in λ:

$$\begin{aligned}\langle 0 | T\Big(\phi(x)\phi(y)e^{-i\int_{-\infty}^{\infty} d\tau H_I(\tau(1-i\delta))}\Big) | 0 \rangle &= \langle 0 | T\Big(\phi(x)\phi(y)e^{-i\frac{\lambda}{4!}\int d^4z \phi^4(z)}\Big) | 0 \rangle \\ &= \langle 0 | T(\phi(x)\phi(y)) | 0 \rangle - \\ &\quad i\frac{\lambda}{4!}\int d^4z \langle 0 | T\Big(\phi(x)\phi(y)\frac{d^4q}{(2\pi)^4}\phi^4(z)\Big) | 0 \rangle + O(\lambda^2)\end{aligned}$$

The 0^{th} order term simply is

$$\langle 0 | T(\phi(x)\phi(y)) | 0 \rangle = D_F(x-y)$$

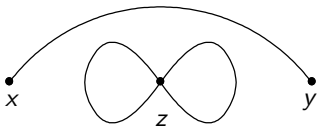
which has the simple graphical representation

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

For the first order term, we have a larger number of contractions because there are 6 fields involved that are at 3 different points (x, y and z). First, we can have contractions of the form

$$\overset{1}{\varphi}(x)\overset{1}{\varphi}(y)\overset{2}{\varphi}(z)\overset{2}{\varphi}(z)\overset{3}{\varphi}(z)\overset{3}{\varphi}(z)$$

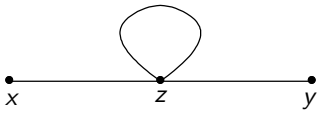
which can be graphically represented as



It is easy to enumerate the number of contractions of this form. First of all, the contraction between x and y is unique. For the first $\phi(z)$ there are then 3 possibilities left: it can either contract with the second, third or fourth $\phi(z)$. In each of these cases, there are two $\phi(z)$ left that have to be contracted with each other. We therefore have a total of 3 possibilities. Explicitly written, the Wick contractions are

$$\begin{array}{l}\overset{1}{\varphi}(x)\overset{1}{\varphi}(y)\overset{2}{\varphi}(z)\overset{2}{\varphi}(z)\overset{3}{\varphi}(z)\overset{3}{\varphi}(z)\\ \overset{1}{\varphi}(x)\overset{1}{\varphi}(y)\overset{2}{\varphi}(z)\overset{3}{\varphi}(z)\overset{2}{\varphi}(z)\overset{3}{\varphi}(z)\\ \overset{1}{\varphi}(x)\overset{1}{\varphi}(y)\overset{2}{\varphi}(z)\overset{3}{\varphi}(z)\overset{3}{\varphi}(z)\overset{2}{\varphi}(z)\end{array}$$

The other possibility is to contract $\phi(x)$ with one of the $\phi(z)$ - for which there are 4 possibilities. Consequently, $\phi(y)$ has to be contracted with one of the remaining $\phi(z)$, for which there are 3 possibilities left. The remaining contraction of two $\phi(z)$ is then unique, so that there are all together 12 such terms. Diagramatically, they can be represented like this:

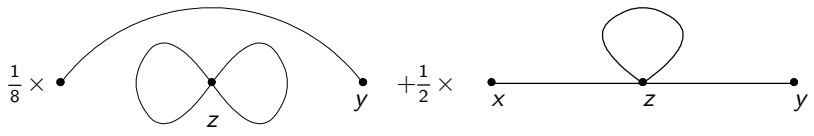


For the first order in λ expression we thus have

$$\begin{aligned}-i\frac{\lambda}{4!}\langle 0|T(\phi(x)\phi(y)\phi^4(z))|0\rangle &= -i\frac{\lambda}{4!}(3D_F(x-y)D_F^2(z-z)+12D_F(x-z)D_F(z-y)D_F(z-z))\\ &= -\frac{i\lambda}{8}D_F(x-y)D_F^2(0)-\frac{i\lambda}{2}D_F(x-z)D_F(z-y)D_F(0)\end{aligned}$$

or, diagrammatically

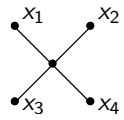
$-i\frac{\lambda}{4!}\int d^4z\,\langle 0|T(\phi(x)\phi(y)\phi^4(z))|0\rangle =$



11 Feynman graphs

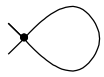
We have already seen that Wick contractions can be conveniently displayed graphically. This correspondence can be made exact with a set of Feynman rules and the corresponding graphs are called Feynamn graphs. Each graph corresponds exactly to one contraction with each element having a precise meaning.

Let us start with the simplest element, an external vertex: •. These are points where propagators "start" and "end" and they simply contribute a trivial factor 1 to a Feynman graph. The next element we know already is the propagator _____ running between two vertices x and y. It represents the Feynman propagator between these points $D_F(x-y)$. As we have seen above, an internal vertex ✕ always comes with the coupling $-i\lambda$ as well as an integration over its coordinate, z in the examples above. In addition, an internal vertex comes with a symmetry factor. This symmetry factor originates from the perturbative expansion of $e^{-i\frac{\lambda}{4!}\int d^4z\phi^4(z)}$. Expanding this expression to order n we obtain n internal vertices with a total symmetry factor $\frac{1}{n!}\left(\frac{1}{4!}\right)^n$. The $\frac{1}{n!}$ factor is cancelled simply by not distinguishing the dummy integration variables in a Feynman diagram, i.e. by a permutation of the internal vertices. The remaining $\frac{1}{4!}$ factor for each internal vertex is then cancelled by permuting the "leg" of the vertex that attaches to an incoming propagator. In the simple case depicted here



there are 4 different legs to choose from for attaching to x_1, 3 remaining legs for x_2 and 2 for x_3. All together we have 4! different possibilities, which exactly cancels the $\frac{1}{4!}$ factor in the expansion. (Note that this argument extends trivially to any number of equivalent legs attached to an internal vertex.)

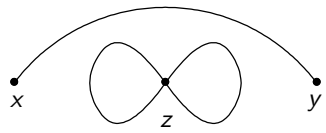
It is evident that the last argument only holds if the legs attach to distinguishable points in the graph. We can imagine two kinds of situatuions where this is violated: First, a propagator might start and end in the same vertex:



In this case, the naive argument from above is overcounting by a factor of 2 the number of graphs because there is only a single contraction between the two legs, not two. Similarly, if more than one propagator is connecting two vertices as in

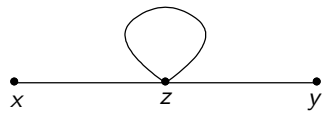
the number of distinct graphs is overcounted by a factor $k!$ with the simple argument above. With these two exceptions however, the cancellation of the $\frac{1}{4!}$ is valid, so we can formulate our Feynman rule for an internal vertex as giving a factor $-i\lambda$, integrating over the vertex position and incorporating a symmetry factor $\frac{1}{2}$ for every propagator staring and ending in the same vertex and $\frac{1}{k!}$ for each instance of k propagators that have the same start and ending vertices.

Example 5. As an example for the symmetry factors, we may look at the two Feynman graphs we have already drawn:



In this case, we have two quark lines starting and ending in the same point, so we get a factor $\frac{1}{2}$ for each of them. In addition, there are 2 propagators with the same starting- and ending vertices resulting in a third factor $\frac{1}{2}$. The total symmetry factor of this graph is therefore $\frac{1}{8}$, in accordance to what we have found by counting Wick contractions.

The second graph



has one propagator starting and ending in the same point, so its symmetry factor is $\frac{1}{2}$, also in accordance with the result from counting Wick contractions

As a summary, let us list the coordinate space Feynman rules for ϕ^4 theory that we have just found:

x $\bullet$ y	$D_F(x-y)$
z ✕	$-i\lambda\int d^4z$
each line starting and ending at the same point	$\frac{1}{2}$
k lines between same points	$\frac{1}{k!}$

11.1 Momentum space Feynman rules

Let us look at an arbitrary propagator in a Feynman diagram that starts at a point x and ends in y. The propagator is

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

Since the propagator always starts and ends in an (external or internal) vertex, we can associate the Fourier factor e^{-ipx} with the vertex at x (where the momentum is outgoing) and the factor e^{ipy} with the vertex at y (where the momentum is incoming) instead of with the propagator. A propagator _____ will thus denote $\frac{i}{p^2-m^2+i\varepsilon}$, i.e. the momentum space propagator at momentum p, together with an implied integral $\int\frac{d^4p}{(2\pi)^4}$ over the momentum. At any internal vertex z however we will have - if we count all momenta p_i as incoming

$$-i\lambda\int d^4z\,e^{iz\sum_i p_i}=-i\lambda(2\pi)^4\delta^4\Big(\sum_i p_i\Big)$$

The factor $(2\pi)^4\delta^4(\sum_i p_i)$ demands momentum conservation at the vertex by eliminating one of the $\int\frac{d^4p_i}{(2\pi)^4}$. Thus in momentum space, an internal vertex represents a factor $-i\lambda$ and momentum conservation. For an external vertex at x, we simply have e^{-ipx} where the momentum is treated as outgoing from the external vertex.

This would complete our set of momentum space Feynman rules, but it is slightly awkward to have represented everything in momentum space except for the external vertices, which are still at a definite space-time point. We recall that these external vertices correspond to the points x_i at which fields are placed in the original vacuum expectation value of the full theory

$$\langle\Omega|\,T(\phi(x_1)...\phi(x_N))|\Omega\rangle$$

thus they are the creation or annihilation points of particles. In reality, we will more frequently be confronted with a situation where particles will be created in (approximate) momentum rather than position states. Although there are some subtelties in relating these amplitudes to experimental observables that we will investigate later, we can for now simply take the Fourier transform of the external vertex positions

$$\langle\Omega|\,T(\hat\phi(k_1)...\hat\phi(k_N))|\Omega\rangle:=\int\prod_{i=1}^N d^4x_i e^{ik_ix_i}\langle\Omega|\,T(\phi(x_1)...\phi(x_N))|\Omega\rangle$$

Recalling that for every propagator attached to an external vertex on one side we have an integral $\int\frac{d^4p_i}{(2\pi)^4}$ remaining and from that vertex we have an $e^{-ip_ix_i}$ factor, an external momentum line _ thus simply represents

$$\int\frac{d^4p_i}{(2\pi)^4}d^4x_i e^{-i(p_i-k_i)x_i}\;=\;\int d^4p_i\delta^4(p_i-k_i)\;=\;1$$

There is a special case where a propagator is connected to two external and no internal vertices. In this case we can rewrite the propagator as

$$\begin{aligned} D_F(x-y) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{i}{p^2-m^2+i\varepsilon} \\ &= \int \frac{d^4p_1}{(2\pi)^4} e^{-ip_1\cdot x} \int \frac{d^4p_2}{(2\pi)^4} e^{ip_2\cdot y} \frac{i}{p^2-m^2+i\varepsilon} (2\pi)^4\delta^4(p_1-p_2) \end{aligned}$$

which serves to show that also in this case we can shift an overall $\int \frac{d^4p_i}{(2\pi)^4} e^{-ip_ix_i}$ to every external vertex for the price of an additional momentum conservation $(2\pi)^4\delta^4(p_1-p_2)$.

Note that in the deriavation above we have cheated a bit. We assumed that single particle momentum states are created by the Fourier transform of the fields $\hat{\phi}(k_i)$. This was true in the free theory, but a priori it is not clear at all whether this is also true for the interacting theory. Intuitively we guess that if the state is sufficiently far away from everything else (e.g. an incoming paricle in a scattering experiment in the asymptotic past) this should be true. It is in fact true under some reasonable assumptions and the corresponding theorem is the Lehman-Symanzik-Zimmermann (LSZ) reduction formula that we will investigate in some more detail later.

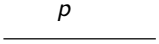
Now that we have external momentum lines, it is even possible to hook them up to an internal vertex directly instead of to an external vertex. If the original vertex position was labeled by z, such an external momentum line at momentum p simply had a factor of e^{ipz}, which gets absorbed by the vertex. Therefore, a momentum line hooked up to an internal vertex directly also just gives a factor 1. Feynman diagrams that have momentum lines hooked up directly to internal vertices are called amputated and are very practical in a number of calculations.

We can now summarize the momentum space Feynman rules. To evaluate a vacuum expectation value of the Heisenberg picture fields

$$\langle\Omega|T(\hat{\phi}(k_1)...\hat{\phi}(k_N))|\Omega\rangle$$

draw all diagrams possible with external momentum lines k_i, sum them and divide them by the sum over all diagrams with no external momentum lines, where the diagrams correspond to expressions according to the following rules:

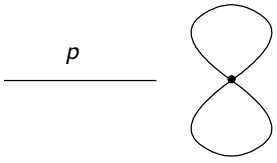
	1
	$\frac{i}{p^2-m^2+i\varepsilon}$
	$-i\lambda$
each line starting and ending at the same point	$\frac{1}{2}$
k lines between same points	$\frac{1}{k!}$
impose momentum conservation at each vertex	
impose momentum conservation on every subgraph	
integrate over all open (loop) momenta k	$\int \frac{d^4k}{(2\pi)^4}$

Example 6. Let us compute as a the trivial exapmle the matrix element of the leading order propagator  . According to our momentum space Feynman rules, we simply have

$$\langle 0|T(\hat{\phi}(k_1)\hat{\phi}(k_2))|0\rangle = (2\pi)^4\delta^4(k_1-k_2)\frac{i}{k_1^2-m^2+i\varepsilon}$$

12 Vacuum subtraction

As we have already seen, Feynman diagrams can have disconnected pieces. Let us take as one example the diagram that appears in the first order expansion of the propagator

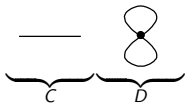


Using momentum space Feynman rules, this diagram corresponds to the expression

$$\frac{i}{p^2-m^2+i\varepsilon}(-i\lambda)\frac{1}{8}\left(\int\frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i\varepsilon}\right)^2$$

It is evident that both the symmetry factor and the expression for the "figure 8 piece" of the diagram are independent of the first "propagator" part $\frac{i}{p^2-m^2+i\varepsilon}$. This is a generic feature of disconnected diagrams which can easily be deduced from the Feynman rules: Disconnected parts of Feynman diagrams give entirely independent factors.

When looking at a generic diagram F, we can thus factor it into a connected piece C and a disconnected piece D

$F =$ 

Let us now consider all diagrams F_C^i for which the connected pieces C are identical (including the one with trivial $D=1$) and sum over them. Because of the factorization, this sum will be

$$\sum_i F_C^i = C\sum_i D^i$$

In the special case where $C=1$, this will be the sum over all "vacuum diagrams" i.e. diagrams with no external legs, which is nothing else but the denominator expression in (24). We see that

$$\langle 0|T\left(e^{-i\int_{-\infty}^{\infty}d\tau H_I(\tau(1-i\delta))}\right)|0\rangle = \langle 0|T\left(e^{-i\frac{\lambda}{4!}\int d^4z\phi^4(z)}\right)|0\rangle = \sum_i D^i$$

and thus, if we resum the numerator of (24) to group all expressions with equal connected pieces, the denominator serves to exactly cancel out the disconnected diagrams in the numerator. This cancellation is a huge simplification as it allows us to both ignore disconnected diagrams and the division of our result by the sum of vacuum diagrams. Our rule of computing vacuum expectation value of time ordered products of Heisenberg picture fields in ϕ^4 theory therefore reduces to

$$\langle\Omega|T(\hat\phi(k_1)...\hat\phi(k_N))|\Omega\rangle = \sum_i C_i$$

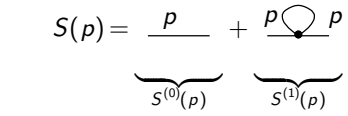
where the C_i are the connected Feynman diagrams with the correct number of external legs.

13 The propagator of ϕ^4 theory to 1 loop order

We finally have all ingredients in place to actually compute the perturbative expansion of the propagator

$$\langle\Omega|T(\hat\phi(k_1)\hat\phi(k_2))|\Omega\rangle = (2\pi)^4\delta^4(k_1-k_2)S(k_1)$$

Diagrammatically we can write

$S(p) =$ 

$+ O(\lambda^2)$

According to the Feynman rules, we have

$$S^{(0)}(p) = \frac{i}{p^2-m^2+i\varepsilon}$$

and

$$S^{(1)}(p) = -\frac{i\lambda}{2}\left(\frac{i}{p^2 - m^2 + i\varepsilon}\right)^2 \int_{\mathbb{R}^4} \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon}$$

which, for later convenience, we will write as

$$S^{(1)}(p) = -\frac{i\lambda}{2}(S^{(0)}(p))^2 \Pi$$

with

$$\Pi = \int_{\mathbb{R}^4} \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon}$$

Evaluationg $S^{(1)}(p)$ we encounter a divergence: the integral $\int_{\mathbb{R}^4} \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon}$ is not well defined. The problem arises from the integration over the large $|q|$ region. One can see this easily by rewriting (15,19)

$$\begin{aligned}
\int_{\mathbb{R}^4} \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon} &= \int_{\mathbb{R}^3} \frac{d^3q}{(2\pi)^3} \frac{1}{2\sqrt{\vec{q}^2 + m^2}} \\
&= \int_0^\infty \frac{dq}{(2\pi)^3} \frac{4\pi q^2}{2\sqrt{q^2 + m^2}} \\
&= \lim_{\Lambda \to \infty} \int_0^\Lambda \frac{dq}{4\pi^2} \frac{q^2}{\sqrt{q^2 + m^2}} \\
&= \lim_{\Lambda \to \infty} \int_0^\Lambda \frac{dq}{4\pi^2} \left(q - \frac{1}{2}\frac{m}{q} + O((m/q)^2) \right) \\
&= \lim_{\Lambda \to \infty} \frac{\Lambda^2}{8\pi^2} + \dots
\end{aligned}$$

Although this divergence appears in a physical quantity, it is not as desastrous as it appears to be on first sight. The reason for its apearance is in fact a remnant of the free KG theory that has not been eliminated in our expression: the mass m. The mass m appearing in the expressions above is still the mass of the free, $\lambda = 0$ theory. A priori, we do not know that this mass is in any way connected to with the mass of the full ϕ^4 theory. In fact, we do not even know that the ϕ^4 theory has a properly defined mass.

Let us therefore atempt to define what the mass of the ϕ^4 theory is. In the KG theory, it was simply the position of the pole in the propagator. So let us try using this definition here and see if we can find the mass m_r (the renormalized mass) of the ϕ^4 theory with it.

It should be noted that this step, as trivial as it might seem, is indeed very subtle. Generically, it is not at all clear that the propagator of a ϕ^4 theory should have a single (up to the sign) pole. It might have multiple poles, a continuous spectrum or even a superposition of both. But since we are in a perturbative framework, we will simply assume that the pole structure of the interacting theory is sufficiently close to that of the free theory so that we can carry over our notion of mass of a particle and continue on.

We thus want to identify the poles of the propagator, working to $O(\lambda)$ in perturbation theory. Since the poles are the zeroes of the inverse propagator, it is easier to look at the inverse

$$\begin{aligned}
(S^{(0)} + S^{(1)})^{-1} &= \left(S^{(0)} \left(1 - \frac{i\lambda}{2} S^{(0)} \Pi \right) \right)^{-1} \\
&= {S^{(0)}}^{-1} \left(1 - \frac{i\lambda}{2} S^{(0)} \Pi \right)^{-1} \\
&= {S^{(0)}}^{-1} \left(1 + \frac{i\lambda}{2} S^{(0)} \Pi \right) + O(\lambda^2) \\
&= {S^{(0)}}^{-1} + \frac{i\lambda}{2} \Pi + O(\lambda^2) \\
&= -i(p^2 - m^2 + i\varepsilon) + \frac{i\lambda}{2} \Pi + O(\lambda^2) \\
&= -i\left(p^2 - m^2 - \frac{\lambda}{2} \Pi + i\varepsilon \right) + O(\lambda^2)
\end{aligned}$$

where we have suppressed the argument p. Ignoring for the moment that Π is divergent and noting that it is independent of p, we can easily identify the pole by the condition

$$p^2 - m^2 - \frac{\lambda}{2}\Pi = 0$$

and thus, according to our definition, we find the renormalized mass to be

$$m_r^2 = m^2 + \frac{\lambda}{2}\Pi = m^2 + \frac{\lambda}{16\pi^2}\lim_{\Lambda\to\infty}\Lambda + \ldots \tag{26}$$

In terms of this renormalized mass, the propagator of the full theory is simply

$$S = S^{(0)} + S^{(1)} + O(\lambda^2) = \frac{i}{p^2 - m_r^2 + i\varepsilon} + O(\lambda^2)$$

and thus has the same form as the free propagator albeit with a different mass.

14 Renormalization

As we have seen in the preceding section, it is possible in principle to absorb a certain class of divergent results into a redefinition of the parameters of our theory. This procedure is not always applicable though and we have not yet defined a theory that is entirely finite. In this section we will outline this procedure, which is called renormalization generically and see where it is applicable.

14.1 Mass and wave function renormalization

Let us start with a theory that is finite for large momenta (UV finite). Such a theory is easily obtained e.g. by modifying the Feynman rules such that a momentum cutoff at a scale Λ is implied in all of integrals. The Lagrangean of the theory is

$$\mathscr{L}_0 = \frac{1}{2}(\partial_\mu\phi_0\partial^\mu\phi_0 - m_0^2\phi_0^2) - \frac{\lambda_0}{4!}\phi_0^4$$

where $\phi_0(\Lambda)$, $m_0(\Lambda)$ and $\lambda_0(\Lambda)$ are called bare quantities that have a yet unspecified dependence on the cutoff Λ. With this bare Lagrangean, we construct the usual perturbative expansion in the parameter λ_0. If we look at the resulting propagator of momentum p to a certain order in perturbation theory, we will obtain corrections that are generically functions of both p and the cutoff Λ. In addition, the symmetry under $p \to -p$ will be preserved, so that the corrections are generic functions of p^2 and Λ. Separating out the part independent of p^2, we can write the propagator as

$$\frac{i}{A(\Lambda, p^2)p^2 - m_0^2(\Lambda) - B(\Lambda)}$$

In our example above, we had $A(\Lambda, p^2) = 1 + O(\lambda_0^2)$ and $B(\Lambda) = \lambda_0\Pi(\Lambda) + O(\lambda_0^2)$ where

$$\Pi(\Lambda) = \int_0^\Lambda \frac{dq}{4\pi^2}\frac{q}{\sqrt{1 + \left(\frac{m}{q}\right)^2}} = \frac{\Lambda^2}{8\pi^2} + \ldots$$

Whenever $A = 1$, we can define

$$m_r^2 = m_0^2(\Lambda) + B(\Lambda) \tag{27}$$

and obtain a propagator of the form

$$\frac{i}{p^2 - m_r^2}$$

Identifying m_r with the actually observable, physical mass of the particle we can use (27) as definition equation for the bare mass $m_0^2(\Lambda)$. The propagator is now expressed in terms of the physical m_r and independent of the cutoff Λ. We can thus take $\Lambda \to \infty$ without causing any physical quantity to diverge. What diverges is the bare coupling $m_0(\Lambda)$, but this does not cause any problems because no physical quantity depends on it.

When we have an $A(\Lambda)$ which does not depend on p^2, we could make a similar argument. In such a case our propagator expression will read

$$\frac{i}{A(\Lambda)p^2 - m_0^2(\Lambda) - B(\Lambda)} = \frac{iA^{-1}(\Lambda)}{p^2 - m_r^2}$$

with

$$m_r^2 = \frac{m_0^2(\Lambda) + B(\Lambda)}{A(\Lambda)}$$

Remembering that the propagator is quadratic in the fields ϕ_0

$$\phi_0\phi_0 \propto \frac{iA^{-1}(\Lambda)}{p^2 - m_r^2}$$

we can define a renormalized field $\phi_r = \sqrt{A(\Lambda)}\,\phi_0$ for which the propagator again takes on the free field form

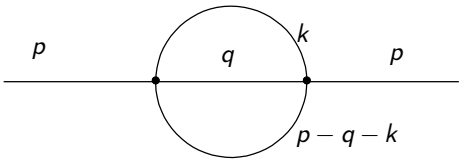
$$\phi_r\phi_r \propto \frac{i}{p^2 - m_r^2}$$

Generically, if we can cast a propagator into the form

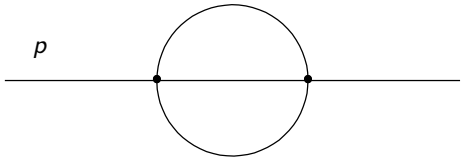
$$\phi_0\phi_0 \propto \frac{iZ}{p^2 - m_r^2}$$

we have reobtained the free field behaviour in the renormalized field $\phi_r = \phi_0/\sqrt{Z}$.This procedure is called wave function renormalization and Z is the pertaining renormalization constant.

One might be tempted by the preceeding analysis to conclude that no Feynman diagrams that correct the leading order propagator do have any physical consequences. This is not the case and it can be seen quite easily. The diagram we have investigated thus far - which is called the "tadpole" diagram as it has has reminded some people of a tadpole - has a loop integral that does not depend on the external momenumm. Such integrals thus generically contribute to $B(\Lambda)$ only. In ϕ^4 theory this is the only diagram at $O(\lambda_0)$, but at $O(\lambda_0^2)$ we have e.g.



where one of the intermediate propagators has a momentum dependent on p. We will not delve into the technicalities of computing these diagrams here as they are quite lengthy. For the diagram above we simply note that it evaluates to

$S^{(2p)} =$  $= -S^{(0)2} A\lambda_0^2 p^2 \left(\ln\left(\frac{\Lambda^2}{p^2}\right) + c \right)$

with some constants A and c. If we start from the free propagator

$$S^{(0)} = \frac{i}{p^2 - m_0^2 + i\varepsilon}$$

we can add to it the above diagram (which is more convincingly called the sunset diagram) and look at the inverse propagator

$$\begin{aligned}
(S^{(0)}+S^{(2p)})^{-1} &= \left(S^{(0)}\left(1-\lambda_0^2 S^{(0)} A p^2\left(\ln\left(\frac{\Lambda^2}{p^2}\right)+c\right)\right)\right)^{-1} \\
&= S^{(0)^{-1}}\left(1-\lambda_0^2 S^{(0)} A p^2\left(\ln\left(\frac{\Lambda^2}{p^2}\right)+c\right)\right)^{-1} \\
&= S^{(0)^{-1}}\left(1+\lambda_0^2 S^{(0)} A p^2\left(\ln\left(\frac{\Lambda^2}{p^2}\right)+c\right)\right)+O(\lambda_0^4) \\
&= S^{(0)^{-1}}+\lambda_0^2 A p^2\left(\ln\left(\frac{\Lambda^2}{p^2}\right)+c\right)+O(\lambda_0^4) \\
&= -i(p^2-m^2+i\varepsilon)+\lambda_0^2 A p^2\left(\ln\left(\frac{\Lambda^2}{p^2}\right)+c\right)+O(\lambda_0^4) \\
&= -i\left(p^2\left(1+\lambda_0^2 A\left(\ln\left(\frac{\Lambda^2}{p^2}\right)+c\right)\right)-m^2+i\varepsilon\right)+O(\lambda_0^4)
\end{aligned}$$

This expression clearly does not have the form of a free propagator anymore. The additional term $\propto p^2 \ln(\Lambda^2/p^2)$ is new and arises only in the interacting theory. It can not simply be absorbed by the wave function renormalization, but we can nonetheless define physical quantities that are free of divergencies. The key insight is that we do not necessarily need to absorb the entirety of the new terms into our redefinitions of m and ϕ. It is sufficient to absorb the divergent part. Once we have absorbed the divergent part, the remaining finite part will actually give us nontrivial corrections in the interacting theory.

Before we proceed to actually perform the renormalization for a generic $A(\Lambda, p^2)$, we need to expose one subtelty of this general, p^2 dependent case that was not obvious in the p^2 independent cases discussed above. In the p^2 independent scenario, we had a clear guide as to what the renormalized propagator should look like: It should simply be in the form of the free propagator. This was a logical choice, but it nonetheless was a choice. Imagine that for the tadpole diagram, where we chose the renormalized mass to be $m_r = m_0^2(\Lambda) + B(\Lambda)$ with the corresponding propagator

$$\frac{i}{p^2-m_r^2}$$

we would instead have decided to use $m_r'^2 = m_0^2(\Lambda) + B(\Lambda) - 1\text{eV}^2$. With this choice, the propagaror would have been

$$\frac{i}{p^2-m_r'^2-1\text{eV}^2}$$

which might not be pretty but still is a perfectly fine definition. The renormalized mass m_r' as well as the propagator are independent of the cutoff Λ and all physical quantities are still finite. Of course the propagator would not have a pole at $p^2 = m_r'^2$ but instead at $p^2 = m_r'^2 + 1\text{eV}^2$, i.e. m_r' is not what we conventionally understand as the mass of a particle, but no matrix element and thus no physical prediction would be affected by this.

For the generic, p^2 dependent case $A(\Lambda, p^2)$ we have seen that there is no obvious choice anymore as the form of the propagator is different from the free one. Any such choice is referred to as a renormalization scheme and intermediate results will depend on this scheme although of course physical quantities will ultimately be unaffected. We will use here as an example the on-shell renormalization scheme. Its defining property, the renormalization condition, is that when a particle is "on the mass shell", i.e. when $p^2 = m_r^2$ then the propagator should have the form of the free propagator. This condition can be expressed as

$$\frac{i}{A(\Lambda,p^2)p^2-m_0^2(\Lambda)-B(\Lambda)} \xrightarrow{p^2\to m_r^2} \frac{iZ}{p^2-m_r^2}$$

which implies

$$m_r^2 = \frac{m_0^2(\Lambda)+B(\Lambda)}{A(\Lambda,m_r^2)}$$

for the correct location of the pole. Expanding the denominator of the r.h.s. around $p^2 = m_r^2$ we then obtain

$$\begin{aligned} A(\Lambda, p^2)p^2 - A(\Lambda, m_r^2)m_r^2 &= (A(\Lambda, m_r^2) + A'(\Lambda, m_r^2)(p^2 - m_r^2) + O((p^2 - m_r^2)^2))p^2 - A(\Lambda, m_r^2)m_r^2 \\ &= A(\Lambda, m_r^2)(p^2 - m_r^2) + A'(\Lambda, m_r^2)(p^2 - m_r^2)p^2 + p^2 O((p^2 - m_r^2)^2) \\ &= (p^2 - m_r^2)(A(\Lambda, m_r^2) + A'(\Lambda, m_r^2)m_r^2) + p^2 O((p^2 - m_r^2)^2) \\ &\stackrel{!}{=} (p^2 - m_r^2)/Z \end{aligned}$$

which implies

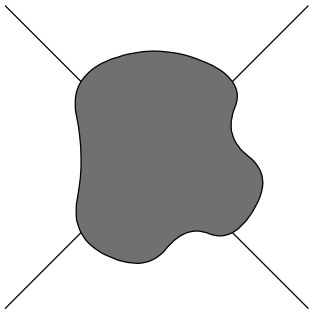
$$Z = \frac{1}{A(\Lambda, m_r^2) + A'(\Lambda, m_r^2)m_r^2}$$

where $A'(\Lambda, m_r^2) = \frac{\partial}{\partial p^2}\Big|_{m_r^2} A(\Lambda, p^2)$.

14.2 Renormalizability

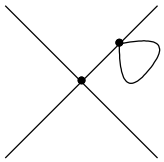
We now know in principle how to absorb divergent integrals in the perturbative expansion of the propagator into the bare mass m_0 and the wave function renormalization constant Z. But how can we check whether more parameters of the theory need to be renormalized?

It seems that renormalization is necessary whenever a Feynman diagram contributing to a certain n-point function shows a divergence. So any divergence in a diagram of the form



should contribute to a renormalization of the coupling λ_0. Even worse, it seems that if there is a divergence in any diagram with e.g. 24 external legs, then the coupling of the term ϕ^{24} needs to be renormalized even if such a coupling is not present in the bare Lagrangean at all. Worse still, an infinite number of divergence absorbtions might be necessary imposing in principle an infinite number of renormalization conditions on our theory. As each of these conditions ultimately requires experimental input to link the theory to experiment, predictability of the theory might be lost.

Let us examine first whether we can draw a Feynman diagram that leads to a divergent integral for any number of external legs. This is in fact easy to do: We already know that the tadpole diagram is divergent, so we can simply replace a leg of an arbitrary diagram with a tadpole and obtain a divergence, e.g.



Because of the tadpole subdiagram, this diagram certainly contains a divergent integral. However, we have already learned how to deal with this particular type of divergence. We know that we can absorb it by a mass renormalization and therefore do not really need to treat it separately again. All we need to do is to note that if we replace the bare propagator _____ in any part of the Feynman diagram with the renormalized propagator, the tadpole and all other diagrams contributing to the renormalized propagator will already be included and the renormalized propagator is finite. We generically denote a renormalized propagator by

With this in mind, let us look at whether there are any nontrivial UV divergencies left in other matrix elements than just the propagator. At first this task seems hopeless, but in fact it is relatively straightforward. All we need to do is consider that the UV divergences arise from the momentum integrals represented by the loops in a Feynman diagram. Every such integral goes over d^4p, so without anything damping its UV behaviour it will diverge as Λ^4 where again Λ is our generic cutoff scale. On the other hand, every internal propagator goes for large momenta as p^{-2} so it is UV convergent as Λ^{-2}. If we now take an arbitrary Feynman diagram with L loops and P propagators, we get a naive dimensionality of the expression in terms of the cutoff Λ^D with

$$D = 4L - 2P$$

The dimensionality D is called the superficial degree of divergence of a diagram. If $D > 0$ it means the diagram is naively expected to be power divergent as Λ^D, if $D < 0$, it is expected to be convergent as Λ^D and for $D = 0$ it is expected to be logarithmically divergent as $\ln(\Lambda)$. There are two cases where this naive estimate fails:

1. When the diagram consists of two or more independent integrals, the power counting needs to be done for each one of them separately. Denoting the superficial degrees of divergence for each independent subdiagram as D_i, it might happen that one $D_i > 0$ although the sum $\sum_i D_i < 0$. A good example of this is the above mentioned tadpole on a leg of the diagram. In this case, the renormalization on the subdiagram has to be performed first. Once this is properly done however, the resulting diagram is even more convergent and D is thus a good conservative estimate for the worst divergence possible.

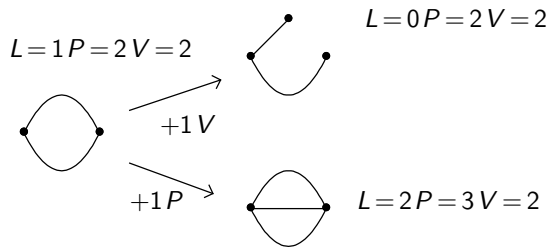
2. Symmetries may reduce the actual degree of divergnce. Evidently, this makes the estimate more conservative, too.

Thus we see, that the superficial degree of divergence is actually usefull as an upper bound on the worst possible divergence for classes of diagrams and we will try to bring it into a more useful form. For this we note that every propagator attaches to two vertices and every external line attaches to one vertex, whereas every vertex has four legs attached. Denoting the number of vertices with V and the number of external lines with E, we thus have the relation

$$4V = 2P + E$$

As a side note, this relation changes into $NV = 2P + E$ for a generic ϕ^N theory.

Finally, we observe that when there is one internal propagator there need to be at least 2 vertices so that there is no loop. If there are more than two vertices the diagram is disconnected, which we disregard. So one internal propagator and one vertex means one loop while one internal propagator and two vertices means no loops. Adding one internal propagator without adding a vertex always produces one more loop while adding one vertex without adding an internal propagator will destroy an existing loop or produce a disconnected diagram, which again we disregard.



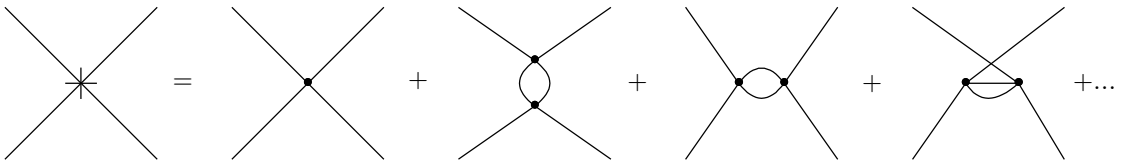
In summary, we therefore have the number of loops equal to the number of internal propagators minus the number of vertices plus one or

$$L = P - V + 1$$

and we obtain as our final result for the superficial degree of divergence

$$D = 4L - 2P = 2P - 4V + 4 = 4 - E$$

This result is extremely useful as it tells us that the only divergent Feynman diagrams are those with two or four external legs (again disregarding the vacuum bubbles with no external legs). Thus in addition to renormalizing m and ϕ we need to renormalize the coupling λ but nothing more. Without going into any detail, we can diagrammatically represent the renormalized ϕ^4 vertex by



When doing a systematic perturbative expansion for E external legs to $O(\lambda^n)$ with Feynman diagrams, one can therefore adopt the following strategy: First compute the renormalized propagator and the renormalized coupling to $O(\lambda^{n-1})$. Then draw all connected Feynman diagrams with E external legs, replacing each propagator and 4-vertex subdiagram by the properly renormalized one. Although there are many technical subtelties, this procedure will in principle yield a finite result for all correlation functions.

As a final remark, let us compute the superficial degree of divergence for a generic ϕ^N theory. As we have seen above, we obtain

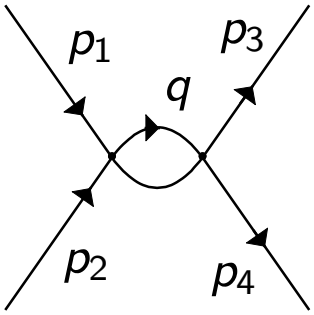
$$D = 4L - 2P = 2P - 4V + 4 = (N-4)V + 4 - E$$

so for any $N > 4$ the superficial degree of divergence increases with the number of vertives. Thus any such theory needs renormalization of an infinite number of couplings. Such theories are called (perturbatively) non-renormalizable and it is unclear whether they make any sense as candidates of fundamentally correct theories.

15 One loop renormalization of ϕ^4 theory

15.1 Evaluating a logarithmically divergent diagram

While it is beyond this lecture to introduce the standard techniques to compute loop diagrams fully, let us look at one example, namely the second order vertex correction diagram



Denoting the sum of incoming momenta by $p = p_1 + p_2$, the amputated diagram translates to the integral

$$I = \frac{1}{2}(-i\lambda)^2 \int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2} \frac{i}{(q-p)^2 - m^2}$$

First we rewrite this integral using the Feynman identity

$$\frac{1}{A}\frac{1}{B} = \int_0^1 dx \, \frac{1}{(xA + (1-x)B)^2}$$

into

$$I = \frac{\lambda^2}{2}\int_0^1 dx \int \frac{d^4q}{(2\pi)^4} \frac{1}{(x(q^2-m^2) + (1-x)((q-p)^2-m^2))^2}$$

which, after a variable substitution

$$r = (q - (1-x)p) \qquad D = m^2 - x(1-x)p^2$$

that completes the squares simplifies into

$$I = \frac{\lambda^2}{2}\int_0^1 dx \int \frac{d^4r}{(2\pi)^4} \frac{1}{(r^2-D)^2}$$

The integral above contains an r_0 integration along the real axis. Regarding this as an integral in the complex plane and remembering the position of the poles of r_0, we are allowed to rotate the integration contour so that it runs along the imaginary axis, as long as the integrand falls off sufficiently fast (i.e. faster than $1/r_0$) at infinity such as not to contribute to the total integral.

Performing this so-called Wick rotation amounts to making the substitution $r_0 \to \mathrm{i}r_0$. This would however leave us with a negative metric, so we would have the slightly paradoxical notation

$$r^2 = r_\mu r^\mu = -r^\mu r^\mu$$

We circumvent this by additionally multiplying all components with $-i$, so in total we replace the spatial components $r_i \to -ir_i$ to have a positively defined metric. This results in

$$I = i\frac{\lambda^2}{2}\int_0^1 dx \int \frac{dr}{(2\pi)^4} \frac{r^3}{(r^2+D)^2} \int d\Omega_4$$

where we have explicitly split the radial and angular part of the integration. As one can see, this integral is logarithmically UV divergent, we therefore need to renormalize it. The currently most commonly used procedure for this is dimensional regularization, where the four-dinemsional integral is replaced by an integral in $4-\varepsilon$ dimensions, where we ultimately send $\varepsilon \to 0$. The resulting integral

$$I_\varepsilon = i\frac{\lambda^2}{2}\int_0^1 dx \int \frac{dr}{(2\pi)^{4-\varepsilon}} \frac{r^{3-\varepsilon}}{(r^2+D)^2} \int d\Omega_{4-\varepsilon}$$

is then convergent for any $\varepsilon > 0$. Noting that there is no angular dependence, we can perform the integral over the surface of the $4-\varepsilon$ sphere explicitly

$$\int d\Omega_{4-\varepsilon} = \frac{2\pi^{2-\frac{\varepsilon}{2}}}{\Gamma\left(2-\frac{\varepsilon}{2}\right)} = \pi^2(2 + \varepsilon(1 - \gamma - \ln(\pi))) + O(\varepsilon^2)$$

Similarly, the radial integral can be performed

$$\int dr \frac{r^{3-\varepsilon}}{(r^2+D)^2} = \frac{\pi}{4}(2-\varepsilon)D^{-\frac{\varepsilon}{2}} \frac{1}{\sin\left(\frac{\varepsilon\pi}{2}\right)} = \frac{1}{\varepsilon} - \frac{1}{2}(1 + \ln(D)) + O(\varepsilon)$$

and for the prefactor we obtain

$$(2\pi)^{\varepsilon-4} = (2\pi)^{-4}(1 + \varepsilon\ln(2\pi)) + O(\varepsilon^2)$$

Notice that D is a dimensionful quantity and we are taking the logarithm of it. This slightly awkward expression is clarified by noting that, due to the change in the dimensionality the coupling constant λ is no more dimensionless. More precisely, the requirement of a dimensionless action implies that

$$[x^{4-\varepsilon}\partial^2\phi^2]=1 \quad \Rightarrow \quad [\phi]=[x]^{\frac{\varepsilon}{2}-1}=[m]^{1-\frac{\varepsilon}{2}}$$

and thus

$$[x^{4-\varepsilon}\lambda\phi^4]=1 \quad \Rightarrow \quad [\lambda]=[x]^{\varepsilon-4}[\phi]^{-4}=[x]^{\varepsilon-4-4(\frac{\varepsilon}{2}-1)}=[x]^{-\varepsilon}=[m]^{\varepsilon}$$

so that the coupling λ has positive mass dimension ε. We therefore introduce a scale parameter μ with mass dimension 1 (i.e. $[\mu]=[m]$) and define a dimensionless coupling $\bar{\lambda}$ via

$$\lambda=\bar{\lambda}\tilde{\mu}^{\varepsilon}=\bar{\lambda}(1+\varepsilon\ln(\tilde{\mu}))+O(\varepsilon^2)$$

Plugging all these explicit expressions in, we arrive at

$$\begin{aligned}
I_\varepsilon &= i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{2}\pi^2\int_0^1 dx(2+\varepsilon(1-\gamma-\ln(\pi)))\left(\frac{1}{\varepsilon}-\frac{1}{2}(1+\ln(D))+O(\varepsilon)\right)\times \\
&\times\ (2\pi)^{-4}(1+\varepsilon\ln(2\pi))(1+\varepsilon\ln(\tilde{\mu}))+O(\varepsilon^2) \\
&= i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{16\pi^2}\frac{1}{\varepsilon}\int_0^1 dx\left(1+\frac{\varepsilon}{2}(1-\gamma-\ln(\pi))\right)\left(1-\frac{\varepsilon}{2}(1+\ln(D))+O(\varepsilon)\right)\times \\
&\times\ (1+\varepsilon\ln(2\pi))(1+\varepsilon\ln(\tilde{\mu}))+O(\varepsilon) \\
&= i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{16\pi^2}\int_0^1 dx\left(\frac{1}{\varepsilon}+\frac{1}{2}(1-\gamma-\ln(\pi)-1-\ln(D)+2\ln(2\pi)+2\ln(\tilde{\mu}))\right)+O(\varepsilon) \\
&= i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{16\pi^2}\int_0^1 dx\left(\frac{1}{\varepsilon}+\frac{1}{2}\ln(4\pi\tilde{\mu}^2/e^{\gamma}D)\right)+O(\varepsilon) \\
&= i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{16\pi^2}\int_0^1 dx\left(\frac{1}{\varepsilon}+\frac{1}{2}\ln(4\pi/e^{\gamma})-\frac{1}{2}\ln(D/\tilde{\mu}^2)\right)+O(\varepsilon)
\end{aligned}$$

where we kept one factor $\tilde{\mu}^{\varepsilon}$ to account for the explicit dimensionality of the integral. Performing the integration over the Feynman variable x we obtain

$$I_\varepsilon=i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{16\pi^2}\left(\frac{1}{\varepsilon}+\frac{1}{2}\ln(4\pi/e^{\gamma})-\frac{1}{2}F(p^2/\tilde{\mu}^2,m^2/\tilde{\mu}^2)\right)+O(\varepsilon)$$

where the function

$$F(a,b)=\int_0^1 dx\ln(b-x(1-x)a)$$

evaluates to

$$F(a,b)=-y\ln\left(\frac{y-1}{y+1}\right)+\ln(b)-2 \qquad y=\sqrt{1-\frac{4b}{a}}$$

Note that for finite b $F(a,b)$ has a well defined limit

$$F(a,b)\xrightarrow{a\to 0}\ln(b)$$

Puting everything together we eventually obtain

$$I_\varepsilon=i\frac{\bar{\lambda}^2\tilde{\mu}^{\varepsilon}}{16\pi^2}\left(\frac{1}{\varepsilon}+\frac{1}{2}\left(\ln(4\pi m^2/\tilde{\mu}^2)-\gamma-2-\frac{\sqrt{p^2-4m^2}}{p}\ln\left(\frac{\sqrt{p^2-4m^2}-p}{\sqrt{p^2-4m^2}+p}\right)\right)\right)+O(\varepsilon)$$

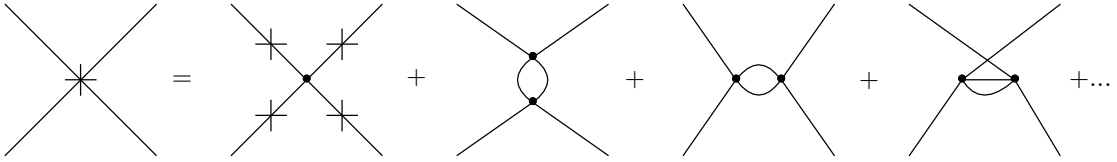
which, in the limit $p^2 \to 0$ simplifies to

$$I_\varepsilon \xrightarrow{p^2\to 0} i\frac{\bar{\lambda}^2\tilde{\mu}^\varepsilon}{16\pi^2}\left(\frac{1}{\varepsilon}+\frac{1}{2}(\ln(4\pi m^2/\tilde{\mu}^2)-\gamma)\right)+O(\varepsilon)$$

Note that for $p<2m$ the integral is real whereas for $p>2m$ it has an imaginary component. This is a mathematical reflection of the physical fact that the two particles can be real only for $p>2m$. A similar situation occurs in quantum mechanics where the behaviour of the wave function s substantially different in the classically allowed and the classically forbidden regions.

15.2 Vertex renormalization

In order to proceed further, we write down the relevant diagrams with four external legs up to second order:



Note the renormalized propagator in the first order diagram which is present according to our prescription of treating divergent subdiagrams seperately. The three nontrivial second order diagrams are all of the same type which we have computed above, the only difference being the value of p. Labeling p_1 and p_2 as incoming momenta, we define $s=(p_1+p_2)^2$, $t=(p_1-p_3)^2$ and $u=(p_1-p_4)^2$, the so called Mandelstam variables. The second order contribution to the diagram with four external vertices is therefore given by

$$I_\varepsilon(p^2=s)+I_\varepsilon(p^2=t)+I_\varepsilon(p^2=u)$$

and the total contribution up to second order is therefore

$$\begin{aligned}
\Lambda_\varepsilon^{(2)} &= i\lambda + I_\varepsilon(p^2=s)+I_\varepsilon(p^2=u)+I_\varepsilon(p^2=u) \\
&= i\bar{\lambda}\tilde{\mu}^\varepsilon\left(1+\frac{\bar{\lambda}}{16\pi^2}\sum_{p^2=s,t,u}\left(\frac{1}{\varepsilon}+\frac{1}{2}\left(\ln(4\pi m^2/\tilde{\mu}^2)-\gamma-2-y_{p^2}\ln\left(\frac{y_{p^2}-1}{y_{p^2}+1}\right)\right)\right)\right) \qquad (28)\\
&= i\lambda\left(1+\frac{\bar{\lambda}}{16\pi^2}\left(\frac{3}{\varepsilon}+\frac{3}{2}(\ln(4\pi m^2/\tilde{\mu}^2)-\gamma-2)-(L(s)+L(t)+L(u))\right)\right)
\end{aligned}$$

with

$$L(q)=\frac{\sqrt{q-4m^2}}{2\sqrt{q}}\ln\left(\frac{\sqrt{q-4m^2}-\sqrt{q}}{\sqrt{q-4m^2}+\sqrt{q}}\right)$$

which has the property

$$L(4m^2)=0 \qquad L(0)=-1$$

Let us now perform the actual renormalization. We see that $\Lambda_\varepsilon^{(2)}$ contains an explicit divergence, which can be absorbed into the coupling constant λ. How this absorbtion is done however is not unique. We can demand e.g. that at a certain scale $\tilde{\mu}$ and for certain s, t and u the renormalized coupling has some specific value. To keep things simple, let us demand that when all external lines are on shell and the spatial momenta vanish ($\vec{p}_i=0$, $s=4m^2$,) the total contribution $\Lambda^{(2)}$ is just the renormalized coupling λ_r.

To formalize this procedure, we say that we start from the bare Lagrangean

$$\mathscr{L}_0=\frac{1}{2}(\partial_\mu\phi_0\partial^\mu\phi_0-m_0^2\phi_0^2)-\frac{\lambda_0}{4!}\phi_0^4$$

to which we add counterterms, so that the full Lagrangean reads

$$\mathscr{L}=\frac{1}{2}(\partial_\mu\phi_r\partial^\mu\phi_r - m_r^2\phi_r^2) - \frac{\lambda_r}{4!}\phi_r^4$$

What we sloppily labeled "adding counterterms" can in fact be formalized, too. Crucially, we wanted these counterterms to cancel the divergencies that appear in bare quantities. We thus write the bare Lagrangean in terms of the renormalized quantities

$$\mathscr{L}_0=\frac{1}{2}(Z_\phi\partial_\mu\phi_r\partial^\mu\phi_r - Z_m m_r^2\phi_r^2) - Z_\lambda\frac{\lambda_r}{4!}\phi_r^4$$

where we introduced the renormalization constants Z_ϕ, Z_m and Z_λ. Comparing coefficients we thus have

$$\begin{aligned}\phi_0^2 &= Z_\phi\phi_r^2\\ m_0^2 &= Z_\phi^{-1}Z_m m_r^2\\ \lambda_0 &= Z_\phi^{-2}Z_\lambda m_r\lambda_r\end{aligned}$$

Quite generically, we know that without loop diagrams (at tree level) bare and renormalized quantities are equal (in fact, one could choose a constant factor between bare and renormalized quantities even at tree level, but a very convincing choice for these constants is just 1). Adding loops can only be achieved by adding vertices though, so again quite generically, one can expand the renormalization constants in powers of λ_r

$$Z_x=1+\sum_{n=1}^{\infty}\lambda_r^n Z_x^{(n)}$$

We now perform this expansion up to first order in λ_r. From our discussion of the tadpole diagram, we already know that $Z_\phi^{(1)}=0$ so $Z_\phi=1+O(\lambda_r^2)$. From (28) and our convention of the renormalized coupling we can see that

$$\lambda_r=-i\Lambda_\varepsilon^{(2)}\big|_{s=4m^2,t=u=0}=\lambda_0\left(1+\frac{\bar{\lambda}_0}{16\pi^2}\left(\frac{3}{\varepsilon}+\frac{3}{2}(\ln(4\pi m^2/\tilde{\mu}^2)-\gamma)-1\right)\right)$$

and we read off the divergent renormalization constant

$$Z_\lambda^{(1)}=\frac{\tilde{\mu}^{-\varepsilon}}{16\pi^2}\left(\frac{3}{\varepsilon}+\frac{3}{2}(\ln(4\pi m^2/\tilde{\mu}^2)-\gamma)-1\right)$$

Note that this renormalization constant also depends on the energy scale $\tilde{\mu}$ that was introduced in the dimensional regularization procedure.

15.3 Mass and wave function renormalization

Although we already looked at the leading order correction to the propagator in section 13, we have not regularized it dimensionally, but rather with a simple momentum cutoff. For a consistent treatment let us now do the integral in dimensional regularisation.

Written in terms of the renormalized field, the tree level propagator is

$$S^{(0)}=\frac{iZ_\phi}{p^2-m_r^2}$$

and we have seen in section 13 that in terms of the bare field the propagator is

$$S=\frac{i}{p^2-m_0^2-\frac{\lambda_0}{2}\Pi}+O(\lambda_0^2)=\frac{i}{p^2-m_0^2-\frac{\lambda_r}{2}\Pi}+O(\lambda_r^2)$$

where

$$\Pi = \int_{\mathbb{R}^4} \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m_0^2 + i\varepsilon}$$

We can Wick-rotate this to

$$\Pi = \int_0^\infty \frac{dq}{(2\pi)^4} \frac{q^3}{q^2 + m_0^2} \int d\Omega_4$$

which is of course quadratically divergent. Going to $4-\varepsilon$ dimensions, we have

$$\Pi = \int_0^\infty \frac{dq}{(2\pi)^{4-\varepsilon}} \frac{q^{3-\varepsilon}}{q^2 + m_0^2} \int d\Omega_{4-\varepsilon}$$

The angualr integration again results in

$$\int d\Omega_{4-\varepsilon} = \frac{2\pi^{2-\frac{\varepsilon}{2}}}{\Gamma\left(2-\frac{\varepsilon}{2}\right)} = \pi^2(2 + \varepsilon(1 - \gamma - \ln(\pi))) + O(\varepsilon^2)$$

while the radial one gives

$$\int_0^\infty dq \frac{q^{3-\varepsilon}}{q^2 + m^2} = \frac{\pi}{2\sin\left(\frac{\varepsilon\pi}{2}\right)} m_0^{2+\varepsilon} \xrightarrow{\varepsilon \to 0} m^2\left(\frac{1}{\varepsilon} + \ln(m_0)\right) + O(\varepsilon)$$

So in total we get

$$\begin{aligned}
\Pi &= m_0^2 \frac{\pi^2}{16\pi^4}(1 + \varepsilon\ln(2\pi))(2 + \varepsilon(1 - \gamma - \ln(\pi)))\left(\frac{1}{\varepsilon} + \ln(m_0)\right) \\
&= \frac{m_0^2}{16\pi^2}\left(\frac{2}{\varepsilon} + 2\ln(2\pi) + 1 - \gamma - \ln(\pi) + 2\ln(m_0)\right) \\
&= \frac{m_0^2}{16\pi^2}\left(\frac{2}{\varepsilon} + \ln(4\pi m_0^2) + 1 - \gamma\right)
\end{aligned}$$

which again contains a slightly awkward dimensionful logarithm that we will deal with shortly.

The renormalization condition we chose was that near the pole of the propagator the full propagator should look like a free particle propagator, so we demand

$$S|_{p^2 \sim m_r^2} = S^{(0)}|_{p^2 \sim m_r}$$

which can be easily satisfied by demanding

$$\frac{iZ_\phi}{p^2 - m_r^2} = \frac{i}{p^2 - m_0^2 - \frac{\bar{\lambda}_r}{2}\Pi} + O(\lambda_r^2)$$

Notice that here we have to introduce $\bar{\lambda}_r = \tilde{\mu}^{-\varepsilon}\lambda_r$ to have a consistent mass dimension 2 in the denominator. From the renormalization condition it therefore follows that

$$\begin{aligned}
m_r^2 &= m_0^2 + \frac{\bar{\lambda}_r}{2}\frac{m_0^2}{16\pi^2}\left(\frac{2}{\varepsilon} + \ln(4\pi m_0^2) + 1 - \gamma\right) \\
&= m_0^2 + \frac{\lambda_r}{2}\frac{m_0^2}{16\pi^2}\left(\frac{2}{\varepsilon} + \ln(4\pi m_0^2) + 1 - \gamma\right)(1 - \varepsilon\ln(\tilde{\mu})) \\
&= m_0^2\left(1 + \frac{\lambda_r}{32\pi^2}\left(\frac{2}{\varepsilon} + \ln(4\pi m_0^2/\tilde{\mu}^2) + 1 - \gamma\right)\right)
\end{aligned}$$

15.4 The running coupling

From the vertex correction (28) together with the explicit expressions of the mass, wave function and coupling constant renormalization, we can conclude that the vertex, up to second order in λ_r, can be written as

$$\Lambda^{(2)} = i\lambda_r\left(1 - \frac{\lambda_r}{16\pi^2}(L(s) + L(t) + L(u) + 2)\right) \tag{29}$$

Note that this is a finite quantity even in the $\varepsilon \to 0$ limit, so we have taken this limit. Interestingly, this relation implies that the strenght of the ϕ^4 interaction depends on the external momenta through the Mandelstam variables s, t, and u. Since we have chosen a renormalization condition that refers to physical variables only, the interaction does not depend on any unphysical parameters. This is not generically so. We could e.g. have chosen a renormalization procedure that just eliminates the divergent $3/\varepsilon$ term in (28). This renormalization scheme is known as the minimal subtraction scheme and since it does not absorb the dependence on the unphysical parameter $\tilde{\mu}$ into Z_λ, the vertex is

$$\Lambda^{(2)} = i\lambda_r\left(1 - \frac{\lambda_r}{16\pi^2}\left(\frac{3}{2}\left(\gamma - \ln\frac{4\pi m^2}{\tilde{\mu}^2}\right) + L(s) + L(t) + L(u) + 2\right)\right)$$

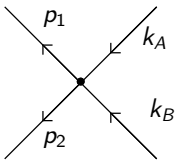
In this scheme, the relation between λ_r and a physical amplitude is not obvious and will depend on the choice of $\tilde{\mu}$, which is called the renormalization scale. Of course physical quantities are independent of the choice of renormalization scheme and thus λ_r has to depend on $\tilde{\mu}$. This phenomenon is labeled the "running coupling" that is a typical phenomenon of quantum field theories.

16 Physical predictions

We are finally in a position to relate amplitudes we compute in ϕ^4 theory to experimentally measured quantities. Let us recall that with Feynman rules we are actually computing quantum mechanical matrix elements, which are probability amplitudes. We need to take their absolute square to find actual transition probabilities between states.

16.1 Computing transition rates

Let us now try to compute a quantity that gives us an experimental prediction for how particles interact with each other in our ϕ^4 theory. Before we actually define this quantity, let us try to see what we can compute already that could ultimately lead to such a prediction. The obvious candidate is a 4-point function, diagramatically expressed, to leading order, as



with two ingoing momenta k_A and k_B and two outgoing momenta p_1 and p_2. This amputated diagram represents a first order contribution to a probability amplitude that the incoming states will go over into the outgoing states (elastic scattering). Our momentum space Feynman rules tell us that this diagram corresponds to an expression

$$-(2\pi)^4\delta^4(p_1 + p_2 - k_A - k_B)i\lambda$$

As we have seen in the previous chapter, going to next order in perturbation theory replaces $i\lambda$ with $\Lambda^{(2)}$ which in the on-shell renormalization scheme is given by (29). We would now like to interpret this as a probability amplitude. This interpretation however has the obvious problem that the probability, which is the square of the probability amplitude, can not readily be computed for an expression containing a Dirac δ. So what can we do?

The underlying reason for this problem is not too difficult to spot: We are trying to construct transition probabilities between states with definite 4-momentum. Strictly speaking, such states can not be normalized in infinite volume. One has to either construct wave packets with them or to go to a finite volume for intermediate steps and only later take the infinite volume limit. We choose the later route, since it is technically easier.

Let us first consider the 1D case. Suppose we have functions $f(x)$ on a finite interval $x \in [-L/2, L/2)$. We define the Fourier transform and its inverse are

$$\begin{aligned} \hat{f}(p) &:= \frac{1}{L}\int_{-L/2}^{L/2} dx f(x) e^{ipx} \\ f(x) &= \sum_p \hat{f}(p) e^{-ipx} \end{aligned}$$

where the momenta p can only take discrete values $p = 2\pi n / L$ and thus the finite difference between momenta is

$$\Delta p = \frac{2\pi}{L} \tag{30}$$

In this setup we have

$$\delta_{p,q} = \frac{1}{L}\int_{-L/2}^{L/2} dx e^{i(p-q)x}$$

as a definition of a Kronecker-δ. In the limit $L \to \infty$ we have already seen that we obtain the Dirac-δ as

$$\delta(p-q) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx e^{i(p-q)x}$$

Comparing these two expressions we obtain

$$2\pi\delta(p-q) = \lim_{L\to\infty} L\delta_{p,q}$$

In 4D with volume V and time T we thus have

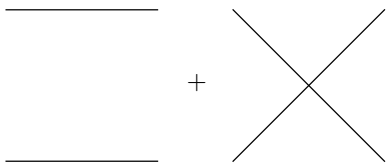
$$(2\pi)^4\delta^4(p-q) = \lim_{V,T\to\infty} VT\delta_{p,q}$$

While we can not square the expression on the l.h.s. directly, we can square the r.h.s. before taking the limit. This results in

$$((2\pi)^4\delta^4(p-q))^2 = \lim_{V,T\to\infty} V^2T^2\,\delta_{p,q}$$

16.2 The S-matrix and the T-matrix

Knowing now how to compute transition probabilities let us come back to the elastic scattering probability and let us start by investigating if we drew the leading order contribution. We in fact can easily see that we did not. The leading order contribution is $O(1)$ and looks like this:



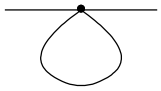
This is not an interesting contribution however. The two particles don't ineract and we will simply get momentum conservation.

Let us formalize this statements a little bit. In any process, we generically have an ingoing state and an outgoing state, both consisting of a number of particles. During a scattering process, there is a time evolution that transforms the ingoing into the outgoing states. This time evolution is governed by the interaction Hamiltonean of the system H_I and thus is unitary. We can thus generically transform any incoming state into an outgoing state with this unitary time evolution operator, which we call S, the scattering matrix. Unitarity means $S^\dagger S = SS^\dagger = 1$.

In the free KG theory or, equivalently, at $\lambda = 0$, this time evolution is trivial. No interaction takes place and particles simply propagate independently of each other, thus the S-matrix is trivial $S = 1$. In fact, even if we switch on an interaction, since it is only a small perturbation, it will affect the S-matrix only slightly. Therefore, one defines the T-matrix, which remover the trivial part of the S-matrix as $S = 1 - iT$. Unitarity of S implies

$$T^\dagger T = i(T - T^\dagger)$$

The $O(1)$ diagrams at the beginning of this subsection are there in the free theory already and therefore contribute to S but not to T. The interaction diagram however does contribute to T as well and we therefore want to compute it. Notice that this is the only nontrivial diagram contributing to this process at $O(\lambda)$. Other diagrams such as e.g.



only contribute to the renormalization of the mass and are therefore irrelevant for scattering amplitudes as we have seen.

Let us now suppose that we have sucessfully computed the T-matrix element of a certain process

$$\langle p_1 ... p_n | T | k_A k_B ... \rangle$$

in terms of (amputated) Feynman diagrams and let us collectively denote the sum of the incoming momenta as k and the outgoing as p. As we have seen, this will generically have the structure

$$\langle p_1 ... p_n | T | k_A k_B ... \rangle = (2\pi)^4 \delta^4(k - p) M(k_A k_B ... \to p_1 ... p_n)$$

where M is the regular part of the expression. We now compute the probability by taking the absolute square of the matrix element in a finite volume

$$\begin{aligned} p(k \to p) &= |\langle p_1 ... p_n | T | k_A k_B ... \rangle|^2 \frac{1}{N^2} \\ &= (VT)^2 \delta_{k,p} |M|^2 (k_A k_B ... \to p_1 ... p_n) \frac{1}{N^2} \end{aligned}$$

where N is the proper normalization factor of all states. For momentum states in infinite volume we had

$$\langle p | q \rangle = (2\pi)^3 \delta^3(\vec{p} - \vec{q}) 2E_p$$

which in finite volume becomes

$$\langle p | q \rangle = 2E_p V \delta_{\vec{p},\vec{q}}$$

and thus

$$N^2 = \prod_i 2E_i V \prod_f 2E_f V$$

where the first product runs over the initial states k and the second runs over the final states p. For the case of one and two incoming particles respectively we thus have

$$p(k \to p) = VT^2 \delta_{k,p} |M|^2 (k_A \to p_1 \dots p_n) \frac{1}{2E_A} \prod_f \frac{1}{2E_f V}$$

and

$$p(k \to p) = T^2 \delta_{k,p} |M|^2 (k_A k_B \to p_1 \dots p_n) \frac{1}{2E_A 2E_B} \prod_f \frac{1}{2E_f V}$$

Remembering (30), we can rewrite

$$V = \frac{(2\pi)^3}{\Delta p^3}$$

and thus get

$$p(k \to p) = VT^2 \delta_{k,p} |M|^2 (k_A \to p_1 \dots p_n) \frac{1}{2E_A} \prod_f \frac{\Delta p^3}{2E_f (2\pi)^3}$$

for the case of one incoming particle and

$$p(k \to p) = T^2 \delta_{k,p} |M|^2 (k_A k_B \to p_1 \dots p_n) \frac{1}{2E_A 2E_B} \prod_f \frac{\Delta p^3}{2E_f (2\pi)^3}$$

for the case of two.

16.3 Scattering cross sections and decay rates

Having computed a transition probability, we are finally ready to connect our result to an experimental observable. In scattering experiments, one usually measures cross sections σ. A cross section is a very intuitive quantity: When you shine geometric light onto a classical object, the cross section is simply the area of the shadow. More generically in a scattering experiment, the cross section is the area for which particles get deflected. There is also a differential cross section $d\sigma$, which is the part of the cross section that gets deflected at a certain angle or more generally, a certain momentum.

Let us now set up our scattering experiment in a finite volume box. We put particle A at rest into the box. Particle B will be placed somewhere at random on a surface of area F of the box and shot orthogonal to that surface with velocity v. The probability that an interaction will take place while B crosses the box from one surface to the opposite one is therefore obviously the ratio between the cross section and the total area A. We thus have

$$\frac{d\sigma}{F} = p(k \to p) = T^2 \delta_{k,p} |M|^2 (k_A k_B \to p_1 \dots p_n) \frac{1}{2E_A 2E_B} \prod_f \frac{\Delta p^3}{2E_f (2\pi)^3}$$

To traverse the box of length L, particle B will need a time $T = L/v$. Also, the volume of the box is simply $V = FL$. Thus we can write

$$\begin{aligned} d\sigma &= FT^2 \delta_{k,p} |M|^2 (k_A k_B \to p_1 \dots p_n) \frac{1}{2E_A 2E_B} \prod_f \frac{\Delta p^3}{2E_f (2\pi)^3} \\ &= VT \delta_{k,p} \frac{1}{2E_A 2E_B v} |M|^2 (k_A k_B \to p_1 \dots p_n) \prod_f \frac{\Delta p^3}{2E_f (2\pi)^3} \end{aligned}$$

We can now take the limit $V, T \to \infty$ and obtain our final result

$$d\sigma = (2\pi)^4 \delta^4 (k - p) \frac{1}{4E_A E_B v} |M|^2 (k_A k_B \to p_1 \dots p_n) \prod_f \frac{d^3 p}{2E_f (2\pi)^3} \tag{31}$$

For the case of elastic scattering of two identical particles in the center of mass frame, this expression simplifies to

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{|M|^2}{64\pi^2 E_{\text{CM}}^2} \tag{32}$$

where Ω is the solid angle of one of the outgoing particles (the other one leaves in exactly opposite direction of course).

In the case of a single incoming particle, we can only compute a sensible quantity when the particle is unstable and decays. In this case, the relevant quantity is the differential decay width $d\Gamma$. The decay width is simply the probability per time that the particle decays, thus

$$d\Gamma = \frac{p(k \to p)}{T} = VT\delta_{k,p}|M|^2(k_A \to p_1 \ldots p_n)\frac{1}{2E_A}\prod_f \frac{\Delta p^3}{2E_f(2\pi)^3}$$

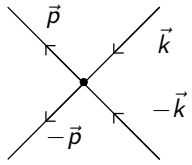
For a particle of mass m at rest, $E_A = m$ and taking the $V, T \to \infty$ limit results in

$$d\Gamma = (2\pi)^4\delta^4(k-p)\frac{1}{2m}|M|^2(k_A \to p_1 \ldots p_n)\prod_f \frac{d^3p}{2E_f(2\pi)^3}$$

Integrating over all final states one obtains the total decay width Γ, which is related to the lifetime of a particle $\tau = 1/\Gamma$

16.4 Tree level scattering crossection in ϕ^4 theory

We are now finally in a position to calculate a physical quantity, the differential scattering cross-section to $O(\lambda)$ in perturbation theory. We will compute this cross section in the center of mass frame, which allows us to use (32). All we need to do is to compute the T-matrix element M. We have already identified the only nontrivial contribution, namely the amputated diagram



The momentum asignments simplify since we are in the center of mass frame, but in any case, the diagram itself is just $-i\lambda$. Since the definition of the T-matrix contains a $-i$, the matrix elemet is simply λ and $|M|^2 = \lambda^2$. The differential cross section to leading order is thus

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{\lambda^2}{64\pi^2 E_{\text{CM}}^2}$$

which is an even angular distribution. Computing the total cross section is therefore (almost) trivial - one simply integrates over half the solid angle 2π, so

$$\sigma = \frac{\lambda^2}{32\pi E_{\text{CM}}^2}$$

The reason that the integration only runs over half the solid angle and not the full one is the fact that the two outgoing particles are indistinguishable. Field quanta do not carry a tag - they are literally indistinguishable. Integrating one final state over half the solid angle is therefore sufficient, since the second, indistinguishabe particle has then covered the other half. Going one order higher in perturbation theory we have to remember (29) and obtain

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{\Lambda^{(2)2}}{64\pi^2 E_{\text{CM}}^2}$$

which has a nontrivial dependence on the momenta of the outgoing particles.

17 The origin of field equations

We have started this course by constructing the Klein-Gordon equation as a relativistic generalization of the Schrödinger equation. Subsequently we saw that we first want to treat it as a classical field equation and then quantize the field. We thus started out with a classical, scalar field and it is natural to ask how we could possibly generalize our treatment to other kinds of fields. In order to investigate this, let us first see, what physical information the KG equation actualy encodes.

17.1 Translational invariance

We obtained the KG equation by demanding the relativistic dispersion relation

$$E^2 = \vec{p}^2 + m^2$$

which can copmactly be written as

$$p_\mu p^\mu = m^2$$

on a plane wave of 4-momentum p^μ resulting in the KG equation

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

Because we are for the moment looking at a classical field theory, $p^\mu = i\partial^\mu$ can not be interpreted as a momentum operator. It is however still related to a spacetime translation

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

with constant a^μ. Under this translation, ϕ transforms as

$$\begin{aligned} \phi(x) \to \phi(x-a) &= \phi(x) - a^\mu \partial_\mu \phi(x) + O(a^2) \\ &= \phi(x) + i a^\mu (i\partial_\mu)\phi(x) + O(a^2) \end{aligned}$$

If we make the translation a infinitesimal, we can identify the generator of the translation on the scalar field

$$P_\mu = i\partial_\mu$$

in the sense that a unitary translation operation $U(a)$ acting on the field ϕ is given by

$$U(a) = 1 + i a^\mu P_\mu = e^{i a^\mu P_\mu}$$

17.2 A field equation in a single dimension

Before we investigate the 3 + 1 dimensional case in detail, let us look at a simple example of 0 + 1 dimensions, i.e. a scalar "field" that has no spatial but only temporal extension. We call this field $q(t)$. If we demand time translation invariance, $E = p^0$ is constant, so

$$\dot{q} = \text{const.}$$

or, similarly

$$\ddot{q} = 0$$

which happens to be the classical equation of motion of a one dimensional mass point (at coordinate q) without external forces. Viewed in this way, the equation of motion is a simple statement about a quantity (E in this case) that is constant under the entire space-time symmetry group.

17.3 Field equations

Trying to generalize from the simple $0+1$ dimensional case, we try to find quantities that are invariant under the full space-time symmetry group in $3+1$ dimensions. In more mathematical terms, we are looking for quantities that commute with all generators of the full space-time symmetry group. It is not difficult to find such quantities: every constant c trivially commutes with all symmetry generators. The equation $c = \text{const.}$ however is trivial and tells us nothing about the dynamics of the field. To obtain an interesting relation, we therefore limit ourselves to quantities that are themselves nontrivial functions of the symmetry generators. Such combinations of symmetry generators that commute with all generators of the symmetry group are known as Casimir operators.

Let us try to find the Casimir operators of a $3+1$ dimensional scalar field. Looking at the $0+1$ dimensional example and noting that in $3+1$ dimensions all momenta commute

$$[p_\mu, p_\nu] = 0$$

one might be tempted to write $p^\mu = \text{const.}$, but this is not correct. The reason is, that while indeed the p^μ do all commute among one another, translation is not the full space-time symmetry group. In addition to translations we have rotations and Lorentz-boosts. These, together with the translations, form the full symmetry group of space-time, the Poincare group.

Anticipating a result that we will derive more rigorously in the next sections, we can already guess one Casimir operator of the Poincare group: While momenta p^μ are not conserved under rotations and Lorentz-boosts, we already know that the square $p^\mu p_\mu$ is invariant. If we call the pertaining constant m^2, we arrive at the relation

$$p^\mu p_\mu = m^2$$

which for a scalar field ϕ indeed implies the KG equation

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

The free scalar field equation is thus a statement about the space-time symmetry, namely the expression that the Casimir operator of the Poincare group is indeed constant.

Clearly it is interesting to explore whether $p^\mu p_\mu$ is the only Casimir operator of the Poincare group and how these Casimir operators look like for more general types of fields.

17.4 The Poincare group

The action of a 3-dimensional rotation on the coordinates x^μ may be expressed by an orthogonal rotation matrix Λ

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu$$

In more mathematical terms, the rotation matrices Λ form a representation of the group $O(3)$, the group of 3-dimensional orthogonal matrices. There are other representations of this group, e.g. the representation on scalar fields. All these representations share the common local structure of the group which is given by the commutation relations between their generators.

Let us first look at a rotation by an angle φ around the $x^3 = z$ axis. Such a rotation is given by

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\varphi) & \sin(\varphi) & 0 \\ 0 & -\sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

For infinitesimal φ, we can write

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \varphi (j_z)^{\mu}{}_{\nu}$$

with

$$(j_z)^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Defining a rotation by its axis is a peculiarity that applies to 3 dimensions only. In order to be more general, we therefore identify the rotation by its plane. A rotation around the $x^3 = z$ axis is a rotation in the x^1-x^2 plane and we therefore write

$$(j_z)^{\mu}{}_{\nu} = (j^{12})^{\mu}{}_{\nu} = g^{2\mu}\delta^1_{\nu} - g^{1\mu}\delta^2_{\nu}$$

which easily generalizes to

$$(j^{kl})^{\mu}{}_{\nu} = g^{l\mu}\delta^k_{\nu} - g^{k\mu}\delta^l_{\nu}$$

for the generators of general spatial rotations. In fact, we can generalize this expression even further to include "space-time rotations", i.e. Lorentz boosts

$$(j^{\alpha\beta})^{\mu}{}_{\nu} = g^{\beta\mu}\delta^{\alpha}_{\nu} - g^{\alpha\mu}\delta^{\beta}_{\nu}$$

Thus, e.g. a Lorentz-boost in $x = x^1$ direction has a generator

$$(j^{10})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and thus a rotation by a finite "angle" θ is given by

$$\Lambda^{\mu}{}_{\nu} = e^{\theta (j^{10})^{\mu}{}_{\nu}} = \begin{pmatrix} \cosh(\theta) & \sinh(\theta) & 0 & 0 \\ \sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

giving indeed the well-known Lorentz-boost.

Acting on scalar fields, we find that for infinitesimal ε

$$x^{\mu} \to (\delta^{\mu}_{\nu} + \varepsilon j^{\alpha\beta})^{\mu}{}_{\nu} x^{\nu} \quad \longrightarrow \quad \phi(x) \to \phi((\delta^{\mu}_{\nu} - \varepsilon j^{\alpha\beta})^{\mu}{}_{\nu} x^{\nu})$$

and

$$\begin{aligned} \phi((\delta^{\mu}_{\nu} - \varepsilon j^{\alpha\beta})^{\mu}{}_{\nu} x^{\nu}) &= \phi(x^{\mu} - \varepsilon (j^{\alpha\beta})^{\mu}{}_{\nu} x^{\nu}) \\ &= \phi(x) - \varepsilon (j^{\alpha\beta})^{\mu}{}_{\nu} x^{\nu} \partial_{\mu}\phi(x) \\ &= \phi(x) + i\varepsilon i(g^{\beta\mu}\delta^{\alpha}_{\nu} - g^{\alpha\mu}\delta^{\beta}_{\nu}) x^{\nu}\partial_{\mu}\phi(x) \\ &= \phi(x) + i\varepsilon i(x^{\alpha}\partial^{\beta} - x^{\beta}\partial^{\alpha})\phi(x) \end{aligned}$$

so that in the scalar field representation the generators are given by

$$J^{\alpha\beta} = i(x^\alpha \partial^\beta - x^\beta \partial^\alpha)$$

These are the generators of the Lorentz-group and, together with the already known generators of translation

$$P_\mu = i\partial_\mu$$

they form the Poincare group.

Exercise 5. Show that the generators of the Poincare group fulfill the commutation relations

$$\begin{array}{lcl} [P^\mu, P^\nu] & = & 0 \\ [P^\mu, J^{\alpha\beta}] & = & i(g^{\mu\alpha}P^\beta - g^{\mu\beta}P^\alpha) \\ [J^{\mu\nu}, J^{\alpha\beta}] & = & i(g^{\nu\alpha}J^{\mu\beta} + g^{\mu\beta}J^{\nu\alpha} - g^{\mu\alpha}J^{\nu\beta} - g^{\nu\beta}J^{\mu\alpha}) \end{array} \tag{33}$$

17.5 The Pauli-Lubanski Pseudovector

With the commutation relations (33) it is trivial to show that indeed $P^\mu P_\mu$ is a Casimir operator. There is also a second Casimir operator, which is less obvious, namely $W^\mu W_\mu$ where W^μ is the Pauli-Lubanski pseudovector

$$W_\mu := -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}J^{\nu\alpha}P^\beta$$

Exercise 6. Use the commutation relations (33) to prove that $W^\mu W_\mu$ is a Casimir operator of the Poincare group.

The Pauli-Lubanski pseudovector is orthogonal to the momentum

$$W_\mu P^\mu = -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}J^{\nu\alpha}P^\beta P^\mu = 0$$

Let us now imagine that we are in a quantum theory (without specifying in detail what kind of field we have) and P^μ and W^μ are promoted to operators. Let us further imagine that the field we are describing has a mass m and let us look at a momentum eigenstate of a single particle in its rest frame. the momentum of the particle is thus given by

$$k^\mu = \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and because of the orthogonality relation just derived

$$W^\mu = \begin{pmatrix} 0 \\ \vec{W} \end{pmatrix}$$

The Pauli-Lubanski vector is thus

$$W^i = \frac{1}{2}\varepsilon_{ijk0}J^{jk}m = mJ^i$$

where we have used the standard 3-dimensional indexing of the angular momentum by its axis

$$J^i = \frac{1}{2}\varepsilon_{ijk}J^{kj}$$

Thus for a massive particle at rest, $\vec{W}$ is just the mass times the angular momentum (spin) of the particle that we already know from nonrelativistic quantum mechanics. We conclude that for a massive particle, the nonrelativistic quantum mechanical description of the spin carries over. Furthermore, the second Casimir operator is just

$$W^\mu W_\mu = -m^2 \vec{J}\cdot\vec{J}$$

We know from nonrelativistic QM that the squre of the angular momentum of a state is given by

$$\vec{J}\cdot\vec{J} = (s+1)s$$

where s is the intrinsic angular momentum or spin of a particle and is restricted to be a positive half integer (or 0). We thus have found a completely generic reason for the quantization of spin for a massive particle that arises from just the symmetry properties of spacetime.

We can write the second equation of motion

$$W^\mu W_\mu = -m^2(s+1)s \tag{34}$$

and find for the special case of a scalar field with $s=0$ that it encodes a trivial $0=0$. For higher spin representations however we can expect this relation to result in an additional equation of motion.

We can also conclude that in a quantum field theory spin is an additional quantum number. Thus, for generic fiends, a momentum eigenstate $|k\rangle$ can be further characterized by e.g. the $z=x^3$ component of the spin $|k,s_3\rangle$ where $s_3 \in \{-s,-s+1,...,s-1,s\}$.

17.6 Massless particles

For the massless case, it is imposible to go to the rest frame. The best one can do to simplify the discussion is to choose the momentum of the particle to be along a certain axis, e.g. $z=x^3$. in this case, the momentum is

$$k^\mu = \begin{pmatrix} k \\ 0 \\ 0 \\ \pm k \end{pmatrix}$$

We construct the Pauli-Lubanski vector on an eigenstate of this momentum

$$\begin{aligned} W_\mu|k\rangle &= -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}J^{\nu\alpha}P^\beta|k\rangle \\ &= -\frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}J^{\nu\alpha}k^\beta|k\rangle \\ &= -\frac{1}{2}k(\varepsilon_{\mu\nu\alpha 0} \pm \varepsilon_{\mu\nu\alpha 3})J^{\nu\alpha}|k\rangle \end{aligned}$$

and thus

$$\begin{aligned} W_0|k\rangle &= \mp k J^{12}|k\rangle \\ W_1|k\rangle &= k(J^{23} \pm J^{02})|k\rangle \\ W_2|k\rangle &= k(J^{31} \pm J^{10})|k\rangle \\ W_3|k\rangle &= k J^{12}|k\rangle \end{aligned}$$

for the individual components. From the generic commutation relations (33) we find that

$$\begin{aligned} [J^{23} \pm J^{02}, J^{31} \pm J^{10}] &= [J^{23},J^{31}] \pm [J^{02},J^{31}] \pm [J^{23},J^{10}] + [J^{02},J^{10}] \\ &= -iJ^{21} + iJ^{21} = 0 \end{aligned}$$

and

$$\begin{aligned}[J^{23} \pm J^{02}, J^{12}] &= [J^{23}, J^{12}] \pm [J^{02}, J^{12}] = i(J^{13} \pm J^{01}) \\ [J^{31} \pm J^{10}, J^{12}] &= [J^{31}, J^{12}] \pm [J^{10}, J^{12}] = i(J^{23} \pm J^{02})\end{aligned}$$

so using the shorthand

$$\boldsymbol{p}_x = J^{23} \pm J^{02} \qquad \boldsymbol{p}_y = J^{31} \pm J^{10} \qquad \boldsymbol{j} = J^{12}$$

we find a group structure

$$\begin{aligned}[\boldsymbol{p}_x, \boldsymbol{p}_y] &= 0 \\ [\boldsymbol{p}_x, \boldsymbol{j}] &= -i\boldsymbol{p}_y \\ [\boldsymbol{p}_y, \boldsymbol{j}] &= i\boldsymbol{p}_x\end{aligned}$$

which is not the structure of the 3-dimensional angular momentum group. In fact, it is quite easy to see that these commutation relations encode the continuous spacetime symmetry of a 2-dimensional plane, i.e. translations in x and y direction and one rotation:

$$\begin{aligned}\boldsymbol{p}_x &= i\partial_x \\ \boldsymbol{p}_y &= i\partial_y \\ \boldsymbol{j} &= i(y\partial_x - x\partial_y)\end{aligned}$$

There is no quantization of angular momentum in this group. In fact, it is a mystery why massless particles in nature, as far as we know, do not appear with arbirtary spins. Maybe this is a hint that either the flat-space assumption underlying our symmetry considerations is not good enough or that fundamentally there are no massless particles and corresponding symmetries are only approximate. In any case, within the framework of quantum field theory in Minkovski spacetime the apparent quantization of spin of massless particles remains a mystery.

18 Vector fields

18.1 Equations of motion

Let us now look at a vector field $A^\mu(x)$. First we have to find the generators of the Poincare group in this representation. Let us start with a translation

$$\begin{aligned}x^\nu &\to x^\nu + a^\nu \\ A^\mu(x) &\to A^\mu(x-a) = A^\mu(x) - a^\nu \partial_\nu A^\mu(x) + O(a^2) = A^\mu(x) + i a^\nu i\partial_\nu A^\mu(x) + O(a^2)\end{aligned}$$

so we find that the generator of translations

$$P^\nu = i\partial^\nu$$

is the same as for a scalar field. From the first Casimir operator we thus find

$$P^\nu P_\nu A^\mu(x) = m^2 A^\mu(x)$$

or

$$(\partial^\nu \partial_\nu + m^2) A^\mu(x) = 0$$

so that the KG equation is fulfilled for every component of a vector field.

Next we investigate infinitessimal generalized rotations

$$x^\sigma \to \Lambda^\sigma{}_\nu x^\nu \qquad \Lambda^\sigma{}_\nu = \delta^\sigma_\nu + \varphi (j^{\alpha\beta})^\sigma{}_\nu$$

Since A^μ is a vector field, its components transform as a vector per definition and we find

$$\begin{aligned}
A^\mu(x) &\to \Lambda^\mu{}_\tau A^\tau((\Lambda^{-1})^\sigma{}_\nu x^\nu) \\
&= (\delta^\mu_\tau + \varphi(j^{\alpha\beta})^\mu{}_\tau)A^\tau((\delta^\sigma_\nu - \varphi(j^{\alpha\beta})^\sigma{}_\nu)x^\nu) \\
&= A^\mu(x) + \varphi((j^{\alpha\beta})^\mu{}_\tau A^\tau(x) - (j^{\alpha\beta})^\sigma{}_\nu x^\nu \partial_\sigma A^\mu(x)) \\
&= A^\mu(x) + i\varphi i(-(g^{\beta\mu}\delta^\alpha_\tau - g^{\alpha\mu}\delta^\beta_\tau)A^\tau(x) + (g^{\beta\sigma}\delta^\alpha_\nu - g^{\alpha\sigma}\delta^\beta_\nu)x^\nu\partial_\sigma A^\mu(x)) \\
&= A^\mu(x) + i\varphi i(-(g^{\beta\mu}\delta^\alpha_\tau - g^{\alpha\mu}\delta^\beta_\tau) + (g^{\beta\sigma}\delta^\alpha_\nu - g^{\alpha\sigma}\delta^\beta_\nu)x^\nu\partial_\sigma\delta^\mu_\tau)A^\tau(x) \\
&= A^\mu(x) + i\varphi i(-(g^{\beta\mu}\delta^\alpha_\tau - g^{\alpha\mu}\delta^\beta_\tau) + (x^\alpha\partial^\beta - x^\beta\partial^\alpha)\delta^\mu_\tau)A^\tau(x)
\end{aligned}$$

which implies that the generalized rotation generator is

$$(J^{\alpha\beta})^\mu{}_\tau = i(-(g^{\beta\mu}\delta^\alpha_\tau - g^{\alpha\mu}\delta^\beta_\tau) + (x^\alpha\partial^\beta - x^\beta\partial^\alpha)\delta^\mu_\tau)$$

and thus the Pauli-Lubanski vector is given by

$$\begin{aligned}
(W_\sigma)^\mu{}_\tau &= -\frac{1}{2}\varepsilon_{\sigma\alpha\beta\gamma}(J^{\alpha\beta})^\mu{}_\tau P^\gamma \\
&= \frac{1}{2}\varepsilon_{\sigma\alpha\beta\gamma}(-(g^{\beta\mu}\delta^\alpha_\tau - g^{\alpha\mu}\delta^\beta_\tau) + (x^\alpha\partial^\beta - x^\beta\partial^\alpha)\delta^\mu_\tau)\partial^\gamma
\end{aligned}$$

The last two terms are symmetric under the exchange $\beta \leftrightarrow \gamma$ resp. $\alpha \leftrightarrow \gamma$ and thus vanish because of the antisymmetry implied by ε. We therefore get

$$\begin{aligned}
(W_\sigma)^\mu{}_\tau &= -\frac{1}{2}\varepsilon_{\sigma\alpha\beta\gamma}(g^{\beta\mu}\delta^\alpha_\tau - g^{\alpha\mu}\delta^\beta_\tau)\partial^\gamma \\
&= \frac{1}{2}(\varepsilon_{\sigma\tau\gamma\beta}g^{\beta\mu} + \varepsilon_{\sigma\tau\gamma\alpha}g^{\alpha\mu})\partial^\gamma \\
&= \varepsilon_{\sigma\tau\gamma}{}^\mu\partial^\gamma
\end{aligned}$$

and thus the second Casimir operator is

$$\begin{aligned}
(W^\sigma)^\mu{}_\tau(W_\sigma)^\tau{}_\nu &= \varepsilon^{\tau\gamma\sigma\mu}\partial_\gamma\varepsilon_{\nu\alpha\sigma\tau}\partial^\alpha \\
&= \varepsilon^{\sigma\tau\gamma\mu}\partial_\gamma\varepsilon_{\sigma\tau\nu\alpha}\partial^\alpha \\
&= -2(\delta^\gamma_\nu\delta^\mu_\alpha - \delta^\gamma_\alpha\delta^\mu_\nu)\partial_\gamma\partial^\alpha \\
&= 2(\delta^\mu_\nu\partial^\alpha\partial_\alpha - \partial^\mu\partial_\nu)
\end{aligned}$$

The second equation of motion can thus be found from (34) with $s=1$

$$((W^\sigma)^\mu{}_\tau(W_\sigma)^\tau{}_\nu + 2m^2\delta^\mu_\nu)A^\nu(x) = 0$$

which results in

$$(\delta^\mu_\nu(\partial^\alpha\partial_\alpha + m^2) - \partial^\mu\partial_\nu)A^\nu(x) = 0$$

or equvalently

$$\partial_\nu(\partial^\nu A^\mu(x) - \partial^\mu A^\nu(x)) + m^2 A^\mu(x) = 0 \tag{35}$$

This second equation of motion for a vector field is known as the Proca equation. In the special case of $m=0$ we obtain

$$\partial_\nu(\partial^\nu A^\mu(x) - \partial^\mu A^\nu(x)) = 0$$

which is the free Maxwell equation. It can be more compactly written in terms of the antisymmetric field strength tensor

$$F^{\nu\mu}(x) := \partial^\nu A^\mu(x) - \partial^\mu A^\nu(x)$$

as

$$\partial_\nu F^{\mu\nu}(x) = 0$$

18.2 Lagrangean

Following the example of the scalar theory, we want to construct the Lagrangean as the next step. The generic Euler-Lagrange equations are

$$\frac{\partial \mathscr{L}}{\partial A_\mu} - \partial_\nu \frac{\partial \mathscr{L}}{\partial(\partial_\nu A_\mu)} = 0$$

From a term in the Lagrangean

$$\mathscr{L}_m = \frac{m^2}{2} A^\alpha A_\alpha$$

we obtain the mass term

$$\frac{\partial \mathscr{L}_m}{\partial A_\mu} = m^2 A^\mu$$

To correctly reproduce the kinetic term of (35), we need

$$-\frac{\partial \mathscr{L}}{\partial(\partial_\nu A_\mu)} = \partial^\nu A^\mu - \partial^\mu A^\nu$$

which can be achieved by a term in the Lagrangean

$$\begin{aligned}
\mathscr{L}_k &= -\frac{1}{2}\partial_\nu A_\mu (\partial^\nu A^\mu - \partial^\mu A^\nu) \\
&= -\frac{1}{4}(\partial_\nu A_\mu - \partial_\mu A_\nu)(\partial^\nu A^\mu - \partial^\mu A^\nu) \\
&= -\frac{1}{4} F_{\nu\mu} F^{\nu\mu} \\
&= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}
\end{aligned}$$

Both parts together form the Proca Lagrangean

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{m^2}{2} A^\mu A_\mu$$

of a vector field theory. In order to make it more accessible for the Hamiltonean approach that we will follow, let us explicitly decompose it into spacial and temporal components

$$\begin{aligned}
\mathscr{L} &= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{m^2}{2} A^\mu A_\mu \\
&= -\frac{1}{2}\partial_\nu A_\mu (\partial^\nu A^\mu - \partial^\mu A^\nu) + \frac{m^2}{2} A^\mu A_\mu \\
&= -\frac{1}{2}\partial_0 A_i (\partial^0 A^i - \partial^i A^0) \\
&\quad -\frac{1}{2}\partial_i A_0 (\partial^i A^0 - \partial^0 A^i) \\
&\quad -\frac{1}{2}\partial_j A_i (\partial^j A^i - \partial^i A^j) + \frac{m^2}{2} A^\mu A_\mu \\
&= \partial_0 A_i \partial^i A^0 - \frac{1}{2}(\partial_0 A_i \partial^0 A^i + \partial_i A_0 \partial^i A^0 + \partial_j A_i (\partial^j A^i - \partial^i A^j) - m^2 A^\mu A_\mu) \\
&= \frac{1}{2}(\dot{A}^i (\dot{A}^i + 2\partial_i A^0) + \partial_i A^0 \partial_i A^0 - \partial_j A^i (\partial_j A^i - \partial_i A^j) + m^2 A^\mu A_\mu)
\end{aligned}$$

Expressing the fields in terms of Fourier components

$$A^\nu(t,\vec{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \hat{A}^\nu(t,\vec{p})$$

we find for the Lagrangean

$$\begin{aligned}
L &= \int d^3x\, \mathscr{L}(x) \\
&= \int d^3x \Big(\partial_0 A_i \partial^i A^0 - \frac{1}{2}(\partial_0 A_i \partial^0 A^i + \partial_i A_0 \partial^i A^0 + \partial_j A_i (\partial^j A^i - \partial^i A^j) - m^2 A^\mu A_\mu) \Big) \\
&= \frac{1}{2}\int d^3x \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3q}{(2\pi)^3} e^{i(\vec{p}+\vec{q})\cdot\vec{x}} \big(\dot{\hat{A}}^i(t,\vec{p})\big(\dot{\hat{A}}^i(t,\vec{q}) + 2iq^i \hat{A}^0(t,\vec{q})\big) \\
&\quad -\vec{p}\cdot\vec{q}\,\hat{A}^0(t,\vec{p})\hat{A}^0(t,\vec{q}) + p^i\hat{A}^i(t,\vec{p})(q^j\hat{A}^i(t,\vec{q}) - q^i\hat{A}^j(t,\vec{q})) + m^2\hat{A}^\mu(t,\vec{p})\hat{A}_\mu(t,\vec{q})) \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3} \big(\dot{\hat{A}}^i\big(\dot{\hat{A}}^{*i} - 2ip^i\hat{A}^{*0}\big) + (m^2+\vec{p}^2)\hat{A}^\mu\hat{A}^*_\mu + p^jp^j\hat{A}^i\hat{A}^{*j}\big)
\end{aligned}$$

where in the last line we have used the shorthand notation $\hat{A}_\mu = \hat{A}_\mu(t,\vec{p})$ and $\hat{A}^{*\mu}{=}\hat{A}^\mu(t,-\vec{p})$.

18.3 Hamiltonean formulation

From the Proca Lagrangean we find the canonical field momentum as

$$\Pi^\alpha := \frac{\partial \mathscr{L}}{\partial(\partial_0 A_\alpha)} = \partial^i A^0 - \partial^0 A^\alpha = F^{\alpha 0}$$

which implies in particular $\Pi^0 = 0$. Thus, the Hamiltonean density is

$$\begin{aligned}
\mathscr{H} &= \partial_0 A_\alpha \Pi^\alpha - \mathscr{L} \\
&= \partial_0 A_\alpha \Pi^\alpha - \frac{1}{2}\partial_\mu A_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu) + \frac{m^2}{2} A^\mu A_\mu \\
&= \frac{\bar{\hat{\psi}}(p)1}{2}(\partial_0 A_\alpha \Pi^\alpha - \partial_i A_\nu(\partial^\nu A^i - \partial^i A^\nu) + m^2 A^\mu A_\mu) \\
&= \frac{1}{2}(\partial_0 A_i \Pi^i - \partial_i A_0(\partial^0 A^i - \partial^i A^0) - \partial_i A_j(\partial^j A^i - \partial^i A^j) + m^2 A^\mu A_\mu) \\
&= \frac{1}{2}(\partial_0 A_i \Pi^i + \partial_i A_0 \Pi^i + \partial_i A_j(\partial^i A^j - \partial^j A^i) + m^2 A^\mu A_\mu) \\
&= \frac{1}{2}(-\Pi^i(\Pi_i - 2\partial_i A_0) + \partial_i A_j(\partial^i A^j - \partial^j A^i) + m^2 A^\mu A_\mu) \\
&= \frac{1}{2}(\Pi^i(\Pi^i - 2\partial^i A^0) + \partial_i A_j(\partial^i A^j - \partial^j A^i) + m^2 A^\mu A_\mu)
\end{aligned}$$

Following our derivation for the scalar field, we will however contiune in momentum space. We construct the canonically conjugate momenta to the fields $(dp/2\pi)^{\frac{3}{2}}\hat{A}^\alpha(t,\vec{p})$ as

$$(dp/2\pi)^{\frac{3}{2}}\hat{\Pi}_\alpha := (dp/2\pi)^{\frac{3}{2}}\frac{\delta L}{\delta \dot{\hat{A}}^\alpha}$$

First we see that no time derivatives of A^0 appear in the Lagrangean, thus $\hat{\Pi}_0 = 0$. This makes explicit that A^0 is not a dynamical variable. For the spatial components we have

$$(dp/2\pi)^{\frac{3}{2}}\hat{\Pi}_i := (dp/2\pi)^{\frac{3}{2}}\frac{\delta L}{\delta \dot{\hat{A}}^\alpha} = (dp/2\pi)^{\frac{3}{2}}\big(\dot{\hat{A}}^{*i} - ip^i\hat{A}^{*0}\big)$$

and thus

$$\dot{\hat{A}}^i = \hat{\Pi}^*_i - ip^i\hat{A}^0 \qquad \dot{\hat{A}}^{*i} = \hat{\Pi}_i + ip^i\hat{A}^{*0}$$

and thus the Hamiltonian in momentum space is

$$\begin{aligned}
H &= \int \frac{d^3p}{(2\pi)^3}\hat{\Pi}_i(t,\vec{p})\dot{\hat{A}}^i(t,\vec{p}) - L \\
&= \int \frac{d^3p}{(2\pi)^3}\left(\hat{\Pi}_i(\hat{\Pi}_i^* - ip^i\hat{A}^0) - \frac{1}{2}\big(\dot{\hat{A}}^i\big(\dot{\hat{A}}^{*i} - 2ip^i\hat{A}^{*0}\big) + (m^2+\vec{p}^2)\hat{A}^\mu\hat{A}_\mu^* + p^jp^j\hat{A}^i\hat{A}^{*j}\big)\right) \\
&= \int \frac{d^3p}{(2\pi)^3}\left(\hat{\Pi}_i(\hat{\Pi}_i^* - ip^i\hat{A}^0) - \frac{1}{2}((\hat{\Pi}_i^* - ip^i\hat{A}^0)(\hat{\Pi}_i - ip^i\hat{A}^{*0}) + (m^2+\vec{p}^2)\hat{A}^\mu\hat{A}_\mu^* + p^jp^j\hat{A}^i\hat{A}^{*j})\right) \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}(\hat{\Pi}_i(\hat{\Pi}_i^* - ip^i\hat{A}^0) + (\hat{\Pi}_i^* - ip^i\hat{A}^0)ip^i\hat{A}^{*0} - (m^2+\vec{p}^2)\hat{A}^\mu\hat{A}_\mu^* - p^jp^j\hat{A}^i\hat{A}^{*j}) \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}((\hat{\Pi}_i + ip^i\hat{A}^{*0})(\hat{\Pi}_i^* - ip^i\hat{A}^0) - (m^2+\vec{p}^2)\hat{A}^\mu\hat{A}_\mu^* - p^jp^j\hat{A}^i\hat{A}^{*j}) \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}((\hat{\Pi}_i + ip^i\hat{A}^{*0})(\hat{\Pi}_i^* - ip^i\hat{A}^0) + (m^2+\vec{p}^2)(\hat{A}^i\hat{A}^{*i} - \hat{A}^0\hat{A}^{*0}) - p^jp^j\hat{A}^i\hat{A}^{*j}) \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}(\hat{\Pi}_i\hat{\Pi}_i^* + ip^i(\hat{A}^{*0}\hat{\Pi}_i^* - \hat{A}^0\hat{\Pi}_i) + (m^2+\vec{p}^2)\hat{A}^i\hat{A}^{*i} - m^2\hat{A}^0\hat{A}^{*0} - p^jp^j\hat{A}^i\hat{A}^{*j})
\end{aligned}$$

Contrary to the case of the scalar field, it is not obvious that the Hamiltonean is bounded from below. We will come back to this issue later, for now let us write down the full set of equations of motion. For the time component we have

$$\begin{aligned}
\frac{\partial H}{\partial\Big((dp/2\pi)^{\frac{3}{2}}\hat{\Pi}_0\Big)} &= 0 &&= (dp/2\pi)^{\frac{3}{2}}\dot{\hat{A}}^0 \\
\frac{\partial H}{\partial\Big((dp/2\pi)^{\frac{3}{2}}\hat{A}^0\Big)} &= -(dp/2\pi)^{\frac{3}{2}}(m^2\hat{A}^{*0} - ip^i\hat{\Pi}_i) &&= -(dp/2\pi)^{\frac{3}{2}}\dot{\hat{\Pi}}_0
\end{aligned}$$

which, remembering that $\hat{\Pi}_0 = 0$, implies

$$ip^i\hat{\Pi}_i = m^2\hat{A}^{*0} \tag{36}$$

and specifically for $m = 0$

$$ip^i\hat{\Pi}_i = 0$$

For the spatial components we have

$$\begin{aligned}
\frac{\partial H}{\partial\Big((dp/2\pi)^{\frac{3}{2}}\hat{\Pi}_i\Big)} &= (dp/2\pi)^{\frac{3}{2}}(\hat{\Pi}_i^* - ip^i\hat{A}^0) &&= (dp/2\pi)^{\frac{3}{2}}\dot{\hat{A}}^i \\
\frac{\partial H}{\partial\Big((dp/2\pi)^{\frac{3}{2}}\hat{A}^i\Big)} &= (dp/2\pi)^{\frac{3}{2}}\Big((m^2+\vec{p}^2)\hat{A}^{*i} - p^ip^j\hat{A}^{*j}\Big) &&= -(dp/2\pi)^{\frac{3}{2}}\dot{\hat{\Pi}}_i
\end{aligned}$$

which implies

$$\ddot{\hat{A}}^i = \dot{\hat{\Pi}}_i^* - ip^i\dot{\hat{A}}^0 = (m^2+\vec{p}^2)\hat{A}^i - p^ip^j\hat{A}^j \tag{37}$$

In the massless case this simplifies to

$$\ddot{\hat{A}}^i = \vec{p}^2\hat{A}^i - p^ip^j\hat{A}^j$$

As in the scalar case, the momentum components again decouple.

18.4 Polarization

Although we have disentangled momentum components in the Hamiltonean, we have not fully diagonalized the canonical equations of motion. To do so, we write (37) in the form

$$\ddot{\hat{A}}^i \;=\; -((\vec{p}^2+m^2)\delta^{ij}-p^ip^j)\hat{A}^j$$

We thus need to diagonalize the matrix

$$M^{ij}=(\vec{p}^2+m^2)\delta^{ij}-p^ip^j$$

which, written explicitly, looks like this

$$M=\begin{pmatrix} (p^2)^2+(p^3)^2+m^2 & p^1p^2 & p^1p^3 \\ p^1p^2 & (p^1)^2+(p^3)^2+m^2 & p^2p^3 \\ p^1p^3 & p^2p^3 & (p^1)^2+(p^2)^2+m^2 \end{pmatrix}$$

One eigenvector, corresponding to longitudinal polarization (i.e. the component of the vector field along the momentum direction) is $\vec{p}$. We have

$$M^{ij}p^j=((\vec{p}^2+m^2)\delta^{ij}-p^ip^j)p^j=m^2p^i$$

As M is a symmetric matrix, the remaining two eigenvectors need to be orthogonal to $\vec{p}$ and orthogonal to each other. Let us take an arbitrary unit vector $\vec{\varepsilon}\perp\vec{p}$. Applying M on any such vector results in

$$M^{ij}\varepsilon^j=((\vec{p}^2+m^2)\delta^{ij}-p^ip^j)\varepsilon^j=(\vec{p}^2+m^2)\varepsilon^i$$

so the eigenvalue of M on all these vectors is $(\vec{p}^2+m^2)$. We can thus choose any two mutually orthogonal unit vectors $\vec{p}\perp\vec{\varepsilon}^{(1)}\perp\vec{\varepsilon}^{(2)}\perp\vec{p}$ representing transverse polarization modes of the field that span a 2-dimensional subspace of degenerate eigenmodes of M. Defining a third, longitudinal unit vector $\vec{\varepsilon}^{(3)}:=\vec{p}/\sqrt{\vec{p}^2}$, we can decompose the spatial components of our vector field in momentum space

$$(dp/2\pi)^{3/2}\hat{A}^i=\varepsilon^{(k)i}\varphi^k$$

with an implied sum over $k=1,2,3$. This allows us to cast (37) into the form

$$\begin{aligned}\ddot{\varphi}^k&=-(\vec{p}^2+m^2)\varphi^k \qquad k=1,2\\ \ddot{\varphi}^3&=-m^2\varphi^3\end{aligned}$$

exposing the full dynamical content of the vector theory. The two transverse modes are again harmonic oscillators with an angular frequency $\omega^2=\vec{p}^2+m^2$. Thus we can carry out the quantization procedure of the scalar field for these two modes unchanged. The longitudinal mode is more tricky however. In the massive case it is simply an additional oscillator with $\omega_l^2=m^2$. In case $m=0$ though we have $\ddot{\varphi}^3=0$ and thus a "free particle" instead of an oscillator and the energy spectrum is continuous. Without discretized energy levels a particle interpretation of the longitudinal mode is no longer possible though, exposing its unphysical nature in the massless case. Indeed this behaviour just reflects the fact that a wave traveling at light speed can not have a longitudinal structure. Already in classical electrodynamics an electromagnetic wave does only have the two transverse polarization directions and not the longitudinal one.

Skipping all the intermediate steps that are equivalent to the scalar case, we quantize our vector field and arrive at a normal ordered Hamiltonean

$$H=\int\frac{d^3p}{(2\pi)^3}\left(\sum_{l=1}^{2}(\vec{p}^2+m^2)a^{(l)\dagger}(t,\vec{p})a^{(l)}(t,\vec{p})+m^2a^{(3)\dagger}(t,\vec{p})a^{(3)}(t,\vec{p})\right) \tag{38}$$

for the case $m^2 > 0$ or

$$H = \int \frac{d^3p}{(2\pi)^3} \sum_{l=1}^{2} \vec{p}^2 a^{(l)\dagger}(t,\vec{p}) a^{(l)}(t,\vec{p}) \tag{39}$$

for the case $m^2 = 0$ where the $a^{(i)\dagger}(\vec{p})$ and $a^{(i)}(\vec{p})$ are creation and annihilation operators for transverse ($i = 1,2$) and longitudinal ($i = 3$) polarization particles with the usual commutation relations

$$[a^{(k)}(t,\vec{p}), a^{(l)\dagger}(t,\vec{q})] = (2\pi)^3 \delta^3(\vec{p}-\vec{q})\delta^{kl}$$

Note that the Hamiltoneans (38,39) are indeed positive definite. In terms of these, the fields and conjugate momenta for $m > 0$ are (compare (10))

$$\begin{aligned} A^j(t,\vec{x}) &= \int \frac{d^3p}{(2\pi)^3} \frac{e^{i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega(\vec{p})}} \sum_{i=1}^{n} \left(\varepsilon^{(i)j} a^{(i)}(t,\vec{p}) + \varepsilon^{*(i)j} a^{(i)\dagger}(t,-\vec{p}) \right) \\ \Pi_j(t,\vec{x}) &= i \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \sqrt{\frac{\omega(\vec{p})}{2}} \sum_{i=1}^{n} \left(\varepsilon^{*(i)j} a^{(i)\dagger}(t,-\vec{p}) - \varepsilon^{(i)j} a^{(i)}(t,\vec{p}) \right) \end{aligned}$$

with the number of polarization states $n = 3$ for $m^2 > 0$ and $n = 2$ for $m = 0$. Consequently we have the commutation relations

$$\begin{aligned} [A^j(t,\vec{x}), A^k(t,\vec{y})] &= 0 \\ [\Pi_j(t,\vec{x}), \Pi_k(t,\vec{y})] &= 0 \\ [A^j(t,\vec{x}), \Pi_k(t,\vec{y})] &= \delta^3(\vec{x}-\vec{y}) \varepsilon^{*(i)j} \varepsilon^{(i)k} \end{aligned}$$

Because of the orthonormality of the $\varepsilon^{(i)}$, we have in the massive case (where the implicit sum over i runs from 1 to 3)

$$\varepsilon^{*(i)j} \varepsilon^{(i)k} = \delta^{jk}$$

while for $m = 0$ we have

$$\varepsilon^{*(i)j} \varepsilon^{(i)k} = \delta^{jk} - \frac{p^j p^k}{\vec{p}^2}$$

18.5 Propagator of the massless vector field

We now specialise our discussion to the massless case. We see from (12) and (39) that in the Heisenberg picture the field and conjugate momenta can be expressed in terms of creation and annihilation operators at a reference time as

$$\begin{aligned} A^j(x) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p^0}} \sum_{i=1}^{2} \left(\varepsilon^{(i)j} e^{-ip\cdot x} a^{(i)}(\vec{p}) + \varepsilon^{*(i)j} e^{ip\cdot x} a^{(i)\dagger}(\vec{p}) \right) \\ \Pi_j(x) &= i \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{p^0}{2}} \sum_{i=1}^{2} \left(\varepsilon^{*(i)j} e^{ip\cdot x} a^{(i)\dagger}(\vec{p}) - \varepsilon^{(i)j} e^{-ip\cdot x} a^{(i)}(\vec{p}) \right) \end{aligned}$$

In complete analogy to the derivation of (14,19) we thus find the propagator to be

$$\begin{aligned} \langle 0|A^a(x)A^b(y)|0\rangle &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2p^0} e^{-ip\cdot(x-y)} \varepsilon^{*(i)a} \varepsilon^{(i)b} \\ &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{i}{p^2 + i\varepsilon} \varepsilon^{*(i)a} \varepsilon^{(i)b} \end{aligned} \tag{40}$$

with a sum over polarization states (i) implied.

Exercise 7. Derive the expression (40)

In the derivation of the vector field propagator we have silently disposed of unneeded field components. In the massless case there were two such components: the timelike A^0 and the longitudinal $\vec{p}\cdot\vec{A}$. This corresponds to imposing two conditions on our quantum field

$$A^0 = 0 \qquad \vec{p}\cdot\vec{A} = 0$$

which resulted in the deletion of the respective timelike and longitudinal components from the polarization matrix

$$M^{\alpha\beta} = \varepsilon^{*(i)\alpha}\varepsilon^{(i)\beta}$$

in (40). The remaining, transverse components fulfilled the standard KG equation

$$\partial_\mu \partial^\mu M_\alpha{}^\beta A^\alpha = 0$$

Remember that the entire quantization procedure was based on using the Maxwell equation which resulted from the second Casimir operator. We now saw that this implies the KG equation to be fulfilled for the transverse polarization components. The first Casimir operator on the other hand was giving us the KG equation for all polarization components. Thus the behaviour of the two transverse components is unique, but there seems to be some kind of ambiguity as to the longitudinal and the timelike component and eliminating them both entirely might be too strict.

To see what is happening here, let us simply assume that we quantized using only the first Casimir operator, i.e. the KG equation on each component individually. This would result in the propagator

$$\langle 0|A^\alpha(x)A^\beta(y)|0\rangle \;=\; \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{-i}{p^2+i\varepsilon} g^{\alpha\beta}$$

and the polarization matrix is simply $g^{\alpha\beta}$. This however can not be the full story: the propagation of the longitudinal and timelike modes violate in general the constraint from the second Casimir operator, i.e. the Maxwell equation. We thus have to impose those additional constraints as boundary conditions.

Comparing the Maxwell equation

$$\partial_\mu(\partial^\mu A^\alpha - \partial^\alpha A^\mu) = 0$$

to the KG equation

$$\partial_\mu \partial^\mu A^\alpha = 0$$

we see that the additional constraint we have to impose is

$$\partial^\alpha \partial_\mu A^\mu = 0$$

which is fulfilled e.g. by

$$A^0 = 0 \qquad \vec{p}\cdot\vec{A} = 0$$

We can however be more sloppy and just demand

$$\partial_\mu A^\mu = 0 \tag{41}$$

which in momentum space reads

$$p_\mu A^\mu = 0$$

which implies that the transverse 4-momentum component is missing from the polarization tensor in the propagator, i.e.

$$\langle 0|A^\alpha(x)A^\beta(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{-i}{p^2+i\varepsilon} P^{\alpha\beta} \tag{42}$$

with

$$P^{\alpha\beta} = g^{\alpha\beta} - \frac{p^\alpha p^\beta}{p^\mu p_\mu}$$

We will use this expression for the propagaror of a massless scalar field for the remainder of the term, keeping in mind that external states always come in just the two transverse polarization states.

The propagator (42) is not unique. In fact, massless vector fields exhibit a gauge symmetry and the condition (41) choses a specific gauge, namely the Landau gauge. Correspondingly, (42) is the propagator of a free massless vector field in Landau gauge. A lot of interesting things could be said about gauge symmetry and fixing a gauge which we have to omit for the reason of time constraints.

19 Fermions

In quantum mechanics, we have encountered fermions that have the strange property that they anti-commute instead of commute. To construct a quantum field theory, we need to take the classical limit of fermionic variables first and make a field theory from them. The classical limit of fermionic variables are Grassmann variables that have the generic property that they anticommute. So if ψ and θ are two Grassmann numbers, they fulfill

$$\{\psi, \theta\} := \psi\theta + \theta\psi = 0$$

which trivially implies

$$\begin{aligned} \psi^2 &= 0 \\ \psi\theta &= -\theta\psi \\ [\psi\theta, \psi\theta] &= 0 \end{aligned}$$

Because of the third property we can define a scalar field as a product of two Grassmann fields

$$\phi(x) = \psi(x)\theta(x)$$

Let us assume that ϕ fulfills the KG equation

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

or, in momentum space

$$(p_\mu p^\mu - m^2)\hat{\phi}(p) = 0$$

at a single momentum we thus have

$$p_\mu p^\mu \hat{\psi}(p)\hat{\theta}(p) = m^2 \hat{\psi}(p)\hat{\theta}(p)$$

and we may ask how to split up the operators $p_\mu p^\mu$ and m^2 to act on one Grassmann component only. For the r.h.s. this is trivial as we can write (omitting the momentum arguments for better readability)

$$p_\mu p^\mu \hat{\psi}\hat{\theta} = (m\hat{\psi})(m\hat{\theta})$$

If we straightforwardly try to do the same for the l.h.s. we have a problem: p^μ is not a scalar, so the simple split

$$p_\mu \hat{\psi} = m\hat{\psi} \qquad p_\mu \hat{\theta} = m\hat{\theta}$$

does not make any sense. Similarly, using the scalar $\sqrt{p_\mu p^\mu}$ results in a nonlocal operator $\sqrt{\partial_\mu \partial^\mu}$ which would not allow to formulate equal time commutation relations.

To get a clue how to proceed, we look at the trivial case of $1+0$ dimensions. There we simply have

$$(p\hat{\psi})(p\hat{\theta}) = (m\hat{\psi})(m\hat{\theta})$$

which can be decomposed into

$$p\hat{\psi} = m\hat{\psi} \qquad p\hat{\theta} = m\hat{\theta}$$

which are sensible equations for both $\hat{\psi}$ and $\hat{\theta}$. In order to disentangle the $p_\mu p^\mu$ in the general case, we would need a vector γ^μ with the property that

$$\frac{1}{2}\{\gamma^\mu, \gamma^\nu\} = g^{\mu\nu} \tag{43}$$

With its help we can write

$$p_\mu p^\mu = p_\mu \gamma^\mu p_\nu \gamma^\nu$$

and thus disentangle the general KG equation as

$$(p_\mu p^\mu - m^2)\hat{\psi}\hat{\theta} = -\hat{\theta}(p_\mu p^\mu - m^2)\hat{\psi} = \hat{\theta}(-\not{p}\not{p} + m^2)\hat{\psi} = 0$$

where we have defined $\not{p} := p_\mu \gamma^\mu$. We can further rewrite this as

$$\hat{\theta}(-\not{p} - m)(\not{p} - m)\hat{\psi} = 0$$

Let us therefore propose the equations of motion

$$\begin{aligned} (\not{p} - m)\hat{\psi} &= 0 \\ (-\not{p} - m)^\dagger \hat{\theta}^\dagger &= 0 \end{aligned}$$

An explicit representation of the γ^μ that satisfies (43) is

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \gamma^i &= \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \end{aligned}$$

from which we immediately see that

$$\gamma^{0\,\dagger} = \gamma^0 \qquad \gamma^{i\,\dagger} = -\gamma^i$$

This in turn implies

$$\gamma^{\mu\,\dagger}\gamma^0 = \gamma^0\gamma^\mu$$

and therefore

$$\gamma^0(-\not{p} - m)^\dagger \hat{\theta}^\dagger = (-\not{p} - m)\gamma^0\hat{\theta}^\dagger = 0$$

Comparing this to

$$(\not{p} - m)\hat{\psi}(p) \;=\; 0 \tag{44}$$

suggests that we can identify

$$\hat{\psi}(p) = \gamma^0 \hat{\theta}^\dagger(-p)$$

We define

$$\hat{\bar{\psi}}(p) := \hat{\psi}^\dagger(-p)\gamma^0 = \hat{\theta}(p) \tag{45}$$

so in fact (44), which is the momentum space Dirac equation, implies the KG equation.

19.1 Poincare group for Grassmann fields

For the moment the Dirac equation is just a candidate for a valid equation of motion for Grassmann fields. What we still lack is a proof that it is invariant under the Poincare group, i.e. that $\not{p}$ is a Casimir operator. In order to show this, we must establish the generators of the Poincare group in the Grassmann representation. As usual we start with translations

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

As in the case of scalar and vector fields we have

$$\begin{aligned} \psi(x) \to \psi(x-a) &= \psi(x) - a^\mu \partial_\mu \psi(x) + O(a^2) \\ &= \psi(x) + i a^\mu (i\partial_\mu)\psi(x) + O(a^2) \end{aligned}$$

and the generator of a translation in μ direction is readily identified as

$$P_\mu = i\partial_\mu$$

which does indeed commute with $\not{p}$.

In order to understand how a Grassmann field transforms under generalized rotations, we remember that γ^μ transforms as a contravariant vector. Thus under a transformation

$$x^\mu \to (\delta^\mu_\nu + \varepsilon j^{\alpha\beta})^\mu{}_\nu x^\nu \tag{46}$$

with infinitesimal ε and

$$(j^{\alpha\beta})^\mu{}_\nu = g^{\beta\mu}\delta^\alpha_\nu - g^{\alpha\mu}\delta^\beta_\nu$$

we find that

$$\gamma^\mu \to (\delta^\mu_\nu - \varepsilon j^{\alpha\beta})^\mu{}_\nu \gamma^\nu = \gamma^\mu - \varepsilon(g^{\beta\mu}\gamma^\alpha - g^{\alpha\mu}\gamma^\beta) = \gamma^\mu + \varepsilon g^{\mu[\alpha}\gamma^{\beta]}$$

where the brackets on the indices indicate antisymmetrization. Let us now denote the transformation of the Grassmann field under (46) by Λ so that a Grassmann constant κ transforms as

$$\kappa \to \Lambda\kappa$$

From the previous section we already know that for a constant κ (i.e. at p=0) $\kappa^\dagger\kappa$ is a scalar and $\kappa^\dagger\gamma^\mu\kappa$ is a vector. Thus under (46) we have

$$\kappa^\dagger\gamma^\mu\kappa \to \kappa^\dagger(\gamma^\mu + \varepsilon g^{\mu[\alpha}\gamma^{\beta]})\kappa = \kappa^\dagger\Lambda^\dagger\gamma^\mu\Lambda\kappa$$

and if we make the ansatz

$$\Lambda = 1 + \varepsilon J^{\alpha\beta}$$

we obtain

$$-g^{\beta\mu}\gamma^{\alpha} + g^{\alpha\mu}\gamma^{\beta} = -[J^{\alpha\beta}, \gamma^{\mu}]$$

which suggests that

- J consists of γ's
- J is a sum of products of 2 γ's as it has two indices
- One of the γ's must coincide with γ^{μ} for the comutator not to vanish, so the γ's are γ^{α} and γ^{β}
- $J^{\alpha\beta}$ is antisymmetric in α and β

We therefore make the ansatz

$$J^{\alpha\beta} = c[\gamma^{\alpha}, \gamma^{\beta}]$$

and

$$\begin{aligned}
[[\gamma^{\alpha}, \gamma^{\beta}], \gamma^{\mu}] &= \gamma^{\alpha}\gamma^{\beta}\gamma^{\mu} - \gamma^{\beta}\gamma^{\alpha}\gamma^{\mu} - \gamma^{\mu}\gamma^{\alpha}\gamma^{\beta} + \gamma^{\mu}\gamma^{\beta}\gamma^{\alpha} \\
&= -\gamma^{\beta}\{\gamma^{\alpha}, \gamma^{\mu}\} + 2g^{\alpha\beta}\gamma^{\mu} + 2g^{\mu\beta}\gamma^{\alpha} + \gamma^{\alpha}\{\gamma^{\beta}, \gamma^{\mu}\} - 2g^{\alpha\beta}\gamma^{\mu} - 2g^{\alpha\mu}\gamma^{\beta} \\
&= 4g^{\mu\beta}\gamma^{\alpha} - 4g^{\alpha\mu}\gamma^{\beta}
\end{aligned}$$

implies that $c = \frac{1}{4}$ so that we finally have

$$J^{\alpha\beta} = \frac{1}{4}[\gamma^{\alpha}, \gamma^{\beta}]$$

Exercise 8. Compute the full generators of Lorentz transformations for Grassmann fields and show that $\not{\partial}$ is a Casimir operator.

19.2 The Dirac Lagrangean

The Dirac equation in coordinate space reads

$$(i\not{\partial} - m)\psi(x) = 0$$

To construct a scalar, we can left multiply with $\bar{\psi}(x)$ and it turns out that this expression

$$\mathscr{L}(x) = \bar{\psi}(x)(i\not{\partial} - m)\psi(x)$$

is already the proper Lagrangean density. Remembering that $\psi(x)$ and $\bar{\psi}(x)$ are independent variables, we simply get

$$\frac{\partial \mathscr{L}}{\partial \bar{\psi}} = (i\not{\partial} - m)\psi = \partial_{\nu}\frac{\partial \mathscr{L}}{\partial(\partial_{\nu}\bar{\psi})} = 0$$

and

$$\frac{\partial \mathscr{L}}{\partial \psi} = -\bar{\psi}m = \partial_{\nu}\frac{\partial \mathscr{L}}{\partial(\partial_{\nu}\psi)} = i\partial_{\nu}\bar{\psi}\gamma^{\nu}$$

While the first of the above equations is simply the Dirac equation, we can write the second one as

$$(i\not{\partial}^{\dagger} - m)(\bar{\psi})^{\dagger} = 0$$

which, remembering that $\bar{\psi} = \psi^\dagger \gamma^0$, can be rewritten as

$$(i\not{\partial}^\dagger - m)(\bar{\psi})^\dagger = (i\not{\partial}^\dagger - m)\gamma^0\psi = \gamma^0(i\not{\partial} - m)\psi = 0$$

which again is equivalent to the dirac equation.

Rewriting the Lagrangean in Fourier space we obtain

$$\begin{aligned} L &= \int d^3x\, \bar{\psi}(x)(i\not{\partial} - m)\psi(x) \\ &= \int d^3x \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3q}{(2\pi)^3} e^{i(\vec{p}+\vec{q})\cdot\vec{x}}\, \hat{\bar{\psi}}(t,\vec{q})(i\partial_0\gamma^0 + p^i\gamma^i - m)\hat{\psi}(t,\vec{p}) \\ &= \int \frac{d^3p}{(2\pi)^3}\, \hat{\bar{\psi}}(t,-\vec{p})(i\partial_0\gamma^0 + p^i\gamma^i - m)\hat{\psi}(t,\vec{p}) \\ &= \int \frac{d^3p}{(2\pi)^3}\, \hat{\psi}^\dagger(t,\vec{p})(i\partial_0 + \gamma^0 p^i\gamma^i - \gamma^0 m)\hat{\psi}(t,\vec{p}) \end{aligned}$$

which also follows directly from (44) with the identification (45).

19.3 The Dirac Hamiltonean

Following our standard construction, we find the canonical field momenta (omitting phase space factors of $(dp/2\pi)^{\frac{3}{2}}$)

$$\hat{\Pi}(t,\vec{p}) = \frac{\delta L}{\delta\dot{\hat{\psi}}(p)} = i\hat{\bar{\psi}}(t,-\vec{p})\gamma^0 = i\hat{\psi}^\dagger(t,\vec{p})$$

while the canonical field momentum for the $\bar{\psi}$ vanishes. (Note that this again is a convention dependent statement as we could have moved the time derivative to the $\bar{\psi}$ when constructing the Lagrangean). We thus find for the Hamiltonean

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3}\hat{\Pi}(t,\vec{p})\partial_0\hat{\psi}(t,\vec{p}) - L \\ &= \int \frac{d^3p}{(2\pi)^3}\hat{\psi}^\dagger(t,\vec{p})(-\gamma^0 p^i\gamma^i + \gamma^0 m)\hat{\psi}(t,\vec{p}) \\ &= i\int \frac{d^3p}{(2\pi)^3}\hat{\Pi}(t,\vec{p})(\gamma^0 p^i\gamma^i - \gamma^0\, m)\hat{\psi}(t,\vec{p}) \end{aligned}$$

and the canonical equations of motion for a single component are

$$\dot{\hat{\psi}} = \frac{\partial H}{\partial\hat{\Pi}} = i\gamma^0(p^i\gamma^i - m)\hat{\psi} \qquad \dot{\hat{\Pi}} = -\frac{\partial H}{\partial\hat{\psi}} = -i\hat{\Pi}\gamma^0(p^i\gamma^i - m)$$

which is again just the Dirac equation. These equations do not look like they describe a classical harmonic oscillator, but let us investigate them a bit more closely. The structure of these equations is

$$\dot{\hat{\psi}} = iM\hat{\psi}$$

where the 4×4 matrix M is given by

$$M = \gamma^0(p^i\gamma^i - m) = \begin{pmatrix} -\vec{\sigma}\cdot\vec{p} & -m \\ -m & \vec{\sigma}\cdot\vec{p} \end{pmatrix}$$

To diagonalize this matrix, we first note that

$$(\vec{\sigma}\cdot\vec{p})^2 = \sigma^i\sigma^j p^i p^j = \delta^{ij}p^i p^j = \vec{p}\cdot\vec{p}$$

and thus $\vec{\sigma}\cdot\vec{p}$ has eigenmodes $\pm|\vec{p}|$.

Exercise 9. Show explicitly that in its eigenbasis

$$\vec{\sigma}\cdot\vec{p} = \begin{pmatrix} |\vec{p}| & 0 \\ 0 & -|\vec{p}| \end{pmatrix}$$

Thus in the proper basis we can write

$$M = \begin{pmatrix} \mp|\vec{p}| & -m \\ -m & \pm|\vec{p}| \end{pmatrix}$$

and we find the eigenvalues by the characteristic equation

$$\begin{vmatrix} \mp|\vec{p}| - \lambda & -m \\ -m & \pm|\vec{p}| - \lambda \end{vmatrix} = \lambda^2 - |\vec{p}|^2 - m^2 = 0$$

so the eigenvalues of M are simply given by the energy $E = \sqrt{|\vec{p}|^2 + m^2}$. Denoting the eigenvectors corresponding to the eigenvalues $+E$ as $v_{1/2}$ and the ones corresponding to the negative eigenvalues $-E$ as $u_{1/2}$ and demanding the normalization convention

$$v_i^\dagger v_j = u_i^\dagger u_j = 2E\delta_{ij}$$

we can write

$$M = \frac{1}{2}(u_i u_i^\dagger - v_i v_i^\dagger)$$

and decompose the field and conjugate momentum into

$$\begin{aligned} \hat{\psi} &= \frac{1}{\sqrt{2E}}(u_i a_i + v_i b_i^\dagger) \\ -i\hat{\Pi}\gamma^0 = \hat{\bar{\psi}} &= \frac{1}{\sqrt{2E}}(\bar{u}_i a_i^\dagger + \bar{v}_i b_i) \end{aligned}$$

We can now rewrite the total hamiltonean as

$$\begin{aligned} H &= i\int \frac{d^3p}{(2\pi)^3}\hat{\Pi}(t,\vec{p})M(\vec{p})\hat{\psi}(t,\vec{p}) \\ &= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{2E}(u_i^\dagger a_i^\dagger + v_i^\dagger b_i)(u_j u_j^\dagger - v_j v_j^\dagger)(u_k a_k + v_k b_k^\dagger) \\ &= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{2E}(2E u_i^\dagger a_i^\dagger - 2E v_i^\dagger b_i)(u_k a_k + v_k b_k^\dagger) \\ &= \int \frac{d^3p}{(2\pi)^3}E(\vec{p})(a_i^\dagger(\vec{p})a_i(\vec{p}) - b_i(\vec{p})b_i^\dagger(\vec{p})) \end{aligned}$$

where in the last line we have reinserted the momentum dependence in the notation.

20 Quantization of the Grassmann field

The Hamiltonean we derived in the previous section already looks conspicuously like the Hamiltonean of a quantum mechanical system although we did not impose any quantization condition yet. In addition, if we try to interpret it classically, the basic anticommutation property of Grassmann variables implies

$$a_i^\dagger a_i = -a_i^\dagger a_i = 0$$

and thus $H=0$. So it seems that a classical theory of anticommuting field does not make sense, which is not really surprising given that only quantum mechanics allowed for them to appear in the first place.

The only way to avoid the triviality of the Hamiltonean is to demand a nonvanishing anticommutator between the a_i and $a_i^\dagger$ resp. b_i and $b_i^\dagger$. Since we want our field quanta to have the proper energy $E=\sqrt{\vec{p}^2+m^2}$, it is obvious that we want

$$\begin{array}{rclcrcl} \{a_i(\vec{p}),a_j^\dagger(\vec{q})\} &=& \{b_i(\vec{p}),b_j^\dagger(\vec{q})\} &=& (2\pi)^3\delta(\vec{p}-\vec{q})\delta_{ij} \\ \{a_i(\vec{p}),a_j(\vec{q})\} &=& \{b_i(\vec{p}),b_j(\vec{q})\} &=& 0 \\ \{a_i(\vec{p}),b_j(\vec{q})\} &=& \{a_i(\vec{p}),b_j^\dagger(\vec{q})\} &=& 0 \end{array}$$

as our quantization condition. The normal ordered Hamiltonean then reads

$$\begin{aligned} H &= :\int\frac{d^3p}{(2\pi)^3}E(\vec{p})(a_i^\dagger(\vec{p})a_i(\vec{p})-b_i(\vec{p})b_i^\dagger(\vec{p})): \\ &= \int\frac{d^3p}{(2\pi)^3}E(\vec{p})(a_i^\dagger(\vec{p})a_i(\vec{p})+b_i^\dagger(\vec{p})b_i(\vec{p})) \end{aligned}$$

where we have made the obvious generalization that for fermionic variables the normal ordering anticommutes creation operators to the left and annihilation operators to the right. We also note in passing that the Wick theorem holds for fermionic fields with the obvious replacement of commutators by anticommutators.

20.1 The fock space

We can interpret $a_i^\dagger(\vec{p})$, $b_i^\dagger(\vec{p})$ as creation operators of Dirac particles and $a_i(\vec{p})$, $b_i(\vec{p})$ as the respective annihilation operators. The anticommutation relations require $(a_i^\dagger(\vec{p}))^2=(b_i^\dagger(\vec{p}))^2=0$ which implies that the maximum occupation number of each state is 1, in contrast to scalar and vector fields, where we could have an arbitrary number of particles with the same momentum. This is the well-known Pauli principle. This property in fact allowed us to obtain a positive normal ordered Hamiltonean even though the eigenvalues of the classical Hamiltonean were $\pm E$ in the following way: Since the maximum occupation number of a state is 1, we can redefine creation and annihilation operators. For a single mode, we can denote an occupied state by

$$|1\rangle=\begin{pmatrix}1\\0\end{pmatrix}$$

and an empty state by

$$|0\rangle=\begin{pmatrix}0\\1\end{pmatrix}$$

In this basis, creation and annihilation operators are given by

$$a^\dagger=\begin{pmatrix}0&1\\0&0\end{pmatrix}\qquad a=\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

with $\{a^\dagger,a\}=1$ and the full set of relations is

$$\begin{array}{rclcrcl} a^\dagger|0\rangle &=& |1\rangle & \quad & a^\dagger|1\rangle &=& 0 \\ a|1\rangle &=& |0\rangle & & a|0\rangle &=& 0 \end{array}$$

which is obviously symmetric under

$$\begin{array}{rcl} a^\dagger &\Leftrightarrow& a \\ |0\rangle &\Leftrightarrow& |1\rangle \end{array}$$

In the above derivation, we have used this liberty to identiy the annihilation operator of a negative energy state $b_i^\dagger$ with the creation operator of a positive energy state.

20.2 The Dirac propagator

We are now ready to compute the propagator of a Dirac particle. The Green's function of the Dirac equation is defined via

$$(i\partial\!\!\!/ - m)G(x) = i\delta^4(x)$$

which in momentum space reads

$$(p\!\!\!/ - m)\hat{G}(p) = i$$

so

$$\hat{G}(p) = \frac{i}{p\!\!\!/ - m} = i\frac{p\!\!\!/ + m}{p^2 - m^2}$$

and, using the Feynman prescription to circumvent the poles, we have the Feynman propagator for the free fermion field

$$\langle 0|\, T(\bar{\psi}(x)\psi(y))|0\rangle = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{i(p\!\!\!/ + m)}{p^2 - m^2 + i\varepsilon}$$

21 Quantum Electrodynamics (QED)

We have now seen that we can write the Lagrangean density of a free, massless vector field (photon) and a free Dirac field (electron) as

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(x)(i\partial\!\!\!/ - m)\psi(x)$$

Let us now, following our strategy in the scalar case, try to find sensible interaction terms. In order to obtain a scalar quantity, we must use at least two fermion fields. As a fermion field squares to zero, it is natural to look for a bilinear expression such as $\bar{\psi}\psi$. We could couple this to the square of the vector field, but there is in fact a simpler possibility: If we can construct a vector out of two fermion fields, we could couple it to the vector field A_μ itself. We know how to construct a vector out of a fermion bilinear and thus the suggested form of the interaction term would be

$$\mathscr{L}_{\text{int}} = e\bar{\psi}\gamma^\mu\psi A_\mu$$

with an arbitrary coupling constant e. Simple dimensional analysis tells us that $[e] = 1$, so e is dimensionless and the interaction term does not obviously destroy renormalizability. In addition, the interaction term has the very curious property of enhancing the symmetry of the theory. While the free Lagrangean is symmetric under the global replacement

$$\begin{aligned} \psi(x) &\to e^{i\alpha}\psi(x) \\ \bar{\psi}(x) &\to \bar{\psi}(x)e^{-i\alpha} \end{aligned}$$

the full Lagrangean, including the interaction term, is invariant under the local (gauge) symmetry

$$\begin{aligned} \psi(x) &\to e^{i\alpha(x)}\psi(x) \\ \bar{\psi}(x) &\to \bar{\psi}(x)e^{-i\alpha(x)} \\ A_\mu(x) &\to A_\mu(x) - \frac{i}{e}\partial_\mu\alpha(x) \end{aligned}$$

One can write the full Lagrangean density in a compact form

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(x)(iD\!\!\!/ - m)\psi(x)$$

where we have defined the covariant derivative

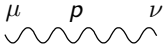
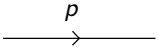
$$D_\mu := \partial_\mu - ieA_\mu$$

21.1 Feynman rules

The definition of the interaction Lagrangean and the resulting interaction Hamiltonean

$$\mathscr{H}_{\text{int}} = -e\bar{\psi}\gamma^\mu\psi A_\mu$$

finally allow us to complete the full set of Feynman rules for QED:

Incoming electron line (momentum p, spin i)	$u_i(p)$
Outgoing electron line (momentum p, spin i)	$\bar{u}_i(p)$
Incoming positron line (momentum p spin i)	$\bar{v}_i(p)$
Outgoing positron line (momentum p spin i)	$v_i(p)$
Incoming photon (momentum p, polarization i)	$\varepsilon^{(i)\mu}(p)$
Outgoing photon (momentum p, polarization i)	$\varepsilon^{*(i)\mu}(p)$
photon propagator 	$\frac{p^{\mu\nu}}{p^2 - m^2 + i\varepsilon}$
fermion propagator 	$\frac{i}{\not{p} - m}$
interaction vertex 	$-ie\gamma^\mu$
each line starting and ending at the same point	$\frac{1}{2}$
k lines between same points	$\frac{1}{k!}$
impose momentum conservation at each vertex	
impose momentum conservation on every subgraph	
a factor -1 for every closed fermion line	
integrate over all open (loop) momenta k	$\int \frac{d^4k}{(2\pi)^4}$