Quantum Field Theory – Part 1

Rainer F. Hauser
rainer.hauser@gmail.com

May 28, 2021

Abstract

Quantum Field Theory combines Special Relativity and Quantum Mechanics and was initially developed to quantize the electromagnetic field. It later turned out that all fundamental interactions are necessarily described by Quantum Field Theory. Tobias Osborne from the Leibniz Universität Hannover held two times 18 lectures in 2016 and 2017, respectively, covering this topic. His lectures were available on YouTube at the time this transcript has been assembled and may as "Quantum Field Theory" and "Advanced Quantum Field Theory" still be available today.

1 Introduction to Quantum Field Theory

1.1 Quantum Field Theory as an Effective Theory

Physicists want a theory of everything. It has not yet been found, and some people believe that it will be string theory. It is generally assumed that it has to do with quantum mechanics and gravity and is therefore a quantum gravitational theory for several reasons. One argument for this assumption is that if one wants to measure something with exact certainty in quantum mechanics, one has to invest so much energy that gravity becomes relevant. Quantum mechanics is a statistical theory and delivers only relative probabilities for the outcomes of experiments. If one wants to estimate a physical quantity such as the expectation value $\langle H \rangle$ of the energy of a system to an arbitrarily accurate level, one has to repeat the experiment very very often to get better and better statistics. Therefore, one runs into physical limits because one has to store the results, for example, in a computer. There are no limits coming from quantum mechanics on how many times an experiment can be repeated, but the data has to be stored somewhere. Eventually, one has to store so much data that one creates a black hole. This is a rather fancy argument, but it is not so far from the arguments convincing physicists that quantum mechanics and gravity must be merged somehow to obtain the final theory of everything.

Human beings cannot directly test and interact with the fundamental degrees of freedom of the universe and can only perceive the universe by low energy and large scale experiments. Only regimes are accessible where relativity and quantum mechanics are important. In this sense relativistic quantum field theory is an effective theory for this theory of everything and only describes the degrees of freedom accessible either in the laboratories on Earth or, indirectly, by observing astronomical events. Thus, it is not possible to distinguish between the universe being described by continuous quantum fields or whatever the building blocks are for the theory of everything. Relativistic quantum field theory is the first effective field theory of the universe. It is the one that seems to work best, and it describes every experiment made so far.

For even lower energies and larger scales there is another effective theory called non-relativistic quantum field theory. It is quite useful for describing condensed matter systems and other regimes with quantum effects but extremely cold temperatures. It is not possible to make good experiments probing relativistic or non-relativistic quantum field theory but one can only access these theories through the lens of decoherence – meaning noise. When there is enough noise there is the even better approximation called classical field theory. A hundred years ago the experiments had so much noise in them that classical field theory was a good description of nature.

The word *quantization* is used very often but suggests the wrong thing. Physics does not take a classical theory and uses it to build a quantum theory. The goal is always the theory of everything, and this theory was always a quantum theory, but because of decoherence one does not always see the quantum character. Quantizing means guessing a quantum theory (or actually the final theory of everything) such that its low energy, large scale, high decoherence limit is the classical theory. Physics is always going backwards from the classical theories through effective theories to the theory of everything and is therefore undoing the approximations and guessing what is inside the box of the nature. Physics is not yet there but is currently stuck at the effective theory called relativistic quantum field theory.

1.2 Notations and Conventions

The units used are such that $\hbar = c = 1$. Therefore, the dimensions of length and time are the same, the dimensions of the energy and mass are the same, and the dimension of energy is the reciprocal dimension of length, or more formally, $[\text{length}] = [\text{time}] = [\text{energy}]^{-1} = [\text{mass}]^{-1}$.

As a convention, 3-vectors are written as $\vec{x} = (x_1, x_2, x_3)$ and 4-vectors as $X = (X_0, X_1, X_2, X_3)$ where X_0 is the temporal component. Sometimes the temporal component is called t and the 4-vector X will be written as $X = (t, \vec{x})$. Greek indices such as μ represent the coordinates of a 4-vector, and Latin indices such as j indicate the coordinates of a 3-vector, with the summation convention in both cases. Usually the notation for an operator is $\boldsymbol{O}$ if not other conventions for special operators have historically established in the literature.

Note that in lectures and scripts covering quantum field theory symbols may be used within one equation for different purposes. The symbol μ, for example, is sometimes used as an index and as a mass or the symbol e is used for exponentials and for the fine structure constant. The same can happen with other symbols such as π used as the number 3.14... and as an operator or as a permutation. Different usages are usually not a problem because the meaning is clear but it is still not very nice.

1.3 Poincaré Group

For a quantum theory to be relativistic, it is certainly needed that it has to be symmetric under the Poincaré group. If $X = (x_0, x_1, x_2, x_3)$ are spacetime coordinates in an inertial reference frame, then the coordinates $X' = (x'_0, x'_1, x'_2, x'_3)$ in any other reference frame must satisfy

$$\eta_{\mu\nu}\, dX'^{\mu}\, dX'^{\nu} = \eta_{\mu\nu}\, dX^{\mu}\, dX^{\nu} \qquad \text{or} \qquad \eta_{\mu\nu} \frac{dX'^{\mu}}{dX^{\rho}} \frac{dX'^{\nu}}{dX^{\sigma}} = \eta_{\rho\sigma} \tag{1.1}$$

where

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.2}$$

is the metric in spacetime with +1 for the temporal coordinate as generally used in particle physics and quantum field theory. Any coordinate transformation satisfying (1.1) is linear and of the form

$$X'^{\mu} = \Lambda^{\mu}{}_{\nu} X^{\nu} + A^{\mu} \tag{1.3}$$

where $\Lambda^{\mu}_{\ \nu}$ are the components of a Lorentz transformation Λ and A^{μ} are the components of a shift A in spacetime. The Lorentz transformation Λ is a 4×4 matrix obeying

$$\eta_{\mu\nu}\, \Lambda^{\mu}_{\ \rho}\, \Lambda^{\nu}_{\ \sigma} = \eta_{\rho\sigma} \qquad \text{or} \qquad \Lambda^{T}\, \eta\, \Lambda = \eta \tag{1.4}$$

and A is a constant 4-vector.

These transformations (Λ, A) form a group $\mathfrak{P}_4$ with the binary operation $*$ whose meaning is that one transformation is applied after the other. The identity is obviously $(\mathbb{I}, 0)$, for every transformation (Λ, A) there is an inverse (Λ', A') such that $(\Lambda', A') * (\Lambda, A) = (\mathbb{I}, 0)$, and the result of two transformations (Λ, A) and $(\bar{\Lambda}, \bar{A})$ applied one after the other is the transformation $(\bar{\Lambda}, \bar{A}) * (\Lambda, A) = (\bar{\Lambda}\,\Lambda, \bar{\Lambda} A + \bar{A})$. In addition, as a group it also satisfies associativity.

1.4 Properties of a Relativistic Quantum Theory

Symmetries in quantum mechanics are represented by either unitary or antiunitary operators. The theory of angular momentum, for example, is the theory of symmetry under rotations. Similarly, any relativistic quantum theory must be a theory of symmetry under Lorentz transformations.

To build a relativistic quantum theory the following ingredients are needed:

1. a (separable) Hilbert space $\mathfrak{H}$
2. to every $(\Lambda, A) \in \mathfrak{P}_4$ one can find a unitary transformation $\boldsymbol{U}(\Lambda, A) : \mathfrak{H} \to \mathfrak{H}$ such that
   (i) $\boldsymbol{U}(\mathbb{I}, 0) = e^{i\phi}\,\mathbf{1}$ with a phase $\phi \in \mathbb{R}$
   (ii) $\boldsymbol{U}(\bar{\Lambda}, \bar{A})\boldsymbol{U}(\Lambda, A) = e^{i\phi((\Lambda,A),(\bar{\Lambda},\bar{A}))}\,\boldsymbol{U}(\bar{\Lambda}\,\Lambda, \bar{\Lambda}A + \bar{A})$ with a phase $\phi((\Lambda, A), (\bar{\Lambda}, \bar{A})) \in \mathbb{R}$

where there is an infinite number of equations in (ii). Any family $\boldsymbol{U}(\Lambda, A)$ satisfying (i) and (ii) is called a projective unitary representation of $\mathfrak{P}_4$. However, there are no non-trivial finite-dimensional unitary representations of $\mathfrak{P}_4$ in contrast to the group of rotations SO(3) and the theory of angular momentum. The mathematical reason is that SO(3) is a compact group but $\mathfrak{P}_4$ is not.

The subgroup $\mathfrak{V}_\mathrm{t} = \{(\mathbb{I}, (t, 0, 0, 0) \,|\, t \in \mathbb{R}\}$ of shifts in time is called the time translation subgroup of $\mathfrak{P}_4$. If $\boldsymbol{U}$ is a unitary representation of $\mathfrak{P}_4$ then $\boldsymbol{V}(t) \equiv \boldsymbol{U}(\mathbb{I}, (t, 0, 0, 0))$ for $t \in \mathbb{R}$ is a one-parameter family of unitaries $\boldsymbol{V}(s)\boldsymbol{V}(t) = \boldsymbol{V}(s + t)$ which solve Schrödinger's equation

$$\frac{d\boldsymbol{V}(t)}{dt} = i\,\boldsymbol{H}\,\boldsymbol{V}(t)$$

for some self-adjoint $\boldsymbol{H}$. Schrödinger's equation tells how time passes in quantum mechanics. Thus, having a unitary representation of the Poincaré group is certainly stronger than the Schrödinger equation. Physics requires that $\boldsymbol{U}(\Lambda, A)$ is "positive energy" and therefore the spectrum of $\boldsymbol{H}$ is contained in $\mathbb{R}^+$. Otherwise the system becomes unstable.

All such single-particle unitary representations of $\mathfrak{P}_4$ have been classified by Wigner and are labeled by the two numbers mass m and helicity/spin s. In other words, one cannot only find unitary transformations which obey all the constraints above but one can even classify them and know all of them in the case of singe-particle theories. The bad news, however, is that nature chooses not to be a theory of single particles. There is a tension between locality and interactions. It turns out that as soon as one imposes the constraint of locality one is forced to allow the theory to have not a definite number of particles. As a though experiment one can imagine to put a particle into a box such that the location is constraint and the momentum is a bit unclear due to the uncertainty principle. By squeezing the box it becomes more and more certain where the particle is, but the momentum becomes more uncertain. By more and more squishing it, the momentum and therefore the energy reaches the threshold to create more particles via interactions. Thus, by localizing a single particle one allows quantum mechanics to tunnel into configurations with more than one particle.

2 Classical Field Theory

2.1 Quantization as an Educated Guess

As discussed above classical field theory is an extremely good effective theory for relativistic quantum field theory describing those degrees of freedom measurable in case of low energy, large scale and in the presence of decoherence. Similarly, water looks like a continuum when observed with the eye despite the fact that it consists of molecules, and if one wants to model the physics of a fluid a continuum classical field is a very good effective theory.

Classical field theory here is used to help guessing a quantum field theory. This process is often called quantization, but quantization means an educated guess and not a magical procedure to build quantum theories. Thus, quantization consists of two steps. Firstly, the classical theory is used to guess a quantum theory, and secondly, it has to be checked that the low energy, large scale and decoherence limit gives back the original classical theory. This was the path from classical mechanics to quantum mechanics but also the path from classical field theory to quantum field theory.

2.2 Dynamics of a Classical Field

A field ϕ is a quantity defined at every point in a manifold $\mathfrak{M}$. Examples of such a quantity are density, spin and charge, and the manifold is here Minkowski spacetime $\mathbb{M}_{1,3} = \mathbb{R}^1 \times \mathbb{R}^3$ but may be something else in other contexts. More formally, a field is a function $\phi \in C^2(\mathfrak{M}, \mathfrak{S})$ from a manifold to a target space $\phi : \mathfrak{M} \to \mathfrak{S}$. Examples of target spaces are $\mathbb{R}$ (scalar field), $\mathbb{R}^N$ (vector fields), $\mathbb{S}^2$ (surface of a sphere) and $\mathbb{S}^1 \times \mathbb{S}^1$ (surface of a torus). A vector field $\phi(X) \in \mathbb{R}^N$ for all $X \in \mathfrak{M}$ is often written in components as $\phi_a(X)$ with $a = 1, ..., N$ where $\phi_a(X) \in \mathbb{R}$ are scalar fields.

The development in time of a system of classical scalar fields $\phi_a(X)$ restricted to fields with dynamics obtained by means of variational principle applied to an action functional can be described by a Lagrangian density $\mathcal{L}$ which is in principle a function of ϕ_a, $\partial_\mu \phi_a$, $\partial_\mu \partial_\nu \phi_a$ and so on. The action functional is

$$S(\Omega) = \int_\Omega \mathcal{L}(\phi_a, \partial_\mu \phi_a)\, d^4X \tag{2.1}$$

where Ω is a measurable region in spacetime. It is assumed that $\mathcal{L}$ is only a function of ϕ_a and $\partial_\mu \phi_a$ and not also of $\partial_\mu \partial_\nu \phi_a$ and higher order derivatives. One reason for this restriction to first order derivatives is that it gives all the equations of motion needed for what has been observed in nature.

It is assumed that the functional $S(\Omega)$ in (2.1) is stationary under variations $\phi_a(X) \to \phi_a(X) + \delta\phi_a(X)$. This means that $S(\Omega)$ changes only in second order for small changes to the configuration. It is further assumed that these variations vanish on the boundary such that $\delta\phi_a(X) = 0$ on $\delta\Omega$. Thus, with a Taylor expansion to first order one gets

$$\begin{aligned}
\delta S(\Omega) &= \int_\Omega d^4X \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a} \delta\phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta(\partial_\mu \phi_a) \right\} = 0 \\
&= \int_\Omega d^4X \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a} - \frac{\partial}{\partial X^\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \right\} \delta\phi_a + \int_\Omega d^4X \, \frac{\partial}{\partial X^\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \, \delta\phi_a \right) \\
&= \int_\Omega d^4X \left\{ \frac{\partial \mathcal{L}}{\partial \phi_a} - \frac{\partial}{\partial X^\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \right\} \delta\phi_a = 0
\end{aligned}$$

using integration by parts first and Green's theorem afterwards to get rid of the surface term. Therefore

$$\frac{\partial \mathcal{L}}{\partial \phi_a} - \frac{\partial}{\partial X^\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) = 0 \tag{2.2}$$

must hold for $a = 1, ..., N$ because $S(\Omega)$ is stationary for all $\delta\phi_a$. These equations are the equations of motion of the fields ϕ_a and are called the Euler-Lagrange equations. Thus, the Lagrangian density $\mathcal{L}$ encodes the equations of motion of a field ϕ in a very compact form.

If the equations of motions should be symmetric under some kind of transformations such as the Poincaré transformations, one way of doing so is by designing the Lagrangian density which is a scalar under these symmetry transformations. Therefore, the equations of motion are guaranteed to be symmetric under these symmetry transformations.

2.3 Klein-Gordon Field

The Klein-Gordon field ϕ as an example of a relativistically invariant classical field theory has

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

as its Lagrangian density. One can list all Lagrangian densities up to second order in ϕ and second order in the derivatives of ϕ. The Lagrangian density (2.3) is basically the most general object that is invariant under Lorentz transformations.

Using

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

the actual equations of motions become

$$\partial_\mu \partial^\mu \phi + m^2 \phi = \Box \phi + m^2 \phi = 0 \tag{2.4}$$

from (2.2). The resulting equation (2.4) is called the Klein-Gordon equation. This equation has to be brought into a form allowing to guess the corresponding quantum field theory.

2.4 Hamiltonian Formalism

One of the best ways to guess a quantum theory from a classical theory is not via the Lagrangian formalism but via the Hamiltonian formalism. The Lagrangian formalism would be the better way to start when using the method of path integral instead but this is still a part of the theory not yet in a rigorous state. Therefore the way via the Hamiltonian formalism is used here.

To guess quantum theories with the classical limit determined by a Lagrangian density $\mathcal{L}$ one has to find pairs of conjugate variables and then impose canonical commutation relations. If $\phi_a(X)$ is the canonical position then one defines via

$$\pi_a(X) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a} \tag{2.5}$$

the conjugate momentum density.

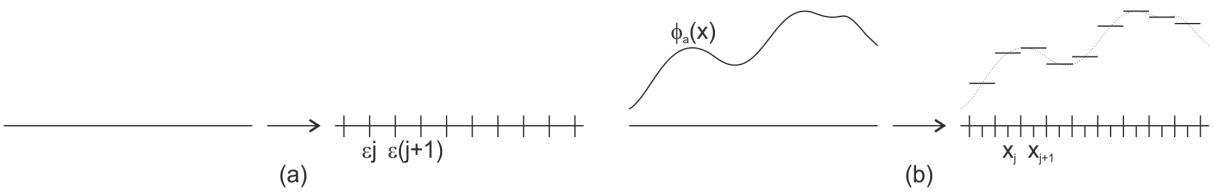


Figure 1: Discretization of a one-dimensional field

There is a justification for formula (2.5) coming from how one might study fields in general. To solve complicated physical systems usually computers are needed. Since fields are infinite, space has to be discretized into a regular lattice of spacing ε. Restricted to $\mathbb{M}_{1,1}$ this means that space is sampled as illustrated in figure 1 (a) resulting in discrete points $..., \varepsilon(j-1), \varepsilon j, \varepsilon(j+1), ...$ for $j \in \mathbb{Z}$. The generalized coordinates $q_j^a(t) = \phi_a(t, x_j)$ are sampled at the midpoints between two discrete points in space as shown in figure 1 (b). The question is how the equations of motion look like in this discretized version of the system. In order to determine them the action is discretized and the variational principle is applied. The term $\partial_\mu \phi_a(t, \vec{x})$ becomes approximated

$$\partial_x \phi_a(t, x_j) \approx \frac{\phi_a(t, x_j + \varepsilon) - \phi_a(t, x_j)}{\varepsilon} = \frac{q_{j+1}^a(t) - q_j^a(t)}{\varepsilon} \approx \dot{q}_j^a$$

for ε as small as possible. The Lagrangian in $\mathbb{M}_{1,3}$ is

$$L(t) = \int d^3\vec{x}\, \mathcal{L}(\phi_a(t, \vec{x}), \partial_\mu \phi_a(t, \vec{x}))$$

in terms of the Lagrangian density $\mathcal{L}$. Discretized in $\mathbb{M}_{1,1}$ it becomes

$$L(t) = \int dx\, \mathcal{L}\big(\phi_a(t,x), \partial_\mu \phi_a(t,x)\big) \approx \sum_j \delta x_j\, \mathcal{L}\big(\phi_a(t,x_j), \partial_\mu \phi_a(t,x_j)\big) = L\big(q_j^a(t), \dot{q}_j^a(t)\big)$$

and is a function of $q_j^a(t)$ and $\dot{q}_j^a(t)$ only. As in classical mechanics, if q_j^a is the canonical position variable then

$$p_j^a = \frac{\partial L}{\dot{q}_j^a} = \sum_k \delta x_k \frac{\partial \mathcal{L}}{\dot{q}_j^a}$$

is the conjugate momentum. It makes sense to define a conjugate momentum density $\pi^a(t, x_j)\, \delta x_j$ for a continuous field instead of the discrete conjugate momentum p_j^a.

As an example the probably simplest scalar field with no mass term in $\mathbb{M}_{1,1}$ defined through the Lagrangian density $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi\,\partial^\mu\phi)$ is used to demonstrate the procedure. Discretizing this Lagrangian density gives

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi\,\partial^\mu\phi) = \frac{1}{2}(\partial_t\phi)^2 - \frac{1}{2}(\partial_x\phi)^2 \qquad\qquad L(t) \approx \frac{1}{2}\sum_{j=-\infty}^{\infty}\varepsilon\left(\frac{dq_j}{dt}\right)^2 - \frac{(q_{j+1}-q_j)^2}{\varepsilon}$$

as an approximation of the Lagrangian $L(t)$. With $p_j(t) = \varepsilon\cdot\frac{dq_j(t)}{dt} \equiv \pi(t,x_j)\,\delta x_j$ and using $\delta x_j = \varepsilon$ one gets

$$q_j(t) \to \phi(t,x) \qquad\qquad p_j(t) \to \dot\phi(t,x)$$

for $\varepsilon \to 0$.

For a discrete system the Hamiltonian is $H = \sum_j p_j^a\,\dot q_j^a - L$. For the discretized system in $\mathbb{M}_{1,1}$ the Hamiltonian can be written as

$$H = \sum_j p_j^a\,\dot q_j^a - L = \sum_j \delta x_j\left\{\pi_a(t,x_j)\,\dot\phi_a(t,x_j) - \mathcal{L}_j\right\}$$

and $\varepsilon \to 0$ allows to write down the proposal

$$H = \int d^3\vec{x}\,\mathcal{H}(X) \qquad\qquad \mathcal{H}(X) = \pi_a(t,\vec{x})\,\dot\phi_a(t,\vec{x}) - \mathcal{L}(\phi_a,\partial_\mu\phi_a) \tag{2.6}$$

for the Hamiltonian H and the Hamiltonian density $\mathcal{H}$ of a classical field ϕ. As an example one gets

$$\mathcal{H}(X) = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 \tag{2.7}$$

for the Klein-Gordon field.

2.5 Symmetries and Noether's Theorem

A symmetry is some transformation of the variables that leaves the equations of motion invariant. Suppose that $\mathcal{L}(\phi_a,\partial_\mu\phi_a)$ is the Lagrangian density for some fields ϕ_a and that $\phi_a(X) \to \phi'_a(X) = \phi_a(X) + Y_a[\phi_a]$ is an infinitesimal transformation then this is called a symmetry when, in words, the Lagrangian density in the new coordinates equals the Lagrangian density in the old coordinates plus optionally some total derivative, or formally, $\mathcal{L}(\phi'_a,\partial_\mu\phi'_a) = \mathcal{L}(\phi_a,\partial_\mu\phi_a) + \partial_\mu F^\mu$. The total derivative is allowed because the action $S[\phi_a] \equiv \int d^4X\,\mathcal{L} = (\int d^4X(\mathcal{L} + \partial_\mu F^\mu)$ does not change.

A fundamental theorem here is Noether's theorem with very powerful consequences for all of modern physics. It states that every continuous symmetry of a Lagrangian implies the existence of a conserved current $j^\mu(X)$ where $\partial_\mu j^\mu = 0$. To prove it let $\delta\phi_a$ be an arbitrary infinitesimal change in the fields such that

$$\begin{aligned}\delta\mathcal{L}\big(\phi_a,\partial_\mu\phi_a\big) &= \mathcal{L}\big(\phi_a+\delta\phi_a,\partial_\mu(\phi_a+\delta\phi_a)\big) - \mathcal{L}\big(\phi_a,\partial_\mu\phi_a\big) \approx \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta(\partial_\mu\phi_a)\\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right)\end{aligned}$$

is the change of the Lagrangian density up to first order where the last step assumes that ϕ_a obeys the Euler-Lagrange equations (2.2). With $\delta\phi_a = Y_a[\phi_a]$ for an infinitesimal symmetry, the change of the Lagrangian density can only be a total derivative $\delta\mathcal{L} = \partial_\mu F^\mu$ leading to

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}Y_a[\phi_a] - F^\mu\right) = 0 \qquad\qquad j^\mu(X) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}Y_a[\phi_a] - F^\mu \tag{2.8}$$

where $j^\mu(X)$ defines the conserved current density.

The conserved current $\partial_\mu j^\mu$ implies the existence of a conserved charge. Because $\partial_\mu j^\mu = 0$ is true locally everywhere one can define Q_V for any measurable region in space as $Q_V = \int_V d^3\vec{x}\, j^0(\vec{x})$. The calculation

$$\frac{dQ_V}{dt} = \int_V d^3\vec{x}\, \partial_0 j^0 = -\int_V d^3\vec{x}\, \partial_k j^k = -\int_{\partial V} \vec{j}\cdot d\vec{s}$$

shows that the charge does not change in time. Assuming that $V = \mathbb{R}^3$ (and that $j \to 0$ as $|\vec{x}| \to \infty$)

$$\frac{dQ_V}{dt} = 0$$

follows. Therefore, there is a conserved charge for every infinitesimal symmetry. Since quantization is a form of educated guessing, knowing the symmetries is a great way to guess quantum theories.

2.6 Spacetime Translations as Symmetries

There are four spacetime translations, three in spatial directions and one in time. The active translation is $X^\mu \to X^\mu - \varepsilon^\mu$ where the field is moved by ε^μ. Thus, the new field $\phi'_a(X^\mu)$ and the old field $\phi_a(X^\mu)$ are related by $\phi'_a(X^\mu) = \phi_a(X^\mu + \varepsilon^\mu)$. This infinitesimal transformation is in first order

$$\phi'_a(X^\mu) = \phi_a(X^\mu + \varepsilon^\mu) \approx \phi_a(X^\mu) + \varepsilon^\mu \partial_\mu \phi_a(X^\mu)$$

and $\varepsilon^\mu \partial_\mu \phi_a(X^\mu)$ corresponds to the above $Y_a[\phi_a]$. The Lagrangian density $\mathcal{L}(\phi_a, \partial_\mu \phi_a)$ can be viewed as a function $\mathcal{L}(X)$ with the argument X such that $\mathcal{L}(X') = \mathcal{L}(X + \varepsilon)$ and

$$\mathcal{L}(X^\mu + \varepsilon^\mu) \approx \mathcal{L}(X^\mu) + \varepsilon^\nu \partial_\nu \mathcal{L}(X^\mu)$$

in first order. As shown above $\varepsilon^\mu \partial_\mu \mathcal{L}(X^\mu)$ must be a total derivative $\partial_\mu F^\mu$.

Applying Noether's theorem to $\varepsilon^\mu \in \varepsilon\{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\} = \varepsilon\{\hat{V}^0, \hat{V}^1, \hat{V}^2, \hat{V}^3\}$ (or simply $\hat{V}^\mu{}_\nu = \delta^\mu{}_\nu$) gives

$$j^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - F^\mu{}_\nu \equiv T^\mu{}_\nu \qquad\qquad T^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta^\mu{}_\nu \mathcal{L} \tag{2.9}$$

for $\hat{V}^\nu$ according to (2.8) where $F^\mu{}_\nu$ turns out to be $F^\mu{}_\nu = \delta^\mu{}_\nu \mathcal{L}$. The quantity $T^\mu{}_\nu$ containing sixteen values is called the energy-momentum tensor and represents four conserved currents

$$\partial_\mu T^\mu{}_\nu = 0 \tag{2.10}$$

packed together.

Corresponding to these conserved currents there are the four conserved charges

$$E = \int d^3\vec{x}\; T^{00} \qquad\qquad p^j = \int d^3\vec{x}\; T^{0j} \tag{2.11}$$

where $T^{\mu\nu} = \eta^{\nu\rho} T^\mu{}_\rho$. They are energy E and momentum $\vec{p}$, and they are called the generators of translations in time and space, respectively. If the equations of motion are invariant under shifts in time, the energy remains invariant. Thus, the quantity that does not change in a closed system as the time passes is the total energy. If a system is invariant under a shift along the x-axis, the quantity that does not change is the momentum along the x-axis.

As an example the Klein-Gordon Lagrangian density (2.3) is used. The energy-momentum tensor is

$$T^\mu{}_\nu = \partial^\mu \phi\, \partial_\nu \phi - \delta^\mu{}_\nu \left(\frac{1}{2}(\partial_\rho \phi)(\partial^\rho \phi) - \frac{1}{2} m^2 \phi^2\right) \qquad T^{\mu\nu} = \partial^\mu \phi\, \partial^\nu \phi - \eta^{\mu\nu} \left(\frac{1}{2}(\partial_\rho \phi)(\partial^\rho \phi) - \frac{1}{2} m^2 \phi^2\right)$$

and the conserved charges can be calculated. The energy and momenta become

$$E = \int d^3\vec{x} \left(\frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2} m^2 \phi^2\right) = \int d^3\vec{x}\; \mathcal{H}(\vec{x}) \qquad\qquad p^j = \int d^3\vec{x}\, \dot{\phi} \partial^j \phi$$

where $\mathcal{H}(\vec{x})$ is the Hamiltonian density (2.6).

2.7 Lorentz Transformations as Symmetries

Lorentz transformations in the usual form $X^\mu \to \Lambda^\mu_{\ \nu} X^\nu$ are not infinitesimal. An infinitesimal Lorentz transformation can be written in the form $\Lambda^\mu_{\ \nu} = \delta^\mu_{\ \nu} + \omega^\mu_{\ \nu}$ where $\delta^\mu_{\ \nu}$ is the identity and $\omega^\mu_{\ \nu}$ is infinitesimal. Because of $\Lambda^T \eta \Lambda = \eta$ given in (1.4) as the defining property of a Lorentz transformation, $\omega^\mu_{\ \nu}$ is constrained by $(\delta^\mu_{\ \sigma} + \omega^\mu_{\ \sigma})(\delta^\nu_{\ \tau} + \omega^\nu_{\ \tau})\eta^{\sigma\tau} = \eta^{\mu\nu}$. This means $\omega^{\mu\nu} + \omega^{\nu\mu} = 0$ up to first order in ω such that ω contains six independent variables as needed for the three rotations and three boosts. Thus, there must be six conserved currents and six conserved charges and they all come from the energy-momentum tensor. Because the energy-momentum tensor is the response of the field to a deformation of spacetime this is comprehensible.

A field ϕ transforms as

$$\phi_a(X) \to \phi'_a(X) = \phi_a(\Lambda^{-1}X) = \phi_a(X^\mu_{\ \nu} - \omega^\mu_{\ \nu} X^\nu) \approx \phi_a(X) - \omega^\mu_{\ \nu} X^\nu\, \partial_\mu \phi_a(X)$$

up to first order in ω. This gives

$$\delta\phi_a = -\omega^\mu_{\ \nu} X^\nu \partial_\mu \phi_a = Y_a[\phi_a] \qquad\qquad \delta\mathcal{L} = -\omega^\mu_{\ \nu} X^\nu \partial_\mu \mathcal{L} = -\partial_\mu(\omega^\mu_{\ \nu} X^\nu \mathcal{L})$$

because $\omega^{\mu\nu}$ is antisymmetric.

The conserved currents due to Noether's theorem are

$$j^\mu_\omega = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \omega^\rho_{\ \nu}\, X^\nu\, \partial_\rho \phi_a + \omega^\mu_{\ \nu}\, X^\nu\, \mathcal{L}$$

and there are six of them. They are called – in a different notation – $(\mathcal{J}^\mu)^{\rho\sigma}$ where the $\rho\sigma$ entry is the one that is non-zero, and they are

$$(\mathcal{J}^\mu)^{\rho\sigma} = X^\rho\, T^{\mu\sigma} - X^\sigma\, T^{\mu\rho} \qquad\qquad \partial_\mu (\mathcal{J}^\mu)^{\rho\sigma} = 0 \tag{2.12}$$

with their conservation law expressed in terms of the energy-momentum tensor $T^{\mu\nu}$ defined in (2.9). The corresponding conserved charges are

$$Q^{jk} = \int d^3\vec{x}\, \left(X^j\, T^{0k} - X^k\, T^{0j}\right) \qquad\qquad Q^{0j} = \int d^3\vec{x}\, \left(X^0\, T^{0j} - X^j\, T^{00}\right) \tag{2.13}$$

where Q^{jk} are called the generators of rotations and Q^{0j} are called the generators of boosts.

Given the pairs (q_j, p_j) of canonical coordinates in Hamiltonian mechanics for $j = 1, 2, ...$ and two functions $f, g : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ in phase space, then

$$\{f, g\} = \sum_{j=1}^{N} \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} \tag{2.14}$$

defines the Poisson brackets. The corresponding definition is

$$\{F, G\} = \int d^3\vec{x}\, \frac{\partial F}{\partial \phi(X)}\frac{\partial G}{\partial \pi(X)} - \frac{\partial G}{\partial \pi(X)}\frac{\partial G}{\partial \phi(X)} \tag{2.15}$$

for two functionals F and G of classical field theory.

It turns out that $\frac{df}{dt} = \{f, H\}$. Replacing H with a conserved charge $Q^{\rho\sigma}$ gives

$$\frac{\partial f}{\partial s^{\rho\sigma}} = \{f, Q^{\rho\sigma}\}$$

generates corresponding symmetry transformation. Putting in one of the momenta, for example, then $\frac{\partial f}{\partial s^j} = \{f, p^j\}$ generates a shift in phase space. That is why these conserved charges are called the generators of transformations. Further, the conserved charges $Q^{\mu\nu}$ obey the Lie algebra of the Poincaré group. Thus, one can start from a continuous symmetry, obtain the conserved charges from Noether's theorem, and then one can rebuild the symmetry from the conserved charges.

As quantization is an educated guess, one can never be sure that the resulting quantum theory is correct but the above results help. The standard dogma of quantization states: Functions in phase space become observables (linear operators on a Hilbert space), and Poisson brackets become commutation relations.

3 The Free Quantum Scalar Field

3.1 Canonical Quantization of Discrete Systems

Suppose there are n (classical) degrees of freedom, and the position and momentum variables are q_j and p_j, respectively, for $j = 1, ..., n$. There is a standard Poisson bracket $\{q_j, p_k\} = \delta_{jk}$ with the Kronecker delta, and the Hamiltonian is

$$H = \sum_{j=1}^{n} \frac{p_j^2}{2m} + \frac{1}{2}\, m \sum_{j,k=1}^{n} q_j\, Q_{jk}\, q_k$$

with a symmetric matrix $\underline{\underline{Q}}$. It represents a set of interacting harmonic oscillators.

The question is what is the right quantum system that has this system as its classical limit, and there is not a unique answer because at least two completely different quantum system with this Hamiltonian as classical limit exist where one is comprised of a set of quantum spins, and the other is comprised of a set of harmonic oscillators. Using canonical quantization, (q_j, p_j) are promoted to self-adjoint operators $(\boldsymbol{q}_j, \boldsymbol{p}_j)$ representing observables, and the Poisson bracket $\{q_j, p_k\}$ becomes the commutation relation $[\boldsymbol{q}_j, \boldsymbol{p}_k] = i\,\delta_{jk}$. The Hamiltonian is the trickiest part because the operators no longer commute with each other. In the case of the above Hamiltonian, H is simply turned into

$$\boldsymbol{H} = \sum_{j=1}^{n} \frac{\boldsymbol{p}_j^2}{2m} + \frac{1}{2}\, m \sum_{j,k=1}^{n} \boldsymbol{q}_j\, Q_{jk}\, \boldsymbol{q}_k \tag{3.1}$$

because the problem that operators no longer commute does not occur in this case. Quantizing a field instead of this finite problem follows the same path except that there is now an infinite number of degrees of freedom. Much experience shows that this approach of quantization works. The fact that it is working for fields means that one can give a unitary representation of the Poincaré group.

3.2 Quantum Harmonic Oscillator

Solving the quantum harmonic oscillator (3.1) helps later solving the free scalar quantum field. The matrix $\underline{\underline{Q}}$ is symmetric and one can diagonalize $\underline{\underline{Q}}$ with an orthogonal matrix $\underline{\underline{O}}$ to get $\underline{\underline{O}}\,\underline{\underline{Q}}\,\underline{\underline{O}}^T = \underline{\underline{D}}$. Further, one can shift the zero of energy to make all the diagonal elements of $\underline{\underline{D}}$ positive. This leads to new operators $\boldsymbol{q}'_j = \sum_k O_{jk}\, \boldsymbol{q}_k$ as well as, in order to keep the canonical commutation relations, $\boldsymbol{p}'_j = \sum_k O_{jk}\, \boldsymbol{p}_k$. The Hamiltonian operator becomes

$$\boldsymbol{H} = \sum_{j=1}^{n} \frac{\boldsymbol{p'}_j^2}{2m} + \frac{1}{2}\, m \sum_{j=1}^{n} \omega_j^2\, \boldsymbol{q'}_j^2 = \sum_{j=1}^{n} \left(\frac{\boldsymbol{p'}_j^2}{2m} + \frac{1}{2}\, m\, \omega_j^2\, \boldsymbol{q'}_j^2 \right)$$

where ω_j^2 are the eigenvalues of matrix $\underline{\underline{D}}$. This is the sum of n independent harmonic oscillators.

The Hamiltonian $\boldsymbol{H}$ becomes

$$\begin{aligned} \boldsymbol{a}_j &= \sqrt{\frac{m\omega}{2}} \left(\boldsymbol{q}_j + \frac{i}{m\omega}\, \boldsymbol{p}_j \right) \quad \boldsymbol{a}_j^\dagger = \sqrt{\frac{m\omega}{2}} \left(\boldsymbol{q}_j - \frac{i}{m\omega}\, \boldsymbol{p}_j \right) \\ \boldsymbol{H} &= \sum_{j=1}^{n} \omega_j \left(\boldsymbol{a}_j^\dagger \boldsymbol{a}_j + \frac{1}{2} \right) \end{aligned} \tag{3.2}$$

expressed with using the ladder operators $\boldsymbol{a}_j$ and $\boldsymbol{a}_j^\dagger$. The creation or raising operator $\boldsymbol{a}_j^\dagger$ increases the energy level by 1 and the annihilation or lowering operator $\boldsymbol{a}_j$ decreases the energy level by 1 where the Hamiltonian with its eigenvalues (energy levels) are

$$\boldsymbol{H}_j = \omega_j \left(\boldsymbol{a}^\dagger \boldsymbol{a} + \frac{1}{2} \right) \qquad\qquad E_{j,n} = \omega_j \left(n + \frac{1}{2} \right)$$

for the j^{th} harmonic oscillator. The lowest energy level is $E_{j,0} = 0.5\,\omega_j$.

3.3 Canonical Quantization of Fields

The canonical variables q_j and p_j in the discrete case correspond to $\phi(X)$ and $\pi(X)$ for fields and are therefore promoted to $\boldsymbol{\phi}(X)$ and $\boldsymbol{\pi}(X)$. The equal time Poisson bracket $\{\phi(t,\vec{x}),\pi(t,\vec{y})\} = \delta^{(3)}(\vec{x}-\vec{y})$ where the Kronecker delta has become a delta function gets turned into $[\boldsymbol{\phi}(\vec{x}),\boldsymbol{\pi}(\vec{y})] = i\,\delta^{(3)}(\vec{x}-\vec{y})$ as commutation relations. Finally, the classical Klein-Gordon Hamiltonian (2.7) gets turned into

$$\boldsymbol{H} = \frac{1}{2}\int d^3\vec{x}\;\boldsymbol{\pi}(\vec{x})^2 + \big(\nabla\boldsymbol{\phi}(\vec{x})\big)^2 + m^2\,\boldsymbol{\phi}(\vec{x})^2 \tag{3.3}$$

as its quantum version.

Thus, the discrete label j becomes the continuous label X when going from the discrete system of the harmonic oscillators to the Klein-Gordon field. For the diagonalization $\boldsymbol{q}'_j = \sum_k O_{jk}\,\boldsymbol{q}_k$ of the matrix $\underline{\underline{M}}$ in the discrete case the analog is $\boldsymbol{\phi}(\vec{x}) = \int d^3\vec{y}\,K(\vec{x},\vec{y})\,\boldsymbol{\phi}(\vec{y})$ in the case of the quantum field. The kernel function $K(\vec{x},\vec{y})$ diagonalizes what comes from the term $(\nabla\boldsymbol{\phi}(X))^2$ in the Hamiltonian, and the Fourier transformation is the way to diagonalize this Hamiltonian.

If a matrix is circulant such that the entry jk only depends on the difference $j-k$ then it is diagonalized by the discrete Fourier transform. Any circulant matrix of the form $\underline{\underline{M}}$ is diagonalized by the matrix $\underline{\underline{U}}$

$$\underline{\underline{M}} = \begin{pmatrix} c & d & e & ... \\ b & c & d & ... \\ a & b & c & ... \\ ... & ... & ... & ... \end{pmatrix} \qquad \underline{\underline{U}} = \frac{1}{\sqrt{n}}\begin{pmatrix} \mu^0 & \mu^0 & \mu^0 & ... \\ \mu^0 & \mu^1 & \mu^2 & ... \\ \mu^0 & \mu^2 & \mu^4 & ... \\ ... & ... & ... & \mu^{jk} \end{pmatrix} \qquad \mu = e^{i\,\pi/n}$$

where the rows j and columns k of $\underline{\underline{U}}$ are counted from 0 to $n-1$. Note that matrix $\underline{\underline{U}}$ is unitary but not orthogonal. It has complex entries and one has to worry about real and imaginary parts. The matrix $\underline{\underline{Q}}$ in the system of multiple harmonic oscillators above is circulant. This makes clear why the Fourier transformation is used to diagonalize the kernel function K in the continuous case of the field.

With

$$O_{jk} = \frac{1}{\sqrt{n}} e^{\frac{2p\,i}{n}\,jk} \qquad\qquad K(\vec{x},\vec{y}) = \frac{e^{i\,\vec{x}\cdot\vec{y}}}{2\pi}$$

one can guess that the Klein-Gordon Hamiltonian is diagonalized by

$$\boldsymbol{\phi}(\vec{x}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\,e^{i\,\vec{p}\cdot\vec{x}}\,\hat{\boldsymbol{\phi}}(\vec{p}) \tag{3.4}$$

where $\hat{\boldsymbol{\phi}}(\vec{p})$ represents the same field as $\boldsymbol{\phi}(\vec{x})$ but in momentum space. Note that $\hat{\boldsymbol{\phi}}^\dagger(\vec{p}) = \hat{\boldsymbol{\phi}}(-\vec{p})$ and $\hat{\boldsymbol{\phi}}(\vec{p})$ is therefore not Hermitian. The Klein-Gordon Hamiltonian (3.3) is indeed diagonalized by (3.4) as can be shown by introducing implicitly operators $\boldsymbol{a}_{\vec{p}}$ and $\boldsymbol{a}^\dagger_{\vec{p}}$

$$\boldsymbol{\phi}(\vec{x}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{\omega_{\vec{p}}}}\left(\boldsymbol{a}_{\vec{p}}\,e^{i\,\vec{p}\cdot\vec{x}} + \boldsymbol{a}^\dagger_{\vec{p}}\,e^{-i\,\vec{p}\cdot\vec{x}}\right) \quad \boldsymbol{\pi}(\vec{x}) = \int \frac{d^3\vec{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\vec{p}}}{2}}\left(\boldsymbol{a}_{\vec{p}}\,e^{i\,\vec{p}\cdot\vec{x}} - \boldsymbol{a}^\dagger_{\vec{p}}\,e^{-i\,\vec{p}\cdot\vec{x}}\right) \tag{3.5}$$

where $\omega_{\vec{p}} = \sqrt{|\vec{p}|^2 + m^2}$. These operators correspond to the ladder operators defined in (3.2) for the discrete system of harmonic oscillators. Inserting (3.5) into the commutation relations should give

$$\big[\boldsymbol{\phi}(\vec{x}),\boldsymbol{\pi}(\vec{y})\big] = \int \frac{d^3\vec{p}\,d^3\vec{q}}{(2\pi)^6}\,\frac{(-i)}{2}\,\sqrt{\frac{\omega_{\vec{q}}}{\omega_{\vec{p}}}}\left(\big[\boldsymbol{a}^\dagger_{\vec{p}},\boldsymbol{a}_{\vec{q}}\big] - \big[\boldsymbol{a}_{\vec{p}},\boldsymbol{a}^\dagger_{\vec{q}}\big]\right)e^{-i(\vec{p}\cdot\vec{x}+\vec{q}\cdot\vec{y})} = i\,\delta^{(3)}(\vec{x}-\vec{y}) \tag{3.6}$$

at equal time and this is indeed the case if

$$\big[\boldsymbol{a}_{\vec{p}},\boldsymbol{a}^\dagger_{\vec{q}}\big] = (2\pi)^3\,\delta^{(3)}(\vec{p}-\vec{q})\,\mathbf{1} \tag{3.7}$$

are the commutation relations of the creation operator $\boldsymbol{a}^\dagger_{\vec{p}}$ and the annihilation operator $\boldsymbol{a}_{\vec{p}}$. The operator $\mathbf{1}$ is the identity operator.

Substituting (3.5) into the Klein-Gordon Hamiltonian (3.3) gives

$$\boldsymbol{H} = \int d^3\vec{x} \int \frac{d^3\vec{p}\, d^3\vec{q}}{(2\pi)^6}\, e^{-i(\vec{p}+\vec{q})\cdot\vec{x}} \left\{ \frac{-\sqrt{\omega_{\vec{p}}\,\omega_{\vec{q}}}}{4} \left(\boldsymbol{a}_{\vec{p}} - \boldsymbol{a}^\dagger_{-\vec{p}}\right)\left(\boldsymbol{a}_{\vec{q}} - \boldsymbol{a}^\dagger_{-\vec{q}}\right) + \right.$$
$$\left. \frac{(-\vec{p}\cdot\vec{q}+m^2)}{4\sqrt{\omega_{\vec{p}}\,\omega_{\vec{q}}}} \left(\boldsymbol{a}_{\vec{p}} + \boldsymbol{a}^\dagger_{-\vec{p}}\right)\left(\boldsymbol{a}_{\vec{q}} + \boldsymbol{a}^\dagger_{-\vec{q}}\right) \right\}$$
$$= \int \frac{d^3\vec{p}}{(2\pi)^3}\, \omega_{\vec{p}} \left(\boldsymbol{a}^\dagger_{\vec{p}}\,\boldsymbol{a}_{\vec{p}} + \frac{1}{2}\big[\boldsymbol{a}_{\vec{p}}, \boldsymbol{a}^\dagger_{\vec{p}}\big]\right) = \int \frac{d^3\vec{p}}{(2\pi)^3}\, \omega_{\vec{p}} \left(\boldsymbol{a}^\dagger_{\vec{p}}\,\boldsymbol{a}_{\vec{p}} + \frac{1}{2}(2\pi)^3\, \delta^{(3)}(\vec{0})\, \mathbf{1}\right)$$

and this shows that it is indeed diagonal. However, the result is infinite, and there is no way to get rid of this infinity. Since it is not possible to measure absolute energy shifts (except maybe for the gravitational field) and one can only measure energy differences the infinite shift will never appear in any physical calculation. Thus, one can ignore the delta function times the identity operator, and the Klein-Gordon Hamiltonian has been diagonalized with the $\omega_{\vec{p}}$ as eigenvalues.

3.4 Generators of Spacetime Translations

The quantum Klein-Gordon Hamiltonian (3.3) and the commutation relations (3.6) at equal time have been derived by analogy with the classical Klein-Gordon field. It had been solved by

$$\boldsymbol{H} = \int \frac{d^3\vec{p}}{(2\pi)^3}\, \omega_{\vec{p}}\, \boldsymbol{a}^\dagger_{\vec{p}}\, \boldsymbol{a}_{\vec{p}} \tag{3.8}$$

ignoring the infinite part coming from $\frac{1}{2}(2\pi)^3\, \delta^{(3)}(\vec{0})\, \mathbf{1}$. The creation and annihilation operators obey the canonical commutation relations (3.7).

The question remains whether the field defined by the Klein-Gordon Hamiltonian is a relativistic quantum field theory. There is still much that can go wrong. A projective unitary representation of the Poincaré group is needed. The solution so far covers certainly the one-parameter subgroup of the Poincaré group

$$\boldsymbol{U}\big((t,0,0,0)\big) \equiv e^{-i\,t\,\boldsymbol{H}}$$

consisting of the translations in time. Thus, the goal is to find generators for all possible transformations of the Poincaré group. To do so one turns all conserved charges given by Noether's theorem into operators and determines their commutation relations. The conserved charge for $X^\mu \to X^\mu + (A^0,0,0,0)$ is the Hamiltonian $H = \int T^{00}\, d^3\vec{x} = \int \mathcal{H}\, d^3\vec{x}$, and the quantum version is just $\boldsymbol{H}$ as shown above.

Also for the generators of the spatial translations is the classical energy-momentum tensor (2.9)

$$T^\mu{}_\nu = \partial^\mu\phi\, \partial_\nu\phi - \delta^\mu{}_\nu \left(\partial_\rho\phi\, \partial^\rho\phi - \frac{1}{2}m^2\phi^2\right)$$

needed where $p^j = \int T^{0j}\, d^3\vec{x} = \int \dot{\phi}\, \partial_j\phi\, d^3\vec{x} = \int \pi\, \partial_j\phi\, d^3\vec{x}$. Quantizing this naively to $\int \boldsymbol{\pi}\, \partial_j\boldsymbol{\phi}\, d^3\vec{x}$ could cause a problem because the order of the two operators matters. However, this version works. This is

$$\boldsymbol{p}^j = \int \boldsymbol{\pi}\, \partial_j\boldsymbol{\phi}\, d^3\vec{x} = \int \frac{d^3\vec{p}}{(2\pi)^3}\, p^j\, \boldsymbol{a}^\dagger_{\vec{p}}\, \boldsymbol{a}_{\vec{p}} \tag{3.9}$$

in terms of creation and annihilation operators. The classical Poisson brackets $\{Q_\alpha, Q_\beta\} = f_{\alpha\beta}{}^\gamma\, Q_\gamma$ should become the canonical commutation relations $[\boldsymbol{Q}_\alpha, \boldsymbol{Q}_\beta] = i\, f_{\alpha\beta}{}^\gamma\, \boldsymbol{Q}_\gamma$ through quantization. Thus, the question is whether the operators $\boldsymbol{H}$ and $\boldsymbol{p}^j$ from (3.8) and (3.9) obey the right algebra. The Poisson brackets give $\{p^\mu, p^\nu\} = 0$ and also $[\boldsymbol{Q}_\alpha, \boldsymbol{Q}_\beta] = 0$ is satisfied.

The result is a projective unitary representation of the translation subgroup of the Poincaré group, and the next goal is to explicitly produce this projective unitary representation. A unitary representation is the same as finding a Hilbert space upon which these operators $\boldsymbol{H}$ in (3.8) and $\boldsymbol{p}^j$ in (3.9) act as generators of unitaries.

3.5 Fock Space

One starts with the vacuum state $|\Omega\rangle$ defined by $\boldsymbol{a}_{\vec{p}}\,|\Omega\rangle = 0$ for all $\vec{p}$. The Hilbert space is built as the space generated by all the finite linear combinations of vectors of the form

$$|\vec{p}_1, \vec{p}_2, ..., \vec{p}_n\rangle = \boldsymbol{a}^\dagger_{\vec{p}_1}\boldsymbol{a}^\dagger_{\vec{p}_2}...\boldsymbol{a}^\dagger_{\vec{p}_n}\,|\Omega\rangle \tag{3.10}$$

completed with respect to the inner product. Because the commutation relations between creation and annihilation operators (3.7) are delta functions, these states are not normalizable, and the inner product cannot be defined easily. These states can never be prepared in the laboratory with exact momenta, and the solution is to smear these improper states out a bit such that they become proper physical states

$$|\psi\rangle \equiv \int \frac{d^3\vec{p}}{(2\pi)^3}\,\psi(\vec{p})\;|\vec{p}\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3}\,\psi(\vec{p})\,\boldsymbol{a}^\dagger_{\vec{p}}\,|\Omega\rangle \tag{3.11}$$

where ψ is a smooth (or L^2) function.

There is also a problem with this approach because the smearing process is not obviously Lorentz invariant. The difficulty comes from the normalization of $|\vec{p}\rangle$ because

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

is not Lorentz invariant. However, there is a neat way to come up with other improper vectors which are Lorentz invariant. One can build a projector onto a single-particle space by

$$\mathbf{1}_{\mathrm{S}} \equiv \int \frac{d^3\vec{p}}{(2\pi)^3}\;|\vec{p}\rangle\,\langle\vec{p}|$$

because the Hilbert space spanned by all single particle vectors is a subspace, and independent of the reference frame, observers agree on whether a vector is either in this subspace or not. Whether there is a single particle or not is an invariant concept. The two components on the other hand

$$\int \frac{d^3\vec{p}}{(2\pi)^3} \qquad\qquad |\vec{p}\rangle\,\langle\vec{p}|$$

alone are both not Lorentz invariant. The trick is to use coordinates where both components are Lorentz invariant.

Obviously, $\int d^4P$ is invariant under Poincaré transformations, but P is constrained by the mass hyperboloid $P^\mu P_\mu = m^2$ meaning $P_0^2 = \omega_{\vec{p}}^2 = |\vec{p}|^2 + m^2$. With $P_0 > 0$

$$\int d^4P\,\delta\big(P_0^2 - |\vec{p}|^2 - m^2\big)\Big|_{P_0>0} = \int \frac{d^3\vec{p}}{2P_0}\Big|_{P_0=\omega_{\vec{p}}}$$

is invariant. Thus, the delta normalization for 3-vectors is defined as $2\omega_{\vec{p}}\,\delta^{(3)}(\vec{p}-\vec{q})$, and the normalized single-particle state $|P\rangle$ as well as the inner product become

$$|P\rangle \equiv \sqrt{2\omega_{\vec{p}}}\;|\vec{p}\rangle = \sqrt{2\omega_{\vec{p}}}\;\boldsymbol{a}^\dagger_{\vec{p}}\,|\vec{\Omega}\rangle \qquad\qquad \langle P|Q\rangle = (2\pi)^3\,2\omega_{\vec{p}}\;\delta^{(3)}(\vec{p}-\vec{q}) \tag{3.12}$$

expressed in 3-vectors. These definitions are independent of the reference frame.

3.6 Translation Invariance in the Schrödinger and the Heisenberg Picture

To determine the action of $\boldsymbol{P}^\mu$ on $|P_1, P_2, ..., P_n\rangle$ the commutation relations

$$[\boldsymbol{H}, \boldsymbol{a}_{\vec{p}}] = -\omega_{\vec{p}}\,\boldsymbol{a}_{\vec{p}} \qquad\qquad [\boldsymbol{P}^j, \boldsymbol{a}_{\vec{p}}] = P^j\,\boldsymbol{a}_{\vec{p}} \tag{3.13}$$

and the corollary

$$\boldsymbol{P}^\mu\,|P_1, P_2, ..., P_n\rangle = \left(\sum_{j=1}^{n} P_j^\mu\right)|P_1, P_2, ..., P_n\rangle$$

where $P_0 = \omega_{\vec{p}}$ are needed. This follows from $\boldsymbol{P}^\mu |\Omega\rangle = 0$. As a consequence, a translation $X \to X + A$ is implemented by

$$\boldsymbol{U}(A) \equiv e^{-i\,A_\mu\,\boldsymbol{P}^\mu} \tag{3.14}$$

quantumly. This is in the Schrödinger picture.

The Heisenberg picture where the operators change according to

$$\boldsymbol{O}_{\rm H} \equiv e^{i\,\boldsymbol{H}\,t}\,\boldsymbol{O}_{\rm S}\,e^{-i\,\boldsymbol{H}\,t} \qquad\qquad \frac{d\boldsymbol{O}_{\rm H}}{dt} = i[\boldsymbol{H}, \boldsymbol{O}_{\rm H}]$$

becomes especially useful when doing perturbation theory for quantum fields. The commutation relations are

$$[\boldsymbol{\phi}(t,\vec{x}), \boldsymbol{\phi}(t,\vec{y})] = [\boldsymbol{\pi}(t,\vec{x}), \boldsymbol{\pi}(t,\vec{y})] = 0 \qquad\qquad [\boldsymbol{\phi}(t,\vec{x}), \boldsymbol{\pi}(t,\vec{y})] = i\,\delta^{(3)}(\vec{x}-\vec{y}) \tag{3.15}$$

in the Heisenberg picture where it is easy to describe spatially local processes because the field operator $\boldsymbol{\phi}(X) = \boldsymbol{\phi}(t,\vec{x})$ is defined at a point in spacetime. The equations of motion are

$$\begin{aligned}
\frac{d\boldsymbol{\phi}(t,\vec{x})}{dt} &= i\,[\boldsymbol{H}, \boldsymbol{\phi}(t,\vec{x})] = \frac{i}{2}\,\Big[\int d^3\vec{y}\;\boldsymbol{\pi}(t,\vec{y})^2 + \big(\nabla\boldsymbol{\phi}(t,\vec{y})\big)^2 + m^2\,\boldsymbol{\phi}(t,\vec{y})^2, \boldsymbol{\phi}(t,\vec{x})\Big] \\
&= \boldsymbol{\pi}(t,\vec{x}) \\
\frac{d\boldsymbol{\pi}(t,\vec{x})}{dt} &= i\,[\boldsymbol{H}, \boldsymbol{\pi}(t,\vec{x})] = \frac{i}{2}\,\Big[\int d^3\vec{y}\;\boldsymbol{\pi}(t,\vec{y})^2 + \big(\nabla\boldsymbol{\phi}(t,\vec{y})\big)^2 + m^2\,\boldsymbol{\phi}(t,\vec{y})^2, \boldsymbol{\pi}(t,\vec{x})\Big] \\
&= \frac{i}{2}\int d^3\vec{y}\,\Big\{2i\big(\nabla_j\boldsymbol{\phi}(t,\vec{y})\big)\big(\nabla_j\delta^{(3)}(\vec{x}-\vec{y})\big) + 2m^2\big(i\,\delta^{(3)}(\vec{x}-\vec{y})\,\boldsymbol{\phi}(t,\vec{y})\big)\Big\} \\
&= \nabla^2\boldsymbol{\phi}(t,\vec{x}) - m^2\,\boldsymbol{\phi}(t,\vec{x}) = \frac{d^2\boldsymbol{\phi}(t,\vec{x})}{dt^2}
\end{aligned}$$

where the derivative of $\boldsymbol{\pi}(t,\vec{x})$ is trickier because of the commutator $[(\nabla\boldsymbol{\phi}(t,\vec{y}))^2, \boldsymbol{\pi}(t,\vec{x})]$. The combined result is just the Klein-Gordon equation $(\partial^\mu\partial_\mu + m^2)\,\boldsymbol{\phi}(t,\vec{x}) = 0$.

The Klein-Gordon quantum Hamiltonian (3.3) with the commutation relations (3.15) has been solved by (3.8), and $\boldsymbol{U}(t) = e^{-i\,t\,\boldsymbol{H}}$ states the dymnamics as the solution to the Schrödinger equation. In non-relativistic quantum mechanics this would be all there is because one can put in any initial state and find out what the state is some time later. In relativistic quantum mechanics this is only part of the story because a projective unitary representation of the full Poincaré group is needed. The generators of rotations and boosts are still missing.

3.7 Causality

The spatial nature of the solutions for the Klein-Gordon Hamiltonian is easier to study in the Heisenberg picture where

$$\boldsymbol{\phi}(t,\vec{x}) = e^{i\,\boldsymbol{H}\,t}\,\boldsymbol{\phi}(\vec{x})\,e^{-i\,\boldsymbol{H}\,t} \tag{3.16}$$

shows the spacial field observable translated in time. The question is whether the observable $\boldsymbol{\phi}(t,\vec{x})$ obeys causality. Two spacelike events should not influence each other while for two timelike events correlations are fine. The commutation relation $[\boldsymbol{H}, \boldsymbol{a}_{\vec{p}}] = -\omega_{\vec{p}}\,\boldsymbol{a}_{\vec{p}}$ in (3.13) implies

$$e^{i\,\boldsymbol{H}\,t}\,\boldsymbol{a}_{\vec{p}}\,e^{-i\,\boldsymbol{H}\,t} = e^{-i\,\omega_{\vec{p}}\,t}\,\boldsymbol{a}_{\vec{p}} \tag{3.17}$$

and after substituting (3.16) and (3.17) into $\boldsymbol{\phi}(\vec{x})$ given by (3.5) the field operator in the Heisenberg picture becomes

$$\boldsymbol{\phi}(t,\vec{x}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{\omega_{\vec{p}}}}\Big(\boldsymbol{a}_{\vec{p}}\,e^{-i\,P\cdot X} + \boldsymbol{a}^\dagger_{\vec{p}}\,e^{i\,P\cdot X}\Big)$$

where $P = (\omega_{\vec{p}}, \vec{p})$ and $X = (t,\vec{x})$. The smeared versions of $\boldsymbol{\phi}(X)$ are observables, and it should be possible to attach these observables to some kind of experiments. One can interpret $\boldsymbol{\phi}(X)$ as an observable that probes the field at exactly the spacetime location X although they are not really observables.

It is assumed that one makes an experiment where one interrogates the field at some location $X = (t, \vec{x})$ such that an apparatus has been set up at location $\vec{x}$ in space to measures the field at the exact point in time t. There is a back reaction on the field due to the measurement of the operator $\boldsymbol{\phi}(X)$ as always in quantum mechanics. As time passes this disturbance of the field propagates outwards with the speed of light depicted in figure 2 (a) as one would expect. However, if it is not causal the disturbance would instantaneously propagate throughout the field as illustrated in (b). It turns out that the disturbance does fly instantaneously, and this fact looks like a big problem. Thus, if one measures a quantum field it changes everywhere in the instant of time t where the measurement is performed.

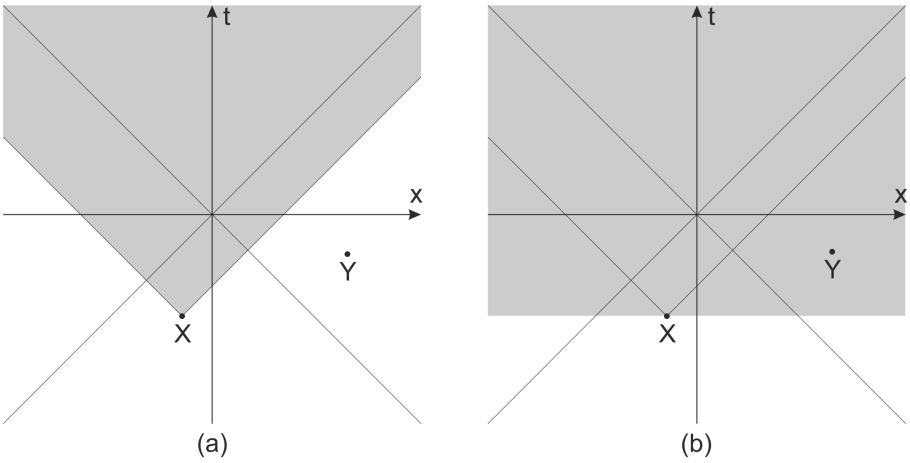
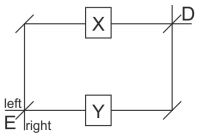


Figure 2: Region of disturbance through a measurement of the field

This fact does not violate the principle of relativity because it is not possible to send information faster than the speed of light. Although the field is disturbed there is no way to figure out at Y what has been done at X if $(X - Y)^2 < 0$. This is not surprising similarly as a situation of entangled states.

There is a natural guess to study the problem of causality by looking at the two-point correlation function $\langle 0 \,|\boldsymbol{\phi}(X)\, \boldsymbol{\phi}(Y)|\, 0\rangle = D(X - Y)$ but this guess is wrong because it has no operational meaning. There is no experiment to measure this function, and $\boldsymbol{\phi}(X)\, \boldsymbol{\phi}(Y)$ is not even Hermitian. However, one can find measurable correlation function coming from interference experiments. The following thoughts are useful and could be made rigorous but are only outlined here.

The setup for the experiment is a Klein-Gordon field with some auxiliary modes of light. There must be something outside of the field in an interference experiment to capture the result of the experiment. The auxiliary modes of light are two lasers. In addition there are two unitary operators



$$\boldsymbol{U}(X) = e^{i\,\varepsilon\,\boldsymbol{\phi}(X)} \qquad\qquad \boldsymbol{U}(Y) = e^{i\,\varepsilon\,\boldsymbol{\phi}(Y)}$$

representing quantum mechanical processes. They are used to gently push the field causing the creation of a particle at position X or Y, respectively. The source of light is marked E. If it takes the upper branch of the interferometer a particle is created at X, and if it takes the lower path a particle at Y is created.

This is an experiment one can imagine, in principle, to be possible. It needs some preparation at E and some detection on the opposite side marked as D. The field is everywhere, and there is one observer at spacetime location X and one at spacetime location Y. One shares some state of light between these two observers and puts this shared auxiliary mode in some superposition state like 1 for the observer at X and 0 for the observer at Y respectively 0 for the observer at X and 1 for the observer at Y. In superposition, they create their particle conditioned on the value of this mode of light. Then one brings these auxiliary modes together and looks at incidences of whether or not there is a particle in the left mode or in the right mode.

The mathematical description of the different steps of this experiment:

1. $|0\rangle_{\text{field}} \quad |0\rangle_{\text{left}} \quad |1\rangle_{\text{right}}$

2. $|0\rangle_{\text{field}}$ $(\frac{1}{\sqrt{2}}(|0\rangle\,|1\rangle + |1\rangle\,|0\rangle))$ by Hadamard/beam splitter
3. Apply: $\boldsymbol{U}(X) \otimes |1\rangle\,\langle 1|_{\text{left}} \otimes \mathbf{1}_{\text{right}} + \boldsymbol{U}(Y) \otimes \mathbf{1}_{\text{left}} \otimes |1\rangle\,\langle 1|_{\text{right}}$
$\frac{1}{\sqrt{2}}(|0\rangle\,|1\rangle + |1\rangle\,|0\rangle)$ evolves to $\frac{1}{\sqrt{2}}\boldsymbol{U}(Y)\,|0\rangle\;|0\,1\rangle + \frac{1}{\sqrt{2}}\boldsymbol{U}(X)\,|0\rangle\;|1\,0\rangle$
4. Apply Hadamard/beam splitter: $(\frac{1}{2}\boldsymbol{U}(X)\,|0\rangle + \frac{1}{2}\boldsymbol{U}(Y)\,|0\rangle)\,|0\,1\rangle + (\frac{1}{2}\boldsymbol{U}(X)\,|0\rangle - \frac{1}{2}\boldsymbol{U}(Y)\,|0\rangle)\,|1\,0\rangle$
(final state)
5. Measure: $P(''01'') = \frac{1}{4}\,\langle 0|(\boldsymbol{U}^\dagger(X) + \boldsymbol{U}^\dagger(Y))(\boldsymbol{U}(X) + \boldsymbol{U}(Y))|0\rangle = \frac{1}{2} + \frac{1}{2}\,\mathrm{Re}\,\langle 0|\boldsymbol{U}^\dagger(X)\boldsymbol{U}(Y)|0\rangle$

Whether X and Y are related spacelike or timelike presumably has some consequences for the interference pattern $\langle 0|\boldsymbol{U}^\dagger(X)\boldsymbol{U}(Y)|0\rangle$. This expression is

$$\langle 0|\boldsymbol{U}^\dagger(X)\boldsymbol{U}(Y)|0\rangle = \langle 0|e^{-i\varepsilon\boldsymbol{\phi}(X)}\,e^{i\varepsilon\boldsymbol{\phi}(Y)}|0\rangle$$

and it is not clear whether $\boldsymbol{\phi}(X)$ and $\boldsymbol{\phi}(Y)$ commute. Because of the Baker-Campbell-Hausdorff formula this expression becomes

$$\langle 0|e^{-i\varepsilon\boldsymbol{\phi}(X)}\,e^{i\varepsilon\boldsymbol{\phi}(Y)}|0\rangle = \langle 0|e^{-i\varepsilon\boldsymbol{\phi}(X)+i\varepsilon\boldsymbol{\phi}(Y)+\frac{\varepsilon^2}{2}[\boldsymbol{\phi}(X),\boldsymbol{\phi}(Y)]}|0\rangle$$

where the equation is exact and there are no higher order terms of the form $[\boldsymbol{\phi}(X),[\boldsymbol{\phi}(X),\boldsymbol{\phi}(Y)]]$ and so on. This has to be proven.

Thus, if X and Y are spacelike one expects $[\boldsymbol{\phi}(X),\boldsymbol{\phi}(Y)] = 0$ such that the observable $\boldsymbol{\phi}(X)$ has no influence on the observable $\boldsymbol{\phi}(Y)$. Calculations show

$$\begin{aligned}[\boldsymbol{\phi}(X),\boldsymbol{\phi}(Y)] &= \int\frac{d^3\vec{p}}{(2\pi)^3}\int\frac{d^3\vec{q}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}\,2\omega_{\vec{q}}}}\left\{[\boldsymbol{a}_{\vec{p}},\boldsymbol{a}^\dagger_{\vec{q}}]\,e^{-i\,P\cdot X+i\,Q\cdot Y} + [\boldsymbol{a}^\dagger_{\vec{p}},\boldsymbol{a}_{\vec{q}}]\,e^{i\,P\cdot X-i\,Q\cdot Y}\right\}\\ &= \int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2\omega_{\vec{p}}}\left(e^{-i\,P\cdot(X-Y)} - e^{i\,P\cdot(X-Y)}\right)\mathbf{1} = \Delta(X-Y)\,\mathbf{1}\end{aligned}$$

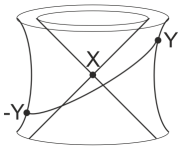
using (3.7), and this function $\Delta(X-Y)$ is Lorentz invariant.

If $(X-Y)^2 > 0$ and therefore X and Y are timelike separated then one can go to a reference frame with $X - Y = (t,0,0,0)$. In this case

$$\begin{aligned}\Delta(X-Y) &= \int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2\omega_{\vec{p}}}\left(e^{-i\,\omega_{\vec{p}}\,t} - e^{i\,\omega_{\vec{p}}\,t}\right)\\ &= \frac{1}{4\pi^2}\int_m^\infty dE\,\sqrt{E^2+m^2}\,e^{-i\,E\,t} \sim e^{-i\,m\,t} - e^{i\,m\,t} \neq 0\end{aligned}$$

with a simple variable substitution.

If $(X-Y)^2 < 0$ and therefore X and Y are spacelike separated then one can go to a reference frame with $X - Y = (0,\vec{x}-\vec{y})$. In this case



$$\begin{aligned}\Delta(X-Y) &= \frac{1}{2}\int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{|\vec{p}|^2+m^2}}\left(e^{i\,\vec{p}\cdot(\vec{x}-\vec{y})} - e^{-i\,\vec{p}\cdot(\vec{x}-\vec{y})}\right)\\ &= \frac{1}{2}\int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{|\vec{p}|^2+m^2}}\,((I)-(II))\end{aligned}$$

where the integral is obviously Lorentz invariant but $(I)-(II)$ is not manifestly Lorentz invariant. With a Lorentz transformation of $(\vec{x}-\vec{y}) \to -(\vec{x}-\vec{y})$ for (II) allowed for spacelike vectors it follows that $\Delta(X-Y) = 0$ as necessary for spacelike separated points.

The Klein-Gordon theory is therefore causal. The two operators $\boldsymbol{\phi}(X)$ and $\boldsymbol{\phi}(Y)$ at two spacelike separated points X and Y in spacetime are independent, and it is not possible to send a signal faster than the speed of light. However, the same calculations should be done for $\boldsymbol{\pi}(X)$ and all the other possible observables of the theory. This has not yet be done. There are two ways to deal with this problem. One is to find a full unitary representation of the Poincaré group and the other is to declare that the field operators $\boldsymbol{\phi}(X)$ are the only observables which matter.

3.8 Propagators

The quantity

$$D(X-Y) = \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} e^{-i\,P\cdot(X-Y)} = \langle 0\,|\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)|\,0\rangle \tag{3.18}$$

is called propagator. If $(X-Y)^2 < 0$ and therefore X and Y are spacelike separated then

$$D(X-Y) = \frac{1}{2}\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{\sqrt{|\vec{p}|^2+m^2}} e^{i\,\vec{p}\cdot(\vec{x}-\vec{y})} = \frac{m}{4\pi^2}\frac{1}{|\vec{x}-\vec{y}|} K_1(m\,|\vec{x}-\vec{y}|) \neq 0 \tag{3.19}$$

in polar coordinates with the Hankel function K_1. This quantity is therefore not measurable because one could send a signal faster than the speed of light, but this quantity is rather small as it is asymptotically $e^{m\,|\vec{x}-\vec{y}|}$.

The Feynman propagator is defined as

$$\begin{aligned}\Delta_{\mathrm{F}}(X-Y) &= \begin{cases} D(X-Y) & \text{if } X^0 > Y^0 \\ D(Y-X) & \text{if } X^0 < Y^0 \end{cases} \\ &= \langle 0\,|\mathcal{T}[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)]|\,0\rangle \end{aligned} \tag{3.20}$$

where $\mathcal{T}$ acts as

$$\langle 0\,|\mathcal{T}[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)]|\,0\rangle = \begin{cases} \boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y) & \text{if } X^0 \geq Y^0 \\ \boldsymbol{\phi}(Y)\,\boldsymbol{\phi}(X) & \text{if } X^0 < Y^0 \end{cases} \tag{3.21}$$

and is called the time ordering operator.

The Feynman propagator is later needed in the form

$$\Delta_{\mathrm{F}}(X-Y) = \int \frac{d^4P}{(2\pi)^4} \frac{i\,e^{-i\,P\cdot(X-Y)}}{P^2-m^2+i\,\varepsilon} = \int \frac{d^3\vec{p}}{(2\pi)^3} \left(\int_{\mathbb{C}} \frac{dP_0}{(2\pi)} \frac{i\,e^{-i\,P\cdot(X-Y)}}{P^2-m^2+i\,\varepsilon} \right) \tag{3.22}$$

where P^0 is interpreted as a complex variable along the contour running along the x-axis. The value ε is infinitesimal. The contour integral in the complex plane over P_0 has the two poles where $P^2-m^2+i\,\varepsilon = 0$ as shown in figure 3. The contour has to be closed.

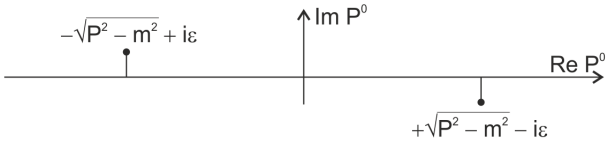


Figure 3: The two poles of contour integral in the complex plane

The Feynman propagator satisfies

$$\left(\partial_0^2 - \nabla^2 + m^2\right) \Delta_{\mathrm{F}}(X-Y) = -i\,\delta^{(4)}(X-Y) \tag{3.23}$$

in the form of a Green's function. This becomes clear in the context of quantum field theory with the path integral. Roughly speaking a Green's function is the inverse of a differential operator. The Feynman propagator is somehow the inverse of the Klein-Gordon equation, and the delta function plays the role of the identity.

3.9 Lie Groups and Lie Algebras

The Poincaré group is a continuous group, and there is a beautiful theory for continuous groups that can here only be given as a physicist's highlights and not as a full mathematical theory as it would deserve.

It is assumed that the group is also a manifold $\mathfrak{M}$ and the group operation is continuous. There is a point in $\mathfrak{M}$ acting as the identity of the group, and the group multiplication is a function $\mathfrak{M} \times \mathfrak{M} \to \mathfrak{M}$ mapping the pair g, h to $g * h$. This mapping and the inverse g^{-1} as a mapping $\mathfrak{M} \to \mathfrak{M}$ are continuous.

The manifold is illustrated in figure 4 as a curved surface with the special element called identity marked as $\mathbb{I}$ in (a). The idea of Lie groups is to study elements very close to the identity $\mathbb{I}$. The tangent space at $\mathbb{I}$ is depicted in (b) as a tangent plane with a small neighborhood. Taylor's theorem ensures that the closer one gets to the identity the better this tangent space is as an approximation for the neighborhood of $\mathbb{I}$.

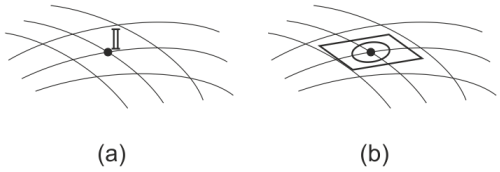


Figure 4: The Lie group as a manifold with the tangent space at the identity element

Thus, to understand Lie groups such as the Poincaré group one studies elements $g \in \mathfrak{M}$ infinitesimally close to the identity element $\mathbb{I}$. In order to talk about $g - \mathbb{I} \approx O(\varepsilon)$ where ε is infinitesimal, a metric is needed which is always obvious in the examples used in the following. The space of such infinitesimal elements is given by the tangent space $T_{\mathbb{I}}\mathfrak{M}$. This is a linear space and is called Lie algebra. It has a dimension and one can define a basis since the Lie algebra is a vector space.

A simple first example of a Lie group is U(1) containing all 1×1 unitary matrices. It consists of all complex numbers with norm 1. The identity is $\mathbb{I} = 1$, and the tangent space is $T_1\mathrm{U}(1) = \{z \,|\, (1+\varepsilon z)^*(1+\varepsilon z) = 1\}$ to $O(\varepsilon)$. Thus, $\mathrm{Re}\, z = 0$ and $T_1\mathrm{U}(1) = \{z = ix \,|\, x \in \mathbb{R}\}$.

The second example is the Lie group O(3) of the rotations in space. O(3) is the set of 3×3 matrices M with $M^T M = \mathbb{I}$. The tangent space is $T_{\mathbb{I}}\mathrm{O}(3) = \{X \,|\, (\mathbb{I} + \varepsilon X)^T(\mathbb{I} + \varepsilon X) = \mathbb{I}\} = \{X \,|\, X + X^T = 0\}$ to $O(\varepsilon)$. Thus, these are the antisymmetric matrices $X^T = -X$. The dimension of $T_{\mathbb{I}}\mathrm{O}(3)$ is 3, and the basis

$$T_{\mathbb{I}}\mathrm{O}(3) = \left\{ X = \sum_{j=1}^{3} x_j J^j \,|\, J^1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, J^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, J^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \right\} \tag{3.24}$$

is a convenient choice.

The third and most important example in the context of quantum field theory is the group G of Lorentz transformations that consists of all 4×4 matrices Λ satisfying $\Lambda^T \eta \Lambda = \eta$ according to (1.4). The group G and its tangent space $T_{\mathbb{I}}G$ are

$$G = \left\{\Lambda \,|\, \eta_{\mu\nu} \Lambda^{\mu}_{\ \rho} \Lambda^{\nu}_{\ \sigma} = \eta_{\rho\sigma}\right\} \qquad T_{\mathbb{I}}G = \left\{\omega \,|\, \eta_{\mu\nu} (\mathbb{I} + \varepsilon\omega)^{\mu}_{\ \rho} (\mathbb{I} + \varepsilon\omega)^{\nu}_{\ \sigma} = \eta_{\rho\sigma}\right\} = \left\{\omega \,|\, \omega^{\mu\nu} = -\omega^{\nu\mu}\right\}$$

to $O(\varepsilon)$. Thus, the tangent space to the Lorentz group is a set of 4×4 antisymmetric matrices with two covariant indices. It can be written in the form

$$T_{\mathbb{I}}G = \left\{ \omega \,|\, \omega = \sum_{0 \le \rho < \sigma \le 3} \frac{1}{2} \Omega_{(\rho\sigma)} J^{(\rho\sigma)} \right\} \qquad \left[J^{(\rho\sigma)}\right]^{\mu}_{\ \nu} = \eta^{\rho\mu} \delta^{\sigma}_{\ \nu} - \eta^{\sigma\mu} \delta^{\rho}_{\ \nu} \tag{3.25}$$

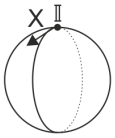
with the basis $J^{(\rho\sigma)}$.

The Lie algebra structure gives a Lie group structure and vice versa. This will only be sketched but not properly proven.

It will be shown that the infinitesimally small linear space near the identity $\mathbb{I}$ tells a lot about the whole Lie group. This comes from the fact that the group multiplication is continuous. This strongly determines the shape of the manifold. Any element of $T_{\mathbb{I}}\mathfrak{M}$ determines an element of the group $\mathfrak{M}$ because the element $\mathbb{I} + \varepsilon X$ built from $X \in T_{\mathbb{I}}\mathfrak{M}$ is a group element in $\mathfrak{M}$ up to $O(\varepsilon)$ and this can be written in the

form $\mathbb{I} + \varepsilon X \approx e^{\varepsilon X} \in \mathfrak{M}$ still up to first order $O(\varepsilon)$. Defining the group element $g(s)$ as the limes

$$g(s) = \lim_{n\to\infty} \left(\mathbb{I} + \frac{s}{n} X\right)^n = e^{sX} \tag{3.26}$$



then $\frac{s}{n}$ becomes ε as n gets bigger and bigger, and $g(s)$ becomes an honest member of the Lie group $\mathfrak{M}$. Thus, every element of the tangent space determines an element of the Lie group. The definition (3.26) maps a vector X in the tangent space to a path along the geodesic of the manifold in the direction of this vector with length s. In other words, every pair of tangent vector X and parameter s gives an element of the manifold.

For sufficiently nice Lie groups that are connected one can go back from the element of the Lie group to the Lie algebra. The logarithm of a group element gives the corresponding element of the Lie algebra. The manifold $\mathfrak{M}$ as a group has a multiplication and the question is what consequences this has for the tangent space. Because with $g, h \in \mathfrak{M}$ also the group commutator $[g, h] \equiv ghg^{-1}h^{-1} \in \mathfrak{M}$, the group commutator is a map $[\ ,\ ] : \mathfrak{M} \times \mathfrak{M} \to \mathfrak{M}$ with $[\mathbb{I}, \mathbb{I}] = \mathbb{I}$. This commutator map of the manifold pushes forward to a map on the tangent space $[\ ,\ ] : T_{\mathbb{I}}\mathfrak{M} \times T_{\mathbb{I}}\mathfrak{M} \to T_{\mathbb{I}}\mathfrak{M}$. Therefore $(\mathbb{I} + \delta X)(\mathbb{I} + \varepsilon Y)(\mathbb{I} - \delta X)(\mathbb{I} - \varepsilon Y) \in \mathfrak{M}$ inserting the first order approximations of the group elements into $ghg^{-1}h^{-1}$. The commutator becomes $[\mathbb{I} + \delta X, \mathbb{I} + \varepsilon Y] = \mathbb{I} + \delta\varepsilon(XY - YX) + O(\delta^2) + O(\varepsilon^2)$. Thus, $[X, Y] = XY - YX$ is also an element of the tangent space $T_{\mathbb{I}}\mathfrak{M}$.

The commutators for the above three examples are as follows. For U(1) it is trivial, and for O(3) the commutator is $[J^j, J^k] = -\varepsilon^{jk}_{\ \ l}\, J^l$. The Lorentz group as the third example has

$$[J^{\mu\nu}, J^{\rho\sigma}] = \eta^{\nu\rho} J^{\mu\sigma} - \eta^{\mu\rho} J^{\nu\sigma} + \eta^{\mu\sigma} J^{\nu\rho} - \eta^{\nu\sigma} J^{\mu\rho} \tag{3.27}$$

as its commutation relations. One case as an illustration is $[J^{01}, J^{12}] = -J^{13}$. Because the Lorentz group is so important, special notations have been introduced such as $K^1 = J^{01}$, $K^2 = J^{02}$, $K^3 = J^{03}$ for the generators of boosts, and $e^{\frac{1}{2}s_j K^j}$ gives a pure boost. Similarly, the generators of rotations are $J^1 = J^{12}$, $J^2 = J^{13}$, $J^3 = J^{23}$. The rotations are a subgroup of the Lorentz group, and they are contained as 3×3 matrices in the 4×4 matrices of the Lorentz group. By exponenting linear combinations of these six generators one can get any Lorentz transformation.

If one finds a representation of the Lie algebra $T_{\mathbb{I}}\mathfrak{M}$ as a mapping π from the tangent space to the bounded linear operators $\mathfrak{B}$ on some Hilbert space $\mathfrak{H}$

$$\pi : T_{\mathbb{I}}\mathfrak{M} \to \mathfrak{B}(\mathfrak{H})$$

such that $[\pi(X), \pi(Y)] = \pi([X, Y])$, one gets a full representation of $\mathfrak{M}$

$$\pi(g = e^{sX}) = e^{s\pi(X)}$$

via exponentiation. This is the fundamental property of Lie groups and Lie algebras. Thus, one can either study the representation of a Lie group or one can focus on the Lie algebra of the group and find matrices that do the same thing as $[\pi(X), \pi(Y)] = \pi([X, Y])$. That is often the easier task.

3.10 Generators of Lorentz Transformations

As discussed above a unitary representation of the Poincaré group $\boldsymbol{U}(\Lambda, A)$ is needed for a relativistic quantum field theory. Because one is allowed to ignore a phase in quantum mechanics there is a little bit more freedom and a projective unitary representation is sufficient.

According to Noether's theorem there is a conserved current $(\mathcal{J}^\mu)^{\rho\sigma}$ for each symmetry $X \to e^{\varepsilon J^{\mu\nu}} X$, one for each of the six $J^{\mu\nu}$. The conserved currents and the conserved charges are

$$(\mathcal{J}^\mu)^{\rho\sigma} \equiv X^\rho\, T^{\mu\sigma} - X^\sigma\, T^{\mu\rho} \qquad\qquad Q^{(\rho\sigma)} = \int d^3\vec{x}\, \left(X^\rho\, T^{0\sigma} - X^\sigma\, T^{0\rho}\right)$$

as shown in (2.12) and (2.13).

The inverse of Noether's theorem is surprisingly powerful but rarely discussed in text books. It states that these conserved charges are the generators of exactly the symmetry transformations they come from. In

other words, the conserved charges represent the Lie algebra of the symmetry transformations. One uses the symmetry transformation to construct the conserved current and throws out a lot of information by integrating over all space focusing on the time component, and these functions on phase space somehow still contain the knowledge of the structure of the symmetry transformation. They do so according to the Poisson bracket

$$\{Q^{\mu\nu}, Q^{\rho\sigma}\} = \eta^{\nu\rho} Q^{\mu\sigma} - \eta^{\mu\rho} Q^{\nu\sigma} + \eta^{\nu\sigma} Q^{\mu\rho} - \eta^{\mu\sigma} Q^{\nu\rho}$$

in classical field theory. The approach of turning Poisson brackets into commutators and converting quantities to operators works for free quantum field theories but falls apart for interacting theories. The proposal is to define

$$\boldsymbol{Q}^{\mu\nu} = \int d^3\vec{x} \left(X^\mu \, \boldsymbol{T}^{0\nu} - X^\nu \, \boldsymbol{T}^{0\mu} \right)$$

but one has to check that the $\boldsymbol{Q}^{\mu\nu}$ obey the Lie algebra of the Lorentz group.

Supposedly

$$\boldsymbol{K}^j = \int d^3\vec{x} \left(X^j \, \boldsymbol{T}^{00} - t \, \boldsymbol{T}^{0j} \right) = \boldsymbol{Q}^{0j} = -t\boldsymbol{P}^j + \int d^3\vec{x} \left(X^j \, \mathcal{H}(t, \vec{x}) \right)$$

are the generators of boosts. However, one has to be careful because both parts depend on time. Because $\boldsymbol{K}^j$ is time independent the time dependence of the two parts must cancel and it does so as

$$\frac{d\boldsymbol{K}^j}{dt} = -\boldsymbol{P}^j + i \left[\boldsymbol{H}, \int d^3\vec{x} \left(X^j \, \mathcal{H}(t, \vec{x}) \right) \right]$$

shows because an operator $\boldsymbol{O}$ evolves in quantum mechanics according to the Schrödinger equation as $[\boldsymbol{H}, \boldsymbol{O}]$. Note that $\boldsymbol{H}$ and $\mathcal{H}(t, \vec{x})$ do not commute. Only $\boldsymbol{H}$ and the integral of $\mathcal{H}(t, \vec{x})$ over all of space commute. The Hamilton operator commutes with the momentum operator such that one can write

$$0 = -\boldsymbol{P}^j + i \left[\boldsymbol{H}, \boldsymbol{K}^j \right]$$

and this means that $[\boldsymbol{H}, \boldsymbol{K}^j] = -i\boldsymbol{P}^j$ and therefore as expected that the boost operator commutes with the Hamilton operator to give a translation.

After similar calculations for the generators of rotations one gets

$$\boldsymbol{J}^{jk} = \int d^3\vec{x} \, \boldsymbol{\pi}(X) \left(X^j \, \partial_k - X^k \, \partial_j \right) \boldsymbol{\phi}(X) \tag{3.28}$$

and all Lorentz transformations can be expressed with

$$\boldsymbol{U}(\Lambda) = e^{-\frac{1}{2} \, \Omega_{\mu\nu} \, \boldsymbol{J}^{\mu\nu}} \tag{3.29}$$

as unitary operators. The question remains whether these unitary operators obey the right commutation relations. They are

$$\begin{aligned}
&[\boldsymbol{J}^j, \boldsymbol{J}^k] = -i \, \varepsilon^{jk}_{\ \ l} \, \boldsymbol{J}^l \quad [\boldsymbol{J}^j, \boldsymbol{K}^k] = -i \, \varepsilon^{jk}_{\ \ l} \, \boldsymbol{K}^l \quad [\boldsymbol{K}^j, \boldsymbol{K}^k] = +i \, \varepsilon^{jk}_{\ \ l} \, \boldsymbol{J}^l \\
&[\boldsymbol{J}^j, \boldsymbol{P}^k] = -i \, \varepsilon^{jk}_{\ \ l} \, \boldsymbol{P}^l \quad [\boldsymbol{K}^j, \boldsymbol{P}^k] = i \, \boldsymbol{H} \, \delta^{jk} \qquad [\boldsymbol{J}^j, \boldsymbol{H}] = [\boldsymbol{P}^j, \boldsymbol{H}] = [\boldsymbol{P}^j, \boldsymbol{P}^k] = 0 \\
&[\boldsymbol{K}^j, \boldsymbol{H}] = i \, \boldsymbol{P}^j
\end{aligned} \tag{3.30}$$

as one can show using (3.28). This is the full Lie algebra of the Poincaré group.

Thus, by exponentiating the generators (3.28), unitary operators for all transformations in the Poincaré group have been found for the Klein-Gordon field. To summarize, they are

$$\boldsymbol{U}(0, A) = e^{-i \, A_\mu \, \boldsymbol{P}^\mu} \qquad\qquad \boldsymbol{U}(\Lambda, 0) = e^{-\frac{1}{2} \, \Omega_{\mu\nu} \, \boldsymbol{J}^{\mu\nu}}$$

as in (3.14) for spacetime translations and in (3.29) for Lorentz transformations, and the Klein-Gordon quantum field $\boldsymbol{\phi}(t, \vec{x})$ in the Heisenberg picture, which evolves according to (3.16) in the Schrödinger picture where $\boldsymbol{H}$ is the Klein-Gordon Hamiltonian (3.3), is indeed a relativistic quantum field.

4 Interactions

4.1 Non-Interacting Particles in the Klein-Gordon Quantum Field

The Klein-Gordon quantum field is an example of a free theory, and almost all free theories are easy to solve. But they allow no interactions. In a free field with one particle the presence of another particle does not influence the dynamics of the first particle. These two particles evolve completely independent of each other. In the Hilbert space spanned by $\boldsymbol{a}^\dagger_{\vec{p}}\boldsymbol{a}^\dagger_{\vec{q}}|0\rangle$ the state of two particles

$$\int \phi_X(\vec{p})\,\phi_Y(\vec{q})\,\boldsymbol{a}^\dagger_{\vec{p}}\boldsymbol{a}^\dagger_{\vec{q}}|0\rangle\,\frac{d^3\vec{p}\,d^3\vec{q}}{(2\pi)^6} = |\Phi_2(0)\rangle$$

with the momentum distribution $\phi_X(\vec{p})$ and $\phi_Y(\vec{q})$ localized at X and Y evolves as

$$|\Phi_2(t)\rangle = e^{-i\boldsymbol{H}t}\,|\Phi_2(0)\rangle = \int \frac{d^3\vec{p}\,d^3\vec{q}}{(2\pi)^6}\,\phi_X(\vec{p})\,\phi_Y(\vec{q})\,e^{-i\boldsymbol{H}t}\,\boldsymbol{a}^\dagger_{\vec{p}}\,e^{i\boldsymbol{H}t}\,e^{-i\boldsymbol{H}t}\,\boldsymbol{a}^\dagger_{\vec{q}}\,e^{i\boldsymbol{H}t}\;|0\rangle$$

in time with the Klein-Gordon Hamiltonian $\boldsymbol{H}$ according to the Schrödinger equation. $\boldsymbol{H}$ is quadratic in $\boldsymbol{a}_{\vec{p}}, \boldsymbol{a}^\dagger_{\vec{p}}$ and therefore

$$e^{-i\boldsymbol{H}t}\,\boldsymbol{a}^\dagger_{\vec{p}}\,e^{i\boldsymbol{H}t}$$

is linear in $\boldsymbol{a}^\dagger_{\vec{p}}$, and the different $\boldsymbol{a}^\dagger_{\vec{p}}$ and $\boldsymbol{a}^\dagger_{\vec{q}}$ commute with each other. Therefore whether there is another creation operator or not does not play a role in this calculation, and the two particles evolve independently of each other.

Therefore, the Klein-Gordon quantum field alone cannot be a good model for nature because in nature particles interact with each other. In the next step interactions will be introduced.

4.2 Local Interactions

The main goal of physics is to model nature, and it is therefore necessary to maintain Lorentz invariance because nature is assumed to be Lorentz invariant. There is an uncountable nightmare of possible interactions in nature and it is not possible to prove that non-local interactions do not exist, but it seems that one can achieve this goal by assuming that interactions are local. It is much easier to deal with local interactions for Lorentz invariance than with non-local interactions. Together with the requirement that the theory must be renormalizable one can study theories given by their Lagrangian density such as

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

which are manifestly Lorentz invariant and where the infinite sum is a Taylor approximation of $V(\phi)$. The theory is local because the argument of ϕ is always the same location X. It actually turns out that the terms ϕ^n are irrelevant for $n > 4$, and $n = 3$ makes the theory unstable because it has no ground state. The only interesting theory has the Langrangian density

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

and is called for obvious reasons ϕ^4-theory. The number λ is the so-called coupling constant.

4.3 Dealing with Interactions

There are four ways of dealing with interactions in quantum mechanics:

1. Perturbation theory
2. Variational method

3. Monte Carlo sampling
4. Exact solution

and the last way is obviously the best way but it does not always seem possible. Here mainly perturbation theory is used but all four ways had a role to play in explaining quantum field theory. Perturbation theory was so successful that for a long time no other method was employed to study quantum fields, and it is still the dominant method. In relativistic quantum field theory the variational method had no great successes. Monte Carlo sampling is the only way for solving the equations of motion in nature, and the standard model of particle physics has been explored with it. There are exact solution in quantum field theory in lower dimensions. Perturbation theory mostly gives results that are so good that one does not see the difference between an exact solution and the solution of the perturbation theory.

In the following it is assumed that the Hamiltonian has the form $\boldsymbol{H} = \boldsymbol{H}_0 + \boldsymbol{H}_{\text{int}}$ where it is known how to calculate $e^{-i\boldsymbol{H}_0 t}$ and where $\boldsymbol{H}_{\text{int}}$ is small. Free quantum fields can be calculated rigorously, but as soon as interactions appear this is no longer possible. Technically all the error estimates that govern how much error is made by the expansions are dominated by the infinity norm $\|\boldsymbol{H}_{\text{int}}\|_\infty$ (the largest eigenvalue), and this is the reason why quantum field perturbation theory is not rigorous while normal perturbation theory is. For interacting quantum fields this norm is infinity, independent of how small the interaction is. However, this number should be smaller than one, and this is why perturbation theory in quantum field theory technically does not work.

However, this does not mean that it practically does not work. If one just assumes that $\|\boldsymbol{H}_{\text{int}}\|_\infty$ is smaller than one and keeps calculating Taylor series then one gets answers that work extremely well. The technical estimates are a worst case estimate, and nature seems to have chosen $\boldsymbol{H}_{\text{int}}$ such that it has so much more structure in it that the perturbation theory arguments work even though the worst case error estimate is infinity.

4.4 Interaction Picture

The goal in perturbation theory of quantum field theory is exactly the same as the goal for quantum theory namely solve the Schrödinger equation

$$i\,\frac{d}{dt}\,|\psi\rangle = \boldsymbol{H}\,|\psi\rangle$$

to figure out what the object will do in the future given that it has been set up in the initial state. A very good starting point for this goal is to go to the interaction picture. It means removing the easy part solved by $e^{-i\boldsymbol{H}_0 t}$ and concentrating on the difficult part $\boldsymbol{H}_{\text{int}}$. One takes the solution of the Schrödinger equation

$$|\psi_{\text{I}}(t)\rangle = e^{i\boldsymbol{H}_0 t}\,|\psi_{\text{S}}(t)\rangle$$

and evolves backwards in time according to the free Hamiltonian. Thus, one tries to undo as much of the free evolution as possible. Operators in the interaction picture, because one is in a rotating reference frame, also evolve

$$\boldsymbol{O}_{\text{I}}(t) = e^{i\boldsymbol{H}_0 t}\,\boldsymbol{O}_{\text{S}}\,e^{-i\boldsymbol{H}_0 t}$$

but they do not evolve according to the Heisenberg picture but to the Heisenberg picture just due to the free evolution. This is just a change of the reference frame. Instead of being stationary one moves with the free solution. The time evolution is

$$\begin{aligned} i\,\frac{d}{dt}\,|\psi_{\text{I}}(t)\rangle &= e^{i\boldsymbol{H}_0 t}\,(-\boldsymbol{H}_0 + \boldsymbol{H}_0 + \boldsymbol{H}_{\text{int}})\,|\psi_{\text{S}}(t)\rangle = e^{i\boldsymbol{H}_0 t}\,\boldsymbol{H}_{\text{int}}\,|\psi_{\text{S}}(t)\rangle \\ &= e^{i\boldsymbol{H}_0 t}\,\boldsymbol{H}_{\text{int}}\,e^{-i\boldsymbol{H}_0 t}\,e^{i\boldsymbol{H}_0 t}\,|\psi_{\text{S}}(t)\rangle = (\boldsymbol{H}_{\text{int}})_{\text{I}}(t)\,|\psi_{\text{I}}(t)\rangle \end{aligned}$$

and this is the Schrödinger equation in the interaction picture. It is as if the system evolves only according to the interaction term but the interaction term is time dependent now. Thus, one got rid of the easy part and is left with the difficult part, but the difficult part got harder.

The resulting equation

$$i\,\frac{d}{dt}\,|\psi_{\text{I}}(t)\rangle = (\boldsymbol{H}_{\text{int}})_{\text{I}}(t)\,|\psi_{\text{I}}(t)\rangle \tag{4.1}$$

is formally equivalent to the time dependent Schrödinger equation in quantum mechanics. The equation (4.1) can be solved by introducing a propagator-type approach, and this is exactly the same as used to solve the time dependent Schrödinger equation but with a time independent Hamiltonian. The solution with the so-called propagator $\boldsymbol{U}(t,t_0)$ starts from

$$|\psi_{\mathrm{I}}(t)\rangle = \boldsymbol{U}(t,t_0)\,|\psi_{\mathrm{I}}(t_0)\rangle \qquad\qquad i\,\frac{d}{dt}\boldsymbol{U}(t,t_0) = (\boldsymbol{H}_{\mathrm{int}})_{\mathrm{I}}(t)\,\boldsymbol{U}(t,t_0)$$

for some initial time t_0. The left side is a solution to (4.1) if the propagator is a solution to the right side. The propagator is found as

$$\boldsymbol{U}(t,t_0) = \mathbf{1} - i\int_{t_0}^{t} (\boldsymbol{H}_{\mathrm{int}})_{\mathrm{I}}(t')\,\boldsymbol{U}(t',t_0)\,dt'$$

by integration. This equation looks easy but it is expressed in terms of itself. One can iteratively guess and improve the guess until a solution is found. This fixed point iteration gives

$$\begin{aligned}\boldsymbol{U}(t,t_0) &= \mathbf{1} - i\int_{t_0}^{t} \boldsymbol{H}_{\mathrm{int}}(t')\,\boldsymbol{U}(t',t_0)\,dt' \\ &= \mathbf{1} - i\int_{t_0}^{t} \boldsymbol{H}_{\mathrm{int}}(t')\,dt' + (-i)^2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\,\boldsymbol{H}_{\mathrm{int}}(t')\,\boldsymbol{H}_{\mathrm{int}}(t'')\,\boldsymbol{U}(t'',t_0) \\ &= \ldots\end{aligned}$$

where $\boldsymbol{H}_{\mathrm{int}}$ here and from now on always means $(\boldsymbol{H}_{\mathrm{int}})_{\mathrm{I}}$ in the interaction picture. This procedure can be repeated, and each iteration is correct, independent of how small or big $\boldsymbol{H}_{\mathrm{int}}$ is. The structure of the n^{th} term is

$$(-i)^n \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \ldots \int_{t_0}^{t^{(n-1)}} dt^{(n)}\,\boldsymbol{H}_{\mathrm{int}}(t')\,\boldsymbol{H}_{\mathrm{int}}(t'')\ldots\boldsymbol{H}_{\mathrm{int}}(t^{(n)}) \tag{4.2}$$

with $\mathbf{1}$ as the 0^{th} term. This is the Dyson series. The norm of the n^{th} term is $\|.\|_\infty \leq \frac{(t-t_0)^n}{n!}\|\boldsymbol{H}_{\mathrm{int}}\|_\infty^{*n}$ where $\|\boldsymbol{H}_{\mathrm{int}}\|_\infty^* \equiv \sup_{t'\in[t,t_0]}\|\boldsymbol{H}_{\mathrm{int}}(t')\|_\infty = \|\boldsymbol{H}_{\mathrm{int}}\|_\infty$ because the norm is unitarily invariant. The upper bound is the Taylor series of an exponential. (The argument does not work for quantum fields because the norm is always infinity, but this fact is here ignored.)

There is a theorem stating

$$\boldsymbol{U}(t,t_0) = \mathcal{T}\left[e^{-i\int_{t_0}^{t}\boldsymbol{H}_{\mathrm{int}}(t')\,dt'}\right] \tag{4.3}$$

and this is certainly a nicer result than the infinite sum of integrals. Time ordering is defined here as

$$\mathcal{T}\big(\boldsymbol{H}(t_1)\,\boldsymbol{H}(t_2)\big) = \begin{cases}\boldsymbol{H}(t_1)\,\boldsymbol{H}(t_2) & \text{if } t_1 > t_2 \\ \boldsymbol{H}(t_2)\,\boldsymbol{H}(t_1) & \text{if } t_1 \leq t_2\end{cases}$$

similarly to (3.21).

4.5 Observables in Quantum Field Theory

There are many observables in quantum field theory but the focus here is on scattering experiments. They are the primary means for testing relativistic quantum fields. The S-matrix tells how the initial states are related to the final states, and it is determined by Green's functions. These Green's functions have the form

$$G^{(n)}(X_1,X_2,\ldots,X_n) = \langle\Omega|\mathcal{T}\left[\boldsymbol{\phi}_{1\mathrm{H}}\,\boldsymbol{\phi}_{2\mathrm{H}}\ldots\boldsymbol{\phi}_{n\mathrm{H}}\right]|\Omega\rangle \qquad\qquad \boldsymbol{\phi}_{j\mathrm{H}} = \boldsymbol{\phi}(X_j) = \boldsymbol{\phi}(t_j,\vec{x}_j) \tag{4.4}$$

where Ω is the full ground state of $\boldsymbol{H}$ and $\boldsymbol{\phi}_{j\mathrm{H}}$ is the j^{th} field operator in the Heisenberg picture.

The objective here is to show how one can calculate these Green's functions using perturbation theory. This will allow to predict the outcome of scattering experiments. The $\boldsymbol{\phi}_{j\mathrm{H}}$ are in the Heisenberg picture but they are supposed to be in the interaction picture to be used by perturbation theory.

The claim is that the n-point correlation functions can be calculated in terms of the ground state expectation value and the S-matrix as

$$\begin{aligned} G^{(n)}(X_1, X_2, ..., X_n) &= \frac{\langle 0 \,|\mathcal{T}\,[\boldsymbol{\phi}_{1\mathrm{I}}\,\boldsymbol{\phi}_{2\mathrm{I}} ... \,\boldsymbol{\phi}_{n\mathrm{I}}\,\mathsf{S}]\,|\,0\rangle}{\langle 0\,|\mathsf{S}|\,0\rangle} \\ \langle\phi\,|\mathsf{S}|\,\psi\rangle &= \lim_{t_\pm\to\pm\infty} \langle\phi\,|\boldsymbol{U}(t_+,t_-)|\,\psi\rangle \qquad \forall\,|\phi\rangle\,|\psi\rangle \end{aligned} \tag{4.5}$$

where $\boldsymbol{U}(t_1,t_2)$ in the S-matrix is defined in (4.3), where $\boldsymbol{H}_0\,|0\rangle = 0$, and where $|0\rangle$ is the ground state of the free Hamiltonian $\boldsymbol{H}_0$ which is completely known while the ground state $|\Omega\rangle$ of the full Hamiltonian is a nightmare full of entanglement. The proof is easy assuming $t_1 > t_2 > ... > t_n$ such that the terms are already in the right order. The numerator on the right-hand side is

$$\langle 0\,|\,\boldsymbol{U}(+\infty,t_1)\,\boldsymbol{\phi}_{1\mathrm{I}}\,\boldsymbol{U}(t_1,t_2)\,\boldsymbol{\phi}_{2\mathrm{I}} ... \,\boldsymbol{\phi}_{n\mathrm{I}}\,\boldsymbol{U}(t_n,-\infty)\,|\,0\rangle = \langle 0\,|\,\boldsymbol{U}(+\infty,t_0)\,\boldsymbol{\phi}_{1\mathrm{H}}\,\boldsymbol{\phi}_{2\mathrm{H}} ... \,\boldsymbol{\phi}_{n\mathrm{H}}\,\boldsymbol{U}(t_0,-\infty)\,|\,0\rangle$$

because $\boldsymbol{\phi}_{\mathrm{H}}(t,\vec{x}) = \boldsymbol{U}^\dagger(t,t_0)\,\boldsymbol{\phi}_{\mathrm{I}}(t,\vec{x})\,\boldsymbol{U}(t,t_0)$ where t_0 is the initial time. This is almost right except for

$$\begin{aligned} \boldsymbol{U}(t_0,-\infty) &= \lim_{t'\to-\infty}\lim_{t\to t_0} e^{i\boldsymbol{H}_0(t-t_0)}\,e^{-i\boldsymbol{H}(t_0-t')}\,e^{-i\boldsymbol{H}_0(t'-t)}\,|\,0\rangle = \lim_{t'\to-\infty} e^{-i\boldsymbol{H}(t_0-t')}\,|\,0\rangle \\ &= \lim_{t'\to-\infty}\left(|\Omega\rangle\,\langle\Omega| + \sum_{n>0} e^{-iE_n(t_0-t')}\,|E_n\rangle\,\langle E_n|\right)|\,0\rangle \end{aligned}$$

written in the eigenbasis of $\boldsymbol{H}$. The term $e^{i\boldsymbol{H}_0(t-t_0)}$ disappears because of the limes $t\to t_0$ and the term $e^{-i\boldsymbol{H}_0(t'-t)}$ disappears because $\boldsymbol{H}_0\,|0\rangle = 0$. This looks as if there is a discrete number of eigenvalues E_n but quantum fields have a continuous spectrum. Thus, the sum is more an integral and it follows

$$\lim_{t'\to-\infty}\int dE\,\langle E|0\rangle\;e^{-iE(t'-t_0)}\;\langle\phi|E\rangle \to 0$$

for some arbitrary ϕ due to the Riemann-Lebesgue lemma. Therefore, the numerator is

$$\langle 0|\Omega\rangle\,\langle\Omega|0\rangle\,\langle\Omega|\boldsymbol{\phi}_{1\mathrm{H}}\,\boldsymbol{\phi}_{2\mathrm{H}} ... \,\boldsymbol{\phi}_{n\mathrm{H}}|\Omega\rangle$$

and the denominator is $\langle 0|\Omega\rangle\,\langle\Omega|0\rangle$ such that this term cancels.

As a result, all relevant quantities have been converted to objects that can be handled by perturbation theory, and equation (4.5) is the starting point for all calculations in perturbation theory. The calculations can be done automatically with Taylor series for $\boldsymbol{\phi}_{\mathrm{I}}$ and Taylor series for S.

4.6 Time Ordering, Normal Ordering and Contractions

The S-matrix must be inside the time ordering operator in (4.5) because it is some limit of the operator

$$\begin{aligned} e^{-i\int_{-T}^{T}\boldsymbol{H}_{\mathrm{int}}(t)\,dt} &= e^{-i\int_{-T}^{T}(\int\frac{\lambda}{4!}\boldsymbol{\phi}_{\mathrm{I}}^4(t,\vec{x})\,d^3\vec{x})\,dt} \\ &= \mathbf{1} - i\int\frac{\lambda}{4!}\boldsymbol{\phi}_{\mathrm{I}}^4(X)\,d^4X + \frac{(-i)^2}{2!}\int_{X^0>Y^0}\frac{\lambda^2}{(4!)^2}\,\boldsymbol{\phi}_{\mathrm{I}}^4(X)\,\boldsymbol{\phi}_{\mathrm{I}}^4(Y)\,d^4X\,d^4Y + \\ &\qquad \frac{(-i)^2}{2!}\int_{X^0\le Y^0}\frac{\lambda^2}{(4!)^2}\,\boldsymbol{\phi}_{\mathrm{I}}^4(Y)\,\boldsymbol{\phi}_{\mathrm{I}}^4(X)\,d^4X\,d^4Y + ... \end{aligned}$$

and already the order two terms show the time dependence. The time ordering operator in the numerator on the right side of (4.5) is horrendously complicated as the terms in this Taylor expansion get mixed with the terms $\boldsymbol{\phi}_{j\mathrm{I}}$, and one has to deal with expressions of the form $\langle 0\,|\mathcal{T}\,[\boldsymbol{\phi}_{\mathrm{I}}(X_1)\,\boldsymbol{\phi}_{\mathrm{I}}(X_2) ... \,\boldsymbol{\phi}_{\mathrm{I}}(X_m)]\,|\,0\rangle$ where $m \ge n$. However, one can actually compute such expressions very accurately.

The starting point is

$$\begin{aligned} \boldsymbol{\phi}_{\mathrm{I}}(X) &= \int\frac{d^3\vec{p}}{\omega_{\vec{p}}}\left(\boldsymbol{a}_{\vec{p}}\,e^{-i\,P\cdot X} + \boldsymbol{a}_{\vec{p}}^\dagger\,e^{i\,P\cdot X}\right) = \boldsymbol{\phi}_{\mathrm{I}}^+(X) + \boldsymbol{\phi}_{\mathrm{I}}^-(X) \\ \boldsymbol{\phi}_{\mathrm{I}}^+(X) &= \int\frac{d^3\vec{p}}{\omega_{\vec{p}}}\,\boldsymbol{a}_{\vec{p}}\,e^{-i\,P\cdot X} \qquad\qquad \boldsymbol{\phi}_{\mathrm{I}}^-(X) = \int\frac{d^3\vec{p}}{\omega_{\vec{p}}}\,\boldsymbol{a}_{\vec{p}}^\dagger\,e^{i\,P\cdot X} \end{aligned} \tag{4.6}$$

where $P=(\omega_{\vec{p}},\vec{p})$ and $\omega_{\vec{p}}=\sqrt{|\vec{p}|^2+m^2}$ as usual. The notation with the $\boldsymbol{\phi}_{\mathrm{I}}^{+}(X)$ and $\boldsymbol{\phi}_{\mathrm{I}}^{-}(X)$ is confusing but for historical reasons this is what is used in the literature and what therefore also will be used here. (Note that $\boldsymbol{\phi}_{\mathrm{I}}^{\dagger}=\boldsymbol{\phi}_{\mathrm{I}}$ and thus $\boldsymbol{\phi}_{\mathrm{I}}^{+}$ cannot be mistaken for $\boldsymbol{\phi}_{\mathrm{I}}^{\dagger}$ that is never used.)

In the following $\boldsymbol{\phi}(X)$ and so on always means the operator $\boldsymbol{\phi}_{\mathrm{I}}(X)$ in the interaction picture. If occasionally the Heisenberg picture is used then it will clearly be indicated as $\boldsymbol{\phi}_{\mathrm{H}}(X)$, and there are no reasons to use the Schrödinger picture anymore.

The reason for splitting $\boldsymbol{\phi}(X)$ into $\boldsymbol{\phi}^{+}(X)$ and $\boldsymbol{\phi}^{-}(X)$ is that $\boldsymbol{\phi}^{+}(X)\,|\,0\rangle=0$ and $\langle 0\,|\,\boldsymbol{\phi}^{-}(X)=0$. The expressions of the form $\langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}_{\mathrm{I}}(X_1)\,\boldsymbol{\phi}_{\mathrm{I}}(X_2)\,...\,\boldsymbol{\phi}_{\mathrm{I}}(X_m)\right]|\,0\rangle$ could be evaluated without the time operator by moving $\boldsymbol{a}^{\dagger}$ and $\boldsymbol{a}$ past each other. By looking at one expression as an example

$$\mathcal{T}\left[\boldsymbol{\phi}(X)\boldsymbol{\phi}(Y)\right]\overset{X^0\geqq Y^0}{=}\boldsymbol{\phi}(X)\boldsymbol{\phi}(Y)\overset{X^0\geqq Y^0}{=}\boldsymbol{\phi}^{+}(X)\boldsymbol{\phi}^{+}(Y)+\boldsymbol{\phi}^{+}(X)\boldsymbol{\phi}^{-}(Y)+\boldsymbol{\phi}^{-}(X)\boldsymbol{\phi}^{+}(Y)+\boldsymbol{\phi}^{-}(X)\boldsymbol{\phi}^{-}(Y)$$

one can see what changes due to the time ordering operator. A term $\boldsymbol{\phi}^{+}$ on the right side or a term $\boldsymbol{\phi}^{-}$ on the left side simplify the expression because they get destroyed by $|0\rangle$. Thus, using the commutator

$$\begin{aligned}\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\right]\overset{X^0\geqq Y^0}{=}{}&\boldsymbol{\phi}^{+}(X)\,\boldsymbol{\phi}^{+}(Y)+\boldsymbol{\phi}^{-}(X)\,\boldsymbol{\phi}^{+}(Y)+\boldsymbol{\phi}^{-}(X)\,\boldsymbol{\phi}^{+}(Y)+\boldsymbol{\phi}^{-}(X)\,\boldsymbol{\phi}^{-}(Y)\\&+\left[\boldsymbol{\phi}^{+}(X),\boldsymbol{\phi}^{-}(Y)\right]\end{aligned}$$

one can move all $\boldsymbol{\phi}^{+}$ to the right. The commutator is $[\boldsymbol{\phi}^{+}(X),\boldsymbol{\phi}^{-}(Y)]=D(X-Y)\mathbf{1}$. Therefore,

$$\langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\right]|\,0\rangle\overset{X^0\geqq Y^0}{=}D(X-Y)\qquad\qquad\langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\right]|\,0\rangle\overset{X^0\leqq Y^0}{=}D(Y-X)$$

because all terms except for the commutator are 0. The result is

$$\langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\right]|\,0\rangle=\Delta_{\mathrm{F}}(X-Y)=i\int\frac{d^4P}{(2\pi)^4}\frac{e^{-i\,P\cdot(X-Y)}}{P^2-m^2+i\,\varepsilon}$$

and is called the Feynman propagator that has been introduced in (3.20) and in (3.22) for $\varepsilon>0$. This is the case for two operators but there can be an arbitrarily large number of operators in these expressions. The result for any number of operators becomes clear when using path integrals but they are not yet introduced here.

Before the general result is shown new notations have to be introduced. The Feynman propagator can be written as

$$\Delta_{\mathrm{F}}(X-Y)\,\mathbf{1}\equiv\overbrace{\boldsymbol{\phi}(X)\,\boldsymbol{\phi}}(Y)\tag{4.7}$$

with so-called contractions. This notation extends to

$$\boldsymbol{A}\,\overbrace{\boldsymbol{\phi}(X)\,\boldsymbol{B}\,\boldsymbol{\phi}}(Y)\,\boldsymbol{C}\equiv\Delta_{\mathrm{F}}(X-Y)\,\boldsymbol{ABC}$$

for all operators $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$. Another concept used here is the normal ordering. Instead of defining normal ordering rigorously, the two examples

$$\mathcal{N}(\boldsymbol{a}_{\vec{p}}\,\boldsymbol{a}_{\vec{q}}^{\dagger}\,\boldsymbol{a}_{\vec{r}})=\boldsymbol{a}_{\vec{q}}^{\dagger}\,\boldsymbol{a}_{\vec{p}}\,\boldsymbol{a}_{\vec{r}}\qquad\qquad\mathcal{N}(\boldsymbol{a}_{\vec{p}}^{\dagger}\,\boldsymbol{a}_{\vec{q}}\,\boldsymbol{a}_{\vec{r}}^{\dagger}\,\boldsymbol{a}_{\vec{s}})=\boldsymbol{a}_{\vec{p}}^{\dagger}\,\boldsymbol{a}_{\vec{r}}^{\dagger}\,\boldsymbol{a}_{\vec{q}}\,\boldsymbol{a}_{\vec{s}}$$

should illustrate the concept clear enough. Operators with daggers are moved to the left and operators without dagger are moved to the right. (Note that another notation often used instead of $\mathcal{N}(...)$ is **:**(...)**:** with two colons.) The expression $\mathcal{T}[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)]$ can be written as

$$\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\right]=\mathcal{N}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)+\Delta_{\mathrm{F}}(X-Y)\,\mathbf{1}\right]=\mathcal{N}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)+\overbrace{\boldsymbol{\phi}(X)\,\boldsymbol{\phi}}(Y)\,\mathbf{1}\right]$$

in these notations. Another more complicated example is

$$\begin{aligned}\mathcal{T}\left[\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4\right]=\mathcal{N}\Big[&\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4+\overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4+\overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}}_3\,\boldsymbol{\phi}_4+\overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}}_4\\&+\boldsymbol{\phi}_1\,\overbrace{\boldsymbol{\phi}_2\,\boldsymbol{\phi}}_3\,\boldsymbol{\phi}_4+\boldsymbol{\phi}_1\,\overbrace{\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}}_4+\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\overbrace{\boldsymbol{\phi}_3\,\boldsymbol{\phi}}_4\\&+\overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2}\,\overbrace{\boldsymbol{\phi}_3\,\boldsymbol{\phi}}_4+\overbrace{\boldsymbol{\phi}_1\,\overbrace{\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3}\,\boldsymbol{\phi}}_4+\overbrace{\boldsymbol{\phi}_1\,\overbrace{\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}}_4}\Big]\end{aligned}$$

where $\boldsymbol{\phi}_j=\boldsymbol{\phi}(X_j)$ is used to simplify the expressions. Thus, the time ordering of an expression is the same as the normal ordering of it together with all contractions.

4.7 Wick's Theorem and Feynman Propagators

Wick's theorem states that time ordering of n terms $\boldsymbol{\phi}_j$ is the same as normal ordering of the n terms together with all contractions. The proof by induction starts with $n = 2$ shown above. Thus, knowing Wick's theorem is correct for $n-1$ then one can assume without loss of generality $X_1^0 > X_2^0 > ... > X_n^0$ and knows the time-ordered expression. Further

$$\begin{aligned}\mathcal{T}\left[\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,...\,\boldsymbol{\phi}_n\right] &= (\boldsymbol{\phi}_1^+ + \boldsymbol{\phi}_1^-)\,\mathcal{N}\left[\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n + \text{all contractions not including } \boldsymbol{\phi}_1\right]\\ &= \mathcal{N}\left[\boldsymbol{\phi}_1^-\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n + \text{all contractions not including } \boldsymbol{\phi}_1\right]\\ &\quad + \boldsymbol{\phi}_1^+\,\mathcal{N}\left[\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n + \text{all contractions not including } \boldsymbol{\phi}_1\right]\end{aligned}$$

where the first summand is already in normal order. The second summand without contractions becomes

$$\begin{aligned}\boldsymbol{\phi}_1^+\,\mathcal{N}\left[\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n\right] &= \mathcal{N}\left[\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n\right]\,\boldsymbol{\phi}_1^+ + \left[\boldsymbol{\phi}_1^+, \mathcal{N}\left[\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n\right]\right]\\ &= \mathcal{N}\left[\boldsymbol{\phi}_1^+\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n\right]\\ &\quad + \mathcal{N}\left[\left[\boldsymbol{\phi}_1^+, \boldsymbol{\phi}_2\right]\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n + \boldsymbol{\phi}_2\left[\boldsymbol{\phi}_1^+, \boldsymbol{\phi}_3\right]...\,\boldsymbol{\phi}_n + ... + \boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_{n-1}\left[\boldsymbol{\phi}_1^+, \boldsymbol{\phi}_n\right]\right]\end{aligned}$$

because the commutators $[\boldsymbol{\phi}_1^+, \boldsymbol{\phi}_j]$ are proportional to the identity operator $\mathbf{1}$, and this expression can be written in the form

$$\boldsymbol{\phi}_1^+\,\mathcal{N}\left[\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n\right] = \mathcal{N}\left[\boldsymbol{\phi}_1^+\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n + \overbrace{\boldsymbol{\phi}_1^+\,\boldsymbol{\phi}_2}\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n + \overbrace{\boldsymbol{\phi}_1^+\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3}\,...\,\boldsymbol{\phi}_n + ... + \overbrace{\boldsymbol{\phi}_1^+\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,...\,\boldsymbol{\phi}_n}\right]$$

with $n-1$ contractions. The last step is similar and includes all contractions including those with $\boldsymbol{\phi}_1^+$.

The good news is that Wick's theorem simplifies all these calculations but the bad news is that it is still exponentially complex. The example shows

$$\begin{aligned}&\langle 0\,|\mathcal{T}[\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4]|\,0\rangle = \langle 0\,|\mathcal{N}\left[\overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2}\,\overbrace{\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4} + \overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3}\,\boldsymbol{\phi}_4 + \overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4}\right]|\,0\rangle\\ &= \Delta_{\mathrm{F}}(X_1 - X_2)\,\Delta_{\mathrm{F}}(X_3 - X_4) + \Delta_{\mathrm{F}}(X_1 - X_3)\,\Delta_{\mathrm{F}}(X_2 - X_4) + \Delta_{\mathrm{F}}(X_1 - X_4)\,\Delta_{\mathrm{F}}(X_2 - X_3)\end{aligned}$$

that all terms $\langle 0\,|\mathcal{N}[\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4]|\,0\rangle$, $\langle 0\,|\mathcal{N}[\overbrace{\boldsymbol{\phi}_1\,\boldsymbol{\phi}_2}\,\boldsymbol{\phi}_3\,\boldsymbol{\phi}_4]|\,0\rangle$ and so on with at least one $\boldsymbol{\phi}_j$ not bound in a contraction are zero because all annihilation operators go to the right, all creation operators go to the left in normal ordering and that results in $\boldsymbol{\phi}^+(X)\,|\,0\rangle = 0$ and $\langle 0\,|\,\boldsymbol{\phi}^-(X) = 0$.

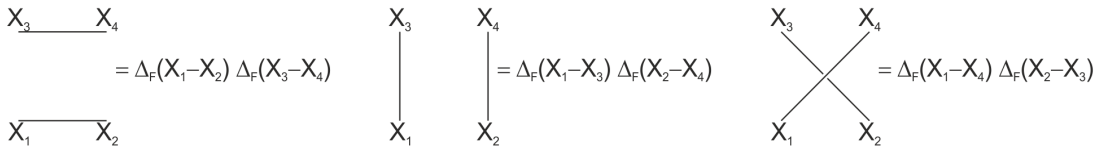


Figure 5: Feynman propagators as connections between points

Feynman propagators can be drawn in graphical form as shown in figure 5. In the following the opposite direction from pictures to formulas is used. A calculus for these diagrams is introduced because their manipulation is simpler than these horrible Taylor series. Only at the very end the pictures are converted back into expressions of Feynman propagators. It turns out that every process is the sum of integrals of products of Feynman propagators.

4.8 Taylor Expansion of the Green's Functions

The objects studied with perturbative quantum field theory in the interaction picture are the n-point or Green's functions (4.5) for the interacting quantum fields. The focus here is on the ϕ^4-theory where $\boldsymbol{H}_0$ is the Klein-Gordon field and where the coupling constant λ in the interaction term

$$\boldsymbol{H}_{\mathrm{I}} = \frac{\lambda}{4!}\int d^3\vec{x}\,\boldsymbol{\phi}^4(X) \tag{4.8}$$

is supposed to be very small. It is difficult to make rigorous sense of this theory because of infinite norms but this fact will be ignored.

To calculate $G^{(n)}(X_1, X_2, ..., X_n)$ the numerator and denominator in (4.5) must be expanded in λ such that the result is of the form

$$G^{(n)}(X_1, X_2, ..., X_n) = \frac{A_0 + \lambda A_1 + \lambda^2 A_2 + ...}{B_0 + \lambda B_1 + \lambda^2 B_2 + ...} = C_0 + \lambda C_1 + \lambda^2 C_2 + ...$$

with $A_j, B_j, C_j \in \mathbb{C}$. For the expansion into a single Taylor series $C_0 = \frac{A_0}{B_0}$ is obvious but already C_1 is less clear. With $1/(1+\lambda X) = 1 - \lambda X + \lambda^2 X^2 - ...$ one can at least calculate C_1 and there is hardly a reason to go beyond first order. The Taylor expansion is

$$\begin{aligned} G^{(n)}(X_1, X_2, ..., X_n) &= \frac{A_0 + \lambda A_1 + ...}{B_0\left(1 + \frac{\lambda B_1}{B_0} + ...\right)} = \frac{1}{B_0}\Big(A_0 + \lambda A_1 + ...\Big) \cdot \Big(1 - \frac{\lambda B_1}{B_0} + ...\Big) \\ &\approx \frac{1}{B_0}\Big(A_0 + \lambda A_1\Big) \cdot \Big(1 - \frac{\lambda B_1}{B_0}\Big) \approx \frac{A_0}{B_0} + \frac{\lambda A_1}{B_0} - \frac{\lambda A_0 B_1}{B_0^2} \end{aligned}$$

approximated to $O(\lambda)$. Thus, $C_0 = \frac{A_0}{B_0}$ and $C_1 = (\frac{A_1}{B_0} - \frac{A_0 B_1}{B_0^2})$ up to $O(\lambda)$, and the next task is to determine A_0, A_1, B_0, B_1.

The focus is first on $G^{(2)}(X, Y)$ as the simplest of the non-trivial Green's functions. This is

$$G^{(2)}(X,Y) = \lim_{T\to\infty} \frac{\langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\,e^{-i\int_{-T}^{T}\boldsymbol{H}_{\rm I}(t)\,dt}\right]|\,0\rangle}{\langle 0\,|\mathcal{T}\left[e^{-i\int_{-T}^{T}\boldsymbol{H}_{\rm I}(t)\,dt}\right]|\,0\rangle} = \frac{(*)}{(**)} \tag{4.9}$$

where

$$e^{-i\int_{-T}^{T}\boldsymbol{H}_{\rm I}(t)\,dt} = \mathbf{1} - i\int_{-T}^{T}\boldsymbol{H}_{\rm I}(t)\,dt + \frac{(-i)^2}{2!}\int_{-T}^{T}\int_{-T}^{T}\boldsymbol{H}_{\rm I}(t_1)\,\boldsymbol{H}_{\rm I}(t_2)\,dt_1\,dt_2 + ...$$

is the expansion of the exponentials to $O(\lambda)$. The numerator $(*)$ of $G^{(2)}(X,Y)$ becomes

$$\begin{aligned} (*) &\approx \langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\left\{\mathbf{1} - \frac{i\lambda}{4!}\int d^4Z\,\boldsymbol{\phi}^4(Z)\right\}\right]|\,0\rangle \\ &\approx \langle 0\,|\mathcal{T}\left[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\right]|\,0\rangle - \frac{i\lambda}{4!}\int d^4Z\,\langle 0\,|\mathcal{T}\Big[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\Big]|\,0\rangle \\ &\approx \Delta_{\rm F}(X-Y) - \frac{i\lambda}{4!}\int d^4Z\,\langle 0\,|\mathcal{N}\Big[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z) + \text{all contractions}\Big]|\,0\rangle \end{aligned}$$

using Wick's theorem. As mentioned above all terms with uncontracted field operators are 0. The two examples of fully contracted terms

$$\langle 0\,|\mathcal{N}\Big[\overbrace{\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)}\,\overbrace{\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)}\,\overbrace{\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)}\Big]|\,0\rangle \qquad \langle 0\,|\mathcal{N}\Big[\overbrace{\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)\,\boldsymbol{\phi}(Z)}\Big]|\,0\rangle$$

show that an even number of field operators is needed for full contraction. They also make clear that some clever combinatorics must simplify the calculations. There are $\binom{6}{2} = 15$ fully contracted summands but only the two

$$\Delta_{\rm F}(X-Y)\Delta_{\rm F}(Z-Z)\Delta_{\rm F}(Z-Z) = \Delta_{\rm F}(X-Y)\Delta_{\rm F}(Z-Z)^2 \qquad \Delta_{\rm F}(X-Z)\Delta_{\rm F}(Y-Z)\Delta_{\rm F}(Z-Z)$$

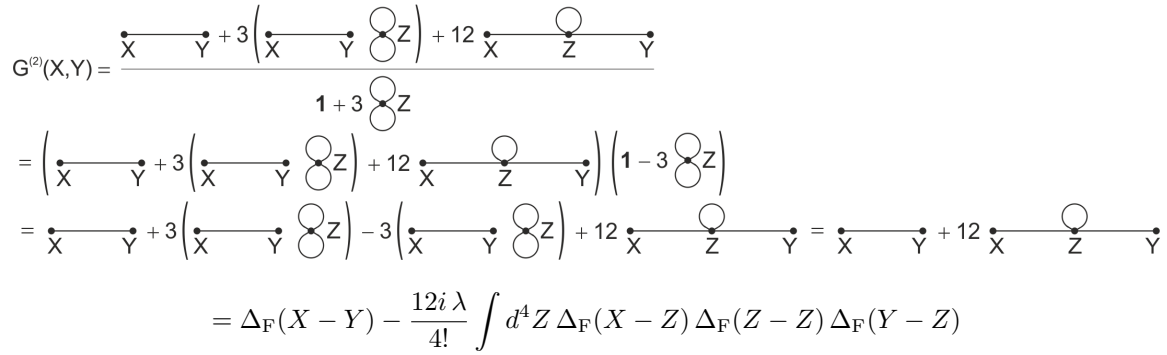
are different because $\boldsymbol{\phi}(X)$ must either be contracted with $\boldsymbol{\phi}(Y)$ or with $\boldsymbol{\phi}(Z)$. There are 3 contractions of the first type and 12 contractions of the second type. Thus, the numerator $(*)$ of $G^{(2)}(X,Y)$ is

$$\begin{aligned} (*) &\approx \Delta_{\rm F}(X-Y) \\ &\quad - \frac{3i\lambda}{4!}\int d^4Z\,\Delta_{\rm F}(X-Y)\Delta_{\rm F}(Z-Z)^2 - \frac{12i\lambda}{4!}\int d^4Z\,\Delta_{\rm F}(X-Z)\Delta_{\rm F}(Y-Z)\Delta_{\rm F}(Z-Z) \end{aligned}$$

to order $O(\lambda)$. The fact that the term $\Delta_{\rm F}(Z-Z)$ is infinite is ignored, and it is assumed to be finite.

4.9 Feynman Diagrams and Feynman Rules

These propagators of the numerator together with the denominator are illustrated in figure 6 in diagrammatic form. The diagrams are called Feynman diagrams, and the points in them are either called external or internal. External vertices are X and Y in this figure. They appear in the Green's function $G^{(n)}(X_1, ..., X_n)$ as field operators while all the other points are internal vertices. In this figure point Z is the only internal vertex.

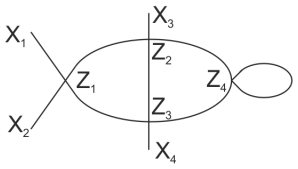


$$= \Delta_{\mathrm{F}}(X-Y) - \frac{12i\,\lambda}{4!}\int d^4Z\,\Delta_{\mathrm{F}}(X-Z)\,\Delta_{\mathrm{F}}(Z-Z)\,\Delta_{\mathrm{F}}(Y-Z)$$

Figure 6: The value of $G^{(2)}(X,Y)$ represented as a Feynman diagram to first order

Translating Feynman diagrams to equations is guided by Feynman rules. Each vertex is labeled, and for every edge one writes down a Feynman propagator. If the vertex is external then one does nothing. If the vertex is internal then one has to integrate over that argument and multiply it by $\frac{-i\lambda}{4!}$. In the ϕ^4-theory one encounters diagrams of this type and only this type.

The following example shows the translation from a complex Feynman diagram with the four external vertices X_j and the four internal vertices Z_j according to the Feynman rules for the ϕ^4-theory.



The value of the diagram

$$= \left(\frac{-i\lambda}{4!}\right)^4 \iiiint \Delta_{\mathrm{F}}(X_1-Z_1)\,\Delta_{\mathrm{F}}(X_2-Z_1)\,\Delta_{\mathrm{F}}(X_3-Z_2)\,\Delta_{\mathrm{F}}(X_4-Z_3)$$
$$\Delta_{\mathrm{F}}(Z_1-Z_2)\,\Delta_{\mathrm{F}}(Z_1-Z_3)\,\Delta_{\mathrm{F}}(Z_2-Z_3)\,\Delta_{\mathrm{F}}(Z_2-Z_4)$$
$$\Delta_{\mathrm{F}}(Z_3-Z_4)\,\Delta_{\mathrm{F}}(Z_4-Z_4)\,d^4Z_1\,d^4Z_2\,d^4Z_3\,d^4Z_4$$

contains only terms inside the integral because all edges have at least one internal vertex.

Going to second and third order make these diagrams more complicated. The world record is by hand probably around third or fourth order and by computer around order thirty. There are even worse problems. These asymptotic series have a zero radius of convergence.

To calculate

$$G^{(2)}(X,Y) = \frac{\langle 0\,|\mathcal{T}\,[\boldsymbol{\phi}(X)\,\boldsymbol{\phi}(Y)\,\mathsf{S}]\,|\,0\rangle}{\langle 0\,|\mathsf{S}|\,0\rangle}$$

given by (4.5) to arbitrary order one has to sum all possible diagrams with two external points. If one starts to expand the numerator and denominator to higher and higher orders each of them has products of field operators, and each of those terms were simplified by using Wick's theorem. Only the fully connected terms were kept. Associated with each of these fully contracted terms one can draw a Feynman diagram according to the Feynman rules in position space for the ϕ^4-theory:

1. For each propagator the edge connecting X and Y becomes $\Delta_{\mathrm{F}}(X-Y)$.
2. For each internal vertex Z insert integral $(-i\lambda)\int d^4Z$.
3. For each external vertex X insert 1.
4. Divide by symmetry factor.

Different theories have different rules, and higher order terms in this Taylor series expanding $G^{(2)}(X,Y)$ become successively more complicated but many diagrams are the same.

The Feynman diagram on the right side, for example, can occur in $4!\,8$ ways. This redundancy is called the symmetry factor of the diagram. What one does in a full computation is only write down distinct diagrams and then determine overcounting via symmetry factor. However, it is often a non-trivial task to determine the symmetry factor.

To summarize, the observable quantity $G^{(n)}(X_1, ..., X_n)$ from equation (4.5) has been studied via a perturbation expansion in the smallness of λ within the context of the Klein-Gordon Hamiltonian extended by the interaction term (4.8). Note that the interaction term may not be small even if λ is small but it is assumed here that it is small. The weak metric for success is that the result match the experiments. The interaction term is the sum of all possible diagrams with n external vertices subject to the Feynman rules of the given theory.

The above Feynman rules for the ϕ^4-theory are in position space but there is nothing that stops one from expressing them in momentum space. The Feynman propagator defined in (3.20) can be expressed in momentum space

$$\Delta_{\mathrm{F}}(X - Y) = \int \frac{d^4P}{(2\pi)^4} \frac{e^{-iP\cdot(X-Y)}}{P^2 - m^2 + i\varepsilon}$$

via a Fourier transform as shown in (3.22). By expressing all quantities in terms of P instead of X one could equivalently calculate the perturbation expansion for $G^{(n)}(X_1, ..., X_n)$ as follows: Wherever one sees a Feynman propagator $\Delta_{\mathrm{F}}(X - Y)$ in the perturbation expansion for the Green's function one can stick in the integral (3.22), do the integral in the Feynman rules with respect to the position space variable Z and one is left with integrals over momentum variables. One does not win very much because one just has a different way to write exactly the same terms, but it might occasionally be helpful. The Green's function $G^{(n)}(X_1, ..., X_n)$ is the sum of all diagrams with n external vertices subject to the Feynman rules for the ϕ^4-theory in momentum space:

1. For each propagator the oriented edge labeled P becomes $\frac{i}{P^2-m^2+i\varepsilon}$.
2. For each vertex where all four directed edges are pointing in insert $-i\lambda$.
3. For each external vertex X with an incoming edge labeled P insert $e^{-i\,P\cdot X}$.
4. Impose momentum conservation at each vertex.
5. Integrate over underdetermined momenta.
6. Divide by symmetry factor.

This is convenient, for example, in scattering theory where one has states with a definite (nearly definite) momentum coming in. Sometimes it is also easier to see cancellations. To make sure that all of the four directed edges are pointing in, one can change the direction of the edge and replace P by $-P$.

4.10 Vacuum Bubbles

One encounters soon infinities, but there are no infinities in quantum field theory. One only gets infinities if one asks wrong questions. Infinite ground-state energy, for example, cannot be measured because only energy differences can be measured. Therefore, the ground-state energy is not an observable quantity. A lot of infinities are of this kind, and they are not controversial infinities. One can shift them away as one can shift the ground-state energy away. They come from vacuum bubbles. However there are also more troublesome infinities that need a bit deeper physics.

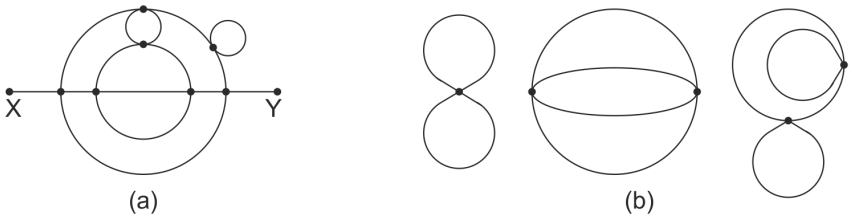
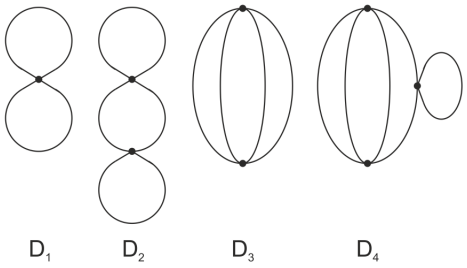


Figure 7: Examples of connected and unconnected parts in Feynman diagrams

The fact that the two terms in the numerator and denominator of figure 6 above cancel is positive because evaluating the diagram in the form of the number eight represents an infinite quantity. The question is

whether this cancellation is a lucky coincidence or a more general result. It turns out to be a very useful general result because it allows to eliminate a lot of unpleasant infinities.

A typical diagram has a standard structure. As illustrated in figure 7 there are parts that are connected to the external vertices X and Y as depicted in (a) for the case of $n = 2$, and there are parts that are not connected to the external vertices visualized in (b). The external vertices X and Y are always connected. It is therefore not possible that X is connected to a part of the diagram, Y is connected to another part of the diagram, and these two parts are not connected. The reason is that such degrees would be odd. In the case of the ϕ^4-theory each internal vertex is involved in four edges and each external vertex is involved in only one edge.



All possible disconnected pieces representing vacuum bubbles can be listed up to combinatorial equivalence. This set is infinite but countable where the number of edges and vertices help listing them. The figure on the right side shows the first four diagrams labeled D_1 to D_4. The value of diagram D_j is called v_j. All these values are infinite, and this is a serious point. In order to make sense of these numbers one has to set a cutoff into the theory. The only way one can interpret the equations coming up is to take them as a recipe for a quantity that depends on a cutoff. One does all the manipulations in the presence of this cutoff and only in the end one insists that the answers for operationally defined quantities do not depend on this cutoff.

Supposing that a given diagram D has n_j pieces of type D_j as well as the connected piece then the value of the diagram D is the value of the connected piece times the product of the disconnected pieces. This is

$$v_{\mathrm{D}} = v_{\mathrm{C}} \cdot \prod_{j=1}^{\infty} \frac{1}{n_j!} (v_j)^{n_j}$$

where v_{D} is the value of the diagram D and v_{C} the value of the connected piece. The factor $\frac{1}{n_j!}$ is the symmetry factor.

Checking this formula with the example in figure 7 shows that the value of this diagram D is the value of the connected piece times the value of the three disconnected pieces. The first one is of type D_1, the second of type D_3 and the third – after a bit of rearrangement – is of type D_2. Thus, one gets

$$v_{\mathrm{D}} = v_{\mathrm{C}} \cdot v_1 \cdot v_3 \cdot v_2$$

as the value of diagram D. This is correct because $n_j = 1$ for $j \in \{1, 2, 3\}$ and $n_j = 0$ for $j > 3$.

The numerator of $G^{(2)}(X, Y)$ in (4.9) is

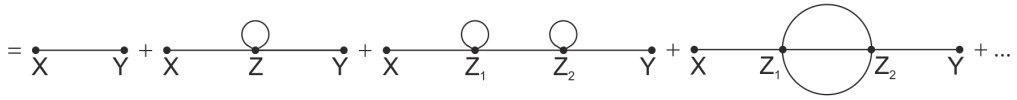
$$\begin{aligned}
(*) &= \sum_{\substack{\text{all} \\ \text{connected} \\ \text{pieces}}} \sum_{\text{all } n_j} (\text{value of connected piece}) \cdot \left(\prod_{j=1}^{\infty} \frac{1}{n_j!} (v_j)^{n_j} \right) \\
&= \sum_{\substack{\text{all} \\ \text{connected} \\ \text{pieces}}} (\text{value of connected piece}) \sum_{\text{all } n_j} \prod_{j=1}^{\infty} \frac{1}{n_j!} (v_j)^{n_j} \\
&= \sum_{\substack{\text{all} \\ \text{connected} \\ \text{pieces}}} (\text{value of connected piece}) \times \prod_{j=1}^{\infty} \left(\sum_{n_j} \frac{1}{n_j!} (v_j)^{n_j} \right) \\
&= \sum_{\substack{\text{all} \\ \text{connected} \\ \text{pieces}}} (\text{value of connected piece}) \cdot e^{\sum_j v_j}
\end{aligned}$$

assuming that the values v_j are finite, and the denominator of $G^{(2)}(X, Y)$ in (4.9) is

$$(**) = e^{\sum_j v_j}$$

by similar manipulations. Thus, this exponential factor cancels exactly. The final result is

$$G^{(2)}(X,Y) = \frac{(*)}{(**)} = \Big(\text{Sum of all connected diagrams subject to Feynman rules}\Big)$$



and all disconnected pieces disappeared. The same is true for $G^{(n)}(X_1, ..., X_n)$ in general.

4.11 Infinities and Cutoffs in Quantum Field Theory

One encounters many infinities in quantum field theory, and cutoffs help to handle this problem. They already appeared before interactions have been introduced. The Klein-Gordon Hamiltonian (3.3)

$$\boldsymbol{H} = \frac{1}{2}\int d^3\vec{x}\ \boldsymbol{\pi}(\vec{x})^2 + \big(\nabla\boldsymbol{\phi}(\vec{x})\big)^2 + m^2\,\boldsymbol{\phi}(\vec{x})^2$$

is an operator on the physical level but it also assumes the existence of infinitely many degrees of freedom. Roughly speaking, there is one quantum degree of freedom per location in spacetime. The smallest distance one can resolve today is pretty small but it is not that small compared to zero. Thus, the question is with what right does one assume an infinite number of degrees of freedom. Physicists extrapolate from things they have seen to things they have not yet seen. If one does an experiment with more and more degrees of freedom it is convincing that the same is true for even more degrees of freedom. However, fairly natural thought experiments one can do show that probing things at the Planck scale makes physics break down. It would be difficult to imagine what spacetime means at length scales below the Planck scale.

Theories such as the Klein-Gordon theory are effective theories for something more fundamental. The Navier-Stokes equation, for example, is an effective theory for quantum chromodynamics. It is very convenient to assume that a liquid is a continuous medium when doing experiments even though it is well-known that it consists of molecules. An effective theory is a description for the degrees of freedom one can measure at a certain scale. In other words, an effective theory is a description which explains all observations up to a given cutoff Λ usually expressed in inverse length. This also means that one accepts that there may be a scale beyond which this description no longer applies.

When a theory with Hamiltonian $\boldsymbol{H}$ breaks down above cutoff Λ and gives infinities, a possible solution is to add a correction term $\boldsymbol{H}_\Lambda$ to get the new Hamiltonian $\boldsymbol{H}' = \boldsymbol{H} + \boldsymbol{H}_\Lambda$ that makes the theory well behaved. An obvious constraint is that $\boldsymbol{H}'$ must match predictions of $\boldsymbol{H}$ up to an observational scale S. Because there is an infinite number of degrees of freedom left there is an infinite number of possible correction terms $\boldsymbol{H}_\Lambda$. However, many of the possible theories have no observational effect at lower scales.

The renormalization theory by Wilson starts by parametrizing

$$\boldsymbol{H}_\Lambda = \sum_{\text{polynomials}} P(\boldsymbol{\phi}, \nabla\boldsymbol{\phi}, \nabla^2\boldsymbol{\phi}, ..., \boldsymbol{\pi}, \nabla\boldsymbol{\pi}, ...) = c_1\boldsymbol{\phi} + c_2\boldsymbol{\phi}^2 + ...d_1\boldsymbol{\pi} + d_2\boldsymbol{\pi}^2 + ...e_j(\nabla\boldsymbol{\phi})^j + ...$$

to get this infinite number of possible ways to make the theory better. The question is what effect each term has on observations up to scale S. What effect does $d_2\boldsymbol{\pi}^2$, for example, have up to scale S? The remarkable result is that only a finite number of such terms has an effect on observations below S where having an effect means that the value is not suppressed by an inverse power of Λ. Such terms are called relevant. In dimension 4, the list of relevant terms is $\boldsymbol{\phi}, \boldsymbol{\phi}^2, (\nabla\boldsymbol{\phi})^2, \boldsymbol{\pi}, \boldsymbol{\pi}^2, \boldsymbol{\phi}^3, \boldsymbol{\phi}^4$.

A cutoff is any way one likes to stop the theory pretending that there are degrees of freedom beyond this value. The momentum cutoff as an example breaks Lorentz invariance but is otherwise very useful. It means for the Klein-Gordon Hamiltonian (3.8) as an integral over all momenta

$$\boldsymbol{H} = \int_{|\vec{p}|<\infty} \frac{d^3\vec{p}}{(2\pi)^3}\,\omega_{\vec{p}}\,\boldsymbol{a}^\dagger_{\vec{p}}\,\boldsymbol{a}_{\vec{p}} \qquad \xrightarrow{\boldsymbol{H}+\boldsymbol{H}_\Lambda} \qquad \boldsymbol{H} = \int_{|\vec{p}|<\Lambda} \frac{d^3\vec{p}}{(2\pi)^3}\,\omega_{\vec{p}}\,\boldsymbol{a}^\dagger_{\vec{p}}\,\boldsymbol{a}_{\vec{p}}$$

that momenta beyond the cutoff Λ are eliminated. Similarly, the interaction term in the ϕ^4-theory is constrained in this way by Λ in momentum space. From now on all Hamiltonians are written down with a cutoff in mind and this ensures that the theory is finite. Thus, a Hamiltonian is actually a family of Hamiltonians parameterized by the cutoff Λ. The dependence on the cutoff is chosen such that the experiments observed up to a given scale S are all modeled well by this Hamiltonian with the given cutoff.

4.12 Practical Computations and Scattering Theory

So far, the Green's functions $G^{(n)} = \langle \Omega \,|\mathcal{T}\{\boldsymbol{\phi}(X_1)...\boldsymbol{\phi}(X_n)\}|\, \Omega\rangle$ has been approximated in perturbation theory as the sum of all connected diagrams with n external vertices. The restriction to connected diagrams comes from the cancellation of the vacuum bubbles. These Green's functions are not directly observable. Quantities obtainable by experiments with particles are more like scattering cross-sections. Thus, the goal is to find connections between these Green's functions and observable quantities.

One can either promote this sum of all connected diagrams with n external vertices to the status of a definition and see whether predictions match experiments, or one can use Wilson's theory of renormalization and effective theories in a more rigorous way. These two paths are related but here the first approach will be followed.

Some of the diagrams lead to infinities. Within the ϕ^4-theory a few diagrams for the Green's function $G^{(4)}(X_1, X_2, X_3, X_4)$ are shown in figure 8. The first three diagrams already presented in figure 5 together with the Feynman propagators are zeroth order. Note that they are connected but not fully connected. The next two diagrams are examples for first order. They have a loop on one propagator and evaluate to infinity but play no particular role in computing scattering amplitudes for low orders as will be proven below. One can get rid of them with an argument that is very similar to the vacuum bubble cancellation introduced above.

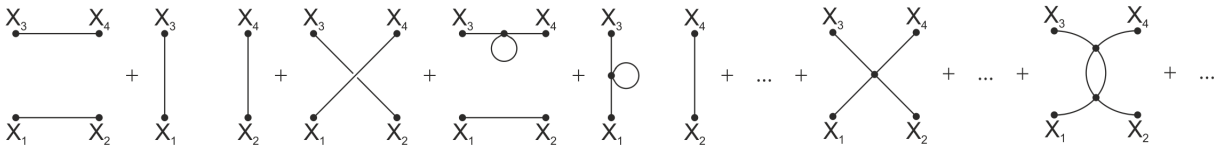


Figure 8: Connected Feynman diagrams in $G^{(4)}(X_1, X_2, X_3, X_4)$ of the ϕ^4-theory

However, those infinities coming from higher order terms such as the last diagram in figure 8 are not covered here. They typically come from integrals in momentum space and can be forced to be finite by imposing a cutoff Λ. The result obviously may depend on the chosen cutoff Λ but it turns out that one can always compensate arbitrary choices in Λ with arbitrary choices in coupling constants. This explains why Λ does not appear in any physical quantity that actually can be observed.

The probability of a scattering event is related to

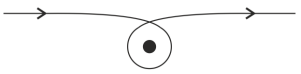
$$ {}_{\text{out}}\langle P_1 P_2 ... P_n | Q_A Q_B \rangle_{\text{in}} = \langle P_1 P_2 ... P_n | \, \mathsf{S} \, | Q_A Q_B \rangle \tag{4.10} $$

in the case of well collimated beams with momenta Q_A, Q_B where the S-matrix here is for the ϕ^4-theory. There is one collimated beam of particles with momentum Q_A and another collimated beam of particles with momentum Q_B usually in opposite direction, and they are well separated in the far past and can be described in the non-interacting theory. For a short intermediate time around $t = 0$ there is a scattering event where these two beams of particles collide and then there is debris flying off in various directions with momenta P_j. At very late times the resulting particles are also assumed to be so well separated that one can approximate the state of the system with the non-interacting theory. There are the two limits $t = \pm\infty$. Scattering theory is an effective theory for limits $|t| \gg c$ for some constant c.

The formula

$$ |\vec{q}_A \, \vec{q}_B\rangle_{\text{in}} = \sqrt{2\omega_{\vec{q}_A} \, 2\omega_{\vec{q}_B}} \, \boldsymbol{a}^\dagger_{\vec{q}_A} \, \boldsymbol{a}^\dagger_{\vec{q}_B} \, |\, 0\rangle $$

is based on the crucial assumption that $|...\rangle_{\text{in}}$ and $|...\rangle_{\text{out}}$ are states in the free theory. Therefore $|0\rangle$ is the vacuum and not $|\Omega\rangle$ as in the case of interactions.



The figure on the right side illustrates a scattering event in the center of mass coordinates. A single particle in the presence of a scattering potential drawn as a big black dot behaves more and more as a free particle the further away it is from this scattering potential. The state of the particle very early in time is $|\psi\rangle_{\text{in}}$, and the state very late in time is $|\psi\rangle_{\text{out}}$ such that the task of scattering theory is to say how a given interacting trajectory asymptotic to some $|\psi\rangle_{\text{in}}$ is transformed to a given outgoing trajectory asymptotic to some $|\psi\rangle_{\text{out}}$. This corresponds to only asking questions about observables defined for very large times $|t|$ and ignoring everything that happens at intermediate times.

Since it is known how to expand S into a Taylor series the amplitude for the scattering to occur in (4.10) can be calculated. To remove the possible case represented by $\mathbf{1}$ where the particle passes through without interacting the operator $\boldsymbol{T}$ is defined by $\mathsf{S} = \mathbf{1} + i\boldsymbol{T}$ representing only the interacting cases. It is also convenient to define $\mathcal{M}$ to factor out conservation of momentum as a delta function. In

$$\langle P_1 ... P_n |\, i\,\boldsymbol{T}\, | Q_A Q_B \rangle = (2\pi)^4\, \delta^{(4)} \left(Q_A + Q_B - \sum_{f=1}^{n} P_f \right) \cdot i\,\mathcal{M}\big((Q_A Q_B) \to (P_1 ... P_n) \big) \tag{4.11}$$

is $\mathcal{M}$ the non-trivial quantity that collects the actual scattering process. Because all particles are physical particles all momenta are on shell such that $P^0 = \omega_{\vec{p}}$. After some work the differential cross-section for $n = 2$ and all masses equal is

$$\left(\frac{d\sigma}{d\Omega} \right)_{\text{CM}} = \frac{|\mathcal{M}|^2}{64\,\pi^2\, E_{\text{CM}}^2} \tag{4.12}$$

in the center of mass frame CM. This equation expressed in terms of the asymptotic empirical distribution σ with respect to the angle Ω in the outcome of the scattering experiment justifies the focus on $\mathcal{M}$.

The next task is to compute equation (4.10) as best as possible by taking

$$|\vec{q}_A\, \vec{q}_B\rangle \propto \lim_{T\to\infty} e^{-i\,\boldsymbol{H}\,T}\, |\vec{q}_A\, \vec{q}_B\rangle_0$$

on faith as a basis for the computation. This is related to

$$|\Omega\rangle = \lim_{T\to\infty} e^{-i\,\boldsymbol{H}\,T}\, |0\rangle$$

stating that the ground state of the free theory is projected onto the ground state of the interacting theory in the given large time limit but is far more radical. It is certainly not true for large interactions but is likely to be true for small interactions. Under this assumption one can write

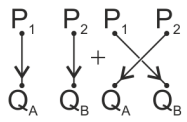
$$\lim_{T\to\infty} {}_0\langle P_1 ... P_n | e^{-i\,\boldsymbol{H}(2T)} | Q_A Q_B \rangle_0 \propto \lim_{T\to\infty} {}_0\langle P_1 ... P_n | \mathcal{T} \left\{ \exp\left[-i \int_{-T}^{T} \boldsymbol{H}_{\text{int}}(t)\, dt \right] \right\} | Q_A Q_B \rangle_0$$

where the proportionality factor is difficult to determine but it can be eliminated the same way as the vacuum bubbles. The formula

$$\langle P_1 ... P_n |\, i\,\boldsymbol{T}\, | Q_A Q_B \rangle = \lim_{T\to\infty} \left({}_0\langle P_1 ... P_n | \mathcal{T} \left\{ \exp\left[-i \int_{-T}^{T} \boldsymbol{H}_{\text{int}}(t)\, dt \right] \right\} | Q_A Q_B \rangle_0 \right)$$

holds but is only informally justified here but not rigorously proven.

The first term in the Taylor expansion for ${}_{\text{out}}\langle P_1 P_2 | Q_A Q_B \rangle_{\text{in}}$ is



$$\begin{aligned} O(\lambda^0) &= {}_0\langle P_1 P_2 | Q_A Q_B \rangle_0 = \sqrt{2\omega_{\vec{p}_1}\, 2\omega_{\vec{p}_2}\, 2\omega_{\vec{q}_A}\, 2\omega_{\vec{q}_B}}\; \boldsymbol{a}_{\vec{p}_1}\, \boldsymbol{a}_{\vec{p}_2}\, \boldsymbol{a}^\dagger_{\vec{q}_A}\, \boldsymbol{a}^\dagger_{\vec{q}_B}\, |\,0\rangle \\ &= 2\omega_{\vec{q}_A}\, 2\omega_{\vec{q}_B}\, (2\pi)^6 \big(\delta(P_1 - Q_A)\delta(P_2 - Q_B) + \delta(P_1 - Q_B)\delta(P_2 - Q_A) \big) \end{aligned}$$

after bringing it into normal order. It corresponds to the contribution of $\mathbf{1}$ in $\mathsf{S} = \mathbf{1} + i\boldsymbol{T}$ and is shown diagrammatically in the figure on the right side.

The first order term in λ is the next term to be determined. One begins to see the difference between the Green's function case and the scattering amplitude case. The vacuum in the Green's function case is

now replaced by momentum eigenstates and therefore applying Wick's theorem will presumably lead to a different result. The second term in the Taylor expansion is

$$\begin{aligned} O(\lambda^1) &= {}_0\langle P_1 P_2|\mathcal{T}\left\{\frac{-i\lambda}{4!}\int d^4X\,\boldsymbol{\phi}_{\mathrm{I}}^4(X)\right\}|Q_A Q_B\rangle_0 \\ &= {}_0\langle P_1 P_2|\mathcal{N}\left\{\frac{-i\lambda}{4!}\int d^4X\,\boldsymbol{\phi}_{\mathrm{I}}^4(X) + \text{all contractions}\right\}|Q_A Q_B\rangle_0 \end{aligned}$$

where an important point is that in the Green's function case only the fully contracted pieces mattered while here the field operators in normal order may not annihilate the state $|Q_A Q_B\rangle$ in this case.

There are therefore additional terms as the simplest possible example shows. Looking at the positive frequency operator applied to some momentum eigenstate $|\,\vec{p}\,\rangle$ in the free theory gives

$$\boldsymbol{\phi}_{\mathrm{I}}^+(X)\,|P\rangle_0 = \frac{1}{(2\pi)^3}\int\frac{d^3\vec{q}}{\sqrt{2\omega_{\vec{q}}}}\,e^{-i\,K\cdot X}\,\boldsymbol{a}_{\vec{q}}\,|\,\vec{p}\,\rangle = e^{-i\,P\cdot X}\,|0\rangle$$

because of

$$\boldsymbol{a}_{\vec{q}}\,|\,\vec{p}\,\rangle = \boldsymbol{a}_{\vec{q}}\,\sqrt{2\omega_{\vec{p}}}\;\boldsymbol{a}_{\vec{p}}^\dagger\,|0\rangle = (2\pi)^3\,\delta(\vec{q}-\vec{p})\,\sqrt{2\omega_{\vec{p}}}\,|0\rangle$$

by creating $|\vec{p}\rangle$ from $|0\rangle$ first and by using the commutation relations (3.7) afterwards.

The way to deal with $|Q_A Q_B\rangle$ is to include them in contractions by defining extended contractions of the form

$$\overbrace{\boldsymbol{\phi}_{\mathrm{I}}(X)\,|\,\vec{p}\,\rangle} = e^{-i\,P\cdot X} \qquad\qquad \langle\,\vec{p}\,|\,\overbrace{\boldsymbol{\phi}_{\mathrm{I}}(X)} = e^{i\,P\cdot X} \tag{4.13}$$

such that

$${}_0\langle P_1 P_2|\mathcal{T}\Big\{\text{field operators}\Big\}|Q_A Q_B\rangle_0 = \sum\Big\{\text{all possible full contractions including external states}\Big\}$$

where also the extended contractions in (4.13) are included. This allows to approximate the S-matrix transition amplitudes.

4.13 Infinities on External Edges

When calculating an amplitude for a scattering event one can reduce the problem to calculating terms like

$${}_0\langle P_1 ... P_n|\mathcal{T}\left[\frac{-\lambda}{4!}\int d^4X\,\boldsymbol{\phi}_{\mathrm{I}}^4(X)\right]|Q_A Q_B\rangle_0$$

with the help of the generalized Wick's theorem. There are fully contracted terms, partially contracted terms plus one uncontracted term

$$\mathcal{T}\left[\boldsymbol{\phi}^4(X)\right] = \mathcal{N}\left[\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}} + \overbrace{\boldsymbol{\phi}\,\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\boldsymbol{\phi}} + ... + \overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\boldsymbol{\phi}\,\boldsymbol{\phi} + \overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}} + ... + \boldsymbol{\phi}\,\boldsymbol{\phi}\,\boldsymbol{\phi}\,\boldsymbol{\phi}\right]$$

in Wick's theorem at $O(\lambda)$, and they lead to three different kinds of diagrams. The fully, partially and not contracted terms as Feynman diagrams are shown in figure 9 for $n = 2$. The fully connected terms presented diagrammatically in (a) of the figure give contributions such as

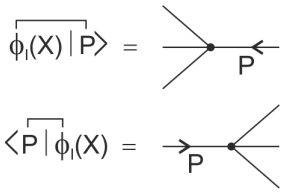
$$\frac{-\lambda}{4!}\int d^4X\,{}_0\langle P_1 P_2|\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}|Q_A Q_B\rangle_0$$

where $\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}$ is a complex number making the integration easy. The partially contracted terms depicted in (b) of the figure lead to contributions of the form

$$\frac{-\lambda}{4!}\int d^4X\,{}_0\langle P_1 P_2|\overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,\mathcal{N}\left[\boldsymbol{\phi}\,\boldsymbol{\phi}\right]|Q_A Q_B\rangle_0 = \frac{-\lambda}{4!}\int d^4X\,\big((*)\big)$$

where the two uncontracted $\boldsymbol{\phi}$ may be contracted to the right or to the left. The three possible cases are

$$
\begin{aligned}
(*) = \overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,{}_0\langle P_1P_2|\overbrace{\boldsymbol{\phi}\,\overbrace{\boldsymbol{\phi}|Q_AQ_B\rangle}}\,{}_0 + ... \\
+ \overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,{}_0\langle P_1P_2|\overbrace{\boldsymbol{\phi}\,\overbrace{\boldsymbol{\phi}|Q_AQ_B\rangle}}\,{}_0 + ... \\
+ \overbrace{\boldsymbol{\phi}\,\boldsymbol{\phi}}\,{}_0\langle P_1P_2|\overbrace{\boldsymbol{\phi}\,\overbrace{\boldsymbol{\phi}|Q_AQ_B\rangle}}\,{}_0 + ...
\end{aligned}
$$



where the first line contains all contractions to the right, the second all contractions to the left, and the third contractions to the left as well as to the right, and where the extended contractions (4.13) are shown in the figure on the right side. The uncontracted term

$$
\begin{aligned}
\frac{-\lambda}{4!}\int d^4X\,{}_0\langle P_1P_2|\mathcal{N}\left[\boldsymbol{\phi}^4\right]|Q_AQ_B\rangle_0 &= (4!)\frac{-i\lambda}{4!}\int d^4X\, e^{-i(Q_A+Q_B-P_1-P_2)\cdot X} \\
&= -i\lambda(2\pi)^4\,\delta^{(4)}(Q_A+Q_B-P_1-P_2)
\end{aligned}
$$

is an easy integral despite all the possible contractions and corresponds to the diagram (c) in the figure.

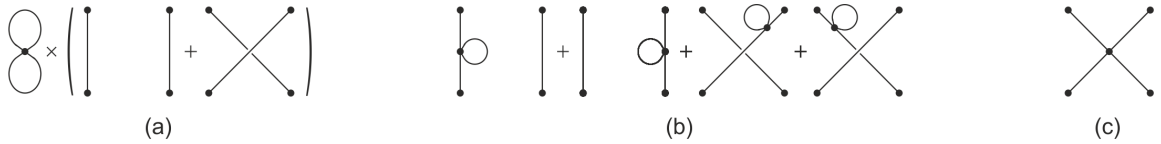


Figure 9: Feynman diagrams for fully contracted, partially contracted and uncontracted terms

The contributions in (a) are the standard fully contracted contributions one expects. The four diagrams in (b) represent the terms for the partially contracted terms with symmetry factors. The diagram (c) just imposes momentum conservation.

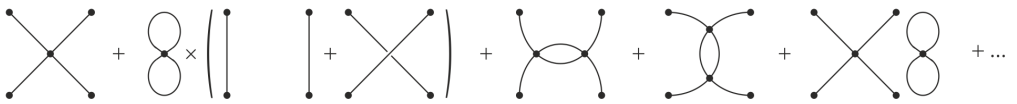
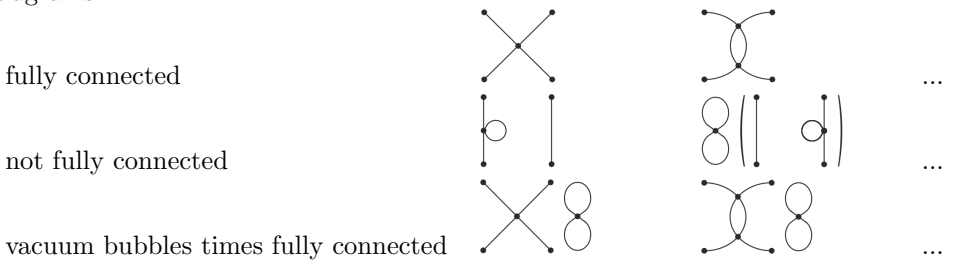
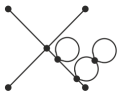


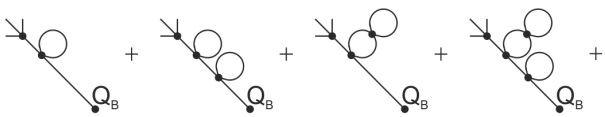
Figure 10: Full diagram expansion for $\langle P_1P_2|i\boldsymbol{T}|Q_AQ_B\rangle$

Putting all together means the sum of integrals for the operator $\boldsymbol{T}$ and gives the diagrams in figure 10 as the full diagram expansion for $\langle P_1P_2|i\boldsymbol{T}|Q_AQ_B\rangle$. In this expansion one keeps adding up terms in the Taylor series applying the generalized Wick's theorem. In this expansion one encounters three types of diagrams:



An observation without proof here is that if one sums up just the diagrams of vacuum bubbles times fully connected then the vacuum bubbles exponentiate. Therefore, the only diagrams which contribute to the S-matrix are the fully connected diagrams without vacuum bubbles.





However, there are still problems because the Feynman diagram just above on the right side results in infinity. It will turn out that such diagrams do not contribute to the S-matrix. All diagrams with external leg corrections such as the ones in the figure on the left side can be cut from the diagram with one cut. They represent the process where one projects the external momentum eigenstate

$|Q\rangle_0$ (just as for the $|\Omega\rangle$ case) onto the interacting external eigenstate $|Q\rangle_B$, the single particle state of the interacting theory. Because they factorize one can cut them off. One does that not in the sense of the cutoffs defined above to get rid of infinities but according to the amputation procedure in figure 11.

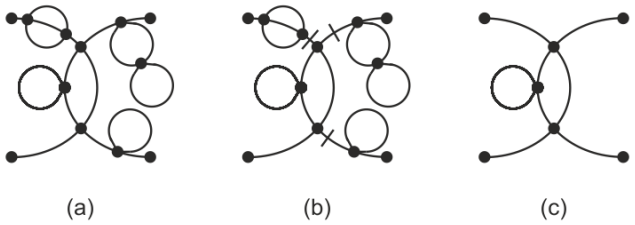


Figure 11: Amputation procedure

The amputation procedure starts from a diagram and determines the point for each external leg where one can cut the diagram to separate the external leg. For the initial diagram shown in (a) the cutting points have been marked in (b) and the diagram in (c) is the result after the separation. One leg does not need amputation, and one of the bubbles remains because it cannot be removed.

This leads to the rule for calculating scattering amplitudes. The value of

$$i\,\mathcal{M}\cdot(2\pi)p^4\cdot\delta^{(4)}\left(Q_A+Q_B-\sum_{f=1}^{n}P_f\right)$$

in (4.11) is the sum of all connected and amputated Feynman diagrams with Q_A, Q_B incoming and $P_1, ..., P_n$) outgoing. This is the final rule for how to compute a scattering process with Feynman diagrams. One could have evaluated the terms of the Taylor series introduced above and cancel quotients to get the same answer but the approach shown here is a much more economical way.

5 The Dirac Field

5.1 Relativistic Classical Theory of Fermions

In order to build a relativistic quantum theory of fermions one has to go back to the representation theory for the Lorentz group and try to find representations with an intrinsic spin. The Lie algebra for the Lorentz group is given in (3.27). The strategy to build a relativistic quantum theory used for the scalar field starts from a relativistic classical field theory, continues with a guess called quantization to get a quantum field theory and ends with searching for a unitary representation of the Lorentz group as the final test. This strategy is also used for a fermionic field called the Dirac field.

To do so one needs a classical field in the form of some kind of Lagrangian density. Given a set of objects $J^{\mu\nu}$ in form of 4×4 matrices obeying the commutation relations (3.27) one can get a representation of the Lorentz group by exponentiating these matrices

$$g = e^{-\frac{i}{2}\,\omega_{\mu\nu}\,J^{\mu\nu}}$$

and can look for Lagrangian densities invariant under these transformations g. Thus, the question is how to build these matrices $J^{\mu\nu}$. The solution is here simply given without showing the ideas needed by Dirac to developed it.

A set of four $n\times n$ matrices γ^μ satisfying the anticommutation relations on the left side gives the solution $S^{\mu\nu}$ on the right side

$$\{\gamma^\mu,\gamma^\nu\}=2\,\eta^{\mu\nu}\,\mathbb{I} \qquad\qquad S^{\mu\nu}=[\gamma^\mu,\gamma^\nu] \tag{5.1}$$

that satisfies (3.27). Because $S^{\mu\nu}$ is antisymmetric the number of degrees of freedom must be six as required by the Lorentz group. The number 4 is the smallest n leading to a non-trivial solution for the

Lorentz group in spacetime $\mathbb{M}_{1,3}$ with three plus one dimensions. The matrices

$$\gamma^0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \quad \gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} \qquad \gamma^j = \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} \tag{5.2}$$

used in the following to get the representation of the Lorentz group are called the Dirac matrices and are built from

$$\sigma^1 = \sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma^2 = \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma^3 = \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.3}$$

called the Pauli matrices σ^i which are both Hermitian and unitary and satisfy the commutation relations $[\sigma^j, \sigma^k] = 2i\,\varepsilon^{jkl}\,\sigma^l$ and the anticommutation relations $\{\sigma^j, \sigma^k\} = 2\,\delta^{jk}\,\mathbb{I}$. The quantities $S^{\mu\nu}$ in (5.1) are

$$\begin{aligned} S^{0j} &\equiv \frac{i}{4}\left[\begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix}\right] = -\frac{i}{2}\begin{pmatrix} \sigma^j & 0 \\ 0 & -\sigma^j \end{pmatrix} \\ S^{jk} &\equiv \frac{i}{4}\left[\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}\right] = \frac{1}{2}\varepsilon^{jkl}\begin{pmatrix} \sigma^l & 0 \\ 0 & -\sigma^l \end{pmatrix} \end{aligned} \tag{5.4}$$

where S^{0j} give the boost generators and S^{jk} the rotation generators.

The Jordan-Wigner transformations help to find γ matrices satisfying the anticommutation relations in (5.1). Fermions obey anticommutation relations and the question is how to represent $\{\mu^a, (\mu^b)^\dagger\} = \delta^{ab}$ where $a, b = 1, ..., n$. Using

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

to start with a simpler case, one can build any μ^j up to μ^n:

$n = 1$ $\quad \mu^1 = \sigma^+$

$n = 2$ $\quad \mu^1 = \sigma^+ \otimes \mathbb{I}$ $\qquad \mu^2 = \sigma^z \otimes \sigma^+$

$n = 3$ $\quad \mu^1 = \sigma^+ \otimes \mathbb{I} \otimes \mathbb{I}$ $\qquad \mu^2 = \sigma^z \otimes \sigma^+ \otimes \mathbb{I}$ $\qquad \mu^3 = \sigma^z \otimes \sigma^z \otimes \sigma^+$

$n = 4$ $\quad \mu^1 = \sigma^+ \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I}$ $\quad \mu^2 = \sigma^z \otimes \sigma^+ \otimes \mathbb{I} \otimes \mathbb{I}$ $\quad \mu^3 = \sigma^z \otimes \sigma^z \otimes \sigma^+ \otimes \mathbb{I}$ $\quad \mu^4 = \sigma^z \otimes \sigma^z \otimes \sigma^z \otimes \sigma^+$

...

The procedure for the γ matrices is not much harder. The trick is to find operators that anticommute with each other, and that is the difficult part where the Jordan-Wigner transformations become useful. Starting from $\gamma^0 = \sigma^x \otimes \mathbb{I}$ one can build $\gamma^j = i\,\sigma^y \otimes \sigma^j$ in the Minkowski spacetime, but the procedure generalizes to arbitrary dimensions. The Pauli matrices are basically the atoms of these fermionic anticommutation relations. (They also give a representation of the quaternions. Note that quaternions give a very efficient and compact way to work with rotations in three dimensions.)

5.2 Spinors, Spinor Fields and the Dirac Equation

An object ψ with four components transforming according to

$$\psi \to \psi' = e^{-\frac{i}{2}\,\omega_{\mu\nu}\,S^{\mu\nu}}\,\psi \tag{5.5}$$

is called a Dirac spinor. Spinors are vectors in the sense that they belong to a vector space but they are not vectors according to their transformation properties under the Lorentz group. They are defined by the way they look in different reference frames.

A spinor field is a map $\psi_a : \mathbb{M}_{1,3} \to \mathbb{C}^4$ transforming for $\Lambda \in$ SO(1,3) as

$$\psi_a(X) \to \psi' = \sum_{b=0}^{3} [\Lambda_{\frac{1}{2}}]_{ab}\,\psi_b(\Lambda^{-1}X) \tag{5.6}$$

where the difference between $\Lambda_{\frac{1}{2}}$ and Λ is

$$\Lambda_{\frac{1}{2}} = e^{-\frac{i}{2}\,\omega_{\mu\nu}\,S^{\mu\nu}} \qquad\qquad \Lambda = e^{-\frac{i}{2}\,\omega_{\mu\nu}\,J^{\mu\nu}} \tag{5.7}$$

with $S^{\mu\nu}$ defined in (5.1) as the newly introduced representation of the Lorentz group and $J^{\mu\nu}$ defined in (3.27) as the standard representation of the Lorentz group. Note that both use the same values $\omega_{\mu\nu}$.

Every representation of the Lorentz group can be used to build Lorentz invariant field equations $\mathcal{D}\psi = 0$. This is the ultimate aim because done so one can write down a Lagrangian density and try to quantize, and a first guess might be the Klein-Gordon equation $(\partial^2 + m^2)\psi_a(X) = 0$ for all a. It is indeed invariant but there is a far more remarkable equation due to Dirac. The equation

$$[\gamma^\mu, S^{\rho\sigma}] = (J^{\rho\sigma})^\mu{}_\nu\ \gamma^\nu$$

is equivalent to the equation

$$\left(1 + \frac{i}{2}\omega_{\rho\sigma}\,S^{\rho\sigma}\right)\gamma^\mu\left(1 - \frac{i}{2}\omega_{\rho\sigma}\,S^{\rho\sigma}\right) = \left(1 - \frac{i}{2}\omega_{\rho\sigma}\,S^{\rho\sigma}\right)^\mu{}_\nu\ \gamma^\nu$$

to first order in ω. Exponentiating this gives

$$e^{\frac{i}{2}\omega_{\rho\sigma}\,S^{\rho\sigma}}\,\gamma^\mu\,e^{-\frac{i}{2}\omega_{\rho\sigma}\,S^{\rho\sigma}} = \Lambda^\mu{}_\nu\,\gamma^\nu \qquad\qquad \Lambda_{\frac{1}{2}}^{-1}\,\gamma^\mu\,\Lambda_{\frac{1}{2}} = \Lambda^\mu{}_\nu\,\gamma^\nu$$

and that shows that the γ matrices transform exactly like 4-vectors under Lorentz transformations. This leads to the famous Dirac equation

$$(i\,\gamma^\mu\,\partial_\mu - m)\,\psi = 0 \tag{5.8}$$

because one gets a Lorentz invariant object if one contracts a vector against another vector. The object $(i\,\gamma^\mu\,\partial_\mu - m)$ is a 4×4 matrix of partial derivative operators and the Dirac equation can be written as

$$(i\,\gamma^0\,\partial_0 + i\,\gamma^1\,\partial_1 + i\,\gamma^2\,\partial_2 + i\,\gamma^3\,\partial_3 - m)\,\psi = 0$$

in expanded form and the term $i\,\gamma^0\,\partial_0$ is

$$i\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix}\partial_0 = \begin{pmatrix}0&0&i\,\partial_0&0\\0&0&0&i\,\partial_0\\i\,\partial_0&0&0&0\\0&i\,\partial_0&0&0\end{pmatrix}$$

for example such that

$$\psi = \begin{pmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{pmatrix} \qquad (i\,\gamma^\mu\,\partial_\mu - m)\,\psi = \begin{pmatrix}-m&0&i\,\partial_0 + i\,\partial_3&i\,\partial_1 + \partial_2\\0&-m&i\,\partial_1 - \partial_2&i\,\partial_0 - i\,\partial_3\\i\,\partial_0 - i\,\partial_3&-i\,\partial_1 - \partial_2&-m&0\\-i\,\partial_1 + \partial_2&i\,\partial_0 + i\,\partial_3&0&-m\end{pmatrix}\begin{pmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{pmatrix} \tag{5.9}$$

is the Dirac equation in the fully expanded form.

The Dirac equation as a scalar is obviously invariant under Lorentz transformations but it is worthwhile to show it explicitly. Substituting the transformation laws

$$\psi(X) \to \Lambda_{\frac{1}{2}}\,\psi(\Lambda^{-1}X) \qquad\qquad \partial_\mu \to (\Lambda^{-1})^\nu{}_\mu\,\partial_\nu$$

into the Dirac equation gives

$$\begin{aligned}(i\,\gamma^\mu\,\partial_\mu - m)\,\psi \to &\big(i\,\gamma^\mu\,(\Lambda^{-1})^\nu{}_\mu\,\partial_\nu - m\big)\Lambda_{\frac{1}{2}}\,\psi(\Lambda^{-1}X)\\ &= \Lambda_{\frac{1}{2}}\,\Lambda_{\frac{1}{2}}^{-1}\,\big(i\,\gamma^\mu\,(\Lambda^{-1})^\nu{}_\mu\,\partial_\nu - m\big)\Lambda_{\frac{1}{2}}\,\psi(\Lambda^{-1}X)\\ &= \Lambda_{\frac{1}{2}}\,\Big(i\,\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu\Lambda_{\frac{1}{2}}\,(\Lambda^{-1})^\nu{}_\mu\,\partial_\nu - m\Big)\,\psi(\Lambda^{-1}X)\\ &= \Lambda_{\frac{1}{2}}\,\big(i\,\Lambda^\mu{}_\sigma\gamma^\sigma\,(\Lambda^{-1})^\nu{}_\mu\,\partial_\nu - m\big)\,\psi(\Lambda^{-1}X)\\ &= \Lambda_{\frac{1}{2}}\,(i\,\gamma^\nu\,\partial_\nu - m)\,\psi(\Lambda^{-1}X) = 0\end{aligned}$$

showing that the Dirac equation is indeed Lorentz invariant. The derivation used the transformation properties of a spinor (5.5) and the definition (5.7) of $\Lambda_{\frac{1}{2}}$.

The Dirac equation contains the Klein-Gordon equation and is strictly stronger. Given that ψ is a solution to the Dirac equation then one is free to multiply it with any operator and therefore also the equation

$$(-i\,\gamma^\mu\,\partial_\mu - m)\,(i\,\gamma^\mu\,\partial_\mu - m)\,\psi = 0$$

is satisfied. Further steps using (5.1)

$$\left(\frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu + m^2\right)\psi = 0 \qquad \left(\eta^{\mu\nu}\partial_\mu\partial_\nu + m^2\right)\psi = 0 \qquad \left(\partial_\mu\partial^\mu + m^2\right)\psi = 0$$

show that each component of ψ satisfies the Klein-Gordon equation. The important difference is that the Dirac equation is a first order differential equation while the Klein-Gordon equation is second order.

5.3 Solutions to the Dirac Equation

To build a Lagrangian density a Lorentz scalar is needed. A first guess which turns out to be wrong is to take the adjoint of the Dirac spinor $\psi^\dagger(X)\,\psi(X)$ that would transform according to

$$\psi^\dagger(X)\,\psi(X) \to \psi^\dagger(X)\,\Lambda^\dagger_{\frac{1}{2}}\Lambda_{\frac{1}{2}}\,\psi(X)$$

by a Lorentz transformation. However, as shown above a representation of the Lorentz group cannot be unitary if it is finite dimensional, and $\Lambda_{\frac{1}{2}}$ defined in (5.7) is a representation of the Lorentz group. Thus, it cannot be unitary and it is usually not the case that

$$\Lambda^\dagger_{\frac{1}{2}}\Lambda_{\frac{1}{2}}$$

is the identity. The quantity that is a Lorentz scalar is $\psi^\dagger(X)\,\gamma^0\,\psi(X)$ such that

$$\bar\psi(X) = \psi^\dagger(X)\,\gamma^0 \tag{5.10}$$

is defined for spinors. Therefore, $\bar\psi(X)$ is used for fermionic fields instead of $\phi^\dagger(X)$ used for bosonic fields. The proof that $\bar\psi(X)\,\psi(X)$ is a scalar is based on the lemma

$$\Lambda^\dagger_{\frac{1}{2}}\gamma^0 = \gamma^0\,\Lambda^{-1}_{\frac{1}{2}}$$

that shows how $\Lambda^\dagger_{\frac{1}{2}}$ can be turned into $\Lambda^{-1}_{\frac{1}{2}}$. Similarly, $V^\mu = \bar\psi(X)\,\gamma^\mu\,\psi$ transforms as a Lorentz vector.

The Lagrangian density for the Dirac equation is

$$\mathcal{L} = \bar\psi\big(i\,\gamma^\mu\,\partial_\mu - m\big)\psi \tag{5.11}$$

and one can show that it is a scalar. When one applies the Euler-Lagrange equation one gets back the Dirac equation. Thus, the Lagrangian density (5.11) and the Dirac equation (5.8) build a classical theory of the Dirac field.

In order to guess a quantum version of the Dirac field the Hamiltonian density is helpful. The conjugate momentum to ψ is $i\,\psi^\dagger$, because γ^0 squared is the identity, and leads to

$$H = \int d^3\vec{x}\;\bar\psi\left(-i\,\vec\gamma\cdot\vec\nabla + m\right)\psi = \int d^3\vec{x}\;\psi^\dagger\left(-i\,\gamma^0\,\vec\gamma\cdot\vec\nabla + m\,\gamma^0\right)\psi \tag{5.12}$$

as the Hamiltonian where $\vec\gamma\cdot\vec\nabla = \gamma^j\,\partial_j$ means the sum over the spatial indices j according to Einstein's summation convention.

To study the solutions of the Dirac equation the zero-mass limit is first examined. It is obvious from (5.9) for $m = 0$ that a solution for ψ_1 and ψ_2 decouples from ψ_3 and ψ_4 because only the mass term couples

the upper two components of ψ with the lower two components of ψ. The two so-called Weyl spinors $\psi_{\mathrm{L}}(X) \in \mathbb{C}^2$ and $\psi_{\mathrm{R}}(X) \in \mathbb{C}^2$ defined as

$$\psi(X) = \begin{pmatrix} \psi_1(X) \\ \psi_2(X) \\ \psi_3(X) \\ \psi_4(X) \end{pmatrix} = \begin{pmatrix} \psi_{\mathrm{L}}(X) \\ \psi_{\mathrm{R}}(X) \end{pmatrix} \qquad \psi_{\mathrm{L}}(X) = \begin{pmatrix} \psi_1(X) \\ \psi_2(X) \end{pmatrix} \qquad \psi_{\mathrm{R}}(X) = \begin{pmatrix} \psi_3(X) \\ \psi_4(X) \end{pmatrix} \tag{5.13}$$

transform under Lorentz transformations as

$$\psi_{\mathrm{L}} \to \left(\mathbb{I} - i\,\vec{\theta} \cdot \frac{\vec{\sigma}}{2} - \vec{\beta} \cdot \frac{\vec{\sigma}}{2} \right) \psi_{\mathrm{L}} \qquad \psi_{\mathrm{R}} \to \left(\mathbb{I} - i\,\vec{\theta} \cdot \frac{\vec{\sigma}}{2} + \vec{\beta} \cdot \frac{\vec{\sigma}}{2} \right) \psi_{\mathrm{R}}$$

into themselves where $\vec{\theta} = (\theta_1, \theta_2, \theta_3)$ are the three parameters of a rotation and $\vec{\beta} = (\beta_1, \beta_2, \beta_3)$ are the three parameters of a boost. One gets the law for ψ_{R} from ψ_{L} by taking the complex conjugate and multiplying by σ^2 using the operation

$$\sigma^2\,\vec{\sigma}^* = -\vec{\sigma}\,\sigma^2$$

such that $\sigma^2\,\psi_{\mathrm{L}}^*$ transforms like ψ_{R}. This operation is called the spin flipper on the Bloch sphere because it takes some Bloch vector pointing in one direction on the Bloch sphere and gives the inverse vector pointing in the opposite direction. The Dirac equation in terms of the Weyl spinors

$$(i\,\gamma^\mu \partial_\mu - m)\psi = \begin{pmatrix} -m & i(\partial_0 + \vec{\sigma} \cdot \vec{\nabla}) \\ i(\partial_0 - \vec{\sigma} \cdot \vec{\nabla}) & m \end{pmatrix} \begin{pmatrix} \psi_{\mathrm{L}} \\ \psi_{\mathrm{R}} \end{pmatrix} \tag{5.14}$$

also shows that the equations of motion for the two Weyl spinors are coupled for $m \neq 0$ and decoupled for $m = 0$ such that there are two independent solutions. In this zero-mass limit one has obtained a representation of the Poincaré group that is 2-dimensional. Thus, there exist representations of the Poincaré group with a dimension less than 4 but only for $m = 0$. The two decoupled equations are

$$i(\partial_0 - \vec{\sigma} \cdot \vec{\nabla})\,\psi_{\mathrm{L}} = 0 \qquad i(\partial_0 + \vec{\sigma} \cdot \vec{\nabla})\,\psi_{\mathrm{R}} = 0 \tag{5.15}$$

for $m = 0$. Thus, one could build field theories with just ψ_{L} or just ψ_{R}. Using the definitions $\sigma^\mu = (\mathbb{I}, \vec{\sigma})$ and $\bar{\sigma}^\mu = (\mathbb{I}, -\vec{\sigma})$, the γ matrices (5.2) and the decoupled Dirac equations (5.15) can be written as

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \qquad \begin{aligned} i(\bar{\sigma} \cdot \partial)\,\psi_{\mathrm{L}} &= 0 \\ i(\sigma \cdot \partial)\,\psi_{\mathrm{R}} &= 0 \end{aligned}$$

in a simple form, respectively.

Solving the Dirac equation for free particles with mass $m \neq 0$ starts from the observation that the Dirac equation is a linear system of partial differential equations such that a Fourier transformation can be used. The solution can be expressed in terms of plane waves

$$\psi(X) = e^{-i\,P \cdot X}\,u(P)$$

for $u(P) \in \mathbb{C}^4$, and the general solution is a linear combination of these plane waves. The two constraints are $P^2 = m^2$ such that the solution is on shell and $P^0 > 0$ such that the energy is positive. (Later also solutions with negative energy will be considered.) Inserting these general plane wave solution into the Dirac equation (5.8) gives the equation

$$(P_\mu \gamma^\mu - m)\ u(P) = 0 \tag{5.16}$$

which is of the form matrix times vector equal zero. Because this is a covariant equation one can solve it in the rest frame $P = (m, \vec{0})$ and transform it back to a general reference frame. Equation (5.16) in the rest frame becomes

$$\left[m \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} - m \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix} \right] \begin{pmatrix} u_{\mathrm{L}} \\ u_{\mathrm{R}} \end{pmatrix}$$

and the solutions are of the form $u_{\mathrm{L}} = u_{\mathrm{R}} = \sqrt{m}\xi$ where ξ is a fixed vector with norm 1 and $\sqrt{m}$ will later turn out to be convenient. There are two linearly independent vectors ξ because this is taking place in a 2-dimensional complex vector space. Thus, the solutions in the rest frame with $P = (m, \vec{0})$ are

$$e^{-i\,P\cdot X}\,u^s(P) = \sqrt{m}\begin{pmatrix}\xi^s\\ \xi^s\end{pmatrix}e^{-i\,P\cdot X} \qquad \xi^1 = \begin{pmatrix}1\\0\end{pmatrix} \qquad \xi^2 = \begin{pmatrix}0\\1\end{pmatrix}$$

for $s = 1, 2$.

To demonstrate the solution in a general reference frame a boost along the z-direction is shown as an example. The corresponding Lorentz transformation is

$$\Lambda = \begin{pmatrix}\cosh(\eta) & 0 & 0 & \sinh(\eta)\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \sinh(\eta) & 0 & 0 & \cosh(\eta)\end{pmatrix} = e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}}$$

where η is called the rapidity. Only one of the $\omega_{\mu\nu}$ for this boost transformation is non-zero and this is $\omega_{03} = 2\eta$. To carry out this transformation on a spinor

$$\Lambda_{\frac{1}{2}} = e^{-\frac{i}{2}\omega_{03}S^{03}} = e^{-\frac{i}{2}2\eta\,S^{03}} = e^{-\frac{1}{2}\eta\begin{pmatrix}\sigma^3 & 0\\ 0 & \sigma^3\end{pmatrix}} = \begin{pmatrix}e^{-\frac{1}{2}\eta\,\sigma^3} & 0\\ 0 & e^{-\frac{1}{2}\eta\,\sigma^3}\end{pmatrix}$$

from (5.7) with (5.4) has to be used. The solutions to the Dirac equation look like

$$\psi(X) = e^{-i\,P\cdot X}\,u^s(P) = e^{-i\,P\cdot X}\sqrt{m}\begin{pmatrix}e^{-\frac{1}{2}\sigma^3}\xi^s\\ e^{\frac{1}{2}\sigma^3}\xi^s\end{pmatrix} = e^{-i\,P\cdot X}\begin{pmatrix}\sqrt{P\cdot\sigma}\,\xi^s\\ \sqrt{P\cdot\bar{\sigma}}\,\xi^s\end{pmatrix}$$

in the $P = \Lambda(m, \vec{0})$ frame. The last step uses $(P\cdot\sigma)(P\cdot\bar{\sigma}) = m^2 = P^2$ and gives the general solutions

$$\psi(X) = e^{-i\,P\cdot X}\,u^s(P) = e^{-i\,P\cdot X}\begin{pmatrix}\sqrt{P\cdot\sigma}\,\xi^s\\ \sqrt{P\cdot\bar{\sigma}}\,\xi^s\end{pmatrix}$$

for any Lorentz transformation and not only for the boost along the z-axis.

These solutions are two linearly independent solutions to the Dirac equation but there must be four. The other two come from negative frequency solutions. Similarly to the positive frequency solutions $u(P)$ one defines

$$\psi(X) = e^{+i\,P\cdot X}\,v(P) \qquad \xi^1 = \begin{pmatrix}1\\0\end{pmatrix} \qquad \xi^2 = \begin{pmatrix}0\\1\end{pmatrix}$$

with $P^2 = m^2$ and $P^0 > 0$. The four linearly independent solutions to the Dirac equation are

$$\psi(X) = \begin{cases} e^{-i\,P\cdot X}\,u^s(P) & = e^{-i\,P\cdot X}\begin{pmatrix}\sqrt{P\cdot\sigma}\,\xi^s\\ \sqrt{P\cdot\bar{\sigma}}\,\xi^s\end{pmatrix}\\ e^{+i\,P\cdot X}\,v^s(P) & = e^{+i\,P\cdot X}\begin{pmatrix}\sqrt{P\cdot\sigma}\,\eta^s\\ -\sqrt{P\cdot\bar{\sigma}}\,\eta^s\end{pmatrix}\end{cases} \tag{5.17}$$

for $s = 1, 2$.

The definition (5.10) has been introduce such that $\bar{\psi}\,\psi$ is Lorentz invariant. This means

$$\bar{\psi}\,\psi = \bar{u}^s(P)\,u^s(P) = 2m\left(\xi^s\right)^\dagger\left(\xi^s\right)$$

for the normalization of $\psi(X) = e^{-i\,P\cdot X}\,u^s(P)$ and similarly for $\psi(X) = e^{+i\,P\cdot X}\,v^s(P)$. This shows with

$$\begin{aligned} &\bar{u}^r(P)\,u^s(P) = 2m\,\delta^{rs} \quad && \bar{v}^r(P)\,v^s(P) = -2m\,\delta^{rs}\\ &\bar{u}^r(P)\,v^s(P) = 0 && \bar{v}^r(P)\,u^s(P) = 0\end{aligned} \tag{5.18}$$

that the four solutions of the classical Dirac field are indeed linearly independent.

5.4 The Free Quantum Dirac Field

The quantization of the Klein-Gordon field above was based on placing a harmonic oscillator on every degree of freedom. Similarly, the strategy for the quantization of the Dirac field is to guess a quadratic quantum theory via plane wave solutions. In the case of the Klein-Gordon field a bosonic degree of freedom had been attached to each plane wave. Trying to do the same in the case of the Dirac field fails because one has to attach a fermionic degree of freedom to each plane wave.

The bosonic approach fails but it is interesting to see how it fails because it works in some sense. The failure of the approach tells what is considered a valid quantization. Thus, assuming that the quantum Dirac field is a theory of many bosons means that the commutation relations

$$[\boldsymbol{\psi}_a(t,\vec{x}),\boldsymbol{\psi}_b^\dagger(t,\vec{y})]=\delta^{(3)}(\vec{x}-\vec{y})\,\delta_{ab}$$

are imposed at equal times t with the spinor indices $a,b\in\{1,2,3,4\}$. Using $e^{i\,P\cdot X}\,u^s(P)$ and $e^{-i\,P\cdot X}\,v^s(P)$ as the plane-wave basis to expand

$$\boldsymbol{\psi}(X)=\int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\,e^{i\,P\cdot X}\sum_{s=1,2}\left(\boldsymbol{a}^s_{\vec{p}}\,u^s(\vec{p})+\boldsymbol{b}^s_{-\vec{p}}\,v^s(-\vec{p})\right)$$

leads to the annihilation operators $\boldsymbol{a}^s_{\vec{p}}$ and $\boldsymbol{b}^s_{\vec{p}}$ for $s=1,2$ with

$$\left[\boldsymbol{a}^r_{\vec{p}},\left(\boldsymbol{a}^s_{\vec{q}}\right)^\dagger\right]=\left[\boldsymbol{b}^r_{\vec{p}},\left(\boldsymbol{b}^s_{\vec{q}}\right)^\dagger\right]=(2\pi)^3\,\delta^{(3)}(\vec{p}-\vec{q})\,\delta^{rs}$$

as their commutation relations. All the other commutation relations are zero. The classical Hamiltonian H turned into the quantum Hamiltonian operator $\boldsymbol{H}$ becomes

$$\boldsymbol{H}=\int\frac{d^3\vec{p}}{(2\pi)^3}\sum_s\left(\omega_{\vec{p}}\,\left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger\left(\boldsymbol{a}^s_{\vec{p}}\right)-\omega_{\vec{p}}\,\left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger\left(\boldsymbol{b}^s_{\vec{p}}\right)\right)$$

after some derivations. The problem is that this Hamiltonian $\boldsymbol{H}$ has no ground state because it is not bounded from below. It is possible to create states with arbitrarily low energy by simply occupying the modes with the $\boldsymbol{b}$ operators. It creates infinitely many bosons in all these $\boldsymbol{b}$ modes, and this system is unstable.

Many things can be wrong here. One problem could be that the quantization from H to $\boldsymbol{H}$ was too naive. Because there are infinitely many Hamiltonians it might be possible that there is one for the bosonic quantization of the Dirac field. Because the goal is not only to find a Hamiltonian but to find a unitary representation of the Poincaré group, and because there is the other option of a fermionic quantization matching the experiment, the Dirac field is assumed to be fermionic.

The successful quantization of the Dirac field imposes the anticommutation relations

$$\left\{\boldsymbol{\psi}_a(t,\vec{x}),\boldsymbol{\psi}_b^\dagger(t,\vec{y})\right\}=\delta^{(3)}(\vec{x}-\vec{y})\,\delta_{ab}\qquad\left\{\boldsymbol{\psi}_a(t,\vec{x}),\boldsymbol{\psi}_b(t,\vec{y})\right\}=\left\{\boldsymbol{\psi}_a^\dagger(t,\vec{x}),\boldsymbol{\psi}_b^\dagger(t,\vec{y})\right\}=0\tag{5.19}$$

at equal times t. The next steps are very similar to the Klein-Gordon field. The field operators defined in terms of position are expressed in terms of momentum, the creation and annihilation operators are introduced and the Hamiltonian is defined. In the first step the field operators are expanded

$$\begin{aligned}\boldsymbol{\psi}(t,\vec{x})&=\int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\sum_{s=1}^2\left(\boldsymbol{a}^s_{\vec{p}}\;u^s(\vec{p})\,e^{-i\,\vec{p}\cdot\vec{x}}+\left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger v^s(\vec{p})\,e^{i\,\vec{p}\cdot\vec{x}}\right)\\\bar{\boldsymbol{\psi}}(t,\vec{x})&=\int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\sum_{s=1}^2\left(\boldsymbol{b}^s_{\vec{p}}\;\bar{v}^s(\vec{p})\,e^{-i\,\vec{p}\cdot\vec{x}}+\left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger\bar{u}^s(\vec{p})\,e^{i\,\vec{p}\cdot\vec{x}}\right)\end{aligned}\tag{5.20}$$

in terms of creation and annihilation operators. The anticommutation relations (5.19) impose the anti-commutation relations

$$\left\{\boldsymbol{a}^r_{\vec{p}},\left(\boldsymbol{a}^s_{\vec{q}}\right)^\dagger\right\}=\left\{\boldsymbol{b}^r_{\vec{p}},\left(\boldsymbol{b}^s_{\vec{q}}\right)^\dagger\right\}=(2\pi)^3\,\delta^{(3)}(\vec{p}-\vec{q})\,\delta^{rs}\tag{5.21}$$

on the creation and annihilation operators. All other anticommutators are zero. Finally, the quantum Hamiltonian becomes

$$\boldsymbol{H} = \int \frac{d^3\vec{p}}{(2\pi)^3} \sum_{s=1}^{2} \omega_{\vec{p}} \left(\left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger \boldsymbol{a}^s_{\vec{p}} + \left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger \boldsymbol{b}^s_{\vec{p}} \right) \tag{5.22}$$

up to an infinite shift. It shows that it is the same as the Hamiltonian first guessed for the bosonic Dirac field except for the negative sign. Thus, $\boldsymbol{H}$ is a positive operator and has a ground state $|\Omega\rangle$ defined by $\boldsymbol{a}^s_{\vec{p}}\,|\Omega\rangle = \boldsymbol{b}^s_{\vec{p}}\,|\Omega\rangle = 0$.

5.5 Unitary Representation of the Poincaré Group

This gives a Hilbert space $\mathfrak{H}$ and an operator $\boldsymbol{H}$ generating time translations. However, to call this a relativistic quantum field theory a full representation of the Poincaré group on the Hilbert space of this theory is needed. Only the generator of the time translation has been determined so far.

The generators of the spatial translations are the momentum operators of the Dirac field, and with the momentum density $\boldsymbol{\psi}^\dagger(-i\,\vec{\nabla})\boldsymbol{\psi}$ the momentum operators are

$$\vec{\boldsymbol{p}} = \int d^3\vec{x}\; \boldsymbol{\psi}^\dagger(\vec{x})(-i\,\vec{\nabla})\boldsymbol{\psi}(\vec{x}) = \int \frac{d^3\vec{p}}{(2\pi)^3} \sum_{s=1}^{2} \left[\vec{p} \left\{ \left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger \boldsymbol{a}^s_{\vec{p}} + \left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger \boldsymbol{b}^s_{\vec{p}} \right\} \right]$$

similar to the Klein-Gordon case. The commutation relations are $[\boldsymbol{p}_j, \boldsymbol{H}] = 0$ for the components of $\vec{\boldsymbol{p}}$.

Thus, $\left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger$ and $\left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger$ create particles of energy $\omega_{\vec{p}}$ and momentum $\vec{p}$ because due to

$$\boldsymbol{H}\Big(\left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger |\Omega\rangle \Big) = \omega_{\vec{p}} \left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger |\Omega\rangle \qquad\qquad \boldsymbol{p}_j \Big(\left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger |\Omega\rangle \Big) = p_j \left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger |\Omega\rangle$$

the single particle eigenstates have the right energy-momentum 4-vector. The representation theory of the Lorentz group in the Dirac field is given as a shortcut only by introducing normalized single particle states

$$|\vec{p}, s\rangle = \sqrt{2\omega_{\vec{p}}} \left(\boldsymbol{a}^s_{\vec{p}}\right)^\dagger |\Omega\rangle$$

although these states are actually not really normalizable but are called here normalized because of the definition of the inner product

$$\langle \vec{p}, r | \vec{q}, s \rangle = 2\omega_{\vec{p}}\, (2\pi)^3\, \delta^{(3)}(\vec{p} - \vec{q}) \delta^{rs}$$

in (3.12) which is Lorentz invariant. The unitary representation of the Lorentz group is just defined via

$$\boldsymbol{U}(\Lambda)\, \boldsymbol{a}^s_{\vec{p}}\, \boldsymbol{U}^\dagger(\Lambda) = \sqrt{\frac{\omega_{\vec{\Lambda p}}}{\omega_{\vec{p}}}}\, \boldsymbol{a}^s_{\vec{\Lambda p}} \qquad\qquad \boldsymbol{U}(\Lambda)\, \boldsymbol{b}^s_{\vec{p}}\, \boldsymbol{U}^\dagger(\Lambda) = \sqrt{\frac{\omega_{\vec{\Lambda p}}}{\omega_{\vec{p}}}}\, \boldsymbol{b}^s_{\vec{\Lambda p}}$$

and not deduced from Noether's theorem as above in the case of the Klein-Gordon field. The expression $\vec{\Lambda p}$ means to take the three spatial components as a 3-vector from the 4-vector ΛP. This gives the full unitary representation of the Poincaré group because the Fock space is generated by products of these operators on the vacuum.

To calculate the actions of the Poincaré transformations on the spinor field itself one can insert these definitions into (5.20)

$$\begin{aligned}
\boldsymbol{U}(\Lambda)\, \boldsymbol{\psi}(t, \vec{x})\, \boldsymbol{U}^\dagger(\Lambda) &= \boldsymbol{U}(\Lambda) \left\{ \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \sum_{s=1}^{2} \left(\boldsymbol{a}^s_{\vec{p}}\; u^s(\vec{p})\, e^{-i\,\vec{p}\cdot\vec{x}} + \left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger v^s(\vec{p})\, e^{i\,\vec{p}\cdot\vec{x}} \right) \right\} \boldsymbol{U}^\dagger(\Lambda) \\
&= \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} \sqrt{2\omega_{\vec{p}}} \sum_{s=1}^{2} \left(\sqrt{\frac{\omega_{\vec{\Lambda p}}}{\omega_{\vec{p}}}}\, \boldsymbol{a}^s_{\vec{\Lambda p}}\, u^s(\vec{p})\, e^{-i\,\vec{p}\cdot\vec{x}} + \sqrt{\frac{\omega_{\vec{\Lambda p}}}{\omega_{\vec{p}}}} \left(\boldsymbol{b}^s_{\vec{\Lambda p}}\right)^\dagger v^s(\vec{p})\, e^{i\,\vec{p}\cdot\vec{x}} \right) \\
&= \int \frac{d^3\vec{q}}{(2\pi)^3} \frac{1}{2\omega_{\vec{q}}} \sqrt{2\omega_{\vec{q}}} \sum_{s=1}^{2} \left(\boldsymbol{a}^s_{\vec{q}}\, u^s(\Lambda^{-1}\vec{q})\, e^{-i\,\vec{q}\cdot\Lambda^{-1}\vec{x}} + \left(\boldsymbol{b}^s_{\vec{q}}\right)^\dagger v^s(\Lambda^{-1}\vec{q})\, e^{i\,\vec{q}\cdot\Lambda^{-1}\vec{x}} \right)
\end{aligned}$$

and change the integration variable from $P = (\omega_{\vec{p}}, \vec{p})$ to $\Lambda P = Q = (\omega_{\vec{q}}, \vec{q})$. Note that

$$\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}}$$

is invariant under Lorentz transformations. The result after a few additional steps is

$$\boldsymbol{U}(\Lambda)\, \boldsymbol{\psi}(X)\, \boldsymbol{U}^\dagger(\Lambda) = \Lambda_{\frac{1}{2}}^{-1} \boldsymbol{\psi}(\Lambda X)$$

and this is therefore the way the field operator transforms as a quantum operator. This is good evidence that this unitary operator is indeed a representation of the Lorentz group. The really convincing argument, however, needs a check that all the commutation relations of the Lie algebra are obeyed.

The angular momentum (as the generator of rotations) transforms $\psi(X) \to \psi'(X) = \Lambda_{\frac{1}{2}} \psi(\Lambda^{-1} X)$ where Λ is a rotation. An infinitesimal rotation in the xy-plane (about the z- or 3-axis) is

$$\Lambda_{\frac{1}{2}} \approx \mathbb{I} - \frac{i}{2}\, \omega_{\mu\nu}\, S^{\mu\nu} = \mathbb{I} - \frac{i}{2}\, \theta \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} = \mathbb{I} - \frac{i}{2}\, \theta\, \Sigma^3$$

where a quantum version of Σ^3 is needed. This gives applied to a spinor

$$\begin{aligned} \delta\psi(X) = \psi'(X) - \psi(X) &= (\mathbb{I} - \frac{i}{2}\, \theta\, \Sigma^3)\, \psi(t, x + \theta y, y - \theta x, z) - \psi(X) \\ &= -\theta\, (x\, \partial_y - y\, \partial_x + \frac{1}{2}\Sigma^3)\, \psi(X) \equiv \theta\, \Delta\phi(X) \end{aligned}$$

with Taylor expansion to first order. The time component of the conserved current is

$$j^0 = \frac{\partial \mathcal{L}}{\partial(\partial_0 \psi)}\, \Delta\psi = -i\, \bar{\psi}\, \gamma^0 \left(x\, \partial_y - y\, \partial_x + \frac{1}{2}\Sigma^3 \right) \psi$$

from applying Noether's theorem. Integration of j^0 over the 3-dimensional space gives the generator of rotations about the 3-axis such that

$$J_k = \int d^3\vec{x}\ \psi^\dagger \left(\left[\vec{x} \times (-i\, \vec{\nabla}) \right]_k + \frac{1}{2}\Sigma^k \right) \psi$$

are the generators of the rotation about the k-axis for $k = 1, 2, 3$. There are two contributing components where the Klein-Gordon case only had the cross-product term due to orbital angular momentum. The contribution of Σ^k in the Dirac case is due to spin. The quantum generators of rotations are

$$\boldsymbol{J} = \int d^3\vec{x}\ \boldsymbol{\psi}^\dagger(t, \vec{x}) \left(\vec{x} \times (-i\, \vec{\nabla}) + \frac{1}{2}\vec{\Sigma} \right) \boldsymbol{\psi}(t, \vec{x}) \tag{5.23}$$

at equal times t. (Note that it is not easy to write the angular momentum operator in terms of creation and annihilation operators.)

To complete the proof that the theory of the free quantum Dirac field is Lorentz invariant one has to show that all the generators of the Poincaré group including the four generators of the translations, the three generators of the rotations and the three generators of the boosts obey the right commutation relations from the Lie algebra of the Poincaré group. However, this is not done here because it is a rather difficult task for the fermions.

5.6 The Spin of a Dirac Particle

The expectation value of the three angular momentum operators $\boldsymbol{J}$ in (5.23) for $\vec{p} = \vec{0}$ can be calculated using the identity $[AB, C] = ABC - CAB = A\{B, C\}$ if $\{AC\} = 0$ and using

$$\left\{ \boldsymbol{\psi}(\vec{x}), \left(\boldsymbol{a}_{\vec{0}}^s\right)^\dagger \right\} = \frac{u^s(\vec{0})}{\sqrt{2\omega_{\vec{0}}}} = \frac{u^s(\vec{0})}{\sqrt{2m}}$$

in order to avoid having to express $\boldsymbol{J}$ in terms of creation and annihilation operators. The spin of the free Dirac particle as the angular momentum at rest is

$$\begin{aligned}\boldsymbol{J}\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}|\Omega\rangle &= \left[\boldsymbol{J},\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}\right]|\Omega\rangle = \int d^{3}\vec{x}\left[\boldsymbol{\psi}^{\dagger}(\vec{x})\left(\vec{x}\times(-i\,\vec{\nabla})+\frac{1}{2}\vec{\Sigma}\right)\boldsymbol{\psi}(\vec{x}),\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}\right]|\Omega\rangle \\ &= \int d^{3}\vec{x}\;\boldsymbol{\psi}^{\dagger}(\vec{x})\left(\vec{x}\times(-i\,\vec{\nabla})+\frac{1}{2}\vec{\Sigma}\right)\left\{\boldsymbol{\psi}(\vec{x}),\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}\right\}|\Omega\rangle \\ &= \int d^{3}\vec{x}\;\boldsymbol{\psi}^{\dagger}(\vec{x})\left(\vec{x}\times(-i\,\vec{\nabla})+\frac{1}{2}\vec{\Sigma}\right)\frac{u^{s}(\vec{0})}{\sqrt{2m}}|\Omega\rangle = \int d^{3}\vec{x}\;\boldsymbol{\psi}^{\dagger}(\vec{x})\left(\frac{1}{2}\vec{\Sigma}\frac{u^{s}(\vec{0})}{\sqrt{2m}}\right)|\Omega\rangle\end{aligned}$$

where $\vec{x}\times(-i\,\vec{\nabla})$ and $\vec{\Sigma}$ are 4×4 matrices and $u^{s}(\vec{0})$ is a 4×1 spinor. As a differential operator applied to something that does not depend on X, the orbital part $\vec{x}\times(-i\,\vec{\nabla})$ of angular momentum is zero. One can do the integral

$$\int d^{3}\vec{x}\;\boldsymbol{\psi}^{\dagger}(\vec{x}) = \frac{1}{\sqrt{2m}}\left(\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}\left(u^{s}(\vec{0})\right)^{\dagger}+\left(\boldsymbol{b}_{\vec{0}}^{s}\right)\left(v^{s}(\vec{0})\right)^{\dagger}\right)$$

because the other terms do not depend on X. With $(v^{r}(\vec{0}))^{\dagger}\,(\frac{1}{2}\vec{\Sigma})\,u^{s}(\vec{0})=0$ for all r,s

$$\begin{aligned}\boldsymbol{J}^{k}\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}|\Omega\rangle &= \sum_{r=1}^{2}\left(\left(u^{r}(\vec{0})\right)^{\dagger}\frac{\Sigma^{k}}{2}\frac{u^{s}(\vec{0})}{\sqrt{2m}}\right)\left(\boldsymbol{a}_{\vec{0}}^{r}\right)^{\dagger}|\Omega\rangle \\ &\overset{k=3}{=} \sum_{r=1}^{2}\left(\left(\xi^{r}\right)^{\dagger}\frac{\sigma^{3}}{2}\,\xi^{s}\right)\left(\boldsymbol{a}_{\vec{0}}^{r}\right)^{\dagger}|\Omega\rangle = \frac{1}{2}(-1)^{s-1}\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}|\Omega\rangle\end{aligned}$$

the eigenvalue equation for the z component becomes

$$\boldsymbol{J}^{3}\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}|\Omega\rangle = \frac{1}{2}(-1)^{s-1}\left(\boldsymbol{a}_{\vec{0}}^{s}\right)^{\dagger}|\Omega\rangle \tag{5.24}$$

showing that the spin is either $+\frac{1}{2}$ or $-\frac{1}{2}$.

There is a fundamental connection between spin and statistics. The spin-statistics theorem states that a particle excitation is a fermion if it has half-integer spin and that a particle excitation is a boson if it has an integer spin. Thus, the fact that a half-integer spin has been found here is not a consequence of how the field has been quantized but is a fundamental property coming from representation theory.

5.7 Solving the Quantum Dirac Field

In order to solve the quantum Dirac field the propagator is needed together with Wick's theorem. The equation for the covariance metric

$$\begin{aligned}\langle\Omega|\boldsymbol{\psi}_{a}(X)\,\bar{\boldsymbol{\psi}}_{b}(Y)|\Omega\rangle &= \int\frac{d^{3}\vec{p}}{(2\pi)^{3}}\frac{1}{2\omega_{\vec{p}}}\sum_{s=1}^{2}u_{a}^{s}(\vec{p})\,\bar{u}_{b}^{s}(\vec{p})\,e^{-i\,P\cdot(X-Y)} \\ &= (i\,\not{\partial}_{X}+m)_{ab}\int\frac{d^{3}\vec{p}}{(2\pi)^{3}}\frac{1}{2\omega_{\vec{p}}}\,e^{-i\,P\cdot(X-Y)}\end{aligned}$$

captures all the information about the dynamics of this quantum field theory. The slash notation means $\not{A}\equiv A_{\mu}\,\gamma^{\mu}$ summed over μ according to Einstein's summation convention for a 4-vector A and therefore $\not{A}$ is a 4×4 matrix. Note that in $\not{\partial}_{X}$ the subscript X indicates that the derivative takes place with respect to X and not with respect to Y. Since $(i\,\not{\partial}_{X}+m)_{ab}$ with $a,b\in\{1,2,3,4\}$ is a 4×4 matrix and the integral is a scalar the equation above consists of 16 equations. Also

$$\langle\Omega|\bar{\boldsymbol{\psi}}_{b}(Y)\,\boldsymbol{\psi}_{a}(X)|\Omega\rangle = (-i\,\not{\partial}_{X}+m)_{ab}\int\frac{d^{3}\vec{p}}{(2\pi)^{3}}\frac{1}{2\omega_{\vec{p}}}\,e^{-i\,P\cdot(Y-X)}$$

can be calculated similarly.

The field operators (5.20) have been defined for a fixed equal time slice but they have to be extended to different times for the above two results. To solve the Schrödinger equation

$$\frac{\partial \boldsymbol{\psi}(X)}{\partial t} = i\,[\boldsymbol{H}, \boldsymbol{\psi}(X)]$$

the field operators at X become

$$\boldsymbol{\psi}(X) = \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \sum_{s=1}^{2} \left(\boldsymbol{a}^s_{\vec{p}}\, u^s(P)\, e^{-i\,P\cdot X} + \left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger v^s(P)\, e^{i\,P\cdot X} \right)$$

because of $\left[\boldsymbol{H}, \left(\boldsymbol{a}^s_{\vec{p}}\right)\right] = \omega_{\vec{p}} \left(\boldsymbol{a}^s_{\vec{p}}\right)$ and $\left[\boldsymbol{H}, \left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger\right] = -\omega_{\vec{p}} \left(\boldsymbol{b}^s_{\vec{p}}\right)^\dagger$.

The next task is to construct retarded Green's functions

$$S^{\text{ret}}_{ab}(X-Y) = \theta(X^0 - Y^0)\, \langle \Omega | \boldsymbol{\psi}_a(X)\, \bar{\boldsymbol{\psi}}_b(Y) | \Omega \rangle = (i\partial\!\!\!/_X + m)\, D^{\text{ret}}(X-Y)$$

with the bosonic propagator (using the fact $\partial\!\!\!/\,\partial\!\!\!/ = \Box$). One could also use Fourier transform to solve

$$(i\partial\!\!\!/_X - m)\, S^{\text{ret}} = i\,\delta^{(4)}(X-Y) \cdot \mathbb{I}_{4\times 4}$$

because a Green's function is a solution to a differential equation in the presence of a delta function on the right-hand side. In Fourier space this is

$$i\,\delta^{(4)}(X-Y) \cdot \mathbb{I}_{4\times 4} = \int \frac{d^4P}{(2\pi)^4} (P\!\!\!/ - m)\, e^{-i\,P\cdot(X-Y)}\, \tilde{S}^{\text{ret}}(P)$$

such that by using $P\!\!\!/^2 = P^2$

$$\tilde{S}^{\text{ret}}(P) = \frac{i}{P\!\!\!/ - m} = \frac{i(P\!\!\!/ + m)}{P^2 - m^2}$$

is the Fourier transform of the retarded propagator. It is a 4×4 matrix. One can do the same for the advanced Green's function. It turns out that one does not need the advanced or retarded Green's function for Wick's theorem but a combination of both. Using $\tilde{S}^{\text{ret}}$ and $\tilde{S}^{\text{adv}}$ one can define

$$S_{\text{F}}(X-Y) = \lim_{\varepsilon\to 0} \int \frac{d^4P}{(2\pi)^4} \frac{i(P\!\!\!/ + m)}{P^2 - m^2 + i\varepsilon}\, e^{-i\,P\cdot(X-Y)} = \langle \Omega \,|\, \mathcal{T}\left\{ \boldsymbol{\psi}(X)\, \bar{\boldsymbol{\psi}}(Y) \right\} |\, \Omega \rangle \tag{5.25}$$

as the Feynman propagator for the quantum Dirac field.

6 Quantum Field Theory for Interacting Fermions and Bosons

6.1 Observables and Predictions

The observables one has access to in quantum field theory by scattering experiments are the time-ordered correlation functions

$$G^{(n)}_{\vec{\alpha}}(X_1, ..., X_n) = \langle \Omega \,|\, \mathcal{T}\left[\boldsymbol{\Phi}_{\alpha_1}(X_1) ... \boldsymbol{\Phi}_{\alpha_n}(X_n) \right] |\, \Omega \rangle \tag{6.1}$$

in the Heisenberg picture where Ω is the full interacting vacuum and $\boldsymbol{\Phi}_\alpha(X)$ is a list of all quantum fields for $\alpha \in \{1, ..., n\}$ at various spacetime locations. A theory of scalar bosons and fermions, for example, consists of $\boldsymbol{\Phi}_0(X) = \boldsymbol{\phi}(X)$ for a scalar boson and $\boldsymbol{\Phi}_\alpha(X) = \boldsymbol{\psi}_\alpha(X)$ for the four fields $\alpha \in \{1, 2, 3, 4\}$ of a Dirac spinor. Of course, there are more complicated cases with multiple scalar boson fields as in quantum electrodynamics where a vector boson field represents the electromagnetic field or with more than one type of Dirac spinors as in quantum chromodynamics.

A statistical physical theory has a set of observables and gives a set of predictions which are expectation values for these observables. This is already built in the equation (6.1). Making predictions can either be

performed using computer simulations on a lattice or using perturbation theory in the interaction picture with $\boldsymbol{H} = \boldsymbol{H}_{\text{free}} + \boldsymbol{H}_{\text{int}}$ where $\boldsymbol{H}_{\text{free}}$ consists of copies of the free scalar bosons or the free Dirac fermions. The free Hamiltonian, for example, could be

$$\boldsymbol{H}_{\text{free}} = \frac{1}{2}\int d^3\vec{x}\ \boldsymbol{\pi}(\vec{x})^2 + \left(\nabla\boldsymbol{\phi}(\vec{x})\right)^2 + m^2\,\boldsymbol{\phi}(\vec{x})^2 + \int d^3\vec{x}\ \boldsymbol{\psi}^\dagger\left(-i\,\gamma^0\,\vec{\gamma}\cdot\vec{\nabla} + m\,\gamma^0\right)\boldsymbol{\psi} \tag{6.2}$$

where the first integral comes from the Klein-Gordon Hamiltonian (3.3) and the second one from the quantized version of the Dirac Hamiltonian in (5.12). The interaction part could look like

$$\boldsymbol{H}_{\text{int}} = \int d^3\vec{x}\,\frac{\lambda}{4!}\boldsymbol{\phi}^4 + g\,\boldsymbol{\phi}(X)\bar{\boldsymbol{\psi}}(X)\boldsymbol{\psi}(X) + ... + \left(\bar{\boldsymbol{\psi}}(X)\boldsymbol{\psi}(X)\right)^2 + ...$$

and its terms have the form of point interactions of low power due to locality. An interaction part with only the first term would be boring because only the bosons would interact with each other but there would be no interactions between the bosons and the fermions, and therefore bosons and fermions would be decoupled. A term such as $\boldsymbol{\phi}(X)\boldsymbol{\psi}(X)$ however is not a scalar and $\boldsymbol{\phi}(X)\bar{\boldsymbol{\psi}}(X)\boldsymbol{\psi}(X)$ turns out to be the simplest Lorentz scalar connecting bosons and fermions. A term such as $(\bar{\boldsymbol{\psi}}(X)\boldsymbol{\psi}(X))^2$ is interesting but the interaction Hamiltonian used in the following is for simplicity

$$\boldsymbol{H}_{\text{int}} = \int d^3\vec{x}\,g\,\boldsymbol{\phi}(X)\bar{\boldsymbol{\psi}}(X)\boldsymbol{\psi}(X) \tag{6.3}$$

with the coupling constant g. One cannot measure the fermion field by itself but only densities $\bar{\boldsymbol{\psi}}(X)\boldsymbol{\psi}(X)$ because the fermion field has four components and is not Hermitian. The first term with $\boldsymbol{\phi}^4$ is not included because it has been studied in the context of the ϕ^4-theory above.

Equation (4.5) allows to relate the interacting quantities and the free quantities with the help of the S-matrix. The Dyson series for $\boldsymbol{U}(t,-t)$ gives

$$\boldsymbol{U}(t,-t) = \mathbf{1} - i\int_{-t}^{t} dt'\,\boldsymbol{H}_{\text{int}}(t') + ...$$

where this infinite series of the interacting term of the Hamiltonian evolved in the interaction picture according to the free propagator. At that point the problem has just been reduced to summing up an infinite series of progressively more difficult terms. By substituting $\boldsymbol{U}$ into S defined in (4.5) one can calculate order by order. However, the Hamiltonian (6.3) is too strong as an assumption for the degrees of freedom of the universe. To get rid of the resulting infinities one has to introduce a cutoff Λ and replace $\boldsymbol{H}_{\text{int}}$ by $\boldsymbol{H}_{\text{int}}(\Lambda)$. The higher order terms of the Dyson series are time-ordered products of $\boldsymbol{H}_{\text{int}}$ of the form (4.2) and therefore terms that appear in series for $G^{(n)}$ are all time-ordered products of field operators $\langle 0\,|\,\mathcal{T}[\boldsymbol{\Phi}(X_1)...\boldsymbol{\Phi}(X_n)]\,|\,0\rangle$ and integrals thereof in the interaction picture. In order to calculate these terms Wick's theorem must be generalized to fermions.

6.2 Generalization of Wick's Theorem to Fermions

The Dirac field is a quadratic theory which is completely solved with a diagonalized Hamiltonian and one would expect that there is a way to calculate time-ordered products of Dirac field operators. The difference between bosons and fermions are the anticommutation relations instead of the commutation relations. Thus, whenever one moves operators past each other there is an additional minus sign, and the time-ordering symbol

$$\langle 0\,|\mathcal{T}[\boldsymbol{\psi}_\alpha(X)\,\bar{\boldsymbol{\psi}}_\beta(Y)]|\,0\rangle = \begin{cases}\boldsymbol{\psi}_\alpha(X)\,\bar{\boldsymbol{\psi}}_\beta(Y) & \text{if } X^0 \geq Y^0\\ -\bar{\boldsymbol{\psi}}_\beta(Y)\,\boldsymbol{\psi}_\alpha(X) & \text{if } X^0 < Y^0\end{cases} \tag{6.4}$$

is similar to the one for bosons (3.21) but has a minus sign. However, these are actually 16 equations because of the indices α and β. The Feynman propagator

$$S_{\text{F}}(X,Y) = \int \frac{d^4P}{(2\pi)^4}\,\frac{i(\not{P}+m}{P^2-m^2+i\varepsilon}\,e^{-i\,P\cdot(X-Y)} = \langle 0\,|\mathcal{T}[\boldsymbol{\psi}(X)\,\bar{\boldsymbol{\psi}}(Y)]|\,0\rangle = \overbrace{\boldsymbol{\psi}(X)\bar{\boldsymbol{\psi}}}(Y)$$

as defined in (5.25) obviously consists of 16 equations as well such that one should write $S_{\rm F}^{\alpha\beta}(X,Y)$. Feynman propagators are important because they turn up when one works with a Dyson series.

Wick's theorem converts time-ordered products into normal-ordered products and this is very useful because something in normal order becomes instantly zero when put into a vacuum expectation value. In order to generalize Wick's theorem for fermions the time order operator $\mathcal{T}$ has to be generalized first

$$\mathcal{T}\left\{\psi(X_1)\,\bar{\psi}(X_2)\,...\,\psi(X_n)\right\} = (-1)^{\mathrm{sgn}(\pi)}\,\psi(X_{\pi^{-1}(1)})\,\bar{\psi}(X_{\pi^{-1}(2)})\,...\,\psi(X_{\pi^{-1}(n)})$$

to handle many field operators such that $X^0_{\pi^{-1}(1)} > X^0_{\pi^{-1}(2)} > ... > X^0_{\pi^{-1}(n)}$ where the different field operators may have a bar or no bar. Thus, the goal is to find the permutation π that puts the terms into the right order and to determine the sign $\mathrm{sgn}(\pi)$ of π. Normal ordering for fermions is also defined similarly as for bosons. The example

$$\mathcal{N}\left\{\boldsymbol{a}_{\vec{p}}\,\boldsymbol{a}_{\vec{q}}\,\boldsymbol{a}^\dagger_{\vec{r}}\right\} = (-1)\,\boldsymbol{a}_{\vec{p}}\,\boldsymbol{a}^\dagger_{\vec{r}}\,\boldsymbol{a}_{\vec{q}} = (-1)^2\,\boldsymbol{a}^\dagger_{\vec{r}}\,\boldsymbol{a}_{\vec{p}}\,\boldsymbol{a}_{\vec{q}} = (-1)^3\,\boldsymbol{a}^\dagger_{\vec{r}}\,\boldsymbol{a}_{\vec{q}}\,\boldsymbol{a}_{\vec{p}}$$

shows that each move of an operator past a neighbor changes the sign.

The base case for Wick's theorem with two field operators is

$$\mathcal{T}\left\{\psi(X)\,\bar{\psi}(Y)\right\} = \mathcal{N}\left\{\psi(X)\,\bar{\psi}(Y)\right\} + \overbrace{\psi(X)\bar{\psi}}(Y)$$

where the contraction can be written as

$$\overbrace{\psi(X)\bar{\psi}}(Y) = S_{\rm F}(X,Y) = \begin{cases} \left\{\psi^+(X), \bar{\psi}^-(Y)\right\} & \text{if } X^0 \geq Y^0 \\ -\left\{\bar{\psi}^+(Y), \psi^-(X)\right\} & \text{if } X^0 < Y^0 \end{cases}$$

explicitly in terms of the positive and negative frequency parts of the field operators. The contractions

$$\overbrace{\psi(X)\psi}(Y) = \overbrace{\bar{\psi}(X)\bar{\psi}}(Y) = 0$$

are all zero. One can extend normal ordering to account for operator interchange such as

$$\mathcal{N}\left\{\overbrace{\psi(X_1)\psi(X_2)\bar{\psi}(X_3)\bar{\psi}}(X_4)\right\} \overset{-1}{\to} \mathcal{N}\left\{\overbrace{\psi(X_1)\psi(X_2)\bar{\psi}(X_4)\bar{\psi}}(X_3)\right\} \overset{-1}{\to} \mathcal{N}\left\{\overbrace{\psi(X_1)\bar{\psi}}(X_4)\psi(X_2)\bar{\psi}(X_3)\right\}$$

where each move of a field operator past a neighbor gives a minus sign. Thus, this means

$$\begin{aligned}\mathcal{N}\left\{\overbrace{\psi(X_1)\psi(X_2)\bar{\psi}}(X_3)\bar{\psi}(X_4)\right\} &= (-1)\overbrace{\psi(X_1)\bar{\psi}}(X_3)\,\mathcal{N}\left\{\psi(X_2)\bar{\psi}(X_4)\right\} \\ &= (-1)\,S_{\rm F}(X_1,X_3)\,\mathcal{N}\left\{\psi(X_2)\bar{\psi}(X_4)\right\}\end{aligned}$$

for one interchange of neighboring field operators.

The general case of Wick's theorem for fermions is

$$\mathcal{T}\left\{\psi(X_1)\,\bar{\psi}(X_2)\,...\right\} = \mathcal{N}\left\{\psi(X_1)\,\bar{\psi}(X_2)\,... + \text{all contractions}\right\}$$

and states therefore as in the case of bosons that time ordering is normal ordering plus all possible contractions. The difference is the additional step that one has to move the contracted operators next to each other in order to evaluate a normal ordered contraction.

The schematic of a full quantum field calculation

$$\begin{aligned}G^{(n)}_{\vec{\alpha}}(X_1,...,X_n) &= \frac{\langle 0\,|\,\mathcal{T}\left[\boldsymbol{\Phi}_{\alpha_1}(X_1)...\boldsymbol{\Phi}_{\alpha_n}(X_n)\,\mathsf{S}\right]|\,0\rangle}{\langle 0\,|\,\mathsf{S}|\,0\rangle} \overset{\text{Dyson}}{=} \frac{\langle 0\,|\boldsymbol{\Phi}_{\alpha_1}(X_1)...\left(\mathbf{1} - i\int dt'...\right)|\,0\rangle}{\langle 0\,|\mathbf{1} - i\int dt'...\,|\,0\rangle} \\ &\overset{\text{Wick}}{=} 1 + \int dZ...\overbrace{\boldsymbol{\Phi}(Z)}\,\boldsymbol{\Phi}(Z) + ... + \iiint \overbrace{\boldsymbol{\Phi}}\,\overbrace{\boldsymbol{\Phi}}\,\overbrace{\boldsymbol{\Phi}}\,\overbrace{\boldsymbol{\Phi}} + ...\end{aligned}$$

shows that one starts with the observables for the n-point functions, expresses them as time-ordered interaction picture observables with the S-matrix inside, uses Dyson to expand numerator and denominator, does some Taylor series manipulations and applies Wick's theorem to reduce the entire calculation to the integrals of normal ordered contractions. Finally, one draws Feynman diagrams to represent these integrals because they are much easier to cope with. The end result is a sum of integrals of contractions. However, the number of these integrals is enormous. To cope with all terms and to exploit patterns and cancellations one codifies them in terms of Feynman rules which change from theory to theory because they depend on $\boldsymbol{H}_{\rm int}$.

6.3 Feynman Rules and Feynman Diagrams for the Yukawa Theory

The Yukawa theory is the simplest theory in which one type of bosons interacts with one type of fermions and where the Hamiltonian consists of the free Klein-Gordon and Dirac Hamiltonians (6.2) and the interaction Hamiltonian (6.3). There are bosonic and fermionic propagators, and there are always two fermionic edges and one bosonic edge per internal vertex where dashed lines represent bosonic and solid lines fermionic edges. The mass of the bosons is m_ϕ, and s and m are spin and mass of the fermion.

The Feynman rules in momentum space are:

1. Propagators
   - boson: $\overbrace{\phi(X)\,\phi(Y)} = \dots\dots\dots\rightarrow_{P}\dots\dots\dots = \frac{i}{P^2 - m_\phi^2 + i\varepsilon}$
   - fermion: $\overbrace{\psi(X)\,\bar{\psi}(Y)} = \longrightarrow_{P} = \frac{i(\not{P} + m)}{P^2 - m^2 + i\varepsilon}$
2. Vertices: $= -ig$
3. External legs
   - boson: $\overbrace{\phi\,|\,Q\rangle} = \dots\dots\leftarrow_{Q} = 1$ $\qquad \overbrace{\langle Q\,|\,\phi} = \dots\dots\leftarrow_{Q} = 1$
   - fermion: $\overbrace{\psi\,|\,P,s\rangle} = \longleftarrow_{P} = u^s(P)$ $\qquad \overbrace{\langle P,s\,|\,\bar{\psi}} = \longleftarrow_{P} = \bar{u}^s(P)$
   - antifermion: $\overbrace{\bar{\psi}\,|\,K,s\rangle} = \longleftarrow_{K} = \bar{v}^s(K)$ $\qquad \overbrace{\langle K,s\,|\,\psi} = \longrightarrow_{K} = v^s(K)$
4. Impose momentum conservation at each vertex.
5. Integrate over loop momenta.
6. Calculate the signs of the diagram by keeping track of factors (-1).
7. Divide by symmetry factor.

According to these rules one can calculate the n-point Green's function $G^{(n)}_{\vec{\alpha}}(X_1, ..., X_n)$ as the sum of all connected and amputated Feynman diagrams with n external legs subject to these Feynman rules.

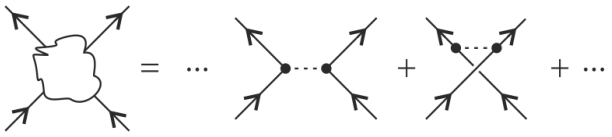


Figure 12: Fermion scattering process in the Yukawa theory

If one applies these rules to a fermion scattering process where two fermions collide and separate in the form of two fermions as illustrated in figure 12 one encounters many diagrams with four legs. Only two Feynman diagrams are shown to highlight the role of the sign of the diagram. The diagram on the left side is an integral over two variables X and Y but they are actually fixed by momentum conservation such that there is nothing to integrate over. The diagram on the right side is basically the same diagram but there is a factor (-1) coming from Wick's theorem due to the crossing legs.