Quantum Field Theory

Lecture Notes

Lecturer: Petr Jizba
Notes by:

Preface

Here will come the preface...

Detailed Contents

1 Non-Relativistic Wave Function

1.1 Transformations

In quantum mechanics the physical state of a particle is represented by wave function $\psi(\boldsymbol{x}, t)$. In the Schrödinger picture we have $\psi(\boldsymbol{x}, t) = \langle \boldsymbol{x}|\psi(t)\rangle$, i.e. time evolution in this picture is contained in the *state vectors* and *base vectors* are time independent (basis is rigid). On the other hand, in the Heisenberg picture we have $\psi(\boldsymbol{x}, t) = \langle \boldsymbol{x}, t|\psi\rangle$, i.e. *state vectors* are time independent and *base vectors* evolve (basis evolves in time). Consequently, in the Schrödinger picture the dynamics is given by the Schrödinger equation, which prescribes evolution of state vectors, while in the Heisenberg picture the evolution is given by the Heisenberg equation which prescribes the evolution of the complete set of observables (i.e., Hermitian operators). The later, in turn, imply the time dependent base vectors (via corresponding eigenfunctions).

Choice of the representation is in quantum mechanics simply matter of convenience (though work with Schrödinger equation is often simpler due to its linearity) since these representations are unitarily equivalent. This is essence of the so-called Stone–von Neumann uniqueness theorem. We will see that the S-vN theorem is typically broken in quantum theories with infinitely many degrees of freedom. This will have important consequences for the entire structure of Quantum Field Theory (e.g., renormalization, non-trivial vacuum condensates, etc.).

Let us now explore behaviour of the quantum-mechanical wave function under two important transformations, namely **rotation**

$$\psi(\boldsymbol{x}) \stackrel{\boldsymbol{R}}{\rightarrow} \psi_{\boldsymbol{R}}(\boldsymbol{x}) = \psi(\boldsymbol{R}^{-1}\boldsymbol{x})\,, \tag{1.1}$$

and **translation**

$$\psi(\boldsymbol{x}) \stackrel{\boldsymbol{a}}{\rightarrow} \psi_{\boldsymbol{a}}(\boldsymbol{x}) = \psi(\boldsymbol{x} - \boldsymbol{a})\,. \tag{1.2}$$

Here we have omitted time argument as it is immaterial for this discussion. Time argument will, however, be important in the next chapter where relativistic transformation of state vectors will be considered.

In this text we will deal with active transformation only. In those transformation the base space is fixed and only states change in rigid basis.

Symmetry transformations are for (compact) groups implemented via unitary operations:

$$\psi_{\boldsymbol{a}}(\boldsymbol{x}) = U_{\boldsymbol{a}}(\boldsymbol{a})\psi(\boldsymbol{x}) = \psi(\boldsymbol{x} - \boldsymbol{a}), \tag{1.3}$$

$$\psi_{\boldsymbol{R}}(\boldsymbol{x}) = U_{\boldsymbol{R}}(\theta)\psi(\boldsymbol{x}) = \psi(\boldsymbol{R}^{-1}(\theta)\boldsymbol{x}). \tag{1.4}$$

If particles have spin (or other internal indices associated with them) i.e.

$$\psi(\boldsymbol{x}) \rightarrow \psi_{\alpha}(\boldsymbol{x}),\ \alpha \in \boldsymbol{I}\,. \tag{1.5}$$

In this case such wave function will change under rotation according to

$$\psi_\alpha(\boldsymbol{x}) \xrightarrow{\boldsymbol{R}} \psi_{\boldsymbol{R},\alpha}(\boldsymbol{x}) = D_{\alpha\beta}(\boldsymbol{R})\psi_\beta(\boldsymbol{R}^{-1}\boldsymbol{x})\,, \tag{1.6}$$

here $D(\boldsymbol{R})$ is an appropriate representation of the group element $\boldsymbol{R} \in$ SO(3), which acts on indices.

Example 1.1.1 For example, consider spin-$\frac{1}{2}$ particle, in this case wave function index takes values $\alpha = \pm\frac{1}{2}$. From quantum mechanics we know, that our transformation can be written as

$$D(\boldsymbol{R}) = e^{-i\theta\boldsymbol{n}\cdot\boldsymbol{s}}\,,$$

which is a matrix that acts on Pauli spinors. Here θ is the angle of rotation, $\boldsymbol{n}$ is axis of the rotation, and $\boldsymbol{s} = \frac{1}{2}\boldsymbol{\sigma} = \frac{1}{2}(\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli matrices, which are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1. \end{pmatrix}$$

In this representation $s_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is diagonal, and so α has interpretation of being the (eigen)value of s_3.

2 Relativistic Wave Function

2.1 Relativistic Conventions

In this text we will assume so-called **natural units**, in which $c = \hbar = 1$. In this system, E, p have units of $\frac{1}{\text{lenght}} = \frac{1}{\text{time}}$. Also the following relativistic conventions are used:

- Space-time 4-vector will be denoted by $x^\mu = (x^0, \boldsymbol{x}) = (ct, \boldsymbol{x})$.
- 4-momentum will be denoted by $p^\mu = (p^0, \boldsymbol{p}) = (E, \boldsymbol{p})$.
- Scalar product is given by

$$a \cdot b = \boldsymbol{g}_{\mu\nu} a^\mu b^\nu = a^\mu b_\mu = a_\mu b^\mu, \tag{2.1}$$

where $\boldsymbol{g}_{\mu\nu} = \boldsymbol{g}^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \text{diag}(1,-1,-1,-1)$ is a metric tensor, with $\boldsymbol{g}^{\mu\nu}$ being inverse to the $\boldsymbol{g}_{\mu\nu}$. We can immediately derive a simple relation between metric tensors and Kronecker delta:

$$\boldsymbol{g}_{\mu\nu} \boldsymbol{g}^{\nu\sigma} = \delta_\mu{}^\sigma. \tag{2.2}$$

A **Lorentz transformation** $\boldsymbol{L}$ maps 4-vectors according to the relation

$$x^\mu \xrightarrow{\boldsymbol{L}} x'^\mu = \boldsymbol{L}^\mu{}_\nu x^\nu. \tag{2.3}$$

Here $\boldsymbol{L}^\mu{}_\nu \in \text{SO}(1,3)$, so called Lorentz group. From that we can see that

$$x^\nu = \boldsymbol{L}_\mu{}^\nu x'^\mu \tag{2.4}$$

Thus $\boldsymbol{L}_\mu{}^\nu$ and $\boldsymbol{L}^\mu{}_\nu$ are inverse to each other. From the fact that Lorentz transformation preserve scalar product (as can be easily shown by the reader), i.e. $a' \cdot b' = a \cdot b$, the following relations hold

$$\boldsymbol{g}^{\mu\mu'} \boldsymbol{L}^\mu{}_\nu \boldsymbol{L}^{\mu'}{}_{\nu'} = \boldsymbol{g}^{\nu\nu'}, \tag{2.5}$$

$$\boldsymbol{g}^{\mu\mu'} \boldsymbol{L}_\mu{}^\nu \boldsymbol{L}_{\mu'}{}^{\nu'} = \boldsymbol{g}^{\nu\nu'}. \tag{2.6}$$

By taking determinant of both sides we arrive to the fact that

$$\det^2 L = 1. \tag{2.7}$$

Hence we can classify two classes of Lorentz transformations - proper, for which $\det L = 1$ and improper, for which $\det L = -1$.

Differentials

We will define differential to be given by

$$\mathrm{d} \equiv \mathrm{d}x^{\mu}\frac{\partial}{\partial x^{\mu}}. \tag{2.8}$$

In relativistic notation we define $\partial_{\mu} \equiv \frac{\partial}{\partial x^{\mu}}$, then differential can be written as

$$\mathrm{d} = \mathrm{d}x^{\mu}\partial_{\mu}. \tag{2.9}$$

Such definition of d is Lorentz invariant (as can be shown from the fact that $\partial'_{\mu} = \boldsymbol{L}_{\mu}^{\ \nu}\partial_{\nu}$). We define covariant 4-gradient operator to be

$$\partial_{\mu} = \left(\frac{\partial}{\partial x_0}, \boldsymbol{\nabla}\right), \tag{2.10}$$

and contravariant operator to be

$$\partial^{\mu} = \left(\frac{\partial}{\partial x_0}, -\boldsymbol{\nabla}\right). \tag{2.11}$$

Using those we can define d'Alambertian operator via

$$\boldsymbol{g}^{\mu\nu}\partial_{\mu}\partial_{\nu} = \partial_{\mu}\partial^{\mu} = \left(\frac{\partial^2}{\partial x_0^2}, -\nabla^2\right) \equiv \Box. \tag{2.12}$$

In the context of special relativity it is common to denote the metric tensor $\boldsymbol{g}^{\mu\nu}$ as $\boldsymbol{\eta}^{\mu\nu}$ or simply $\eta^{\mu\nu}$.

2.2 Structure of Lorentz Transformation

We begin our study of Lorentz transformations by taking infinitesimal limit of such transformation

$$\boldsymbol{L}^{\mu}_{\ \nu} = \delta^{\mu}_{\ \nu} + \omega^{\mu}_{\ \nu}, \tag{2.13}$$

which can be thought of as infinitesimal deformation from identical transformation. Starting from (2.5) and raising indices by metric tensor we get

$$\boldsymbol{L}^{\mu\nu'}\boldsymbol{L}_{\mu}^{\ \nu} = \boldsymbol{g}^{\nu'\nu}. \tag{2.14}$$

Then acting on it by $\boldsymbol{g}_{\nu\alpha}$ from right

$$\boldsymbol{L}^{\mu\nu'}\boldsymbol{L}_{\mu\alpha} = \boldsymbol{g}^{\nu'}_{\ \alpha} = \delta^{\nu'}_{\ \alpha}. \tag{2.15}$$

This can be equivalently rewritten as

$$\boldsymbol{L}_{\mu}^{\ \nu'}\boldsymbol{L}^{\mu}_{\ \alpha} = \delta^{\nu'}_{\ \alpha}. \tag{2.16}$$

By using (2.13) we finally arrive to the following equation

$$(\delta_{\mu}^{\ \nu'} + \omega_{\mu}^{\ \nu'})(\delta^{\mu}_{\ \alpha} + \omega^{\mu}_{\ \alpha}) = \delta^{\nu'}_{\ \alpha}. \tag{2.17}$$

If we restrict ourselves only to the first order in ω

$$\delta^{\nu'}_{\alpha} + \omega_{\alpha}{}^{\nu'} + \omega^{\nu'}{}_{\alpha} = \delta^{\nu'}{}_{\alpha}. \tag{2.18}$$

Subtracting Kronecker deltas from both sides and lowering all indices one gets

$$\omega_{\alpha\nu'} + \omega_{\nu'\alpha} = 0. \tag{2.19}$$

This statement implies that ω is a 4×4 antisymmetric matrix, which has 6 independent parameters in the case of infinitesimal Lorentz transformation. This fact also holds for finite Lorentz transformations.

Properties of Lie groups

The transformation laws of continuous groups (Lie groups) such as rotation or Lorentz group are typically conveniently expressed in an infinitesimal form. By combining successively many infinitesimal transformations it is always possible to reconstruct from these the finite transformation laws. This is a consequence of the fact, that exponential function e^x can always be obtained by a product of many small-x approximations: $e^{\delta\alpha x} \approx 1 + \delta\alpha x$, where $\delta\alpha = \alpha/M,\ M \gg 1$. Taking multiple product of infinitesimal steps we obtain $(1+\alpha x/M)^M$. Here x are so-called **group generators**. The finite group transformation is then given by $L(\alpha) = e^{\alpha x}$. One can also reverse the process later on by taking $\left.\frac{\mathrm{d}L(\alpha)}{\mathrm{d}\alpha}\right|_{\alpha=0} = x$, which recovers the group generator.

Transitioning from infinitesimal to finite transformation we have the form

$$\boldsymbol{L}^{\rho}{}_{\tau} = \left(e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}\right)^{\rho}{}_{\tau}. \tag{2.20}$$

We can fix $\boldsymbol{M}^{\mu\nu}$ by comparing expression (2.20) in the limit of $\omega \to 0$ ($\omega_{\mu\nu} = -\omega_{\nu\mu}$) with the infinitesimal form:

$$\begin{aligned}\boldsymbol{L}^{\rho}{}_{\tau} &\underset{\omega_{\mu\nu}\to 0}{\approx} \delta^{\rho}{}_{\tau} - \frac{i}{4}(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau}\omega_{\mu\nu} = \delta^{\rho}{}_{\tau} + \omega^{\rho}{}_{\tau} \\ &= \delta^{\rho}{}_{\tau} + g^{\rho\mu}g_{\tau}{}^{\nu}\omega_{\mu\nu} = \delta^{\rho}{}_{\tau} + \frac{1}{2}g^{\rho\mu}\delta_{\tau}{}^{\nu}(\omega_{\mu\nu} - \omega_{\nu\mu}) \\ &= \delta^{\rho}{}_{\tau} + \frac{1}{2}(g^{\rho\mu}\delta^{\nu}_{\tau} - g^{\rho\nu}\delta^{\mu}_{\tau}).\end{aligned} \tag{2.21}$$

From this we have

$$(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau} = 2i\left(g^{\rho\mu}\delta^{\nu}_{\tau} - g^{\rho\nu}\delta^{\mu}_{\tau}\right). \tag{2.22}$$

2.3 Relativistic Wave Equations

A spinless relativistic particle can be described in terms of a scalar wave function $\phi(\boldsymbol{x}, t)$. This wave function can't posses any internal index, which can bear information about other degrees of freedom, such as spin. Relativistic particles satisfy the energy-momentum dispersion relation

$$E = \sqrt{m^2 + \boldsymbol{p}^2}. \tag{2.23}$$

In classical relativity we do not consider negative sign in the dispersion relation.

Recall that $p^\mu = (E, \boldsymbol{p})$ and that there exists a relativistic invariant given by

$$p^\mu p_\mu = p_0^2 - \boldsymbol{p}^2 = m^2. \tag{2.24}$$

In the formalism of first quantization, quantum mechanics is brought about by identifying operators with dynamical quantities

$$\boldsymbol{p} \rightarrow -i\boldsymbol{\nabla},\ E \rightarrow i\frac{\partial}{\partial t}. \tag{2.25}$$

Applying this process to the relativistic invariant in (2.24) we arrive to the following equation

$$\left(-\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2\right)\phi(x) = m^2\phi(x). \tag{2.26}$$

From the fact that $\partial_\mu = \left(\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right)$ we obtain another version of the same equation

$$\partial^\mu\partial_\mu\phi = \Box\phi = -m^2\phi. \tag{2.27}$$

Klein-Gordon equation is just a reflection of energy conservation (similarly as Schrödinger equation so, all relativistic wave functions satisfy this equation. It enforces the relativistic relationship between energy and momentum.

Finally, we arrive at the relativistic wave equation known as the Klein-Gordon equation, given by

$$(\Box + m^2)\phi(x) = 0. \tag{2.28}$$

If we accept this equation and seek solution of a definite energy and momentum, we get

$$\phi(x) \propto e^{-ipx} = e^{-iEt + i\boldsymbol{p}\cdot\boldsymbol{x}} = e^{-ip_0x_0 + i\boldsymbol{p}\cdot\boldsymbol{x}}. \tag{2.29}$$

Adopting $\partial_\mu\phi = -ip_\mu\phi$ we get that $\Box\phi = -p^2\phi$ and then

$$(-p^2 + m^2)\phi = 0. \tag{2.30}$$

Both positive and negative energy solutions are relevant in relativistic quantum theory!

So if $\phi \neq 0$ we have condition that $p^2 = m^2$ and hence

$$E = \pm\sqrt{\boldsymbol{p}^2 + m^2}. \tag{2.31}$$

Why can't we directly quantize relativistic energy relation?

A question may rise, why can't we directly quantize dispersion relation $\omega_{\boldsymbol{p}} = E_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2}$ using fact that $\boldsymbol{p} \rightarrow -i\boldsymbol{\nabla}$? To make sense to such a function of operator we have to interpret it in terms of the Taylor expansion:

$$\boldsymbol{H_p} = \sqrt{\boldsymbol{p}^2 + m^2} = m\left(1 + \frac{\boldsymbol{p}^2}{m^2}\right)^{\frac{1}{2}} = m + \frac{\boldsymbol{p}^2}{2m} - \frac{\boldsymbol{p}^4}{8m^3} + \cdots \tag{2.32}$$

Unfortunately we can not form covariant wave equation, i.e. if we formed a coordinate space (or momentum space) representation of a state vector $|\psi\rangle$, the resulting wave equation would have one time derivative and **infinite** series of increasing spatial derivatives. There is no way to put time and space on an "equal footing". Nonetheless, let us go ahead and try to build a wave equation

$$\begin{aligned} i\frac{\partial}{\partial t}\langle \boldsymbol{x}|\psi(t)\rangle &= \int \mathrm{d}^3\boldsymbol{p}\,\langle \boldsymbol{x}|\boldsymbol{p}\rangle\,\langle \boldsymbol{p}|\boldsymbol{H_p}|\psi(t)\rangle \\ &= \int \mathrm{d}^3\boldsymbol{x}'\int \mathrm{d}^3\boldsymbol{p}\,\langle \boldsymbol{x}|\boldsymbol{p}\rangle\,\langle \boldsymbol{p}|\boldsymbol{x}'\rangle\,\langle \boldsymbol{x}'|\boldsymbol{H_p}|\psi(t)\rangle \\ &= \int \mathrm{d}^3\boldsymbol{x}'\int \mathrm{d}^3\boldsymbol{p}\,\frac{e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')}}{(2\pi)^3}\,\langle \boldsymbol{x}'|\boldsymbol{H_p}|\psi(t)\rangle\,. \end{aligned} \tag{2.33}$$

The matrix element $\langle \boldsymbol{x}'|\boldsymbol{H_p}|\psi(t)\rangle$ is proportional to the infinite sum of $\langle \boldsymbol{x}'|\boldsymbol{p}^n|\psi(t)\rangle = (-i)^n \frac{\partial^n}{\partial x'^n}\langle \boldsymbol{x}'|\psi(t)\rangle$ elements. This in turn renders wave function to be **non-local**. Since it must reach further and further away from the region near $\boldsymbol{x}'$ in order to evaluate the time derivative. Eventually, the causality will be violated for any spatially localized function $\langle \boldsymbol{x}|\psi(t)\rangle$. Because of that we must abandon this approach and work with square of $\boldsymbol{H_p}$, i.e. ω_p^2 instead. This will remove the problem of the square root, but will introduce a different problem - negative energies. This will still prove to be more useful way to proceed.

Let us look at non-relativistic limit of Klein-Gordon equation. A mode with $E = m + \varepsilon$ would oscillate in time as $\phi \propto e^{-iEt}$. In the non-relativistic regime ε is much smaller than the rest mass m. We can factor-out the fast-oscillating part of the ϕ away and rewrite it

$$\phi(x) = \phi(\boldsymbol{x}, t) = e^{-imt}\varphi(\boldsymbol{x}, t). \tag{2.34}$$

Field φ is oscillating much more slowly that e^{-imt} in time. By inserting this into Klein-Gordon equation and using the fact that $\frac{\partial}{\partial t}e^{-imt}\varphi(\boldsymbol{x},t) = e^{-imt}\left(-im + \frac{\partial}{\partial t}\right)\varphi(\boldsymbol{x},t)$:

$$\begin{aligned} \frac{\partial}{\partial t}\left[e^{-imt}\left(-im+\frac{\partial}{\partial t}\right)\varphi\right] - e^{-imt}\boldsymbol{\nabla}^2\varphi(\boldsymbol{x},t) + m^2e^{-imt}\varphi = 0 \\ e^{-imt}\left(-im+\frac{\partial}{\partial t}\right)\left(-im+\frac{\partial}{\partial t}\right)\varphi - e^{-imt}\boldsymbol{\nabla}^2\varphi + m^2e^{-imt}\varphi = 0 \\ \left(-m^2 - 2im\frac{\partial}{\partial t} + \frac{\partial^2}{\partial t^2}\right)\varphi - \boldsymbol{\nabla}^2\varphi + m^2\varphi = 0. \end{aligned} \tag{2.35}$$

Dropping $\frac{\partial^2\phi}{\partial t^2}$ as small compared to $-imt\frac{\partial\phi}{\partial t}$ we find that

By the way, the Klein-Gordon equation was actually discovered before Schrödinger equation (by Erwin Schrödinger).

$$i\frac{\partial}{\partial t}\varphi = -\frac{\boldsymbol{\nabla}^2}{2m}\varphi, \tag{2.36}$$

and hence we recover the Schrödinger equation.

Let us focus on general solution to the Klein-Gordon equation, $\phi(x)$. Using Fourier decomposition it can be written as

$$\phi(x) = \int \frac{\mathrm{d}^4p}{(2\pi)^3}e^{-ipx}\widetilde{\phi}(p). \tag{2.37}$$

To find the solution we need to solve the Klein-Gordon equation in momentum space, i.e.

$$(p^2 - m^2)\widetilde{\phi}(p) = 0. \tag{2.38}$$

Equations of this form are solved by the Dirac δ-functions. Using this knowledge, we can write our solution in momentum space as

Here we use the well known property of Dirac δ-function, namely that $\delta(f(x)) = \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}$, where x_i are roots of f.

$$\begin{aligned}\widetilde{\phi}(p) &= f(p)\delta(p^2 - m^2) \\ &= \frac{f(p)\delta(p_0 + \sqrt{\boldsymbol{p}^2 + m^2})}{2\sqrt{\boldsymbol{p}^2 + m^2}} + \frac{f(p)\delta(p_0 - \sqrt{\boldsymbol{p}^2 + m^2})}{2\sqrt{\boldsymbol{p}^2 + m^2}}\end{aligned} \tag{2.39}$$

Using this knowledge and denoting $\omega_p = \sqrt{\boldsymbol{p}^2 + m^2}$, we can write the full solution as

$$\begin{aligned}\phi(x) &= \int \frac{\mathrm{d}^4 p}{(2\pi)^3} \frac{1}{2\omega_p} e^{-ipx}[f(p)\delta(p_0 + \omega_p) + f(p)\delta(p_0 - \omega_p)] \\ &= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2\omega_p}[e^{-i\omega_p t + i\boldsymbol{p}\cdot\boldsymbol{x}} f(\omega_p, \boldsymbol{p}) + e^{i\omega_p t + i\boldsymbol{p}\cdot\boldsymbol{x}} f(-\omega_p, \boldsymbol{p})] \\ &= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2\omega_p}[e^{-ipx} \underbrace{f(\omega_p, \boldsymbol{p})}_{f(\boldsymbol{p})} + e^{ipx} \underbrace{f(-\omega_p, -\boldsymbol{p})}_{g(\boldsymbol{p})}]\end{aligned} \tag{2.40}$$

Here $p^\mu = (\omega_p, \boldsymbol{p})$.

From that we can see, that general solution of the Klein-Gordon equation is a superposition of positive and negative energy eigenstate solutions. If we want to interpret $\phi(x)$ as a wave function, we have to find a non-negative norm, which will be Lorentz invariant and conserved by time evolution. Let's define norm of $\phi(x)$ to be

$$||\phi||^2 = (\phi|\phi) = i\int \mathrm{d}^3\boldsymbol{x} \left[\phi^* \frac{\partial\phi}{\partial x^0} - \left(\frac{\partial\phi}{\partial x^0}\right)^* \phi\right]. \tag{2.41}$$

The naturalness of this choice comes from the analogy with quantum mechanics - continuity equation, which defines the probability density. Consider the 4-current

$$J_\mu(x) = \frac{i}{2m}\left[\phi^*\partial_\mu\phi - (\partial_\mu\phi)^*\phi\right], \tag{2.42}$$

where the factor $1/2m$ is only a convention that ensures a correct non-relativistic limit [see Eq. (2.45)]. We know that each 4-current $J_\mu = (\rho, \boldsymbol{J})$ and hence it can be rewritten as

$$\begin{aligned}\boldsymbol{J}(x) &= \frac{i}{2m}[\phi^*\boldsymbol{\nabla}\phi - (\boldsymbol{\nabla}\phi)^*\phi], \\ \rho(x) &= \frac{i}{2m}[\phi^*\partial_0\phi - (\partial_0\phi)^*\phi].\end{aligned} \tag{2.43}$$

From Noether's theorem we know, that each conserved current has a zero 4-divergence. Computing it we obtain

$$\begin{aligned}
\partial^\mu J_\mu(x) &= i\left[\partial_\mu(\phi^*\partial^\mu\phi) - \partial_\mu(\phi\partial^\mu\phi^*)\right] \\
&= i[(\partial_\mu\phi^*)(\partial^\mu\phi) + \underbrace{\phi^*\partial^2\phi}_{-m^2\phi^*\phi} - (\partial_\mu\phi)(\partial^\mu\phi^*) - \underbrace{\phi\partial^2\phi^*}_{-m^2\phi\phi^*}] \\
&= 0. \qquad (2.44)
\end{aligned}$$

And hence from Noether's theorem $J_\mu(x)$ is a conserved current.

Current in non-relativistic limit

In non-relativistic limit, where we assume that $\phi(x) = e^{-imt}\varphi(\boldsymbol{x}, t)$ beforementioned form of current will reduce to the well know form for Schrödinger equation, namely:

$$\begin{aligned}
\boldsymbol{J}_{NR}(x) &= \frac{i}{2m}\left[\varphi^*\boldsymbol{\nabla}\varphi - (\boldsymbol{\nabla}\varphi)^*\varphi\right] \\
\rho_{NR}(x) &= \frac{i}{2m}\left[(-im)\varphi\varphi^* + \varphi^*\partial_0\varphi - (im)\varphi\varphi^*\right] \\
&= \frac{i}{2m}\left[(-i2m)\varphi\varphi^*\right] = \varphi\varphi* . \qquad (2.45)
\end{aligned}$$

This norm is also time independent, since

$$-\underbrace{\frac{\partial}{\partial t}\int_V \mathrm{d}^3x\rho}_{\substack{\text{Change in total}\\ \text{probability inside } V}} = \int_V \mathrm{d}^3x\boldsymbol{\nabla}\cdot\boldsymbol{J} = \underbrace{\int_{\partial V}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J}}_{\substack{\text{Flux of } \boldsymbol{J}\\ \text{through the } \partial V}} \to 0. \qquad (2.46)$$

Are our integrals convergent?

We want to show that $\int_V \mathrm{d}^3x\rho$ is finite. Since $\rho = \phi^*\phi$, we can rewrite our integral in spherical coordinates as $\int \mathrm{d}\omega\mathrm{d}r r^2|\phi|^2$. Since our fields behave as $\phi \sim \frac{1}{r^{3/2+\varepsilon}}$. Our current then has to behave like $\boldsymbol{J} \sim \phi\boldsymbol{\nabla}\phi \sim \frac{1}{r^{3/2+\varepsilon}}\frac{1}{r^{5/2+\varepsilon}} = \frac{1}{r^{4+\varepsilon}}$. Since our $\partial V \sim R^2$, our $\boldsymbol{J} \sim \frac{1}{R^{4+\varepsilon}}$ and total integral of $\lim_{R\to\infty}\int_{\partial V}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J} = 0$.

The fact that our norm is relativistically invariant can be shown as follows:

$$\begin{aligned}
||\phi||^2 = \int \mathrm{d}^3\boldsymbol{x}\rho(x) = \int \mathrm{d}^4x \underbrace{\delta(t_0)}_{\frac{\partial}{\partial x^0}\theta(x_0)} \rho(x) = \\
= \int \mathrm{d}^4x J^\alpha \frac{\partial}{\partial x^\alpha}\theta(n^\beta x_\beta).
\end{aligned} \qquad (2.47)$$

Here $n^\beta = (1\ 0\ 0\ 0)$. Define another wave function norm, $||\widetilde{\phi}||^2$ as

$$||\widetilde{\phi}||^2 = \int \mathrm{d}^4x J^\alpha \frac{\partial}{\partial x^\alpha}\theta(n'^\beta x_\beta). \qquad (2.48)$$

Here n' is a generic time-like vector. Taking difference between two norm we obtain

$$||\phi||^2 - ||\widetilde{\phi}||^2 = \int d^4x J^\alpha \frac{\partial}{\partial x^\alpha}\Big(\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\Big), \tag{2.49}$$

with $n'^\beta = \boldsymbol{L}^\beta{}_\gamma n^\gamma$. Because $\partial_\alpha J^\alpha = 0$ (as was shown before) we can rewrite expression under the integral sign as

$$||\phi||^2 - ||\widetilde{\phi}||^2 = \int d^4x \frac{\partial}{\partial x^\alpha}\left[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\right], \tag{2.50}$$

and using 4-dimensional version of Gaussian theorem we obtain that

$$||\phi||^2 - ||\widetilde{\phi}||^2 = \int dS_\alpha \left[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\right]. \tag{2.51}$$

To show that this is zero, consider two possibilities:

1. J^α is presumed to vanish if $|\boldsymbol{x}| \to \infty$ with fixed t.
2. $\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)$ vanishes for $|t| \to \infty$ with $\boldsymbol{x}$ fixed.

Hence, the difference is zero and norm is relativistically invariant.

Let us return to the general solution

$$\phi(x) = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_p}[f(\boldsymbol{p})e^{-ipx} + g(\boldsymbol{p})e^{ipx}], \tag{2.52}$$

and explore, what its norm would look like:

$$\begin{aligned}||\phi||^2 = i\int d^3x \Bigg\{\Bigg[\int &\frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_p}\left(f^*(\boldsymbol{p})e^{ipx} + g^*(\boldsymbol{p})e^{-ipx}\right)\\ \times \int &\frac{d^3\boldsymbol{q}}{(2\pi)^3 2\omega_q}\left(f(\boldsymbol{q})(-i\omega_q)e^{-iqx} + g(\boldsymbol{q})(i\omega_q)e^{iqx}\right)\Bigg]\\ - \Bigg[\int &\frac{d^3\boldsymbol{q}}{(2\pi)^3 2\omega_q}\left(f(\boldsymbol{q})e^{-iqx} + g(\boldsymbol{q})e^{iqx}\right)\\ \times \int &\frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_q}\left(f^*(\boldsymbol{p})(i\omega_p)e^{ipx} + g^*(\boldsymbol{p})(-i\omega_p)e^{-ipx}\right)\Bigg]\Bigg\}.\end{aligned} \tag{2.53}$$

Since our norm is time independent, elements of type $e^{\pm i(\omega_p+\omega_q)t}$ must cancel, and only terms of the type $e^{\pm i(\omega_p-\omega_q)t}$ should be considered. Continuing

$$\begin{aligned}||\phi||^2 = i\int d^3\boldsymbol{x}\int \frac{d^3\boldsymbol{p}d^3\boldsymbol{q}}{(2\pi)^6 4\omega_p\omega_q}\Big[&\Big(f^*(\boldsymbol{p})f(\boldsymbol{q})(-i\omega_q)e^{ix(p-q)}\\ &+ g^*(\boldsymbol{p})g(\boldsymbol{q})(i\omega_q)e^{-ix(p-q)}\Big)\\ &- \Big(f(\boldsymbol{q})f^*(\boldsymbol{p})(i\omega_p)e^{ix(p-q)}\\ &+ g(\boldsymbol{q})g^*(\boldsymbol{p})(-i\omega_p)e^{-ix(p-q)}\Big)\Big]\end{aligned} \tag{2.54}$$

And hence our total norm is

$$||\phi||^2 = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\left[|f(\boldsymbol{p})|^2 - |g(\boldsymbol{p})|^2\right]. \tag{2.55}$$

Negative energy states are culprits of problems in relativistic quantum mechanics.

But this norm is not generally positive definite! However, if we restrict our attention to positive energy only, $g(\boldsymbol{p}) = 0$, then $||\phi||^2$ is positive definite.

In this spirit, the general scalar product between two states is

$$(\psi, \phi) = i \int d^3x \, [\psi^* \partial_0 \phi - \phi(\partial_0 \psi)^*] \; . \tag{2.56}$$

Generally, ρ cannot represent probability density, it may well be considered (while satisfying continuity equation) as the density of charge (or any other conserved quantity).

There is a second problem with Klein-Gordon equation. Any plane wave function, i.e.

$$\psi(x) = N e^{\pm i(\omega_p t - \boldsymbol{p}\cdot\boldsymbol{x})} \, , \tag{2.57}$$

satisfies Klein-Gordon equation, provided that $E^2 = \omega_p^2 = \boldsymbol{p}^2 + m^2$. Thus negative energies $E = -\sqrt{\boldsymbol{p}^2 + m^2}$ are on the same footing as the physical ones $E = \sqrt{\boldsymbol{p}^2 + m^2}$. This gives a problem - energy spectrum is unbounded from bellow. Of course even in classical physics, the relativistic relation $E^2 = \boldsymbol{p}^2 c^2 + m^2 c^4$ has two solutions $E = \pm\sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4}$. However, in classical physics we can simply assume that the only physical particles are those with $E \geq 0$. This is because the positive-energy solutions have $E > mc^2$, while the negative ones have $E \leq mc^2$. Hence there is a finite gap between them and in classical (non-quantum) physics there does not exist any continuous process that can take a particle from positive to negative energy.

In relativistic quantum mechanics the problem is more pressing. As Dirac pointed out in 1928 paper [P. A. M. Dirac, Proc. Roy. Soc. A117, 610 (1928)] the interaction of electrons with radiation can produce transition, in which a positive energy electrons falls into a negative energy state, with the energy carried off by two or more photons. This brings about a problem. If we have a quantum particle whose state satisfies the Klein–Gordon equation it is possible to extract an arbitrary amount of energy from it (in the form of photons). This, in turn, will lead to the perpetum mobile of first kind. In addition, when particle reaches the negative energy states there is nothing that would prevent it to decay to even lower energy state. Consequently, the matter (together with us) would be unstable!

2.4 Dirac Equation

Klein-Gordon equation is a second order equation in time derivative, which leads to the norm not being positive definite. Dirac sought an equation, that would remedy these "difficulties". It turned out, that by "linearizing" relativistic wave equation, Dirac arrived (by coincidence) on the wave equation for electron. Since the spins is involved, the wave function is not anymore scalar (recall Pauli equation, where solutions are two-component spinor wave function that is not scalar wrt. Galileo group).

Dirac had two goals:

1. Equation for wave function that is linear and first order in time derivative. Relativistic invariance then suggests that the equation will also be of first order in spatial derivatives.
2. Positive definite norm of a solution.

Assume that this equation has the form

$$\left(i\gamma^0 \frac{\partial}{\partial x^0} + i\boldsymbol{\gamma}\cdot\boldsymbol{\nabla}\right)\psi(x) = m\psi(x), \tag{2.58}$$

simplifying this leads to the

$$\left(i\gamma_\mu \partial^\mu - m\right)\psi(x) = 0, \tag{2.59}$$

and defining Feynman's slash notation $\not{\partial} = \gamma_\mu \partial^\mu$ this equation reduces to the

$$\left(i\not{\partial} - m\right)\psi(x) = 0. \tag{2.60}$$

Here $\{\gamma^\mu\} = \{\gamma^0, \boldsymbol{\gamma}\}$ are some unspecified numbers or matrices. We require that $\psi(x)$ should also satisfy Klein–Gordon equation, since Klein-Gordon equation is just statement of $p_\mu p^\mu = m^2$. Multiplying (2.59) by $(i\gamma_\nu \partial^\nu + m)$ we get

$$\begin{aligned} (i\gamma_\nu \partial^\nu + m)\left(i\gamma_\mu \partial^\mu - m\right)\psi(x) = 0 \\ \left(-\gamma_\mu \gamma_\nu \partial^\mu \partial^\nu - m^2\right)\psi(x) = 0. \end{aligned} \tag{2.61}$$

We can rewrite $\gamma_\mu \gamma_\nu \partial^\mu \partial^\nu$ as

$$\gamma_\mu \gamma_\nu \partial^\mu \partial^\nu = \frac{1}{2}\gamma_\mu \gamma_\nu \partial^\mu \partial^\nu + \frac{1}{2}\gamma_\nu \gamma_\mu \partial^\nu \partial^\mu = \frac{1}{2}\{\gamma_\mu, \gamma_\nu\}\partial^\mu \partial^\nu, \tag{2.62}$$

where $\{\gamma_\mu, \gamma_\nu\} = \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu$. To obtain Klein–Gordon equation we must impose condition

$$\{\gamma_\mu, \gamma_\nu\} = 2\boldsymbol{g}_{\mu\nu}, \ \{\gamma^\mu, \gamma^\nu\} = 2\boldsymbol{g}^{\mu\nu}. \tag{2.63}$$

And because $\gamma^\mu \gamma^\nu \partial^\mu \partial^\nu = \partial_\nu \partial^\nu$ we get $(\Box + m^2)\psi(x) = 0$.

Dirac's derivation

Dirac started with the following ansatz:

$$i\frac{\partial\psi}{\partial t} = \left(\frac{1}{i}\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m\right)\psi = \boldsymbol{H}_D\psi, \tag{2.64}$$

where $\boldsymbol{H}_D$ is Dirac's Hamiltonian, which should be Hermitan (and hence $\boldsymbol{\alpha}$ and β are Hermitan). Klein-Gordon equation implies that $\{\alpha_i, \alpha_k\} = 0$, $\{\alpha_i, \beta\} = 0$ and $\alpha_i^2 = \beta^2 = 1$ for $i \neq k$. Here $\{A, B\}$ is a symmetric combination of A and B. This operation is called anti-commutator. By rewriting Dirac equation explicitly

$$\left(i\gamma^0\partial_0 + \gamma^i\partial_i - m\right)\psi = 0, \tag{2.65}$$

multiplying by the inverse of γ^0 we get

$$\left(i\gamma^{0,-1}\gamma_0\partial_0 + i\gamma^{0,-1}\gamma^i\partial_i - \gamma^{0,-1}m\right)\psi = 0. \tag{2.66}$$

And finally this is equivalent to the

$$i\partial_0\psi = \left(\frac{1}{i}\gamma^{0,-1}\gamma^i\partial_i + \gamma^{0,-1}m\right)\psi. \tag{2.67}$$

And hence we can see that $\boldsymbol{\alpha} = \gamma_0^{-1}\boldsymbol{\gamma}$ and $\beta = \gamma_0^{-1}$. Because $\{\gamma^0, \gamma^0\} = 2$, we see that $\gamma^0 = \gamma_0^{-1}$. From anti-commutation relation for γ^0 and γ^i we have that $\gamma^0\gamma^i = -\gamma^i\gamma^0$ and from Hermiticity we also have that $\gamma^0\gamma^i = (\gamma^0\gamma^i)^\dagger = \gamma^{i\dagger}\gamma^{0\dagger}$. And from that we see that $\gamma^0\gamma^i\gamma^0 = \gamma^{i\dagger} = -\gamma^i(\gamma^0)^2$ and hence we get another important condition, that $(\gamma^0)^2 = \mathbf{1}$.

Relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\boldsymbol{g}^{\mu\nu} \tag{2.68}$$

is known as Clifford algebra $\mathrm{CL}_{1,3}(\mathbb{R})$. This particular algebra is also known as the Dirac algebra. We can ask ourselves, what is the smallest dimension of γ^μ in 4-dimensional space. In fact, matrices α^i and β have eigenvalues equal to ± 1 (i.e. they are Hermitan). For $i \neq j$ we have

$$\det\left(\alpha^i\alpha^j\right) = \det\left(-\alpha^j\alpha^i\right) = (-1)^\alpha \det\left(\alpha^j\alpha^i\right) \tag{2.69}$$

$$\det\left(\alpha^i\beta\right) = (-1)^\alpha \det\left(\beta\alpha^i\right). \tag{2.70}$$

The dimension of α^i, $i = 1, 2, 3$ and β must be even. Since for $d = 2$ there exists only 3 anti-commuting Hermitian matrices - Pauli matrices, we have $d \geq 4$. There are many representations of $\mathrm{CL}_{1,3}(\boldsymbol{R})$ with $d > 4$ (altought they are rarely used in practice). An explicit representation with $d = 4$ is provided by matrices

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \ \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \tag{2.71}$$

Pauli matrices by themselves generate Clifford algebra $\mathrm{CL}_{0,3}(\mathbb{R})$ via relation $\{\sigma_a, \sigma_b\} = 2\delta_{ab}\mathbf{1}$. This algebra is known as Pauli algebra.

Here $\boldsymbol{\sigma}$ are Pauli matrices. This representation is known as Dirac representation and is useful when discussing non-relativistic limits of the theory. The useful technical trick, for calculating gamma matrices in Dirac's representation is based on properties of tensor product ⊗ on matrices. Because

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \cdot (\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{A} \cdot \boldsymbol{C}) \otimes (\boldsymbol{B} \cdot \boldsymbol{D}) \tag{2.72}$$

and from Pauli matrix identity

$$\sigma^i \sigma^j = \delta^{ij} + i\varepsilon^{ijk}\sigma^k \tag{2.73}$$

we can rewrite our gamma matrices as

$$\gamma^0 = \sigma^3 \otimes \mathbf{1}, \ \boldsymbol{\gamma} = i\sigma^2 \otimes \boldsymbol{\sigma}. \tag{2.74}$$

2.5 Lorentz Invariance of Dirac Equation

Recall that a non-relativistic particle with spin has a wave function $\psi_\alpha(x)$ (Weyl spinor) which transforms under rotation $\boldsymbol{R}$ as

$$\psi_\alpha(\boldsymbol{x}) \xrightarrow{\boldsymbol{R}} D_{\alpha\beta}(\boldsymbol{R})\psi_\beta(\boldsymbol{R}^{-1}\boldsymbol{x}). \tag{2.75}$$

In a similar way, the Dirac wave function under a Lorentz transformation $\boldsymbol{L}$ transforms as

$$\psi(x) \xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) = S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x), \tag{2.76}$$

where $S(\boldsymbol{L})$ is an appropriate representation of the Lorentz group, that acts in the vector space, in which the Dirac wave function takes its values. From this we have that $S(\boldsymbol{L})$ should be a 4×4 matrix. We want to show that if we apply Lorentz transformation on Dirac equation, that we will get

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi_{\boldsymbol{L}}(x) = 0. \tag{2.77}$$

Let us rewrite this expression explicitly:

$$\begin{aligned} \text{L.H.S.} &= \left(i\gamma^\mu \partial_\mu - m\right) S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x) \\ &= S(\boldsymbol{L}) \left[iS^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) - m\right] \psi(\boldsymbol{L}^{-1}x) \\ &= S(\boldsymbol{L}) \left[iS^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) - m\right] \psi(x'), \end{aligned} \tag{2.78}$$

where $x' = \boldsymbol{L}^{-1}x$. If we can find an appropriate matrix $S(\boldsymbol{L})$ such that

$$S^{-1}(\boldsymbol{L})\gamma^\mu \partial_\mu S(\boldsymbol{L}) = \gamma^\mu \partial'_\mu, \tag{2.79}$$

then

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi_{\boldsymbol{L}}(x) = S(\boldsymbol{L})(i\gamma^\mu \partial'_\mu - m)\psi(x'). \tag{2.80}$$

Recall that $\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = \boldsymbol{L}_\mu{}^\nu \frac{\partial}{\partial x^\nu}$.

Since ψ is a Dirac wave function, then

$$\left(i\gamma^\mu \partial'_\mu - m\right) \psi(x') = 0 \implies \left(i\gamma^\mu \partial_\mu - m\right) \psi_{\boldsymbol{L}}(x) = 0. \tag{2.81}$$

So also $\psi_{\boldsymbol{L}}$ is Dirac's wave function and thus Dirac's equation is relativistically invariant. To find such $S(\boldsymbol{L})$ we have to explore properties of Lorentz group in greater details.

Lorentz group

From the condition (2.5) we know that there exist 2 kinds of Lorentz transformations - proper (those with $\det L = 1$) and improper (those with $\det L = -1$). We can expand this classification even further by following approach:

$$1 = \boldsymbol{L}^0{}_\nu \boldsymbol{L}^0{}_{\nu'} \boldsymbol{g}^{\nu\nu'} = (\boldsymbol{L}^0{}_0)^2 - \sum_{i=1}^{3} (\boldsymbol{L}^0{}_i)^2. \tag{2.82}$$

Rewriting we arrive to the condition

$$(\boldsymbol{L}^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\boldsymbol{L}^0{}_i)^2. \tag{2.83}$$

Lorentz transformations for which $\boldsymbol{L}^0{}_0 \geq 1$ are called orthochronous transformations, those with $\boldsymbol{L}^0{}_0 < -1$ are called non-orthochronous. We can't switch between those transformations using continuous process, only via discrete transformations such as **parity** or **time** reversal. Recall from section 2.2. that Lorentz group has 6 independent parameters in infinitesimal transformation form. Using those we can construct finite Lorentz transformation as

$$\boldsymbol{L}^\rho{}_\tau = \left(e^{-\frac{i}{4} \boldsymbol{M}^{\mu\nu} \omega_{\mu\nu}} \right)^\rho{}_\tau. \tag{2.84}$$

Here $M^{\mu\nu}$ are so-called generators of Lorentz transformation fixed by comparison with the infinitesimal transformation when $\omega_{\mu\nu} \to 0$:

$$(\boldsymbol{M}^{\mu\nu})^\rho{}_\tau = 2i \left(g^{\rho\mu} \delta^\nu_\tau - g^{\rho\nu} \delta^\mu_\tau \right). \tag{2.85}$$

Let us review rotation group first. Element of a rotational group are defined by

$$\boldsymbol{R}_{ij} = \left(e^{-i\theta \boldsymbol{n}_k \boldsymbol{J}_k} \right)_{ij} = \left(e^{-i\omega_k \boldsymbol{J}_k} \right)_{ij}. \tag{2.86}$$

Here $\boldsymbol{n}$ is a unit vector defining the axis of rotation. J_i here satisfy angular momentum commutation relations

$$\left[\boldsymbol{J}_i, \boldsymbol{J}_j \right] = i\varepsilon_{ijk} \boldsymbol{J}_k. \tag{2.87}$$

Recall from quantum mechanics that vector operators are sets of 3 operators, rotated according to

$$U(\boldsymbol{R})^\dagger V_i(\boldsymbol{x}) U(\boldsymbol{R}) = \sum_j \boldsymbol{R}_{ij} V_j(\boldsymbol{x}). \tag{2.88}$$

Here $U(\boldsymbol{R})$ is a representation of rotation group which acts state space (e.g. $L^2(\mathbb{R})$) and $\boldsymbol{R}_{ij}$ is a representation of rotation group that acts on the operator indices.

For infinitesimal rotation one obtains

$$(1 + i\omega_k \boldsymbol{J}_k) \boldsymbol{V}_i (1 - i\omega_k \boldsymbol{J}_k) = (1 - i\omega_k (\boldsymbol{J}_k)_{ij}) \boldsymbol{V}_j. \tag{2.89}$$

Here J_k are operators of angular momentum, acting on state space (e.g.

$L^2(\mathbb{R})$) and $(\boldsymbol{J}_k)_{ij}$ is a vector representation of angular momentum in 3D, which is defined by $(\boldsymbol{J}_j)_{ik} = i\varepsilon_{ijk}$, i.e.:

$$\boldsymbol{J}_1 = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \ \boldsymbol{J}_2 = i\begin{pmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \ \boldsymbol{J}_3 = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.90}$$

After algebraic manipulations in (2.89) we obtain

$$i\omega_k \left[\boldsymbol{J}_k, \boldsymbol{V}_i\right] = -i\omega_k (\boldsymbol{J}_k)_{ij} \boldsymbol{V}_j, \tag{2.91}$$

which then simplifies to

$$[\boldsymbol{J}_k, \boldsymbol{V}_i] = -(\boldsymbol{J}_k)_{ij}\boldsymbol{V}_j = -i\varepsilon_{ikj}\boldsymbol{V}_j = i\varepsilon_{kij}\boldsymbol{V}_j. \tag{2.92}$$

From this follows the fact, that Lie generators themselves are vector operators.

Adjoint representation

Representation where elements of algebra (generators) are defined via structure constants as $(a_i)_{jk} = c_{ikj}$ are called **adjoint** or **regular** representation.

From the above representation of the generators we can deduce the commutation relations determining the Lie algebra. We find

$$\left[\boldsymbol{M}^{\mu\nu}, \boldsymbol{M}^{\alpha\beta}\right] = 2i\left\{\boldsymbol{g}^{\mu\beta}\boldsymbol{M}^{\nu\alpha} + \boldsymbol{g}^{\nu\alpha}\boldsymbol{M}^{\mu\beta} - \boldsymbol{g}^{\mu\alpha}\boldsymbol{M}^{\nu\beta} - \boldsymbol{g}^{\nu\beta}\boldsymbol{M}^{\mu\alpha}\right\}. \tag{2.93}$$

For $\boldsymbol{M}^{ij}$, $i, j = 1, 2, 3$ we have

$$\begin{aligned} \left[\boldsymbol{M}^{12}, \boldsymbol{M}^{13}\right] &= 2i\boldsymbol{M}^{23} \\ \left[\boldsymbol{M}^{23}, \boldsymbol{M}^{12}\right] &= -2i\boldsymbol{M}^{31} \end{aligned} \tag{2.94}$$

etc. Defining

$$\boldsymbol{J}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{M}^{jk} \iff \boldsymbol{M}^{jk} = 2\varepsilon^{jki}\boldsymbol{J}_i. \tag{2.95}$$

Then we have

$$\begin{aligned} [\boldsymbol{J}_3, (-\boldsymbol{J}_2)] &= i\boldsymbol{J}_1, \\ [\boldsymbol{J}_2, \boldsymbol{J}_3] &= i\boldsymbol{J}_1, \\ [\boldsymbol{J}_1, \boldsymbol{J}_3] &= -i\boldsymbol{J}_2. \end{aligned} \tag{2.96}$$

Generally we have

$$[\boldsymbol{J}_i, \boldsymbol{J}_j] = i\varepsilon_{ijk}\boldsymbol{J}_k. \tag{2.97}$$

From this we can see, that $\boldsymbol{J}_i$ are generators of rotations in 3D, since they close the familiar algebra of rotations, i.e., SO(3) ~ SU(2) algebra. Similarly we can define

$$\boldsymbol{M}^{0i} = 2\boldsymbol{K}^i. \tag{2.98}$$

From this we can see that

$$[\boldsymbol{M}^{01}, \boldsymbol{M}^{02}] = -2i\boldsymbol{M}^{12} \implies [\boldsymbol{K}^1, \boldsymbol{K}^2] = -i\boldsymbol{J}^3, \tag{2.99}$$

and hence generally

$$[\boldsymbol{K}^i, \boldsymbol{K}^j] = -i\varepsilon^{ijk}\boldsymbol{J}_k \iff [\boldsymbol{K}_i, \boldsymbol{K}_j] = -i\varepsilon_{ijk}\boldsymbol{J}_k. \tag{2.100}$$

Here $\boldsymbol{K}_i$ are generators of boosts in i direction. To close the algebra we also need commutators of the type $[\boldsymbol{K}, \boldsymbol{J}]$. It can be shown that

$$[\boldsymbol{J}_i, \boldsymbol{K}_j] = i\varepsilon_{ijk}\boldsymbol{K}_k. \tag{2.101}$$

So we can equivalently rewrite the algebra (2.93) as

$$\begin{aligned} [\boldsymbol{J}_i, \boldsymbol{J}_j] &= i\varepsilon_{ijk}\boldsymbol{J}_k, \\ [\boldsymbol{J}_i, \boldsymbol{K}_j] &= i\varepsilon_{ijk}\boldsymbol{K}_k, \\ [\boldsymbol{K}_i, \boldsymbol{K}_j] &= -i\varepsilon_{ijk}\boldsymbol{J}_k. \end{aligned} \tag{2.102}$$

Those commutation relations define SO(3, 1) algebra. First commutation relation forms a subalgebra of SO(3, 1) algebra, namely algebra SO(3). But boosts do not form a subalgebra of SO(3, 1), so we need both boosts and rotations to form closed algebra.

Group of Lorentz transformations

Defining or **fundamental** representation of SO(3, 1) group is given by

$$x^\intercal x = (\boldsymbol{L}x)^\intercal(\boldsymbol{L}x) = x^\intercal \boldsymbol{L}^\intercal \boldsymbol{L} x = \text{invariant}. \tag{2.103}$$

From this we can explicitly write that $\boldsymbol{L}^\intercal\boldsymbol{L} = 1$ ($\boldsymbol{L}_\mu{}^\nu \boldsymbol{L}^\mu{}_\alpha = \delta^\nu{}_\alpha$) and hence $\boldsymbol{L}^\intercal = \boldsymbol{L}^{-1}$. From this we can see that $\boldsymbol{L}$ are orthogonal matrices preserving structure $x_0^2 - \sum_{i=1}^3 x_i^2$. Those matrices are of O(3, 1) type and since $\det \boldsymbol{L} = 1$ for proper transformations, we stress this extra fact in "S" in SO(3, 1) group.

Commutation relations can be diagonalized via transformation

This diagonalization doesnot provide Hermitian generators!

$$\boldsymbol{N}_i = \frac{1}{2}(\boldsymbol{J}_i + i\boldsymbol{K}_i), \quad \boldsymbol{N}_i^\dagger = \frac{1}{2}(\boldsymbol{J}_i - i\boldsymbol{K}_i).$$

From this follows that

$$[\boldsymbol{N}_i, \boldsymbol{N}_j^\dagger] = 0, \quad [\boldsymbol{N}_i, \boldsymbol{N}_j] = i\varepsilon_{ijk}\boldsymbol{N}_k, \quad [\boldsymbol{N}_i^\dagger, \boldsymbol{N}_j^\dagger] = i\varepsilon_{ijk}\boldsymbol{N}_k^\dagger. \tag{2.104}$$

The relation $[\boldsymbol{N}_i, \boldsymbol{N}_j] = i\varepsilon_{ijk}\boldsymbol{N}_k$ (and the same for $\boldsymbol{N}^\dagger$) belongs to the SU(2) algebra. Hence we can see that SO(3, 1) $\sim$ SU(2) $\oplus$ SU(2) $\sim$ SL(2, $\mathbb{C}$). For SU(2) algebra we can define Casimir operator. For $\boldsymbol{N}_i$ ($\boldsymbol{N}_i^\dagger$) it can be defined as $\sum_{i=1}^3 \boldsymbol{N}_i^2 = n(n+1)$ (or $\sum_{i=1}^3 \boldsymbol{N}_i^{\dagger 2} = m(m+1)$). Constants n (m) describe nothing but the size of the angular momenta (or spin). The representation of SO(3, 1) can be then denoted with the pair (n, m). Note in particular that the transformation with respect to spatial parity is given by

$$\underbrace{\boldsymbol{J}_i \to \boldsymbol{J}_i}_{\text{Pseudovector}} \quad \text{and} \quad \underbrace{\boldsymbol{K}_i \to -\boldsymbol{K}_i}_{\text{Vector}}, \tag{2.105}$$

and hence

$$\boldsymbol{N}_i \overset{P}{\leftrightarrow} \boldsymbol{N}_i^\dagger \implies (n,m) \overset{P}{\leftrightarrow} (m,n). \tag{2.106}$$

Generally representations of Lorentz group are not parity invariant (for example parity is violated in weak interactions). In addition, since $\boldsymbol{J}_i = \boldsymbol{N}_i + \boldsymbol{N}_i^\dagger$ we can identify the spin of representation $(n+m)$, e.g. spin 1 particle can have representations $\underbrace{(1/2,1/2)}_{\text{Parity invariant}}$ or $\underbrace{(1,0) \text{ or } (0,1)}_{\text{Parity non-invariant}}$.

Lorentz Invariance of Dirac Equation Continued

$S(\boldsymbol{L})S(\boldsymbol{L}^{-1}) = S(\boldsymbol{1}) = \boldsymbol{1}$. Hence $S(\boldsymbol{L}^{-1}) = S^{-1}(\boldsymbol{L})$.

Our goal is to show that the Dirac equation is invariant under

$$\psi(x) \xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) = S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x), \tag{2.107}$$

where $S(\boldsymbol{L})$ is a representation of the Lorentz transformation on the space of ψ, i.e. $S(\boldsymbol{L}_1)S(\boldsymbol{L}_2) = S(\boldsymbol{L}_1\boldsymbol{L}_2)$. The invariance of the Dirac equation depends on showing that

$$S^{-1}(\boldsymbol{L})\gamma^\mu S(\boldsymbol{L}) = \boldsymbol{L}^\mu{}_\nu\gamma^\nu. \tag{2.108}$$

Again we are dealing with connected part of the Lorentz group.

If we consider an infinitisemal Lorentz transformation

$$\boldsymbol{L} = \boldsymbol{1} - \frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}, \tag{2.109}$$

then correspondingly

$$S(\boldsymbol{L}) = \boldsymbol{1} - \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}, \tag{2.110}$$

where $\boldsymbol{\sigma}^{\mu\nu}$ are the generators of the Lorentz group in the representation that is appropriate to Dirac space (space of Dirac's wave functions — bispinors) and of course

$$S^{-1}(\boldsymbol{L}) = S(\boldsymbol{L}^{-1}) = \boldsymbol{1} + \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}. \tag{2.111}$$

Putting this into (2.108) we obtain that

$$\left(\boldsymbol{1} + \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}\right)\gamma^\rho\left(\boldsymbol{1} - \frac{i}{4}\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}\right) = \left(\delta^\rho_\tau + \omega^\rho_\tau\right)\gamma^\tau. \tag{2.112}$$

This should determine the Dirac representation generators $\boldsymbol{\sigma}^{\mu\nu}$.

Hence

$$\frac{i}{4}[\boldsymbol{\sigma}^{\mu\nu}\omega_{\mu\nu}, \gamma^\rho] = \omega^\rho_\tau\gamma^\tau, \tag{2.113}$$

or by writing

$$\omega^\rho{}_\tau\gamma^\tau = \omega_{\mu\nu}\boldsymbol{g}^{\mu\rho}\gamma^\nu = \frac{1}{2}\omega_{\mu\nu}\left(\boldsymbol{g}^{\mu\rho}\gamma^\nu - \boldsymbol{g}^{\nu\rho}\gamma^\mu\right). \tag{2.114}$$

This reduces to the following relation

$$\frac{i}{4}[\sigma^{\mu\nu},\gamma^{\rho}] = \frac{1}{2}\left(g^{\mu\rho}\gamma^{\nu} - g^{\nu\rho}\gamma^{\mu}\right). \tag{2.115}$$

This condition is satisfied if $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^{\mu},\gamma^{\nu}]$. This can be shown using well known commutation identity $[AB,C] = A\{B,C\} - \{A,C\}B$:

$$\begin{aligned}
&\frac{i}{4}\left(\frac{i}{2}[\gamma^{\mu}\gamma^{\nu},\gamma^{\rho}] - \frac{i}{2}[\gamma^{\nu}\gamma^{\mu},\gamma^{\rho}]\right) = \\
&= -\frac{1}{8}\left(\gamma^{\mu}\{\gamma^{\nu},\gamma^{\rho}\} - \{\gamma^{\mu},\gamma^{\rho}\}\gamma^{\nu} - \gamma^{\nu}\{\gamma^{\mu},\gamma^{\rho}\} + \{\gamma^{\nu},\gamma^{\rho}\}\gamma^{\mu}\right) \\
&= -\frac{1}{8}\left(2\gamma^{\mu}g^{\nu\rho} - 2g^{\mu\rho}\gamma^{\nu} - 2\gamma^{\nu}g^{\mu\rho} + 2g^{\nu\rho}\gamma^{\mu}\right) \\
&= \frac{1}{2}\left(\gamma^{\nu}g^{\mu\rho} - \gamma^{\mu}g^{\nu\rho}\right).
\end{aligned} \tag{2.116}$$

We also can show, that condition $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^{\mu},\gamma^{\nu}]$ is appropriate for Lorentz generators, i.e. we can show that it satisfies the correct algebra

$$[\sigma^{\mu\nu},\sigma^{\alpha\beta}] = 2i\left\{g^{\mu\beta}\sigma^{\nu\alpha} + g^{\nu\alpha}\sigma^{\mu\beta} - g^{\mu\alpha}\sigma^{\nu\beta} - g^{\nu\beta}\sigma^{\mu\alpha}\right\}. \tag{2.117}$$

From all of that, for finite Lorentz transformation in Dirac space we have

$$S(L) = e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}},\; L = e^{-\frac{i}{4}M^{\mu\nu}\omega_{\mu\nu}}. \tag{2.118}$$

And hence (2.108) is Lorentz invariant.

2.6 Dirac Bilinears

Dirac bilinears are relevant for construction of observables in quantum field theory. First, general Dirac wave function has form

$$\psi(x) = \begin{pmatrix}\psi_1(x)\\ \psi_2(x)\\ \psi_3(x)\\ \psi_4(x)\end{pmatrix},\; \psi^{\dagger}(x) = (\psi_1^{\dagger}(x),\psi_2^{\dagger}(x),\psi_3^{\dagger}(x),\psi_4^{\dagger}(x)). \tag{2.119}$$

We define spinor adjoint to ψ as

$$\overline{\psi}(x) = \psi^{\dagger}(x)\gamma^0. \tag{2.120}$$

How $\overline{\psi}(x)$ transforms under the Lorentz transformation? To understand this we use two simple facts, namely

$$\begin{aligned}
\psi(x) &\xrightarrow{L} \psi_{L}(x) = S(L)\psi(L^{-1}x) \\
\psi^{\dagger}(x) &\xrightarrow{L} \psi_{L}^{\dagger}(x) = \psi^{\dagger}(L^{-1}x)S^{\dagger}(L).
\end{aligned} \tag{2.121}$$

Now, we can multiply the second equation by γ^0 on the right, i.e.

$$\overline{\psi}(x) = \psi^{\dagger}(x)\gamma^0 \xrightarrow{L} \psi_{L}^{\dagger}(x)\gamma^0 = \psi^{\dagger}(L^{-1}x)S^{\dagger}(L)\gamma^0. \tag{2.122}$$

We might now use the fact that since $S(\boldsymbol{L}) = e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}}$, then $S^\dagger(\boldsymbol{L}) = e^{\frac{i}{4}(\sigma^{\mu\nu})^\dagger\omega_{\mu\nu}}$, where $(\sigma^{\mu\nu})^\dagger$ is

$$(\sigma^{\mu\nu})^\dagger = \left(\frac{i}{2}[\gamma^\mu, \gamma^\nu]\right)^\dagger = -\frac{i}{2}[\gamma^{\nu\dagger}, \gamma^{\mu\dagger}] = \frac{i}{2}[\gamma^{\mu\dagger}, \gamma^{\nu\dagger}]. \tag{2.123}$$

Problem, however is, how do we "rotate" $\gamma^{\nu\dagger}$ to γ^ν?

By writing Dirac equation in Schrödinger like form, i.e.

$$i\partial_t\psi = -i\gamma^0\gamma^i\partial_i\psi + \gamma^0 m\psi \equiv \boldsymbol{H}_D\psi, \tag{2.124}$$

($\boldsymbol{H}_D$ is a Dirac Hamiltonian) we get from the presumed hermiticity of $\boldsymbol{H}_D$ (as already seen in "Dirac's derivation" note) that

$$\gamma^0 = \gamma^{0\dagger} \tag{2.125}$$

$$\gamma^0\gamma^i = \gamma^{i\dagger}\gamma^{0\dagger} = \gamma^{i\dagger}\gamma^0 \tag{2.126}$$

$$\gamma^{i\dagger} = \gamma^0\gamma^i\gamma^0 = -\gamma^i. \tag{2.127}$$

Those identities are valid irrespective of chosen representation and a particular example of γ-matrices that satisfy these conditions is provided by the oldest representation of Dirac matrices that is due to Dirac himself, i.e.

$$\gamma^0 = \sigma^3 \otimes \mathbf{1} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \; \gamma^i = i\sigma^2 \otimes \sigma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{2.128}$$

Properties of γ-matrices

Here we summaries some important properties of γ-matrices.

- $\gamma^0 = \gamma^{0,-1}$, $\gamma^{0,2} = \mathbf{1}$ and $(\gamma^0)^\dagger = \gamma^0$
- $(\gamma^i)^\dagger = -\gamma^i$
- $\gamma^0(\gamma^0)^\dagger\gamma^0 = \gamma^0\gamma^0\gamma^0 = \gamma^0$
- $\gamma^0(\gamma^i)^\dagger\gamma^0 = -\gamma^0\gamma^i\gamma^0 = \gamma^i$
- $\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu$
- $\gamma^0(\sigma^{\mu\nu})^\dagger\gamma^0 = \gamma^0\frac{1}{2}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger]\gamma^0 = \frac{i}{2}\{\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\} = \sigma^{\mu\nu}$

Now consider

$$\gamma^0 S(\boldsymbol{L})^\dagger\gamma^0 = e^{\frac{i}{4}\omega_{\mu\nu}\gamma^0(\sigma^{\mu\nu})^\dagger\gamma^0} = e^{\frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu}} = (S(\boldsymbol{L}))^{-1}. \tag{2.129}$$

So $\gamma^0 S(\boldsymbol{L})^\dagger\gamma^0 = S^{-1}(\boldsymbol{L}) = S(\boldsymbol{L}^{-1})$. From this it follows that

$$\begin{aligned} \overline{\psi}(x) = \psi^\dagger(x)\gamma^0 \xrightarrow{\boldsymbol{L}} \; & \psi^\dagger(\boldsymbol{L}^{-1}x)S^\dagger(\boldsymbol{L})\gamma^0 = \psi^\dagger(\boldsymbol{L}^{-1}x)\gamma^0\gamma^0 S(\boldsymbol{L})^\dagger\gamma^0 \\ = \; & \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L}). \end{aligned} \tag{2.130}$$

So finally we have the following transformation rules

$$\psi(x) \xrightarrow{\boldsymbol{L}} \psi_{\boldsymbol{L}}(x) = S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\,,$$
$$\overline{\psi}(x) \xrightarrow{\boldsymbol{L}} \overline{\psi}_{\boldsymbol{L}}(x) = \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})\,. \tag{2.131}$$

These relations are key in forming bilinears.

Classification of bilinears

First, we begin with **scalar bilinears**

$$\begin{aligned}\overline{\psi}(x)\psi(x) \xrightarrow{\boldsymbol{L}} \overline{\psi}_{\boldsymbol{L}}(x)\psi_{\boldsymbol{L}}(x) &= \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x)\\ &= \overline{\psi}(\boldsymbol{L}^{-1}x)\psi(\boldsymbol{L}^{-1}x) \equiv s(x).\end{aligned} \tag{2.132}$$

Which can be rewritten as

$$s(x) \xrightarrow{\boldsymbol{L}} s_{\boldsymbol{L}}(x) = s(\boldsymbol{L}^{-1}x). \tag{2.133}$$

This is transformation property of scalar field. Next, the **vector fields** (or vector currents) are defined as

$$J^{\mu}(x) = \overline{\psi}(x)\gamma^{\mu}\psi(x). \tag{2.134}$$

Transforming this current we obtain

$$J^{\mu}(x) \xrightarrow{\boldsymbol{L}} J^{\mu}_{\boldsymbol{L}}(x) = \overline{\psi}_{\boldsymbol{L}}(x)\gamma^{\mu}\psi_{\boldsymbol{L}}(x) = \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})\gamma^{\mu}S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x). \tag{2.135}$$

Recall that from the proof of Lorentz invariance for the Dirac equation the following holds

$$S^{-1}(\boldsymbol{L})\gamma^{\mu}S(\boldsymbol{L}) = \boldsymbol{L}^{\mu}{}_{\nu}\gamma^{\nu}. \tag{2.136}$$

Thus we can rewrite

$$J^{\mu}_{\boldsymbol{L}}(x) = \overline{\psi}(\boldsymbol{L}^{-1}x)\boldsymbol{L}^{\mu}{}_{\nu}\gamma^{\nu}\psi(\boldsymbol{L}^{-1}x) = \boldsymbol{L}^{\mu}{}_{\nu}\overline{\psi}(\boldsymbol{L}^{-1}x)\gamma^{\nu}\psi(\boldsymbol{L}^{-1}x). \tag{2.137}$$

Therefore

$$J^{\mu}_{\boldsymbol{L}} = \boldsymbol{L}^{\mu}{}_{\nu}\overline{\psi}(\boldsymbol{L}^{-1}x)\gamma^{\nu}\psi(\boldsymbol{L}^{-1}x) = \boldsymbol{L}^{\mu}{}_{\nu}J^{\nu}(\boldsymbol{L}^{-1}x). \tag{2.138}$$

This is the correct transformation law for a **vector field**.

In order to discuss pseudoscalars and pseudovectors, we will introduce a new γ-matrix

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \sigma^1 \otimes \mathbf{1} = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}. \tag{2.139}$$

Properties of γ^5

- $(\gamma^5)^{\dagger} = \gamma^5$

- $(\gamma^5)^2 = \mathbf{1}$
- $\{\gamma^5, \gamma^\mu\} = 0$

Now, the bilinear

$$P(x) = \overline{\psi}(x)\gamma^5\psi(x) \tag{2.140}$$

is a **pseudoscalar**, i.e. under standard (orthochronous) Lorentz transformation it behaves like a scalar and changes sign under the parity transformation. To see this, let us realize that

$$\gamma^5 = -i\gamma^1\gamma^0\gamma^2\gamma^3 = i\gamma^1\gamma^0\gamma^3\gamma^2 = \frac{i}{4!}\varepsilon_{\mu\nu\sigma\tau}\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau, \tag{2.141}$$

and also that

$$\varepsilon^{\mu\nu\sigma\tau}\gamma^5 = i\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau. \tag{2.142}$$

Here $\varepsilon_{\mu\nu\sigma\tau}$ is a Levi-Civita symbol. $\varepsilon_{\mu\nu\sigma\tau} = 1$ if $(\mu\nu\sigma\tau)$ is even permutation of (0123) and $\varepsilon_{\mu\nu\sigma\tau} = -1$ if $(\mu\nu\sigma\tau)$ is odd permutation of (0123), otherwise $\varepsilon_{\mu\nu\sigma\tau} = 0$. Continuing from (2.140) we have

$$\begin{aligned} P(x) = \overline{\psi}(x)\gamma^5\psi(x) \xrightarrow{\boldsymbol{L}} & \quad \overline{\psi}_{\boldsymbol{L}}(x)\gamma^5\psi_{\boldsymbol{L}}(x) \\ = & \quad \overline{\psi}(\boldsymbol{L}^{-1}x)S^{-1}(\boldsymbol{L})\gamma^5 S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x). \end{aligned} \tag{2.143}$$

We can rewrite term $S^{-1}(\boldsymbol{L})\gamma^5 S(\boldsymbol{L})$ as follows:

$$\begin{aligned} S^{-1}(\boldsymbol{L})\gamma^5 S(\boldsymbol{L}) &= \frac{i}{4!}\varepsilon_{\mu\nu\sigma\tau}S^{-1}(\boldsymbol{L})\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\tau S(\boldsymbol{L}) \\ &= \frac{i}{4}\varepsilon_{\mu\nu\sigma\tau}(S^{-1}(\boldsymbol{L})\gamma^\mu S(\boldsymbol{L}))\dots(S^{-1}(\boldsymbol{L})\gamma^\tau S(\boldsymbol{L})) \\ &= \frac{i}{4!}\varepsilon_{\mu\nu\sigma\tau}\boldsymbol{L}^\mu{}_{\mu'}\boldsymbol{L}^\nu{}_{\nu'}\boldsymbol{L}^\sigma{}_{\sigma'}\boldsymbol{L}^\tau{}_{\tau'}\gamma^{\mu'}\gamma^{\nu'}\gamma^{\sigma'}\gamma^{\tau'} \\ &= \gamma^5(\det\boldsymbol{L}), \end{aligned} \tag{2.144}$$

$\det\boldsymbol{A} = \sum_{J_1,\dots,J_n}\varepsilon_{J_1,\dots,J_n}A_{1J_1}\dots A_{nJ_n}$.

and hence

$$P(x) = \overline{\psi}(x)\gamma^5\psi(x) \xrightarrow{\boldsymbol{L}} \det\boldsymbol{L}\,\overline{\psi}_{\boldsymbol{L}}(x)\gamma^5\psi_{\boldsymbol{L}}(x) = \det\boldsymbol{L}\,P(\boldsymbol{L}^{-1}x). \tag{2.145}$$

So, the function $P(x)$ is a Lorentz scalar for proper Lorentz transformations ($\det\boldsymbol{L} = 1$).

Now, for the Lorentz transformations involving parity reversal

$$\boldsymbol{L}^P = \mathrm{diag}(1, -1, -1, -1), \tag{2.146}$$

the transformation changes sign ($\det\boldsymbol{L} = -1$). Complete set of bilinears is given in the following table.

Table 2.1: Table of bilinears

Bilinear	Transformation properties
$\overline{\psi}(x)\psi(x)$	Scalar
$\overline{\psi}(x)\gamma^5\psi(x)$	Pseudoscalar
$\overline{\psi}(x)\gamma^\mu\gamma^5\psi(x)$	Pseudovector field
$\overline{\psi}(x)\gamma^\mu\psi(x)$	Vector field
$\overline{\psi}(x)[\gamma^\mu, \gamma^\nu]\psi(x)$	Antisymmetric tensor field

All bilinears have the form $\overline{\psi}(x)\Gamma\psi(x)$, where Γ is one of 16 possible matrices - $\mathbf{1}$, γ^μ, $[\gamma^\mu, \gamma^\nu]$, γ^5 and $\gamma^5\gamma^\mu$.

2.7 Current for a Dirac wave function

Main motivation of Dirac was to have consistent probability current with positive definite probability. There is now a natural candidate for a probability current, namely

$$J_\mu(x) = \overline{\psi}(x)\gamma_\mu\psi(x). \tag{2.147}$$

The norm can be defined via current as

$$(\psi, \psi) = ||\psi||^2 = \int \mathrm{d}^3x J_0(x) = \int \mathrm{d}V n^\mu J_\mu(x). \tag{2.148}$$

The last integral is over a space-like slice orthogonal to the 4-vector n^μ (time-like). We want to show the following:

- This is a Lorentz invariant norm.
- Such norm is time independent - $\partial_\mu J^\mu(x) = 0$.

Proving that this is the case will show, that this norm is independent of the space-like slice (i.e. Lorentz invariant).

First, let us show that $\partial_\mu J^\mu(x) = 0$:

$$\partial_\mu\left(\overline{\psi}(x)\gamma^\mu\psi(x)\right) = \left(\partial_\mu\overline{\psi}(x)\right)\gamma^\mu\psi(x) + \overline{\psi}(x)\gamma^\mu\left(\partial_\mu\psi(x)\right). \tag{2.149}$$

Recalling that $\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = 0$ and hence

$$\gamma^\mu\partial_\mu\psi(x) = -im\psi(x) \tag{2.150}$$

For the adjoint wave function we obtain an equation of motion

$$\left[\left(i\gamma^\mu\partial_\mu - m\right)\psi(x)\right]^\dagger = 0 \tag{2.151}$$

$$-i\partial_\mu\psi^\dagger(x)\left(\gamma^\mu\right)^\dagger - m\psi^\dagger(x) = 0. \tag{2.152}$$

Taking advantage of the fact that

$$\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu \iff (\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0, \tag{2.153}$$

we can rewrite

$$\begin{aligned}\underbrace{\left(i\partial_\mu\psi^\dagger(x)(\gamma^\mu)^\dagger + m\psi^\dagger(x)\right)}_{0}\gamma^0 &= 0\\ \left(i\partial_\mu\psi^\dagger(x)\gamma^0\gamma^\mu + m\psi^\dagger(x)\gamma^0\right) &= 0\\ \left(i\partial_\mu\overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x)\right) &= 0\\ \overline{\psi}(x)\left(i\gamma^\mu\overleftarrow{\partial_\mu} + m\right) &= 0.\end{aligned} \tag{2.154}$$

From this we see that

$$\left(\partial_\mu \overline{\psi}(x)\right)\gamma^\mu = im\overline{\psi}. \tag{2.155}$$

Plugging this into the (2.149) we obtain

$$\left(\partial_\mu \overline{\psi}(x)\right)\gamma^\mu \psi(x) + \overline{\psi}(x)\gamma^\mu\left(\partial_\mu \psi(x)\right) = im\overline{\psi}(x)\psi(x) - im\overline{\psi}(x)\psi(x) = 0. \tag{2.156}$$

Another possible current is

$$J^\mu(x) = \overline{\psi}(x)\gamma^\mu\gamma^5\psi(x). \tag{2.157}$$

It can be shown that in this case $\partial_\mu J^\mu(x) = 2im\overline{\psi}(x)\gamma^5\psi(x)$. If $m = 0$, this current (often called the axial vector current) is also conserved.

Let's return back to the norm. Choose $n = (1, 0, 0, 0)$, then

$$\begin{aligned}(\psi,\psi) &= \int \mathrm{d}^3x J_0(x) = \int \mathrm{d}^3x \overline{\psi}(x)\gamma^0\psi(x) = \\ &= \int \mathrm{d}^3x\psi^\dagger(x)\gamma^0\gamma^0\psi(x) = \int \mathrm{d}^3x\psi^\dagger(x)\psi(x) \geq 0.\end{aligned} \tag{2.158}$$

From this we can see, that such norm is a positive definite. Norm is also clearly time independent, since

$$-\underbrace{\frac{\partial}{\partial t}\int_V \mathrm{d}^3x\rho}_{\substack{\text{Change in total}\\ \text{probability inside } V}} = \int_V \mathrm{d}^3x\boldsymbol{\nabla}\cdot\boldsymbol{J} = \underbrace{\int_{\partial V}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J}}_{\substack{\text{Flux of } \boldsymbol{J}\\ \text{through the } \partial V}} \to 0. \tag{2.159}$$

Here we use same argument as in Klein-Gordon equation case. At last, norm is relativistically invariant

$$\begin{aligned}||\psi||^2 &= \int \mathrm{d}^3\boldsymbol{x}\rho(x) = \int \mathrm{d}^4x \underbrace{\delta(t_0)}_{\frac{\partial}{\partial x^0}\theta(x_0)} \rho(x) = \\ &= \int \mathrm{d}^4x J^\alpha \frac{\partial}{\partial x^\alpha}\theta(n^\beta x_\beta).\end{aligned} \tag{2.160}$$

Defining

$$||\widetilde{\psi}||^2 = \int \mathrm{d}^4x J^\alpha \frac{\partial}{\partial x^\alpha}\theta(n'^\beta x_\beta), \tag{2.161}$$

we finally obtain

$$\begin{aligned}||\psi||^2 - ||\widetilde{\psi}||^2 &= \int \mathrm{d}^4x J^\alpha \frac{\partial}{\partial x^\alpha}\left(\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\right) \\ &\overset{\partial_\alpha J^\alpha = 0}{=} \int \mathrm{d}^4x \frac{\partial}{\partial x^\alpha}\left[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\right] \\ &= \int \mathrm{d}S_\alpha \left[J^\alpha\{\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)\}\right].\end{aligned} \tag{2.162}$$

To show that this is zero, consider two possibilities:

1. J^α is presumed to vanish if $|\boldsymbol{x}| \to \infty$ with fixed t.
2. $\theta(n^\beta x_\beta) - \theta(n'^\beta x_\beta)$ vanishes for $|t| \to \infty$ with $\boldsymbol{x}$ fixed.

Hence, the difference is zero and norm is relativistically invariant. To define related scalar product we take

$$(\psi_1, \psi_2) = \int \mathrm{d}V n_\mu J^{\mu(1,2)}(x), \tag{2.163}$$

where $J^{\mu(1,2)}(x) = \overline{\psi}_1(x)\gamma^\mu\psi_2(x)$.

Plane Wave Solutions of Dirac Equation

We know that because $\psi(x)$ satisfies Dirac equation, it also satisfies Klein-Gordon equation. Wave functions of definite energy and momentum are plane waves of the form

$$\begin{aligned}\psi_p^+(x) &= u(p)e^{-ipx} = u(p)e^{-i\omega_{\boldsymbol{p}}t+i\boldsymbol{p}\cdot\boldsymbol{x}} \text{ (positive energy)},\\ \psi_p^-(x) &= v(p)e^{ipx} = v(p)e^{i\omega_{\boldsymbol{p}}t-i\boldsymbol{p}\cdot\boldsymbol{x}} \text{ (negative energy)},\end{aligned} \tag{2.164}$$

where $p_0 = \omega_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2} > 0$. For the plane wave $i\partial_\mu \to p_\mu$ and given that $\left(i\gamma^\mu\partial_\mu - m\right)\psi_p^+ = 0$ we obtain

$$\left(\gamma^\mu p_\mu - m\right) u(p) = 0 \iff \left(\not{p} - m\right) u(p) = 0, \tag{2.165}$$

$$\left(\gamma^\mu p_\mu + m\right) v(p) = 0 \iff \left(\not{p} + m\right) v(p) = 0. \tag{2.166}$$

In order to have a non-trivial solutions we require $\det\left(\not{p} - m\right) = 0$ and $\det\left(\not{p} + m\right) = 0$. We will use the following trick for computation:

$$\begin{aligned}\det\left(\gamma^\mu p_\mu - m\right) &= \det\left(\gamma^5\gamma^5\left(\gamma^\mu p_\mu - m\right)\right)\\ &= \det\left(\gamma^5\left(\gamma^\mu p_\mu - m\right)\gamma^5\right)\\ &= \det\left(\left(-\gamma^\mu p_\mu - m\right)\gamma^5\gamma^5\right)\\ &= \det\left(\gamma^\mu p_\mu + m\right).\end{aligned} \tag{2.167}$$

The second equation follows from property of determinant and the third from anticommutativity. From this we can see that

$$\det\left[\left(\gamma^\mu p_\mu - m\right)\left(\gamma^\mu p_\mu + m\right)\right] = \det{}^2\left(\gamma^\mu p_\mu \pm m\right). \tag{2.168}$$

Using properties of gamma-matrices we can further show that

$$\gamma^\nu\gamma^\mu p_\mu p_\nu - m^2 = \frac{1}{2}\left\{\gamma^\mu, \gamma^\nu\right\} p_\mu p_\nu - m^2 = p^2 - m^2, \tag{2.169}$$

and hence

$$\det{}^2\left(\gamma^\mu p_\mu \pm m\right) = (p^2 - m^2)^4 = 0. \tag{2.170}$$

Last condition is known as on mass shell condition.

Positive Energy Solutions

From the representation of the γ^μ we find that

$$\left(\gamma^\mu p_\mu - m\right) = \begin{pmatrix} (E-m)\mathbf{1} & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E+m)\mathbf{1} \end{pmatrix}. \tag{2.171}$$

Take $u(p)$ to have form of

$$u(p) = \begin{pmatrix} \chi \\ \varphi \end{pmatrix}, \tag{2.172}$$

here χ and φ have 2 components each. From this we can rewrite equation (2.166) as a system of two coupled equations

$$(E-m)\chi - \boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi = 0\,, \tag{2.173}$$
$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi - (E+m)\varphi = 0\,. \tag{2.174}$$

The condition $p^2 = m^2$ ensures, that these two equations are equivalent. Indeed multiplying (2.173) by $(E+m)$ we obtain

$$\underbrace{(E^2-m^2)}_{\boldsymbol{p}^2}\chi - \boldsymbol{\sigma}\cdot\boldsymbol{p}(E+m)\varphi = 0 \tag{2.175}$$

Similarly, by multiplying (2.174) by $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ we get

$$\sigma^i\sigma^j p_i p_j \chi - \boldsymbol{\sigma}\cdot\boldsymbol{p}(E+m)\varphi = 0. \tag{2.176}$$

Using the fact that $\sigma^i\sigma^j p_i p_j$ can be rewritten as $\frac{1}{2}\{\sigma^i,\sigma^j\}p_i p_j = \delta^{ij}p_i p_j = \boldsymbol{p}^2$ we get that those two equations are indeed equivalent and any of the two equations is okay to use. From the second equation we obtain

$$\varphi = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi, \tag{2.177}$$

which implies that

$$u(p) \propto \begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix}. \tag{2.178}$$

We fix the normalization of $u(p)$ so that

$$u(p) = \sqrt{E+m}\begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix}. \tag{2.179}$$

Other Normalizations

Often the normalization is chosen differently, namely

$$u(p) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{E+m}{2m}}\chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{\sqrt{2m(E+m)}}\chi \end{pmatrix}. \tag{2.180}$$

This gives nicer normalization for $\bar{u}u$ and $\bar{v}v$.

There is another, more physical way of solving $\left(\not{p} - m\right)u(p) = 0$ and

$(\not{p}+m)\, v(p) = 0$. Let assume that $m \neq 0$. In the rest frame of the particle $p^\mu = (m, \mathbf{0})$. Then Dirac equation reduces to

$$\left(\gamma^0 - \mathbf{1}\right) u(m, \mathbf{0}) = 0\,, \tag{2.181}$$

$$\left(\gamma^0 + \mathbf{1}\right) v(m, \mathbf{0}) = 0\,. \tag{2.182}$$

There are clearly 2 linearly independent solutions for both u and v, namely

$$u^{(1)}(m,\mathbf{0}) = \begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ u^{(2)}(m,\mathbf{0}) = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ v^{(1)}(m,\mathbf{0}) = \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ v^{(2)}(m,\mathbf{0}) = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{2.183}$$

We could now boost these solutions from rest up to a velocity $|\boldsymbol{v}| = \frac{|\boldsymbol{p}|}{p_0}$ by a pure Lorentz transformation $S(\Lambda) = e^{-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}}$.

Relativistic Velocity

From $p^\mu = \left(\frac{E}{c}, \boldsymbol{p}\right) = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}(c, \boldsymbol{v})$. Setting $c = 1$ we finally have that $|\boldsymbol{v}| = \frac{|\boldsymbol{p}|}{p_0}$.

It is simpler to observe that

$$(\not{p}-m)\,(\not{p}+m) = p^2 - m^2 = (\not{p}+m)\,(\not{p}-m) = 0. \tag{2.184}$$

Hence

$$(\not{p}-m)\,(\not{p}+m)\, u^{(\lambda)} = 0,\ \lambda = 1, 2. \tag{2.185}$$

This implies that

$$\begin{aligned}(\not{p}+m)\, u^{(\lambda)}(m,\mathbf{0}) &= \begin{pmatrix} E+m & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E-m)\end{pmatrix} u^{(\lambda)}(m,\mathbf{0}) \\ &= \begin{pmatrix}(E+m)\varphi^{(\lambda)} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi^{(\lambda)}\end{pmatrix}.\end{aligned} \tag{2.186}$$

Here we take $u^{(\lambda)}(m,\mathbf{0}) = \begin{pmatrix}\varphi^{(\lambda)}(m,\mathbf{0})\\0\\0\end{pmatrix}$ and $\varphi^{(\lambda)}$ is either $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$. Negative energy solution is obtained similarly

$$\begin{aligned}(\not{p}-m)\, v^{(\lambda)}(m,\mathbf{0}) &= \begin{pmatrix} E-m & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E+m)\end{pmatrix} v^{(\lambda)}(m,\mathbf{0}) \\ &= \begin{pmatrix}-\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi^{(\lambda)} \\ -(E+m)\chi^{(\lambda)}\end{pmatrix}.\end{aligned} \tag{2.187}$$

Where in analogy with positive energy solution, we have $v^{(\lambda)}(m,\mathbf{0}) =$

$\begin{pmatrix} 0 \\ 0 \\ \chi^{(\lambda)}(m, \mathbf{0}) \end{pmatrix}$ and $\chi^{(\lambda)}$ is again $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Relation with our former normalization is obtained when

$$u^{(\lambda)}(p) = \frac{\not{p} + m}{\sqrt{E + m}} u^{(\lambda)}(m, \mathbf{0}), \tag{2.188}$$

$$v^{(\lambda)}(p) = \frac{-\not{p} + m}{\sqrt{E + m}} v^{(\lambda)}(m, \mathbf{0}). \tag{2.189}$$

Relation to the Schrödinger Equation

In this part we will assume that we work with free particle. Recall that positive and negative energy solutions to the massive Dirac equation have the form

$$u_\lambda = \sqrt{E + m} \begin{pmatrix} \chi_\lambda \\ \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E+m} \chi_\lambda \end{pmatrix} = \begin{pmatrix} u_L \\ u_S \end{pmatrix}. \tag{2.190}$$

Note that $u_S = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E+m} u_L$. Similarly for negative energy solution we have

$$v_\lambda = \sqrt{E + m} \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E+m} \varphi_\lambda \\ \varphi_\lambda \end{pmatrix} = \begin{pmatrix} v_S \\ v_L \end{pmatrix}, \tag{2.191}$$

and $v_S = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E+m} v_L$. In the non-relativistic limit $|\boldsymbol{p}| \ll m$ and so $u_S \ll u_L$ and $v_S \ll v_L$. Hence the subscript "S" references to the **small component** and the subscript "L" to the large component.

The positive energy solutions satisfy

$$\begin{pmatrix} m\mathbf{1} & \boldsymbol{\sigma} \cdot \boldsymbol{p} \\ \boldsymbol{\sigma} \cdot \boldsymbol{p} & -m\mathbf{1} \end{pmatrix} \begin{pmatrix} u_L \\ u_S \end{pmatrix} = E \begin{pmatrix} u_L \\ u_S \end{pmatrix}. \tag{2.192}$$

This equation can be rewritten as

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} u_S = (E - m) u_L \,, \tag{2.193}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} u_L = (E + m) u_S \,. \tag{2.194}$$

Substituting (2.194) to (2.193) gives

$$\frac{(\boldsymbol{\sigma} \cdot \boldsymbol{p})(\boldsymbol{\sigma} \cdot \boldsymbol{p})}{E + m} u_L = (E - m) u_L \,. \tag{2.195}$$

Since $|\boldsymbol{p}| \ll m$ and $E \approx m$ we obtain

$$\frac{\boldsymbol{p}^2}{2m} u_L = (E - m) u_L = E_{NR} u_L. \tag{2.196}$$

This is the usual Schrödinger equation for a free non-relativistic particle.

Applications of Lorentz Transformations on Dirac Wave Functions

We will begin with **rotations of Dirac wave functions**. We know that $\{R\} \subset \{L\}$. Further on $\omega_{\mu\nu}(R)$ will denote parameters of the rotation. We have seen that $L = e^{-\frac{i}{4}\boldsymbol{M}^{\mu\nu}\omega_{\mu\nu}}$ where the generator of rotations was connected with $\boldsymbol{M}^{ij}$ $(i, j = 1, 2, 3)$ as

$$\boldsymbol{J}_i = \frac{1}{4}\varepsilon_{ijk}\boldsymbol{M}^{jk}, \tag{2.197}$$

Do not forget that $[\boldsymbol{J}_i, \boldsymbol{J}_j] = i\varepsilon_{ijk}J_k$ - i.e. there is an algebra of rotations.

$$\boldsymbol{M}^{jk} = 2\varepsilon^{jki}\boldsymbol{J}_i. \tag{2.198}$$

In particular, when M^{jk} acts on 4-vectors

$$(\boldsymbol{M}^{\mu\nu})^{\rho}{}_{\tau} = 2i\left[\boldsymbol{g}^{\mu\rho}\delta^{\nu}{}_{\tau} - \boldsymbol{g}^{\nu\rho}\delta^{\mu}{}_{\tau}\right]. \tag{2.199}$$

This gives us

$$(\boldsymbol{J}_i)^{\rho}{}_{\tau} = \frac{i}{2}\varepsilon_{ijk}\left[\boldsymbol{g}^{\rho j}\delta^{k}{}_{\tau} - \boldsymbol{g}^{\rho k}\delta^{j}{}_{\tau}\right]. \tag{2.200}$$

E.g. for third component

$$(\boldsymbol{J}_3)^{\rho}{}_{\tau} = \frac{i}{2}\varepsilon_{3jk}\left[\boldsymbol{g}^{\rho j}\delta^{k}{}_{\tau} - \boldsymbol{g}^{\rho k}\delta^{j}{}_{\tau}\right]. \tag{2.201}$$

From this we see, that j has can be either 1 or 2, k can be 2 or 1. Hence non trivial contributions to the $(\boldsymbol{J}_3)^{\rho}{}_{\tau}$ come only from components $\rho = 1, 2$, $\tau = 2, 1$. Namely

$$(\boldsymbol{J}_3)^{1}{}_{2} = \frac{i}{2}\varepsilon_{312}\boldsymbol{g}^{11}\delta^{2}_{2} - \frac{i}{2}\varepsilon_{321}\boldsymbol{g}^{11}\delta^{2}_{2} = -i. \tag{2.202}$$

Upper index signifies rows, lower index signifies columns.

$$(\boldsymbol{J}_3)^{2}{}_{1} = \frac{i}{2}\varepsilon_{321}\boldsymbol{g}^{22}\delta^{1}_{1} - \frac{i}{2}\varepsilon_{312}\boldsymbol{g}^{22}\delta^{1}_{1} = i. \tag{2.203}$$

Thus we can finally write our $\boldsymbol{J}_3$ in explicit form as

$$\boldsymbol{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \boldsymbol{R}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.204}$$

Similarly we obtain

$$\boldsymbol{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \to \boldsymbol{R}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}, \tag{2.205}$$

$$\boldsymbol{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} \to \boldsymbol{R}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{2.206}$$

Recall that $\boldsymbol{J}_k = i\left.\frac{\partial R}{\partial \boldsymbol{\theta}^k}\right|_{\boldsymbol{\theta}=0}$.

On the other hand

$$L \supset R = e^{-\frac{i}{4}\omega_{ij}\boldsymbol{M}^{ij}} = e^{-\frac{i}{2}\varepsilon^{ijk}\omega_{ij}\boldsymbol{J}_k} = e^{-i\boldsymbol{\theta}^k\boldsymbol{J}_k}. \tag{2.207}$$

Where we have defined $\boldsymbol{\theta}^k = \frac{1}{2}\varepsilon^{ijk}\omega_{ij}$. From this we can see that

$$\boldsymbol{R}_3 = e^{-i\boldsymbol{\theta}^3\boldsymbol{J}_3}, \tag{2.208}$$

and generally

$$\boldsymbol{R} = e^{-i\boldsymbol{\theta}^i\boldsymbol{J}_i}. \tag{2.209}$$

We can equivalently write that

$$\boldsymbol{\theta}^i\boldsymbol{J}_i = \theta\boldsymbol{n}^i\boldsymbol{J}_i, \tag{2.210}$$

where $\boldsymbol{n}$ is a unit 3-vector and θ is rotation angle around that direction.

Deriving Explicit Form of Matrix Exponential

We can obtain explicit form of matrix exponential by summing up all terms in Taylor expansion

$$\begin{aligned} e^{-i\theta^3 J_3} &= \sum_{n=0}^{\infty} \frac{(-i\theta^3)^n}{n!} J_3^n = \sum_{n=0}^{\infty} \frac{(-i\theta^3)^{2n}}{2n!} J_3^{2n} + \sum_{n=0}^{\infty} \frac{(-i\theta^3)^{2n+1}}{(2n+1)!} J_3^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} (\theta^3)^{2n} \mathbf{1} - i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (\theta^3)^{2n+1} J_3 \\ &= \cos\theta^3 \mathbf{1} - i \sin\theta^3 J_3 . \end{aligned} \tag{2.211}$$

The spinor representation of rotation is given by

$$L = e^{-\frac{i}{4} M^{\mu\nu}\omega_{\mu\nu}} \leftrightarrow S(L) = e^{-\frac{i}{4}\sigma^{\mu\nu}\omega_{\mu\nu}} . \tag{2.212}$$

$$R = e^{-\frac{i}{4} M^{ij}\omega_{ij}} = e^{-i\theta n^i J_i} \leftrightarrow S(R) = e^{-\frac{i}{4}\sigma^{ij}\omega_{ij}} = e^{-i\theta n^i \hat{\sigma}_i} . \tag{2.213}$$

Here $\hat{\sigma}_i = \frac{1}{4}\varepsilon_{ijk}\sigma^{jk} \iff \sigma^{jk} = 2\varepsilon^{ijk}\hat{\sigma}_i$. Let us find explicit form of $\hat{\sigma}_i$. Starting from the definition of σ^{ij} for $i \neq j$:

$$\begin{aligned} \sigma^{ij} &= \frac{i}{2}\left[\gamma^i, \gamma^j\right] = i\gamma^i\gamma^j \\ &= i \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} = i \begin{pmatrix} -\sigma_i\sigma_j & 0 \\ 0 & -\sigma_i\sigma_j \end{pmatrix} \\ &= \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \varepsilon_{ijk} . \end{aligned} \tag{2.214}$$

Using definition of $\hat{\sigma}_k$ we get

$$\begin{aligned} \hat{\sigma}_k &= \frac{1}{4}\varepsilon_{klm}\sigma^{lm} = \frac{1}{4}\varepsilon_{klm}\varepsilon^{lmp} \begin{pmatrix} \sigma_p & 0 \\ 0 & \sigma_p \end{pmatrix} = \\ &= \frac{1}{4}\varepsilon_{lmk}\varepsilon^{lmp} \begin{pmatrix} \sigma_p & 0 \\ 0 & \sigma_p \end{pmatrix} = \\ &= \frac{1}{4}(\delta_m^m \delta_k^p - \delta_m^p \delta_k^m) \begin{pmatrix} \sigma_p & 0 \\ 0 & \sigma_p \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} . \end{aligned} \tag{2.215}$$

Rotation in Bispinor Space

Taking commutator

$$[\hat{\sigma}_k, \hat{\sigma}_l] = \begin{pmatrix} \frac{1}{2}[\sigma_k, \sigma_l] & 0 \\ 0 & \frac{1}{2}[\sigma_k, \sigma_l] \end{pmatrix} = \frac{i}{2}\varepsilon_{klm} \begin{pmatrix} \sigma_m & 0 \\ 0 & \sigma_m \end{pmatrix} = i\varepsilon_{klm}\hat{\sigma}_m . \tag{2.216}$$

Hence $\hat{\sigma}_i$ are generators of rotations in bispinor space.

By denoting

$$2\hat{\sigma}_k = \Sigma_k = \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}, \tag{2.217}$$

we obtain that

$$S(\boldsymbol{R}) = e^{-\frac{i}{2}\theta \boldsymbol{n}^k \Sigma_k} = \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{n}_k \sigma_k} & 0 \\ 0 & e^{-\frac{i}{2}\boldsymbol{n}_k \sigma_k} \end{pmatrix}. \tag{2.218}$$

Now, consider action of $S(\boldsymbol{R})$ on $u(p)$:

$$\begin{aligned} S(\boldsymbol{R})u(p) &= \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{n}_k \sigma_k} & 0 \\ 0 & e^{-\frac{i}{2}\boldsymbol{n}_k \sigma_k} \end{pmatrix} u(p) \\ &= \sqrt{E+m} \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{n}_k \sigma_k} & 0 \\ 0 & e^{-\frac{i}{2}\boldsymbol{n}_k \sigma_k} \end{pmatrix} \begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix} \\ &= \sqrt{E+m} \begin{pmatrix} e^{-\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi \\ \frac{e^{-\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}} \boldsymbol{\sigma}\cdot\boldsymbol{p} e^{\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}}}{E+m} e^{-\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi \end{pmatrix}. \end{aligned} \tag{2.219}$$

Define $\chi_{\boldsymbol{R}} = e^{-\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}}\chi$, where $\frac{\boldsymbol{\sigma}}{2}$ are generators of rotations in representation with spin 1/2. At this stage we can recognize correctly rotated 2-component spinor (Pauli spinor) with spin 1/2. We have the following transformation relationship

Compare this relation to the now know relation $S^{-1}(\boldsymbol{L})\gamma^\mu S(\boldsymbol{L}) = \boldsymbol{L}^\mu{}_\nu \gamma^\nu = \gamma^\nu \boldsymbol{L}^\mu{}_\nu$.

$$e^{-\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}} \boldsymbol{\sigma}\cdot\boldsymbol{p} e^{\frac{i}{2}\theta \boldsymbol{n}\cdot\boldsymbol{\sigma}} = \boldsymbol{\sigma} e^{-i\theta \boldsymbol{n}\cdot\boldsymbol{J}} \cdot \boldsymbol{p}. \tag{2.220}$$

From this we finally get rotated spinor as

$$S(\boldsymbol{R})u(p) = \begin{pmatrix} \chi_{\boldsymbol{R}} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}_{\boldsymbol{R}}}{E+m}\chi_{\boldsymbol{R}} \end{pmatrix} = u_{\boldsymbol{R}}(p). \tag{2.221}$$

$u_{\boldsymbol{R}}(p)$ is constructed from the rotated spinor $\chi_{\boldsymbol{R}}$ and $\boldsymbol{p}_{\boldsymbol{R}}$ is the rotated 4-momentum. Let's introduce $\chi_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (Pauli spinors) and denote them generally as χ_λ, $\lambda = \pm\frac{1}{2}$. Then

$$u_\lambda(p) = \sqrt{E+m} \begin{pmatrix} \chi_\lambda \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \end{pmatrix}, \tag{2.222}$$

and $\lambda = \pm\frac{1}{2}$ has meaning of the "third component of spin".

Action of Rotation Generators on Free Solutions of Dirac Equation

Let us look what happens, when we apply generators of rotations (angular momentum operators) on a free solutions of Dirac equation in the rest frame. We know that those solutions have general form of

$$\psi_\lambda^+(x) = \sqrt{2m} \begin{pmatrix} \chi_\lambda \\ 0 \end{pmatrix} \frac{e^{-ik_0 t + i\boldsymbol{k}\cdot\boldsymbol{x}}}{\sqrt{2\pi}^3}, \tag{2.223}$$

$$\psi_\lambda^-(x) = -\sqrt{2m}\begin{pmatrix}0\\ \chi_\lambda\end{pmatrix}\frac{e^{ik_0t-i\boldsymbol{k}\cdot\boldsymbol{x}}}{\sqrt{2\pi}^3}, \tag{2.224}$$

with $\boldsymbol{k}\cdot\boldsymbol{x}$ being zero at rest and $k_0 = m$. We also know that

$$\boldsymbol{J}_i \leftrightarrow \hat{\boldsymbol{\sigma}} = \frac{1}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}, \tag{2.225}$$

and in particular

$$\sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{2.226}$$

This leads to

$$\boldsymbol{J}_3\psi^+_{\frac{1}{2}} = \frac{1}{2}\psi^+_{\frac{1}{2}},\ \boldsymbol{J}_3\psi^+_{-\frac{1}{2}} = -\frac{1}{2}\psi^+_{-\frac{1}{2}}, \tag{2.227}$$

$$\boldsymbol{J}_3\psi^-_{\frac{1}{2}} = -\frac{1}{2}\psi^-_{\frac{1}{2}},\ \boldsymbol{J}_3\psi^-_{-\frac{1}{2}} = \frac{1}{2}\psi^-_{-\frac{1}{2}}. \tag{2.228}$$

This shows that indeed Dirac's bispinors describe particle with spin 1/2.

2.8 Lorentz Boosts and Dirac Wave Function

Lorentz boost in x can be written as

$$\begin{aligned} t' &= \gamma(t - vx), &\Leftrightarrow\quad x_0' &= \gamma(x_0 - \beta x),\\ x' &= \gamma(x - vt), &\Leftrightarrow\quad x' &= \gamma(x - \beta x_0),\\ y' &= y, & &\\ z' &= z, & & \end{aligned} \tag{2.229}$$

where $v = \beta$ and $-1 < \beta < 1$ (note that we use the notation $c = 1$).

The Lorentz transformations are often also written in a way that resembles rotations in 3D using hyperbolic functions. This is possible, because β and γ satisfy identity

$$\gamma^2 - \gamma^2\beta^2 = \gamma^2(1-\beta^2) = 1, \tag{2.230}$$

which allows us to take $\gamma = \cosh\zeta$ and $\gamma\beta = \sinh\zeta$. Then we get

$$\begin{aligned} x_0' &= x_0\cosh\zeta - x\sinh\zeta,\\ x' &= x\cosh\zeta - x_0\sinh\zeta,\\ y' &= y,\\ z' &= z. \end{aligned} \tag{2.231}$$

Note that

$$\begin{aligned} \cosh\zeta = \gamma &= \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\tanh^2\zeta}},\\ \tanh\zeta &= \frac{\sinh\zeta}{\cosh\zeta} = \beta \Rightarrow \zeta = \tanh^{-1}\beta \end{aligned} \tag{2.232}$$

Rapidity is a standard parameter to measure relativistic velocities in particle accelerators. In 1D motion, rapidities are additive whereas velocities must be combined by Einstein's velocity-addition formula.

and since $\beta \in (-1, 1)$, we get $\zeta \in (-\infty, \infty)$. The new variable ζ is called **rapidity**.

Similar equations can be also gained for boost in z direction

$$\begin{aligned} x_0' &= x_0 \cosh\zeta - z \sinh\zeta, \\ x' &= x, \\ y' &= y, \\ z' &= z\cosh\zeta - x_0 \sinh\zeta. \end{aligned} \tag{2.233}$$

Transformations can be written using boost matrices

$$\boldsymbol{L}_x = \begin{pmatrix} \cosh\zeta & -\sinh\zeta & 0 & 0 \\ -\sinh\zeta & \cosh\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.234}$$

$$\boldsymbol{L}_z = \begin{pmatrix} \cosh\zeta & 0 & 0 & -\sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix} \tag{2.235}$$

and $\boldsymbol{L}_y$ similarly.

Consider now standard four-momentum of a particle with mass m in its rest frame, i.e. $p^\mu = (m, 0, 0, 0)$. After boosting the four-momentum in z direction, we get

$$p' = \boldsymbol{L}_z p = \begin{pmatrix} \cosh\zeta & 0 & 0 & -\sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix} \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} E \\ 0 \\ 0 \\ q \end{pmatrix}, \tag{2.236}$$

where $E = m\cosh\zeta = m\gamma$ and $q = -m\sinh\zeta = -mv\gamma$ (relativistic three-momentum). So, in this example, $\tanh\zeta = -q/E$.

If ζ is infinitesimal ($\ll 1$), it allows us to simplify the equations (e.g. for a boost in z direction)

$$\boldsymbol{L}_z \approx \begin{pmatrix} 1 & 0 & 0 & -\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\zeta & 0 & 0 & 1 \end{pmatrix} = \mathbf{1} + \zeta \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{2.237}$$

At the same time, we can use the Eq. 2.109 to compare the result above to

$$\boldsymbol{L}_z \approx \mathbf{1} - \frac{i}{4}\omega_{\mu\nu}\boldsymbol{M}^{\mu\nu} = \mathbf{1} - \frac{i}{4}(\omega_{03}\boldsymbol{M}^{03} + \omega_{30}\boldsymbol{M}^{30}) = \mathbf{1} - \frac{i}{2}\omega_{03}\boldsymbol{M}^{03} \tag{2.238}$$

because $\omega_{\mu\nu} = 0$ except for $\omega_{03} = -\omega_{30} = \zeta$. We remind that $\boldsymbol{K}^i = \frac{1}{2}\boldsymbol{M}^{0i}$

holds. Therefore, it is obvious that

$$\boldsymbol{K}^3 = i\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{M}^{03} = 2i\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{2.239}$$

We can go back to the finite transformations in the form

$$\begin{aligned} \boldsymbol{L}_z = \exp\left(-\frac{i}{2}\omega_{03}\boldsymbol{M}^{03}\right) &= \exp\left\{\zeta\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}\right\} = \\ &= \begin{pmatrix} \cosh\zeta & 0 & 0 & -\sinh\zeta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\zeta & 0 & 0 & \cosh\zeta \end{pmatrix}, \end{aligned} \tag{2.240}$$

where in the last step we used a fact that the matrix $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ is involutive.

This derivation served as an independent control that the result is consistent with Eq. 2.235.

In the spinor space then the corresponding matrix is

$$S(\boldsymbol{L}) = \exp\left\{-\frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu}\right\}, \tag{2.241}$$

which means that specifically in the z direction

$$S(\boldsymbol{L}_z) = \exp\left\{-\frac{i}{2}\zeta\sigma^{03}\right\}, \tag{2.242}$$

where $\sigma^{03} = \frac{i}{2}\left[\gamma^0, \gamma^3\right] = i\gamma^0\gamma^3$, so

$$S(\boldsymbol{L}_z) = \exp\left\{-\frac{i}{2}\zeta i\gamma^0\gamma^3\right\} = \exp\left\{\frac{\zeta}{2}\gamma^0\gamma^3\right\}. \tag{2.243}$$

The matrix $S(\boldsymbol{L}_z)$ is non-unitary. This is closely connected to a fact that there are no finite dimensional unitary representations for non-compact groups of which boosts are an example.

Let us try to compute the explicit form of $S(\boldsymbol{L}_z)$. Firstly, we can calculate $(\gamma^0\gamma^3)^2 = \gamma^0\gamma^3\gamma^0\gamma^3 = -\gamma^0\gamma^0\gamma^3\gamma^3 = -\mathbf{1}(-\mathbf{1}) = \mathbf{1}$, which means that $\gamma^0\gamma^3$ is an involutive matrix. We will now use the important identity

$$e^{x\boldsymbol{N}} = (\cosh x)\,\mathbf{1} + (\sinh x)\,\boldsymbol{N} \tag{2.244}$$

which is valid for the matrix $\boldsymbol{N}$ which is involutive ($\boldsymbol{N}^2 = \mathbf{1}$). Hence

$$S(\boldsymbol{L_z}) = \cosh\frac{\zeta}{2} + \gamma^0\gamma^3\sinh\frac{\zeta}{2}. \tag{2.245}$$

Note that in spinor representation $S(\boldsymbol{L})$, the arguments of hyperbolic sine and cosine are $\zeta/2$ whereas in fundamental representation $\boldsymbol{L}$, the arguments are just ζ.

Secondly, using geometric half-angle formulae, one can write

$$\cosh\frac{\zeta}{2} = \sqrt{\frac{\cosh\zeta + 1}{2}} = \sqrt{\frac{E/m + 1}{2}} = \sqrt{\frac{E+m}{2m}} \tag{2.246}$$

and

$$\sinh\frac{\zeta}{2} = \sqrt{\frac{\cosh\zeta - 1}{2}} = \sqrt{\frac{E-m}{2m}} = \sqrt{\frac{E^2-m^2}{2m(E+m)}}$$
$$= \sqrt{\frac{q^2}{2m(E+m)}}, \tag{2.247}$$

since $\cosh\zeta = \gamma$ and $E = \gamma m$ we have that $\cosh\zeta = E/m$. Here m is rest mass and q relativistic three-momentum. From these results

$$S(\boldsymbol{L}_z) = \sqrt{\frac{E+m}{2m}}\left[\mathbf{1} + \frac{q}{E+m}\gamma^0\gamma^3\right]. \tag{2.248}$$

Thirdly, we can compute

$$\gamma^0\gamma^3 = (\sigma^3\otimes\mathbf{1})(i\sigma^2\otimes\sigma^3) = (\sigma^3 i\sigma^2\otimes\sigma^3) = \begin{pmatrix}\mathbf{0} & \sigma^3\\ \sigma^3 & \mathbf{0}\end{pmatrix}. \tag{2.249}$$

Thus we can finally write the explicit form of boost in z direction in spinor representation as

$$S(\boldsymbol{L}_z) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix}\mathbf{1} & \frac{q}{E+m}\sigma^3\\ \frac{q}{E+m}\sigma^3 & \mathbf{1}\end{pmatrix}. \tag{2.250}$$

This might be generalized for a boost in general direction with velocity $\boldsymbol{v}$ for which the matrix takes the form

Now the $\boldsymbol{q}$ is a three-vector. In the previous calculations q was just its z component (the direction in which we boosted the system).

$$S(\boldsymbol{L}) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix}\mathbf{1} & \frac{\boldsymbol{q}\boldsymbol{\sigma}}{E+m}\\ \frac{\boldsymbol{q}\boldsymbol{\sigma}}{E+m} & \mathbf{1}\end{pmatrix}. \tag{2.251}$$

As an example, consider the rest frame of a particle where its four-momentum is again $\overline{p}^\mu = \begin{pmatrix}m\\ \mathbf{0}\end{pmatrix}$ (now the line above denotes the rest frame), so the Dirac spinor for the particle is (in form of plane waves)

$$u_\lambda(\overline{p}) = \sqrt{2m}\begin{pmatrix}\chi_\lambda\\ 0\end{pmatrix}. \tag{2.252}$$

When boosting system by a velocity v in z direction we get

$$p = \boldsymbol{L}_z\overline{p} = \begin{pmatrix}E\\ 0\\ 0\\ q\end{pmatrix}, \tag{2.253}$$

then we claim that

$$u_\lambda(p) = S(\boldsymbol{L}_z)u_\lambda(\overline{p}) = \sqrt{\frac{E+m}{2m}}\begin{pmatrix}\mathbf{1} & \frac{q\sigma^3}{E+m}\\ \frac{q\sigma^3}{E+m} & \mathbf{1}\end{pmatrix}\sqrt{2m}\begin{pmatrix}\chi_\lambda\\ 0\end{pmatrix}, \tag{2.254}$$

so we have

$$u_\lambda(P) = \sqrt{E+m}\begin{pmatrix}\chi_\lambda\\ \frac{q\sigma^3}{E+m}\chi_\lambda\end{pmatrix}, \tag{2.255}$$

as we know it should be.

2.9 Spin Sums and Projection Operators

First of all, we remind that negative energy solution were defined as

$$v_\lambda(p)e^{ipx} = v_\lambda e^{i\omega_p t - i\boldsymbol{p}\cdot\boldsymbol{x}}, \tag{2.256}$$

where amplitude $v(p)$ satisfies

$$(\gamma^\mu p_\mu + m)v(p) = 0 \quad \Rightarrow \quad v(p) = \sqrt{E+m}\begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\lambda \\ \chi_\lambda \end{pmatrix}. \tag{2.257}$$

Index λ denotes different spin projection ("degrees of freedom"). We have $\lambda = \pm\frac{1}{2}$ with $\chi_{1/2} = \begin{pmatrix}1\\0\end{pmatrix}$ and $\chi_{-1/2} = \begin{pmatrix}0\\1\end{pmatrix}$ being eigenvectors of $S_3 = \frac{1}{2}\sigma_3$ because

$$\frac{1}{2}\sigma_3\chi_\lambda = \frac{\lambda}{2}\chi_\lambda. \tag{2.258}$$

Depending on literature, one can also finds notation $\lambda = \pm 1$ or $\lambda = \{1, 2\}$.

We have the following "ortho-normality" relations

$$\overline{u}_\lambda(p)u_{\lambda'}(p) = 2m\delta_{\lambda\lambda'}, \tag{2.259}$$

$$\overline{v}_\lambda(p)v_{\lambda'}(p) = -2m\delta_{\lambda\lambda'}, \tag{2.260}$$

$$\overline{u}_\lambda(p)v_{\lambda'}(p) = 0, \tag{2.261}$$

$$\overline{v}_\lambda(p)u_{\lambda'}(p) = 0, \tag{2.262}$$

where $\overline{u}_\lambda(p) = u^\dagger_\lambda(p)\gamma^0$, $\overline{v}_\lambda(p) = v^\dagger_\lambda(p)\gamma^0$.

We already know that $\overline{\psi}(x)\psi(x)$ is a Lorentz scalar, and so is $\overline{u_\lambda}(p)u_\lambda(p)$, $\overline{v_\lambda}(p)v_\lambda(p)$. Indeed,

$$\begin{aligned}\overline{u_\lambda}(p)v_\lambda(p) &= \overline{u_\lambda}(\boldsymbol{L}^{-1}p)S^{-1}(\boldsymbol{L})S(\boldsymbol{L})v_\lambda(\boldsymbol{L}^{-1}p) \\ &= \overline{u_\lambda}(\boldsymbol{L}^{-1}p)v_\lambda(\boldsymbol{L}^{-1}p).\end{aligned} \tag{2.263}$$

Hence to find the result we can go to rest frames where

$$\begin{aligned} u_{1/2}(m,\boldsymbol{0}) &= \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \; u_{-1/2}(m,\boldsymbol{0}) = \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \\ v_{1/2}(m,\boldsymbol{0}) &= \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \; v_{-1/2}(m,\boldsymbol{0}) = \sqrt{2m}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \end{aligned} \tag{2.264}$$

These relations imply that 4-spinors are linearly independent. They form a complete basis for any Dirac bi-spinor.

Consider the operators

$$\tilde{\Lambda}^+(p) = \sum_\lambda u_\lambda(p)\overline{u_\lambda}(p) \tag{2.265}$$

$$\tilde{\Lambda}^-(p) = \sum_\lambda v_\lambda(p)\overline{v_\lambda}(p). \tag{2.266}$$

We will show some properties of these operators now.

$$\tilde{\Lambda}^+(p)u_\lambda(p) = \sum_{\lambda'} u_{\lambda'}(p)\overline{u_{\lambda'}}(p)u_\lambda(p) = 2m\sum_{\lambda'} u_{\lambda'}(p)\delta_{\lambda\lambda'} = 2mu_\lambda(p).$$
$$\tilde{\Lambda}^+(p)v_\lambda(p) = \sum_{\lambda'} u_{\lambda'}(p)\underbrace{\overline{u_{\lambda'}}(p)v_\lambda(p)}_{0} = 0. \tag{2.267}$$

By comparing this with action of the operators $(\gamma^\mu p_\mu \pm m)$ we find out that following identities hold

$$\tilde{\Lambda}^+(p) = \sum_{\lambda'} u_{\lambda'}(p)\overline{u_{\lambda'}}(p) = \gamma^\mu p_\mu + m, \tag{2.268}$$
$$\tilde{\Lambda}^-(p) = \sum_{\lambda} v_\lambda(p)\overline{v_\lambda}(p) = \gamma^\mu p_\mu - m. \tag{2.269}$$

Let us now define the operators

$$\begin{aligned} \frac{\gamma^\mu p_\mu + m}{2m} &= \Lambda^+(p), \\ \frac{-(\gamma^\mu p_\mu - m)}{2m} = -\frac{\tilde{\Lambda}^-(p)}{2m} &= \Lambda^-(p). \end{aligned} \tag{2.270}$$

These operators are projection operators as they fulfill the necessary conditions for projection operators, in particular:

$$\begin{aligned} \left[\Lambda^\pm(p)\right]^2 &= \left[\frac{\pm(\gamma^\mu p_\mu \pm m)}{2m}\right]^2 = \frac{(\gamma^\mu p_\mu)^2 \pm 2\gamma^\mu p_\mu m + m^2}{4m^2} \\ &= \frac{\pm(\gamma^\mu p_\mu \pm m)}{2m} = \Lambda^\pm(p)\,, \end{aligned} \tag{2.271}$$

$$\Lambda^+(p) + \Lambda^-(p) = \mathbf{1}, \tag{2.272}$$

$$\Lambda^\pm\Lambda^\mp = \mathbf{0} \tag{2.273}$$

Note also that

$$\mathrm{Tr}\left(\Lambda^\pm(p)\right) = \frac{\pm\mathrm{Tr}(\gamma^\mu p_\mu)}{2m} + 2$$
$$= \frac{\pm p_\mu \mathrm{Tr}\,\gamma^\mu}{2m} + 2 = 2,$$

where we used the fact that

$$\mathrm{Tr}\,\gamma^\mu = 0.$$

Above projection operators project over positive and negative energy states.

2.10 Electromagnetic Coupling of Electrons

Non-relativistic Charged Particle

For a free particle of a mass m, the Hamiltonian is given as $H_0 = \boldsymbol{p}^2/2m$. If the particle has a charge q, then in the presence of an electromagnetic field the Hamiltonian becomes

$$H = \frac{\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right)^2}{2m} + q\phi, \tag{2.274}$$

Using *minimal substitution* (*minimal coupling*).

where $\boldsymbol{A}$ (ϕ) is the vector (scalar) potential.

Moreover, we also have to insist on a gauge restriction $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$ (so called Coulomb gauge). The magnetic field is connected with the vector potential via $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$, so a situation of zero electric field and constant magnetic field can be expressed for example as $\phi = 0$, $\boldsymbol{A} = \frac{1}{2}B(-y, x, 0)$ and $\boldsymbol{B} = (0, 0, B)$.

Since $\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}$, $\boldsymbol{\nabla} \times E = -\frac{\partial \boldsymbol{B}}{\partial t}$ thanks to the fact that $\boldsymbol{\nabla} \times (\boldsymbol{\nabla}\phi) = 0$.

To the first order in q, the Hamiltonian becomes

$$H = \frac{1}{2m}\left[\boldsymbol{p}^2 - \frac{q}{c}(\boldsymbol{pA} + \boldsymbol{Ap})\right] + O(q^2). \tag{2.275}$$

In quantum mechanics, $\boldsymbol{pA} \neq \boldsymbol{Ap}$, but we can relate these two expressions via $p_i A_i = A_i p_i + [p_i, A_i]$, where $[p_i, A_i] = -i\hbar\,[\nabla_i, A_i] = -i\hbar\boldsymbol{\nabla}\cdot\boldsymbol{A} = 0$ (using Coulomb gauge condition). Then

$$H = \frac{\boldsymbol{p}^2}{2m} - \frac{q}{mc}\boldsymbol{Ap} + O(q^2). \tag{2.276}$$

Generally, the commutator of a function f with the derivation $\left[\frac{d}{dx}, f\right]$ is equal to $\frac{df}{dx}$ since (operating on a function u)

$$\left[\frac{d}{dx}, f\right] u = \left(\frac{d}{dx}(fu)\right) - f\frac{du}{dx} =$$
$$= \left(\frac{df}{dx}\right) u + f\frac{du}{dx} - f\frac{du}{dx} = \left(\frac{df}{dx}\right) u.$$

Evaluating explicitly

$$\boldsymbol{A}\cdot\boldsymbol{p} = \frac{1}{2}B(-y, x, 0)(p_x, p_y, p_z) = \frac{B}{2}(xp_y - yp_x) = \frac{B}{2}L_z. \tag{2.277}$$

Thus we can write the Hamiltonian as composed of two parts $H = H_0 + H_{\text{EM}}$, where

$$H_{\text{EM}} = -\frac{q}{2mc}\boldsymbol{B}\cdot\boldsymbol{L}. \tag{2.278}$$

It is conventional to write this as

$$H_{\text{EM}} = -\frac{q\hbar}{2mc}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} = -g\mu_{\text{B}}\frac{\boldsymbol{L}}{\hbar}\cdot\boldsymbol{B} = -\boldsymbol{\mu}\cdot\boldsymbol{B}, \tag{2.279}$$

where $\mu_{\text{B}} = e\hbar/(2m_e c)$ is the Bohr magneton (m_e denotes the electron mass). We call g the g-factor (here $g = 1$) and μ the orbital magnetic moment of an electron. The term $\gamma = g\mu_{\text{B}}$, which relates magnetic moment and angular momentum of a particle, is called the gyromagnetic ratio.

Atomic physics tells us that (i) electrons have spin and (ii) the gyromagnetic ratio for spin is twice the gyromagnetic ratio for orbital angular momentum (i.e. $g = 2$). As we will see later in this section, the great achievement of Dirac was to show that his equation predicted these results.

Minimal Electromagnetic Coupling

In the classical (non-relativistic) physics, we "substitute" a momentum $\boldsymbol{p}$ with $\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}$ (if $\hbar = c = 1$, $-i\boldsymbol{\nabla} \to -i\boldsymbol{\nabla} - q\boldsymbol{A}$, i.e. $\boldsymbol{\nabla} \to \boldsymbol{\nabla} - iq\boldsymbol{A}$).

When considering a relativistic situation, we have to work with the four-potential A_μ. The minimal coupling prescription is then given by

$$p_\mu \to p_\mu - qA_\mu, \quad \partial_\mu \to \partial_\mu + iqA_\mu, \tag{2.280}$$

where $p^\mu = i\frac{\partial}{\partial x_\mu}$ or $p_\mu = i\frac{\partial}{\partial x^\mu}$, respectively.

We can impose the gauge condition $\partial_\mu A^\mu = 0$ (Lorentz gauge), thanks to which $\left[\partial_\mu, A^\mu\right] = \partial_\mu A^\mu = 0$.

The Dirac equation for a particle of a charge q becomes $i\gamma\partial \to i\gamma(\partial + iqA)$, therefore

$$[i\gamma(\partial + iqA) - m]\,\psi(x) = 0 \quad \Leftrightarrow \quad [\gamma(p - qA) - m]\,\psi(x) = 0. \tag{2.281}$$

This equation is still Lorentz covariant since both vector expressions transform in a same way, i.e.

$$\partial'_\mu = L_\mu^{\ \nu}\partial_\nu \quad A'_\mu(x') = L_\mu^{\ \nu}A_\nu(x). \tag{2.282}$$

If we multiply (2.281) by $[\gamma(p - qA) + m]\,\psi(x) = 0$, we get

$$\begin{aligned} &\left\{[\gamma(p - qA)]^2 - m^2\right\}\psi(x) \\ = &\left[\gamma^\mu\gamma^\nu(p_\mu - qA_\mu)(p_\nu - qA_\nu) - m^2\right]\psi(x) = 0. \end{aligned} \tag{2.283}$$

We can rewrite the product as

$$\begin{aligned} &\gamma^\mu\gamma^\nu(p_\mu - qA_\mu)(p_\nu - qA_\nu) \\ = &\left(\frac{1}{2}\gamma^\mu\gamma^\nu - \frac{1}{2}\gamma^\nu\gamma^\mu + \frac{1}{2}\gamma^\mu\gamma^\nu + \frac{1}{2}\gamma^\nu\gamma^\mu\right)\left(p_\mu p_\nu + q^2 A_\mu A_\nu - qp_\mu A_\nu - qA_\mu p_\nu\right) \\ = &\left(\frac{1}{2}\{\gamma^\mu,\gamma^\nu\} + \frac{1}{2}[\gamma^\mu,\gamma^\nu]\right)\Big[p_\mu p_\nu + q^2 A_\mu A_\nu - \frac{1}{2}(qp_\mu A_\nu + qA_\mu p_\nu) - \\ &-\frac{1}{2}(qp_\nu A_\mu + qA_\nu p_\mu) - \frac{1}{2}(qp_\mu A_\nu - qA_\mu p_\nu) + \frac{1}{2}(qp_\nu A_\mu - qA_\nu p_\mu)\Big], \end{aligned} \tag{2.284}$$

where we have split each term into its symmetric and anti-symmetric parts. Now, since for a symmetric matrix $\boldsymbol{A} = (a_{ij})$ and an anti-symmetric matrix $\boldsymbol{B} = (b_{kl})$, the relation $a_{ij}b_{ij} = 0$ holds, and using the defining relation for the Dirac algebra (2.68), the product further equals to

$$\begin{aligned} &g^{\mu\nu}\left(\frac{1}{2}(p_\mu - qA_\mu)(p_\nu - qA_\nu) + \frac{1}{2}(p_\nu - aA_\nu)(p_\mu - qA_\mu)\right) + \\ &q\frac{1}{2}[\gamma^\mu,\gamma^\nu]\left(-\frac{1}{2}\left[p_\mu, A_\nu\right] + \frac{1}{2}\left[p_\nu, A_\mu\right]\right) = \\ &(p - qA)^2 - \frac{iq}{4}[\gamma^\mu,\gamma^\nu]\left(\left[\partial_\mu, A_\nu\right] - \left[\partial_\nu, A_\mu\right]\right) = \\ &(i\partial - qA)^2 - \frac{iq}{4}[\gamma^\mu,\gamma^\nu](\partial_\mu A_\nu - \partial_\nu A_\mu). \end{aligned}$$

If we recall the definition of the electromagnetic tensor $F_{\mu\nu}$, we can finally write Eq. (2.283) as

$$\left\{(i\partial - qA)^2 - \frac{iq}{4}\left[\gamma^\mu, \gamma^\nu\right] F_{\mu\nu} - m^2\right\} \psi(x)$$
$$= \left\{(i\partial - qA)^2 - m^2 - \frac{q}{2}\sigma^{\mu\nu} F_{\mu\nu}\right\} \psi(x) = 0, \qquad (2.285)$$

where we have used the expression for the commutator of two gamma matrices, i.e. $\frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right] = \sigma^{\mu\nu}$.

The first term is what would be the energy from minimal coupling to the Klein-Gordon equation. The second term reflects the spin of the electron (which is not present in Klein-Gordon equation). The term $\sigma^{\mu\nu} F_{\mu\nu}$ is a spin coupling to the electromagnetic field.

If we take $F^{oi} = E^i$ and $F^{kl} = \varepsilon^{klm} B_m$ (where klm is an even permutation of 1,2,3), we get

$$\left\{\left[(p - qA)^2 - m^2\right] - q\boldsymbol{\Sigma B} + iq\boldsymbol{\alpha} \cdot \boldsymbol{E}\right\} \psi(x) = 0, \qquad (2.286)$$

where $\boldsymbol{\alpha} = \gamma^0 \boldsymbol{\gamma}$.

As a non-relativistic limit, consider now slow electrons passing through weak, slowly varying electromagnetic fields. Assume no high momenta in the Dirac equation, then

$$(p - qA)^2 = (p^0 - qA^0)^2 - (\boldsymbol{p} - q\boldsymbol{A})^2 = \left\{p^0 = H + m\right\}$$
$$= (H - qA^0 + m)^2 - (\boldsymbol{p} - q\boldsymbol{A})^2 \approx m^2 + 2m(H - qA^0) - (\boldsymbol{p} - q\boldsymbol{A})^2,$$

where the non-relativistic Hamiltonian H is effectively small in comparison to m, and so we neglected the second order term H^2.

Then for $\phi = A^0$ we insert the previous relation to our equation (2.286) and get

$$\left\{2m(H - q\phi) - (\boldsymbol{p} - q\boldsymbol{A})^2 - q\boldsymbol{\Sigma B} + iq\boldsymbol{\alpha} \cdot \boldsymbol{E}\right\} \psi(x) = 0,$$

which implies

$$H\psi = \left[\frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} + q\phi + \frac{q}{2m}\boldsymbol{\Sigma B} - i\boldsymbol{\alpha E}\right] \psi. \qquad (2.287)$$

This is effectively the non-relativistic Hamiltonian which emerges from the Dirac equation. In the absence of an electric field and in the presence of a static magnetic field, the Hamiltonian takes the form

$$H = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} + \frac{q}{2m}\boldsymbol{\Sigma B}. \qquad (2.288)$$

By realizing that the large component of the Dirac bispinor dominates in the non-relativistic limit,

$$H\big|_{\text{large comp.}} = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} + \frac{q}{2m}\boldsymbol{\sigma B}, \qquad (2.289)$$

where we recall that the spin operator is equal to $\boldsymbol{S} = \frac{1}{2}\boldsymbol{\sigma}$.

Recall the origin of the non-relativistic calculation, where we assumed $H = H_0 + H_{\text{EM}}$. Now we have

$$H_{\text{EM}} = -\frac{q}{2m}\boldsymbol{B}\boldsymbol{L} + 2\frac{q}{2m}\boldsymbol{S}\boldsymbol{B}, \tag{2.290}$$

where the two factors in front of the scalar products are the gyromagnetic ratios for the orbital angular momentum and for the spin-angular momentum, respectively. Neglecting sign, the orbital gyromagnetic ratio for the electron is equal to $e/2m$ (the Bohr magneton), whereas the spin gyromagnetic ratio equal to e/m (twice the Bohr magneton)!

Exercise: Consider a relativistic particle of a charge q moving in a central potential for which the Dirac Hamiltonian is given by

$$\frac{\hat{H}}{c} = \hat{p}_0 = \boldsymbol{\alpha}\boldsymbol{p} + \beta + \frac{V(r)}{c}, \tag{2.291}$$

where $\boldsymbol{\alpha} = \gamma^0\boldsymbol{\gamma}$ and $\beta = mc\gamma_0$ and $V(r) = q\phi(r)$ for a particle in an electrostatic field. Show that $\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}} \times \hat{\boldsymbol{p}}$ is not a constant of the motion, but $\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{S}}$ is, where $S_i = \hat{\sigma}_i = \frac{\hbar}{4}\varepsilon_{ijk}\sigma_{jk} = \frac{\hbar}{2}\begin{pmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{pmatrix}$.

Various Important notes related to Dirac Equation

In this section, another approach to the description of coupling of a relativistic electron to an electromagnetic field will be given. We will start directly from the Dirac equation.

The minimal coupling prescription tells us to replace

$$\hat{p}_\mu \to \hat{p}_\mu - qA_\mu = \hat{\Pi}_\mu \quad \Longleftarrow \quad \partial_\mu \to \partial_\mu + iqA_\mu, \tag{2.292}$$

where $\hat{p}_\mu$ is referred to as the canonical momentum and $\hat{\Pi}_\mu$ as the kinetic momentum. The Dirac equation then takes the form

$$(i\not{\partial} - q\not{A} - m)\psi(x) = 0, \tag{2.293}$$

where $q = -|e|$ is the electron charge.

A more general coupling leads to dipole and higher multi-pole terms. If we would add e.g. term $\sigma^{\mu\nu}F_{\mu\nu}$, it would provide non-minimal coupling.

This can be written in the Schrödinger form

$$\begin{aligned} i\frac{\partial\psi}{\partial t} &= [\boldsymbol{\alpha}(-i\boldsymbol{\nabla} - q\boldsymbol{A}) + \beta m + q\phi]\,\psi \\ &= [\boldsymbol{\alpha}\boldsymbol{p} + \beta m]\,\psi + [-q\boldsymbol{\alpha}\boldsymbol{A} + q\phi]\,\psi \\ &= (H_0 + H_{\text{int}})\,\psi\,. \end{aligned} \tag{2.294}$$

To extract more physics, let us concentrate at the non-relativistic limit. We write $\psi = \begin{pmatrix}\varphi\\ \chi\end{pmatrix}$ and use the Dirac's representation. Then from

$$i\frac{\partial\psi}{\partial t} = \left[\begin{pmatrix}0 & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & 0\end{pmatrix}(-i\boldsymbol{\nabla} - q\boldsymbol{A}) + \begin{pmatrix}q\phi + m & 0\\ 0 & q\phi - m\end{pmatrix}\right]\psi, \tag{2.295}$$

we get two coupled equations

$$i\frac{\partial\varphi}{\partial t} = \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\chi + q\phi\varphi + m\varphi, \tag{2.296}$$

$$i\frac{\partial\chi}{\partial t} = \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\varphi + q\phi\chi - m\chi. \tag{2.297}$$

Similarly as for the free particle we pass to the limit by factoring from

ψ the fast oscillating factor, i.e.

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = e^{-imt} \begin{pmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{pmatrix}, \tag{2.298}$$

where we expect that $\tilde{\varphi}$ and $\tilde{\chi}$ are changing slowly in time. So the equation (2.297) reduces to

$$i\frac{\partial \tilde{\varphi}}{\partial t} = \boldsymbol{\sigma} \cdot \boldsymbol{\Pi} \tilde{\chi} + q\phi\tilde{\varphi}, \tag{2.299}$$

$$i\frac{\partial \tilde{\chi}}{\partial t} = \boldsymbol{\sigma} \cdot \boldsymbol{\Pi} \tilde{\varphi} + q\phi\tilde{\chi} - 2m\tilde{\chi}, \tag{2.300}$$

where in the second equation the $2m\tilde{\chi}$ term dominates over $i\partial\tilde{\chi}/\partial t$. So, the second equation may be approximated for kinetic energies

$$\tilde{\chi} = \frac{\boldsymbol{\sigma\Pi}}{2m}\tilde{\varphi} + \frac{q\phi}{2m}\tilde{\chi} \sim \frac{\boldsymbol{\sigma\Pi}}{2m}\tilde{\varphi}, \tag{2.301}$$

where the latter approximation holds thanks to the fact that the interaction energy $q\phi$ is much smaller than the rest-mass energy mc^2, so $q\phi/2mc^2 \ll 1$.

Inserting this into the first equation, we get

$$i\frac{\partial \tilde{\varphi}}{\partial t} = \left(\frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\Pi})(\boldsymbol{\sigma} \cdot \boldsymbol{\Pi})}{2m}\right)\tilde{\varphi} + q\phi\tilde{\varphi} = \left(\frac{(\boldsymbol{\sigma} \cdot \boldsymbol{\Pi})(\boldsymbol{\sigma} \cdot \boldsymbol{\Pi})}{2m} + q\phi\right)\tilde{\varphi}. \tag{2.302}$$

This can further be reduced by using an analogue of the well known identity identity

$$\begin{aligned}
(\boldsymbol{\sigma} \cdot \boldsymbol{a})(\boldsymbol{\sigma} \cdot \boldsymbol{b}) &= \sigma^i \sigma^j a_i b_j = (\delta^{ij} + i\varepsilon^{ijk}\sigma^k) a_i b_j \\
&= \boldsymbol{a} \cdot \boldsymbol{b} + i\sigma_k (\boldsymbol{a} \times \boldsymbol{b})_k \\
&= \boldsymbol{a} \cdot \boldsymbol{b} + i\boldsymbol{\sigma} \cdot (\boldsymbol{a} \times \boldsymbol{b}),
\end{aligned} \tag{2.303}$$

which is true only if $\boldsymbol{a}$ and $\boldsymbol{b}$ are c-numbered vectors. For operators, this identity must be modified. In particular, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are generic vector operators we should write

$$(\boldsymbol{\sigma} \cdot \boldsymbol{a})(\boldsymbol{\sigma} \cdot \boldsymbol{b}) = \frac{1}{4}\left\{\sigma^i, \sigma^j\right\}\left\{a_i, b_j\right\} + \frac{1}{4}\left[\sigma^i, \sigma^j\right]\left[a_i, b_j\right], \tag{2.304}$$

Notice that $(\nabla \times A)_k = \varepsilon_{klm}\partial_l A_m$, i.e. $2(\nabla \times A)_k = \varepsilon_{klm}(\partial_l A_m - \partial_m A_l)$.

where we used the decomposition into a symmetric and an anti-symmetric parts. Specifically for the scalar products of $\boldsymbol{\sigma}$ with the kinetic momentum, we get

$$\begin{aligned}
(\boldsymbol{\sigma} \cdot \boldsymbol{\Pi})(\boldsymbol{\sigma} \cdot \boldsymbol{\Pi}) &= \frac{1}{4}2\delta^{ij}\left\{\Pi_i, \Pi_j\right\} + \frac{1}{4}2i\varepsilon_{ijk}\sigma_k\left[\Pi_i, \Pi_j\right] \\
&= \boldsymbol{\Pi}^2 + \frac{1}{2}i\varepsilon_{ijk}\sigma_k\left[p_i - qA_i, p_j - qA_j\right] \\
&= \boldsymbol{\Pi}^2 - \frac{1}{2}i\varepsilon_{ijk}q\left(\left[p_i, A_j\right] + \left[A_i, p_j\right]\right) \\
&= \boldsymbol{\Pi}^2 - \frac{1}{2}q\sigma_k\varepsilon_{ijk}\left(\nabla_i A_j - \nabla_j A_i\right) \\
&= \boldsymbol{\Pi}^2 - q\sigma_k(\nabla \times A)_k = \boldsymbol{\Pi}^2 - q\boldsymbol{\sigma} \cdot \boldsymbol{B}.
\end{aligned} \tag{2.305}$$

Thus, the first equation finally becomes

$$i\frac{\partial\tilde{\varphi}}{\partial t} = \left[\frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} - \frac{q\boldsymbol{\sigma}\cdot\boldsymbol{B}}{2m} + q\phi\right]\tilde{\varphi}\,, \tag{2.306}$$

Note that the only spin dependence is through the interaction with a magnetic field $\boldsymbol{\sigma}\cdot\boldsymbol{B}$.

which is nothing but the Pauli equation of the non-relativistic quantum physics. Hence we gain certain confidence that we are on the right track.

Restoring the factors $\hbar$ and c,

$$H_{\text{MAG}} = -\frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B} = -\boldsymbol{\mu}\cdot\boldsymbol{B}\,, \tag{2.307}$$

where $\boldsymbol{\mu}$ is the magnetic moment. We remind here that $\boldsymbol{\mu} = g\mu_{\text{B}}\frac{\hat{\boldsymbol{\sigma}}}{\hbar}$, where μ_{B} is the Bohr magneton, g is the g-factor and $g\mu_{\text{B}}$ is the gyromagnetic ratio. Since the latter equals $2\mu_{\text{B}}$, $g = 2$ must hold.

The current result is equal to $g/2 = 1,0011596521....$

The fact that $g = 2$ is a nontrivial prediction of Dirac theory derived within the non-relativistic context of the Pauli equation.

Exercise: Determine the energy levels in a uniform magnetic field (relativistic generalization of Landau levels).

The g-factor has now been measured to something like 12 figures of accuracy and it is not precisely 2, it differs by a tiny amount. The understanding of the difference goes beyond simple Dirac theory into quantum field theory.

2.11 Representations of Gamma Matrices

We have the following Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2\boldsymbol{g}^{\mu\nu}. \tag{2.308}$$

Further on, we will explore different representations of gamma matrices. First, we will note that there exists a **fundamental theorem of Clifford algebra**:

Theorem 2.11.1 *If two distinct sets of γ-matrices are given, that both satisfy the Clifford algebra relation*

$$\{\gamma^\mu, \gamma^\nu\} = 2\boldsymbol{g}^{\mu\nu}, \tag{2.309}$$

then they are connected to each other by similarity transformation

$$\gamma'^\mu = \boldsymbol{S}^{-1}\gamma^\mu\boldsymbol{S}. \tag{2.310}$$

If, in addition, the matrices are (anti)unitary (as in our particular case $\gamma^0 = \gamma^{0\dagger}$, $\gamma^0\gamma^i = \gamma^{i\dagger}\gamma^0$, which implies $\gamma^i = -\gamma^{i\dagger}$) then $\boldsymbol{S}$ itself is unitary, i.e.

$$\gamma'^\mu = \boldsymbol{U}^\dagger\gamma^\mu\boldsymbol{U}. \tag{2.311}$$

This transformation is unique up to a multiplicative factor of absolute value 1.

Clearly when a set of γ^μ matrices satisfies Clifford algebra relation and are (anti)hermitan, then also $\gamma'^\mu = \boldsymbol{U}^\dagger \gamma^\mu \boldsymbol{U}$ also satisfies those relation and are (anti)hermitian and vice versa. Problematic part is to show existence and uniqueness of $\boldsymbol{U}$.

Let us now review most typical representations of γ-matrices. Before, let us recall what Pauli matrices look like

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.312}$$

Dirac's Representation

First representation of γ-matrices in known as Dirac's representation and it is given by

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{-1} \end{pmatrix},\ \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix},\ \gamma^5 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}. \tag{2.313}$$

By using properties of tensor product, we can conveniently rewrite those matrices as

$$\gamma^0 = \sigma_3 \otimes \mathbf{1},\ \boldsymbol{\gamma} = i\sigma_2 \otimes \boldsymbol{\sigma},\ \gamma^5 = \sigma_1 \otimes \mathbf{1}. \tag{2.314}$$

This representation is particularly convenient for taking the non-relativistic limit, e.g. for a free particle we have already seen that in this representation

$$\begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi \end{pmatrix} \overset{\text{NR}}{\rightarrow} \begin{pmatrix} \chi_{NR} \\ 0 \end{pmatrix}. \tag{2.315}$$

The corresponding Lorentz group generators, as we know, have the explicit forms

$$S(\boldsymbol{R}) = \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{n}_k\boldsymbol{\sigma}_k} & 0 \\ 0 & e^{-\frac{i}{2}\boldsymbol{n}_k\boldsymbol{\sigma}_k} \end{pmatrix}, \tag{2.316}$$

$$S(\boldsymbol{B}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \mathbf{1} & \frac{\boldsymbol{p}\cdot\boldsymbol{\sigma}}{E+m} \\ \frac{\boldsymbol{p}\cdot\boldsymbol{\sigma}}{E+m} & \mathbf{1} \end{pmatrix}. \tag{2.317}$$

Chiral (Weyl) Representation

Another well known representation is chiral representation given by

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix},\ \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix},\ \gamma^5 = \begin{pmatrix} \mathbf{-1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}. \tag{2.318}$$

And again, we can rewrite this representation in terms of tensor products as

$$\gamma^0 = \sigma_1 \otimes \mathbf{1},\ \boldsymbol{\gamma} = i\sigma_2 \otimes \boldsymbol{\sigma},\ \gamma^5 = \sigma_3 \otimes \mathbf{1}. \tag{2.319}$$

This representation is important for description of massless spin 1/2 particles. Even more, it is important for the discussions about Lorentz group. Recall that

$$S(\boldsymbol{R}) = e^{-\frac{i}{2}\omega_{ij}\boldsymbol{\Sigma}^{ij}} = e^{-\frac{i}{4}\omega_{ij}\boldsymbol{\sigma}^{ij}}, \tag{2.320}$$

with $\boldsymbol{\sigma}^{ij} = \frac{1}{2}[\gamma^i, \gamma^j]$ and $\boldsymbol{\Sigma}^{ij} = \frac{1}{2}\varepsilon_{ijk}\begin{pmatrix}\sigma_k & 0\\ 0 & \sigma_k\end{pmatrix}$. By defining

$$\boldsymbol{\Sigma}^i = \frac{1}{2}\varepsilon_{ijk}\boldsymbol{\Sigma}^{jk} \equiv \boldsymbol{L}^i \tag{2.321}$$

we get

$$S(\boldsymbol{R}) = e^{-i\boldsymbol{\theta}\cdot\boldsymbol{\Sigma}} = \begin{pmatrix} e^{-\frac{i}{2}\boldsymbol{\theta}_k\boldsymbol{\sigma}_k} & 0\\ 0 & e^{\frac{i}{2}\boldsymbol{\theta}_k\boldsymbol{\sigma}_k}\end{pmatrix} = e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}}. \tag{2.322}$$

It has the same structure as in the Dirac's representation, because γ^i, $i = 1, 2, 3$ remain the same.

General boost is given by

$$S(\boldsymbol{B}) = e^{-\frac{i}{4}\omega_{0i}\boldsymbol{\sigma}^{0i}}. \tag{2.323}$$

For instance, for a boost in 3rd direction

Recall that $(\pm\sigma_1 \otimes \mathbf{1})(i\sigma_2 \otimes \sigma_3) = (\pm i\sigma_1\sigma_2 \otimes \sigma_3) = \mp\sigma_3 \otimes \sigma_3$.

$$S(\boldsymbol{B}) = e^{\frac{1}{2}\gamma^0\gamma^3} = \begin{pmatrix} e^{\mp\frac{1}{2}\sigma_3} & 0\\ 0 & e^{\pm\frac{1}{2}\sigma_3}\end{pmatrix}. \tag{2.324}$$

It seems we do not need to work with bispinors, but it would be enough to work with spinors only. This, indeed, is true for massless particles. For massive particles the issue is more complicated and is related to parity and time reversal. Next, we will focus on discrete elements of Lorentz group.

Space Reflection (Parity) Transformation

Parity transformation is acting on 4-vector as follows:

$$x^\mu = (t, \boldsymbol{x}) \xrightarrow{P} x^\mu_p = (t, -\boldsymbol{x}). \tag{2.325}$$

From this we can see, that this transformation has matrix form of

$$\boldsymbol{L}^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix} = \boldsymbol{g}^{\mu\nu} = \boldsymbol{g}_{\mu\nu}. \tag{2.326}$$

This satisfies the defining property of Lorentz group

$$\boldsymbol{L}_\alpha{}^\mu \boldsymbol{g}_{\mu\nu}\boldsymbol{L}^\nu{}_\beta = \boldsymbol{g}_{\alpha\beta}, \tag{2.327}$$

and hence it is an element of Lorentz group. We can also see, that $\det \boldsymbol{L} = -1$. With respect to the Dirac equation covariance, we require that

$$S(\boldsymbol{L}^P)\gamma^\mu S^{-1}(\boldsymbol{L}^P) = \boldsymbol{L}^{P}_\nu{}^{\mu}\gamma^\nu. \tag{2.328}$$

In this case, we cannot solve the equation in terms of infinitesimal transformations. Fortunately, it can be solved directly. In this case note that

$$S\left(\boldsymbol{L}^{p}\right)\gamma^{\mu}S^{-1}\left(\boldsymbol{L}^{P}\right)=-(-1)^{\delta_{\mu 0}}\gamma^{\mu} \tag{2.329}$$

Note that $\delta_{\mu 0}$ and γ^{μ} are not Einstein summed.

Using the fact, that $\{\gamma^0,\gamma^\mu\}=2\boldsymbol{g}^{0\mu}$, we obtain that

$$\begin{aligned}\gamma^0\gamma^\mu &= -\gamma^\mu\gamma^0+2\boldsymbol{g}^{0\mu}\\ &= -\mathbf{1}+2\cdot\mathbf{1}=\mathbf{1},\quad m=0\\ &= -\gamma^i\gamma^0,\quad m\neq 0, m=i.\end{aligned} \tag{2.330}$$

Take $S\left(\boldsymbol{L}^{p}\right)=\gamma^0$, then (2.329) reduces to the

$$\gamma^0\gamma^\mu\gamma^{0,-1}=\gamma^0\gamma^\mu\gamma^0=-(-1)^{\delta_{\nu 0}}\gamma^\nu. \tag{2.331}$$

Most generally, we can chose $S\left(\boldsymbol{L}^{P}\right)$ to be

$$S\left(\boldsymbol{L}^{P}\right)=e^{i\phi}\gamma^0. \tag{2.332}$$

Then transformed wave function takes form

$$\psi'(x')=\psi'(-\boldsymbol{x},t)=e^{i\phi}\gamma^0\psi(\boldsymbol{x},t). \tag{2.333}$$

This can be rewritten as

$$\psi_P(x)=\psi'(x')=e^{i\phi}\gamma^0\psi(\boldsymbol{x},t)=\psi(\boldsymbol{L}^{-1}x). \tag{2.334}$$

Here $\psi_P(x)$ is a spatially reflected bispinor. If one requires that after two reflections one gets to the original state, i.e.

$$S\left(\boldsymbol{L}^{P}\right)\psi(x)=\psi(x) \iff e^{i2\phi}\gamma^{0,2}\psi(x)=e^{i2\phi}\psi(x)=\psi(x), \tag{2.335}$$

then this implies that $\phi=0$ or $\phi=\pi$ (mod 2π). Thus in this case we take

$$S\left(\boldsymbol{L}^{P}\right)=\pm\gamma^0=\eta_P\gamma^0, \tag{2.336}$$

where η_P is an internal parity, another quantum number of particle. For rotations, we know that

$$\psi'(x')=S(\boldsymbol{R})\psi(x)=\begin{pmatrix}e^{-\frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}} & 0\\ 0 & e^{-\frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma}}\end{pmatrix}\psi(x). \tag{2.337}$$

Thus for 2π rotation around second axis we get:

$$S(\boldsymbol{R},\theta=2\pi,\text{around 2-axis})\psi(x)=\begin{pmatrix}e^{-i\pi\sigma_3} & 0\\ 0 & e^{-i\pi\sigma_3}\end{pmatrix}\psi(x). \tag{2.338}$$

Under parity transformation the positive Dirac wave function of momentum p transforms as

$$u(p,\lambda)e^{-ipx}\to\gamma^0u(p,\lambda)e^{-ipx_P}=u(p_P,\lambda)e^{-ip_Px}. \tag{2.339}$$

Rewriting this in more detail and recalling that in Dirac representation

$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}$ and also $u(p,\lambda) \propto \begin{pmatrix} \chi\lambda \\ \frac{\sigma\cdot p}{E+m}\chi\lambda \end{pmatrix}$:

$$\begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \chi\lambda \\ \frac{\sigma\cdot p}{E+m}\chi\lambda \end{pmatrix} = \begin{pmatrix} \chi\lambda \\ \frac{-\sigma\cdot p}{E+m}\chi\lambda \end{pmatrix} = \begin{pmatrix} \chi\varphi \\ \frac{\sigma\cdot p_P}{E+m}\chi\varphi \end{pmatrix}. \tag{2.340}$$

That is, the spatial part of the momentum has been reflected, but the spin state has been unaltered. Which is what we would expect from parity transformation. For negative energy solutions we would get

$$v(p,\lambda)e^{ipx} \to \gamma^0 v(p,\lambda)e^{ipx_P} = -v(p_P,\lambda)e^{ip_P x}. \tag{2.341}$$

Using the same argument as for positive energy solutions we can see, that positive and negative energy solutions have relative opposite parities. After the reinterpretation of negative energy solutions this will mean opposite intrinsic parities for particle and antiparticle.

Parity of a Scalar Particle

For a complex wave function $\phi(x)$ of a relativistic scalar particle

$$\phi(x) \to \eta_P \phi_P(x), \tag{2.342}$$

with $\phi_P(x) = \phi(x_P)^*$. E.g. for a state of definite momentum $\phi(p,x) = e^{-ipx}$ we get that

$$\phi_P(p,x) = \{e^{-ipx_P}\}^* = e^{-ip_P x} = \phi(p_P,x). \tag{2.343}$$

Recall that $e^{-i\pi\sigma_3} = \cos(\pi) + i\sin(\pi)\sigma_3 = -1$.

Now, consider 2π rotation not around second axis:

$$S(\boldsymbol{R},\theta = 2\pi, \text{not around 2-axis}) = \begin{pmatrix} -\mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}\psi(x) \neq \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}\psi(x). \tag{2.344}$$

Hence, one should rotate by 4π around to get original bispinor. 4π rotation can be analogue of 4 reflections by assuming, that only after 4 reflections the electron will be in its original state $\psi(x)$. Thus we finally get that

$$S\left(\boldsymbol{L}^P\right)\psi(x) = \pm i\gamma^0\psi(x) = i\eta_P\gamma^0\psi(x). \tag{2.345}$$

Time Reversal

Time reversal transformation action on 4-vector can be described as

$$x^\mu = (t,\boldsymbol{x}) \xrightarrow{T} x^\mu_T = (-t,\boldsymbol{x}). \tag{2.346}$$

This transformation has matrix form of

$$\boldsymbol{L}^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = -\boldsymbol{g}^{\mu\nu} = -\boldsymbol{g}_{\mu\nu}. \tag{2.347}$$

This again satisfies the defining property of Lorentz group

$$L_\alpha{}^\mu g_{\mu\nu} L^\nu{}_\beta = g_{\alpha\beta}, \tag{2.348}$$

and also $\det L = -1$. Again, we can see that $L_\alpha{}^\mu$ is an element of a Lorentz group. Under the time reversal transformation, the following quantities are transformed as

$$v \xrightarrow{T} v_T = -v, \;\; p \xrightarrow{T} p_T = -p, \;\; J \xrightarrow{T} J_T = -J. \tag{2.349}$$

In non-relativistic quantum mechanics we know that the complete effect of a linear operator can be determined by specifying its action on a basis set of the vector space of physical states and then its application by exploiting the linearity of the maps. Similarly, the complete effect of an antilinear map can be determined by specifying its effect on a basis and extending the results using its antilinearity. Take for instance the momentum basis $|\boldsymbol{p}\rangle$. Then $\hat{\boldsymbol{T}}\,|\boldsymbol{p}\rangle = |-\boldsymbol{p}\rangle$. A generic state then would look like

$$|\psi\rangle = \sum_{\boldsymbol{p}} \widetilde{\psi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle\,. \tag{2.350}$$

From this we can see, that effect of time reversal is then

$$\begin{aligned}\hat{\boldsymbol{T}}\,|\psi\rangle &= \hat{\boldsymbol{T}}\sum_{\boldsymbol{p}} \widetilde{\psi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle = \sum_{\boldsymbol{p}} \widetilde{\psi}^*(\boldsymbol{p})\hat{\boldsymbol{T}}\,|\boldsymbol{p}\rangle \\ &= \sum_{\boldsymbol{p}} \widetilde{\psi}^*(\boldsymbol{p})\,|-\boldsymbol{p}\rangle = \sum_{\boldsymbol{p}} \widetilde{\psi}^*(-\boldsymbol{p})\,|\boldsymbol{p}\rangle \\ &= \sum_{\boldsymbol{p}} \widetilde{\psi}_T(\boldsymbol{p})\,|\boldsymbol{p}\rangle = |\psi_T\rangle\,.\end{aligned} \tag{2.351}$$

Hence we see that $\widetilde{\psi}_T(\boldsymbol{p}) = \widetilde{\psi}^*(-\boldsymbol{p})$ for scalar wave functions. If the state $|\phi\rangle = \sum_{\boldsymbol{p}} \widetilde{\phi}(\boldsymbol{p})\,|\boldsymbol{p}\rangle$ is defined similarly, then the scalar product

$$\begin{aligned}\langle\phi|\psi\rangle &= \sum_{\boldsymbol{p}} \widetilde{\phi}^*(\boldsymbol{p})\widetilde{\psi}(\boldsymbol{p}) = \sum_{\boldsymbol{p}} \widetilde{\phi}^*(-\boldsymbol{p})\widetilde{\psi}(-\boldsymbol{p}) \\ &= \left[\sum_{\boldsymbol{p}} \widetilde{\phi}(-\boldsymbol{p})\widetilde{\psi}^*(-\boldsymbol{p})\right]^* = \left[\sum_{\boldsymbol{p}} \widetilde{\phi}_T(\boldsymbol{p})\widetilde{\psi}_T(\boldsymbol{p})\right]^* \\ &= \langle\phi_T|\psi_T\rangle^*\,.\end{aligned} \tag{2.352}$$

In order to get x-representation of our wave function we can apply Fourier transformation as

$$\begin{aligned}\psi_T(\boldsymbol{x}) &= \int \psi_T(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} = \int \psi^*(-\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} \\ &= \int \psi^*(\boldsymbol{p})e^{i\boldsymbol{p}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{p} = \left[\int \psi(\boldsymbol{p})e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}\right]^* = \psi^*(\boldsymbol{x}).\end{aligned} \tag{2.353}$$

$\psi_T(\boldsymbol{x},t) = \left[e^{-i\boldsymbol{H}t}\psi^*(\boldsymbol{x},0)\right] = \left[e^{i\boldsymbol{H}t}\psi(\boldsymbol{x},0)\right]^* = \psi^*(\boldsymbol{x},-t) = \psi^*(x_T)\,.$

In line with non-relativistic quantum mechanics, the effect of time reversal on the Dirac wave function can be written in the form

$$\psi_T(x) = \boldsymbol{B}\psi^*(x_T), \tag{2.354}$$

where the matrix $\boldsymbol{B}$ acts on spinorial indices.

To find $\boldsymbol{B}$ we complex conjugate Dirac equation and take $x_0 \to -x_0$.

This implies that

$$\left(i\gamma^{0*}\partial_t - i\gamma^{i*}\partial_i - m\right)\psi^*(x_T) = 0. \tag{2.355}$$

If we assume that the matrix $\boldsymbol{B}$ is defined so that

$$\boldsymbol{B}\left(\gamma^{0*}, -\gamma^{i*}\right)\boldsymbol{B}^{-1} = \left(\gamma^0, \gamma^i\right), \tag{2.356}$$

Matrices again satisfy Clifford algebra, hence there exists similarity transformation.

then with the definition of $\psi_T(x)$ in (2.354) we get

$$0 = \boldsymbol{B}\left(i\gamma^{0*}\partial_t - i\gamma^{i*}\partial_i - m\right)\psi^*(x_T) = \left(i\gamma^\mu\partial_\mu - m\right)\psi_T(x). \tag{2.357}$$

So that $\psi_T(x)$ satisfies Dirac equation if $\psi(x)$ does. Using the fact that

$$\gamma^{0\dagger} = \gamma^0 \implies \gamma^{0*} = \gamma^{0\intercal}, \tag{2.358}$$

$$\gamma^{i\dagger} = -\gamma^i \implies \gamma^{i*} = -\gamma^{i\intercal}, \tag{2.359}$$

and

$$\boldsymbol{C}\gamma^{\mu\intercal}\boldsymbol{C}^{-1} = -\gamma^\mu. \tag{2.360}$$

where $\boldsymbol{C} = i\gamma^0\gamma^2$ is a charge conjugation matrix (which will be derived further on) we can write $\boldsymbol{B}$ in the form

$$\begin{aligned}\boldsymbol{B} &= \eta_T\gamma_5\boldsymbol{C} = i\eta_T\gamma^0\gamma^1\gamma^2\gamma^3 i\gamma^0\gamma^2 = \eta_T\gamma^1\gamma^3 \\ &= \eta_T\begin{pmatrix}0 & \sigma^1\\ -\sigma^1 & 0\end{pmatrix}\begin{pmatrix}0 & \sigma^3\\ -\sigma^3 & 0\end{pmatrix} = \eta_T\begin{pmatrix}i\sigma_2 & 0\\ 0 & i\sigma_2\end{pmatrix},\end{aligned} \tag{2.361}$$

where η_T is intrinsic time reversal, i.e., another quantum number of the particle (similarly as is intrinsic angular momentum — spin). Indeed

$$\boldsymbol{B}\gamma^{0*}\boldsymbol{B}^{-1} = \gamma_5\boldsymbol{C}\gamma^{0\intercal}\boldsymbol{C}^{-1}\gamma_5 = \gamma^0, \tag{2.362}$$

$$\boldsymbol{B}\gamma^{i*}\boldsymbol{B}^{-1} = -\gamma_5\boldsymbol{C}\gamma^{i\intercal}\boldsymbol{C}^{-1}\gamma_5 = -\gamma^i. \tag{2.363}$$

Since we require $\hat{\boldsymbol{T}}^2 = \mathbf{1}$ we get

$$(i\boldsymbol{B})^2 = \mathbf{1} = i^2\eta_T^2\begin{pmatrix}i\sigma_2 & 0\\ 0 & i\sigma_2\end{pmatrix}^2 = -\eta_T^2\mathbf{1}, \tag{2.364}$$

where we choose $\eta_T = \pm 1$. The situation is the same both in chiral and Dirac representation. Also, it can be easily shown that

$$\boldsymbol{B}^\intercal = -\boldsymbol{B},\ \boldsymbol{B}^* = -\boldsymbol{B}^{-1}. \tag{2.365}$$

2.12 Charge Conjugation

Last of discrete symmetries is a charge conjugation, which

$$e \xrightarrow{\boldsymbol{C}} -e. \tag{2.366}$$

This is best discussed when electromagnetic field is coupled via minimal coupling, i.e.

$$\left[\left(i\partial\!\!\!/ - eA\!\!\!/\right) - m\right]\psi(x) = 0. \tag{2.367}$$

Charge conjugated wave function $\psi_c(x)$ must satisfy

$$\left[\left(i\partial\!\!\!/ + eA\!\!\!/\right) - m\right]\psi_c(x) = 0. \tag{2.368}$$

Note that from (2.367) follows that

$$\psi^\dagger(x)\left[-i\gamma^{\mu,\dagger}\overleftarrow{\partial}_\mu - e\gamma^{\mu,\dagger}A_\mu - m\right] = 0. \tag{2.369}$$

By multiplying this equation from right by γ^0 and using $\gamma^0\gamma^\dagger\gamma^0 = \gamma$ we get

$$\begin{aligned}\psi^\dagger(x)\gamma^0\left[-i\gamma^\mu\overleftarrow{\partial}_\mu - e\gamma^\mu A_\mu - m\right] &= 0,\\ \overline{\psi}(x)\left[i\gamma^\mu\overleftarrow{\partial}_\mu + e\gamma^\mu A_\mu + m\right] &= 0,\\ \left[i\gamma^{\mu,\mathsf{T}}\partial_\mu + e\gamma^{\mu,\mathsf{T}}A_\mu + m\right]\overline{\psi}^\mathsf{T}(x) &= 0,\\ \left[-i\gamma^{\mu,\mathsf{T}}\partial_\mu - e\gamma^{\mu,\mathsf{T}}A_\mu - m\right]\overline{\psi}^\mathsf{T}(x) &= 0.\end{aligned} \tag{2.370}$$

We might therefore assume that

$$\psi_c(x) = \boldsymbol{C}\overline{\psi}^\mathsf{T}(x). \tag{2.371}$$

$\boldsymbol{C}$ must be chosen so that ψ_c satisfies Dirac equation with opposite charge. From all above we get that

$$\left(-i\boldsymbol{C}(\gamma^\mu)^\mathsf{T}\boldsymbol{C}^{-1}\partial_\mu - e\boldsymbol{C}(\gamma^\mu)^\mathsf{T}\boldsymbol{C}^{-1}A_\mu - m\right)\psi_c(x) = 0. \tag{2.372}$$

Assuming $\boldsymbol{C}$ satisfies

$$\boldsymbol{C}(\gamma^\mu)^\mathsf{T}\boldsymbol{C}^{-1} = -\gamma^\mu, \tag{2.373}$$

we get

$\boldsymbol{C}$ exists due to the Pauli theorem.

$$\begin{aligned}\left(i\gamma^\mu\partial_\mu + e\gamma^\mu\partial_\mu - m\right)\psi_c(x) &= 0\\ \left(i\partial\!\!\!/ + eA\!\!\!/ - m\right)\psi_c(x) &= 0.\end{aligned} \tag{2.374}$$

When $\psi(x)$ satisfies (2.367). It can be checked that

$$\boldsymbol{C} = i\gamma^0\gamma^2 = \begin{pmatrix} 0 & i\sigma_2 \\ i\sigma_2 & 0\end{pmatrix}. \tag{2.375}$$

Often $\boldsymbol{C}$ is taken as $\boldsymbol{C} = i\gamma^2\gamma^0$.

$\boldsymbol{C}$ Matrix in Chiral Representation

In chiral representation $\boldsymbol{C}$ matrix is given by

$$\begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0\end{pmatrix}\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0\end{pmatrix} = \begin{pmatrix} -i\sigma_2 & 0 \\ 0 & i\sigma_2\end{pmatrix}. \tag{2.376}$$

Note that

$$C = -C^{-1} = -C^{\dagger} = -C^{\intercal}. \tag{2.377}$$

$\psi_c(x)$ describes particle with the same mass and the same spin direction, but with opposite charge and energy. Change conjugation is antilinear transformation. Let us compute $\psi_c(x)$ for $\psi(x)$ describing a spin-down negative energy electron at rest in absence of external field. Begin with

$$\psi(x) = e^{imt}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \tag{2.378}$$

and take $C = i\gamma^2\gamma^0$. Then $\psi_c(x)$ can be written as

$$\begin{aligned}\psi_c(x) &= \eta_c C\overline{\psi}^{\intercal} = i\eta_c\gamma^2\gamma^0(\psi^{\dagger}(x)\gamma^0)^{\intercal} = \eta_c C\gamma^{0\intercal}\gamma^*(x) = \eta_c(-\gamma^0)C\psi^*(x)\\ &= \eta_c(-\gamma^0)i\gamma^2\gamma^0 e^{-imt}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \eta_c e^{-imt}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}.\end{aligned} \tag{2.379}$$

2.13 Dirac's Hole Theory and Positron

Although we have shown that the Dirac theory accommodates negative energy solutions whose existence should not be ignored, we have as yet not examined the physical significance of those solutions. Let assume that real electrons are described only by positive energy states. These are the states with $E = \sqrt{\boldsymbol{p}^2 + m^2}$. All states of negative energy are occupied by electrons - one electron in each state of negative energy and given $\boldsymbol{p}$ and spin projection λ. In this way a real electron of positive energy is prevented from falling into energetically lower and lower states by radiation emission. This so called **radiation catastrophe** is averted by Pauli's exclusion principle, which prevents these transitions.

One might ask, what is the meaning of a **hole** in the occupied "sea" of negative states. In absence of any field (electromagnetic, e.t.c.) the **vacuum** represents the lower continuum ("Dirac sea"), whole states are completely occupied with electrons. Occasionally one of the negative-energy electrons in the Dirac sea can absorb a photon of energy $\hbar\omega > 2mc^2$ and transit into positive energy states. As a result, a "hole" is created in the Dirac sea. The observable energy of Dirac sea is

$$E_{obs} = E_{vac} - (-|E_e|) = E_{vac} + |E_e|. \tag{2.380}$$

E_{vac} has increased, hence we expect that the absence of a negative-energy electron appears as the presence of a positive-energy particle, a **hole**.

Similarly, when a hole is created in Dirac's sea, the total charge of the

Dirac's sea becomes

$$Q = Q_{vac} - e = Q_{vac} - (-|e|) = Q_{vac} + |e|. \qquad (2.381)$$

Hole in the sea of negative-energy states looks like a positive-energy particle of charge $|e|$. Once we accept that (a) negative-energy are completely filled under normal conditions, (b) negative-energy electron can absorb a photon of energy $> 2mc^2$ (just as a positive-energy electron can) to become a positive-energy, we are unambiguously led to predict the existence of a particle of a charge $|e|$ with a positive energy. This particle is called **positron**.

We may also consider a closely related process

$$e^-_{E>0} \to e^-_{E<0} + 2\gamma, \qquad (2.382)$$

which is allowed when a hole is present in Dirac's sea. We can relate this to another process

$$e^-_{E>0} + e^+_{E>0} \to 2\gamma, \qquad (2.383)$$

Even though the spectrum of energies is identical for Klein-Gordon particle. Existence of Pauli exclusion principle means, that no Dirac vacuum can be formed.

i.e., process when both electron and hole/positron disappear and two photon quanta are generated. This process, called e^-e^+ annihilation is often observed in solids.

Absence of **momentum** $\boldsymbol{p}^e$ in the Dirac sea appears as a presence of $-\boldsymbol{p}^e$ momentum. Thus

$$\boldsymbol{p} = \boldsymbol{p}_{vac} - \boldsymbol{p}^e, \qquad (2.384)$$

or equivalently the moment of hole/positron is

$$\boldsymbol{p}_{hole} = \boldsymbol{p} - \boldsymbol{p}_{vac} = -\boldsymbol{p}^e. \qquad (2.385)$$

Similarly, the absence of spin up $E < 0$ electron is a presence of spin down $E > 0$ positron. We can summarise the changes in dynamical quantities in the table below.

	Q	E	$\boldsymbol{p}$	$\boldsymbol{S}$	$\boldsymbol{h}$
$E < 0$ Electron state	$-\|e\|$	$-\|E\|$	$\boldsymbol{p}$	$\frac{\hbar}{2}\boldsymbol{\Sigma}$	$\frac{1}{2}\boldsymbol{\Sigma}\cdot\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|}$
Positron state	$+\|e\|$	$+\|E\|$	$-\boldsymbol{p}$	$-\frac{\hbar}{2}\boldsymbol{\Sigma}$	$\frac{1}{2}\boldsymbol{\Sigma}\cdot\frac{\boldsymbol{p}}{\|\boldsymbol{p}\|}$

Table 2.2: Comparison of properties of negative-energy electron and ensuing hole/positron

2.14 Antiparticles

It would seem that only fermions can have antiparticles, in fact relativistic mechanics forces us to introduce antiparticle for any particle (be it boson or fermion). Striking feature of the Lorentz transformations is that they do not leave invariant the order of events.

E.g. suppose that event at $\boldsymbol{x}_2$ occurs later that at $\boldsymbol{x}_1$ (i.e. $\boldsymbol{x}_2^0 > \boldsymbol{x}_0^1$). A second observer who sees the first observer moving with velocity $\boldsymbol{v}$ will see will see the event separated by the difference

$$\boldsymbol{x}_2'^0 - \boldsymbol{x}_1'^0 = \boldsymbol{L}^0{}_i(\boldsymbol{v})(\boldsymbol{x}_2^i - \boldsymbol{x}_1^i). \qquad (2.386)$$

"Lorentz Boost Parametrized by Velocity"

For boost we can show that

$$\boldsymbol{L}^i{}_j = \delta^i_j = \boldsymbol{v}_i \boldsymbol{v}_j \frac{(\gamma - 1)}{\boldsymbol{v}^2}. \tag{2.387}$$

$$\boldsymbol{L}^0{}_i = \Lambda^i{}_0 = \gamma \boldsymbol{v}_i. \tag{2.388}$$

$$\boldsymbol{L}^0{}_0 = \gamma. \tag{2.389}$$

Further on we can see that

$$\boldsymbol{x}_2'^0 - \boldsymbol{x}_1'^0 = \gamma(\boldsymbol{x}_2^0 - \boldsymbol{x}_1^0) + \gamma \boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1). \tag{2.390}$$

This is negative if

$$\boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1) < -(\boldsymbol{x}_2^0 - \boldsymbol{x}_1^0), \tag{2.391}$$

which is a seeming paradox. Suppose that 1. observer sees a radiation decay $A \longrightarrow B + C$ at $\boldsymbol{x}_1$, followed by absorption of particle B, e.g. $B + D \longrightarrow E$ at $\boldsymbol{x}_2$. Does the 2. observer then see B absorbed at $\boldsymbol{x}_2$ before it is emitted at $\boldsymbol{x}_1$?

The paradox disappears if we note that the speed $|\boldsymbol{v}| < 1$, so that

$$\begin{aligned}
&(\boldsymbol{x}_2^0 - \boldsymbol{x}_1^0) < -\boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1)\\
&|\boldsymbol{x}_2^0 - \boldsymbol{x}_1^0| < |\boldsymbol{v}(\boldsymbol{x}_2 - \boldsymbol{x}_1)| < |\boldsymbol{v}||(\boldsymbol{x}_2 - \boldsymbol{x}_1)|\\
&|\boldsymbol{x}_2^0 - \boldsymbol{x}_1^0| < |(\boldsymbol{x}_2 - \boldsymbol{x}_1)|
\end{aligned} \tag{2.392}$$

But this is impossible because to travel from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ would require the average velocity greater than 1 (that is $> c$). Since

$$\frac{|\boldsymbol{x}_2 - \boldsymbol{x}_1|}{|\boldsymbol{x}_2^0 - \boldsymbol{x}_1^0|} > 1. \tag{2.393}$$

Temporal order raises no problem for classical physics, but it plays an important role in quantum theories.

The uncertainty principle tells us that when we specify that a particle is at position $\boldsymbol{x}_1$ at time t_1, we cannot also define its velocity precisely. There is a certain chance of particle getting from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ even if $\boldsymbol{x}_1 - \boldsymbol{x}_2$ is spacelike, i.e. $|\boldsymbol{x}_1 - \boldsymbol{x}_2| > |\boldsymbol{x}_1^0 - \boldsymbol{x}_2^0|$. To be more precise, one can obtain from QM commutation relations (setting $(t_1, \boldsymbol{x}_1) = (0, 0)$ and $(t_2, \boldsymbol{x}_2) = (t, x, 0, 0)$) that

$$ct'^2 - \boldsymbol{x}'^2 = c^2 t^2 - \boldsymbol{x}^2 + \hbar^2 c^2 \frac{\hat{\boldsymbol{H}}^{-2}}{4}, \tag{2.394}$$

here $\hat{\boldsymbol{H}}^2 = \boldsymbol{p}^2 c^2 + m^2 c^4$. Since $\hat{\boldsymbol{H}}^2 \geq m^2 c^4$ (in the sense of eigenvalues), we get for a time-like (or light like) interval $c^2t'^2 - \boldsymbol{x}'^2 \geq 0$ that

$$0 \leq c^2t'^2 - \boldsymbol{x}'^2 = c^2t^2 - \boldsymbol{x}^2 + \hbar^2 c^2 \frac{\hat{\boldsymbol{H}}^{-2}}{4} \leq c^2t^2 - \boldsymbol{x}^2 + \frac{\hbar^2 c^2}{4m^2c^2}$$
$$= c^2t^2 - \boldsymbol{x}^2 + \left(\frac{\lambda\!\!\!^{-}}{2}\right)^2 . \tag{2.395}$$

Here $\lambda\!\!\!^{-} = \frac{\hbar}{mc}$ is the Compton wavelength of the particle.

Consequently, the particle can propagate over space-like interval provided that

$$0 > c^2t^2 - \boldsymbol{x}^2 \geq -\left(\frac{\lambda\!\!\!^{-}}{2}\right)^2 . \tag{2.396}$$

Generally, the so-called Weinberg formula holds in this case

$$0 > c^2(t_2 - t_1)^2 - (\boldsymbol{x}_1 - \boldsymbol{x}_2)^2 \geq -\left(\frac{\lambda\!\!\!^{-}}{2}\right)^2 . \tag{2.397}$$

Such space time intervals are very "narrow" even for elementary particle masses, e.g. if m is the mass of proton, $\lambda\!\!\!^{-} = 2 \cdot 10^{-14}$ cm. We are thus faced again with a paradox; if one observer sees a particle emitted at $\boldsymbol{x}_1$ and absorbed at $\boldsymbol{x}_2$, and if the Weinberg formula is satisfied, then a second observer may see the particle absorbed at $\boldsymbol{x}_2$ at a time t_2 before it is emitted at $\boldsymbol{x}_1$ at time t_1. There is only one known way out of this paradox. The second observer must see a particle emitted at $\boldsymbol{x}_2$ and absorbed at $\boldsymbol{x}_1$. But in general the particle seen by the second observer will necessarily be different from the first one.

For instance, if the first observer sees a proton turn into a neutron and positive π-meson at $\boldsymbol{x}_1$ and then sees π^+ and some other neutron turn into proton at $\boldsymbol{x}_2$, then the second observer must see the neutron at $\boldsymbol{x}_2$ turn into proton and a particle of a **negative** charge, which is then absorbed by a proton at $\boldsymbol{x}_1$ that turns into a neutron. Since the rest mass is a Lorentz invariant, the mass of the negative charged particle seen by the second observer will be equal to π^-.

This conclusion is not obtainable in non-relativistic quantum mechanics, nor in relativistic classical mechanics.

For every type of charged particle there is an oppositely charged particle of equal mass.

"Feynman–Stueckelberg Interpretation"

Uncertainty relations allow a particle tunnel from time-like to space-like regions. In Feynman–Stueckelberg we assume that antiparticle is particle with negative energy, mass, charge and spin moving **backward** in time.

2.15 Central field problem: exact solution (fine structure of atomic spectra)

Let us consider Dirac's hamiltonian describing a particle in a central scalar potential. The energy eigenvalue equation and conserved angular momentum are

$$H_D\Psi = [\underbrace{\alpha\cdot p}_{\text{Dirac rep.}\ \begin{pmatrix}0 & \sigma\\ \sigma & 0\end{pmatrix}} + \underbrace{\beta m}_{\begin{pmatrix}\mathbf{1} & 0\\ 0 & -\mathbf{1}\end{pmatrix}} + V(r)]\Psi = E\Psi \tag{2.398}$$

$$\underbrace{J}_{\text{Conserv. ang. momenta}} = \underbrace{r\times p}_{\text{orbital angular momenta}} + \hat{\sigma}$$

Correct notation $\hat{\sigma}^i = \frac{i}{2}\epsilon^i_{jk}\gamma^j\gamma^k$ ($\epsilon^{123} = 1$, $\epsilon_{123} = -1$) $\Rightarrow$ $\hat{\sigma}^i = \frac{i}{2}\epsilon^i_{jk}(-i)\otimes\epsilon^{jk}_l\sigma^l = \frac{i}{2}\eta^i_l(\mathbb{I}\otimes\sigma^l)$

$$\hat{\sigma}_i = \frac{i}{2}\epsilon_{ijk}\gamma^j\gamma^k = \frac{i}{4}\epsilon_{ijk}\underbrace{[\gamma^j,\gamma^k]}_{\sigma^{jk}}$$

Note

Conserved quantity must transform state vector in the same space-time point namely

$$\Psi'(x) = e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\Psi(x)\underbrace{\sim}_{|\omega^{\mu\nu}|<<1}(\mathbb{I}-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu})\Psi(x) \tag{2.399}$$

We know that

$$\Psi'(x) = S(L)\Psi(L^{-1}x)\underbrace{\longmapsto}_{|\omega|<<1}(\mathbb{I}-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu})\Psi(L^{-1}x)$$

$$= (\mathbb{I}-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu})\underbrace{\Psi(x^\rho-\omega^\rho{}_\nu x^\nu)}_{\text{or equivalently } \Psi(x^\rho-\omega_{\mu\nu}(M^{\mu\nu})^\rho{}_\alpha x^\alpha)=\Psi(x)-\omega^\rho{}_\nu x^\nu\frac{\partial}{\partial x^\rho}\Psi(x)}$$

to linear in ω

$$= (\mathbb{I}-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu} - \underbrace{\omega^\rho{}_\nu x^\nu}_{\omega^{\rho\nu}x_\nu}\frac{\partial}{\partial x^\rho})\Psi(x) \tag{2.400}$$

$$= [1-\frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu} - \frac{1}{2}\omega^{\mu\nu}(x_\nu\partial_\mu - x_\mu\partial_\nu)]\Psi(x)$$

$$= [1-\frac{i}{2}\omega^{\mu\nu}(\frac{1}{2}\sigma_{\mu\nu} + i(x_\mu\partial_\nu - x_\nu\partial_\mu))]\Psi(x)$$

$$\Rightarrow J_{\mu\nu} = \frac{1}{2}\sigma_{\mu\nu} + i(x_\mu\partial_\nu - x_\nu\partial_\mu)$$

Note: Reason why this should be conserved follows from the fact that

$$
\begin{aligned}
H\Psi = E\Psi &\longmapsto e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}} H e^{\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}} e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\Psi = E e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\Psi \\
&\Rightarrow \text{ if } e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}} H e^{\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}} = H \\
&\Rightarrow [J_{\mu\nu}; H] = 0 \\
&\Rightarrow \underbrace{\Psi'(x) = e^{-\frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}}\Psi(x)}_{\text{same argument}} \; (*)
\end{aligned}
\tag{2.401}
$$

so if $J_{\mu\nu}$ is a conserved angular momentum Ψ must transform as (*).

In particular not $\frac{1}{2}\sigma_{\mu\nu}$ nor $i(x_\mu\partial_\nu - x_\nu\partial_\mu)$ are conserved separately.

The first thing to note is that $\{H_D; \boldsymbol{J}^2; J_Z\}$ may be simultaneously diagonalized. In both Weyl Dirac representations of the γ-mat. $\hat{\sigma}$ is diagonal and the angular momentum operator acts in the same way on the upper and lower bispinor components

$$
\boldsymbol{J}\Psi = \begin{pmatrix} \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix} \begin{pmatrix} \Psi^+ \\ \Psi^- \end{pmatrix} \tag{2.402}
$$

Proof:

$$
\begin{aligned}
[H_D; \boldsymbol{J}] &= \left[\begin{pmatrix} [m + V(r)]\mathbb{I}_{2\times 2} & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & [V(r) - m]\mathbb{I}_{2\times 2} \end{pmatrix}; \begin{pmatrix} \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix} \right] \\
&= \underbrace{[\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m; \hat{\boldsymbol{\sigma}}(= \frac{1}{2}\boldsymbol{\Sigma})]}_{A} + \underbrace{[\boldsymbol{\alpha}\cdot\boldsymbol{p} + V(r); \boldsymbol{r}\times\boldsymbol{p}]}_{B}
\end{aligned}
$$

$$
\begin{aligned}
\text{A}: \left[\begin{pmatrix} m & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & m \end{pmatrix}; \begin{pmatrix} \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix} \right] &= \begin{pmatrix} \frac{m}{2}\boldsymbol{\sigma} & \frac{1}{2}(\boldsymbol{\sigma}\cdot\boldsymbol{p})\cdot\boldsymbol{\sigma} \\ \frac{1}{2}(\boldsymbol{\sigma}\cdot\boldsymbol{p})\cdot\boldsymbol{\sigma} & -\frac{m}{2}\boldsymbol{\sigma} \end{pmatrix} - \begin{pmatrix} \frac{m}{2}\boldsymbol{\sigma} & \frac{1}{2}\boldsymbol{\sigma}\cdot(\boldsymbol{\sigma}\cdot\boldsymbol{p}) \\ \frac{1}{2}\boldsymbol{\sigma}\cdot(\boldsymbol{\sigma}\cdot\boldsymbol{p}) & -\frac{m}{2}\boldsymbol{\sigma} \end{pmatrix} \\
&= \begin{pmatrix} 0 & \frac{1}{2}[\boldsymbol{\sigma}\cdot\boldsymbol{p}; \boldsymbol{\sigma}] \\ \frac{1}{2}[\boldsymbol{\sigma}\cdot\boldsymbol{p}; \boldsymbol{\sigma}] & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2}[\sigma_i; \boldsymbol{\sigma}]p_i \\ \frac{1}{2}[\sigma_i; \boldsymbol{\sigma}]p_i & 0 \end{pmatrix} \\
&= i \begin{pmatrix} 0 & \boldsymbol{\sigma}\times\boldsymbol{p} \\ \boldsymbol{\sigma}\times\boldsymbol{p} & 0 \end{pmatrix} = i\boldsymbol{\alpha}\times\boldsymbol{p}
\end{aligned}
$$

$$
\text{B}: [\boldsymbol{\alpha}\cdot\boldsymbol{p} + V(r); \boldsymbol{r}\times\boldsymbol{p}] = \underbrace{[V(r); \boldsymbol{r}\times\boldsymbol{p}]}_{\text{a)}} + \underbrace{[\boldsymbol{\alpha}\cdot\boldsymbol{p}; \boldsymbol{r}\times\boldsymbol{p}]}_{\text{b)}}
$$

$$
\text{a)}: [V(r); \boldsymbol{r}\times\boldsymbol{p}]_k = \epsilon^{ijk}[V(r); r^i p^j] = \epsilon^{ijk} r^i \underbrace{[V(r); p^j]}_{i\partial_j V(r)} = \epsilon^{ijk} r^i \left(i\frac{\partial}{\partial x^j} V(r) \right) = \underbrace{\epsilon^{ijk}}_{antisym.} \underbrace{r^i (-\frac{i r^j}{r}}_{sym.} \frac{\partial V}{\partial r}) = 0
$$

b) : $[V(r); \boldsymbol{r} \times \boldsymbol{p}]_k = \alpha^i \epsilon^{lmk}[p^i; r^l p^m] = \alpha^i \epsilon^{lmk} \underbrace{[p^i; r^l]}_{-i\delta^{il}} p^m = -i\alpha^i \epsilon^{imk} p^m = -i(\boldsymbol{\alpha} \times \boldsymbol{p})_k$

$$\Rightarrow [H_D; \boldsymbol{J}] = 0 \Rightarrow\Rightarrow [H; \boldsymbol{J}^2]\& \Rightarrow [H; J_Z]$$

Also trivially $[J^2; J_Z] = 0$.

Consequently, a Dirac spinor angular momentum eigenstates are

$$\boldsymbol{J}^2\Psi = j(j+1)\Psi; \quad J_3\Psi = J_Z\Psi = m\Psi \quad (m \in (-J, ...J)) \tag{2.403}$$

Must be composed of two-component Pauli spinors with the same angular momentum eigenvalues

$$\Psi = \begin{pmatrix} \varphi^{Pauli}_{j,m} \\ \chi^{Pauli}_{j,m} \end{pmatrix}$$

Remember

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{L} + \frac{1}{2}\boldsymbol{\sigma} \end{pmatrix}$$

There is another operator that commutes with H_D and $\boldsymbol{J}$. Intuitively, we expect that we must be able to specify whether the electron spin is parallel or antiparallel to the total angular momentum. In non-relativistic QM these two possibilities are distinguished by the eigenvalues of

$$\boldsymbol{\sigma} \cdot \boldsymbol{J} = \boldsymbol{\sigma}(\boldsymbol{L} + \frac{\boldsymbol{\sigma}}{2})(*) \tag{2.404}$$

For a relativistic electron we might try 4x4 generalization of (*), namely $\boldsymbol{\Sigma} \cdot \boldsymbol{J}$ or $\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J}$. Commutator of this with H_D is rather involved.

One might thus try $\beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}$, which has the same non-relativistic limit as $\boldsymbol{\Sigma} \cdot \boldsymbol{J}$. Since

$$[H_D; \beta\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J}] = \frac{1}{4}[H_D; \beta] \quad \text{or} \quad [H_D; \boldsymbol{\Sigma} \cdot \boldsymbol{J}] = \frac{1}{2}[H_D; \beta] \tag{2.405}$$

Proof:

$$[H_D; \beta\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J}] = [H_D; \beta]\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J} + \beta[H_D; \hat{\boldsymbol{\sigma}}] \cdot \boldsymbol{J} = -2\beta(\boldsymbol{\alpha} \cdot \boldsymbol{p})(\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{J}) + i\beta(\boldsymbol{\alpha} \times \boldsymbol{p}) \cdot \boldsymbol{J}$$

or

$$[H_D; \beta\boldsymbol{\Sigma} \cdot \boldsymbol{J}] = -2\beta(\boldsymbol{\alpha} \cdot \boldsymbol{p})(\hat{\boldsymbol{\Sigma}} \cdot \boldsymbol{J}) + 2i\beta(\boldsymbol{\alpha} \times \boldsymbol{p}) \cdot \boldsymbol{J}$$

since

$$\begin{aligned}(\boldsymbol{\alpha}\cdot\boldsymbol{p})(\hat{\boldsymbol{\Sigma}}\cdot\boldsymbol{J}) &= (\sigma_1\otimes\boldsymbol{\sigma}\cdot\boldsymbol{p})(\mathbb{I}\otimes\boldsymbol{\sigma}\cdot\boldsymbol{J})\\ &= (\sigma_1\otimes(\boldsymbol{\sigma}\cdot\boldsymbol{p})(\boldsymbol{\sigma}\cdot\boldsymbol{J})) = \sigma_1\otimes\sigma_i\sigma_j p_i \mathbb{I}_j\\ &= \sigma_1\otimes(\frac{1}{2}\underbrace{\{\sigma_i\sigma_j\}}_{2\delta_{ij}}+\frac{1}{2}\underbrace{[\sigma_i\sigma_j]}_{i\epsilon_{ijk}\sigma_k})p_iJ_j\\ &= \underbrace{(\sigma_1\otimes\boldsymbol{p}\cdot\boldsymbol{J})}_{\gamma^5\boldsymbol{p}\cdot\boldsymbol{J}}+i\underbrace{(\sigma_1\otimes\boldsymbol{\sigma}\cdot(\boldsymbol{p}\times\boldsymbol{J}))}_{i\boldsymbol{\alpha}\cdot\boldsymbol{p}\times\boldsymbol{J}}\end{aligned}$$

$$\begin{aligned}\Rightarrow&[H_D;\beta\boldsymbol{\Sigma}\cdot\boldsymbol{J}] = -2\beta\gamma^5\boldsymbol{p}\boldsymbol{J}-2\beta(i\boldsymbol{\alpha}\cdot(\boldsymbol{p}\times\boldsymbol{J}))+2i\beta(\boldsymbol{\alpha}\times\boldsymbol{p})\cdot\boldsymbol{J}\\ &= -2\beta\gamma^5\boldsymbol{p}\boldsymbol{J} = -2\beta\gamma^5(\boldsymbol{p}\cdot(\boldsymbol{L}+\frac{\boldsymbol{\Sigma}}{2}) = -2\beta\gamma^5\boldsymbol{p}\cdot\frac{\boldsymbol{\Sigma}}{2}\\ &= -2\beta(\sigma_1\otimes\boldsymbol{\sigma})\cdot\boldsymbol{p}\frac{1}{2} = -\beta\boldsymbol{\alpha}\cdot\boldsymbol{p} = \frac{1}{2}[H_D;\beta]\end{aligned}$$

Therefore an operator k is defined as

$$\begin{aligned}k &= \beta\boldsymbol{\Sigma}\cdot\boldsymbol{J}-\frac{\beta}{2} = \beta(\boldsymbol{\Sigma}(\boldsymbol{L}+\frac{\boldsymbol{\Sigma}}{2})-\frac{1}{2}) = \beta(\boldsymbol{\Sigma}\boldsymbol{L}+\frac{\boldsymbol{\Sigma}^2}{2}-\frac{1}{2})\\ &= \beta(\boldsymbol{\Sigma}\boldsymbol{L}+\frac{3}{2}-\frac{1}{2}) = \beta(\boldsymbol{\Sigma}\boldsymbol{L}+1)\end{aligned} \tag{2.406}$$

Does commute with H_D: $[H_D;k]=0$. Furthermore, since $\boldsymbol{J}$ commutes with β and $\boldsymbol{\Sigma}\cdot\boldsymbol{L}\Rightarrow[\boldsymbol{J};k]=0\Rightarrow[\boldsymbol{J}^2;k]=0$.

$\Rightarrow$ For Dirac's particle in a central potential we can construct a simultaneous eigenfunctions of H_D; k; $\boldsymbol{J}^2$ and J_z. The corresponding eigenvalues are denoted by E, $-K$, $j(j+1)$ and m.

Note that K and j are not totally independent (similarly as $m\in[-j;j]$).

Consider first

$$\begin{aligned}K^2 &= \beta(\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1)\beta(\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1) = \beta^2(\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1)^2\\ &= (\boldsymbol{\Sigma}\cdot\boldsymbol{L})(\boldsymbol{\Sigma}\cdot\boldsymbol{L})+2\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1 = \begin{pmatrix}\sigma^i & 0\\ 0 & \sigma^i\end{pmatrix}\begin{pmatrix}\sigma^j & 0\\ 0 & \sigma^j\end{pmatrix}L_iL_j+2\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1\\ &= \begin{pmatrix}\sigma^i\sigma^j & 0\\ 0 & \underbrace{\sigma^i\sigma^j}_{\delta^{ij}+i\epsilon^{ijk}\sigma^k}\end{pmatrix}L_iL_j+2\boldsymbol{\Sigma}\cdot\boldsymbol{L}+1 = \boldsymbol{L}^2+i\Sigma^k\epsilon^{ijk}L_iL_j+2\boldsymbol{\Sigma}\boldsymbol{L}+1\\ &= \boldsymbol{L}^2-\boldsymbol{\Sigma}\boldsymbol{L}+2\boldsymbol{\Sigma}\boldsymbol{L}+1 = \boldsymbol{L}^2+\boldsymbol{\Sigma}\boldsymbol{L}+1\end{aligned} \tag{2.407}$$

At the same time, since

$$\boldsymbol{J}^2 = (\boldsymbol{L}+\frac{\boldsymbol{\Sigma}}{2})^2 = \boldsymbol{L}^2+\frac{2}{2}\boldsymbol{\Sigma}\cdot\boldsymbol{L}+\frac{\Sigma^2}{4} = \boldsymbol{L}^2+\boldsymbol{\Sigma}\cdot\boldsymbol{L}+\frac{3}{4} \tag{2.408}$$

We have $K^2 = \boldsymbol{J}^2+\frac{1}{4}$.

⇒ eigenvalues $\boldsymbol{J}^2$ and K^2 are related to each other by

$$K^2 = (j(j+1) + \frac{1}{4} = (j + \frac{1}{2})^2 \tag{2.409}$$

⇒ !! $K = \pm\left(j + \frac{1}{2}\right)$ ⇒ K is a non-zero integer which can be both positive and negative.

K is explicitly given by

$$\hat{K} = \begin{pmatrix} \boldsymbol{\sigma L} + 1 & 0 \\ 0 & -\boldsymbol{\sigma L} + 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\sigma J} - \frac{1}{2} & 0 \\ 0 & -\boldsymbol{\sigma J} + \frac{1}{2} \end{pmatrix} \tag{2.410}$$

Note:

Pictorially speaking, the sign of K determines whether the spin is antiparallel (K > 0) or parallel (K < 0) to $\boldsymbol{J}$ in the non-relativistic limit.

(picture)

⇒ if the four-component wave function Ψ (assumed to be an energy eigenfunction) is a simultaneous eigenfunction of K, $\boldsymbol{J}^2$ and J_z then

$$\underbrace{(\boldsymbol{\sigma L} + 1)\Psi_+ = -K\Psi_+; \quad (\boldsymbol{\sigma L} + 1)\Psi_- = K\Psi_-}_{\hat{k}\Psi = -K\Psi}$$

and

$$\boldsymbol{J}^2\Psi_{+/-} = (\boldsymbol{L} + \boldsymbol{\sigma}/2)^2\Psi_{+/-} = j(j+1)\Psi_{+/-}$$

$$J_3\Psi_{+/-} = (L_3 + \sigma_3/2)^2\Psi_{+/-} = m\Psi_{+/-}$$

Note:

The operator $\boldsymbol{L}^2 = \boldsymbol{J}^2 - \boldsymbol{\Sigma}\cdot\boldsymbol{L} - \frac{3}{4} = -(\boldsymbol{\sigma L} + 1) + \frac{1}{4} \Rightarrow \boldsymbol{L}^2 = \boldsymbol{J}^2 - \boldsymbol{\sigma}\cdot\boldsymbol{L} - \frac{3}{4}$ when it acts on Ψ_+ and Ψ_-.

$$\boldsymbol{L}^2\Psi_+ = j(j+1)\Psi_+ + K\Psi_+ + \frac{1}{4}\Psi_+ \Rightarrow l_+(l_+ + 1) = j(j+1) + k + \frac{1}{4}$$

$$\boldsymbol{L}^2\Psi_- = j(j+1)\Psi_- - K\Psi_- + \frac{1}{4}\Psi_- \Rightarrow l_-(l_- + 1) = j(j+1) - k + \frac{1}{4}$$

⇒ so any two-component eigenfunction of $\boldsymbol{\sigma L} + 1$ and $\boldsymbol{J}^2$ is also an eigenfunction of $\boldsymbol{L}^2$

Thus, although the four-component $\Psi = \binom{\Psi_+}{\Psi_-}$ is not an eigenfunction of $\boldsymbol{L}^2$ (since $\hat{H}_D$ does not commute with $\boldsymbol{L}^2$) Ψ_+ and Ψ_- separately are eigenfunctions of $\boldsymbol{L}^2$ whose eigenvalues are denoted by $l_+(l_+ + 1)$ and $l_-(l_- + 1)$ resp. This implies that since (for each two-component eigenfunction)

$$\boldsymbol{L}^2 = \boldsymbol{J}^2 - \boldsymbol{\sigma} \cdot \boldsymbol{L} - \frac{3}{4} = \boldsymbol{J}^2 - \hat{K} + \frac{1}{4} \Rightarrow \hat{L} = \boldsymbol{J}^2 - \boldsymbol{L}^2 + \frac{1}{4}$$

$$\Rightarrow -K = j(j+1) - l_+(l_+ + 1) + \frac{1}{4}$$
$$K = j(j+1) - l_-(l_- + 1) + \frac{1}{4}$$

using now the fact that $K = \pm\left(j + \frac{1}{2}\right)$ we can determine l_+ and l_- for a given K

Table 2.3: Relations among K, j, l_+ and l_-

	l_+	l_-
$K = j + \frac{1}{2}$	$j + \frac{1}{2}$	$j - \frac{1}{2}$
$K = -(j + \frac{1}{2})$	$j - \frac{1}{2}$	$j + \frac{1}{2}$

For given j; $l_\pm$ we can assume two possible values corresponding to two possible values of K, so instead of K one can use l_+ and l_-.

After this preliminary we can write Ψ as

$$\Psi = \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = \begin{pmatrix} g(r) Y^m_{j,l_+} \\ i f(r) Y^m_{j,l_-} \end{pmatrix} \tag{2.411}$$

where $Y^m_{j,l}$ stands for a normalized spin-angular function (i. e., r-independent eigenfunctions of $\boldsymbol{J}^2$, J_3, L^2 and of course $\hat{\boldsymbol{\sigma}}^2 = (\frac{1}{2}\boldsymbol{\Sigma})^2$) formed by the combination of the Pauli spinor with the spherical harmonics of order l.

Let me remind the Clebsch-Gordan expansion

$$|j, m, l\rangle = \alpha \left| l; m_l = m_j - \frac{1}{2} \right\rangle \oplus \left| \frac{1}{2}; \frac{1}{2} \right\rangle + \beta \left| l; \underbrace{m_l}_{J_3 = L_3 + \hat{\sigma}_3 \Rightarrow L_3 = m_l = m_j \pm \frac{1}{2}} = m_j + \frac{1}{2} \right\rangle \ominus \left| \frac{1}{2}; -\frac{1}{2} \right\rangle \tag{2.412}$$

for any l.

more explicitly, when $j = l + \frac{1}{2}$

$$Y^m_{j,l} = \sqrt{\frac{l + m + \frac{1}{2}}{2l + 1}} Y_l^{m_j - 1/2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sqrt{\frac{l - m + \frac{1}{2}}{2l + 1}} Y_l^{m_j + 1/2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.413}$$

and for $j = l - \frac{1}{2}$

$$Y^m_{j,l} = -\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}Y_l^{m_j-1/2}\begin{pmatrix}1\\0\end{pmatrix} + \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}Y_l^{m_j+1/2}\begin{pmatrix}0\\1\end{pmatrix} \tag{2.414}$$

We would like to substitute our solution to a Dirac equation and solve it for the radial functions and find the spectrum. In Dirac representation Dirac's equation splits as:

$$\begin{aligned} H_D &= [\underbrace{\boldsymbol{\alpha}\cdot\boldsymbol{p}}_{\begin{pmatrix}0 & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & 0\end{pmatrix}} + \underbrace{\beta m}_{\begin{pmatrix}\mathbb{I} & 0\\ 0 & \mathbb{I}\end{pmatrix}} + V(r)] \Rightarrow H_D\Psi = E\Psi \Rightarrow \\ &(\boldsymbol{\sigma p})\Psi_- = (E - V(r) - m)\Psi_+ \\ &(\boldsymbol{\sigma p})\Psi_+ = (E - V(r) + m)\Psi_- \end{aligned} \tag{2.415}$$

Note in this connection that

$$\begin{aligned} \boldsymbol{\sigma}\cdot\boldsymbol{p} &= \underbrace{\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r^2}}_{\sigma_i\sigma_j x_i x_j = \frac{1}{2}\{\sigma_i;\sigma_j\}x_i x_j = \boldsymbol{x}^2 = r^2}(\boldsymbol{\sigma}\cdot\boldsymbol{p}) \\ &= \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r^2}(\underbrace{\sigma^i\sigma^j}_{\delta^{ij}+i\epsilon^{ijk}\sigma^k} x^i p^j) = \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r^2}(\boldsymbol{x}\cdot\boldsymbol{p} + i\boldsymbol{\sigma}\cdot\boldsymbol{L}) \\ &= \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r^2}(-ir\frac{\partial}{\partial r} + i\boldsymbol{\sigma}\cdot\boldsymbol{L}) = (\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}})\left[-i\frac{\partial}{\partial r} + \frac{i\boldsymbol{\sigma}\cdot\boldsymbol{L}}{r}\right] \end{aligned} \tag{2.416}$$

In spherical coord. $\boldsymbol{x}$ has radial dir.
$\Rightarrow \boldsymbol{x} = r e^s_r + 0 e^s_\theta + 0 e^s_\varphi$
$\boldsymbol{\Delta} = e^s_r \frac{\partial}{\partial r} + \frac{1}{r} e^s_\theta \frac{\partial}{\partial\theta} + e^s_\varphi \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}$
$\Rightarrow \boldsymbol{x}\cdot\boldsymbol{p} = -ir\frac{\partial}{\partial r}$

Note:

$\underbrace{\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r}}_{\text{projection of spin into the unit radius vector}}$ is a pseudoscalar (under parity $\boldsymbol{x} \mapsto -\boldsymbol{x}$; $r \mapsto r$; $\boldsymbol{\sigma} \mapsto \boldsymbol{\sigma}$)

$$\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r} \xrightarrow{P} -\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r}$$

$$\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r} = \boldsymbol{\sigma}\hat{\boldsymbol{r}} = \begin{Bmatrix}\hat{r}_x = \sin\theta\cos\varphi \\ \hat{r}_y = \sin\theta\sin\varphi \\ \hat{r}_z = \cos\theta\end{Bmatrix} = \begin{pmatrix}\cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta\end{pmatrix}$$

First, the effect of $(\boldsymbol{\sigma}\cdot\boldsymbol{L}$ on J^m_{jl} is simple. Namely, observe

$$\underbrace{\hat{K}}_{\text{we already know}} = \beta(\boldsymbol{\Sigma L} + 1) = \begin{pmatrix}\boldsymbol{\sigma L} + 1 & 0 \\ 0 & -\boldsymbol{\sigma L} - 1\end{pmatrix} \tag{2.417}$$

$$
\begin{aligned}
\Rightarrow \hat{K}\Psi \mapsto -K\Psi \rightarrow -K\Psi_- &= (-\boldsymbol{\sigma}\boldsymbol{L} - 1)\Psi_- \\
\Rightarrow \boldsymbol{\sigma}\boldsymbol{L}\Psi_- &= (K-1)\Psi_- \\
\text{and } -K\Psi_+ &= (\boldsymbol{\sigma}\boldsymbol{L} + 1)\Psi_+ \\
\Rightarrow \boldsymbol{\sigma}\boldsymbol{L}\Psi_+ &= -(K+1)\Psi_+
\end{aligned}
$$

It is trickier to calculate the effect of the matrix factor

$$
\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{x})}{r} = \boldsymbol{\sigma}\hat{\boldsymbol{r}} = \begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta \end{pmatrix}
$$

on the spinor wave functions. In principle, we can carry out the multiplication directly on J_{jl}^m. Then we would have various identities for spherical harmonics $Y_l^m(\theta,\varphi)$.

Note:
$\langle\hat{\boldsymbol{n}}|l,m\rangle = \Psi^m{}_l(\theta,\varphi) = \Psi^m{}_l|\hat{\boldsymbol{n}}\rangle$
$|\hat{\boldsymbol{n}}\rangle$ = *direction eigen vect.*
$(\langle\boldsymbol{x}|n,l,m\rangle = R_{ml}(r)\Psi^m{}_l(\theta,\varphi))$

There is an easier way, however. Observe that $\boldsymbol{\sigma}\hat{\boldsymbol{r}}$ commutes with $\boldsymbol{J}^2$, $\boldsymbol{L}^2$, $S^2\left(=\frac{\sigma^2}{4}\right)$ and J_3.

Proof:

$$
[\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}};\boldsymbol{J}^2] = \underbrace{[\hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma};\boldsymbol{L}^2]}_{A} + \underbrace{[\hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma};\boldsymbol{\sigma}^2/4]}_{B} \tag{2.418}
$$

$$
\begin{aligned}
\text{A}: \frac{1}{r}\sigma_i[x_i;L^2] &= \frac{\sigma_i}{r}L_j\underbrace{[x_i;L_j]}_{-i\epsilon_{jik}x_k} + \frac{\sigma_i}{r}\underbrace{[x_i;L_j]}_{-i\epsilon_{jik}x_k}L_j \\
&= \frac{\sigma_i}{r}[-i\underbrace{\epsilon_{jik}}_{\text{antisym.}}\underbrace{(L_jx_k + x_kL_j)}_{\text{sym.}}] = 0
\end{aligned}
$$

$$
\begin{aligned}
\text{B}: [\hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma};\boldsymbol{\sigma}^2/4] &= \frac{\hat{r}_i}{4}\left\{[\sigma_i;\sigma_j]\sigma_j + \sigma_j[\sigma_i;\sigma_j]\right\} \\
&= \frac{\hat{r}_i}{4}\left\{2i\epsilon_{ijk}\sigma_k\sigma_j + \sigma_j 2i\epsilon_{ijk}\sigma_k\right\} \\
&= \frac{\hat{r}_i}{4}2i\underbrace{\epsilon_{ijk}}_{\text{antisym.}}\underbrace{\{\sigma_k\sigma_j + \sigma_j\sigma_k\}}_{\text{sym.}} = 0
\end{aligned}
$$

$$
[\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}};J_3] = \underbrace{[\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}};L_3]}_{C} + \underbrace{[\hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma};\sigma_3/2]}_{D} \tag{2.419}
$$

$$\mathrm{C}: \frac{\sigma_i}{r} \underbrace{[x_i ; L_3]}_{-i\epsilon_{3ij}x_j} = -i\epsilon_{3ij}\sigma_i\hat{r}_j$$

$$\mathrm{D}: \frac{\hat{r}_i}{2} \underbrace{[\sigma_i ; \sigma_j]}_{2i\epsilon_{i3k}\sigma_k} = i\epsilon_{i3k}\hat{r}_i\sigma_k = -i\epsilon_{3ik}\hat{r}_i\sigma_k = i\epsilon_{3jk}\sigma_j\hat{r}_k$$

So $\boldsymbol{\sigma}\hat{\boldsymbol{r}}$ is simultaneously diagonal with J^2, J_3. Since $(\boldsymbol{\sigma}\hat{\boldsymbol{r}})^2 = 1$, its eigenvalues are ± 1.

Since $\underbrace{\boldsymbol{\sigma}}_{\text{pseudo-vector}}\underbrace{\hat{\boldsymbol{r}}}_{\text{vector}}$ is a pseudo-scalar under rotation, if we evaluate its effect at one particular $\hat{\boldsymbol{r}}$ (say $\hat{z}$), it should behave that way for all $\hat{\boldsymbol{r}}$. $\hat{z}$ corresponds to $\theta = 0 \Rightarrow$

$$\underbrace{Y_l^m(\theta = 0, \varphi) = \sqrt{\frac{2l+1}{4\pi}}\delta_{m,0}}_{\text{see, e.g. Grandshteyn–Ryzhik}}$$

In which case

$$Y^{m_j}_{j=l\pm 1/2;l} = \frac{1}{\sqrt{2l+1}}\begin{bmatrix} \pm\sqrt{l \pm m + 1/2}\ Y_l^{m_j-1/2} \\ \sqrt{l \mp m + 1/2}\ Y_l^{m_j+1/2} \end{bmatrix} = \frac{1}{\sqrt{4\pi}}\begin{bmatrix} \pm\sqrt{l \pm m + 1/2}\ \delta_{m;1/2} \\ \sqrt{l \mp m + 1/2}\ \delta_{m;-1/2} \end{bmatrix}$$

or

$$J^{m_j}_{j;l=j\mp 1/2}(\theta = 0; \varphi) = \sqrt{\frac{j+1/2}{4\pi}}\begin{bmatrix} \pm\delta_{m;1/2} \\ \delta_{m;-1/2} \end{bmatrix}$$

$\Rightarrow$

$$\begin{aligned} \boldsymbol{\sigma}\cdot\hat{\boldsymbol{z}}\ J^m_{j;l=j\mp 1/2}(\theta = 0; \varphi) &= \sqrt{\frac{j+1/2}{4\pi}}\begin{bmatrix} \pm\delta_{m;1/2} \\ -\delta_{m;-1/2} \end{bmatrix} \\ &= -\sqrt{\frac{j+1/2}{4\pi}}\begin{bmatrix} \mp\delta_{m;1/2} \\ \delta_{m;-1/2} \end{bmatrix} = -Y^m_{j;l=j\pm 1/2}(\theta = 0; \varphi) \end{aligned}$$

or by rotating both sides of the equation to general $\hat{\boldsymbol{r}}$

$$\Rightarrow (\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}})\ J^m_{j;l=j\mp 1/2}(\theta; \varphi) = -Y^m_{j;l=j\pm 1/2}(\theta; \varphi) \tag{2.420}$$

Thus

$$
\begin{aligned}
(\boldsymbol{\sigma r})\Psi_{-} &= \frac{i(\boldsymbol{\sigma x})}{r}\left(-i\frac{\partial}{\partial r}f + \frac{i(k-1)}{r}f\right)J^m_{j,l_-} \\
&= \frac{\boldsymbol{\sigma x}}{r}J^m_{j,l_-}\left(\frac{\partial f}{\partial r} - \frac{(k-1)}{r}f\right) \\
&= -\frac{\partial f}{\partial r}J^m_{j,l_+} - \frac{(1-k)}{r}fJ^m_{j,l_-}
\end{aligned}
$$

similarly

$$
\begin{aligned}
(\boldsymbol{\sigma r})\Psi_{+} &= \frac{(\boldsymbol{\sigma x})}{r}\left(-i\frac{\partial}{\partial r} + i\boldsymbol{\sigma L}\right)gJ^m_{j,l_+} \\
&= \frac{\boldsymbol{\sigma x}}{r}J^m_{j,l_+}\left(-i\frac{\partial g}{\partial r} - i\frac{(k+1)}{r}g\right) \\
&= i\frac{\partial g}{\partial r}J^m_{j,l_-} + i\frac{(k+1)}{r}gJ^m_{j,l_+}
\end{aligned}
$$

By pluging to decoupled Dirac equation we get

Note: The spin-angular functions completely drop out.

$$
-\frac{\mathrm{d}f}{\mathrm{d}r} - \frac{(1-k)}{r}f = (E-V-m)g \quad (\bullet)
$$

$$
-\frac{\mathrm{d}g}{\mathrm{d}r} + \frac{(k+1)}{r}g = (E-V+m)f \quad (\bullet\,\bullet)
$$

As in non-relativistic QM we introduce the ansatzes

$$
\underbrace{F(r) = rf(r); \qquad G(r) = rg(r)}_{\text{are supposed to be non-singular} \to f(r);\; g(r) \sim \frac{1}{r},\; r<<1}
$$

$$
\Rightarrow \frac{\mathrm{d}F(r)}{\mathrm{d}r} = f(r) + r\frac{\mathrm{d}f}{\mathrm{d}r} \Rightarrow \frac{\mathrm{d}f}{\mathrm{d}r} + \frac{f}{r} = \frac{1}{r}\frac{\mathrm{d}F(r)}{\mathrm{d}r}
$$

$$
\Rightarrow \frac{\mathrm{d}g}{\mathrm{d}r} + \frac{g}{r} = \frac{1}{r}\frac{\mathrm{d}G}{\mathrm{d}r}
$$

$$
\begin{aligned}
(\bullet) \;\Rightarrow\; & -\frac{1}{r}\frac{\mathrm{d}F}{\mathrm{d}r} + \frac{kF}{r^2} = (E-V-m)G/r \quad /(\cdot(-r)) \\
\Rightarrow\; & \frac{\mathrm{d}F}{\mathrm{d}r} + \frac{k}{r}F = -(E-V-m)G
\end{aligned}
$$

$$(\bullet\,\bullet) \;\Rightarrow\; -\frac{1}{r}\frac{\mathrm{d}G}{\mathrm{d}r} + \frac{kG}{r^2} = (E - V + m)F/r \quad /(\cdot r)$$
$$\Rightarrow\; \frac{\mathrm{d}G}{\mathrm{d}r} + \frac{k}{r}G = (E - V + m)F$$

- anomalous Zeeman ef.
- free spherical waves
- exact solution to the Coulomb scattering problem
- etc.

With this system of equations a variety of problems can be attacked. We shall consider only one problem, i.f. electron bound to the atomic nucleus by a Coulomb potential. This problem can be solved exactly.

Consider $V(r) = -\frac{Z\boldsymbol{e}^2}{r}$, where $\boldsymbol{e} = \frac{e}{\sqrt{4\pi\epsilon_0}}$.

Solution: Let us first look at the asymptotic behavior of F and G at $r \to \infty$

$$\frac{\mathrm{d}G}{\mathrm{d}r} = (E + m)F; \quad \frac{\mathrm{d}F}{\mathrm{d}r} = -(E - m)G$$

$$\frac{\mathrm{d}^2G}{\mathrm{d}r^2} = -(E^2 - m^2)G \quad \text{and} \quad \frac{\mathrm{d}^2F}{\mathrm{d}r^2} = -(E^2 - m^2)F$$

$E = E_{kin.} + E_{pot.} + m$
classically $E_{kin.} + E_{pot.} = const$ for all $\boldsymbol{p}$, so if $\boldsymbol{p} = 0 \Rightarrow const$ is negativ
$\Rightarrow E - m < 0 \Rightarrow E^2 - m^2 < 0$
or $m^2 - E^2 > 0$

$$\Rightarrow G; F \sim e^{\underbrace{-\sqrt{m^2 - E^2}\,r}_{positive}}$$

for $r \to 0$

$$\frac{kG}{r} = -rF \quad \text{and} \quad \frac{-kF}{r} = rG$$

$$\Rightarrow kG = 2\boldsymbol{e}^2F \quad \text{and} \quad kF = 2\boldsymbol{e}^2G$$

$\Rightarrow G$ and F have the same asymptotic at $r \to 0$, due to regularity G, F $\propto r^{s>0}$

To work with dimensionless quantities we can introduce the following scaled variables

$$\alpha_1 = (m + E),\ \alpha_2 = (m - E),\ \gamma = Z\hat{e}^2 = Z\underbrace{\alpha}_{\text{fine structure constant}} \simeq Z/137,\ \rho = \sqrt{\alpha_1\alpha_2}\,r$$

So, the coupled equations $(\bullet)$ and $(\bullet\,\bullet)$ get a form

$$(\bullet) \Rightarrow \left(\frac{\mathrm{d}}{\mathrm{d}r} - \frac{k}{r}\right) F - \left[-(E-m) - \frac{\gamma}{r}\right] G = 0 \quad / \cdot \frac{1}{\sqrt{\alpha_1 \alpha_2}}$$
$$\Rightarrow \left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{k}{\rho}\right) F - \left[\sqrt{\frac{\alpha_2}{\alpha_1}} - \frac{\gamma}{\rho}\right] G = 0$$

similarly

$$\Rightarrow \left(\frac{\mathrm{d}}{\mathrm{d}\rho} - \frac{k}{\rho}\right) G - \left[\sqrt{\frac{\alpha_1}{\alpha_2}} + \frac{\gamma}{\rho}\right] F = 0$$

As in the non-relativistic treatment of the hydrogen atom, we seek solutions in this form

$$F = \underbrace{e^{-\rho}}_{r\to\infty} \underbrace{\rho^s}_{r\to 0} \sum_{m=0} a_m \rho^m; \quad G = e^{-\rho} \rho^s \sum_{m=0} b_m \rho^m$$

"s" furthere constrained normalization condition for Ψ

Substituting this to ($\bullet$) and ($\bullet\ \bullet$) we get recursion treat on

$$\left(\frac{\mathrm{d}}{\mathrm{d}\rho} F = -F + e^{-\rho} \sum_{m=0} a_m (m+s) \rho^{m+s-1}\right)$$

$$(\bullet) \Rightarrow -e^{-\rho} \sum_{m=0}^{\infty} a_m \rho^{m+s} + e^{-\rho} \sum_{m=0}^{\infty} a_m (m+s) \rho^{m+s-1} - k e^{-\rho} \sum_{m=0}^{\infty} a_m \rho^{m+s-1}$$
$$- e^{-\rho} \sum_{m=0}^{\infty} \left(b_m \sqrt{\frac{\alpha_2}{\alpha_1}}\right) \rho^{m+s} + e^{-\rho} \sum_{m=0}^{\infty} b_m \gamma \rho^{m+s-1} = 0$$

comparin coefficients at ρ^{s+m-1}

for $m \neq 0$ $(a \geqslant 1)$

$$a_m (m+s-k) - a_{m-1} + \gamma b_m - \sqrt{\frac{\alpha_2}{\alpha_1}} b_{m+1} = 0 \tag{2.421}$$

similarly

$$(\bullet) \Rightarrow b_m (m+s+k) - b_{m-1} + \gamma a_m - \sqrt{\frac{\alpha_1}{\alpha_2}} a_{m-1} = 0 \tag{2.422}$$

for $m = 0$

$$(\bullet) \Rightarrow a_0 (s-k) + \gamma b_0 = 0$$

$$(\bullet\ \bullet) \Rightarrow b_0 (s+k) - \gamma a_0 = 0$$

$$\Longleftrightarrow \begin{pmatrix} s-k & \gamma \\ -\gamma & s+k \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = 0$$

Since a_0 and b_0 must have non-trivial solution

$$\Rightarrow \begin{vmatrix} s-k & \gamma \\ -\gamma & s+k \end{vmatrix} = 0 \Rightarrow s^2 - k^2 + \gamma^2 = 0 \Rightarrow s = \pm\sqrt{k^2-\gamma^2} \qquad (2.423)$$

Note:

There is an important restriction on the scaling exponent "s" from the normalizability of Dirac's wave function.

$$\|\Psi\|^2 = \int \overline{\Psi}\gamma^0\Psi = \int \Psi^+\Psi < \infty$$
$$= \int \mathrm{d}\Omega \mathrm{d}r r^2 \begin{pmatrix} gJ \\ ifJ \end{pmatrix}^+ \begin{pmatrix} gJ \\ ifJ \end{pmatrix}$$
$$= \int \underbrace{|g|^2 r^2 \mathrm{d}r}_{|G|^2} \int \underbrace{|J|^2 \mathrm{d}\Omega}_{<\infty} + \int \underbrace{|f|^2 r^2 \mathrm{d}r}_{|F|^2} \int \underbrace{|J|^2 \mathrm{d}\Omega}_{<\infty} < \infty(\text{required})$$

$$\Rightarrow \int |G|^2 \mathrm{d}r < 0;\ \int |F|^2 \mathrm{d}r < 0 \Longleftrightarrow \int |G|^2 \mathrm{d}\rho < 0;\ \int |F|^2 \mathrm{d}\rho < 0$$

$\Rightarrow$ F and G must behave better than $\rho^{-1/2}$ (e.g. $\rho^{-1/2+\epsilon(>0)}$) at the origin $\Rightarrow$ $s > -1/2$

$k = (j+1/2)$ =positive
$(j = 1/2; 3/2; 5/2..)$
$\min(k^2) = 1$

However, since

$$k^2 - \gamma^2 \geqslant \min(k^2) - \gamma^2 \simeq 1 - Z^2\alpha^2 \Rightarrow k^2 - \gamma^2 \geqslant 1 - Z^2\alpha^2$$

Now

$$-\frac{1}{2} < s = \pm\sqrt{k^2-\gamma^2}$$
$$1)\frac{1}{2} > \sqrt{k^2-\gamma^2} \geqslant \sqrt{1-Z^2\alpha^2} \in (0,99..,0,64...)\ (\text{typical } Z < 105)$$
$$2)\frac{1}{2} > -\sqrt{k^2-\gamma^2}$$

1) cannot be solved, 2) is OK

Back to recurence equations.

We know that F and G must be normalizable well behaved functions, so particularly for large m we have

$$
\begin{aligned}
(\bullet) &\Rightarrow ma_m - a_{m-1} + \gamma b_m - \frac{\alpha_2}{\alpha_1} b_{m-1} = 0 \\
(\bullet\,\bullet) &\Rightarrow mb_m - b_{m-1} - \gamma a_m - \frac{\alpha_1}{\alpha_2} a_{m-1} = 0 \;/\cdot\sqrt{\frac{\alpha_2}{\alpha_1}} \text{ andsubtract} \\
&\Rightarrow \left(m + \gamma\sqrt{\frac{\alpha_2}{\alpha_1}}\right) a_m + \left(\gamma - m\sqrt{\frac{\alpha_2}{\alpha_1}}\right) b_m = 0 \;/\cdot\sqrt{\alpha_1} \\
&\Rightarrow (\sqrt{\alpha_1} m + \gamma\sqrt{\alpha_2}) a_m + (\sqrt{\alpha_1}\gamma - m\sqrt{\alpha_2}) b_m = 0 \\
&\Rightarrow a_m = \frac{(m\sqrt{\alpha_2} - \gamma\sqrt{\alpha_1})}{(m\sqrt{\alpha_1} + \gamma\sqrt{\alpha_2})} b_m
\end{aligned}
$$

Plugging this to $(\bullet\,\bullet)$ we get

$$
\begin{aligned}
&mb_m - b_{m-1} - \gamma\frac{(m\sqrt{\alpha_2} - \gamma\sqrt{\alpha_1})}{(m\sqrt{\alpha_1} + \gamma\sqrt{\alpha_2})} b_m - \sqrt{\frac{\alpha_1}{\alpha_2}}\left(\frac{(m-1)\sqrt{\alpha_2} - \gamma\sqrt{\alpha_1}}{(m-1)\sqrt{\alpha_1} + \gamma\sqrt{\alpha_2}}\right) b_{m-1} = 0 \\
&\Rightarrow b_m\left[m - \gamma\frac{(m\sqrt{\alpha_2} - \gamma\sqrt{\alpha_1})}{(m\sqrt{\alpha_1} + \gamma\sqrt{\alpha_2})}\right] - b_{m-1}\left[1 + \frac{(m-1)\sqrt{\alpha_2} - \gamma\sqrt{\alpha_1}}{(m-1)\sqrt{\alpha_1} + \gamma\sqrt{\alpha_2}}\right] = 0
\end{aligned}
$$

For large m (large summation order) we have

$$
\begin{aligned}
&b_m\left[m - \gamma\sqrt{\frac{\alpha_2}{\alpha_1}}\right] - b_{m-1}[1+1] = 0 \\
&\Rightarrow b_m = \frac{2}{m - \underbrace{\gamma\sqrt{\frac{\alpha_2}{\alpha_1}}}_{const.}} b_{m-1} \sim \frac{2}{m} b_{m-1}
\end{aligned}
$$

for large m

$$
\begin{aligned}
a_m &\sim b_m\sqrt{\frac{\alpha_2}{\alpha_1}} \\
&\Rightarrow \text{ since } \frac{b_m}{b_{m-1}} = \frac{2}{m} \Rightarrow \frac{a_m}{a_{m-1}} = \frac{2}{m} \\
&\Rightarrow \sum_{m=0}^{\infty} a_m \rho^m \sim \sum_{m=0}^{\infty} \frac{1}{m!}(2\rho)^m \sim e^{2\rho} \\
&\Rightarrow \sum_{m=0}^{\infty} b_m \rho^m \sim e^{2\rho}
\end{aligned}
$$

$\Rightarrow$ if $\sum_{n=0}$ is infinite $\Rightarrow F \sim e^{\rho}$, $G \sim e^{\rho}$ $\Rightarrow$ diverges at $r \to \infty$ (if wave functions are nor normalizable)

$\Rightarrow \exists n_c$ s. t. for $\forall\, n > n_c$, $a_n;\; b_n = 0$

We know that $a_m = \sqrt{\frac{\alpha_2}{\alpha_1}} b_m$, assume that

$$
n_c = m - 1 \Rightarrow a_{n_c} = -\sqrt{\frac{\alpha_2}{\alpha_1}} b_{n_c} \tag{2.424}
$$

(both series terminate at the same n_c)

Now we know the ratio $\frac{a_{n_c}}{b_{n_c}}$, so no we are in position to use the recurrence equations to find on (**??**), (**??**) to find relations between b_{n_c} and a_{n_c} (whose ratio we know). From this we get equation for E (since E is in α_1 and α_2).

Let us multiply the first recursion relation by α_1 and the second by $\sqrt{\alpha_1\alpha_2}$, set $m = n_c$ and subtract, i.e.

$$\begin{aligned}(*)\ \alpha_1(s+n_c-k)a_{n_c} - \alpha_1 a_{n_c} - \alpha_1 a_{n_c-1} + \alpha_1\gamma b_{n_c} - \sqrt{\alpha_1\alpha_2}b_{n_c-1} &= 0\\ (**)\ \sqrt{\alpha_1\alpha_2}(s+n_c+k)b_{n_c} - \sqrt{\alpha_1\alpha_2}a_{n_c} - \sqrt{\alpha_1\alpha_2}\gamma a_{n_c} - \alpha_1 a_{n_c-1} &= 0\end{aligned}$$

(*) - (**) $\Rightarrow$

$$\begin{aligned}&\alpha_1(s+n_c-k)a_{n_c} + \sqrt{\alpha_1\alpha_2}\gamma a_{n_c} = \sqrt{\alpha_1\alpha_2}(s+n_c+k)b_{n_c} - \alpha_1\gamma b_{n_c}\\ \Leftrightarrow &a_{n_c}[\alpha_1(s+n_c-k) + \sqrt{\alpha_1\alpha_2}\gamma] = b_{n_c}[\sqrt{\alpha_1\alpha_2}(s+n_c+k-\alpha_1\gamma]\\ &[\alpha_1(s+n_c-k)+\sqrt{\alpha_1\alpha_2}\gamma] = \underbrace{[\sqrt{\alpha_1\alpha_2}(s+n_c+k)-\alpha_1\gamma]}_{-[\alpha_1(s+n_c+k)-\alpha_1\sqrt{\frac{\alpha_1}{\alpha_2}}\gamma}\left(-\sqrt{\frac{\alpha_1}{\alpha_2}}\right)\\ &2\alpha_1(s+n_c) = \gamma\left(\alpha_1\sqrt{\frac{\alpha_1}{\alpha_2}} - \sqrt{\alpha_1\alpha_2}\right)\ /\cdot\sqrt{\frac{\alpha_2}{\alpha_1}}\\ &2\sqrt{\alpha_1\alpha_2}(s+n_c) = \gamma(\alpha_1-\alpha_2)\\ &2\sqrt{m^2-E^2}(s+n^c) = \gamma[m+E-(m-E)] = 2\gamma E\\ \Rightarrow &\sqrt{m^2-E^2}(s+n_c) = E\gamma = \text{quadratic eq.}\end{aligned}$$

$\Rightarrow$

negative energies are occupied
$\alpha = \hat{e}^2 = \frac{e^2}{4\pi\epsilon_0} = \frac{1}{137}$

$$E = \frac{m}{\sqrt{1+\frac{\gamma^2}{(s+n_c)^2}}} = \frac{m}{\sqrt{1+\frac{Z^2\alpha^2}{(n_c\sqrt{k^2-\gamma^2})^2}}} = \frac{m}{\sqrt{1+\frac{Z^2\alpha^2}{(n_c\sqrt{(j+1/2)^2-Z^2\alpha^2})^2}}} \tag{2.425}$$

Note that E depends only on n_c and $j+1/2 = |k|$. In order to compare (2.425) with the corresponding expression obtained in the Schrödinger theory, we define $n \equiv n_c + \underbrace{(j+1/2)}_{\text{we know it is integer}} = n_c + |k|$.

Since the minimum value of $n_c = 0$, we have

$$0 \leqslant n_c = n - (j+1/2) \Rightarrow n \geqslant (j+1/2) = |k|$$

which is at least unity. Expanding (2.425) we get

$$E = m\left[1 - \frac{1}{2}\frac{(Z\alpha)^2}{n^2} - \frac{1}{2}\frac{(Z\alpha)^4}{n^3}\underbrace{\left(\frac{1}{j+1/2} - \frac{3}{4n} - ...\right)}_{>0}\right]$$

since

$$\frac{1}{2}\alpha^2 mc^2 = \left\{\hat{e}^2 = \alpha = \frac{e^2}{4\pi\epsilon_0}\right\} = \frac{1}{2}m\frac{e^4}{(4\pi\epsilon_0)^2} = \frac{e^4 m}{32\pi^2\epsilon_0^2} = \frac{e^2}{8\pi\underbrace{a_B}_{\text{Bohrv radius } = \frac{4\pi\epsilon_0}{m_e e^2} = \frac{1}{m_e\alpha}}}$$

$j = n - n_c - 1/2$
$\max J = n - \min(n_c) - 1/2 = n - 1/2$
$n = 1, 2, ..; j = 1/2, 3/2, ...$

We see that "n" is indeed identical with the familiar "principal quantum number" of non-rel. QM.

Note:

A catastrophe occures in the original formula for E when $Z = 137$ $\Rightarrow \sqrt{(j+1/2) - Z^2\alpha^2}$ becomes imaginar.

Note:

First dominant contribution comes from

$$-\frac{1}{2}\frac{mc^2(Z\alpha)^2}{n^2} = \left\{\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}\right\} = -\frac{1}{2}\frac{mc^2 Z^2 e^4}{n^2(4\pi\epsilon_0\hbar c)^2} = -\frac{Z^2}{2n^2}\frac{e^2}{4\pi}\underbrace{\frac{e^2 m}{4\pi\epsilon_0\hbar^2}}_{\frac{1}{a_B}} = -\frac{Z^2 e^2}{8^2 a_B}$$

a_B - most probable distance of electrons from nucleus in its ground state

$\Rightarrow$ Higher order corrections are due to spin and relativity removes the energy degeneracy in an observable way.

For a give n, higher j-states are at higher energy levels.

In the Dirac theory each state of a hydrogen atom can be completely characterized by n_c (or n), $\underbrace{\kappa}_{l \text{ and } J}$ and j_3 (only on the level of statevectors).

One can translate this classification scheme into the more familiar spectroscopic notation.

It should be stressed that even though $\boldsymbol{L}^2$ is not "good" in the relativistic theory, it is customary to use the notation

$$nx_j$$

n - principal quantum number $n = 1, 2, \ldots$
x - orbital quantum number $l_+ = (0, 1, 2, \ldots) = (s, p, d, \ldots)$
j - total angular momenta $j = 1/2;\ 3/2;\ 5/2; \ldots$

Table 2.4

	l_+
$\kappa = j + 1/2$	$j + 1/2$
$\kappa = -(j + 1/2)$	$j - 1/2$

e. g. $2p_{1/2}$

Note:

The orbital angular momentum l_+ of the upper two-component wave function (in non-rel. case will correspond to Schrödinger-Pauli theory) determines the orbital angular momentum in the spectroscopic language.

Table 2.5

n	$n_c = n - \lvert\kappa\rvert \geqslant 0$	$\kappa = \pm(j + 1/2)$	spectroscopic notation
1	0	-1 $j = 1/2;\ l_+ = 0$	$1s_{1/2}$
2	1	-1 $j = 1/2;\ l_+ = 0$	$2s_{1/2}$
2	1	1 $j = 1/2;\ l_+ = 1$	$2p_{1/2}$
2	0	-2 $j = 3/2;\ l_+ = 1$	$2p_{3/2}$
3	2	-1 $j = 1/2;\ l_+ = 0$	$3s_{1/2}$
3	2	1 $j = 1/2;\ l_+ = 1$	$3p_{1/2}$
3	1	-2 $j = 3/2;\ l_+ = 1$	$3p_{3/2}$
3	1	2 $j = 3/2;\ l_+ = 2$	$3d_{3/2}$
3	0	-3 $j = 5/2;\ l_+ = 2$	$3d_{5/2}$

$n = n_c + j + 1/2 \Rightarrow n - j - 1/2 \geqslant 0$
$\Rightarrow j - 1/2$
$\max J = l_+ + 1/2 \leqslant n - 1/2$
$\Rightarrow l_+ \leqslant n - 1$

Fine structure of the spectrum is a new phenomenon w.r.t. s.e. It is the difference between energy levels of different j but identical n.

(*picture*)

For $n = 2$, $Z = 1$

$$E(2p_{3/2} - 2p_{1/2}) \simeq -\frac{1}{2}\frac{\alpha^4 mc^2}{8}\left[\frac{1}{2} - 1\right] = \frac{\alpha^4 mc^2}{32} = 4,53 \cdot 10^{-5}\ \text{eV}$$

2.16 Relativistic wave equations

Apart from K-G wave equation (for spin - 0 part.) and Dirac's wave equation (for spin - 1/2 part.), there exist a number of higher-spin wave equations.

Examples include:

- "Maxwell equation" (for spin - 1 part.)

$$\partial_m \partial^m A^\nu = e\overline{\Psi}\gamma^\nu\Psi \quad (\text{in Lorentz gauge } \partial_m A^m = 0)$$

also (without source)

$$\begin{array}{l} i\frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{i}\boldsymbol{S}\cdot\boldsymbol{\Delta}(i\boldsymbol{B}) \\ i\frac{\partial i\boldsymbol{B}}{\partial t} = \frac{1}{i}\boldsymbol{S}\cdot\boldsymbol{\Delta}(\boldsymbol{E}) \end{array} \longleftrightarrow \underbrace{\begin{array}{l} i\frac{\partial \varphi}{\partial t} = \frac{1}{i}\boldsymbol{\sigma}\cdot\boldsymbol{\Delta}(\Psi) \\ i\frac{\partial \Psi}{\partial t} = \frac{1}{i}\boldsymbol{\sigma}\cdot\boldsymbol{\Delta}(\varphi) \end{array}}_{\text{Weyl equation for massless spin}-1/2\text{ part.}}$$

$$(\boldsymbol{E}i - i\boldsymbol{B}) \Longleftrightarrow (\varphi, \Psi)$$

(spin matrices S_i play for the spin 1 electromagnetic field the same role as Pauli matrices σ_i for spin 1/2)

- Proca equation (for massive spin - 1 part.)

$$\partial_m(\partial^m A^\nu - \partial^\nu A^m) + mA^\nu = 0$$

for example ($W^\pm$, Z bosons)

- Rarita-Schwinger equation (for massive spin - 3/2 part.)

$$\epsilon^{\mu\nu\rho\sigma}\gamma^5\gamma_\nu\partial_\rho\Psi_\sigma + m\Psi^m = 0$$

- Bargmann-Winger equation (for massive arbitrary spin free particle)
  - quite complicated system of equations

Wave equations of these types have a number of conceptual difficulties.

- No simple way to include multi-particle interactions.
- Strictly single-particle description does not allows (is not applicable for) unstable particles or resonances.
- Various paradoxes: "Zitterbewegung", klein paradox, problematic probabilistic interpretations.
- Single-particle picture is not tenable beyond energies allowing pair formation.
- No fundamental particles observed beyond spin 1. (SUSY predict e.g. 3/2-spin gravition).

Good reasons for abandoning wave equations.

Quantum Field Theory 3

3.1 Why Quantum Field Theory?

A1. — First is because the combination of quantum mechanics and special relativity implies that particle number is not conserved. Relativity necessarily brings in the possibility of conversion of mass into energy and vice versa, i.e., the creation and annihilation of particles. E.g., β decay of the neutron via $n \to p + e^- + \bar{\nu}_e$ or positron-electron annihilation $e^+e^- \to 2\gamma$. There are also situations when the number of particles of given species is not conserved, even though the number of particles of all types taken together is conserved.

The creation of particles is impossible to avoid whenever one tries to locate a particle of mass m within its Compton wavelength. Indeed, from Heisenberg uncertainty relation we find that

$$\begin{aligned}
\sigma_E^2\sigma_x^2 &\geq \frac{1}{4}\left|\langle[\hat{H}, x]\rangle_\psi\right| = \frac{1}{4}\left|\left\langle[\sqrt{\boldsymbol{p}_x^2c^2 + m^2c^4}, x]\right\rangle_\psi\right| \\
&= \frac{\hbar^2}{4}\left|\left\langle p_x c^2/\sqrt{p_x c^2 c^2 + m^2 c^4}\right\rangle_\psi\right|^2 \\
&\sim \frac{\hbar^2}{4}\left|c + O\left(\left\langle\frac{m^2c^4}{p_x^2c^2}\right\rangle_\psi\right)c^2\right|^2 .
\end{aligned} \tag{3.1}$$

This leads to the well known relationship $\Delta E \Delta x \geq \frac{\hbar c}{2}$. If we assume that $\Delta x \sim \lambda_C = \frac{\hbar}{mc}$, then we have $\Delta E \geq mc^2$. Therefore, in a relativistic theory, the fluctuations of the energy are enough to allow the creation of particles out of the vacuum. In the case of spin $\frac{1}{2}$ particle, the Dirac sea picture shows clearly how, when the energy fluctuations are of order m, electrons from the Dirac sea can be excited to positive energy states, thus creating electron-positron pairs.

In order to discuss such processes, the usual formalism of many-body quantum mechanics with wave functions of fixed number of particles, has to be augmented by including the possibility of creation and annihilation of particles via interaction.

A 2. — Ordinary (non-relativistic) point-particle QM can deal with the quantum description of a many-body system in terms of many body wave functions. This is important, e.g., in atomic, molecular or condense matter physics. Similar generalization for relativistic particles would be desirable. Problem with this generalization, however, starts already at classical level. There does not exist any generalization to relativistic invariant interacting many-body theory — not even for 2 interacting particles. This is known as Leutwyler's no-interaction theorem (Leutwyler 1965, Makunda 1984):

Any finite number of point particles cannot interact in such a way that the principles of special relativity are respected, i.e. that the system provides a representation of the Poincare algebra. Accordingly, classical relativistic point particles are necessarily free, as a consequence of Poincare invariance.

Note: The only exception are two particles in one spatial dimension confined to each other by a linearly rising potential.

In contrast to point particles, strings can interact relativistically in higher dimensions, without violation of Leutwyler's non-interaction theorem. Hence, it is not surprising that particle physics is based on Quantum Field Theory (i.e., infinite number of degrees of freedom) rather than on relativistic point particle quantum mechanics.

A 3. — We know of classical field that is fundamental in physics — the electromagnetic field. Analyses of Bohr and Rosenfeld show that there are difficulties in having a quantum description of various charged particle phenomena (such as those that occur in atomic physics) while retaining a classical description of the electromagnetic field. One has to quantize the electromagnetic field (e.g., to get Lamb shift correctly); this is independent of any many-particle interpretation that might emerge from quantization.

A 4. — Because all particles of the same type are the same. What we mean by this is that two electrons are identical in every way, regardless of where they came from and what they have been through. The same is true of every other fundamental particle. Let me illustrate this through a rather prosaic story. Suppose we capture a proton from a cosmic ray which we identify as coming from a supernova lying 8 billion light years away. We compare this proton with one freshly created in a particle accelerator here on Earth. And the two are exactly the same! How is this possible? Why are not there errors in proton production? How can two objects, manufactured so far apart in space and time, be identical in all respects? One explanation that might be offered is that there's a sea of proton "stuff" filling the universe and when we make a proton we somehow dip our hand into this stuff and from it mould a proton. Then it's not surprising that protons produced in different parts of the universe are identical: they're made of the same stuff. It turns out that this is roughly what happens. The "stuff" is the proton field or, if you look closely enough, the quark field.

There are two complementary approaches to field theory.

- We can postulate fields as the basic dynamical variables and show that the result can be interpreted in many-body terms.
- One can start with point-particles as the basic objects of interest and derive or construct field operator as an efficient way of organizing the many-particle states.

We will work with the first approach as it brings us faster to the point. The second approach is often a starting point in non-relativistic field theory in condense matter physics. Approach a) is known as Quantum Field Theory (QFT) or Theory of Quantized Fields, b) is known as Second Quantization.

3.2 Some useful background from quantum mechanics

Let us first recall the familiar path to the quantization of a classical dynamical system in particle mechanics. For the purpose of illustration, consider a 1-D motion of a particle in a conservative potential. Let q be the (generalized) coordinate of the particle, $\dot{q} = \frac{\mathrm{d}q}{\mathrm{d}t}$ the velocity, and $L(q, \dot{q})$ the Lagrangian. According to Hamilton's principle, the dynamics of the particle is determined by the condition

$$\delta S[q] = \delta \int_{t_1}^{t_2} \mathrm{d}t L(q, \dot{q}) = 0, \tag{3.2}$$

This is analogous to the case when $\mathrm{d}f(x) = 0$ implies that $f(x_0) = f(x_0 + \mathrm{d}x)$ to first order in $\mathrm{d}x$.

this equation determines the actual physical path $q(t)$ from (q_1, t_1) to (q_2, t_2). Action is stationary around classical trajectory, i.e. small variations from classical path, $q(t) \to q(t) + \delta q(t)$, leaving the action unchanged to the first order in the variation.

Hamilton's principle gives us the well known Euler-Lagrange equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0. \tag{3.3}$$

In order to carry out the formal quantization based on this equation, we rewrite it in the Hamiltonian form, by defining the momentum p conjugate to q as

$$p = \frac{\partial L}{\partial \dot{q}}, \tag{3.4}$$

and introducing the Hamiltonian by the Legendre transformation

$$H(p, q) = p\dot{q} - L(q, \dot{q}). \tag{3.5}$$

Note, that H is not dependent on $\dot{q}$ since

$$\mathrm{d}H = (\mathrm{d}p)\dot{q} + p\mathrm{d}\dot{q} - \frac{\partial L}{\partial q}\mathrm{d}q - \frac{\partial L}{\partial \dot{q}}\mathrm{d}\dot{q}. \tag{3.6}$$

In terms of H, the Euler-Lagrange equation becomes

$$\begin{aligned} \{q, H\}_{PB} &= \frac{\partial H}{\partial q} = \dot{q}, \\ \{p, H\}_{PB} &= -\frac{\partial H}{\partial q} = \dot{p}, \end{aligned} \tag{3.7}$$

where $\{\cdot, \cdot\}_{PB}$ means a Poisson bracket. To quantize (3.7) we let q become a Hermitan operator in a Hilbert space and replace p by $-\frac{\partial}{\partial q}$, so that the conjugate momentum and coordinate satisfy a commutation relation

$$[p, q] = -i\hbar, \ [q, p] = i\hbar, \tag{3.8}$$

corresponding to the classical Poisson bracket $\{p, q\}_{PB} = -1$ or $\{q, p\}_{PB} = 1$. The dynamics of particle is contained in the Schrödinger equation

$$\boldsymbol{H}(p, q)\psi(t) = i\frac{\partial \psi(t)}{\partial t}, \tag{3.9}$$

where $\psi(t)$ is a wave function or state vector in the Hilbert space. In this formulation all time dependence is carried by ψ while p and q are not time dependent. This approach is called the Schrödinger picture. Alternatively, we may transfer whole time dependence to the operators $q(t)$ and $p(t)$, ψ is then time independent. This is called Heisenberg picture. There also exist and intermediate pictures, like Dirac picture or thermo-field dynamics picture, that will be discussed later on. Both pictures are equivalent as they can be connected via unitarity transformation. From Schrödinger equation it follows that

We will assume $\hbar = 1$ further on.

$$\psi_S(t) = e^{-i\boldsymbol{H}t}\psi_S(0) = e^{-i\boldsymbol{H}t}\psi_H. \tag{3.10}$$

Here the value of $\psi_S(t = 0)$ is set to coincide with ψ_H. Similarly, operators are connected according to

$$\mathbf{O}_H(t) = e^{i\boldsymbol{H}t}\mathbf{O}_S e^{-i\boldsymbol{H}t}. \tag{3.11}$$

The unitarity transformations are constructed such, that matrix elements of all observables are identical. Solution of a dynamical problem in quantum mechanics consists in finding, at a later time t, matrix elements of operators which represent physical observables, provided we know the matrix elements at some initial time (say $t = 0$). In Schrödinger picture this is done by solving Schrödinger equation.

In Heisenberg picture, one solves the equation of motion for the Heisenberg operator $\mathbf{O}_H(t)$

$$\frac{\mathrm{d}\boldsymbol{O}_H(t)}{\mathrm{d}t} = i\,[\boldsymbol{H}, \boldsymbol{O}_H(t)]\,. \tag{3.12}$$

As long as we deal with energy eigenfunctions in **non-relativistic** theory, there is a little practical difference, as

$$\boldsymbol{H}_H(t) = \boldsymbol{H}_S \equiv \boldsymbol{H}, \tag{3.13}$$

and in the absence of external time-varying forces we have $\frac{\mathrm{d}\boldsymbol{H}}{\mathrm{d}t} = 0$. For energy eigenfunctions, the Schrödinger wave function is $\psi_H(q,t) = e^{-i\omega_n t}u_n(q)$. In relativistic field theory, we will see that the Heisenberg picture is more convenient, since the explicit representation of the state vector ψ is considerably more complicated than in the non-relativistic case (this ψ is a solution of the so-called functional Schrödinger equation), and the dynamics of operators is easier to prescribe than the dynamics of ψ.

Also, Lorentz invariance can be more readily implemented in the Heisenberg picture, which puts the time together with space coordinates in the field operators. In the Heisenberg picture it follows that the CCR retain the form

$$[q(t), p(t)] = i. \tag{3.14}$$

For an arbitrary t we have

$$\hat{p}(t) = -i\frac{\partial}{\partial q(t)},\ \hat{q}(t) = q(t) \text{ in } q\text{-representation,} \tag{3.15}$$

$$\hat{p}(t) = p(t),\ \hat{q}(t) = i\frac{\partial}{\partial p(t)} \text{ in } p\text{-representation.} \tag{3.16}$$

Equations of motion are of Hamilton-Heisenberg form

$$\frac{\mathrm{d}p(t)}{\mathrm{d}t} = i\,[\boldsymbol{H}, p(t)] \text{ in } q\text{-representation,} \tag{3.17}$$

$$\frac{\mathrm{d}q(t)}{\mathrm{d}t} = i\,[\boldsymbol{H}, q(t)] \text{ in } p\text{-representation.} \tag{3.18}$$

Sub-index $\boldsymbol{H}$ and hat over operators will be suppressed further on.

To completely determine the dynamical problem, we must specify the matrix elements of p and q at the initial time. Let us illustrate that on the following problem. Consider Lagrangian of form

$$L = \frac{1}{2}m\dot{q}^2 - \frac{\omega^2}{2}mq^2 \text{ (set } m = 1 \text{ for simplicity).} \tag{3.19}$$

Then the action is given by

$$S = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{1}{2}m\dot{q}^2 - \frac{\omega^2}{2}mq^2\right] = \int_{t_1}^{t_2} \mathrm{d}t \frac{1}{2}m\left[\dot{q}^2 - \omega^2 q^2\right]. \tag{3.20}$$

Requiring $\delta S = 0$ implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0, \tag{3.21}$$

which reduces to the

$$\ddot{q} + \omega^2 q = 0. \tag{3.22}$$

Which is equation in a configuration space. We get p from

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} = \dot{q}. \tag{3.23}$$

Constructing Hamiltonian from it we get

$$H(p,q) = p\dot{q} - L(q,\dot{q}) = p^2 - \frac{1}{2}p^2 + \frac{\omega^2}{2}q^2 = \frac{1}{2}(p^2 + \omega^2 q^2). \tag{3.24}$$

Which is Hamiltonian of harmonic oscillator. Hamilton equations of motion

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \{q, H\} = \frac{\partial H}{\partial p} = p, \tag{3.25}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \{p, H\} = -\frac{\partial H}{\partial q} = -\omega^2 q, \tag{3.26}$$

will yield us the same equation of motion as the one from Euler-Lagrange equations. We can transfer to quantum mechanics by making change $\{\cdot,\cdot\} \to -\frac{i}{\hbar}[\cdot,\cdot]$, which implies that

$\{f, g\} = d \to [\boldsymbol{O}_f, \boldsymbol{O}_g] = i\hbar\boldsymbol{O}_d$.

$$\boldsymbol{O}_{\{f,g\}} = -\frac{i}{\hbar}\left[\boldsymbol{O}_f, \boldsymbol{O}_g\right]. \tag{3.27}$$

This leads to the

$$\dot{q} = \{q, H\} = p \to \left[\hat{q}, \hat{\boldsymbol{H}}\right] = i\hbar\hat{p}, \tag{3.28}$$

$$\dot{p} = \{p, H\} = -\omega^2 q \to \left[\hat{p}, \hat{\boldsymbol{H}}\right] = i\hbar\omega^2 \hat{q}. \tag{3.29}$$

Further on we again proceed without hats.

This reduces to the operator equation

$$\ddot{\hat{q}} = -\omega^2 \hat{q}. \tag{3.30}$$

To solve for the coordinates we define

$$a(t) = \frac{\omega q(t) + ip(t)}{\sqrt{2\omega}} \text{ Annihilation operator} \tag{3.31}$$

$$a^\dagger(t) = \frac{\omega q(t) - ip(t)}{\sqrt{2\omega}} \text{ Creation operator} \tag{3.32}$$

By simple algebra, we can also show that

$$q(t) = \frac{1}{\sqrt{2\omega}} \left(a^\dagger(t) + a(t)\right), \tag{3.33}$$

$$p(t) = i\sqrt{\frac{\omega}{2}} \left(a^\dagger(t) - a(t)\right), \tag{3.34}$$

The reason for that is the fact, that if a and $a^\dagger$ satisfy relationship $\left[a, a^\dagger\right] = \lambda \in \mathbb{R}^+$, then eigenvalues of $a^\dagger a$ are $0, \lambda, 2\lambda, 3\lambda, \ldots$. We can see that

$$\begin{aligned} \dot{a}^\dagger(t) &= -i\left[a^\dagger(t), \boldsymbol{H}\right] = -i\left[\frac{\omega q(t) - ip(t)}{\sqrt{2\omega}}, \boldsymbol{H}\right] \\ &= \frac{\omega p(t) + i\omega^2 q(t)}{\sqrt{2\omega}} = i\omega a^\dagger(t). \end{aligned} \tag{3.35}$$

This implies that

$$a^\dagger(t) = a^\dagger(0)e^{i\omega t}, \tag{3.36}$$

and similarly that

$$a(t) = a(0)e^{-i\omega t}. \tag{3.37}$$

This means that we can rewrite $q(t)$ as

$$q(t) = \frac{1}{\sqrt{2\omega}}(a^\dagger(0)e^{i\omega t} + a(0)e^{-i\omega t}), \tag{3.38}$$

and similarly for $p(t)$. In terms of $a(t)$ and $a^\dagger(t)$ the Hamiltonian reads

$$\begin{aligned} \boldsymbol{H} &= \frac{1}{2}\omega\left(a^\dagger(t)a(t) + a(t)a^\dagger(t)\right) \\ &= \frac{1}{2}\omega\left(a^\dagger(0)a(0) + a(0)a^\dagger(0)\right) \\ &= \omega\left(a^\dagger(0)a(0) + \frac{1}{2}\right) \end{aligned} \tag{3.39}$$

Suppose now that

$$\boldsymbol{H}\left|n\right\rangle = \omega_n \left|n\right\rangle. \tag{3.40}$$

Applying $a^\dagger(0)$ to $|n\rangle$ we get

$$\begin{aligned} H(a^\dagger(0)\,|n\rangle) &= \left(\left[H, a^\dagger(0)\right] + a^\dagger(0)H\right)|n\rangle \\ &= \left(\omega a^\dagger(0) + a^\dagger(0)\omega_n\right)|n\rangle \\ &= (\omega + \omega_n)a^\dagger(0)\,|n\rangle\,. \end{aligned} \tag{3.41}$$

And similarly

$$H(a(0)\,|n\rangle) = (\omega_n - \omega)a(0)\,|n\rangle\,. \tag{3.42}$$

More generaly we have

$$H(a^m\,|n\rangle) = (\omega_n - m\omega)(a^m\,|n\rangle), \tag{3.43}$$

$$H(a^{\dagger,m}\,|n\rangle) = (\omega_n + m\omega)(a^{\dagger,m}\,|n\rangle). \tag{3.44}$$

For energy to have a lower bound (Hamiltonian is positive definite), there must be a state $|0\rangle$, such that $a\,|0\rangle = 0$. In this case,

$$H\,|0\rangle = \omega\left(a^\dagger a + \frac{1}{2}\right)|0\rangle = \frac{1}{2}\omega\,|0\rangle\,. \tag{3.45}$$

I.e. vacuum state $|0\rangle$ is the **lowest** energy state.

Cute note

$$\begin{aligned} 2\,\langle n|H|n\rangle = 2\omega_n &= \langle n|p^2|n\rangle + \omega^2\,\langle n|q^2|n\rangle \\ &= ||p\,|n\rangle\,||^2 + ||\omega q\,|n\rangle\,||^2 \geq 2||p\,|n\rangle\,|||\omega q\,|n\rangle\,||. \end{aligned} \tag{3.46}$$

Where last inequality follows from triangle inequality.

Virial theorem implies that $\langle n|p|n\rangle = \langle n|q|n\rangle = 0$. The latter is a direct consequence of the fact that for any operator A and any eigenstate $|n\rangle$ of H we have

$$\begin{aligned} \langle n|\,[A, H]\,|n\rangle &= \langle n|AH|n\rangle - \langle n|HA|n\rangle \\ &= \omega_n\,\langle n|A|n\rangle - \omega_n\,\langle n|A|n\rangle = 0. \end{aligned} \tag{3.47}$$

Thus if $A = q$

$$\langle n|\,[q, H]\,|n\rangle = \langle n|ip|n\rangle = 0, \tag{3.48}$$

or if $A = p$

$$\langle n|\,[p, H]\,|n\rangle = -i\omega\,\langle n|q|n\rangle = 0. \tag{3.49}$$

Hence

$$2\,\langle n|H|n\rangle \geq 2||p\,|n\rangle\,|||\omega q\,|n\rangle\,|| \geq 2\omega\Delta p\Delta q \geq \hbar\omega. \tag{3.50}$$

Uncertainty relations prohibit $\langle n|H|n\rangle = \omega_n$ in any energy eigenstate to be smaller that $\frac{\hbar}{2}\omega$. State $\omega_n = \frac{\hbar}{2}\omega$ represent the so-called zero mode of H (in this specific case we call them zero-mode fluctuation or ground state fluctuation). Note, that no other states apart from that $|n\rangle$ with eigenvalues $\omega_n = \hbar(\frac{1}{2} + n)\omega$ is allowed (on account of the theorem $\left[a, a^\dagger\right] = \lambda$, shown above). By applying creation operator on vacuum state we get

$$a^\dagger |0\rangle \propto |1\rangle\,, \tag{3.51}$$

and generally

$$(a^\dagger)^n |0\rangle \propto |n\rangle\,. \tag{3.52}$$

Using the fact that

$$\boldsymbol{H}\,|n\rangle = \left(n + \frac{1}{2}\right)\omega\,|n\rangle\,,\ n = 0, 1, \ldots. \tag{3.53}$$

After normalization of states to the $\langle n|m\rangle = \delta_{nm}$, we get

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle\,. \tag{3.54}$$

This can be easily proven by induction.

It is straightforward to generalize this procedure to n degrees of freedom. We introduce n Hermitan operators $q_i(t)$, $i = 1, \ldots, n$ in the Heisenberg picture and n conjugate momenta $p_i(t)$. The dynamics is again given by the $2n$ classical equations of motion

$$-\frac{\partial \boldsymbol{H}}{\partial q_i} = \dot{p}_i,\ \ \frac{\partial \boldsymbol{H}}{\partial p_i} = \dot{q}_i,\ \ i = 1, \ldots, n. \tag{3.55}$$

Again, our goal is to find the matrix elements of p_i and q_i at an initial time, say $t = 0$ with the restriction that

$$\left[p_i(0), q_j(0)\right] = -i\delta_{ij}, \tag{3.56}$$
$$\left[p_i(0), p_j(0)\right] = 0, \tag{3.57}$$
$$\left[q_i(0), q_j(0)\right] = 0. \tag{3.58}$$

And we can write Heisenberg equations of motion as

$$\dot{p}_i(t) = i\,[\boldsymbol{H}, p_i(t)]\,, \tag{3.59}$$
$$\dot{q}_i(t) = i\,[\boldsymbol{H}, q_i(t)]\,. \tag{3.60}$$

3.3 Fields

A **field** is a quantity defined at every point of space and time $(\boldsymbol{x}, t)$. While classical particle mechanics deals with a finite number of generalized coordinates $q_i(t)$, indexed by a label "i", in field theory we are interested in the dynamics of fields $\varphi_i(\boldsymbol{x}, t)$, where both "$i$" and $\boldsymbol{x}$ are considered as labels. We are thus dealing with a system with an infinite (uncountably infinite) number of degrees of freedom - at least one for each point $\boldsymbol{x}$ in space. Notice, that the concept of position has been relegated from a dynamical variable in particle mechanics (so called wave mechanics or first quantization) to a mere label in field theory. Further we will review example of a well known field.

- **Electromagnetic field**
  $\boldsymbol{E}(\boldsymbol{x}, t)$ and $\boldsymbol{B}(\boldsymbol{x}, t)$, both of these fields are 3 spatial vectors. In covariant treatment of electromagnetism $\boldsymbol{E}, \boldsymbol{B} \to A^\mu(\boldsymbol{x}, t) = (\phi, \boldsymbol{A})$

($\mu = 1, \dots, 4$) so A^μ is a vector in spacetime. Here we can recover $\boldsymbol{E}$ as

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \; \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}. \tag{3.61}$$

We again start with **Lagrangian** (it is a convenient tool for a covariant treatment - next semester and it also provides a simple passage to Hamiltonian, i.e. canonical formalism). In the Lagrangian formalism, the dynamics is governed by a Lagrangian, which is a function of $\phi_i(\boldsymbol{x}, t)$, $\dot{\phi}(\boldsymbol{x}, t)$ and $\boldsymbol{\nabla}\phi_i(\boldsymbol{x}, t)$. We change our Lagrangian (L) to Lagrangian density ($\mathcal{L}$), i.e.

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \int \mathrm{d}^3\boldsymbol{x}\, \mathcal{L}(\phi_i(\boldsymbol{x}, t), \partial_\mu \phi_i(\boldsymbol{x}, t)). \tag{3.62}$$

In principle, we could consider also higher derivative terms (non-local interactions), but in all systems, studied in this course, the Lagrangian is of the form given above. The action is

$$S = \int_{t_1}^{t_2} \mathrm{d}t L = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3\boldsymbol{x}\, \mathcal{L}(\phi_i, \partial_\mu \phi_i) = \int \mathrm{d}^4 x \mathcal{L}(\phi_i, \partial_\mu \phi_i). \tag{3.63}$$

In particle mechanics L depends on q_i and $\dot{q}_i$, but not $\ddot{q}_i$. In field theory we similarly restrict the Lagrangian $\mathcal{L}$ on ϕ_i and $\dot{\phi}_i$. In principle, there is nothing to stop $\mathcal{L}$ from depending on $\boldsymbol{\nabla}\phi$, $\boldsymbol{\nabla}^2\phi$, $\boldsymbol{\nabla}^3\phi, \dots$ In cases when we require Lorentz invariance, we will consider only dependence of $\mathcal{L}$ on $\boldsymbol{\nabla}\phi$ (this is not needed in non-relativistic context). Also we will not consider $\mathcal{L}$ explicitly dependent on x^μ (no external fields).

We can obtain equation of motion via least action principle as

$$\begin{aligned} \delta S &= \int \mathrm{d}^4 x \left[\frac{\partial \mathcal{L}}{\partial \phi_i} \delta\phi_i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \delta(\partial_\mu \phi_i) \right] \\ &\overset{p.p.}{=} \int \mathrm{d}^4 \left[\frac{\partial \mathcal{L}}{\partial \phi_i} \delta\phi_i - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) \delta\phi_i + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \delta\phi_i \right) \right]. \end{aligned} \tag{3.64}$$

This leads us to the Euler-Lagrange equations of motion for field

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0. \tag{3.65}$$

As an example, consider the following Lagrangian

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} m^2 \phi^2 \\ &= \frac{1}{2} \dot{\phi}^2 - \frac{1}{2} (\boldsymbol{\nabla}\phi)^2 - \frac{1}{2} m^2 \phi^2. \end{aligned} \tag{3.66}$$

Using previously derived Euler-Lagrange equation we arrive at the Klein-Gordon equation.

As another example, consider Lagrangian with a complex field that is linear in time derivative (rather than quadratic)

$$\mathcal{L} = \frac{i}{2}\left(\psi^*\dot{\psi} - \dot{\psi}^*\psi\right) - \nabla\psi^*\nabla\psi - m\psi^*\psi. \tag{3.67}$$

This Lagrangian yields us the equation of motion in the form

$$i\frac{\partial\psi}{\partial t} = -\nabla^2\psi + m\psi. \tag{3.68}$$

This equation looks like Schrödinger equation, but it is not. It's interpretation is very different and field ψ is a classical fied with no probabilistic interpretation of wave function. As an exercise, try to show that

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A^\mu)^2 \tag{3.69}$$

reduces to the Maxwell equations.

3.4 Quantization of Scalar Field

We will be particularly interested in relativistic field theories. We have seen that for relativistic (scalar) field theories, equations of motion, i.e. Euler-Lagrange equations read

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right) - \frac{\partial\mathcal{L}}{\partial\phi_i} = 0. \tag{3.70}$$

Should this equation be covariant under Lorentz transformations, $\mathcal{L}$ must transform as scalar density, i.e.

$$\mathcal{L}(x) \equiv \mathcal{L}(\phi_i(x), \partial_\mu\phi_i(x)) \xrightarrow{L} \mathcal{L}_{\boldsymbol{L}}(x) = |\det\boldsymbol{L}|\,\mathcal{L}(\boldsymbol{L}^{-1}x). \tag{3.71}$$

Let us construct the simplest free scalar theory with maximally second derivative in equation of motion.

The simplest $\mathcal{L}$ that is a scalar density and bounded from below potential energy is

$$\mathcal{L} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi(x)\partial_\nu\phi(x) - \frac{1}{2}m^2\phi^2(x). \tag{3.72}$$

Proof of the fact that $\mathcal{L}$ is a scalar density is quite simple. First we realize that

$$\begin{aligned} \phi(x) &\xrightarrow{\boldsymbol{L}} \phi_{\boldsymbol{L}} = \phi(\boldsymbol{L}^{-1}x) \\ \partial_\mu\phi(x) &\xrightarrow{\boldsymbol{L}} \left(\boldsymbol{L}^{-1}\right)^\nu{}_\mu\,\partial_\nu\phi(\boldsymbol{L}^{-1}x)\,. \end{aligned} \tag{3.73}$$

Thus

$$\begin{aligned} \partial_\mu\phi(x)\partial_\nu\phi(x)\eta^{\mu\nu} &\xrightarrow{\boldsymbol{L}} \left(\boldsymbol{L}^{-1}\right)^\alpha{}_\mu\,\partial_\alpha\phi(\boldsymbol{L}^{-1}x)\left(\boldsymbol{L}^{-1}\right)^\beta{}_\nu\,\partial_\beta\phi(\boldsymbol{L}^{-1}x)\boldsymbol{L}^\mu{}_\gamma\eta^{\gamma\delta}\boldsymbol{L}^\nu{}_\delta \\ &= \partial_\gamma\phi(\boldsymbol{L}^{-1}x)\partial_\delta\phi(\boldsymbol{L}^{-1}x)\eta^{\gamma\delta}. \end{aligned} \tag{3.74}$$

The potential term transforms in the same way, with $\phi^2(x) \xrightarrow{\boldsymbol{L}} \phi^2(\boldsymbol{L}^{-1}x)$. Since $|\det \boldsymbol{L}| = 1$, it is indeed scalar density. Action is also invariant under Lorentz transformation.

As before, the link between the Lagrangian formalism and canonical quantum theory is via Hamiltonian formalism (of field theory). We start by defining momentum $\pi(x)$ conjugated to $\phi(x)$ as

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} = \frac{\delta S[\phi]}{\delta \dot{\phi}(x)} . \tag{3.75}$$

This should not be confused with total momentum, which will be defined shortly. The **Hamiltonian density** is given by

$$\mathcal{H}(x) = \pi(x)\dot{\phi}(x) - \mathcal{L}(x) , \tag{3.76}$$

where as in classical mechanics we eliminate $\dot{\phi}(x)$ in favour of $\pi(x)$ everywhere in $\mathcal{H}(x)$. The **Hamiltonian** is the simply

$$H = \int \mathrm{d}^3\boldsymbol{x} \mathcal{H}(x) . \tag{3.77}$$

Quantization starts by identifying commutators via Poisson brackets. We know that

$$-\frac{i}{\hbar}\left[O_f, O_g\right] = O_{\{f,g\}} . \tag{3.78}$$

So $[\phi(\boldsymbol{x}), \pi(\boldsymbol{x}')] = i\delta(\boldsymbol{x} - \boldsymbol{x}')$. In the second step, we construct Hamiltonian

$$H = \int \mathrm{d}^3\boldsymbol{x} \left[\pi(x)\dot{\phi}(x) - \mathcal{L}(x)\right] \tag{3.79}$$

$$= \int \mathrm{d}^3\boldsymbol{x}\, \pi(x)\dot{\phi}(x) - L . \tag{3.80}$$

Since for our case $\pi(\boldsymbol{x}) = \dot{\phi}(\boldsymbol{x})$,

$$H = \frac{1}{2}\int \mathrm{d}^3\boldsymbol{x} \left[\pi(\boldsymbol{x})^2 + (\boldsymbol{\nabla}\phi(\boldsymbol{x}))^2 + m^2\phi(\boldsymbol{x})^2\right] . \tag{3.81}$$

Now we can shift to the Heisenberg picture

$$\phi(\boldsymbol{x}) \to \phi(\boldsymbol{x}, t) \equiv \phi(x) = e^{iHt}\phi(\boldsymbol{x})e^{-iHt} . \tag{3.82}$$

(Assume that Heisenberg and Schrödinger picture coincide at $t = t_0 = 0$). Similarly for $\pi(\boldsymbol{x})$

$$\pi(\boldsymbol{x}) \to \pi(x) = e^{iHt}\pi(\boldsymbol{x})e^{-iHt} . \tag{3.83}$$

Generally, we can derive equal-time commutation relations

$$\begin{aligned} [\phi(\boldsymbol{x}, t), \pi(\boldsymbol{x}', t)] &= i\delta(\boldsymbol{x} - \boldsymbol{x}') , \\ [\phi(\boldsymbol{x}, t), \phi(\boldsymbol{x}', t)] &= 0 . \end{aligned} \tag{3.84}$$

Equations of motion for Heisenberg picture fields are then

$$\dot{\phi}(x) = i\,[H, \phi(x)] = i \int \mathrm{d}^3\boldsymbol{x}'\,[\mathcal{H}(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]\;. \tag{3.85}$$

To this we need to evaluate $[\mathcal{H}(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]$,

$$\begin{aligned}
[\mathcal{H}(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)] &= \frac{1}{2}\left[\pi^2(\boldsymbol{x}', t) + (\boldsymbol{\nabla}\phi(\boldsymbol{x}', t))^2 + m^2\phi^2(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right] \\
&= \frac{1}{2}\left[\pi^2(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right] \\
&= \frac{1}{2}\pi(\boldsymbol{x}', t)\,[\pi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)] \\
&\quad + \frac{1}{2}\,[\pi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)]\,\pi(\boldsymbol{x}', t) \\
&= -i\pi(\boldsymbol{x}', t)\delta(\boldsymbol{x} - \boldsymbol{x}')\,.
\end{aligned} \tag{3.86}$$

Thus

$$[H, \phi(\boldsymbol{x}, t)] = -i \int \mathrm{d}^3\boldsymbol{x}'\pi(\boldsymbol{x}', t)\delta(\boldsymbol{x} - \boldsymbol{x}') = -i\pi(\boldsymbol{x}, t)\,, \tag{3.87}$$

which leads to

$$\dot{\phi}(\boldsymbol{x}, t) = -i(i\pi(\boldsymbol{x}, t)) = \pi(\boldsymbol{x}, t)\,. \tag{3.88}$$

Similarly

$$\dot{\pi}(\boldsymbol{x}, t) = i\,[H, \pi(\boldsymbol{x}, t)] = i \int \mathrm{d}^3\boldsymbol{x}'\,[\mathcal{H}(\boldsymbol{x}', t), \pi(\boldsymbol{x}, t)]\;. \tag{3.89}$$

For this we need

$$\begin{aligned}
[\mathcal{H}(\boldsymbol{x}', t), \pi(\boldsymbol{x}, t)] &= \frac{1}{2}\left[(\boldsymbol{\nabla}\phi(\boldsymbol{x}', t))^2, \pi(\boldsymbol{x}, t)\right] \\
&\quad + \frac{1}{2}m^2\left[\phi^2(\boldsymbol{x}', t), \pi(\boldsymbol{x}, t)\right] \\
&= i\boldsymbol{\nabla}'\phi(\boldsymbol{x}', t)\boldsymbol{\nabla}'\delta(\boldsymbol{x} - \boldsymbol{x}') \\
&\quad + im^2\phi(\boldsymbol{x}', t)\delta(\boldsymbol{x} - \boldsymbol{x}')\,,
\end{aligned} \tag{3.90}$$

and hence

$$\begin{aligned}
[H, \pi(\boldsymbol{x}, t)] &= -i \int \mathrm{d}^3\boldsymbol{x}'\{\boldsymbol{\nabla}\phi(\boldsymbol{x}', t)\boldsymbol{\nabla}\delta(\boldsymbol{x} - \boldsymbol{x}') + m^2\phi(\boldsymbol{x}', t)\delta(\boldsymbol{x} - \boldsymbol{x}')\} \\
&= i\{-i\boldsymbol{\nabla}^2\phi(\boldsymbol{x}, t) + im^2\phi(\boldsymbol{x}, t)\}\,.
\end{aligned} \tag{3.91}$$

This leads to the second equation of motion of the form

$$\dot{\pi}(\boldsymbol{x}, t) = \boldsymbol{\nabla}^2\phi(\boldsymbol{x}, t) - m^2\phi(\boldsymbol{x}, t)\,. \tag{3.92}$$

Recall that $\dot{\phi}(\boldsymbol{x}, t) = \pi(\boldsymbol{x}, t)$, then

$$\ddot{\phi}(\boldsymbol{x}, t) = \dot{\pi}(\boldsymbol{x}, t) = \boldsymbol{\nabla}^2\phi(\boldsymbol{x}, t) - m^2\phi(\boldsymbol{x}, t)\,. \tag{3.93}$$

This is an equation of motion of the Heisenberg field, i.e.

$$\ddot{\phi} - \nabla^2 \phi = -m^2 \phi \;\Leftrightarrow\; \partial_\mu \partial^\mu \phi + m^2 \phi = 0. \tag{3.94}$$

Which can be shortly rewritten as

$$\left(\Box + m^2\right) \phi = 0. \tag{3.95}$$

Energy and Momentum Operators

Our Hamiltonian reads

$$\begin{aligned} H &= \int \mathrm{d}^3\boldsymbol{x} \left[\pi(x)\dot{\phi}(x) - \mathcal{L}(x)\right] \\ &= \int \mathrm{d}^3\boldsymbol{x}\, \dot{\phi}^2(\boldsymbol{x}, t) - L(\boldsymbol{x}, t)\,. \end{aligned} \tag{3.96}$$

Heisenberg field obeys $\phi(\boldsymbol{x}, t) = e^{iHt}\phi(\boldsymbol{x}, 0)e^{-iHt}$, which implies that

$$\phi(\boldsymbol{x}, \tau - t) = e^{-iH\tau}\phi(\boldsymbol{x}, t)e^{iH\tau}. \tag{3.97}$$

The question now is, how does operator $\boldsymbol{P}$, that affects spatial transformation, look like? If we translate the physical system by a spatial displacement $\boldsymbol{a}$, then $\phi(\boldsymbol{x}, t) \to \phi(\boldsymbol{x} - \boldsymbol{a}, t)$. Idea is that $\boldsymbol{P}$ is the generator of these translation. We therefore require

$$e^{i\boldsymbol{P}\cdot\boldsymbol{a}}\phi(\boldsymbol{x}, t)e^{-i\boldsymbol{P}\cdot\boldsymbol{a}} = \phi(\boldsymbol{x} - \boldsymbol{a}, t). \tag{3.98}$$

Let $\boldsymbol{a}$ is infinitesimal. Then we can write

$$\phi(\boldsymbol{x}, t) + i\left[\boldsymbol{P}\cdot\boldsymbol{a}, \phi(\boldsymbol{x}, t)\right] + O(\boldsymbol{a}^2). \tag{3.99}$$

But $\phi(\boldsymbol{x} - \boldsymbol{a}, t) = \phi(\boldsymbol{x}, t) - \boldsymbol{a}\cdot\nabla\phi(\boldsymbol{x}, t) + O(\boldsymbol{a}^2)$. This implies that

$$i\left[\boldsymbol{P}\cdot\boldsymbol{a}, \phi(\boldsymbol{x}, t)\right] = -\boldsymbol{a}\cdot\nabla\phi(\boldsymbol{x}, t). \tag{3.100}$$

Since $\boldsymbol{a}$ is arbitrary, we require

$$i\left[\boldsymbol{P}^k, \phi(\boldsymbol{x}, t)\right] = -\nabla_k\phi(\boldsymbol{x}, t). \tag{3.101}$$

To construct $\boldsymbol{P}^k$, we observe that because of the canonical commutation relations

$$\begin{aligned} \left[\pi(\boldsymbol{x}', t)\nabla'_k\phi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right] &= \left[\pi(\boldsymbol{x}', t), \phi(\boldsymbol{x}, t)\right]\nabla'_k\phi(\boldsymbol{x}', t) \\ &= -i\delta(\boldsymbol{x} - \boldsymbol{x}')\nabla'_k\phi(\boldsymbol{x}', t), \end{aligned} \tag{3.102}$$

we can choose

$$\boldsymbol{P}^k = -\int \mathrm{d}^3\boldsymbol{x}'\pi(\boldsymbol{x}', t)\nabla_k\phi(\boldsymbol{x}', t). \tag{3.103}$$

In such case,

$$\begin{aligned}\left[\boldsymbol{P}^k, \phi(\boldsymbol{x},t)\right] &= -\int \mathrm{d}^3\boldsymbol{x}'[\pi(\boldsymbol{x},t)\boldsymbol{\nabla}_k\phi(\boldsymbol{x}',t),\phi(\boldsymbol{x},t)] \\ &= (-1)(-i)\int \mathrm{d}^3\boldsymbol{x}'\delta(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{\nabla}_k\phi(\boldsymbol{x}',t) = i\boldsymbol{\nabla}_k\phi(\boldsymbol{x},t)\,.\end{aligned}$$

(3.104)This is indeed what we required for $\boldsymbol{P}^k$. This argument fixes $\boldsymbol{P}^k$ modulo additive c-number (this fact will be relevant shortly).

By using the fact that

$$\boldsymbol{\nabla}_k = \partial_k = -\partial^k,\ (k=1,2,3)\ \Rightarrow\ \boldsymbol{P}^k = \int \mathrm{d}^3\boldsymbol{x}\,\pi(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t)\,, \quad (3.105)$$

we can check that $\boldsymbol{P}^k$ is independent of t, by taking commutator of $[H,\boldsymbol{P}]$. Recalling that

$$H = \int \mathrm{d}^3\boldsymbol{x}\left[\frac{1}{2}\pi^2(\boldsymbol{x},t)+\frac{1}{2}(\boldsymbol{\nabla}\phi(\boldsymbol{x},t))^2+\frac{1}{2}m^2\phi^2(\boldsymbol{x},t)\right], \quad (3.106)$$

we can directly calculate our commutator as

$$\begin{aligned}[H,\boldsymbol{P}] &= -\int \mathrm{d}^3\boldsymbol{x}\mathrm{d}^3\boldsymbol{x}'\left[\pi(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t),\left(\frac{1}{2}\pi^2(\boldsymbol{x}',t)+\right.\right.\\ &\qquad \left.\left.\frac{1}{2}(\boldsymbol{\nabla}\phi(\boldsymbol{x}',t))^2 + \frac{1}{2}m^2\phi^2(\boldsymbol{x}',t)\right)\right] \\ &= \int \mathrm{d}^3\boldsymbol{x}'\left[H,\pi(\boldsymbol{x}',t)\right]\partial'^k\phi(\boldsymbol{x}',t)+\pi(\boldsymbol{x}',t)\left[H,\partial'^k\phi(\boldsymbol{x},t)\right] \\ &= \int \mathrm{d}^3\boldsymbol{x} - i\dot{\pi}(\boldsymbol{x},t)\partial^k\phi(\boldsymbol{x},t)+\pi(\boldsymbol{x},t)\partial^k\underbrace{[H,\phi(\boldsymbol{x},t)]}_{=-i\pi(\boldsymbol{x},t)} \\ &= -i\int \mathrm{d}^3\boldsymbol{x}'\left[\left(\boldsymbol{\nabla}'\boldsymbol{\nabla}'\phi(\boldsymbol{x}',t)-m^2\phi(\boldsymbol{x}',t)\right)\partial^k\phi(\boldsymbol{x}',t)\right.\\ &\qquad \left.+\,\pi(\boldsymbol{x}',t)\partial'^k\pi(\boldsymbol{x}',t)\right] \\ &= -i\int \mathrm{d}^3\boldsymbol{x}\,\boldsymbol{\nabla}\boldsymbol{\nabla}\phi(\boldsymbol{x},t)\boldsymbol{\nabla}_k\phi(\boldsymbol{x},t) \\ &= -i\int \mathrm{d}^3\boldsymbol{x}\underbrace{\left\{-\frac{1}{2}\boldsymbol{\nabla}_k\left[\boldsymbol{\nabla}_l\phi(\boldsymbol{x},t)\boldsymbol{\nabla}_l\phi(\boldsymbol{x},t)\right]\right\}}_{\text{surface term}} \\ &= 0.\end{aligned} \quad (3.107)$$

So, for a field that for any fixed t and $|\boldsymbol{x}|\to\infty$ goes quickly to zero, the commutator of $[H,\boldsymbol{P}]=0$. Set $P^0 = H$, then we can combine the above results for H and $\boldsymbol{P}$ to a 4-vector

$$P^\mu = (P^0,\boldsymbol{P}) = \int \mathrm{d}^3\boldsymbol{x}\,T^{0\mu}(x)\,, \quad (3.108)$$

where

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - g^{\mu\nu}\mathcal{L}(\phi(x),\dot{\phi}(x))\,. \quad (3.109)$$

From this last formula we have

$$T^{00} = \partial^0\phi\partial^0\phi - \mathcal{L} = \partial_t\phi\partial_t\phi - \mathcal{L} = \pi\dot{\phi} - \mathcal{L} = \mathcal{H}\,, \tag{3.110}$$

and

$$T^{0k} = \partial^0\phi\partial^k\phi = -\dot{\phi}\boldsymbol{\nabla}_k\phi = -\pi\boldsymbol{\nabla}_k\phi\,, \tag{3.111}$$

which consistently implies

$$\boldsymbol{P}^k = \int \mathrm{d}^3\boldsymbol{x}\, T^{0k}(x)\,. \tag{3.112}$$

In addition, it follows from the equations of motion that

$$\partial_\mu T^{\mu\nu} = 0\,. \tag{3.113}$$

Indeed, expanding it we get

$$\begin{aligned} \partial_\mu T^{\mu\nu} &= \partial_\mu\left[\partial^\mu\phi\partial^\nu\phi - g^{\mu\nu}\left(\frac{1}{2}(\partial_\alpha\phi)^2 - \frac{1}{2}m^2\phi^2\right)\right] \\ &= \underbrace{(\partial_\mu\partial^\mu\phi)}_{-m^2\phi}\partial^\nu\phi + \partial^\mu\phi\partial_\mu\partial^\nu\phi - (\partial^\nu\partial_\alpha\phi)\partial^\alpha\phi \\ &= -m^2\phi\partial^\alpha\phi + m^2\phi\partial^\nu\phi = 0\,, \end{aligned} \tag{3.114}$$

Note that (3.108) can be written in explicitly covariant manner, namely

$$P^\mu = (P^0, \boldsymbol{P}) = \int \mathrm{d}V\, n_\mu T^{\mu\nu}(x)\,. \tag{3.115}$$

Here, the measure dV is over the space-like slice (of the $4d$ spacetime) that is orthogonal to the unit time-like vector n^μ. We claim that P^μ is independent of n^μ. To see this, we can write

$$\begin{aligned} \int \mathrm{d}^3\boldsymbol{x}\, T^{0\mu}(x) &= \int \mathrm{d}^4\boldsymbol{x}\,\delta(x_0)T^{0\mu}(x) = \int \mathrm{d}^4\boldsymbol{x}\left[\frac{\partial}{\partial x^0}\theta(x_0)\right]T^{0\mu}(x) \\ &= \int \mathrm{d}^4x\,\frac{\partial}{\partial x^\nu}\theta(n_0^\alpha x_\alpha)T^{\nu\mu}(x)\,. \end{aligned} \tag{3.116}$$

Here, $n_0^\alpha = (1, 0, 0, 0)$ is a time-like unit vector that is orthogonal to the space-like slice over which we integrate. Now, we relabel x^μ in (3.115) to x'^μ and take the Lorentz transformation $x^\mu \xrightarrow{\boldsymbol{L}} x'^\mu = \boldsymbol{L}^\mu{}_\nu x^\nu$. With this we can write

$$
\begin{aligned}
&\int d^4x' \frac{\partial}{\partial x'^\nu}\theta(n_0^\alpha x'_\alpha)T^{\nu\mu}(x') \\
&= \int d^4x\,|\det \boldsymbol{L}|\boldsymbol{L}_\nu{}^\sigma \frac{\partial}{\partial x^\sigma}\theta(n_{0,\alpha}\boldsymbol{L}^\alpha{}_\beta x^\beta)\boldsymbol{L}^\nu{}_\gamma \boldsymbol{L}^\mu{}_\delta T^{\gamma\delta}(\underbrace{\boldsymbol{L}^{-1}x'}_{x}) \\
&= \int d^4x\,\eta^\sigma_\gamma \frac{\partial(n_{0,\alpha}\boldsymbol{L}^\alpha{}_\beta x^\beta)}{\partial x^\sigma}\delta(n_{0,\alpha}\boldsymbol{L}^\alpha{}_\beta x^\beta)\boldsymbol{L}^\mu{}_\delta T^{\gamma\delta}(x) \\
&= \boldsymbol{L}^\mu{}_\delta \int d^4x\, n'_\beta \delta(n'x)T^{\beta\delta}(x) \\
&= \boldsymbol{L}^\mu{}_\delta \int dV\, n'_\beta T^{\beta\delta}(x)\,,
\end{aligned}
\tag{3.117}
$$

n_β denotes time-like 4-vector orthogonal to space-like slice over which we integrate. $T^{\mu\nu}(x)$ is called **energy-momentum tensor** and satisfies $T^{\mu\nu} = T^{\nu\mu}$.

with $n'^\mu = \boldsymbol{L}^\mu{}_\nu n_0^\nu$, so that

$$
P^\mu(n_0) = P^\mu(L^{-1}n') = L^\mu{}_\alpha \int dV n'_\beta T^{\beta\alpha}(x) = L^\mu{}_\alpha P'^\mu(n')\,. \tag{3.118}
$$

Thus, the above 4-vector (3.115) transforms as a true relativistic 4-vector and hence the actual value of P^μ is independent of the space-like slice over which we integrate.

Particle interpretation

Let us first Fourier decompose $\phi(x)$ as

$$
\phi(x) = \int d^4p e^{-ipx}\tilde{\phi}(p)\,. \tag{3.119}
$$

We get

$$
\begin{aligned}
(\Box + m^2)\phi(x) = 0 \Rightarrow &\int d^4p e^{-ipx}(p^2 - m^2)\tilde{\phi}(p) = 0 \\
\Rightarrow\;& (p^2 - m^2)\tilde{\phi}(p) = 0\,.
\end{aligned}
\tag{3.120}
$$

Solution of this equation has the generic form

$$
\begin{aligned}
\tilde{\phi}(p) &= f(p)\delta(p^2 - m^2) \\
&= f(p)\frac{\delta(p_0 + \sqrt{\boldsymbol{p}^2 + m^2})}{2\underbrace{\sqrt{\boldsymbol{p}^2 + m^2}}_{\omega_{\boldsymbol{p}}}} + f(p)\frac{\delta(p_0 - \sqrt{\boldsymbol{p}^2 + m^2})}{2\sqrt{\boldsymbol{p}^2 + m^2}}\,.
\end{aligned}
\tag{3.121}
$$

This implies that

$$\begin{aligned}\phi(x) &= \int \frac{d^4p}{2\omega_{\boldsymbol{p}}} e^{-ipx}[f(p)\delta(p_0+\omega_{\boldsymbol{p}}) + f(p)\delta(p_0-\omega_{\boldsymbol{p}})] \\ &= \int \frac{d^3\boldsymbol{p}}{2\omega_{\boldsymbol{p}}} \left[e^{-i\omega_{\boldsymbol{p}}t_0+i\boldsymbol{px}} f(\omega_{\boldsymbol{p}},\boldsymbol{p}) + e^{i\omega_{\boldsymbol{p}}t_0+i\boldsymbol{px}} f(-\omega_{\boldsymbol{p}},\boldsymbol{p})\right] \\ &= \int \frac{d^3\boldsymbol{p}}{2\omega_{\boldsymbol{p}}} \left[e^{-ipx} f(p) + e^{ipx} g(p)\right].\end{aligned} \tag{3.122}$$

Here we have set $g(p) = f(-\omega_{\boldsymbol{p}}, -\boldsymbol{p})$. We can rewrite

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \left(a(p)e^{-ipx} + b(p)e^{ipx}\right), \tag{3.123}$$

where

$$\begin{aligned} a(p) &= (2\pi)^3 f(p), \\ b(p) &= (2\pi)^3 g(p). \end{aligned} \tag{3.124}$$

By requiring that ϕ is Hermitian in order to have Hermitian Hamiltonian, we have that $b(p) = a^\dagger(p)$, which gives

$$\phi(x) = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \left[a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right]. \tag{3.125}$$

For conjugated momenta $\pi(x) = \dot{\phi}(x)$, we have

$$\begin{aligned}\pi(x) = \dot{\phi}(x) &= \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \left[(-i\omega_{\boldsymbol{p}})a(p)e^{-ipx} + i\omega_{\boldsymbol{p}} a^\dagger(p)e^{ipx}\right] \\ &= -\frac{i}{2}\int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \left[a(p)e^{-ipx} - a^\dagger(p)e^{ipx}\right].\end{aligned} \tag{3.126}$$

The measure in the integral (3.125) is manifestly Lorentz invariant. Indeed,

$$\frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} = \frac{d^4p}{(2\pi)^4} 2\pi\delta(p^2-m^2)\theta(p^0). \tag{3.127}$$

Canonical commutation relations are defined by

$$\begin{aligned} [\phi(\boldsymbol{x},t), \pi(\boldsymbol{y},t)] &= i\delta(\boldsymbol{x}-\boldsymbol{y}) \\ \Leftrightarrow \left[\phi(\boldsymbol{x},t), \dot{\phi}(\boldsymbol{y},t)\right] &= i\delta(\boldsymbol{x}-\boldsymbol{y}) \\ [\phi(\boldsymbol{x},t), \phi(\boldsymbol{y},t)] = [\pi(\boldsymbol{x},t), \pi(\boldsymbol{y},t)] &= 0. \end{aligned} \tag{3.128}$$

This also provides canonical commutation relations for $a(p)$ and $a^\dagger(p)$.

Let us make a statement about them:

$$\begin{aligned} \left[a(p), a^\dagger(p')\right] &= (2\pi)^3 2\omega_p \delta(\boldsymbol{p}-\boldsymbol{p}') \\ [a(p), a(p')] = \left[a^\dagger(p), a^\dagger(p')\right] &= 0. \end{aligned} \tag{3.129}$$

The consistency check can be done easily.

First, we will use the notation

$$\sum_{\boldsymbol{p}} \leftrightarrow \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p}, \tag{3.130}$$

$$\delta_{\boldsymbol{pp}'} = (2\pi)^3 2\omega_p \delta(\boldsymbol{p} - \boldsymbol{p}'), \tag{3.131}$$

$$\sum_{\boldsymbol{p}'} \delta_{\boldsymbol{pp}'} f(\boldsymbol{p}'), \tag{3.132}$$

$$\left[a(p), a^\dagger(p')\right] = \delta_{\boldsymbol{pp}'}. \tag{3.133}$$

In this notation our field operator reads

$$\phi(x) = \sum_{\boldsymbol{p}} \left(a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right) \tag{3.134}$$

and

$$\pi(x) = \dot{\phi}(x) = \sum_{\boldsymbol{p}} (-i\omega_p) \left(a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right). \tag{3.135}$$

Now set $x \equiv (t, \boldsymbol{x})$ and $y \equiv (t, \boldsymbol{y})$, which gives

$$\begin{aligned}
[\phi(x), \pi(y)] &= \sum_{\boldsymbol{pp}'} (-i\omega_{\boldsymbol{p}'}) \left[a(p)e^{-ipx}, (-a^\dagger(p))e^{ip'y}\right] \\
&\quad + (-i\omega_{\boldsymbol{p}'}) \left[a^\dagger(p)e^{ipx}, a(p)e^{-ip'y}\right] \\
&= \sum_{\boldsymbol{p}} \left[(i\omega_p)e^{i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{y})} + (i\omega_p)e^{-i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{y})}\right] \\
&= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{i2\omega_p}{2\omega_p} e^{i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{y})} \\
&= i\delta(\boldsymbol{x} - \boldsymbol{y}).
\end{aligned} \tag{3.136}$$

Similarly, it can be checked that

$$[\phi(x), \phi(y)] = [\pi(x), \pi(y)] = 0. \tag{3.137}$$

Energy - Hamiltonian

Hamiltonian can be written in the form

It is normally tedious, however, we can drop time-dependent terms.

$$H = \int \mathrm{d}^3\boldsymbol{x}\, \underbrace{\frac{1}{2}\pi^2(x)}_{H_1} + \underbrace{\frac{1}{2}(\boldsymbol{\nabla}\phi(x))^2}_{H_2} + \underbrace{m^2\phi^2}_{H_3}. \tag{3.138}$$

Let us now compute the respective terms explicitly in terms of a and

$a^\dagger$:

$$\begin{aligned}
H_1 &= \int \mathrm{d}^3\boldsymbol{x}\,\frac{1}{2}\pi^2 \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}(-i\omega_p)(-i\omega_{p'})\int \mathrm{d}^3\boldsymbol{x}\, a(p)a(p')e^{-i(p+p')x} - a(p)a^\dagger(p')e^{-i(p-p')x} \\
&\quad + a^\dagger(p)a^\dagger(p')e^{i(p+p')x} - a^\dagger(p)a(p')e^{i(p-p')x} \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}(-\omega_p\omega_{p'})\left[-a(p)a^\dagger(p')(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}') - a^\dagger(p)a(p')(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')\right] \\
&= \frac{1}{4}\sum_{\boldsymbol{p}}\omega_p\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right]. \qquad (3.139)
\end{aligned}$$

$$\begin{aligned}
H_2 &= \frac{1}{2}\int \mathrm{d}^3\boldsymbol{x}\,(\boldsymbol{\nabla}\phi)^2 \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}(i\boldsymbol{p})(-i\boldsymbol{p}')\int \mathrm{d}^3\boldsymbol{x}[a(p)a^\dagger(p')e^{-i(p-p')x} + a^\dagger(p)a(p')e^{i(p-p')x}] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}\boldsymbol{p}'}\boldsymbol{p}\boldsymbol{p}'(2\pi)^3\delta^3(\boldsymbol{p}-\boldsymbol{p}')[a(p)a^\dagger(p') + a^\dagger(p)a(p')] \\
&= \frac{1}{2}\sum_{\boldsymbol{p}}\frac{\boldsymbol{p}^2}{2\omega_p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right], \qquad (3.140)
\end{aligned}$$

$$H_3 = \frac{1}{2}\sum_{p}\frac{m^2}{2\omega_p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right]. \qquad (3.141)$$

which gives

$$H = H_1 + H_2 + H_3 = \frac{1}{2}\sum_{\boldsymbol{p}}\omega_p\left(a(p)a^\dagger(p) + a^\dagger(p)a(p)\right). \qquad (3.142)$$

We could rewrite this as

$$\sum_{\boldsymbol{p}}\omega_p\left(a^\dagger(p)a(p) + \frac{1}{2}\delta_{pp}\right), \qquad (3.143)$$

which mimics the linear harmonic oscillator, though now we have sum of infinitely many of them.
The term $\frac{1}{2}$ will diverge, because at each point in $\boldsymbol{p}$ space we get vacuum oscillation.

Spatial momentum

Spatial momentum can be written as

$$\boldsymbol{P} = -\int \mathrm{d}^3\boldsymbol{x}\,\pi(x)\boldsymbol{\nabla}\phi(x). \qquad (3.144)$$

Analogous calculation as for H gives

$$\boldsymbol{P} = \frac{1}{2}\sum_{\boldsymbol{p}}\boldsymbol{p}\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right], \qquad (3.145)$$

which does not suffer with vacuum divergences. Together with H, we can form four-vector

$$P^\mu = \frac{1}{2}\sum_{\boldsymbol{p}}p^\mu\left[a(p)a^\dagger(p) + a^\dagger(p)a(p)\right]. \qquad (3.146)$$

Particle interpretation

Similarly as in linear harmonic oscillator, the physical states of our quantum system are excitations over a ground state, which we identify with the vacuum state $|0\rangle$. Here $|0\rangle$ satisfies $a(p)|0\rangle = 0$ for all $\boldsymbol{p}$.

Typical states are:

- $|p\rangle = a^\dagger(p)|0\rangle$ - **single-particle state**, it creates excitation with energy $\omega_{\boldsymbol{p}}$ and momenta $\boldsymbol{p}$. In linear harmonic oscillator, $|1\rangle = a^\dagger|0\rangle$ creates excited state with energy ω - where no momenta is included, thus one cannot really speak about particle,
- $|p_1, p_2\rangle = a^\dagger(p_1)a^\dagger(p_2)|0\rangle$ - **two-particle state**, it creates two excitations, one with energy $\omega_{\boldsymbol{p}_1}$ and momenta $\boldsymbol{p}_1$, second with energy $\omega_{\boldsymbol{p}_2}$ and momenta $\boldsymbol{p}_2$.

In linear harmonic oscillator,

$$|2\rangle = \frac{a^\dagger a^\dagger}{\sqrt{2}}|0\rangle\ , \tag{3.147}$$

creates 2 excited states above $|0\rangle$ with energy $\omega + \omega = 2\omega$.

We can construct a **particle number operator**

Clearly an analogue of the energy-level operator in linear harmonic oscillator $\hat{H} = a^\dagger a$.

$$N = \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p} a^\dagger(p)a(p) = \sum_{\boldsymbol{p}} a^\dagger(p)a(p)\,. \tag{3.148}$$

This gives

$$\begin{aligned}[N, a^\dagger(p)] &= \sum_{\boldsymbol{p}}[a^\dagger(p')a(p'), a^\dagger(p)] \\ &= \sum_{\boldsymbol{p}} a^\dagger(p') \underbrace{[a(p'), a^\dagger(p')]}_{\delta_{\boldsymbol{p}\boldsymbol{p}'}} = a^\dagger(p).\end{aligned} \tag{3.149}$$

So,

$$Na^\dagger(p) = a^\dagger N + a^\dagger(p) = a^\dagger(p)(N+1). \tag{3.150}$$

Operator N counts particles in a given state. Indeed,

$$N|0\rangle = \sum_{\boldsymbol{p}} a^\dagger(p)a(p)|0\rangle = 0, \tag{3.151}$$

so for vacuum, $N = 0$. Similarly,

$$N|p\rangle = Na^\dagger(p)|0\rangle = a^\dagger(p)(N+1)|0\rangle = a^\dagger(p)|0\rangle = |p\rangle\,, \tag{3.152}$$

$|p\rangle$ is an eigenstate of N with $N = 1$.

$$\begin{aligned} N|p_1, p_2\rangle &= Na^\dagger(p_1)a^\dagger(p_2)|0\rangle = a^\dagger(p_1)(N+1)a^\dagger(p_2)|0\rangle \\ &= a^\dagger(p_1)a^\dagger(p_2)(N+2)|0\rangle = 2|p_1, p_2\rangle\,, \end{aligned} \tag{3.153}$$

$|p_1 p_2\rangle$ is a 2-particle eigenstate of N.

Occupation number representation

$$|n_{p_1}, n_{p_2}, \cdots\rangle = \frac{(a^\dagger_{\boldsymbol{p}_1})^{n_{\boldsymbol{p}_1}}}{\sqrt{n_{\boldsymbol{p}_1}!}} \frac{(a^\dagger_{\boldsymbol{p}_2})^{n_{\boldsymbol{p}_2}}}{\sqrt{n_{\boldsymbol{p}_2}!}} \cdots |0\rangle\,, \tag{3.154}$$

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle\,. \tag{3.155}$$

- occupation number representation takes into account automatically symmetric exchange of particles (bosonic indistinguishability),
- occupation number basis is an orthonormal basis on the Hilbert space $\mathcal{H}_N$ for each fixed $N = \sum_i n_{p_i}$,
- normalization to 1 is intuitive $\langle n_{p_1} \ldots n_{p_k} \ldots | n'_{p_1} \ldots n'_{p_k} \ldots\rangle =$ 0, where $n' = n$ for all p_i.

The Hilbert space can be combined in the so called **Fock space**

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \ldots = \bigoplus_{N=0}^{\infty} \mathcal{H}_N\,. \tag{3.156}$$

The Hilbert space $\mathcal{H}_0$ only contains one element, the vacuum $|0\rangle = |0, 0, 0, \ldots 0 \ldots\rangle$.

Note: For instance, we have that

$$|p_i, p_j\rangle = |\underbrace{0}_{p_1}\;\underbrace{0}_{p_2}\;\ldots\;\underbrace{1}_{p_i}\;\ldots 0 \ldots\;\underbrace{1}_{p_j}\;0 \ldots 0\rangle\,. \tag{3.157}$$

Note on vacuum energy

We have

$$\frac{1}{2}\left(a(p)a^\dagger(p) + a^\dagger(p)a(p)\right) = a^\dagger(p)a(p) + \frac{1}{2}\delta_{\boldsymbol{pp}'}. \tag{3.158}$$

Now

$$(2\pi)^3\delta(0) = \int \mathrm{d}^3\boldsymbol{x}\, e^{\pm i\boldsymbol{px}}|_{\boldsymbol{p}=0} = \int \mathrm{d}^3\boldsymbol{x} = V, \tag{3.159}$$

this is divergence, given by the infinite volume. Then,

$$\begin{aligned}\delta_{\boldsymbol{pp}'} &= \text{energy density} \cdot V \Rightarrow \frac{1}{2}\sum_{\boldsymbol{p}} \omega_p \delta_{\boldsymbol{pp}'} \\ &= \frac{1}{2}\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3\omega_p^2}(2\pi)^3 2\omega_p \delta(0) \\ &= \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\omega_p V,\end{aligned} \tag{3.160}$$

where $\omega_p V$ is the density of energy with given $\boldsymbol{p}$ in the entire space. Compare:

- LHO
  - ω is $\boldsymbol{p}$ independent,
  - there is only 1 or finite number of oscillators in volume V,
- QFT
  - frequency $\omega_{\boldsymbol{p}}$ depends on $\boldsymbol{p}$ which changes smoothly over $\mathbb{R}^3$,
  - in a given volume V, there is infinity of oscillators with different frequencies.

Vacuum energy density diverges at $|\boldsymbol{p}| \to \infty$. This UV divergence arises because the "oscillators" with large momentum have very large zero-point energy $1/2\omega_p \approx \frac{|\boldsymbol{p}|}{2}$.

Vacuum energy is relevant only for the gravitational field. If we neglect gravity then the presence of vacuum energy cannot be detected in experiment involving only transformation between excited states.

The subtraction of the vacuum energy does not, however, remove the vacuum fluctuations of the quantum field. This can be evaluated from the correlation function

$$\langle 0|\,\psi(\boldsymbol{x};t)\psi(\boldsymbol{y};t)\,|0\rangle\,. \tag{3.161}$$

Similarly as subtraction of vacuum energy in LHO does not remove fluctuation of $\boldsymbol{p}$.

Typically one wishes to ensure that vacuum has 0 energy and momentum. So, we subtract the ground state energy and define the Hamiltonian to be original Hamiltonian minus ground state energy. We wish to set $P^\mu = \sum_{\boldsymbol{p}} P^\mu a^\dagger(p)a(p)$. Now $P^\mu|0\rangle = 0$.

$$\begin{aligned}P^\mu|p\rangle &= \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(p')a(p')a^\dagger(p)|0\rangle = \sum_{\boldsymbol{p}'} p^{\mu\prime}\left[a^\dagger(p')a(p'), a^\dagger(p)\right]|0\rangle \\ &= \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(p')\left[a(p'), a^\dagger(p)\right]|0\rangle = \sum_{\boldsymbol{p}'} p^{\mu\prime} a^\dagger(p')\delta_{\boldsymbol{p}'\boldsymbol{p}}|0\rangle \\ &= p^\mu a^\dagger(p)|0\rangle = p^\mu|p\rangle\,.\end{aligned} \tag{3.162}$$

This means that $|p\rangle$ is an eigenstate of P^μ with eigenvalue p^μ.

For two particle states:

$$\begin{aligned}
P^{\mu}|p_1,p_2\rangle &= \sum_{p'} p^{\mu\prime}(a^{\dagger}(p')a(p')a^{\dagger}(p_1)a^{\dagger}(p_2)|0\rangle \\
&= \sum_{p'} p^{\mu\prime}\left[a^{\dagger}(p')a(p'),a^{\dagger}(p_1)a^{\dagger}(p_2)\right]|0\rangle \\
&= \sum_{p'} p^{\mu\prime}a^{\dagger}(p')\left[a(p'),a^{\dagger}(p_1)a^{\dagger}(p_2)\right]|0\rangle \\
&= \sum_{P'} p^{\mu\prime}a^{\dagger}(p')\left[a^{\dagger}(p_2)\delta_{p'p_1}+a^{\dagger}(p_1)\delta_{pp_2}\right]|0\rangle \\
&= (p_1^{\mu}+p_2^{\mu})a^{\dagger}(p_1)a^{\dagger}(p_2)|0\rangle \\
&= (p_1^{\mu}+p_2^{\mu})|p_1,p_2\rangle\,. \qquad (3.163)
\end{aligned}$$

The change does not destroy the significance of the previous calculations with p^{μ}, since only commutators are involved. Adjustment by a c-number does not affect commutators.

This means that $|p_1,p_2\rangle$ is an eigenstate of P^{μ} with $(p_1^{\mu}+p_2^{\mu})$ as eigenvalue.

Question: What about if some p_i coincide?

The zero-point energy can be removed automatically by interpreting product of field operators (from classical formulas) as normal ordered product.

3.5 Normal ordered product

We have $\phi^{+}(x)+\phi^{-}(x)$ with $\phi^{+}(x)=\sum_{\boldsymbol{P}}a(p)e^{-ipx}$ being positive frequency part and $\phi^{-}(x)=\sum_{\boldsymbol{P}}a^{\dagger}(p)e^{ipx}$ being negative frequency part. That gives us:

For p^i normal ordering is not needed since $\sum_{\boldsymbol{P}}\frac{p^i}{2}=0$.

$$\phi(x)\phi(y)=(\phi^{+}+\phi^{-})_x(\phi^{+}+\phi^{-})_y=\phi_x^{+}\phi_y^{+}+\phi_x^{-}\phi_y^{-}+\phi_x^{-}\phi_y^{+}+\phi_x^{+}\phi_y^{-}. \quad (3.164)$$

Normal ordering is defined as

$$:\phi(x)\phi(y):=\ \phi^{+}(x)\phi^{+}(y)+\phi^{-}(x)\phi^{+}(y)+\phi^{-}(y)\phi^{+}(x)+\phi^{-}(x)\phi^{-}(y) \quad (3.165)$$

The key point in this relation is that $\langle 0|:\phi(x)\phi(y):|0\rangle=0$. Because $\phi^{+}(x)\phi^{-}(y)$ and $\phi^{-}(y)\phi^{+}(x)$ differ only by c-number (not by an operator) and because $\langle 0|\text{c-number}|0\rangle=$ c-number, we get

This Hamiltonian and the one with vacuum-energy contribution can be viewed as an ordering ambiguity in moving from the classical to quantum theory.

$$:\phi(x)\phi(y):=\ \phi(x)\phi(y)-\langle 0|\,\phi(x)\phi(y)\,|0\rangle\,. \quad (3.166)$$

So, in $:\cdots:$ we move all ϕ^{+}'s to the right and all ϕ^{-}'s to the left. Having use of this notation we can write

$$P^{\mu}=\int \mathrm{d}^3\boldsymbol{x}\,:T^{\mu 0}:\,. \quad (3.167)$$

Here $:T^{\mu\nu}(x):$ is defined as

$$:T^{\mu\nu}(x):\ =\ :\partial^{\mu}\psi(x)\partial^{\nu}\psi(x):\ -\eta^{\mu\nu}:\mathcal{L}(x):, \quad (3.168)$$

so, in particular

$$P^{\mu}\ =\ \frac{1}{2}\sum_{\boldsymbol{p}}p^{\mu}:a(p)a^{\dagger}(p)+a^{\dagger}(p)a(p):=\sum_{\boldsymbol{p}}p^{\mu}a_p^{\dagger}a_p. \quad (3.169)$$

Let me mention, as a side note that there is simple prescription for a computation of E-M tensor (also known as stress tensor) that is inspired by general relativity, namely $T_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}}$ — Hilbertian stress tensor. This stress tensor is explicitly symmetric and covariantly conserved $\partial_\mu T^{\mu\nu} = 0$. Let this apply to our scalar-field situation. First recall that

$$\frac{\delta S[g,\phi]}{\delta g^{\mu\nu}(y)} = \left.\frac{\mathrm{d}S[g+\epsilon\delta(x-y)]}{\mathrm{d}\epsilon^{\mu\nu}}\right|_{\epsilon=0}, \tag{3.170}$$

$\det \boldsymbol{A} = e^{\mathrm{Tr}\{\log \boldsymbol{A}\}}$

then

$$\frac{\delta S[g,\phi]}{\delta g^{\mu\nu}(y)} = \int \mathrm{d}^4x \left[\frac{\delta\sqrt{-g_x}}{\delta g^{\mu\nu}(y)}\left(g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi - m^2\phi^2\right) + \sqrt{-g}\frac{\delta g^{\alpha\beta}(x)}{\delta g^{\mu\nu}(y)}\partial_\alpha\phi\partial_\beta\phi\right]. \tag{3.171}$$

Here

$$\begin{aligned}\frac{\delta\sqrt{-g_x}}{\delta g^{\mu\nu}(y)} &= \frac{\delta\sqrt{-\det g_{\mu\nu}(x)}}{\delta g^{\mu\nu}(y)} = \frac{\delta\frac{1}{\sqrt{-\det g^{\mu\nu}(x)}}}{\delta g^{\mu\nu}(y)}\\ &= \frac{1}{2}\frac{1}{\sqrt{-\det g^{\mu\nu}}^3}\frac{\mathrm{d}}{\mathrm{d}\epsilon^{\mu\nu}}\det\left(g_x^{\mu\nu} + \epsilon^{\mu\nu}\delta(x-y)\right)\Big|_{\epsilon=0}\\ &= \frac{1}{2}\sqrt{-g}^3\frac{\mathrm{d}}{\mathrm{d}\epsilon^{\mu\nu}}e^{\mathrm{Tr}\{\log(g^{\mu\nu}+\epsilon^{\mu\nu}\delta(x-y))\}} = \frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta(x-y)\end{aligned} \tag{3.172}$$

and

$$\frac{\delta g^{\alpha\beta}(x)}{\delta g^{\mu\nu}(y)} = \frac{\mathrm{d}}{\mathrm{d}\epsilon^{\mu\nu}}\left[g^{\alpha\beta}(x) + \epsilon^{\alpha\beta}\delta(x-y)\right] = \eta^\alpha_\mu\eta^\beta_\nu\delta(x-y). \tag{3.173}$$

Altogether

$$T^{\mu\nu}(y) = \left[\frac{1}{2}\sqrt{-g}g_{\mu\nu}\left(g_{\alpha\beta}\partial^\alpha\phi(y)\partial^\beta\phi(y) - m^2\phi^2(y)\right) + \sqrt{-g}\partial_\mu\phi(y)\partial_\nu\phi(y)\right]. \tag{3.174}$$

If we set $g^{\mu\nu} = \eta^{\mu\nu}$, then $T^{\mu\nu}(y) = \partial^\mu\phi(y)\partial^\nu\phi(y) - \eta^{\mu\nu}\mathcal{L}(y)$. So $T^{00} = \pi\dot{\phi} = \mathcal{L} = \mathcal{H}$ and $T^{0i} = \pi\partial^i\phi = -\pi\boldsymbol{\nabla}_i\phi$.

3.6 Multiplet of Scalar Fields

Now we have $\phi(x) \to \boldsymbol{\phi}(\boldsymbol{x}) = \phi_r(x)_{r=1}^n$ and $\phi_r^+ = \phi_r$. Then we get

$$\mathcal{L}(\phi_r,\dot{\phi}_r,\boldsymbol{\nabla}\phi_r) = \sum_r \frac{1}{2}(\partial_\mu\phi_r\partial^\mu\phi_r - m^2\phi_r^2), \tag{3.175}$$

and the commutation relations are

$$\begin{aligned}[\phi_r(\boldsymbol{x},t),\pi_s(\boldsymbol{y},t)] &= i\delta_{rs}\delta(\boldsymbol{x}-\boldsymbol{y}),\\ [\phi_r(\boldsymbol{x},t),\phi_s(\boldsymbol{y},t)] &= [\pi_r(\boldsymbol{x},t),\pi_r(\boldsymbol{y},t)] = 0,\end{aligned} \tag{3.176}$$

where

$$\pi_r(\boldsymbol{x},t) = \frac{\delta S[\phi]}{\delta\dot{\phi}_r(\boldsymbol{x},t)} = \frac{\partial\mathcal{L}(\phi,\boldsymbol{\nabla}\phi)_x}{\partial\dot{\phi}_r(\boldsymbol{x},t)}. \tag{3.177}$$

For Hamiltonian we get

$$
\begin{aligned}
H &= \sum_r \int \mathrm{d}^3\boldsymbol{x}\pi_r(\boldsymbol{x},t)\dot{\phi}_r(\boldsymbol{x},t) - L \\
&= \frac{1}{2}\sum_r \int \mathrm{d}^3\boldsymbol{x}\left[\pi_r^2(\boldsymbol{x},t) + (\boldsymbol{\nabla}\phi_r(\boldsymbol{x},t))^2 + m^2\phi^2(\boldsymbol{x},t)\right]. \quad (3.178)
\end{aligned}
$$

By analogy to a single-field case using usual normal expansion

$$\phi_r(x) = \sum_p \left(a_r(p)e^{-ipx} + a_r^\dagger(p)e^{ipx}\right). \quad (3.179)$$

We get commutation relations

$$\left[a_r(p), a_s^\dagger(p)\right] = \delta_{rs}\delta_{pp'}. \quad (3.180)$$

Vacuum state $|0\rangle$ is defined is that

$$a_r(p)\,|0\rangle = 0 \quad (3.181)$$

for $\forall \boldsymbol{p}, \boldsymbol{r}$. After appropriate normal ordering we have

$$\hat{P}^\mu = \sum_{rp} P^\mu a_r^\dagger(p)a_r(p). \quad (3.182)$$

and operator of the particle number

Operator that counts only particles of r-th type.

$$\hat{N} = \sum_{rp} a_r^\dagger(p)a_r(p). \quad (3.183)$$

Eigenstates of $\hat{N}$ and $\hat{P}^\mu$ are typically of the form

$$a_{r_1}^\dagger a_{r_2}^\dagger \ldots |0\rangle = |p_1r_1, p_2r_2, \ldots\rangle. \quad (3.184)$$

Then

$$\hat{P}^\mu\,|p_1r_1, p_2r_2, \ldots\rangle = (p_1 + p_2 + \ldots)\,|p_1r_1, p_2r_2, \ldots\rangle. \quad (3.185)$$

N_{r_1} is set of states $n_{r_1}(p_1), n_{r_1}(p_2), \ldots$.

$$\hat{N}_r\,|N_1, N_2, \ldots, N_r, \ldots, N_{r_n}, \ldots\rangle = \sum_p n_r(p)\,|N_1, N_2, \ldots, N_r, \ldots\rangle. \quad (3.186)$$

Fock space another bite

In QM N-particle state contain clear information about which particle occupies which state. This is unphysical due to the indistinguishability of particles. With the construction of the symmetric (or asymmetric) N-particle state this information is eliminated, and the only information which remains is how many particles n_r occupy single-particle state (say $|\psi_r\rangle$). We thus may use a different notation for the symmetric (antisymmetric) state. Notation in terms of the occupation number n_p of a single particle states.

As an example, consider the (normalize) set of wave functions $|\lambda\rangle$ of some single particle bosonic system. The corresponding N-particle wave function for indistinguishable particles has form

$\frac{1}{\sqrt{N!\Pi_\lambda(n_\lambda)!}}$ is total number of particles in state λ.

$$|\lambda_1, \ldots, \lambda_n\rangle = \frac{1}{\sqrt{N!\Pi_\lambda(n_\lambda)!}} \sum_p |\lambda_1, p_1\rangle \otimes |\lambda_2, p_2\rangle \ldots |\lambda_n p_n\rangle . \quad (3.187)$$

Let us order in the ascending sense moments $|p_1, p_2 \ldots, p_n\rangle$, where $p_1 \leq p_2 \leq p_3 \leq \cdots \leq p_n$. So for instance, the state

$$|p_1, p_1, p_1, p_2, p_2, p_3, p_3, p_3, p_4, p_5, p_6 \ldots\rangle , \quad (3.188)$$

contains a redundancy and a more efficient encoding of the state might read

$$|3p_1, 2p_2, 3p_3, 1p_4, 1p_5, 1p_6, \ldots\rangle \quad \text{or} \quad |3, 2, 3, 1, 1, 1, \ldots\rangle . \quad (3.189)$$

We will more generally denote the state as $|\{n_p\}\rangle$, where $\{n_p\}$ is set of occupation numbers. For free Hamiltonian commutes with the total particle number operator $N = \sum_k a_k^\dagger a_k = \sum_k \hat{n_k}$, where shows $a_k^\dagger a_k$ how many particles occupy the state k and n_k is the particle number operator for k-th moment.

Also $N, \hat{n_k}, \hat{n}_{k'}$ commute pair-wisely e.g.

$$\begin{aligned} [\hat{n_k}, \hat{n_{k'}}] &= \left[a_k^\dagger a_k, a_{k'}^\dagger a_{k'}\right] = a_k^\dagger \left[a_k, a_{k'}^\dagger a_{k'}\right] + \left[a_k^\dagger, a_{k'}^\dagger a_{k'}\right] a_k^\dagger \\ &= a_k^\dagger \left[a_k, a_{k'}^\dagger\right] a_{k'} + a_{k'}^\dagger \left[a_k^\dagger, a_{k'}\right] a_k = 0. \end{aligned} \quad (3.190)$$

In this derivation we have used $\left[a_k, a_{k'}^\dagger\right] = \delta_{kk'}$ and $\left[a_{k'}^\dagger, a_{k'}\right] = -\delta_{kk'}$.

So, it is possible to find a common set of eigenstates for all of those commuting operators. These states are fully characterized by specifying the particle set of occupation number $\{n_p\}$, e.g.

$$\hat{n}_i \, |\{n_k\}\rangle = n_i \, |\{n_k\}\rangle , \quad (3.191)$$

and

$$N \, |\{n_k\}\rangle = \sum_l n_l \, |\{n_k\}\rangle . \quad (3.192)$$

So the states $|\{n_k\}\rangle$ form the basis of the Hilbert space. General structure of $|\{n_p\}\rangle$ is

$$|\{n_p\}\rangle = \Pi_p (n_p!)^{-1/2} \left[a^\dagger(p)\right]^{n_p} |0\rangle . \quad (3.193)$$

Generally, we can express action of $a_k^\dagger = a^\dagger(k)$ on $|n_p\rangle$ as

$$\begin{aligned}
a_k^\dagger \,|\{n_p\}\rangle &= \frac{a_{p_1}^{\dagger n_{p_1}}}{\sqrt{n_{p_1}!}} \dots \frac{a_k^{\dagger n_k+1}}{\sqrt{n_k!}} \dots |0\rangle \\
&= \frac{\sqrt{(n_k+1)!}}{\sqrt{n_k!}} \left[\frac{a_{p_1}^{\dagger n_{p_1}}}{\sqrt{n_{p_1}!}} \dots \frac{a_k^{\dagger n_k+1}}{\sqrt{n_k+1!}} \dots |0\rangle \right] \\
&= \sqrt{n_k+1}\,|n_{p_1}, n_{p_2}, \dots, n_k+1, \dots\rangle \ ,
\end{aligned} \tag{3.194}$$

and operator a_k as

$$\begin{aligned}
a_k\,|\{n_p\}\rangle &= \sum_{\{n'\}} |\{n'\}\rangle \,\langle\{n'\}|\, a_k \,|\{n_p\}\rangle \\
&= \sum_{\{n'\}} |\{n'\}\rangle \,\langle n_{p_1}, \dots|\, a_k^\dagger \,|n'_{p_1}, \dots\rangle^* \\
&= \sum_{\{n'\}} \sqrt{n'_k+1}\delta^{\dot{k}}_{n,n'}\delta_{n_k:n'_k+1}\,|\{n'\}\rangle \\
&= \sqrt{n_k}\,|n_{p_1}, \dots, n_k+1, \dots\rangle \ .
\end{aligned} \tag{3.195}$$

Let us now use the resolution of unity

$$\begin{aligned}
1 &= \sum_{\{n\}} |\{n\}\rangle\,\langle\{n\}| \\
&= |0\rangle\,\langle 0| + \sum_{\sum n_i=1} |\{n\}\rangle\,\langle\{n\}| + \sum_{\sum n_i=2} |\{n\}\rangle\,\langle\{n\}| \dots .
\end{aligned} \tag{3.196}$$

The resolution of unity can be also equivavently written in the form

$$1 = |0\rangle\,\langle 0| + \sum_{n=1}^{\infty} \frac{1}{n!} \prod_{i=1}^{n} \int \frac{d^3p_i}{(2\pi)^3 2\omega_{p_i}} \,|p_1, p_2, \dots, p_n\rangle\,\langle p_1, p_2, \dots, p_n| \ . \tag{3.197}$$

With this any state $|\psi\rangle$ can be written as

$$|\psi\rangle = |0\rangle\,\langle 0\,|\psi\rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \prod_{i=1}^{n} \int \frac{d^3p_i}{(2\pi)^3 \omega_{p_i}} f_\psi^{(n)}(p_1, \dots, p_n)\,|p_1, \dots, p_n\rangle \ , \tag{3.198}$$

Here $f_\psi^{(n)}(p_1, \dots, p_n)$ are n-point wave functions that are (as can be easily checked) symmetric in all their arguments. Note that $|0\rangle$ is a vacuum state, i.e., state without any particle. The later is different from the vacuum in the 1st quantized theory (despite of a similar notation), which is the lowest-energy single-particle state. In particular, for the matrix element $\langle 0|\,\phi\,|\psi\rangle$ of the field operator we have

$$\begin{aligned}\langle 0|\,\phi\,|\psi\rangle &= \langle 0|\,\phi^+(x)\,|\psi\rangle \\ &= \langle 0| \int \frac{d^3k}{(2\pi)^3 2\omega_{p_k}} e^{-ikx} a(k) \int \frac{d^3p}{(2\pi)^3 2\omega_p} |p\rangle\, f_\psi^{(1)}(p) \\ &= \int \frac{d^3k d^3p}{(2\pi)^6 2\omega_k 2\omega_p} e^{-ikx} f_\psi^{(1)}(p) 2\omega_k (2\pi)^3 \delta(\boldsymbol{k}-\boldsymbol{p}) \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-ikx} f_\psi^{(1)}(k) \\ &= \int \frac{d^4k}{(2\pi)^4} e^{-ikx} \delta(\boldsymbol{k}-\boldsymbol{p})\theta(k) f_\psi^{(1)}(k) \\ &= \psi^{(1)}(x)\,. \end{aligned} \tag{3.199}$$

Note that $z\,\langle\psi|\,\hat{\phi}\,|0\rangle$ also satisfies K-G equation but for negative energies.

On the other hand, the field satisfies the equation $(\partial^2 + m^2)\hat{\phi} = 0$ we can write

$$\langle 0|\,(\partial^2 + m^2)\hat{\phi}\,|\psi\rangle = 0\,, \tag{3.200}$$

which implies that $(\partial^2 + m^2)\,\langle 0|\,\hat{\phi}\,|\psi\rangle = (\partial^2 + m^2)\psi^{(1)} = 0$. So, $\psi^{(1)}$ correctly satisfies the Klein–Gordon equation, as the scalar-particle wave function should. This provides a very important lesson. Field itself must satisfy the 1st quantized wave equation because its single-particle matrix element satisfies it. This will give us a useful guide for selecting logically consistent field theories.

'First quantization' means $x, p, \cdots \rightarrow \hat{x}, \hat{p}, \ldots$. 'Second quantization' means $\psi \rightarrow \hat{\phi}$.

3.7 Quantization of the Dirac Field

Dirac field describes Fermions with spin $\frac{1}{2}$ (e.g., for instance, electron or positron), which (as will be seen shortly) posses antisymmetric statistics (justifying thus the spin-statistics theorem). In order to quantize this field, we can start with a classical field Lagrangian. Since we know that the field must satisfy the 1st quantized wave equation (so as to yield the correct one-particle state equation) we should use the Lagrangian that can reproduce Dirac's equation. The later has the form

$$L = \int d^3\boldsymbol{x}\; \underbrace{\overline{\psi}(x)}_{\psi^\dagger(x)\gamma^0} \left(i\gamma^\mu \partial_\mu - m\right)\psi(x)\,. \tag{3.201}$$

In analogy with the complex scalar field we treat $\psi(x)$ and $\overline{\psi}(x)$ as independent fields. Now, we wish to compute fields canonically conjugate to $\psi(x)$ and $\overline{\psi}(x)$. Following the usual recipe, we get

$$\pi_\alpha(x) \equiv \chi_\alpha(x) = \frac{\delta L}{\delta \dot{\psi}_\alpha(x)} = \left(i\overline{\psi}(x)\gamma^0\right)_\alpha = i\overline{\psi}_\beta(x)\gamma^0_{\beta\alpha} = i\psi^\dagger(x)_\alpha\,. \tag{3.202}$$

Note that $\dot{\overline{\psi}}(x)$ does not appear in the Lagrangian (!) and so there is no field conjugated to $\overline{\psi}(x)$. At this stage one may expect, that the commutation rule is of the form

$$\left[\psi_\alpha(\boldsymbol{x}, t), \chi_\beta(\boldsymbol{y}, t)\right] = i\delta(\boldsymbol{x}-\boldsymbol{y})\delta_{\alpha\beta}\,, \tag{3.203}$$

or equivalently

$$\left[\psi_\alpha(\boldsymbol{x},t),\psi^\dagger_\beta(\boldsymbol{y},t)\right] = \delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{y})\,. \tag{3.204}$$

We shall, however, see shortly that one has to use anticommutators for the Dirac theory.

The Hamiltonian is given by

$$\begin{aligned} H &= \int \mathrm{d}^3\boldsymbol{x}(\chi\dot{\psi} + \underbrace{\overline{\chi}\dot{\overline{\psi}}}_{=0}) - L \\ &= \int \mathrm{d}^3\boldsymbol{x}\left(i\overline{\psi}\gamma^0\dot{\psi} - \overline{\psi}i\gamma^\mu\partial_\mu\psi + m\overline{\psi}\psi\right) \\ &= \int \mathrm{d}^3\boldsymbol{x}\left(i\overline{\psi}\boldsymbol{\gamma}\boldsymbol{\nabla}\psi + m\overline{\psi}\psi\right) \\ &= \int \mathrm{d}^3\boldsymbol{x}\,\psi^\dagger\left(i\gamma^0\boldsymbol{\gamma}\boldsymbol{\nabla} + m\gamma^0\right)\psi\,. \end{aligned} \tag{3.205}$$

Here $\boldsymbol{\gamma} = (\gamma_1,\gamma_2,\gamma_3)$. By employing the properties of γ-matrices and integrating per-partes we see from the last identity that the Hamiltonin is clearly Hermitian (as it should). With the Hamiltonian we can write in the Heisenberg-picture the standard Heisenberg equations of motion

$[AB,C] = A[B,C] + [A,C]B = A\{B,C\} - \{A,C\}B.$

$$\dot{\psi} = i\,[H,\psi] = i\int \mathrm{d}^3\boldsymbol{x}\,\left[i\overline{\psi}\boldsymbol{\gamma}\boldsymbol{\nabla}\psi + m\overline{\psi}\psi,\psi\right]\,, \tag{3.206}$$

and similarly for $\overline{\psi}$.

Now, because the fields satisfy the Dirac (and hence the Klein–Gordon) equation, the solution can be written in the form

$$\begin{aligned} \psi(x) &= \sum_{\boldsymbol{p}}\left(A(p)e^{-ipx} + B(p)e^{ipx}\right) \\ &= \sum_{\boldsymbol{p}}\left(\sum_{\lambda\in\{1,2\}} a(p,\lambda)u(p,\lambda)e^{-ipx} + \sum_{\lambda\in\{1,2\}} b^*(p,\lambda)v(p,\lambda)e^{ipx}\right) \\ &= \sum_{\boldsymbol{p}}\sum_{\lambda}\left(a(p,\lambda)u(p,\lambda)e^{-ipx} + b^*(p,\lambda)v(p,\lambda)e^{ipx}\right)\,. \end{aligned} \tag{3.207}$$

Here, λ is the helicity index or alternatively spin index. This immediately implies that

Here $(\not{p} - m)u(p,\lambda) = 0$ and $(\not{p} + m)v(p,\lambda) = 0$.

$$\overline{\psi}(x) = \psi^\dagger(x)\gamma^0 = \sum_{\boldsymbol{p}}\sum_{\lambda}\left(b(p,\lambda)\overline{v}(p,\lambda)e^{-ipx} + a^*(p,\lambda)\overline{u}(p,\lambda)e^{ipx}\right)\,. \tag{3.208}$$

Let us recall that we have already derived the following orthogonality relations:

$$\overline{u}(p,\lambda)u(p,\lambda') = 2m\delta_{\lambda\lambda'}, \tag{3.209}$$
$$\overline{v}(p,\lambda)v(p,\lambda') = -2m\delta_{\lambda\lambda'}, \tag{3.210}$$
$$\overline{u}(p,\lambda)v(p,\lambda') = 0, \tag{3.211}$$
$$\overline{v}(p,\lambda)u(p,\lambda') = 0, \tag{3.212}$$
$$\overline{u}(p,\lambda)\gamma^0 u(p,\lambda') = 2\omega_p\delta_{\lambda\lambda'}, \tag{3.213}$$
$$\overline{v}(p,\lambda)\gamma^0 v(p,\lambda') = 2\omega_p\delta_{\lambda\lambda'}. \text{ (No change of sign!)} \tag{3.214}$$

To quantize, we promote $\psi(x)$ and $\overline{\psi}(x)$ to operators (along with a, a^* and b, b^*, which are promoted to the $a, a^\dagger$ and $b, b^\dagger$). This also promotes Hamiltonian to the operator. The above mode expansion of ψ and $\overline{\psi}$ allows to formulate H entirely in the language of a, a^* and b, b^*. Indeed, H can be then rewritten in terms of "creation" and "annihilation" operators as

$$H = \sum_{\boldsymbol{p},\lambda} \omega_p \left(a^\dagger_\lambda(p) a_\lambda(p) - b_\lambda(p) b^\dagger_\lambda(p) \right) . \tag{3.215}$$

If we would now use the presumed commutation rules for ψ and $\overline{\psi}$, then it can be easily checked that $a, a^\dagger$ and $b, b^\dagger$ would obey the commutation rules as well. This leads, however, to the problem since the ensuing Hamiltonian would not be positive (even worse, it would be unbounded from below). The particles a and b contribute opposite sign to the energy, which means that this theory would not admit a stable ground state. If we however assume that anticommutation relations hold:

$$\left\{a_\lambda(p), a^\dagger_{\lambda'}(p')\right\} = \delta_{\lambda\lambda'}\delta_{\boldsymbol{p}\boldsymbol{p}'}, \tag{3.216}$$
$$\left\{b_\lambda(p), b^\dagger_{\lambda'}(p')\right\} = \delta_{\lambda\lambda'}\delta_{\boldsymbol{p}\boldsymbol{p}'}, \tag{3.217}$$
$$\left\{a^\dagger_\lambda(p), a^\dagger_{\lambda'}(p')\right\} = \{a_\lambda(p), b_{\lambda'}(p')\} = \cdots = 0\,, \tag{3.218}$$

the Hamiltonian can be written as

$$H = \sum_{\boldsymbol{p},\lambda} \omega_p \left(a^\dagger_\lambda(p) a_\lambda(p) + b^\dagger_\lambda(p) b_\lambda(p) \right) - \sum_{\boldsymbol{p}} 2\omega_p \delta_{\boldsymbol{p}\boldsymbol{p}'}\,. \tag{3.219}$$

Now that we needed only anticommutation rule for b particles, but because $\psi(x)$ and $\overline{\psi}(x)$ involve the sum over a and $b^\dagger$ as well as $a^\dagger$ and b, we must take a and $a^\dagger$ to have anticommutation rules as well in order to have anticommutation rule $\psi(x)$ and $\overline{\psi}(x)$. The later is, in turn, consistent with various physical requirements (e.g. to get correct Dirac's equation of motion for Heisenberg fields). It can be shown, that with the rules (3.218) the following anticommutation relations hold for the canonically conjugated fields

$$\begin{aligned} &\left\{\psi_\alpha(\boldsymbol{x},t), \chi_\beta(\boldsymbol{y},t)\right\} = i\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{y}) \\ \Leftrightarrow\quad &\left\{\psi^\dagger_\alpha(\boldsymbol{x},t), \psi_\beta(\boldsymbol{x},t)\right\} = \delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{y})\,, \end{aligned} \tag{3.220}$$

and also

$$\{\psi, \psi\} = \left\{\psi^\dagger, \psi^\dagger\right\} = 0\,. \tag{3.221}$$

As a byproduct we might notice that (3.220) implies that

$$\left\{\psi_\beta(\boldsymbol{x},t),\overline{\psi}_\delta(\boldsymbol{y},t)\right\} = \gamma^0_{\beta\delta}\delta(\boldsymbol{x}-\boldsymbol{y})\,. \tag{3.222}$$

Let us now see that these anticommutation rules provide expected equation of motion. To this end we need to compute the Heisenberg equations of motion for fields, i.e.

$$\dot{\psi} = i\,[H,\psi] = i\int \mathrm{d}^3\boldsymbol{x}'\,\left[i\overline{\psi}\boldsymbol{\gamma}\boldsymbol{\nabla}\psi + m\overline{\psi}\psi,\psi\right]\,. \tag{3.223}$$

First, let us calculate (consider equal time is implicitly used)

$$\begin{aligned}\left[\overline{\psi}(x')\psi(x'),\psi_\beta(x)\right] &= \overline{\psi}_\alpha(x')\{\psi_\alpha(x'),\psi_\beta(x)\} - \{\overline{\psi}_\alpha(x'),\psi_\beta(x)\}\psi_\alpha(x')\\ &= -\,\delta(\boldsymbol{x}'-\boldsymbol{x})\gamma^0_{\beta\alpha}\psi_\alpha(x')\,,\end{aligned} \tag{3.224}$$

(we have used the identity $[AB,C] = A\{B,C\} - \{A,C\}B$). The second commutator we need is

$$\begin{aligned}\left[i\overline{\psi}_\alpha(x')\boldsymbol{\gamma}_{\alpha\delta}\boldsymbol{\nabla}'\psi_\delta(x'),\psi_\beta(x)\right] &= -i\{\overline{\psi}_\alpha(x'),\psi_\beta(x)\}\boldsymbol{\gamma}_{\alpha\delta}\boldsymbol{\nabla}'\psi_\delta(x')\\ &= -i\delta(\boldsymbol{x}'-\boldsymbol{x})\boldsymbol{\gamma}^0_{\beta\alpha}\boldsymbol{\gamma}_{\alpha\delta}\boldsymbol{\nabla}'\psi_\delta(x').\end{aligned} \tag{3.225}$$

This leads to

$$\begin{aligned}\dot{\psi}_\beta &= i\int \mathrm{d}^3\boldsymbol{x}'\,\left[i\overline{\psi}\boldsymbol{\gamma}\boldsymbol{\nabla}\psi + m\overline{\psi}\psi,\psi\right]\\ &= i\int \mathrm{d}^3\boldsymbol{x}'\,\left\{i(-(\gamma^0\boldsymbol{\gamma})_{\beta\delta}\boldsymbol{\nabla}'\psi_\delta(x')) - m\gamma^0_{\beta\alpha}\psi_\alpha(x')\right\}\delta(\boldsymbol{x}'-\boldsymbol{x})\\ &= i\left[-i(\gamma^0\boldsymbol{\gamma}\boldsymbol{\nabla}\psi(x))_\beta - m(\gamma^0\psi(x))_\beta\right]\\ &= \left[\gamma^0\boldsymbol{\gamma}\boldsymbol{\nabla}\psi(x) - im\gamma^0\psi(x)\right]_\beta\,.\end{aligned} \tag{3.226}$$

Which is the Heisenberg equation of motion. Multiplying this by $i\gamma^0$ we get

$$\begin{aligned}&i\gamma^0\dot{\psi}(x) = i\boldsymbol{\gamma}\boldsymbol{\nabla}\psi(x) + m\psi(x)\\ \Leftrightarrow\quad &i(\gamma^0\partial_0\psi(x) - \boldsymbol{\gamma}\boldsymbol{\nabla}\psi(x)) - m\psi(x) = 0\\ \Leftrightarrow\quad &(i\gamma^\mu\partial_\mu - m)\psi(x) = 0\,.\end{aligned} \tag{3.227}$$

Hence, the Heisenberg equation of motion is indeed the Dirac equation.

Note that when the field obeys the equation of motion, the Lagrangian vanishes ($L = 0$). In such a case we can simplify the formula for the Hamiltonian to the form

$$H = \int \mathrm{d}^3\boldsymbol{x}\, i\overline{\psi}(x)\gamma^0\dot{\psi}(x)\,. \tag{3.228}$$

At this point we can introduce the tensor

$$T^{\mu\nu} = i\overline{\psi}(x)\gamma^\mu\partial^\nu\psi(x).\ \ \text{(This is not symmetric in } \mu\nu!)\,, \tag{3.229}$$

and define the 4-vector P^ν as

$$P^\nu = \int d^3\boldsymbol{x}\, T^{0\nu}(x)\,, \tag{3.230}$$

which explicitly means that

$$\begin{aligned} &\boldsymbol{P} = -i \int d^3\boldsymbol{x}\, \overline{\psi}(x)\gamma^0 \boldsymbol{\nabla}\psi(x) \\ \Leftrightarrow \quad &\boldsymbol{P}^i = i \int d^3\boldsymbol{x}\, \psi^\dagger(x)\partial^i\psi(x)\,. \end{aligned} \tag{3.231}$$

It can be checked that P^μ satisfies the relation

$$e^{iPa}\psi(x')e^{-iPa} = \psi(x+a)\,, \tag{3.232}$$

which in turn implies that it must be a generator of space-time translations, or in other words it is the total 4-momentum.

Relation (3.232) can be proved most easily by looking on the infinitesimal version of the later. Namely, for $|a| \ll 1$ we have to the linear order in a^μ

$$\begin{aligned} \partial_\mu\psi(x) &= i\left[P_\mu, \psi(x)\right]\,, \\ \partial_\mu\overline{\psi}(x) &= i\left[P_\mu, \overline{\psi}(x)\right]\,. \end{aligned} \tag{3.233}$$

It can be easily checked that (3.231) satisfies both relations (3.233).

Furthermore, already from 1st quantization we know that we have a conserved current (now the wave function is replaced with quantum field)

$$J_\mu(x) = \overline{\psi}\gamma_\mu\psi(x)\,. \tag{3.234}$$

Again as in 1st quantization, by using the equation of the motion one can show that

$$\partial_\mu J^\mu(x) = 0\,, \tag{3.235}$$

and that

$$Q \equiv \int d^3\boldsymbol{x}\, J^0(\boldsymbol{x}, t)\,, \tag{3.236}$$

is such that

$$[H, Q] = 0\,. \tag{3.237}$$

i.e., Q is associated conserved charge. Note that in the 1st quantization Q represented total probability.

4-Momentum

If we carry out the calculations of P^ν

$$P^\nu = \int d^3\boldsymbol{x}\, T^{0\nu}(x)\,, \tag{3.238}$$

we find that

$$P^\nu = \sum_{\boldsymbol{p},\lambda} p^\nu \{a_\lambda^\dagger(p)a_\lambda(p) - b_\lambda(p)b_\lambda^\dagger(p)\}\,. \tag{3.239}$$

With this we can easily see that

$a^\dagger$ and $b^\dagger$ create particles with the corresponding momenta (and helicity).

a and b annihilate particles with corresponding momenta.

$$\begin{aligned} P^\nu a_\lambda^\dagger(p) &= [P^\nu, a_\lambda^\dagger(p)] + a_\lambda^\dagger(p)P^\nu \\ &= a_\lambda^\dagger(p)(P^\nu + p^\nu). \end{aligned} \tag{3.240}$$

And similarly for $P^\nu b_\lambda^\dagger(p) = b_\lambda^\dagger(p)(P^\nu + p^\nu)$. Hence, $a_\lambda^\dagger(p)$ and $b_\lambda^\dagger(p)$ create particles with the corresponding momenta and helicity.

It is easy to see that we also get

$$\begin{aligned} P^\nu a_\lambda(p) &= a_\lambda(p)(P^\nu - p^\nu)\,, \\ P^\nu b_\lambda(p) &= b_\lambda(p)(P^\nu - p^\nu)\,. \end{aligned} \tag{3.241}$$

Thus, $a_\lambda(p)$ and $b_\lambda(p)$ annihilate particles with the corresponding momenta and helicity.

Again, we wish $|0\rangle$ to have zero energy and momenta, i.e. $\hat{p}^\nu\,|0\rangle = 0$. This can be done via **normal ordering** (which subtracts vacuum energy):

$$P^\nu = \int \mathrm{d}^3\boldsymbol{x}\, T^{0\nu}, \quad \text{where} \quad T^{\mu\nu} = i : \overline{\psi}(x)\gamma^\mu\partial^\nu\psi(x) : \,. \tag{3.242}$$

In the context of Dirac fields, normal ordering requires that we place annihilation operator to the right and creation operators to the left, but insert a factor of -1 for each operator interchange. E.g.

$$\begin{aligned} P^\nu &= : \sum_{p,\lambda} p^\nu \{a_\lambda^\dagger(p)a_\lambda(p) - b_\lambda(p)b_\lambda^\dagger(p)\} : \\ &= \sum_{p,\lambda} p^\nu \{a_\lambda^\dagger(p)a_\lambda(p) + b_\lambda^\dagger(p)b_\lambda(p)\}\,. \end{aligned} \tag{3.243}$$

This should be compared with non normally ordered p^ν

$$P^\nu = \sum_{p,\lambda} p^\nu [a_\lambda^\dagger(p)a_\lambda(p) + b_\lambda^\dagger(p)b_\lambda(p) - \delta_{\boldsymbol{pp}}\delta_{\lambda\lambda}]\,. \tag{3.244}$$

By noting that

$$\begin{aligned} -\sum_{\boldsymbol{p},\lambda} p^\nu \delta_{\boldsymbol{pp}}\delta_{\lambda\lambda} &= -2\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_p} p^\nu (2\pi)^3 2\omega_p V \\ &= -2\int \mathrm{d}^3\boldsymbol{p}\, p^\nu V = \langle 0|P^\nu|0\rangle\;. \end{aligned} \tag{3.245}$$

we see that

$$P^\nu \;\to : P^\nu : = \; P^\nu - \langle 0|P^\nu|0\rangle\;. \tag{3.246}$$

Note that for spatial components $\langle 0|\boldsymbol{P}|0\rangle = 0$, and so we need the normal ordering only for $P^0 = H$.

Strictly speaking the normal ordering should be now employed for any composite operator. In particular, we should also normally order the conserved charge Q (cf. Eq. (3.236)). In this case we have

$$Q = : \int d^3\boldsymbol{x}\, J^0(x) : = : \int d^3\boldsymbol{x} \psi^\dagger(x)\psi(x) : . \tag{3.247}$$

We can again show that

$$Q = \sum_{\boldsymbol{p},\lambda} \left[a_\lambda^\dagger(p) a_\lambda(p) - b_\lambda^\dagger(p) b_\lambda(p) \right] . \tag{3.248}$$

AWith this we have

$$\begin{aligned} Q a_\lambda^\dagger(p) &= a_\lambda^\dagger(p) Q + \left[Q, a_\lambda^\dagger(p) \right] = a_\lambda^\dagger(p)(Q+1) , \\ Q b_\lambda^\dagger(p) &= b_\lambda^\dagger(p) Q + \left[Q, b_\lambda^\dagger(p) \right] = b_\lambda^\dagger(p)(Q-1) . \end{aligned} \tag{3.249}$$

Let us remind that $a_\lambda^\dagger(p)$ and $b_\lambda^\dagger(p)$ create particles of type a and b with the equal 4-momenta (hence also equal rest mass) and helicity. In this respect particles a and b are identical. Now from (3.249) we see that they are not entirely identical because $a_\lambda^\dagger(p)$ and $b_\lambda^\dagger(p)$ create particles with opposite charge. It will be seen that Q can be identified with the electric charge. Consequently a and b are antiparticles to each other. Conventionally we will call particles of the type a as **particles** and particles of type b as **antiparticles**.

Spin-statistics connection

Due to the canonical relations $\{a_p^\dagger, a_q^\dagger\} = \{a_p, a_q\} = \{b_p^\dagger, b_q^\dagger\} = \{b_p, b_q\} = 0$ we see that only possible occupation number of a state $|\cdots p, \lambda \cdots\rangle$ is 0 or 1. Indeed, take

$$|p_1, \lambda_1 \cdots, \cdots p_n, \lambda_n \cdots\rangle = |\cdots \underbrace{0}_{p,\lambda} \cdots\rangle . \tag{3.250}$$

Now by action with creation operator

$$a_\lambda^\dagger(p)\, |\cdots 0 \cdots\rangle = C\, |\cdots \underbrace{1}_{p,\lambda} \cdots\rangle , \tag{3.251}$$

but

$$a_\lambda^\dagger(p)\, |\cdots 2 \cdots\rangle = \frac{1}{\sqrt{2}} \underbrace{a_\lambda^\dagger(p) a_\lambda^\dagger(p)}_{=0} |\cdots 0 \cdots\rangle . \tag{3.252}$$

Thus, we cannot create a state with the occupation number $n_p \geq 2$. Similarly for particle with an opposite-charge state (antiparticle state) generated $b_\lambda^\dagger(p)$.

Field theory of Dirac fields (spin $\frac{1}{2}$) automatically prescribes (or implies) Pauli's exclusion principle, i.e. Dirac fields need to obey Fermi-Dirac statistics.

3.8 Symmetry and Conserved Currents, Noether's Theorem

In field theory, symmetries and conservation laws are related. Our starting point is Lagrangian

$$L = \int \mathrm{d}^3\boldsymbol{x}\,\mathcal{L}(\phi_r(x), \partial_\mu \phi_r(x))\,, \tag{3.253}$$

with the ensuing equation of motion

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_r} = 0\,. \tag{3.254}$$

Any continuous symmetry transformation, which leaves the Lagrangian $\mathcal{L}$ invariant implies the existence of a current satisfying equation of continuity

$$\partial_\mu J^\mu(x) = 0\,, \tag{3.255}$$

This in turn defines the charge

$$Q = \int \mathrm{d}^3\boldsymbol{x}\,J^0(\boldsymbol{x}, t)\,, \tag{3.256}$$

that is conserved, because the surface term at infinity is negligibly small, i.e.

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int \mathrm{d}^3\boldsymbol{x}\,\partial_0 J^0 = \int \mathrm{d}^3\boldsymbol{x}\,\boldsymbol{\nabla}\cdot\boldsymbol{J} = 0\,. \tag{3.257}$$

Particularly important is the case when the Lagrangian is invariant due to internal symmetry, that acts on the internal indices r, i.e. the action is invariant under transformation

$$\phi_r(x) \to \phi'_r(x) = \phi_r(x) + \delta\phi_r(x)\,. \tag{3.258}$$

Let us recall that for Lorentz group we had

$$\begin{aligned} \psi(x) \xrightarrow{\boldsymbol{L}} \psi'(x) &= S(\boldsymbol{L})\psi(\boldsymbol{L}^{-1}x) = S_{\alpha\beta}(\boldsymbol{L})\psi_\beta(\boldsymbol{L}^{-1}x) \\ &= \left(e^{i\omega_{\mu\nu}\boldsymbol{M}^{\mu\nu}}\right)_{\alpha\beta} \psi_\beta(\boldsymbol{L}^{-1}x)\,. \end{aligned} \tag{3.259}$$

Now for $\phi_r(x) \xrightarrow{\boldsymbol{G}} \phi'_r(x)$:

$$\begin{aligned} \phi_r(x) \xrightarrow{\boldsymbol{G}} \phi'_r(x) &= S(\mathbf{G})_{rq}\phi_q(x) = \left(e^{i\epsilon^a \boldsymbol{T}^a}\right)_{rq} \phi_q(x) \\ &\overset{|\epsilon|\ll 1}{=} (\mathbf{1} + i\epsilon^a \boldsymbol{T}^a)_{rq}\,\phi_q(x) \\ &= \phi_r(x) + \underbrace{i\epsilon^a \boldsymbol{T}^a_{rq}\phi_q(x)}_{\delta\phi_r(x)}\,. \end{aligned} \tag{3.260}$$

ϵ^a are (x-independent) small parameters and $\boldsymbol{T}^a$ are generators of the Lie algebra of the group $\boldsymbol{G}$ with

$$\left[\boldsymbol{T}^a, \boldsymbol{T}^b\right] = ic^{abc}\boldsymbol{T}^c\,. \tag{3.261}$$

Here c^{abc} are the so-called structure constants of $\boldsymbol{G}$. If the Lagrangian is unchanged under the action of the group $\boldsymbol{G}$, then $\delta\mathcal{L} = 0$ also for infinitesimal changes $\delta\phi_r(x)$ specified above, i.e.

$$0 = \delta\mathcal{L}_x = \int \mathrm{d}y \frac{\delta\mathcal{L}_x}{\delta\phi_r(y)}\delta\phi_r(y)\,. \tag{3.262}$$

Functional derivative term can be explicitly written as

$$\begin{aligned}\frac{\delta\mathcal{L}_x}{\delta\phi_r(y)} &= \frac{\mathrm{d}}{\mathrm{d}\epsilon_r}\mathcal{L}_x(\phi_r + \epsilon_r\delta(x-y), \partial_\mu\phi_r + \partial_\mu\epsilon_r\delta(x-y))\bigg|_{\epsilon_r=0} \\ &= \frac{\partial\mathcal{L}_x}{\partial\phi_r(x)}\delta(x-y) + \frac{\partial\mathcal{L}_x}{\partial(\partial_\mu\phi_r(x))}\partial_\mu\delta(x-y)\,.\end{aligned} \tag{3.263}$$

So, when we employ the equations of motion we obtain

$$\begin{aligned}0 = \delta\mathcal{L} &= \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r(x))}\right]\delta\phi_r(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r(x))}\partial_\mu\delta\phi_r(x) \\ &= \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r(x))}\delta\phi_r(x)\right] \\ &= \epsilon^a\partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r(x))}i\boldsymbol{T}^a_{rq}\phi_q(x)\right].\end{aligned} \tag{3.264}$$

Since this is true for arbitrary ϵ^a we can identify a conserved 4-vector quantity (conserved current) satisfying continuity equation $\partial_\mu J^\mu(x) = 0$. This suggests defining a conserved current J^μ as

$$J^a_\mu(x) = -i\frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi_r)}\boldsymbol{T}^a_{rq}\phi_q(x)\,. \tag{3.265}$$

The conserved charges are given by

$$Q^a = \int \mathrm{d}^3\boldsymbol{x}\,J^a_0(x)\,, \tag{3.266}$$

and are (apart from being conserved — $\dot{Q}^a = 0$) also generators of the symmetry group. Currents $J^a_\mu(x)$ and charges Q^a are known as **Noether currents** and **charges**, respectively.

Let us further note that

$$J^{a0} = -i\pi_r\boldsymbol{T}^a_{rq}\phi_q\,. \tag{3.267}$$

From

$$[\phi_a(\boldsymbol{x},t), \pi_b(\boldsymbol{y},t)] = i\delta_{ab}\delta(\boldsymbol{x}-\boldsymbol{y})\,. \tag{3.268}$$

it can be then deduced that

Eq. (3.269) is an example of a so-called *current algebra*.

$$\left[J^{a0}(\boldsymbol{x},t), J^{b0}(\boldsymbol{y},t)\right] = ic^{abc}J^{c0}(\boldsymbol{x},t)\delta(\boldsymbol{x}-\boldsymbol{y})\,. \tag{3.269}$$

By integrating this realation twice on LHS and RHS we obtain

$$\begin{aligned}\int \mathrm{d}^3\boldsymbol{x}\mathrm{d}^3\boldsymbol{y}\left[J^{a0}(\boldsymbol{x},t), J^{b0}(\boldsymbol{y},t)\right] &= ic^{abc}\int \mathrm{d}^3\boldsymbol{x}\mathrm{d}^3\boldsymbol{y}\,J^{c0}(x)\delta(\boldsymbol{x}-\boldsymbol{y}) \\ &= ic^{abc}Q^c\,.\end{aligned} \tag{3.270}$$

Which finally yields

$$\left[Q^a, Q^b\right] = i c^{abc} Q^c \,. \tag{3.271}$$

These generators act on a different representation space than T^a, T^b and T^c. While T's act on *internal space*, Q's act directly on the Hilbert space.

So, the charges satisfy the same algebra as the original generators T^a of the symmetry. As an example we consider Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}\left[(\partial_\mu\phi_1)^2 + (\partial_\mu\phi_2)^2\right] - \frac{1}{2}m^2\left(\phi_1^2+\phi_2^2\right) - \frac{1}{4}\lambda\left(\phi_1^2+\phi_2^2\right)^2 . \tag{3.272}$$

This is invariant under 2D-rotation transformation

$$\begin{aligned} \phi_1 &\rightarrow \phi_1' = \phi_1\cos\alpha - \phi_2\sin\alpha, \\ \phi_2 &\rightarrow \phi_2' = \phi_1\sin\alpha + \phi_2\cos\alpha \,. \end{aligned} \tag{3.273}$$

Which is equivalent to

$$\begin{pmatrix}\phi_1' \\ \phi_2'\end{pmatrix} = \begin{pmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix}. \tag{3.274}$$

Here the rotation matrix is element of SO(2) ~ U(1) group. In this case we can write

$$\boldsymbol{G}(\alpha) = e^{-i\alpha\boldsymbol{T}} \,, \tag{3.275}$$

and we can obtain the corresponding generator via the usual prescription

$$T = -i\,\frac{\mathrm{d}G(\alpha)}{\mathrm{d}\alpha}\bigg|_{\alpha=0} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \tag{3.276}$$

For $|\alpha| \ll 1$ we get

$$\phi_r' = \phi_r + i\alpha\boldsymbol{T}_{rq}\phi_q = (\mathbf{1} + i\alpha\boldsymbol{T})_{rq}\,\phi_q \,, \tag{3.277}$$

and hence in this case our current acquires the form

$$J_\mu = -i\frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi_r)}\boldsymbol{T}_{rq}\phi_q \,. \tag{3.278}$$

Knowing that only $\boldsymbol{T}_{12} = -\boldsymbol{T}_{21} \neq 0$ we can write

$$\begin{aligned} J_\mu &= -i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_1)}\boldsymbol{T}_{12}\phi_2 - i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_2)}\boldsymbol{T}_{21}\phi_1 \\ &= (\partial_\mu\phi_1)\phi_2 - (\partial_\mu\phi_2)\phi_1 \,. \end{aligned} \tag{3.279}$$

The previous analysis can be alternatively (and conveniently) formulated in terms of complex (or non-Hermitian) fields defined as

$$\phi = \frac{1}{\sqrt{2}}\left(\phi_1 + i\phi_2\right), \tag{3.280}$$

$$\phi^\dagger = \frac{1}{\sqrt{2}}\left(\phi_1 - i\phi_2\right). \tag{3.281}$$

In this case we can rewrite our original Lagrangian as

$$\mathcal{L} = \left(\partial_\mu\phi^*\right)\left(\partial^\mu\phi\right) - m^2(\phi^*\phi) - \lambda(\phi^*\phi)^2. \tag{3.282}$$

This is clearly invariant under

$$\begin{aligned} \phi &\rightarrow \phi' = e^{i\alpha}\phi\,, \\ \phi^* &\rightarrow \phi^{*\prime} = e^{-i\alpha}\phi^*\,, \end{aligned} \tag{3.283}$$

which we can again rewrite as

$$\begin{pmatrix} \phi' \\ \phi^{*\prime} \end{pmatrix} = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \begin{pmatrix} \phi \\ \phi^* \end{pmatrix}. \tag{3.284}$$

The ensuing generator reads as

$$T = -i\,\frac{\mathrm{d}G(\alpha)}{\mathrm{d}\alpha}\bigg|_{\alpha=0} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.285}$$

With that in mind, we finally arrive at the ensuing conserved current:

$$\begin{aligned} J_\mu &= -i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_1)}\boldsymbol{T}_{11}\phi - i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_2)}\boldsymbol{T}_{22}\phi^* \\ &= i\left[(\partial_\mu\phi)\phi^* - (\partial_\mu\phi^*)\phi\right]\,. \end{aligned} \tag{3.286}$$

Compare this formula with (2.42), which we had for the Klein–Gordon particle. Interpretation is, however, very different. In the first-quantized version J_μ was a probability current with $J_0 = \rho$ representing probability density. In the present (field-theory) case J_μ represents a conserved current with J_0 being density of the conserved charge. The charge associated with the $U(1)$ symmetry is an *electric charge* and hence there is no surprise that conserved charge Q might be negative, positive or zero (as opposed to total probability in the first-quantized theory that must be positive and normalisable to unity).

Note also that ϕ_1 and ϕ_2 (or ϕ and ϕ^*) are degenerate in mass, because of the required U(1) ~ SO(2) symmetry.

Quantum Field Theory 2 | 4

4.1 Noether's Theorem Continued

Apart from the method used in the previous chapter there exists yet another quick way to conserved currents — the so-called **Noether's method** (1918).

Consider the global symmetry transformation

$$\phi(x) \to \phi'(x) = \phi(x) + \varepsilon\overline{\delta}\phi(x) = \phi(x) + (i\varepsilon_a T^a)\phi(x), \tag{4.1}$$

which leaves the Lagrangian density $\mathcal{L}$ invariant, i.e. $\delta\mathcal{L} = 0$. Here $\phi(x)$ is an arbitrary field in our theory and ε is a constant infinitesimal parameter.

We promote now ε to be a small x-dependent parameter, so we consider instead a general transformation

$$\phi \to \phi'(x) = \phi(x) + \varepsilon(x)\overline{\delta}\phi(x). \tag{4.2}$$

Generally, we call the transformations whose parameter ε is constant (not dependent on a position in spacetime) **global**, whereas the transformation with x-dependent parameter $\varepsilon(x)$ are called **local**.

Lagrange density (and hence action S) is not invariant under such transitions for general $\varepsilon(x)$, since the symmetry we are considering is only global symmetry. Since then action would be invariant for constant ε, its variation is proportional to the derivative of $\varepsilon(x)$ and so it can be written in a general form

$$\delta S = \int d^4x[-J_\alpha(x)]\partial^\alpha\varepsilon(x). \tag{4.3}$$

for some current J_α. The current defined in this way is always conserved if the equations of motion are obeyed. The sign of the current is just a convention.

Indeed, when the equations of motion are obeyed, the action is stationary under any variation and in particular under variation given by (4.2). Thus, when the equations of motion are obeyed, i.e. when $\delta S = 0$ in (4.3) is zero for any parameters $\varepsilon(x)$ from which follows that

$$\partial_\alpha J^\alpha = 0. \tag{4.4}$$

As a simple exercise, we will show that for usual charged scalar fields this gives the same current as we have obtained in the last semester.

Here the Lagrangian equals

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi = -\phi^* (\partial^2 + m^2)\phi \,, \tag{4.5}$$

(modulo irrelevant four-divergence term). This describes two non-hermitian particles such as $\pi^\pm$ which are not their own antiparticles.

We have already seen in the previous chapter that the Lagrangian can be equivalently rewritten in terms of two real fields ϕ_1 and ϕ_2 that are related to ϕ via the relation

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \,. \tag{4.6}$$

Ensuing action equals to

$$S = -\sum_{i=1}^{2} \int \mathrm{d}^4 x \left(\frac{1}{2}\phi_i(\partial^2 + m^2)\phi_i \right) = -\int \mathrm{d}^4 x (\phi^*(\partial^2 + m^2)\phi) \,. \tag{4.7}$$

In this case, the Lagrangian is invariant under $\phi \to e^{i\alpha}\phi(x)$. To get Noether current, we promote $\alpha \to \alpha(x)$. For infinitesimal parameter, one then obtains $\phi(x) \to \phi(x) + i\alpha(x)\phi(x)$, which implies

$$\begin{aligned}
\delta S &= -\int \mathrm{d}^4 x \left[(\phi(x) + i\alpha(x)\phi(x))^* (\partial^2 + m^2)(\phi(x) + i\alpha(x)\phi(x)) - \right. \\
&\quad \left. - \phi^*(x)(\partial^2 + m^2)\phi(x) \right] \\
&= -\int \mathrm{d}^4 x \phi^*(x) \partial^2 (+i\alpha(x)\phi(x)),
\end{aligned} \tag{4.8}$$

where we neglect pieces linear in $\alpha(x)$ (invariance under global symmetries). Continuing and integrating per partes we get

$$\begin{aligned}
\delta S &= -\int \mathrm{d}^4 x \phi^*(x) \left[i(\partial^2\alpha(x))\phi(x) + 2i\partial_\mu\alpha(x)\partial^\mu\phi(x) + i\alpha(x)\partial^2\phi(x) \right] \\
&= -\int \mathrm{d}^4 x \left[2i(\phi^*(x)\partial^\mu\phi(x))\partial_\mu\alpha - i\partial^\mu(\phi^*(x)\phi(x))\partial_\mu\alpha \right] \\
&= -i\int \mathrm{d}^4 x \left[\phi^*(x)\partial^\mu\phi(x) - \phi(x)\partial^\mu\phi^*(x) \right] \partial_\mu\alpha(x) \\
&= -\int \mathrm{d}^4 x J^\mu(x)\partial_\mu\alpha(x) \,,
\end{aligned}$$

where we have identified the conserved current. We see the agreement with Eq. (3.286).

Noether Charge for Dirac Field

Let us now consider free Dirac field, then the Lagrangian is equal to

$$\mathcal{L} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi \,. \tag{4.9}$$

This is invariant under transformation

$$\psi \to e^{i\alpha}\psi \;\simeq\; \psi(x) + i\alpha\psi(x)\,,$$
$$\overline{\psi} \to e^{-i\alpha}\overline{\psi} \;\simeq\; \overline{\psi}(x) - i\alpha\overline{\psi}(x)\,, \tag{4.10}$$

where α is a global constant. If we again proceed as $\alpha \to \alpha(x)$ we get

$$\begin{aligned}
\delta\mathcal{L} &= \overline{\psi}(1 - i\alpha(x))(i\gamma^\mu\partial_\mu - m)(1 + i\alpha(x))\psi - \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi \\
&= \overline{\psi}(i\gamma^\mu\partial_\mu)i(\alpha(x)\psi) \;=\; -\overline{\psi}\gamma^\mu\psi\partial_\mu\alpha(x) \\
&= -J^\mu(x)\partial_\mu\alpha(x)\,,
\end{aligned} \tag{4.11}$$

where the parts linear in $\alpha(x)$ were neglected. From this we get the equation for the Dirac current

$$\partial_\mu J^\mu(x) \;=\; \partial_\mu(\overline{\psi}\gamma^\mu\psi) \;=\; 0\,. \tag{4.12}$$

In the first quantization, we arrived at the same expression for the Dirac's probability current. Now we see that it does not reflect a conservation of probability but a conservation of the charge Q (which can be either positive or negative) and has a form

$$Q \;=\; \int d^3\boldsymbol{x}\, J^0(x)\,. \tag{4.13}$$

This charge is time-independent and relativistically invariant.

Up to now, we considered a "semi-classical level" in which the variables were not considered as operators. On the quantized level the operator that generates the corresponding transformation is

$$\hat{Q} = \int d^3\boldsymbol{x}\,\overline{\psi}\gamma^0\psi\,. \tag{4.14}$$

Strictly speaking one should consider operators in the normal-ordered form, but since they differ only by a complex number (infinity), it is not important in context of computing commutation relations.

It is easy to see that (at a given time t)

$$\left[\hat{Q}, \hat{\psi}(y)\right]_t = \int d^3\boldsymbol{x}\,\left[\hat{\overline{\psi}}(x)\gamma^0\hat{\psi}, \hat{\psi}(y)\right]_t = -\hat{\psi}(y)\,, \tag{4.15}$$

where we used

$$\begin{aligned}
\left[\overline{\psi}_\alpha(x)\gamma^0_{\alpha\beta}\psi_\beta, \psi_\gamma(y)\right]_t &= \overline{\psi}_\alpha(x)\gamma^0_{\alpha\beta}\left\{\psi_\beta(x), \psi_\gamma(y)\right\}_t \\
&\quad - \left\{\overline{\psi}_\alpha, \psi_\gamma(y)\right\}_t \gamma^0_{\alpha\beta}\psi_\beta(x) \\
&= -\gamma^0_{\alpha\beta}\gamma^0_{\gamma\alpha}\delta(\boldsymbol{x} - \boldsymbol{y})\psi_\beta(x) = \\
&= -\delta_{\beta\gamma}\delta(\boldsymbol{x} - \boldsymbol{y})\psi_\beta(x) = -\delta(\boldsymbol{x} - \boldsymbol{y})\psi_\gamma(x)\,.
\end{aligned}$$

Remind the identities

$$[AB, C] = A\{B, C\} - \{A, C\}B,$$
$$\left\{\psi_\alpha(x), \overline{\psi}_\beta(y)\right\} = \gamma^0_{\alpha\beta}\delta(\boldsymbol{x} - \boldsymbol{y}).$$

Thus we obtain the following equations (the second one could be computed similarly)

$$\hat{Q}\hat{\psi}(x) = \hat{\psi}(x)(\hat{Q} - 1), \tag{4.16}$$
$$\hat{Q}\hat{\overline{\psi}}(x) = \hat{\overline{\psi}}(x)(\hat{Q} + 1). \tag{4.17}$$

As an exercise, try to show that the following relations hold

$$\hat{Q}\hat{a}_\lambda(p) = \hat{a}_\lambda(p)(\hat{Q} - 1), \tag{4.18}$$
$$\hat{Q}\hat{b}^\dagger_\lambda(p) = \hat{b}^\dagger_\lambda(p)(\hat{Q} - 1), \tag{4.19}$$
$$\hat{Q}\hat{a}^\dagger_\lambda(p) = \hat{a}^\dagger_\lambda(p)(\hat{Q} + 1), \tag{4.20}$$
$$\hat{Q}\hat{b}_\lambda(p) = \hat{b}_\lambda(p)(\hat{Q} + 1). \tag{4.21}$$

Note that since $\hat{Q}\,|0\rangle = 0$, it follows from the second and the third equation (considering $|p,\lambda\rangle = \hat{b}^\dagger_\lambda(p)\,|0\rangle$ and $\hat{a}^\dagger_\lambda(p)\,|0\rangle$, respectively)

$$\hat{Q}\,|p,\lambda\rangle = \hat{Q}\hat{b}^\dagger_\lambda(p)\,|0\rangle = \hat{b}^\dagger_\lambda(p)(\hat{Q} - 1)\,|0\rangle = -\hat{b}^\dagger_\lambda(p)\,|0\rangle = -\,|p,\lambda\rangle\,, \tag{4.22}$$
$$\hat{Q}\,|p,\lambda\rangle = \hat{Q}\hat{a}^\dagger_\lambda(p)\,|0\rangle = \hat{a}^\dagger_\lambda(p)(\hat{Q} + 1)\,|0\rangle = \hat{a}^\dagger_\lambda(p)\,|0\rangle = |p,\lambda\rangle\,, \tag{4.23}$$

which clearly demonstrates the concept of particles and antiparticles with opposite charges.

Moreover, it can be easily shown that the transformation of the Dirac field (4.10) can be rewritten in terms of the conserved charge. In particular

$$\psi \longrightarrow e^{i\alpha}\psi = e^{-i\alpha\hat{Q}}\psi e^{i\alpha\hat{Q}} \tag{4.24}$$
$$\overline{\psi} \longrightarrow e^{-i\alpha}\overline{\psi} = \left(e^{-i\alpha\hat{Q}}\psi e^{i\alpha\hat{Q}}\right)^\dagger \gamma^0 = e^{-i\alpha\hat{Q}}\overline{\psi} e^{i\alpha\hat{Q}}\,. \tag{4.25}$$

To prove this it suffices to prove the identity to the lowest order.

$$\begin{aligned} \text{LHS} \quad & e^{i\alpha}\psi \simeq \psi + i\alpha\psi + O(\alpha^2), \\ \text{RHS} \quad & e^{-i\alpha\hat{Q}}\psi e^{i\alpha\hat{Q}} \simeq (\psi - i\hat{Q}\alpha\psi)(1 + i\hat{Q}\alpha) \\ & = \psi - i\alpha\left[\hat{Q},\psi\right] + O(\alpha^2) = \psi + i\alpha\psi + O(\alpha^2). \end{aligned}$$

Group composition law will then take the infinitesimal transformation to the full one.

In fact, the result (4.24)-(4.25) is very general. One can show that if

$$\phi \longrightarrow e^{i\alpha T}\phi \tag{4.26}$$

is the symmetry of the Lagrangian $\mathcal{L}$, then it can be equivalently rewritten in terms of Noether charges $\hat{Q}$ as

$$e^{i\alpha T}\phi = e^{-i\alpha\hat{Q}}\phi e^{i\alpha\hat{Q}}. \tag{4.27}$$

So symmetry is implemented via unitary transformation as it should be in Quantum Mechanics.

4.2 Space-time symmetries

So far we have dealt with internal symmetries, i.e. symmetries that act on internal indices of fields. Noether theorem is, however, versatile enough to identify conserved quantities related to space-time symmetries, i.e., symmetries that act directly on space-time "indices" rather than internal field indices.

Translationally invariant systems

Consider a system whose Lagrangian density is invariant (up to a 4-divergence) under the rigid space-time translation

$$\phi(x^\mu) \to \phi(x^\mu + a^\mu) \simeq \phi(x^\mu) + a^\mu \partial_\mu \phi(x^\mu), \tag{4.28}$$

where we used the Taylor expansion to the first order in a^μ.

Since we deal with x instead of discrete index, the Lagrangian density can after transformation differ from the original one by 4-divergence and still provide the same equations of motion. On the other hand, when one deals with *internal symmetries*, x-derivative do not enter in the transformation as everything is done at the same point.

We can derive the consequence of this by adopting the strategy as before, and promote a^μ to be position dependent. Then, if we assume that the field ϕ satisfies the equations of motion, then

$$\begin{aligned}
0 = \delta S &= \int d^4x \,[\mathcal{L}(\phi + a\partial\phi, \partial(\phi + a\partial\phi) - \mathcal{L}(\phi, \partial\phi)] \\
&= \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}(a^\mu \partial_\mu \phi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(a^\nu \partial_\nu \phi)\right] \\
&= \int d^4x \left\{\left[\frac{\partial\mathcal{L}}{\partial\phi}\partial_\nu\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\partial_\nu\phi)\right] a^\nu + \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right]\partial_\mu a^\nu\right\} \\
&= \int d^4x \left[(\partial_\mu\mathcal{L})a^\mu + \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right)\partial_\mu a^\nu\right].
\end{aligned}$$

Using the fact that the 4-divergence $\partial_\mu(\mathcal{L}a^\mu) = 0$ (invariance of the Lagrangian for translations with a constant a) one gets

$$\begin{aligned}
0 = \delta S &= \int d^4x \left[\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right)\partial_\mu a^\nu - \mathcal{L}\partial_\mu a^\mu\right] \\
&= \int d^4x \left[\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right) - \mathcal{L}\eta^\mu_\nu\right]\partial_\mu a^\nu.
\end{aligned}$$

If we define the **canonical** (or Noether) **energy-momentum** tensor by

$$T^\mu_{\cdot\,\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \mathcal{L}\eta^\mu_\nu, \tag{4.29}$$

then the invariance of the action under local translation induces the conservation of the tensor as

$$\partial_\mu T^\mu_{\cdot\,\nu} = 0. \tag{4.30}$$

We see that $T^0_{\cdot\,0} = T^{00}$ is just usual definition of the energy density (the hamiltonian density $\mathcal{H}$). Other components have interpretation

$$T^i_{\cdot\,0} = T^{i0} = -T_{i0} \quad \text{(where } i = 1,2,3\text{) is the energy flux,} \tag{4.31}$$
$$T^{0i} = -T^0_{\cdot\,i} \quad \text{(where } i = 1,2,3\text{) is the momentum density.} \tag{4.32}$$

For a real scalar field we find that the (doubly covariant) tensor $T_{\mu\nu} =$

As an example, check that for the Dirac field we get the electromagnetic tensor introduced on the lecture in the last semester.

$\eta_{\mu\alpha}T^{\alpha}_{\nu}$ is symmetric since

$$T_{\mu\nu} = \partial_\mu \phi \partial^\nu \phi - \eta_{\mu\nu}\mathcal{L}. \tag{4.33}$$

Summary of Noether procedures

In case of **internal symmetries**, we consider transformation of a type

$$\phi(x) \longrightarrow \phi'(x) = e^{i\alpha T}\phi(x), \tag{4.34}$$

where the generators satisfy the commutation relations $\left[T_i, T_j\right] = i f_{ijk} T_k$. These induce conserved currents J^μ_i and charges of a form

$$Q_i = \int d^3\boldsymbol{x}\, J^0_i(x), \tag{4.35}$$

where $\left[Q_i, Q_j\right] = i f_{ijk} Q_k$, i.e., Q_i's satisfy the same algebra as T_i's.

The transformation (4.34) can be equivalently rewritten as

$$e^{i\alpha T}\phi(x) = e^{-i\alpha_i Q^i}\phi(x) e^{i\alpha_i Q^i}. \tag{4.36}$$

On the other hand, **space-time symmetries**, such as translational invariance, represent transformations of the type

$$\phi(x^\nu) \longrightarrow \phi'(x) = \phi(x^\nu + a^\nu), \tag{4.37}$$

induce conservation of the energy-momentum tensor $T^{\mu\nu}$ and ensuing conserved vector "charge"

$$P^\nu = \int d^3\boldsymbol{x}\, T^{0\nu}(x), \tag{4.38}$$

which is equal to the total momentum of the system. Again, the transformation can be written as

$$\phi(x+a) = e^{iP^\nu a_\nu}\phi(x) e^{-iP^\nu a_\nu}. \tag{4.39}$$

4.3 Relativistically Invariant Commutation Relations

Consider first:

$$\begin{aligned} i\Delta(x) &= [\phi(x), \phi(0)], \\ i\Delta(\boldsymbol{x}, t) &= [\phi(\boldsymbol{x}, t), \phi(0)], \\ \frac{\partial}{\partial t}\Delta(\boldsymbol{x}, t) &= -i[\dot{\phi}(\boldsymbol{x}, t), \phi(0)]. \end{aligned} \tag{4.40}$$

When $t = 0$ we recover the equal time commutation relations, i.e.

$$\begin{aligned}
&\Delta(\boldsymbol{x},0) = -i[\phi(\boldsymbol{x},0),\phi(0)] = 0\,,\\
&\frac{\partial}{\partial t}\Delta(\boldsymbol{x},t)|_{t=0} = -i[\pi(\boldsymbol{x},0),\phi(0)] = -i(-i)\delta(\boldsymbol{x}) = -\delta(\boldsymbol{x})\,.
\end{aligned} \tag{4.41}$$

Which means that $\Delta(x)|_{t=0}$ and $\dot{\Delta}(x)|_{t=0} = -\delta(\boldsymbol{x})$. These can be considered as initial conditions for $\Delta(x)$. What equation satisfies $\Delta(x)$? Since

$$(\Box + m^2)\phi(x) = 0 \implies (\partial^2 + m^2)\Delta(x) = 0\,, \tag{4.42}$$

the solution is

$$\begin{aligned}
\Delta(x) &= -i\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\left[f(\boldsymbol{p})e^{-ipx} + h(\boldsymbol{p})e^{ipx}\right]\\
&= -i\sum_{\boldsymbol{p}}\left[f(\boldsymbol{p})e^{-ipx} + h(\boldsymbol{p})e^{ipx}\right].
\end{aligned} \tag{4.43}$$

We fix f and h by applying initial conditions on

$$\frac{\partial}{\partial t}\Delta(x) = -\sum_{\boldsymbol{p}}\omega_{\boldsymbol{p}}\left[f(p)e^{-ipx} - h(p)e^{ipx}\right] \tag{4.44}$$

and

$$\begin{aligned}
\Delta(\boldsymbol{x},0) &= -i\int \frac{\mathrm{d}\boldsymbol{p}^3}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\left[f(p)e^{ipx} + h(p)e^{-ipx}\right]\\
&= -i\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\left[f(p) + h(p_P)\right]e^{ipx}.
\end{aligned} \tag{4.45}$$

This then implies that

$$f(p) + h(p_P) = 0\,. \tag{4.46}$$

where $p_P^\mu = (p^0, -\boldsymbol{p})$. Thus

$$\begin{aligned}
\frac{\partial}{\partial t}\Delta(\boldsymbol{x},t)|_{t=0} &= -\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\left[f(p)e^{ipx} - h(p)e^{-ipx}\right]\omega_{\boldsymbol{p}}\\
&= -\delta(\boldsymbol{x})\,.
\end{aligned} \tag{4.47}$$

Employing the property of Fourier transforms we get

$$\frac{1}{2}[f(p) - h(p_P)] = 1\,. \tag{4.48}$$

Together with (4.46) this gives that $f(p) = 1$ and $h(p_P) = h(p) = -1$ and hence the general form of $\Delta(x)$ reads

$$\Delta(x) = -i\int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\left[e^{-ipx} - e^{ipx}\right]. \tag{4.49}$$

Note that $\Delta(x)$ is relativistically invariant, i.e. $\Delta(x) = \Delta(\boldsymbol{L}^{-1}x)$ for any Lorentz transformation $\boldsymbol{L}$ that maps each mass shell to itself, see Fig. 4.1. This can be shown as follows. First we see that the 4-

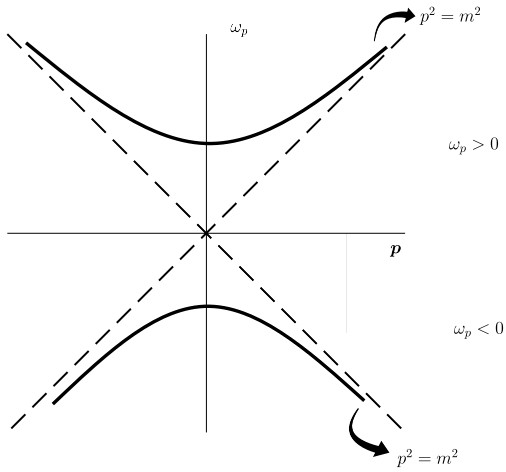


Figure 4.1: Positive and negative energy mass shells.

dimensional momentum measure is Lorentz invariant, indeed

$$\mathrm{d}^4 p' \;=\; \mathrm{d}^4 p \,|\det \boldsymbol{L}|\,, \tag{4.50}$$

here $p' = \boldsymbol{L}p$. The δ-function $\delta(p^2 - m^2)$ is also Lorentz invariant on the mass shell, because $\delta(p'^2 - m^2) = \delta(p^2 - m^2)$. Finally, by defining

$$\varepsilon(p_0) = \begin{cases} 1, & p^0 > 0 \\ -1, & p^0 < 0 \end{cases}. \tag{4.51}$$

we can easily see that also $\varepsilon(p_0)$ is invariant under *orthochronous* Lorentz transformations $\boldsymbol{L}$. Indeed, we know that the defining property of the Lorentz transformation implies that

$$\left(\boldsymbol{L}^0{}_0\right)^2 = \sum_i \left(\boldsymbol{L}^0{}_i\right)^2 + 1\,, \tag{4.52}$$

which gives

$$\left(\boldsymbol{L}^0{}_0\right)^2 > \sum_i \left(\boldsymbol{L}^0{}_i\right)^2\,. \tag{4.53}$$

Similarly, by using the dispersion relation

$$(p^0)^2 \;=\; \sum_i p_i^2 + m^2\,, \tag{4.54}$$

one obtains

$$(p^0)^2 \;>\; \sum_i p_i^2\,. \tag{4.55}$$

Putting all of these together we get

$$\left(\boldsymbol{L}^0{}_0 p^0\right)^2 > \sum_{i,j} (\boldsymbol{L}^0{}_i)^2 (p^j)^2 = \|\boldsymbol{L}^0\|^2 \|\boldsymbol{P}\|^2 \;\geq\; \left(\sum_i (\boldsymbol{L}^0{}_i) p^i\right)^2\,. \tag{4.56}$$

where the last inequality follows from the Schwarz inequality. Eq. (4.56) yields

$$|\boldsymbol{L}^0{}_0 p^0| \;>\; \left|\sum_i \left(\boldsymbol{L}^0{}_i\right) p^i\right|\,. \tag{4.57}$$

Now, $p'_0 = p^0 \boldsymbol{L}^0_{\ 0} + \sum_i p^i \boldsymbol{L}^0_{\ i}$, which means that the sign of p'_0 is fully determined by the $p^0 \boldsymbol{L}^0_{\ 0}$ term. If $\boldsymbol{L}^0_{\ 0} > 1$ (i.e. we consider only orthochronous Lorentz transformations), the signs of p'_0 and p_0 are the same and hence our $\varepsilon(p_0)$ is Lorentz invariant under such transformations.

Consequently

$$\mathrm{d}^4 p\, \varepsilon(p_0)\delta(p^2 - m^2)\,, \tag{4.58}$$

is Lorentz invariant. Let us further realize that

$$\begin{aligned} &\int \mathrm{d}^4 p\, \varepsilon(p_0)\delta(p^2 - m^2)\cdots \\ &= \int \mathrm{d}^4 p\, \varepsilon(p_0)\frac{1}{2\omega_{\boldsymbol{p}}}\left[\delta(p_0 - \omega_{\boldsymbol{p}}) + \delta(p_0 + \omega_{\boldsymbol{p}})\right]\cdots \\ &= \int \frac{\mathrm{d}^4 p}{2\omega_{\boldsymbol{p}}}\left[\delta(p_0 - \omega_{\boldsymbol{p}}) + \delta(p_0 + \omega_{\boldsymbol{p}})\right]\cdots \end{aligned} \tag{4.59}$$

With this we see that $\Delta(x)$ acquires the form

$$\Delta(x) = -i\int \frac{\mathrm{d}^4 p}{(2\pi)^3}\varepsilon(p_0)\delta(p^2 - m^2)e^{-ipx}\,. \tag{4.60}$$

By Lorentz transforming this expression we obtain

$$\Delta(\boldsymbol{L}^{-1}x) = -i\int \frac{\mathrm{d}^4 p}{(2\pi)^4}\varepsilon(p_0)\delta(p^2 - m^2)e^{-ip\boldsymbol{L}^{-1}x}\,, \tag{4.61}$$

But $(p, \boldsymbol{L}^{-1}x) = (\boldsymbol{L}p, x)$ since $(x, \boldsymbol{L}p) = (\boldsymbol{L}^T x, p) = (\boldsymbol{L}^{-1}x, p)$. Here we have used the fact that $\boldsymbol{L} \in \mathrm{SO}(3,1)$. More explicitly, we can use the index notation to write

$$\begin{aligned} x\boldsymbol{L}p &= x_\mu \boldsymbol{L}^\mu_{\ \nu} p^\nu = x_\mu (\boldsymbol{L}^{-1})_\nu^{\ \mu} p^\nu \\ &= p^\nu (\boldsymbol{L}^{-1})_\nu^{\ \mu} x_\mu = p_\nu (\boldsymbol{L}^{-1})^\nu_{\ \mu} x^\mu = p\boldsymbol{L}^{-1}x\,. \end{aligned} \tag{4.62}$$

This leads to (taking $p' = \boldsymbol{L}p$)

$$\Delta(\boldsymbol{L}^{-1}x) = -i\int \frac{\mathrm{d}^4 p'}{(2\pi)^3}\varepsilon(p'_0)\delta(p^{2\prime} - m^2)e^{-ip'x} = \Delta(x)\,. \tag{4.63}$$

Statement: There exists an alternative representation for $\Delta(x)$ given by

$$\Delta(x) = -\int_\gamma \frac{\mathrm{d}^4 p}{(2\pi)^4}\frac{e^{-ipx}}{p^2 - m^2}\,, \tag{4.64}$$

where

$$\int \mathrm{d}^4 p = \int \mathrm{d}^3\boldsymbol{p}\int_\gamma \mathrm{d}p_0\,. \tag{4.65}$$

and γ represents the integration contour depicted on Fig. 4.2. With this

$$\int_\gamma \mathrm{d}p^0 \frac{e^{-ip^0 t}}{(p^0 - \omega_{\boldsymbol{p}})(p^0 + \omega_{\boldsymbol{p}})} = 2\pi i(R_1 + R_2)\,. \tag{4.66}$$

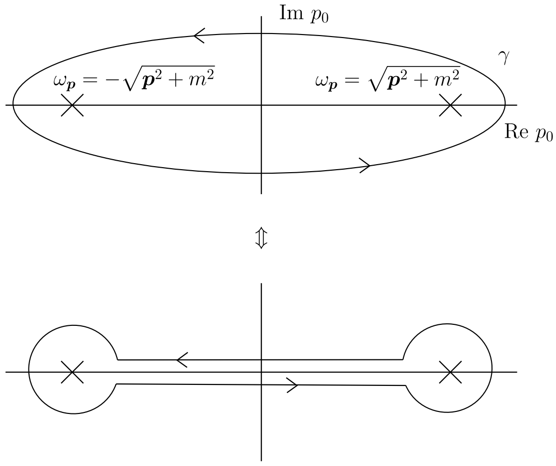


Figure 4.2: Contour of integration in Eq. (4.64)

Where the last step follows from the Cauchy theorem and

$$R_1 = \frac{e^{-i\omega_{\boldsymbol{p}}t}}{2\omega_{\boldsymbol{p}}}, \qquad R_2 = -\frac{e^{i\omega_{\boldsymbol{p}}t}}{2\omega_{\boldsymbol{p}}}. \tag{4.67}$$

This leads to the

$$\begin{aligned} \Delta(x) &= -(2\pi i)\int \frac{d^3\boldsymbol{p}}{(2\pi)^4}\frac{1}{2\omega_{\boldsymbol{p}}}\left(e^{-i\omega_{\boldsymbol{p}}t+i\boldsymbol{p}\boldsymbol{x}} - e^{i\omega_{\boldsymbol{p}}t+i\boldsymbol{p}\boldsymbol{x}}\right) \\ &= -(2\pi i)\int \frac{d^3\boldsymbol{p}}{(2\pi)^4}\frac{1}{2\omega_{\boldsymbol{p}}}\left(e^{-ipx} - e^{ipx}\right) \\ &= -i\int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}}\left(e^{-ipx} - e^{ipx}\right). \end{aligned} \tag{4.68}$$

The c-numbered commutator $[\phi(x), \phi(0)]$ is known as the *Pauli–Jordan commutation function*.

Another useful representation of $\Delta(x)$ is

$$\Delta(x) = -\frac{i}{(2\pi)^3}\int \frac{d^3\boldsymbol{p}}{\omega_{\boldsymbol{p}}} e^{i\boldsymbol{p}\boldsymbol{x}} \sin(t\omega_{\boldsymbol{p}}). \tag{4.69}$$

$J_1(y) \approx \frac{1}{2}y^2 + O(y^3)$ for $|y| \ll 1$.

This can be explicitly Fourier transformed with the result

$$\Delta(x) = -\left[\frac{1}{2\pi}\varepsilon(t)\delta(x^2) - \frac{m\varepsilon(t)}{4\pi\sqrt{x^2}}\Theta(x^2) J_1\left(m\sqrt{x^2}\right)\right], \tag{4.70}$$

where $x^2 = t^2 - \boldsymbol{x}^2$ and J_1 is the Bessel function of the first kind. In the neighborhood of the light cone (i.e. $x^2 \sim 0$) we get

$$\Delta(x) = -\left[\frac{1}{2\pi}\varepsilon(x_0)\delta(x^2) - \frac{m^2}{8\pi}\varepsilon(x_0)\Theta(x^2)\right]. \tag{4.71}$$

In particular, for time-like separation (i.e., $x^2 < 0$) we get $\Delta(x) = 0$. This is called *microcausality*, i.e.

$$[\phi(x), \phi(y)] = 0, \qquad \forall x, y,\ (x-y)^2 < 0. \tag{4.72}$$

So, free fields and any local observables constructed from such fields commute at space-like separated intervals. Consequently, they can be observed/measured independently without influencing each other. Note also that for causality purposes it was indeed enough to study only $[\phi(x), \phi(0)]$, since

$$e^{-iP_\mu y^\mu}[\phi(x), \phi(y)]e^{iP_\mu y^\mu} = [\phi(x-y), \phi(0)]\,. \tag{4.73}$$

It should perhaps be stressed that microcausality does not preclude such non-local quantum effects as **quantum correlations** and ensuing **entanglement**, which result from non-local state (vacuum state) that enters in their definition.

It can be shown, that microcausality holds for all known relativistic fields (with anticommutators in case of fermionic fields).

The Feynman Propagator

The basic building block of the perturbative treatment of scattering problems in particle physics is the so-called Feynman propagator $\Delta_F(x)$, which is defined as

$$i\Delta_F(x-y) \;=\; \langle 0|\, T\,[\phi(x)\phi(y)]\,|0\rangle\;, \tag{4.74}$$

where the *time ordering* (or *time ordered product*) $T(\cdots)$ means

$$T\,(\phi(x)\phi(y)) \;=\; \begin{cases} \phi(x)\phi(y), & x^0 > y^0 \\ \phi(y)\phi(x), & y^0 > x^0 \end{cases}. \tag{4.75}$$

For $x^0 > 0$ we can write

$$\begin{aligned} i\Delta_F(x) &= \langle 0|\,\phi(x)\phi(0)\,|0\rangle \\ &= \sum_{p,p'} \langle 0|\left[a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right]\left[a(p')e^{-ip0} + a^\dagger(p')e^{ip0}\right]|0\rangle \\ &= \sum_{p,p'} \langle 0|\,a(p)a^\dagger(p')\,|0\rangle\, e^{-ipx} \\ &= \sum_{p,p'} \langle 0|\left[a(p), a^\dagger(p')\right]|0\rangle\, e^{-ipx} \\ &= \sum_{p} e^{-ipx}\,. \end{aligned} \tag{4.76}$$

And similarly for $x^0 < 0$ we get

$$i\Delta_F(x) \;=\; \sum_p e^{ipx}\,. \tag{4.77}$$

Thus, generally we can express Feynman propagator as

$$i\Delta_F(x) \;=\; \sum_p \left[\Theta(x_0)e^{-ipx} + \Theta(-x^0)e^{ipx}\right]. \tag{4.78}$$

First term in the sum propagates a particle with positive energy forward in time, while the second one propagates a particle with negative energy backward in time.

There exists a number of useful representations of $\Delta_F(x)$. One very convenient and manifestly Lorentz invariant representation of $\Delta_F(x)$ is

$$\begin{aligned} i\Delta_F(x) &= i\int \frac{d^4p}{(2\pi)^4}\frac{e^{-ipx}}{p^2-m^2+i\varepsilon} \\ &= i\int_{\mathbb{R}^3}\frac{d^3\boldsymbol{p}}{(2\pi)^3}\,e^{i\boldsymbol{p}\cdot\boldsymbol{x}}\int_{\mathbb{R}}\frac{dp^0}{(2\pi)}\underbrace{\frac{e^{-ip_0x_0}}{(p^0)^2-(\omega_p-i\varepsilon)}}_{A}\,. \end{aligned} \tag{4.79}$$

In the complex p_0-plane the pole situation looks like:

$$\begin{aligned} (p^0)^2-(\omega_p^2-i\varepsilon) &= (p^0-\sqrt{\omega_p^2-i\varepsilon})(p^0+\sqrt{\omega_p^2-i\varepsilon}) \\ &\simeq (p_0-\omega_p+\underbrace{\frac{i\varepsilon}{2\omega_p}}_{=i\varepsilon'})(p_0+\omega_p-\underbrace{\frac{i\varepsilon}{2\omega_p}}_{=i\varepsilon'}) \\ &= \left[p_0-(\omega_p-i\varepsilon')\right]\left[p_0+(\omega_p-i\varepsilon')\right]\,. \end{aligned} \tag{4.80}$$

So, the integrand has two poles located at $p_0=\omega_p-i\varepsilon'$ and $p_0=-\omega_p+i\varepsilon'$, cf. Fig 4.3. When $x_0>0$, one can close the contour from

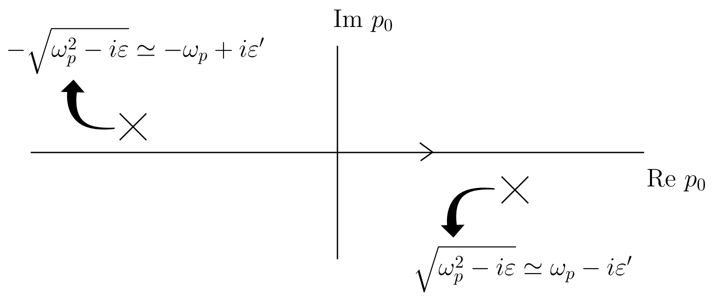


Figure 4.3: Pole structure of the integrand in (4.79).

below by a circle with infinite radius. Indeed, note that in this case the timelike integral in (4.79) over the lover circle is zero. Indeed

$$\begin{aligned} \int_{\circlearrowleft} dp_0\,A &= \{p^0=Re^{i\varphi}\} = \lim_{R\to\infty} iR\int_0^{-\pi}d\varphi e^{i\varphi}\frac{e^{iRe^{i\varphi}}}{(p^0)^2-(\omega_p^2-i\varepsilon)} \\ &= \lim_{R\to\infty} iR\int_0^{-\pi}d\varphi\;e^{i\varphi}\frac{e^{-iRx_0(\cos\varphi+i\sin\varphi)}}{R^2e^{2i\varphi}-(\omega_p^2-i\varepsilon)}\,. \end{aligned} \tag{4.81}$$

This implies that

$$\begin{aligned} &\lim_{R\to\infty} R\left|\int_0^{-\pi}d\varphi e^{i\varphi}\frac{e^{-iRx_0\cos\varphi}e^{Rx_0\sin\varphi)}}{R^2e^{2i\varphi}-(\omega_p^2-i\varepsilon)}\right| \le \lim_{R\to\infty} R\int_0^{-\pi}d\varphi|\ldots| \\ &= \lim_{R\to\infty} R\int_0^{-\pi}d\varphi\;\frac{e^{Rx_0\sin\varphi}}{R^2+\text{bounded}} = 0\,, \end{aligned} \tag{4.82}$$

which shows that $\int_{\circlearrowright} dp_0\, A = 0$. Consequently, for the p^0-integral with $x_0 > 0$ we can write

$$\begin{aligned}
&\int_{\overrightarrow{\circlearrowright}} \frac{\mathrm{d}p_0}{(2\pi)} \frac{e^{-ip_0x_0}}{(p^0)^2 - (\omega_p^2 - i\varepsilon)} \\
&= i \int_{\overrightarrow{\circlearrowright}} \frac{\mathrm{d}p_0}{(2\pi)i} \frac{e^{-ip_0x_0}}{(p^0 - (\omega_p - i\varepsilon'))(p^0 + (\omega_p - i\varepsilon'))} \\
&= -i \frac{e^{-i(\omega_p - i\varepsilon')x_0}}{2(\omega_p - i\varepsilon')} \overset{\varepsilon' \to 0}{\to} -i \frac{e^{-i\omega_p x_0}}{2\omega_p} \, .
\end{aligned} \tag{4.83}$$

Here we used the Cauchy integral formula:

$$f(x_0) = \frac{1}{2\pi i} \oint \mathrm{d}z \frac{f(z)}{z - x_0} \, . \tag{4.84}$$

Thus

$$i\Delta_F(x)|_{x_0>0} = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} \frac{e^{-i\omega_p x_0 - i\boldsymbol{p}x_0}}{2\omega_p} = \sum_p e^{-ipx} . \tag{4.85}$$

Similarly for $x_0 < 0$ we can close the p_0-integral with a large upper circle $\curvearrowleft$. With this we get that

$$i\Delta_F(x)|_{x_0<0} = \sum_p e^{ipx} \, . \tag{4.86}$$

Hence

$$i\Delta_F(x) = \sum_p \left[\Theta(x_0)e^{-ipx} + \Theta(-x^0)e^{ipx}\right] \, . \tag{4.87}$$

Should we have started directly from the form (4.87) we could arrive at the integral representation (4.79) by employing the following representation of Θ function

$$\Theta(x) = \frac{1}{2\pi i} \int_{\boldsymbol{R}^3} \mathrm{d}\tau \frac{e^{ix\tau}}{\tau - i\varepsilon} \, . \tag{4.88}$$

Let us now see that $\Delta_F(x)$ really corresponds to Green's function. To this end we consider

$$(\Box + m^2)\Delta_F(x) = -i(\Box + m^2) \langle 0| T[\phi(x)\phi(0)] |0\rangle \, . \tag{4.89}$$

To do this computation it suffices to concentrate only on the $T(\ldots)$ product. Namely

$$\begin{aligned}
&-i\left(\frac{\partial^2}{\partial(x^0)^2} - \nabla^2 + m^2\right)[\Theta(x_0)\phi(x)\phi(0) + \Theta(-x_0)\phi(0)\phi(x)] \\
&= -i\frac{\partial^2}{\partial(x^0)^2}[\cdots] - i\left[\Theta(x_0)(-\nabla_x^2 + m^2)\phi(x)\phi(0)\right. \\
&\qquad \left. + \ \Theta(-x_0)\phi(0)(-\nabla_x^2 + m^2)\phi(x)\right] .
\end{aligned} \tag{4.90}$$

Rewriting the expressing for $-i\partial^2[\cdots]/\partial(x^0)^2$ as

$$\begin{aligned}
& -i\frac{\partial}{\partial x^0}\,[\delta(x_0)\phi(x)\phi(0) - \delta(x_0)\phi(0)\phi(x) \\
& \qquad + \Theta(x_0)\frac{\partial}{\partial x^0}\phi(x)\phi(0) + \Theta(-x_0)\phi(0)\frac{\partial}{\partial x^0}\phi(x)\Big] \\
& = -i\left[\delta(x_0)\left[\dot{\phi}(x),\phi(0)\right]\right. \\
& \qquad \left. + \Theta(x_0)\frac{\partial^2}{\partial(x^0)^2}\phi(x)\phi(0) + \Theta(-x_0)\phi(0)\frac{\partial^2}{\partial(x^0)^2}\phi(x)\right]. \quad (4.91)
\end{aligned}$$

Thus

$$\begin{aligned}
-i(\Box + m^2)T[\phi(x)\phi(0)] &= -\delta(x) + \Theta(x_0)(\Box + m^2)\phi(x)\phi(0) \\
&\quad + \Theta(-x_0)\phi(0)(\Box + m^2)\phi(x) \\
&= -\delta(x)\,. \qquad (4.92)
\end{aligned}$$

Which means that

$$(\Box + m^2)\Delta_F(x) = -\delta(x)\,. \qquad (4.93)$$

Alternatively, we can prove this directly from the integral representation of the Feynman propagator:

$$\begin{aligned}
-i(\Box + m^2)\langle 0|\,T[\phi(x)\phi(0)]\,|0\rangle &= (\Box + m^2)\int \frac{\mathrm{d}^4p}{(2\pi)^4}\frac{e^{-ipx}}{(p^2 - m^2 + i\varepsilon)} \\
&= -\int \frac{\mathrm{d}^4p}{(2\pi)^4}\frac{(p^2 - m^2)}{(p^2 - m^2 + i\varepsilon)}e^{-ipx} \\
&= -\delta(x)\,. \qquad (4.94)
\end{aligned}$$

The fact that this is indeed equal to the Dirac delta function follows from the properties of generalized functions. We know from Sokhotski formula that

$$\frac{1}{x + i\varepsilon} = \mathcal{P}\frac{1}{x} - i\pi\delta(x)\,, \qquad (4.95)$$

which should be understood in the sense that for any Schwartz test function g one has the scalar product identity

$$\left(\frac{1}{x + i\varepsilon}, g\right) = \left(\mathcal{P}\frac{1}{x}, g\right) - i\pi(\delta, g)\,. \qquad (4.96)$$

Using the fact that $\frac{x}{x+i\varepsilon} = x\mathcal{P}\frac{1}{x} - i\pi x\delta(x) = x\mathcal{P}\frac{1}{x}$ we have

$$\left(\frac{x}{x + i\varepsilon}, g\right) = \lim_{a\to 0}\left(\int_{-\infty}^{-a} dx\,\frac{x}{x}g + \int_{a}^{\infty} dx\,\frac{x}{x}g\right) = (1, g)\,. \qquad (4.97)$$

So, we have that $x/(x + i\varepsilon) = 1$.

One can also calculate the integral (4.79) explicitly. The actual result splits into 3 parts. Light-cone part (i.e., when $x^2 = 0$) that has a simple form $\delta(x^2)/(4\pi)$, part inside of the light cone (i.e., when $x^2 > 0$), which is a combination of Bessel functions $J_1(m\sqrt{x^2})$ and $N_1(m\sqrt{x^2})$, and finally part that corresponds to $x^2 < 0$, which is proportional to $K_1(m\sqrt{-x^2})$. In the neighborhood of the light cone the solution can be expanded as

$$\Delta_F(x) \simeq \frac{1}{4\pi}\delta(x^2) - \frac{i}{4\pi^2 x^2} + \frac{im^2}{8\pi^2}\ln|x| - \frac{m^2}{16\pi}\Theta(x^2)\,. \tag{4.98}$$

So, $\Delta_F(x)$ penetrates also behind the light cone. We will see that $\Delta_F(x)$ basically corresponds to an **amplitude of probability** that a particle starting at $x = (0, \mathbf{0})$ will end up at the point $x = (x_0, \boldsymbol{x})$. In this respect there is a non-zero probability that a quantum particle might evolve into a space-like separated regions.

Let us consider the part of the solution with $x^2 < 0$ more in detail. This has the explicit form

$$\Delta_F(x) \;=\; \frac{im}{4\pi^2}\frac{K_1(m\sqrt{-x^2})}{\sqrt{-x^2}}\,. \tag{4.99}$$

For large $\sqrt{-x^2} \equiv |x|$ the latter has the expansion

$$\Delta_F(x) \;=\; \frac{1}{4}\sqrt{\frac{m}{2(|x|\pi)^3}}\; e^{-m|x|}\;[1 + O(1/|x|)]\;. \tag{4.100}$$

Thus for large $|x|$ the behavior is dominated by the exponential part. Note that when we reintroduce $\hbar$ and c then $e^{-m|x|} \to e^{-\frac{cm}{\hbar}|x|} = e^{-\frac{|x|}{\lambda_C}}$. So a typical distance over which a particle can appreciably "tunel" behind light-cone is $\lambda_C = \frac{\hbar}{mc}$, which is a Compton wave length. We have seen that this was a reason for existence of anti particles.

Notes on microcausality

Despite the microcausality, there is a nontrivial correlation even at space-like distances. This is due to vacuum that can mediate the correlations.

Dirac Field

Recall that we require to use anti-commutation relations for Fermi field instead of commuting ones. We therefore define the time ordering (time ordered product) for $\psi_\alpha(x)$ and $\overline{\psi}_\beta(x)$ to be

$$T[\psi_\alpha(x)\overline{\psi}_\beta(y)] = \begin{cases} \psi_\alpha(x)\overline{\psi}_\beta(y) & x_0 > y_0 \\ -\overline{\psi}_\beta(y)\psi_\alpha(x) & x_0 < y_0 \end{cases}. \tag{4.101}$$

We will see in the following that this definition will also be consistent with other requirements, e.g., it will allow us to get a correct Green's function for Dirac equation.

We define the corresponding Feynman propagator to be

$$i\{S_F(x)\}_{\alpha\beta} = \langle 0|\, T[\psi_\alpha(x)\overline{\psi}_\beta(y)]\, |0\rangle\,. \tag{4.102}$$

It will be this object that will be a basic building block in the perturbative treatment of scattering matrix. Again, it will correspond to the Green function (this time for Dirac equation) with correct pole avoidance prescription.

If we follow through the same type of argument as for scalar field we find that

$$\begin{aligned} \{S_F(x)\}_{\alpha\beta} &= \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{e^{-ipx}(\not{p}+m)_{\alpha\beta}}{p^2-m^2+i\varepsilon} \\ &= -i\,\langle 0|\, T[\psi_\alpha(x)\overline{\psi}_\beta(y)]\,|0\rangle \ . \end{aligned} \quad (4.103)$$

Let us recall some relevant steps in the proof. For $x_0 > 0$ we get

$$\begin{aligned} &\{S_F(x)\}_{\alpha\beta} \\ &\quad = -i \sum_{p,p'} \sum_{\lambda,\lambda'} \langle 0|\, [a(p,\lambda)u_\alpha(p,\lambda)e^{-ipx} + b^\dagger(p,\lambda)v_\alpha(p,\lambda)e^{ipx}] \\ &\qquad \times [b(p',\lambda')\overline{v}_\beta(p',\lambda')e^{-ip0} + a^\dagger(p',\lambda')\overline{u}_\beta(p',\lambda')e^{ip0}]\,|0\rangle \ . \end{aligned}$$

Here we note that only first and fourth term in the sum are relevant. Continuing we get

$$\begin{aligned} \{S_F(x)\}_{\alpha\beta} &= -i \sum_{p,p'} \sum_{\lambda,\lambda'} \langle 0|\, [a(p,\lambda), a^\dagger(p',\lambda')]\,|0\rangle\, u_\alpha(p,\lambda)\overline{u}_\beta(p,\lambda)e^{-ipx} \\ &= -i \sum_{p,\lambda} u_\alpha(p,\lambda)\overline{u}_\beta(p,\lambda)e^{-ipx} \\ &= -i \sum_{p,\lambda} (\not{p}+m)_{\alpha\beta} e^{-ipx} \ . \end{aligned} \quad (4.104)$$

Similarly we can repeat this procedure for $x_0 < 0$. Finally we get that

$$S_F(x) = -i \sum_p \left[\Theta(x_0)(\not{p}+m)e^{-ipx} - \Theta(-x_0)(\not{p}-m)e^{ipx}\right] \ . \quad (4.105)$$

On the other hand,

$$\begin{aligned} &\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{(\not{p}+m)}{p^2-m^2+i\varepsilon} e^{-ipx} \\ &\qquad = i \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{(\not{p}+m)}{[p_0-(\omega_p-i\varepsilon')][p_0+(\omega_p-i\varepsilon')]} e^{-ipx} \ . \end{aligned} \quad (4.106)$$

By closing our contour down (see Fig. (4.4)) we get for $x_0 > 0$ that the integral above is equal to

$$-i \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{(\not{p}+m)}{2\omega_p} e^{i\boldsymbol{px}-i\omega_p x_0} = -i \sum_p (\not{p}+m)e^{-ipx} \ . \quad (4.107)$$

This coincides with Eq. (4.105). Similar reasoning can be done for $x_0 < 0$.

To show that $S_F(x)$ is Green's function of Dirac equation, let us con-

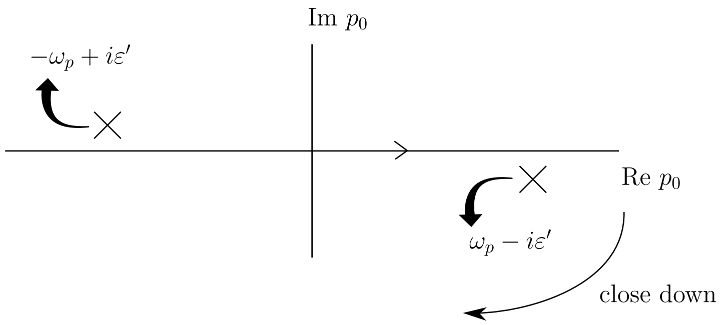


Figure 4.4: Way of closing the contour in integral (4.106) for the case $x_0 > 0$.

sider

$$\begin{aligned}(i\partial\!\!\!/ - m)S_F(x) &= \int \frac{\mathrm{d}^4p}{(2\pi)^4} \frac{(p\!\!\!/ - m)(p\!\!\!/ + m)}{p^2 - m^2 + i\varepsilon} e^{-ipx} \\ &= \int \frac{\mathrm{d}^4p}{(2\pi)^4} \frac{p^2 - m^2}{p^2 - m^2 + i\varepsilon} e^{-ipx} \\ &= \int \frac{\mathrm{d}^4p}{(2\pi)^4} e^{-ipx} = \delta(x)\mathbf{1}\,. \end{aligned} \tag{4.108}$$

Another representation

One often alternatively writes $S_F(x)$ in yeat another form. Take

$$[p\!\!\!/ - (m \pm i\varepsilon)][p\!\!\!/ + (m \pm i\varepsilon)] = p^2 - m^2 \mp 2i\varepsilon m + \varepsilon^2\,. \tag{4.109}$$

Then by inverting this relation

$$[p\!\!\!/ - (m \pm i\varepsilon)]^{-1}[p\!\!\!/ + (m \pm i\varepsilon)]^{-1} = \frac{1}{p^2 - m^2 \mp 2i\varepsilon m + \varepsilon^2}\,. \tag{4.110}$$

By denoting $\varepsilon' = 2i\varepsilon m$ and neglecting ε^2 we get that

$$[p\!\!\!/ - (m \pm i\varepsilon)]^{-1}[p\!\!\!/ + (m \pm i\varepsilon)]^{-1} = \frac{1}{p^2 - m^2 + i\varepsilon'}\,. \tag{4.111}$$

Thus finally

$$\frac{(p\!\!\!/ + m)}{p^2 - m^2 + i\varepsilon'} = (p\!\!\!/ + m)(p\!\!\!/ + (m - i\varepsilon))^{-1}(p\!\!\!/ + m + i\varepsilon)^{-1}\,. \tag{4.112}$$

Or more precisely

$$\frac{(p\!\!\!/ + m)}{p^2 - m^2 + i\varepsilon'} = \frac{1}{p\!\!\!/ - m - i\varepsilon} + \frac{i\varepsilon}{p^2 - m^2 + i\varepsilon'}\,, \tag{4.113}$$

and hence

$$S_F(x) = \int \frac{\mathrm{d}^4p}{(2\pi)^4} \frac{e^{-ipx}}{p\!\!\!/ - m + i\varepsilon} \tag{4.114}$$

It can be shown that this again corresponds to the transitional amplitude and again there is a non-zero contribution from $x^2 < 0$ with effective penetration distance of the order of λ_c.

4.4 Interacting Fields

So far we have dealt with non-interacting particles that were represented through free fields. To include interactions among particles we must introduce interaction terms into the Lagrangian of the field theory.

As a test bed for further applications we start with Hermitian (i.e. uncharged) field Lagrangian. If the field is free we have

$$\mathcal{L}_F = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\varphi^2 . \tag{4.115}$$

The interaction is introduced by making the substitution $\mathcal{L}_F \to \mathcal{L} = \mathcal{L}_F + \mathcal{L}_I$, and requiring this quantity to be the Lorentz density (so that the ensuing equations of the motion are relativistically invariant). Here the term $\mathcal{L}_I$ is the so-called interaction Lagrangian. The simplest form of $\mathcal{L}_I$ that keeps $\mathcal{L}$ to be Lorentz density is the form where $\mathcal{L}_I$ is a local function of fields. Among these, the polynomial functions are the simplest ones. Let us thus consider particularly (and phenomenologically) important case

$$\mathcal{L}_I = -\frac{g}{3!}\phi^3 - \frac{\lambda}{4!}\phi^4 \equiv -V(\phi) . \tag{4.116}$$

The presence of the $\mathcal{L}_I$ in the Lagrangian density means that ϕ no longer obeys the Klein-Gordon equation. If we construct the field $\pi(x)$ conjugate to $\phi(x)$ it is $\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)}$, then

$$\mathcal{L} = \frac{1}{2}\ddot{\phi} - \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}m^2\phi^2 - V(\phi) \tag{4.117}$$

and hence $\pi = \dot{\phi}$ as before.

Canonical commutation relation is

$$[\phi(x), \pi(x')] = [\phi(x), \dot{\phi}(x')] = i\delta(\boldsymbol{x} - \boldsymbol{x}'), \tag{4.118}$$

where $x = (x^0, \boldsymbol{x})$ and $x' = (x^0, \boldsymbol{x}')$. Recall that Hamilton density is defined by

$$\mathcal{H}(x) = \pi(x)\dot{\phi}(x) - \mathcal{L}(x). \tag{4.119}$$

Thus, if we substitute our $\mathcal{L}(x)$ we get

$$\begin{aligned}\mathcal{H} &= \pi^2(x) - \left[\frac{1}{2}\pi^2(x) - \frac{1}{2}(\nabla\phi(x))^2 - \frac{1}{2}m^2\phi^2(x)\right] + V(\phi)\\ &= \frac{1}{2}\pi^2(x) + \frac{1}{2}(\nabla\phi(x))^2 + \frac{1}{2}m^2\phi^2(x) + \frac{g}{3!}\phi^3(x) + \frac{\lambda}{4!}\phi^4(x)\\ &= \mathcal{H}_0(x) + \mathcal{H}_I(x) = \mathcal{H}_0(x) - \mathcal{L}_I(x).\end{aligned} \tag{4.120}$$

Here

$$\mathcal{H}_0(x) = \frac{1}{2}\pi^2(x) + \frac{1}{2}(\nabla\phi(x))^2 + \frac{1}{2}m^2\phi^2(x), \tag{4.121}$$

and

$$\mathcal{H}_I(x) = \frac{g}{3!}\phi^3(x) + \frac{\lambda}{4!}\phi^4(x) = -\mathcal{L}_I(x). \tag{4.122}$$

At a given time t the $H = H_0 + H_I$ and we can calculate them using

$$H_0 = \int d^3\boldsymbol{x}\mathcal{H}_0(x), \tag{4.123}$$

$$H_I = \int d^3\boldsymbol{x}\mathcal{H}_I(x). \tag{4.124}$$

The Heisenberg equations of motion for $\phi(x)$ is given by the Euler-Lagrange equation

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0. \tag{4.125}$$

Which gives us

$$\partial_\mu(\partial^\mu\phi) + m^2\phi + \underbrace{\frac{g}{2}\phi^2 + \frac{\lambda}{3!}\phi^3}_{\frac{\partial V}{\partial\phi}} = 0. \tag{4.126}$$

The extra term $\frac{\partial V}{\partial\phi}$ in the equation of motion prevents the solution being free-field solutions.

4.5 Perturbative theory

We will make split $H = H_0 + H_I$ at time $t = 0$

$$H_0 = \int d^3\boldsymbol{x}\left[\frac{1}{2}\pi^2(0,\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{\nabla}\phi(0,\boldsymbol{x}))^+\frac{1}{2}m^2\phi^2(0,\boldsymbol{x})\right], \tag{4.127}$$

$$\begin{aligned} H_I &= \int d^3\boldsymbol{x}\left[\frac{g}{3!}\phi^3(0,\boldsymbol{x}) + \frac{\lambda}{4!}\phi^4(0,\boldsymbol{x})\right] \\ &= \int d^3\boldsymbol{x}\, V(\phi(0,\boldsymbol{x})). \end{aligned} \tag{4.128}$$

Interaction picture

Although the general theory of fields is best investigated using Heisenberg picture, the fastest route to perturbation theory is via "interaction" picture.

4.6 Interaction (Dirac's) Picture

Schrödinger picture is given by

$$i\frac{d}{dt}\left|\psi(t)\right\rangle_S = H^S\left|\psi(t)\right\rangle_S . \tag{4.129}$$

If H^S is time independent, then

$$|\psi(t)\rangle_S = e^{-iH^S t} |\psi(0)\rangle_S . \tag{4.130}$$

On the quantum level the full information about the interaction is most naturally encoded in fields in the Heisenberg picture. On the other hand, the passage to the perturbation calculus is most easily done via *interaction (or Dirac's) picture*. Let us thus first introduce the interaction picture.

If $H = H_0 + V$ is the full Hamiltonian (and its free and interaction part) in Schrödinger picture, we define interaction picture by the relation

$$|\psi(t)\rangle_I = e^{iH_0{}^S t} |\psi(t)\rangle_S , \tag{4.131}$$

which can be equivalently rewritten as $|\psi(t)\rangle_S = e^{-iH_0{}^S t} |\psi(t)\rangle_I$.

In the other words we peel off the free-theory Schrödinger time evolution so as to be able to concentrate on the effect of the interactions only. The corresponding equation of the motion for $|\psi(t)\rangle_I$ can be directly obtained from the defining relation (4.131), indeed

$$\begin{aligned} i\frac{d}{dt}|\psi(t)\rangle_I &= -H_0{}^S |\psi(t)\rangle_I + e^{-iH_0{}^S t}(i\frac{d}{dt}|\psi(t)\rangle_S) \\ &= -H_0{}^S |\psi(t)\rangle_I + e^{iH_0{}^S t}(H_0{}^S + V^S)e^{-iH_0{}^S t} |\psi(t)\rangle_I \\ &= V^I |\psi(t)\rangle_I = \overline{H}_I |\psi(t)\rangle_I , \end{aligned} \tag{4.132}$$

where

$$\overline{H}_I = e^{iH_0{}^S t} H_I{}^S e^{-iH_0{}^S t} , \tag{4.133}$$

in the interaction part of the Hamiltonian in the interaction picture, which (in contrast to its Schrödinger picture counterpart) is generally time dependent. Note, that for a time-independent $H_I{}^S$, the time dependence of $\overline{H_I}$ is that of a free Heisenberg field. Equation (4.133) provides the defining relation between interaction and Schrödinger picture. So, in general $A_I = e^{iH_0^S t} A^S e^{-iH_0^S t}$.

Note I.

CF. with the usual relation between Heisenberg and Schrödinger picture where

$$A^H = e^{iH^S t} A^S e^{-iH^S t} , \tag{4.134}$$

⇒ in the interaction picture is the evolution of state dictated by the interaction while the evolution of operators via free part of Hamiltonian.

Note II.

Let us look at the equations of motion for $\overline{H}_I$.

$$\begin{aligned} i\frac{d}{dt}\overline{H}_I &= -\, e^{iH_0{}^S t} H_0{}^S H_I{}^S e^{-iH_0{}^S t} \\ &\quad + e^{iH_0{}^S t} H_I{}^S \underbrace{e^{-iH_0{}^S} e^{iH_0{}^S}}_{=1} H_0{}^S e^{-iH_0{}^S t} \\ &= -\overline{H}_0\overline{H}_I + \overline{H}_I\overline{H}_0 = [\overline{H}_I;\overline{H}_0]\,. \end{aligned} \tag{4.135}$$

$\Rightarrow \frac{d}{dt}\overline{H}_I = -i[\overline{H}_I;\overline{H}_0]$.
(i.e., evolution via free-field Hamiltonian, as expected). We assumed that $H_I{}^S$ is time independent and $H_0{}^S = H_0{}^I = \overline{H}_0$.

We now come back to the equation for states. If at $t = t_i$ we have that o$|\psi(t_1)\rangle_I = |\psi_i\rangle$, then one can observe that

$$|\psi(t)\rangle_I = |\psi_i\rangle + \frac{1}{i}\int_{t_1}^{t} dt'\overline{H}_I(t')\,|\psi(t')\rangle_I \ . \tag{4.136}$$

This an integral equation for $|\psi(t)\rangle_I$. In mathematics this is known as the Voltera integral equation of the second type. Indeed, Eq. (4.136) clearly satisfies

$$i\frac{d}{dt}|\psi(t)\rangle_I = \overline{H}_I(t)\,|\psi(t)\rangle_I\,, \quad \text{with} \quad |\psi(t_1)\rangle_I = |\psi_i\rangle\ . \tag{4.137}$$

Let us attempt to solve this equation iteratively

$$\begin{aligned} &0^{th}\text{approx.} && |\psi(t)\rangle_I = |\psi_i\rangle \\ &1^{st}\text{approx.} && |\psi(t)\rangle_I = |\psi_i\rangle + \frac{1}{i}\int_{t_1}^{t} dt'\overline{H}_I(t')\,|\psi_i\rangle \\ &2^{nd}\text{approx.} && |\psi(t)\rangle_I = |\psi_i\rangle + \frac{1}{i}\int_{t_1}^{t} dt'\overline{H}_I(t')\,|\psi_i\rangle \\ & && \quad + \frac{1}{(i)^2}\int_{t_1}^{t} dt'\int_{t'} dt''\overline{H}_I(t')\overline{H}_I(t'')\,|\psi_i\rangle \\ & && \quad \vdots \end{aligned}$$

Generally

$$|\psi(t)1\rangle_I = U(t;t_i)\,|\psi_i\rangle\ , \tag{4.138}$$

where

$$U(t;t_i) = 1 + \sum_{n=1}^{+\infty}(-i)^n\int_{t_i}^{t} dt^1\int_{t_i}^{t^1} dt^2\ldots\int_{t_i}^{t^{n-1}} dt^n\ \overline{H}_I(t^1)\ldots\overline{H}_I(t^n)\,. \tag{4.139}$$

In the integration region ($t^1 \geqslant t^2 \geqslant t^3\ldots$) to put (4.139) into a manageable form one can use Dyson trick that is based on the time-ordering product.

To this end, we define $T\{\overline{H}_I(t^1)\overline{H}_I(t^2)\ldots\}$, which is the usual product of operators but organized so that the operators with higher time argument are more in left, or in other words, the product is such that time arguments of involved operators are in descending order from left to right. From mathematical analysis it is well-known integral identity:

$$\int_{t_1}^{t} dt^{(1)} \int_{t_1}^{t^{(1)}} dt^{(2)} \int_{t_1}^{t^{(3)}} dt^{(3)} \dots \int_{t_1}^{t^{(n-1)}} dt^{(n)} f(t^{(1)}) \dots f(t^{(n)})$$
$$= \frac{1}{n!} \int_{t_1}^{t} dt^{(1)} \int_{t_1}^{t} dt^{(2)} \int_{t_1}^{t} dt^{(3)} \dots \int_{t_1}^{t} dt^{(n)} f(t^{(1)}) \dots f(t^{(n)}) .$$

Since behind the symbol "T" all operators $\overline{H_I}(t^i)$ commute, we can write

$$U(t, t_i) = 1 + \sum_{n=0}^{+\infty} \frac{(-i)^n}{n!} \int_{t_i}^{t} dt^{(1)} \int_{t_i}^{t} dt^{(2)} \dots \int_{t_i}^{t} dt^{(n)} \, T\{\overline{H_I}(t^{(1)}) \dots\} . \tag{4.140}$$

In the absence of the time ordering one could naively write

$$U(t; t_i) = e^{-i \int_{t_i}^{t} dt' \overline{H_I}(t')} \tag{4.141}$$

but this is wrong! In fact correctly we have

$$U(t; t_i) = T \left[e^{-i \int_{t_i}^{t} dt' \overline{H_I}(t')} \right] \tag{4.142}$$

where this time-ordered exponential should be understood as "expand and apply on each monomial separately" as in (4.140).

Note III.

Note the difference between

$e^{\int_a^b dt A(t)} \approx e^{A(t_n = t_b)\Delta t + A(t_{n-1})\Delta t + \dots + A(t_1 = t_a)\Delta t}$,

and

$T \left[e^{\int_a^b dt A(t)} \right] \approx e^{A(t_n = t_b)\Delta t} e^{A(t_{n-1})\Delta t} \dots e^{A(t_2)\Delta t} e^{A(t_1 = t_a)\Delta t}$,

where $t_n > t_{n-1} > \dots > t_1$.

Only when all operators $A(t)$ at different times commute then both expressions are identical.

Note that $U(t; t_1)$ satisfies the composition law (assuming $t_1 < t_2 < t$)

$$T \left[e^{\frac{1}{i} \int_{t_1}^{t_2} dt' \overline{H}_I(t') + \frac{1}{i} \int_{t_2}^{t} dt' \overline{H}_I(t')} \right] = T \left[e^{\frac{1}{i} \int_{t_2}^{t} dt' \overline{H}_I(t')} \right] T \left[e^{\frac{1}{i} \int_{t_1}^{t_2} dt' \overline{H}_I(t')} \right]$$
$$= U(t; t_2) U(t_2; t_1) . \tag{4.143}$$

Interacting fields

Scattering processes are described in terms of transitions between an initial state of free particles far in the distant past and final state of free particles far in the remote future.

Cluster decomposition property

Assumption that a studied interacting system can be described in terms of free fields in asymptotic times is called the **cluster decomposition property**.

Formally, we are thus interested in the limits $t \to +\infty$ and $t_1 \to -\infty$ and therefore the operator

$$S = T\left[e^{-i\int_{-\infty}^{+\infty} dt \overline{H}_I(t)}\right] = T\left[e^{i\int_{-\infty}^{+\infty} dt \overline{L}_I(t)}\right], \tag{4.144}$$

will turn out to be of practical importance. From our construction

$$\overline{H}_I = -\int d^3\boldsymbol{x}\, \overline{\mathcal{L}_I}(t). \tag{4.145}$$

Consider that the interaction part of the Lagrangian has classically the form

$$\mathcal{L}_I = -\frac{g}{3!}\phi^3(x) - \frac{\lambda}{4!}\phi^4(x). \tag{4.146}$$

Quantization is again performed via Schrödinger picture. We can then easily pass to the interaction picture via the usual relation

$$\begin{aligned}\phi^I(x) &= \phi^I(t,\mathbf{x}) = e^{iH_0^S t}\phi^S(0,\mathbf{x})e^{-iH_0^S t} \\ &= e^{iH_0^I t}\phi^I(0,\mathbf{x})e^{-iH_0^S t},\end{aligned} \tag{4.147}$$

Here we have assume that Schrödinger and interaction pictures coincide at the reference time $t = 0$. So, in particular at $t = 0$ we have

$$\overline{H}_I(0) = \int d^3\boldsymbol{x}\left(\frac{g}{3!}\phi^3(0,\boldsymbol{x}) + \frac{\lambda}{4!}\phi^4(0,\boldsymbol{x}) + \ldots\right). \tag{4.148}$$

This structure remains the same for all times. Indeed

$$\begin{aligned}\overline{H}_I(t) &= e^{iH_0^I t}\overline{H}_I(0)e^{-iH_0^I t} \\ &= \int d^3\boldsymbol{x}\left(\frac{g}{3!}\phi^3(t,\boldsymbol{x}) + \frac{\lambda}{4!}\phi^4(t,\boldsymbol{x}) + \ldots\right).\end{aligned} \tag{4.149}$$

As already mentioned, the fields that appear in the interaction-picture based perturbative theory are free fields controlled by H_0. Recall, that for any operator Q

$$Q^I(t_2) = e^{iH_0^I(t_2-t_1)}Q^I(t_1)e^{-iH_0^I(t_2-t_1)}, \tag{4.150}$$

or infinitesimally

$$i\frac{dQ^I(t)}{dt} = \left[Q^I(t); H_0^I\right], \tag{4.151}$$

In the interaction picture we have $(\partial^2 + m^2)\phi^I(x) = 0$ where $\phi^I(x) = \sum_{\boldsymbol{p}}[a(p)e^{-ipx} + a^\dagger(p)e^{ipx}]$.

which is nothing but that free field equation of motion. Moreover, the canonical commutation relations also hold in the interaction picture. Indeed, by introducing the conjugated momenta

$$\Pi^I(x) = e^{iH_0 t}\Pi^I(0,\boldsymbol{x})e^{-iH_0 t} = e^{iH_0 t}\Pi^S(0,\boldsymbol{x})e^{-iH_0 t}, \tag{4.152}$$

we might directly write

$$\begin{aligned} \left[\phi^I(t,\boldsymbol{x});\Pi^I(t;\boldsymbol{x}')\right] &= i\delta(\boldsymbol{x}-\boldsymbol{x}')\,, \\ \left[\phi^I(t,\boldsymbol{x});\phi^I(t;\boldsymbol{x}')\right] &= \left[\Pi^I(t,\boldsymbol{x});\Pi^I(t;\boldsymbol{x}')\right] = 0\,, \\ \Pi^I(t,\boldsymbol{x}) &= \dot{\phi}^I(t,\boldsymbol{x})\,. \end{aligned}$$

These relations are, of course, simple consequence of the fact that the interaction picture fields are connected via unitary transformation with Schrödinger (and also Heisenberg) picture fields.

Note

Let the interaction and Heisenberg pictures coincide at some reference time t_0. We can use their respective evolution equations

$$\begin{aligned} \phi^I(t,\boldsymbol{x}) &= e^{iH_0^I(t-t_0)}\phi^I(t_0,\boldsymbol{x})e^{-iH_0^I(t-t_0)}\,, \\ \phi^H(t,\boldsymbol{x}) &= e^{iH(t-t_0)}\phi^H(t_0,\boldsymbol{x})e^{-iH(t-t_0)}\,, \end{aligned}$$

and the fact that $\phi^H(t_0,\boldsymbol{x}) = \phi^I(t_0,\boldsymbol{x})$ to obtain that

$$\begin{aligned} \phi^I(t,\boldsymbol{x}) &= e^{iH_0^I(t-t_0)}e^{-iH(t-t_0)}\phi^H(t,\boldsymbol{x})e^{iH(t-t_0)}e^{-iH_0^I(t-t_0)} \\ &= \Lambda(t,t_0)\phi^H(t,\boldsymbol{x})\Lambda^{-1}(t,t_0)\,. \end{aligned} \tag{4.153}$$

Here $\Lambda(t,t_0)$ is clearly unitary (i.e., $\Lambda^\dagger(t,t_0) = \Lambda^{-1}(t,t_0)$).

In particular, we have in the interaction picture the usual creation and annihilation operators satisfying the familiar algebra (super-index I is omitted)

$$\begin{aligned} \left[a(p),a^\dagger(p')\right] &= \delta_{pp'}\,, \\ \left[a(p),a(p')\right] &= \left[a^\dagger(p),a^\dagger(p')\right] = 0\,. \end{aligned} \tag{4.154}$$

The eigenstates of H_0 are

$$\begin{aligned} &|0\rangle\,;\;\; a(p)\,|0\rangle = 0;\;\; H_0\,|0\rangle = 0 \quad \text{(by normal ordering)}\,, \\ &|p\rangle = a^\dagger(p)\,|0\rangle\,;\;\; H_0[a^\dagger(p)\,|0\rangle] = E_p\,|p\rangle = \omega_p\,|p\rangle\,, \end{aligned} \tag{4.155}$$

and all other multi-particle states obtained by application of creation operators on $|0\rangle$.

Representation (4.153) for $\Lambda(t,t_0)$ is quite inconvenient for practical purposes as it mixes two distinct representations. In addition, for perturbation purposes it is convenient to work directly with the interaction picture. Fortunately, not difficult to find the form of $\Lambda(t,t_0)$ in the interaction picture.

To this end we again assume that t_0 is the time when both pictures coincide, so that $\Lambda(t_0,t_0) = 1$. We know that the correspond Heisenberg field equations in respective pictures read

$$\begin{aligned}\frac{\partial}{\partial t}\phi_H(t,\boldsymbol{x}) &= i\big[H_H(\phi_H,\Pi_H),\phi_H(t,\boldsymbol{x})\big]\,,\\ \frac{\partial}{\partial t}\phi_I(t,\boldsymbol{x}) &= i\big[H_0(\phi_I,\Pi_I),\phi_I(t,\boldsymbol{x})\big]\,. \end{aligned}\tag{4.156}$$

In the following we will also need the simple identity, namely

$$\Lambda\Lambda^{-1} = 1 \;\;\Rightarrow\;\; \left(\frac{d}{dt}\Lambda\right)\Lambda^{-1} \;=\; -\Lambda\frac{d}{dt}\Lambda^{-1}\,. \tag{4.157}$$

Let us now take the derivative of ϕ_I and use the former identities. This gives

$$\begin{aligned}\frac{\partial}{\partial t}\phi_I(t,\boldsymbol{x}) &= \frac{\partial}{\partial t}\big[\Lambda\phi_H\Lambda^{-1}\big] = \dot{\Lambda}\phi_H\Lambda^{-1} + \Lambda\dot{\phi_H}\Lambda^{-1} + \Lambda\phi_H\dot{\Lambda^{-1}}\\ &= \dot{\Lambda}\big(\Lambda^{-1}\phi_I\Lambda\big)\Lambda^{-1} + i\Lambda\big[H_H\big(\phi_H,\Pi_H\big),\phi_H\big]\Lambda^{-1}\\ &\quad + \Lambda\big(\Lambda^{-1}\phi_I\Lambda\big)\dot{\Lambda^{-1}}\\ &= \dot{\Lambda}\Lambda^{-1}\phi_I + i\big[H\big(\phi_I,\Pi_I\big),\phi_I\big] + \phi_I\,\underbrace{\Lambda\dot{\Lambda^{-1}}}_{-\dot{\Lambda}\Lambda^{-1}}\\ &= \big[\dot{\Lambda}\Lambda^{-1} + i\,\underbrace{H\big(\phi_I,\Pi_I\big)}_{H-H_0+H_0},\phi_I\big]\,. \end{aligned}\tag{4.158}$$

This should be compared with (4.156). In fact, since (4.158) holds for any operator (not necessarily only for $\phi_I(t,\boldsymbol{x})$), we inevitably get that

$$\dot{\Lambda}\Lambda^{-1} + i\big(H - H_0\big)_I \;=\; \text{c-number} \;\;\Rightarrow\;\; \dot{\Lambda} = -i\overline{H}_I\Lambda + c\Lambda\,,$$

(c is some c-numbered time-dependent function). The previous line can thus be equivalently rewritten as

$$i\frac{\partial\Lambda(t,t_0)}{\partial t} \;=\; \Big[\overline{H}_I(t) + ic\Big]\,\Lambda(t,t_0)\,. \tag{4.159}$$

Note that this is the same type of equation we had for states in the interaction picture. So, by using the boundary condition $\Lambda(t_0,t_0) = 1$, we can equivalently rewrite (4.159) as the Volterra integral equation

$$\begin{aligned}\Lambda(t,t_0) &= 1 - i\int_{t_0}^{t} dt_1\big(\overline{H}_I + ic\big)\Lambda(t_1,t_0)\\ &= T\left[e^{-i\int_{t_0}^{t}d\tau\left(\overline{H}_I\left(\tau\right)+ic(\tau)\right)}\right]\underbrace{\Lambda\big(t_0,t_0\big)}_{1}\\ &= e^{\int_{t_0}^{t}c(\tau)d\tau}\,T\left[e^{-i\int_{t_0}^{t}d\tau\left(\overline{H}_I(\tau)\right)}\right]\,. \end{aligned}\tag{4.160}$$

Because both $\Lambda(t,t_0)$ (by its very definition) and $T\big[\dots\big]$ are unitary operators, we have that $\left|e^{\int c\,dt}\right| = 1$. Consequently, such a phase factor will not contribute to normalized matrix elements of Λ (this point will be justified later), and we will discard it in the following considerations.

So, we might finally write that $\Lambda(t,t_0) = U(t,t_0)$.

Note

It is not difficult to generalize $\Lambda(t,t_0)$ by allowing its second argument to take on other values than the "reference time" t_0. The correct form is quite natural

$$\Lambda(t,t') = T\left[e^{-i\int_{t'}^{t} d\tau \overline{H}_I(\tau)}\right], \quad (t' \leqslant t). \tag{4.161}$$

Let us check that this is the correct prescription. First, $\Lambda(t,t')$ satisfies the same differential equation as $\Lambda(t,t_0)$, i.e.

$$i\frac{\partial}{\partial t}\Lambda(t,t') = \overline{H}_I \Lambda(t,t'), \tag{4.162}$$

but now with the initial condition $\Lambda = 1$ for $t = t'$. In addition, it can be seen that

$$\begin{aligned}\Lambda(t,t') &= e^{iH_0(t-t_0)}e^{-iH(t-t')}e^{-iH_0(t'-t_0)} \\ &= \Lambda(t,t_0)\Lambda^{-1}(t',t_0).\end{aligned} \tag{4.163}$$

Indeed,

$$\begin{aligned}i\frac{\partial}{\partial t}\Lambda(t,t') &= -H_0\Lambda(t,t') \\ &\quad + e^{iH_0(t-t_0)}H\,\overbrace{e^{-iH_0(t-t_0)}e^{iH_0(t-t_0)}}^{1}\,e^{-iH(t-t')}e^{-iH_0(t'-t_0)} \\ &= -H_0\Lambda(t,t') + \underbrace{\overline{H}}_{H_0+\overline{H}_I}\Lambda(t,t') \\ &= \overline{H}_I\Lambda(t,t').\end{aligned} \tag{4.164}$$

Finally, for $t' = t_0$ we get back our original $\Lambda(t,t_0)$.

Note that $\Lambda(t,t')$ satisfies the following important properties:

- For $t_1 \geq t_2 \geq t_3$,
$$\begin{aligned}\Lambda(t_1,t_2)\Lambda(t_2,t_3) &= \Lambda(t_1,t_0)\Lambda^{-1}(t_2,t_0)\Lambda(t_2,t_0)\Lambda(t_3,t_0) \\ &= \Lambda(t_1,t_3).\end{aligned} \tag{4.165}$$

- $$\begin{aligned}\left[\Lambda(t_1,t_2)\right]^\dagger &= \left[\Lambda(t_1,t_0)\Lambda^{-1}(t_2,t_0)\right]^\dagger = \Lambda^{-1\dagger}(t_2,t_0)\Lambda^\dagger(t_1,t_0) \\ &= \Lambda(t_2,t_1) = \Lambda^{-1}(t_1,t_2).\end{aligned} \tag{4.166}$$

- $$\begin{aligned}\Lambda(t_1,t_3)[\Lambda(t_2,t_3)]^\dagger &= \Lambda(t_1,t_0)\Lambda^{-1}(t_2,t_0)[\Lambda(t_2,t_0)\Lambda^{-1}(t_3,t_0)]^{-1} \\ &= \Lambda(t_1,t_0)\Lambda^{-1}(t_2,t_0) = \Lambda(t_1,t_2).\end{aligned} \tag{4.167}$$

The actual role of $\Lambda(t,t')$ will become clear shortly.

Since we require (as usual in quantum mechanics) that scalar product should in both pictures coincide, we have for any operators $A(t, \boldsymbol{x})$ and any pair of states $|\psi\rangle$ and $|\psi'\rangle$

$$\begin{aligned} \langle\psi_H| A_H(t,\boldsymbol{x}) |\psi'_H\rangle &= \langle\psi_H| \Lambda^{-1}(t,t_0) A_I(t,\boldsymbol{x}) \Lambda(t,t_0) |\psi'_H\rangle \\ &\stackrel{!}{=} \langle\psi_I(t)| A_I(t,\boldsymbol{x}) |\psi'_I(t)\rangle \ . \end{aligned} \tag{4.168}$$

This implies that

$$\begin{aligned} \lambda^{-1}(t,t_0)\Lambda(t,t_0) |\psi_H\rangle &= |\psi_I(t)\rangle \\ \Leftrightarrow \qquad |\psi_H\rangle &= \lambda(t,t_0)\Lambda^{-1}(t,t_0) |\psi_I(t)\rangle \ . \end{aligned} \tag{4.169}$$

Where λ is a phase factor with $|\lambda| = 1$.

Let us briefly repeat the key results:

Note — Summary of key results

$$\phi_I(t,\boldsymbol{x}) = \Lambda(t,t_0)\phi_H(t,\boldsymbol{x}) \underbrace{\Lambda^{-1}(t,t_0)}_{\Lambda^\dagger(t,t_0)} , \tag{4.170}$$

$$\begin{aligned} \Lambda(t,t_0) &= T\left[e^{-i\int_{t_0}^t d\tau \overline{H_I}(\tau)}\right] = T\left[e^{i\int_{t_0}^t d\tau d^3\boldsymbol{x} \overline{\mathcal{L}_I}(\tau,\boldsymbol{x})}\right] \\ &= e^{i(t-t_0)H_0^I} e^{-i(t-t_0)H} , \end{aligned} \tag{4.171}$$

$$\begin{aligned} \Lambda(t,t') &= T\left[e^{-i\int_{t'}^t d\tau \overline{H_I}(\tau)}\right] = T\left[e^{i\int_{t'}^t d\tau d^3\boldsymbol{x} \overline{\mathcal{L}_I}(\tau,\boldsymbol{x})}\right] \\ &= e^{i(t-t_0)H_0^I} e^{-i(t-t_0)H} e^{-iH_0^I(t'-t_0)} \\ &= \Lambda(t,t_0)\Lambda^{-1}(t',t_0) . \end{aligned} \tag{4.172}$$

These representations of Λ are known as Dyson representations. The time ordered exponential is known as Dyson expansion.

Let us now assume that we can adiabatically switch off the interaction at $t \to -\infty$ so that in the remote past $H_I = 0$ and hence $\phi_I(t,\boldsymbol{x}) = \phi_H(t,\boldsymbol{x})$ for $t \to -\infty$. This allows us to identify t_0 with $-\infty$.

The field $\phi_H(t,\boldsymbol{x})$ contains the information about the interaction, since it evolves over time with the full Hamiltonian. In order to describe the "in" and "out" field operator we can now make the following identifications

$$\begin{aligned} t \to -\infty : &\quad \phi_{in}(\boldsymbol{x},t) = \phi_I(\boldsymbol{x},t) = \phi_H(\boldsymbol{x},t) , \\ t \to +\infty : &\quad \phi_{out}(\boldsymbol{x},t) = \phi_H(\boldsymbol{x},t) . \end{aligned} \tag{4.173}$$

Furthermore, since the fields ϕ_I evolve over time with the free Hamiltonian H_0, they always act in the basis of "in" state vectors, such that

$$\phi_{in}(\boldsymbol{x},t) = \phi_I(\boldsymbol{x},t), \qquad -\infty < t < +\infty . \tag{4.174}$$

Note, when ϕ_I and ϕ_H coincide at different times, say times t_0^1 or t_0^2,

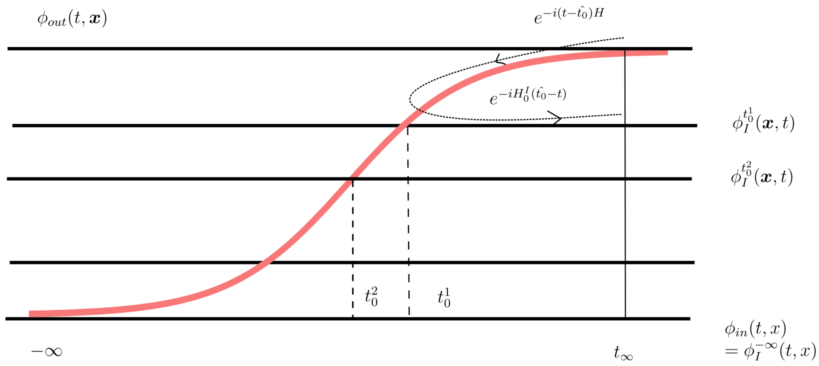


Figure 4.5: Transformation between Heisenberg and interaction pictures and connection with ϕ_{in} and ϕ_{out} fields.

they are related via different unitary transformation, namely

$$\begin{aligned}\phi_I^{t_0^1}(t, \boldsymbol{x}) &= \Lambda(t, t_0^1)\phi_H(t, \boldsymbol{x})\Lambda^\dagger(t, t_0^1), \\ \phi_I^{t_0^2}(t, \boldsymbol{x}) &= \Lambda(t, t_0^2)\phi_H(t, \boldsymbol{x})\Lambda^\dagger(t, t_0^2). \end{aligned} \quad (4.175)$$

Here we use a temporal notation $\phi_I^{t_0}(t, \boldsymbol{x})$ to denote different boundary values for ϕ_I's, see Fig. 4.5. So, in particular, from the relation $\phi_I^{-\infty}(\boldsymbol{x}, t) = \Lambda(t, -\infty)\phi_H(\boldsymbol{x}, t)\Lambda^\dagger(t, -\infty)$ follows that

$$\phi_{in}(\boldsymbol{x}, t) = \Lambda(t, -\infty)\phi_H(\boldsymbol{x}, t)\Lambda^\dagger(t, -\infty). \quad (4.176)$$

As we let $t \to \infty$ we can use the identification (4.173) to write

$$\phi_{in}(\boldsymbol{x}, \infty) = \underbrace{\Lambda(\infty, -\infty)}_{S}\, \phi_{out}(\boldsymbol{x}, \infty)\, \underbrace{\Lambda^\dagger(\infty, -\infty)}_{S^\dagger}. \quad (4.177)$$

Note that ϕ_{in} (as any free field) allows to define corresponding set of creation and annihilation operators $a_{in}(p), a_{in}^\dagger(p)$ and the vacuum state $|0\rangle_{in}$ ($a_{in}(p)\,|0\rangle_{in} = 0$). Similarly, from ϕ_{out} we have creation and annihilation operators $a_{out}(p), a_{out}^\dagger(p)$ and the vacuum state $|0\rangle_{out}$. In addition, from the relation

$${}_{in}\langle 0|\, \phi_{in}^2\, |0\rangle_{in} = {}_{out}\langle 0|\, \phi_{out}^2\, |0\rangle_{out} = {}_{in}\langle 0|S\phi_{out}^2 S^\dagger|0\rangle_{in}, \quad (4.178)$$

we see that $|0\rangle_{out} = S^\dagger|0\rangle_{in}$. By taking into account also the fact that (4.177) implies

$$\begin{aligned} a_{in}(p) &= S a_{out}(p) S^\dagger, \\ a_{in}^\dagger(p) &= S a_{out}^\dagger(p) S^\dagger, \end{aligned} \quad (4.179)$$

we can immediately write

$$a_{in}^\dagger\, |0\rangle_{in} = |p\rangle_{in} = S a_{out}^\dagger(p) S^\dagger S\, |0\rangle_{out} = S\, |p\rangle_{out}, \quad (4.180)$$

and similarly for multi-particle states

$$|p_1, p_2, \ldots\rangle_{in} = S\, |p_1, p_2, \ldots\rangle_{out}. \quad (4.181)$$

We can denote this in a schematic way as

$$|i\rangle_{in} \;=\; S\,|f\rangle_{out}\;, \tag{4.182}$$

(i stands for *initial*-state particle configuration and f for the *final*-state particle configuration). Here the S-operator is better known under the name S-matrix, and it allows for unitary transformation that connects *in*-fields with *out*-fields. In particular, the rate for $|p_1, p_2\rangle_{in} \rightarrow |p_3, p_4, \ldots\rangle_{out}$, is obtained from the matrix element

$${}_{out}\langle f|i\rangle_{in} \;=\; {}_{out}\langle p_3, p_4, p_5, \ldots|p_1, p_2\rangle_{in}\,. \tag{4.183}$$

By noting that

$$|f\rangle_{out} \;=\; S^{\dagger}\,|f\rangle_{in}\,, \quad |i\rangle_{out} \;=\; S^{\dagger}\,|i\rangle_{in}\;, \tag{4.184}$$

we also have

$$S\,|f\rangle_{out} \;=\; |f\rangle_{in}\,, \quad S\,|i\rangle_{out} \;=\; |i\rangle_{in}\;, \tag{4.185}$$

which is equivalent to

$${}_{out}\,\langle Sf| \;=\; {}_{in}\,\langle f|\,, \qquad {}_{out}\,\langle Si| \;=\; {}_{in}\,\langle i|\;. \tag{4.186}$$

The matrix element

$$\begin{aligned}
S_{fi} \;&=\; {}_{out}\langle f|i\rangle_{in} \;=\; {}_{out}\,\langle f|\,S^{\dagger}S\,|i\rangle_{in} \;=\; {}_{out}\,\langle Sf|\,S\,|i\rangle_{in}\\
&=\; {}_{in}\,\langle f|\,S\,|i\rangle_{in} \;=\; {}_{out}\,\langle f|\,SS^{\dagger}\,|i\rangle_{in} \;=\; {}_{out}\,\langle f|\,S\,|S^{\dagger}i\rangle_{in}\\
&=\; {}_{out}\,\langle f|\,S\,|i\rangle_{out}\;,
\end{aligned} \tag{4.187}$$

is known as *scattering transition amplitude*. Generally, one can write the S-matrix in the Dyson expansion form as (cf. Eq. (4.177))

$$\begin{aligned}
S \;&=\; T\left[\exp\left\{i\int_{-\infty}^{\infty}\mathrm{d}t\int_{\mathbb{R}^3}\mathrm{d}^3\boldsymbol{x}\,\mathcal{L}_I\left(\phi(x), \partial_\mu\phi(x)\right)\right\}\right]\\
&=\; T\left[\exp\left\{i\int_{\mathbb{R}^4}\mathrm{d}^4x\,\mathcal{L}_I\left(\phi(x), \partial_\mu\phi(x)\right)\right\}\right]\,.
\end{aligned} \tag{4.188}$$

S-matrix contains all physical information for any scattering process in the theory described by given Lagrangian, since any transition amplitude can be computed from it.

Recall, that the key assumption in our description of scattering processes was the *adiabatic hypothesis*, i.e. the assumption that one can switch off the interaction slowly for large positive and large negative times without changing the physics. For many purposes this is indeed a sensible assumption. However, we will see a bit later (when discussing renormalization) that this description is too simplistic.

Final upshot of this discussion is that in order to describe any realistic scattering, e.g., the scattering process $|p_1, p_2\rangle_{in} \rightarrow |p_3, p_4, p_5, \ldots\rangle_{out}$

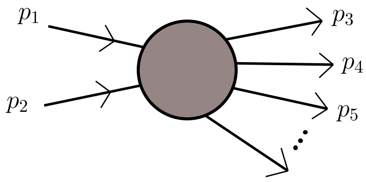


Figure 4.6: Schematic representation of the scaterring process $|p_1, p_2\rangle_{in} \to |p_3, p_4, p_5, \ldots\rangle_{out}$.

we must compute the transition amplitude

$$_{out}\langle p_3, p_4, p_5, \ldots | p_1, p_2\rangle_{in} \; = \; _{in}\langle p_3, p_4, p_5, \ldots | S \, | p_1, p_2\rangle_{in} \,. \tag{4.189}$$

Try to fill the gaps and generalize the present analysis also to non-Hermitian (i.e, charged) scalar fields. Do you find any substantial difference?

Advantage of this form is that fields entering in S are the interaction-picture (i.e. free) fields that coincide with ϕ_H at $t \to -\infty$. So, the entire S matrix is phrased in terms of free fields, and hence in the terms of creation and annihilation operators $a^\dagger_{in}(p)$ and $a_{in}(p)$, respectively.

Time ordered product and Wick's theorem

To compute the S-matrix, we need to know how to systematically compute time ordered products of free fields.

Let us begin with *2 free fields*. In this case

$$T\left[\phi(x)\phi(y)\right] \; = \; \theta(x^0 - y^0)\phi(x)\phi(y) \; + \; \theta(y^0 - x^0)\phi(y)\phi(x)\,, \tag{4.190}$$

where

$$\phi(x) \; = \; \sum_{\boldsymbol{p}} \left[a(p)e^{-ipx} + a^\dagger(p)e^{ipx}\right] \; = \; \phi^{(+)}(x) \; + \; \phi^{(-)}(x)\,. \tag{4.191}$$

For $x^0 > y^0$ we can write

$$\begin{aligned} T\left[\phi(x)\phi(y)\right] \; &= \; \phi(x)\phi(y) \; = \; \phi^{(+)}(x)\phi^{(+)}(y) \; + \; \phi^{(+)}(x)\phi^{(-)}(y) \\ &+ \; \phi^{(-)}(x)\phi^{(+)}(y) \; + \; \phi^{(-)}(x)\phi^{(-)}(y) \\ &= \; \phi^{(+)}(x)\phi^{(+)}(y) \; + \; \phi^{(-)}(x)\phi^{(-)}(y) \; + \; \phi^{(-)}(x)\phi^{(+)}(y) \\ &+ \; \left(\phi^{(+)}(x)\phi^{(-)}(y) - \phi^{(-)}(y)\phi^{(+)}(x)\right) \; + \; \phi^{(-)}(y)\phi^{(+)}(x) \\ &= \; :\phi(x)\phi(y): \; + \; \left[\phi^{(+)}(x), \phi^{(-)}(y)\right]. \end{aligned} \tag{4.192}$$

Since

$$\begin{aligned} \left[\phi^{(+)}(x), \phi^{(-)}(y)\right] \; &= \; \sum_{\boldsymbol{p},\boldsymbol{p}'} \left[a(p)e^{-ipx}, a^\dagger(p')e^{ip'y}\right] \\ &= \; \sum_{\boldsymbol{p},\boldsymbol{p}'} \delta_{pp'}e^{-ipx+ip'y} \; = \; \sum_{\boldsymbol{p}} e^{-ip(x-y)}\,, \end{aligned} \tag{4.193}$$

one can write

$$T\left[\phi(x)\phi(y)\right] = :\phi(x)\phi(y): + \sum_{\boldsymbol{p}} e^{-ip(x-y)} . \qquad (4.194)$$

Similarly, for $x^0 < y^0$ one can easily show that

$$T\left[\phi(x)\phi(y)\right] = \phi(y)\phi(x) = :\phi(x)\phi(y): + \sum_{\boldsymbol{p}} e^{ip(x-y)} . \qquad (4.195)$$

By combining (4.194) and (4.195) together we obtain

$$\begin{aligned} T\left[\phi(x)\phi(y)\right] &= :\phi(x)\phi(y): + \theta(x^0 - y^0)\sum_{\boldsymbol{p}} e^{-ip(x-y)} \\ &\quad + \theta(y^0 - x^0)\sum_{\boldsymbol{p}} e^{ip(x-y)} \\ &= :\phi(x)\phi(y): + i\Delta_F(x-y) , \end{aligned} \qquad (4.196)$$

The later implies, as a byproduct, that the vacuum expectation value of the corresponding time ordered product is

$$\langle 0|\, T\left[\phi(x)\phi(y)\right]|0\rangle = \langle 0|:\phi(x)\phi(y):|0\rangle + i\Delta_F(x-y) , \qquad (4.197)$$

Since the vacuum expectation value of the normal ordered product is zero, we have

$$\langle 0|\, T\left[\phi(x), \phi(y)\right]|0\rangle = i\Delta_F(x-y) . \qquad (4.198)$$

Similarly, for *3 free fields* it can be checked that

$$\begin{aligned} T\left[\phi(x_1)\phi(x_2)\phi(x_3)\right] &= :\phi(x_1)\phi(x_2)\phi(x_3): + \phi(x_1)i\Delta_F(x_2 - x_3) \\ &+ \phi(x_2)i\Delta_F(x_3 - x_1) + \phi(x_3)i\Delta_F(x_1 - x_2) , \end{aligned} \qquad (4.199)$$

and for *4 free fields* one obtains

$$\begin{aligned} T\left[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\right] &= :\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4): \\ &+ :\phi(x_1)\phi(x_2):i\Delta_F(x_3 - x_4) + :\phi(x_1)\phi(x_3):i\Delta_F(x_2 - x_4) \\ &\vdots \\ &+ i\Delta_F(x_1 - x_2)i\Delta_F(x_3 - x_4) + i\Delta_F(x_1 - x_3)i\Delta_F(x_2 - x_4) \\ &+ i\Delta_F(x_1 - x_4)i\Delta_F(x_2 - x_3) . \end{aligned} \qquad (4.200)$$

The Feynman propagator $i\Delta_F(x-y)$ is often referred to as the *contraction* of the fields $\phi(x)$ and $\phi(y)$ and it is sometimes denoted as

$$\ldots\ \overbrace{\phi(x)\ \ldots\ \phi}(y)\ \ldots\ = \ \ldots i\Delta_F(x-y) . \qquad (4.201)$$

The so-called *Wick's theorem* allows us to rephrase the time ordered product of *N free fields* in terms of normal ordering and field contractions, namely

Here we use an abbreviated notation $\phi(x_1)\ldots\phi(x_n) = 1\ldots n$.

$$\begin{aligned}
T\,[123\ldots N] \;=&: 123\ldots N: \;+\; :123\ldots(N-2): i\Delta_F(x_N - x_{N-1}) \\
&+ \text{"all other terms with 1 contraction"} \\
&+ \;:123\ldots(N-4): [i\Delta_F(x_N - x_{N-1}) i\Delta_F(x_{N-2} - x_{N-3}) \\
&+ i\Delta_F(x_N - x_{N-2}) i\Delta_F(x_{N-1} - x_{N-3}) \\
&+ \; i\Delta_F(x_N - x_{N-3}) i\Delta_F(x_{N-1} - x_{N-2})] \\
&+ \text{"all other terms with 2 contractions "} \\
&\quad \vdots \\
&+ \text{"till all contractions are exhausted"}\,.
\end{aligned}$$

Note that if N is *odd* then there is at least one normal ordered field product in each term. Therefore, for N odd

$$\langle 0|\, T\,[\phi(x_1)\ldots\phi(x_N)]\,|0\rangle \;=\; 0\,. \tag{4.202}$$

One can prove Wick's theorem by induction, but this is tedious and unenlightening. It is more instructive to prove the following theorem

Theorem — Generating functional for Wick's theorem

Let $J(x)$ be a c-number function, then

$$\begin{aligned}
T\left[\exp\left(-i\int d^4x\, J(x)\hat{\phi}(x)\right)\right] &= :\exp\left(-i\int d^4x J(x)\hat{\phi}(x)\right): \\
&\quad\times \exp\left(-\frac{1}{2}\int d^4x\, d^4y J(x)\,\langle 0|\, T\,[\phi(x)\phi(y)]\,|0\rangle\, J(y)\right) \\
&= :\exp\left(-i\int d^4x J(x)\hat{\phi}(x)\right): \\
&\quad\times \exp\left(-\frac{i}{2}\int d^4x\, d^4y\, J(x)\Delta_F(x-y)J(y)\right)\,.
\end{aligned} \tag{4.203}$$

Real $J(x)$ ensures that the generating functional in (4.203) is unitary. On the other hand, purely imaginary $J(x) = iK(x)$ allows to better organize terms in the expansion.

Before we prove this theorem, let us begin with a small comment. By replacing $J(x)$ with $iK(x)$, expanding out and comparing coefficients we recover the Wick theorem. For instance, let us restrict ourselves to the second order in K, then

$$\begin{aligned}
&T\left[1 \;+\; \int d^4x K(x)\hat{\phi}(x) \;+\; \frac{1}{2}\int d^4x\, d^4y\, K(x)K(y)\hat{\phi}(x)\hat{\phi}(y) + \ldots\right] \\
&\quad = :1 \;+\; \int d^4x K(x)\hat{\phi}(x) \;+\; \frac{1}{2}\int dx^4\, d^4y\, K(x)K(y)\hat{\phi}(x)\hat{\phi}(y) \;+\; \ldots: \\
&\qquad \times \left(1 + \frac{1}{2}\int d^4x d^4y\, K(x)K(y)\,\langle 0|\, T\left[\hat{\phi}(x)\hat{\phi}(y)\right]|0\rangle \;+\; \ldots\right),
\end{aligned}$$

implies that

$$\frac{1}{2}\int d^4xd^4y\,K(x)K(y)T\left[\phi(x)\phi(y)\right]$$
$$= \frac{1}{2}\int d^4xd^4y\,K(x)K(y)\left[:\phi(x)\phi(y): \;+\; \langle 0|\,T\left[\phi(x)\phi(y)\right]|0\rangle\right], \quad (4.204)$$

which in turn implies that

$$T\left[\phi(x)\phi(y)\right] \;=\; :\phi(x)\phi(y): \;+\, i\Delta_F(x-y)\,. \quad (4.205)$$

One can proceed similarly also for higher orders in K.

This coincides with Eq. (4.196).

Let us now prove the above theorem. We first recall the Baker–Campbell–Hausdorff (BCH) formula

$$e^{\hat{A}}e^{\hat{B}} \;=\; e^{\hat{A}\,+\,\hat{B}\,+\,\frac{1}{2}\left[\hat{A},\hat{B}\right]}, \quad (4.206)$$

which holds for any pair of operators $\hat{A}$ and $\hat{B}$ provided that $\hat{A}$, $\hat{B}$ commute with their commutator $\left[\hat{A},\hat{B}\right]$.

In our case the role of $\hat{A}$, $\hat{B}$ will be played by free fields $\phi(x)$, $\phi(y)$. Since they commute to c-number (Pauli-Jordan function), the assumption will be satisfied. Consider $t_n > t_{n-1} > \cdots > t_1$ and set $\hat{X}(t) = \int d^3x J(x)\hat{\phi}(x)$. With this we can break up the time-ordered product in an approximate way as

$$T\left[e^{-i\int dt\hat{X}(t)}\right] \;\approx\; e^{-i\Delta t\hat{X}(t_{n-1})}e^{-i\Delta t\hat{X}(t_n)}\ldots e^{-i\Delta t\hat{X}(t_1)}$$
$$= \; e^{-i\Delta t\sum_{i=1}^{n}\hat{X}(t_i)\,-\,\frac{1}{2}(\Delta t)^2\sum_{k>l}\left[\hat{X}(t_k),\hat{X}(t_l)\right]}\,. \quad (4.207)$$

Recall that operators after T symbol behave as commuting operators.

where on the second line we used (4.206). Taking the limit as $\Delta t \to dt$, the expression (4.207) turns to

$$T\left[e^{-i\int d^4x\,J(x)\hat{\phi}(x)}\right] \;=\; e^{-i\int d^4xJ(x)\hat{\phi}(x)\,-\,\frac{1}{2}\int d^4xd^4yJ(x)J(y)\left[\hat{\phi}(x),\hat{\phi}(y)\right]\theta(x_0-y_0)}$$
$$= \; e^{-i\int d^4x\;J(x)\hat{\phi}(x)}$$
$$\times\; e^{-\frac{1}{2}\int d^4xd^4y\;J(x)J(y)\left[\hat{\phi}(x),\hat{\phi}(y)\right]\theta(x_0-y_0)}, \quad (4.208)$$

where in the last identity we used the fact that the term with commutator is a c-numbered function and hence it can be factored out from the exponential.

Eq. (4.208) is a nice result in itself —- but not yet what we need. We now note that

$$:e^{-i\int d^4xJ(x)\hat{\phi}(x)}: \;=\; e^{-i\int d^4xJ(x)\hat{\phi}^{(-)}(x)}e^{-i\int d^4xJ(x)\hat{\phi}^{(+)}(x)}$$
$$= \; e^{-i\int d^4xJ(x)\hat{\phi}(x)}$$
$$\times\; e^{-\frac{1}{2}\int d^4xd^4yJ(x)J(y)\left[\hat{\phi}^{(-)}(x),\hat{\phi}^{(+)}(y)\right]}\,. \quad (4.209)$$

The BCH formula was used in the second identity. By combining both

(4.208) and (4.209) we obtain

$$T\left[e^{-i\int d^4x J(x)\hat{\phi}(x)}\right] \;=\; :e^{-i\int d^4x J(x)\hat{\phi}(x)}: \times\; e^{\frac{1}{2}\int d^4x d^4y J_x J_y \left\{\left[\hat{\phi}_x^{(-)},\hat{\phi}_y^{(+)}\right]-\theta(x_0-y_0)\left[\hat{\phi}_x,\hat{\phi}_y\right]\right\}}. \quad (4.210)$$

The expression $\left[\hat{\phi}_x^{(-)}, \hat{\phi}_y^{(+)}\right] - \theta(x_0 - y_0)\left[\hat{\phi}_x, \hat{\phi}_y\right]$ is a c-number and hence it can be conveniently evaluated by taking a vacuum expectation value from it, i.e.

In passing from 1st to 2nd line we employed the fact that $\hat{\phi}_x^{(-)} \sim a^\dagger$ and $\hat{\phi}_y^{(+)} \sim a$. Thus, the only surviving part of the first commutator is $\hat{\phi}_y^{(+)}\hat{\phi}_x^{(-)}$.

$$\begin{aligned}
&\langle 0|\left[\hat{\phi}_x^{(-)}, \hat{\phi}_y^{(+)}\right] - \theta(x_0 - y_0)\left[\hat{\phi}_x, \hat{\phi}_y\right]|0\rangle \\
&= -\langle 0|\hat{\phi}_y^{(+)}\hat{\phi}_x^{(-)}|0\rangle - \theta(x_0-y_0)\langle 0|\hat{\phi}_x\hat{\phi}_y|0\rangle + \theta(x_0-y_0)\langle 0|\hat{\phi}_y\hat{\phi}_x|0\rangle \\
&= -\langle 0|\hat{\phi}_y\hat{\phi}_x|0\rangle \cdot \underbrace{\mathbf{1}}_{=\,\theta(x_0-y_0)\,+\,\theta(y_0-x_0)} - \theta(x_0-y_0)\langle 0|\hat{\phi}_x\hat{\phi}_y|0\rangle + \theta(x_0-y_0)\langle 0|\hat{\phi}_y\hat{\phi}_x|0\rangle \\
&= -\theta(y_0-x_0)\langle 0|\hat{\phi}_y\hat{\phi}_x|0\rangle - \theta(x_0-y_0)\langle 0|\hat{\phi}_x\hat{\phi}_y|0\rangle \\
&= -\langle 0|T\left[\hat{\phi}_x\hat{\phi}_y\right]|0\rangle \;=\; -i\Delta_F(x-y)\,. \qquad (4.211)
\end{aligned}$$

If we now compare (4.210) with (4.211) we obtain the desired generating functional for Wick's theorem.

An important implication of the previous "operatorial" version of Wick's theorem is the weaker version of Wick's theorem for vacuum expectation values (also known as Wick's theorem). Note that

$$\langle 0|T\left[e^{-i\int d^4x J(x)\hat{\phi}(x)}\right]|0\rangle \;=\; \underbrace{\langle 0|:e^{-i\int d^4x J(x)\hat{\phi}(x)}:|0\rangle}_{1} \times\; e^{-\frac{1}{2}\int d^4x d^4y J(x)J(y)\langle 0|T[\phi(x)\phi(y)]|0\rangle}\,. \quad (4.212)$$

Again expansions in K ($J = iK$) provide important relation between $\langle 0|T[12\ldots N]|0\rangle$ and $\langle 0|T[i,j]|0\rangle = i\Delta_F(x_i - x_j)$. Wick's theorem in the form (4.212) will be particularly important in what follows.

For instance, to fourth order in K we get the identity

$$\frac{\delta^4}{\delta K_{y_4}\delta K_{y_3}\delta K_{y_2}\delta K_{y_1}}\langle 0|T\left[e^{-\int d^4x K_x\hat{\phi}_x}\right]|0\rangle\bigg|_{K=0} = \frac{\delta^4}{\delta K_{y_4}\delta K_{y_3}\delta K_{y_2}\delta K_{y_1}} e^{\frac{1}{2}\int d^4x d^4y K_x K_y\langle 0|T\left[\phi_x,\phi_y\right]|0\rangle}\bigg|_{K=0}\,. \quad (4.213)$$

The simplest way to compute the derivatives is to expand each exponent and keep only the fourth order in K since no other term can contribute. The left hand side of (4.213) thus reduces to

$$\frac{\delta^4}{\delta K_{y_4}\delta K_{y_3}\delta K_{y_2}\delta K_{y_1}}\frac{1}{4!}\int d^4x_1\ldots d^4x_4 K_{x_1}\ldots K_{x_4}\langle 0|T[\phi_{x_1}\phi_{x_2}\phi_{x_3}\phi_{x_4}]|0\rangle = \langle 0|T[\phi_{y_1}\phi_{y_2}\phi_{y_3}\phi_{y_4}]|0\rangle\,. \quad (4.214)$$

Here we have used the fact that $\langle 0|\, T[\phi_{x_1}\phi_{x_2}\phi_{x_3}\phi_{x_4}]\,|0\rangle$ is a symmetric function of its arguments.

The right hand side of (4.213) can be then written as

$$\frac{\delta^4}{\delta K_{y_4}\delta K_{y_3}\delta K_{y_2}\delta K_{y_1}}\frac{1}{4}\frac{1}{2!}\int d^4x_1\ldots d^4x_4\, K_{x_1}\ldots K_{x_4} \times \langle 0|\, T\left[\phi_{x_1}\phi_{x_2}\right]|0\rangle\,\langle 0|\, T\left[\phi_{x_3}\phi_{x_4}\right]|0\rangle\;. \qquad (4.215)$$

So, for instance, the first functional derivative gives

$$\begin{aligned}
&\frac{\delta}{\delta K_{y_1}}\frac{1}{4}\frac{1}{2!}\int d^4x_1\ldots d^4x_4\, K_{x_1}\ldots K_{x_4} \times \langle 0|\, T\left[\phi_{x_1}\phi_{x_2}\right]|0\rangle\,\langle 0|\, T\left[\phi_{x_3}\phi_{x_4}\right]|0\rangle\\
&= \frac{1}{8}\int d^4x_2 d^4x_3 d^4x_4\, K_{x_2}K_{x_3}K_{x_4}\,\langle 0|\, T\left[\phi_{y_1}\phi_{x_2}\right]|0\rangle\,\langle 0|\, T\left[\phi_{x_3}\phi_{x_4}\right]|0\rangle\\
&+ \frac{1}{8}\int d^4x_1 d^4x_3 d^4x_4\, K_{x_1}K_{x_3}K_{x_4}\,\langle 0|\, T\left[\phi_{x_1}\phi_{y_1}\right]|0\rangle\,\langle 0|\, T\left[\phi_{x_3}\phi_{x_4}\right]|0\rangle\\
&+ \frac{1}{8}\int d^4x_1 d^4x_2 d^4x_4\, K_{x_1}K_{x_2}K_{x_4}\,\langle 0|\, T\left[\phi_{x_1}\phi_{x_2}\right]|0\rangle\,\langle 0|\, T\left[\phi_{y_1}\phi_{x_4}\right]|0\rangle\\
&+ \frac{1}{8}\int d^4x_1 d^4x_2 d^4x_3\, K_{x_1}K_{x_2}K_{x_3}\,\langle 0|\, T\left[\phi_{x_1}\phi_{x_2}\right]|0\rangle\,\langle 0|\, T\left[\phi_{x_3}\phi_{y_1}\right]|0\rangle\;.
\end{aligned}$$

By proceeding with remaining 3 functional derivatives we arrive at three following types of terms

$$\langle 0|\, T\left[\phi_{y_1}\phi_{y_2}\right]|0\rangle\,\langle 0|\, T\left[\phi_{y_3}\phi_{y_4}\right]|0\rangle\;, \qquad (4.216)$$

$$\langle 0|\, T\left[\phi_{y_1}\phi_{y_3}\right]|0\rangle\,\langle 0|\, T\left[\phi_{y_2}\phi_{y_4}\right]|0\rangle\;, \qquad (4.217)$$

$$\langle 0|\, T\left[\phi_{y_1}\phi_{y_4}\right]|0\rangle\,\langle 0|\, T\left[\phi_{y_2}\phi_{y_3}\right]|0\rangle\;. \qquad (4.218)$$

Since $T\left[\phi_{y_1}\phi_{y_2}\right] = T\left[\phi_{y_2}\phi_{y_1}\right]$, each term of the form (4.216), (4.217) and (4.218) will be generated with the multiplicity of 8.

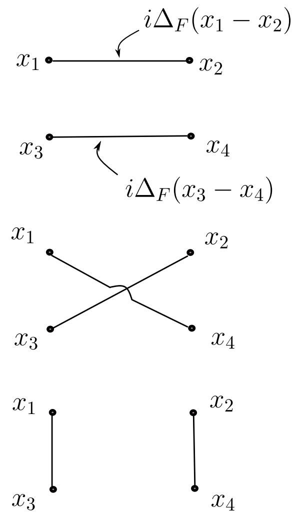
So, we finally obtain that

$$\begin{aligned}
\langle 0|\, T\,[1234]\,|0\rangle &= \langle 0|\, T\,[12]\,|0\rangle\,\langle 0|\, T\,[34]\,|0\rangle\\
&+ \langle 0|\, T\,[13]\,|0\rangle\,\langle 0|\, T\,[24]\,|0\rangle\\
&+ \langle 0|\, T\,[14]\,|0\rangle\,\langle 0|\, T\,[23]\,|0\rangle\;. \qquad (4.219)
\end{aligned}$$

Graphically one can represent $\langle 0|\, T\,[1234]\,|0\rangle$ from the Wick's expansion (4.219) as:

More generally, for any even N we get

$$\begin{aligned}
\langle 0|\, T\,[1\ldots N]\,|0\rangle &= \langle 0|\, T\,[12]\,|0\rangle\,\langle 0|\, T\,[34]\,|0\rangle\ldots\langle 0|\, T\,[(N-1)N]\,|0\rangle\\
&+ \text{“all other distinct contractions”}\,. \qquad (4.220)
\end{aligned}$$



Let us recall that for N odd this would be zero. This formula will be a basis of a perturbation evaluation of the S-matrix elements.

At this stage it is interesting to ask how many distinct terms (i.e., distinct products of $i\Delta_F$'s) can be generated in Wick's expansion from generic $\langle 0| T[1\ldots N] |0\rangle$. Let us set $N = 2M$ with M being a general positive integer. Using again an abbreviated notation $\langle x_1 \ldots x_{2M}\rangle$, the result will be composed of 1st pairing that will comprise $N-1$ contractions, 2nd pairing that will comprise $N-3$ contractions, etc. Schematically

$$\begin{aligned}
&\langle x_1 x_k\rangle\langle x_2 \ldots \hat{x}_k \ldots x_{2M}\rangle && (N-1)\ \text{contractions}\,, \\
&\langle x_1 x_k\rangle\langle x_2 x_l\rangle\langle x_3 \ldots \hat{x}_l \ldots \hat{x}_k \ldots x_{2M}\rangle && (N-3)\ \text{contractions}\,, \\
&\quad\vdots
\end{aligned} \tag{4.221}$$

which together yields the total number of

$$\begin{aligned}
&(2M-1)\times(2M-3)\times(2M-5)\times\cdots\times(2M-(2M-1)) \\
&\quad = \frac{2M\times(2M-1)\times(2M-2)\times(2M-3)\times\cdots\times(2M-(2M-1))}{2M\times 2(M-1)\times 2(M-2)\times\cdots\times 2} \\
&\quad = \frac{(2M)!}{2^M M!}\,,
\end{aligned} \tag{4.222}$$

E.g. for $M = 2$ we see that the number of terms is equal to $4!/(4\cdot 2!) = 3$, in accordance with our previous result (4.219).

of contractions.

$$M = 3 \Rightarrow \frac{6!}{2^3 3!} = 15\,.$$

As an exercise, try to find explicitly all 15 terms in the Wick's expansion of $\langle 0| T[123456] |0\rangle$.

4.7 Green functions — Gell-Mann and Low formula

Experimentalists are typically interested in matrix elements of the S matrix, e.g. $_{in}\langle p_1 p_2| S |p_3 p_4 \ldots\rangle_{in}$. From these elements one can compute directly differential cross-sections in scattering experiments as we will see in chapter dedicated to scattering processes. Such computations are typically done perturbatively in terms of the so-called Feynman diagrams. There exists a very efficient way to the perturbative treatment (and ensuing Feynman diagrams) that is based on the vacuum expectation value of the time-ordered products of Heisenberg fields $\phi_H(x)$, i.e.

These expressions are also known as generalized Green functions or correlators. Here, $|0\rangle$ is a *true ground state* of the interacting system.

$$\begin{aligned}\tau(x_1, x_2, \ldots, x_n) &\equiv \langle x_1 \ldots x_n\rangle \\ &\equiv \langle 0| T\left[\phi_H(x_1)\phi_H(x_2)\ldots\phi_H(x_n)\right]|0\rangle\,. \end{aligned} \tag{4.223}$$

Let us now recall that the Heisenberg field $\phi_H(\boldsymbol{x}, t)$ is related to the in field $\phi_{in}(\boldsymbol{x}, t)$ by [cf. Eq. (4.177)]

$$\phi_H(\boldsymbol{x}, t) = \Lambda^{-1}(t, t_0)\phi_{in}(\boldsymbol{x}, t)\Lambda(t, t_0)\,, \tag{4.224}$$

where $t_0 \to -\infty$.

However, at the moment we only know how to compute

$$\begin{aligned}\tau_0(x_1, x_2, \ldots, x_n) &\equiv \langle x_1 \ldots x_n\rangle_0 \\ &\equiv {}_{in}\langle 0|T\left[\phi_{in}(x_1)\phi_{in}(x_2)\ldots\phi_{in}(x_n)\right]|0\rangle_{\rm in}\,, \end{aligned} \tag{4.225}$$

where $|0\rangle_{\rm in}$ is the ground state of the free Hamiltonian H_0.

Now, from Eq. (4.163) we can recall that $\Lambda(t, t_1)$ satisfies the composition law

$$\Lambda(t_1, t_2) = \Lambda(t_1, t_3)\Lambda(t_3, t_2) = \Lambda(t_1, t_3)\Lambda^{-1}(t_2, t_3)\,. \tag{4.226}$$

So, if we take points x_j, where $j = 1, \ldots, n$ satisfying $x_1^0 > x_2^0 > x_3^0 > \cdots > x_n^0$ (i.e. they are time ordered), then

$$\begin{aligned}&\phi_H(x_1)\phi_H(x_2)\ldots\phi_H(x_n)\\ &= \Lambda(t_1, t_0)^{-1}\phi_{\rm in}(x_1)\Lambda(t_1, t_0)\Lambda(t_2, t_0)^{-1}\phi_{\rm in}(x_2)\ldots\phi_{\rm in}(x_n)\Lambda(t_n, t_0)\\ &= \Lambda(t, t_0)^{-1}\Lambda(t, t_1)\phi_{\rm in}(x_1)\Lambda(t_1, t_2)\phi_{\rm in}(x_2)\ldots\\ &\times\ldots\Lambda(t_{n-1}, t_n)\phi_{\rm in}(x_n)\Lambda(t_n, -t)\Lambda(-t, t_0)\\ &= \Lambda(t)^{-1}T\left[\phi_{\rm in}(x_1)\phi_{\rm in}(x_2)\ldots\phi_{\rm in}(x_n)\exp\left(i\int_{-t}^{t}\overline{\mathcal{L}}_I(x)d^4x\right)\right]\Lambda(-t),\end{aligned} \tag{4.227}$$

where $\Lambda(\pm t) = \Lambda(\pm t, -\infty)$ and $t > x_1^0 > x_2^0 > \cdots > x_n^0 > -t$. We have also used that

$$\Lambda(t_1, t_2) = T\left[\exp\left(i\int_{t_2}^{t_1}\overline{\mathcal{L}}_I(x)d^4x\right)\right], \tag{4.228}$$

involves $\phi_{\rm in}(\boldsymbol{x}, t)$ for times $t \in [t_2, t_1]$. Times t and $-t$ are taken to be times where interaction switches off. So, we adiabatically evolve the non-interacting vacuum state into the true $|0\rangle$ by taking $H = H_0 + \eta(\tau)V$ with $\eta(\tau) = 0$ at $\tau = \pm\infty$ and $\eta = 1$ at $\tau \in [-t, t]$.

Now we want to take vacuum expectation value of the time ordered product (4.227). By denoting the is the ground state of the full Hamiltonian H as $|\Omega\rangle$ (this is more conventional notation that $|0\rangle$) we have

$$\begin{aligned} &\langle\Omega| T\left[\phi_H(x_1)\phi_H(x_2)\ldots\phi_H(x_n)\right]|\Omega\rangle \\ &= \lim_{t_0\to-\infty} \langle\Omega|\,\Lambda(t,t_0)^{-1} T\left[\phi_{in}(x_1), \phi_{in}(x_2)\ldots\phi_{in}(x_n)\right. \\ &\qquad \times \left.\exp\left(i\int_{-t}^{t} \overline{\mathcal{L}}_I(x) d^4x\right)\right] \Lambda(-t,t_0)\,|\Omega\rangle\ . \quad (4.229) \end{aligned}$$

In order to bring (4.229) to a manageable form we need to convert the ground state $|\Omega\rangle$ to the ground state $|0\rangle_{in}$. How these two ground states are connected? We have already seen in Eq. (4.169) that

$$\begin{aligned} |\psi_H\rangle &= \lambda_-(t,t_0)\Lambda^{-1}(t,t_0)\,|\psi_I(t)\rangle\ , \\ \Rightarrow \quad |\psi_H\rangle &= \lambda_-(t,-\infty)\Lambda^{-1}(t,-\infty)\,|\psi(t)\rangle_{in}\ , \end{aligned} \quad (4.230)$$

with $|\lambda_-| = 1$. This implies, in particular, that

$$|\Omega\rangle = \lambda_-(t,-\infty)\Lambda^{-1}(t,-\infty)\,|0\rangle_{in}\,. \quad (4.231)$$

Let us recall that $|\psi(t)\rangle_{in}$ states including the vacuum state $|0\rangle_{in}$ evolve w.r.t free Hamiltonian and $H_0\,|0\rangle_{in} = 0$.

Note that since both $|\Omega\rangle$ and $|0\rangle_{in}$ are time independent, the term $\lambda_-(t,-\infty)\Lambda^{-1}(t,-\infty)$ must also be t independent.

Let us be more specific here and find this relation more explicitly. Take

$$e^{-iHt}\,|0\rangle_{in} = e^{-iE_0t}\,|\Omega\rangle\,\langle\Omega|0\rangle_{in} + \sum_{n\neq0} e^{-iE_nt}|n\rangle\langle n\,|0\rangle_{in}\ . \quad (4.232)$$

States $|n\rangle$ are energy eigenstates of the full Hamiltonian H. We will further assume that the overlap $\langle\Omega|0\rangle_{in} \neq 0$. This is justified in the sense that we would like to use perturbation theory and hence $|0\rangle_{in}$ should not be "too far" from $|\Omega\rangle$. Also, we know that $E_0 = \langle\Omega|\,H\,|\Omega\rangle$. Since $E_n > E_0$ for $\forall n \neq 0$, we can get rid of all $n \neq 0$ terms by sending $t \to \infty(1 - i\varepsilon)$, where $0 < \varepsilon \ll 1$. Then, the exponential factor e^{-iE_nt} dies slowest for $n = 0$. From (4.232) follows that

$$\left(e^{-iE_0t}\langle\Omega\,|0\rangle_{in}\right)^{-1} e^{-iHt}\,|0\rangle_{in} = |\Omega\rangle + \sum_{n\neq0} e^{-i(E_n-E_0)t}\frac{\langle n|0\rangle_{in}}{\langle\Omega|0\rangle_{in}}\,. \quad (4.233)$$

From this we can directly read

$$|\Omega\rangle = \lim_{t\to\infty(1-i\varepsilon)}\left[\left(e^{-iE_0t}\,\langle\Omega|0\rangle_{in}\right)^{-1} e^{-iHt}\,|0\rangle_{in}\right]. \quad (4.234)$$

Now, since t is very large we can shift it by a constant t_0

$$|\Omega\rangle = \lim_{t\to\infty(1-\varepsilon)} \left[\left(e^{-iE_0(t+t_0)}\,\langle\Omega|0\rangle_{in}\right)^{-1} e^{-iH(t_0-(-t))}\,\underbrace{e^{iH_0(t_0-(-t))}\,|0\rangle_{in}}_{\mathbf{1}|0\rangle_{in}} \right]$$

$$= \lim_{t\to\infty(1-i\varepsilon)} \left[\underbrace{\left(e^{-iE_0(t_0-(-t))}\,\langle\Omega|0\rangle_{in}\right)^{-1}}_{\lambda_-(-t,t_0)} \Lambda^{-1}(-t,t_0)\,|0\rangle_{in} \right]. \quad (4.235)$$

Apart from the c-number factor in front, this expression tells us that we can get $|\Omega\rangle$ by simply evolving $|0\rangle_{in}$ from time $-t$ to time t_0 with the operator Λ. In similar way we can express $\langle\Omega|$ as

$$\langle\Omega| = \lim_{t\to\infty(1-i\varepsilon)} {}_{in}\langle 0| \left[\Lambda(t,t_0)\, \underbrace{\left(e^{-iE_0(t-t_0)}\,{}_{in}\langle 0|\Omega\rangle\right)^{-1}}_{\lambda_+(t,t_0)} \right]. \quad (4.236)$$

So the n-point full Green function has the form

$$\begin{aligned}
&\langle\Omega|\,T[\phi_H(x_1)\phi_H(x_2)\ldots\phi_H(x_n)]\,|\Omega\rangle \\
&\quad = \lim_{t_0\to-\infty} \langle\Omega|\,\Lambda^{-1}(t,t_0) T\left[\phi_{in}(x_1)\ldots\phi_{in}(x_n)\right. \\
&\qquad\qquad \times \left.\exp\left(i\int_{-t}^{t} \overline{\mathcal{L}}_I(x) d^4x\right)\right] \Lambda(-t,t_0)\,|\Omega\rangle \\
&\quad = \lim_{t_0\to-\infty}\,\lim_{t\to\infty(1-i\varepsilon)} \lambda_+(t,t_0)\lambda_-(-t,t_0)\,{}_{in}\langle 0|\,T\left[\phi_{in}(x_1)\ldots\phi_{in}(x_n)\right. \\
&\qquad\qquad \times \left.\exp\left(i\int_{-t}^{t} \overline{\mathcal{L}}_I(x) d^4x\right)\right] |0\rangle_{in}\,. \quad (4.237)
\end{aligned}$$

In addition [cf. Eq. (4.172)]

$$\begin{aligned}
1 = \langle\Omega\,|\Omega\rangle &= \lim_{t\to\infty(1-i\varepsilon)} \lambda_+(t,t_0)\lambda_-(-t,t_0)\,{}_{in}\langle 0|\,\Lambda(t,t_0)\Lambda^{-1}(-t,t_0)\,|0\rangle_{in} \\
&= \lim_{t\to\infty(1-i\varepsilon)} \lambda_+(t,t_0)\lambda_-(-t,t_0)\,{}_{in}\langle 0|\,\Lambda(t,-t)\,|0\rangle_{in}\,. \quad (4.238)
\end{aligned}$$

For large (but finite) t we can thus write

$$\lambda_+(t,t_0)\lambda_-(-t,t_0) \simeq \frac{1}{{}_{in}\langle 0|\,\Lambda(t,-t)\,|0\rangle_{in}}\,. \quad (4.239)$$

With this we finally obtain

$$\begin{aligned}
&\langle\Omega|\,T[\phi_H(x_1)\ldots\phi_H(x_n)]\,|\Omega\rangle \\
&= \lim_{t\to\infty(1-i\varepsilon)} \frac{{}_{in}\langle 0|\,T\left[\phi_{in}(x_1)\ldots\phi_{in}(x_n)\exp\left(i\int_{-t}^{t}\overline{\mathcal{L}}_I(x)d^4x\right)\right]|0\rangle_{in}}{{}_{in}\langle 0|\,T\left[\exp\left(i\int_{-t}^{t}\overline{\mathcal{L}}_I(x)d^4x\right)\right]|0\rangle_{in}} \\
&\equiv \langle x_1 x_2\ldots x_n\rangle\,. \quad (4.240)
\end{aligned}$$

This is the so-called *Gell-Mann–Low formula* for the full n-point Green function. So far this expression is exact, but it is ideally suited for perturbative calculations, since we work with free fields and hence we can use a full power of Wick's theorem which (as we already know) boils down to products of $i\Delta_F$.

4.8 Functional Integral Approach

Gell-Mann–Low formula provides a useful starting point for introducing functional integral. There are basically two distinct ways how to arrive at functional integrals:

1. Formulate the so-called path integrals in QM - these represent Green's function for Schrödinger equation and at the same time correspond to transition amplitude $\langle x', t'|x, t\rangle$. One then formally passes to field theory in much the same way as we did when passing from QM to QFT. In this formulation it can be shown that for Klein–Gordon particle $\langle x'^{\mu}, \tau'|x^{\mu}, \tau\rangle \propto \Delta_F(x'-x)$ and similarly also for Dirac's particle (τ represents a proper time that parametrizes particle's wordline).

2. One can use the relation for generating function (4.212), i.e.

$$\begin{aligned}&\langle 0|\, T\left[\exp\left(-i\int d^4x\, J(x)\phi(x)\right)\right]|0\rangle \\ &\quad = \exp\left[-\frac{1}{2}\int \mathrm{d}^4x\mathrm{d}^4y\, J(x)J(y)\,\langle 0|\, T[\phi(x)\phi(y)\,|0\rangle\right], \quad (4.241)\end{aligned}$$

   which encapsulates Wick's theorem.

In this lecture, we will use the second approach because it brings us to functional integrals faster.

Generating Functional for Full Green's Functions

Consider the full n-point Green's function

$$\langle\Omega|\, T[\phi_H(x_1)\ldots\phi_H(x_n)]\,|\Omega\rangle \;\equiv\; \langle x_1x_2\ldots x_n\rangle\,. \quad (4.242)$$

Due to the permutation symmetry of $\langle x_1x_2\ldots x_n\rangle$ one can conveniently combine the entire hierarchy $\{\langle x_1x_2\ldots x_n\rangle, n\in\mathbb{N}\}$ into one generating functional

$$\begin{aligned}Z[J] &= Z[0]\sum_{n=0}^{\infty}\frac{i^n}{n!}\int_{\mathbb{R}^{4n}}\prod_{i=1}^{n}d^4x_i\, J(x_1)\ldots J(x_n)\langle x_1x_2\ldots x_n\rangle \\ &= Z[0]\,\langle\Omega|\, T\left[\exp\left(i\int d^4x\, J(x)\phi_H(x)\right)\right]|\Omega\rangle\,. \quad (4.243)\end{aligned}$$

Here $Z[0]$ is the normalization constant to be fixed shortly. The c-numbered function $J(x)$ is the so-called *Schwinger source term*. With the help of Gell-Mann—Low formula this can be rewritten in terms of free field as (for simplicity we omit limits and we set $|0\rangle \equiv |0\rangle_{in}$ and $\phi(x) \equiv \phi_{in}(x)$)

$$\frac{Z[J]}{Z[0]} = \frac{\langle 0| T\left[e^{i\int d^4x \overline{\mathcal{L}}_I(\phi)+J(x)\phi(x)}\right]|0\rangle}{\langle 0| T\left[e^{i\int d^4x \overline{\mathcal{L}}_I(\phi)}\right]|0\rangle} . \tag{4.244}$$

At this point we set $Z[0] = \langle 0| T\left[e^{i\int d^4x \overline{\mathcal{L}}_I(\phi)}\right]|0\rangle$ so that

$$Z[J] = \langle 0| T\left[e^{i\int d^4x \overline{\mathcal{L}}_I(\phi)+J(x)\phi(x)}\right]|0\rangle . \tag{4.245}$$

It follows from the very definition of $Z[J]$ that $\langle x_1x_2\ldots x_n\rangle$ can be obtained when we n times functionally differentiate $Z[J]$ with respect to $J(x)$, in particular

$$\langle x_1x_2\ldots x_n\rangle = \frac{1}{Z[0]}\left.\frac{(-i)^n\delta^n Z[J]}{\delta J(x_1)\ldots\delta J(x_n)}\right|_{J=0} . \tag{4.246}$$

Generating Functional Note

Generating functional $Z[J]$ is an analogue of the moment generating function (or characteristic function) used in mathematical statistics.

Now, $Z[J]$ in the form given above can be formally rewritten as

$$Z[J] = \exp\left[i\int_{\mathbb{R}^4} d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right]\langle 0| T\left[e^{i\int d^4x\, J(x)\phi(x)}\right]|0\rangle . \tag{4.247}$$

Now, the "overbar" from $\mathcal{L}_I$ was removed since we do not need to emphasize anymore that it is an operator in the interaction representation. This is an analog of the formula

$$f\left(-i\frac{d}{dx}\right)e^{ixp} = f(p)e^{ixp} , \tag{4.248}$$

used, e.g. in Fourier transform. By employing the generating functional for Wick theorem and taking $J \to -J$ in (4.212), we obtain

$$\begin{aligned} Z[J] &= \exp\left\{i\int_{\mathbb{R}^{4n}} d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right\} \\ &\times \exp\left\{-\frac{i}{2}\int d^4y_1 d^4y_2\, J(y_1)J(y_2)\Delta_F(y_1-y_2)\right\} . \end{aligned} \tag{4.249}$$

The Functional Integral and Its Measure

In order to establish contact with functional integrals, let us consider the Fresnel integrals ($a \in \mathbb{R}$)

$$\int_{\mathbb{R}} \frac{dx}{\sqrt{2\pi}} \exp\left(i\frac{a}{2}x^2\right) = \begin{cases} \frac{1}{\sqrt{|a|}} e^{\frac{i\pi}{4}} & a > 0 \\ \frac{1}{\sqrt{|a|}} e^{-\frac{i\pi}{4}} & a < 0 \,. \end{cases} \tag{4.250}$$

Proof of this identity is as follows. We first extend x into $\mathbb{C}$ and evaluate the integral $\int_\gamma e^{iaz^2}\, dz$ for $a > 0$, where the contour γ is depicted on Fig. 4.7.

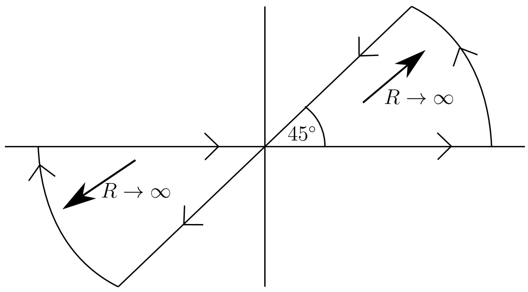


Figure 4.7: Contour γ used in the evaluation of the Fresnel integral (4.250) for $a > 0$.)

Since e^{iaz^2} is an analytic function, it follows from the Cauchy integral theorem that

$$\begin{aligned} 0 = \int_\gamma e^{iaz^2} dz &= \int_{\rightarrow} e^{iaz^2} dz + \int_{\curvearrowleft} e^{iaz^2} dz \\ &+ \int_{\swarrow} e^{iaz^2} dz + \int_{\curvearrowright} e^{iaz^2} dz \,. \end{aligned} \tag{4.251}$$

First notice that

$$\begin{aligned} \int_{\curvearrowleft} e^{iaz^2} dz &= \lim_{R\to+\infty} R \int_0^{\pi/4} d\phi\, i e^{i\phi} e^{aR^2(i\cos 2\phi - \sin 2\phi)} \\ \Rightarrow \quad \left|\int_{\curvearrowleft} e^{iaz^2} dz\right| &\leq \lim_{R\to+\infty} R \int_0^{\pi/4} d\phi\, e^{-aR^2 \sin(2\phi)} = 0 \\ \Rightarrow \quad \int_{\curvearrowleft} e^{iaz^2} dz &= 0 \,, \end{aligned} \tag{4.252}$$

and similarly for $\int_{\curvearrowright} e^{iaz^2} dz$. The integral $\int_{\swarrow} e^{iaz^2} dz$ can be evaluated as follows (consider $a > 0$ first):

$$\begin{aligned} \int_{\swarrow} e^{iaz^2} dz &= \left\{ z = e^{i\pi/4} z',\ dz = dz' e^{i\pi/4} \right\} \\ &= e^{i\pi/4} \int_\infty^{-\infty} dz' e^{-az'^2} = \{a > 0\} = -e^{i\pi/4} \left(\frac{\pi}{a}\right)^{1/2} \\ \Rightarrow \quad - \int_{\swarrow} dz e^{-az^2} &= e^{i\pi/4} \left(\frac{\pi}{a}\right)^{1/2} \,. \end{aligned} \tag{4.253}$$

So, by plugging these results to (4.251) we obtain

$$\int_{-\infty}^{+\infty} e^{iax^2} dx = \{a > 0\} = e^{i\pi/4}\left(\frac{\pi}{a}\right)^{1/2}$$

$$\Rightarrow \quad \int_{-\infty}^{+\infty} \frac{dx}{\sqrt{2\pi}} e^{i\frac{a}{2}x^2} = \{a > 0\} = e^{i\pi/4}\frac{1}{\sqrt{a}}. \qquad (4.254)$$

For the case $a < 0$ we would need to chose a different contour, namely the one where the diagonal line would nor run under the angle $\pi/4$ but $-\pi/4$.

So, from the Fresnel integral (4.250) we have the following N-dimensional generalization

$$\int_{-\infty}^{\infty} \prod_{i=1}^{N} dc_i \exp\left(\frac{i}{2}\sum_{n,m} c_n \mathbb{A}_{nm} c_m\right) = \prod_{i=1}^{N}\sqrt{\frac{2\pi}{|\lambda_i|}} e^{i\pi \mathrm{sign}(\lambda_i)/4}$$

$$= \left|\det\left(\frac{\mathbb{A}}{2\pi}\right)\right|^{-1/2} e^{i\eta\pi/4}. \qquad (4.255)$$

Here $\mathbb{A}$ is real, symmetric (hence diagonalizable) $N \times N$ matrix with eigenvalues $\{\lambda_i, i = 1, ..., N\}$.

Note

Formula (4.255) has sense only if $\mathbb{A}$ has no zero modes. Case with zero modes must be treated independently and it is related to the concept of the so-called *collective coordinates*.

The index $\eta = \sum_i^N \mathrm{sign}(\lambda_i)$ is referred to as the *Morse* or *Maslov* index. The later is mostly important in the context of transition amplitude in QM. For typical applications in QFT (as, e.g. computation of Green functions or S-matrix elements) it is not important as we will see.

In order to establish the connection to fields, let us first observe that any real function $\phi(x)$ can be expanded in term of some real orthonormal basis $\{v_n(x), n \in N\}$, $\phi(x) = \sum_n c_n v_n(x)$, with c_n's being real expansion coefficients.

So, in particular, we can write

$$\int d^4x d^4y\, \phi(x) A(x, y) \phi(y) = \sum_{n,m} c_n \mathbb{A}_{n,m} c_m\,, \qquad (4.256)$$

$A(x, y)$ is some symmetric function or operator in x and y.

with

$$\mathbb{A}_{n,m} = \int d^4x d^4y\, v_n(x) A(x, y) v_m(y) \qquad (4.257)$$

This is a form of similarity transformation where v_x can be viewed as a matrix with discrete and symmetric index.

Since both $A(x, y)$ and $A_{n,m}$ are symmetries, they are diagonalizable, i.e. there exist polar bases $\{u_n(x); n \in N, x \in \mathbb{R}\}$ and $\{u_m^{(n)}; n, m \in N\}$

such that

$$\int d^4y A(x,y)u_n(y) = \lambda_n u_n(x), \tag{4.258}$$

$$\sum_k \mathbb{A}_{mk} u_k^{(n)} = \lambda_n u_m^{(n)}, \tag{4.259}$$

where $u_n(x)$ and $u_k^{(n)}$ are related as

$$u_k^{(n)} = \int d^4x\, u_n(x) v_k(x), \tag{4.260}$$

$$u_n(x) = \sum_k u_k^{(n)} v_k(x). \tag{4.261}$$

Relations (4.258)-(4.261) are simple consequences of the *orthonormality condition*

$$\int d^4x\, v_n(x) v_m(x) = \delta_{nm}, \tag{4.262}$$

and the *completeness relation*

$$\sum_n v_n(x) v_n(y) = \delta(x-y). \tag{4.263}$$

Discretize now points in the spacetime, so that the spacetime is spanned by N points x_i — so-called *Minkowski lattice*. Then any $\{\phi(x_i); i \in N\}$ can be expanded into N base functions $v_n(x_i)$ only. In fact

$$\phi(x) = \sum_{n=1}^{+\infty} c_n v_n(x) \;\Rightarrow\; \phi(x_i) = \sum_{n=1}^{N} c_n v_n(x_i). \tag{4.264}$$

The last equation provides a system of N independent equations for N unknown c_n. Consequently $\{\phi(x_i); i \in \mathbb{N}\}$ is uniquely determined by its expansion modes c_n and vice versa. With this we can formulate the integral measure as

$$\mathcal{D}\phi = \lim_{N\to\infty} \prod_{i=1}^{N} d\phi(x_i) = \lim_{N\to\infty} \prod_{n=1}^{N} dc_n |J^{(N)}|, \tag{4.265}$$

with the Jacobian

$$J^{(N)} = \det \begin{Vmatrix} v_1(x_1) & v_1(x_2) & \cdots & \cdots \\ v_2(x_1) & v_2(x_2) & & \\ \vdots & & \ddots & \\ \vdots & & & v_n(x_n) \end{Vmatrix}. \tag{4.266}$$

The identity in Eq. (4.265) should be understood in the weak sense, namely that the limit $N \to \infty$ stands in front of the corresponding multiple integral.

Note that, due to the orthonormality of the base system, we have that $J^{(N)} \to 1$ in the large N limit (also known as *continuity limit* or *long wave limit*).

Truncation of the base system elements changes the infinite dimensional matrix to $N \times N$ matrix $\mathbb{A}^{(N)}$.

Recalling identity (4.255), we might define the functional integral over ϕ as

$$\int \mathcal{D}\phi \exp\left(\frac{i}{2}\int d^4x d^4y\, \phi(x) A(x,y)\phi(y)\right)$$
$$= \lim_{N\to+\infty}\left[\int \prod_{i=1}^{N} dc_i \exp\left(\frac{i}{2}\sum_{n,m} c_n \mathbb{A}^{(N)}_{n,m} c_m\right)\right] |J^{(N)}|$$
$$= \lim_{N\to+\infty}\left|\det\left(\frac{\mathbb{A}^{(N)}}{2\pi}\right)\right|^{-1/2} J^{(N)} \;=\; N'|\det(A(x,y))|^{-1/2}\,, \qquad (4.267)$$

Both $\mathbb{A}$ and A have identical spectrum.

where on the last line we have neglected the Maslov index. N' is an infinite constant.

At this point we might note the following identity

$$N'|\det(\Delta_F(x,y)|^{1/2} \exp\left(-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x,y)J(y)\right)$$
$$= \int \mathcal{D}\phi \exp\left(\frac{i}{2}\int d^4x d^4y \phi(x)\left[\Delta_F(x,y)\right]^{-1}\phi(y)\right)$$
$$\times \exp\left(-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x,y)J(y)\right). \qquad (4.268)$$

Here $\Delta_F(x,y) = -i\langle 0|T[\phi(x)\phi(y)]|0\rangle$. At this point we use translational invariance of $\mathcal{D}\phi$, i.e.

$$\mathcal{D}\phi = \mathcal{D}(\phi + g) \sim \prod_i d\left(\phi(x_i) + g(x_i)\right)\,, \qquad (4.269)$$

note that $g(x)$ is an arbitrary but fixed function (hence $g(x_i)$ is a constant while $\phi(x_i)$ changes). This implies that (4.8) can be further written as

$$\int \mathcal{D}\phi \exp\left(\frac{i}{2}\int d^4x d^4y \left[\phi_x + (J\Delta_F)_x\right]\Delta_F^{-1}(x,y)\left[\phi_y + (\Delta_F J)_y\right]\right)$$
$$\times \exp\left(-\frac{i}{2}\int d^4x d^4y J(x)\Delta_F(x,y)J(y)\right)$$
$$= \int \mathcal{D}\phi \exp\left(\frac{i}{2}\int d^4x d^4y \left[\phi(x)\Delta_F^{-1}(x,y)\phi(y)\right]\right)$$
$$\times \exp\left(i\int d^4x\, J(x)\phi(x)\right). \qquad (4.270)$$

What is $\Delta_F^{-1}(x,y)$? We know that it is defined so that

$$\int d^4z \Delta_F^{-1}(x,z)\Delta_F(z,y) \;=\; \delta(x-y)\,.$$

Since $(\Box + m^2)_x \Delta_F(x,y) = -\delta(x-y)$ implies that

$$\int d^4 z \,\delta(z-x)(\Box + m^2)_x \Delta_F(z,y) = -\delta(x-y)\,,$$

we see that $\Delta_F^{-1}(x,z) = -\delta(z-x)(\Box + m^2)_x$. With this we can further rewrite (4.270) as

$$\int \mathcal{D}\phi \exp\left(\frac{i}{2}\int d^4x\,\phi(x)\left[-\left(\Box + m^2\right)\right]\phi(x) + i\int d^4x\,J(x)\phi(x)\right)$$
$$= \int \mathcal{D}\phi \exp\left(iS_0[\phi] + i\int d^4x\,J(x)\phi(x)\right).$$

Here S_0 is the action for a free scalar field. Let us put now everything together and rewrite the generator of Green functions in the following way:

$$\begin{aligned}
Z[J] &= \exp\left[i\int_{R^4} d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right]\langle 0|\,T\left[e^{i\int_{R^4} d^4x\,J(x)\phi(x)}\right]|0\rangle \\
&= \exp\left[i\int_{R^4} d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right] e^{-\frac{1}{2}\int d^4x d^4y J_x J_y \langle 0|T[\phi(x)\phi(y)]|0\rangle} \\
&= \exp\left[i\int_{R^4} d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right] \\
&\quad\times \frac{\int \mathcal{D}\phi \exp\left(iS_0[\phi] + i\int d^4x\,J(x)\phi(x)\right)}{N'\,|\mathrm{det}\Delta_F|^{1/2}} \\
&= \exp\left[i\int_{R^4} d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right] \\
&\quad\times \frac{\int \mathcal{D}\phi \exp\left(iS_0[\phi] + i\int d^4x\,J(x)\phi(x)\right)}{\int \mathcal{D}\phi \exp\left(iS_0[\phi]\right)} \\
&= \frac{\int \mathcal{D}\phi \exp\left(iS[\phi] + i\int d^4x J(x)\phi(x)\right)}{\int \mathcal{D}\phi \exp\left(iS_0[\phi]\right)}\,,
\end{aligned} \tag{4.271}$$

where $S[\phi] = S_0[\phi] + \int d^4x\mathcal{L}_I(\phi)$ is the *full action* of an interacting scalar field theory. The corresponding full n-point Green function is [cf. Gell-Mann-Low formula]

$$\begin{aligned}
\langle x_1 \ldots x_n\rangle &= \frac{\langle 0|\,T\left[\phi(x_1)\ldots\phi(x_n)e^{i\int d^4x\,\overline{\mathcal{L}}_I(\phi)}\right]|0\rangle}{\langle 0|\,T\left[e^{i\int d^4x\,\overline{\mathcal{L}}_I(\Phi)}\right]|0\rangle} \\
&= \frac{1}{Z[0]}\frac{(-i)^n\delta^n}{\delta J(x_1)\ldots\delta J(x_n)}Z[J]\bigg|_{J=0} \\
&= \frac{1}{Z[0]}\frac{\int \mathcal{D}\phi\;\phi(x_1)\ldots\phi(x_n)\exp\left(iS[\phi] + i\int d^4x\,J(x)\phi(x)\right)}{\int \mathcal{D}\phi \exp\left(iS_0[\phi]\right)}\Bigg|_{J=0}\,.
\end{aligned} \tag{4.272}$$

In particular, for $n = 0$ we get

$$\langle \Omega | \Omega \rangle = \langle 1 \rangle = 1 = \frac{1}{Z[0]} \frac{\int \mathcal{D}\phi \; e^{iS[\Phi]}}{\int \mathcal{D}\phi \; e^{iS_0[\phi]}} , \tag{4.273}$$

which implies that

$$\langle x_1, \ldots, x_n \rangle = \frac{\int \mathcal{D}\phi \; \phi(x_1) \ldots \phi(x_n) e^{iS[\phi]}}{\int \mathcal{D}\phi \; e^{iS[\phi]}} . \tag{4.274}$$

This is the so-called *functional-integral representation* of the n-point Green function.

So far we have considered only real scalar fields. Extension to complex scalar fields (charged scalar particles) is obtained by means of an analog of Fresnel integral, namely

$$\begin{aligned} \int \frac{dz^* dz}{2\pi i} e^{i\frac{a}{2}|z|^2} &= \int \frac{dxdy}{\pi} e^{i\frac{a}{2}x^2 + i\frac{a}{2}y^2} \\ &= \sqrt{\frac{2}{|a|}} \sqrt{\frac{2}{|a|}} e^{i\frac{\pi}{4}\mathrm{sign}(a)} e^{i\frac{\pi}{4}\mathrm{sign}(a)} \\ &= \frac{2}{|a|} e^{i\frac{\pi}{2}\mathrm{sign}(a)} , \end{aligned} \tag{4.275}$$

where we have employed the complex measure

$$dz^* \wedge dz = (dx - idy) \wedge (dx + idy) = 2idx \wedge dy , \tag{4.276}$$

(we use the notation of differential forms). This can also be alternatively obtained from the usual (real analysis) change of variables

$$dz^* dz = \left| \frac{\partial(z^*, z)}{\partial(x, y)} \right| dxdy = 2idxdy , \tag{4.277}$$

but in the complex calculus the absolute value refers only to the sign $\pm$, not the complex i factor.

By neglecting Morse index we have (set $a/2 \to a$)

$$\int \frac{1}{2\pi i} dz^* dz \; e^{ia|z|^2} = \frac{1}{|a|} . \tag{4.278}$$

More generally

$$\begin{aligned} \int \frac{dz^* dz}{2\pi i} \; e^{ia|z|^2 + ib^* z + ibz^*} &= \int \frac{dz^* dz}{2\pi i} \; e^{ia(z + b/a)(z^* + b^*/a) - i|b|^2/a} \\ &= \frac{1}{|a|} \exp\left(-i \frac{|b|^2}{a} \right) . \end{aligned} \tag{4.279}$$

From these Fresnel integrals we obtain

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\prod_{i=1}^{N}\left[\frac{dz_i^* dz_i}{2\pi i}\right]\exp\left[iz_i^*\mathbb{A}_{ij}z_j + ib_i^* z_i + ib_i z_i^*\right] \\
&= \int_{-\infty}^{\infty}\prod_{i=1}^{N}\left[\frac{dz_i^* dz_i}{2\pi i}\right]\exp\left[i(z^* + b^*\mathbb{A}^{-1})_i\mathbb{A}_{ij}(z + \mathbb{A}^{-1}b)_j - ib_i^*(\mathbb{A}^{-1})_{ij}b_j\right] \\
&= e^{-ib_i^*(\mathbb{A}^{-1})_{ij}b_j}\int_{-\infty}^{\infty}\prod_{i=1}^{N}\left[\frac{dz_i^* dz_i}{2\pi i}\right]\exp\left[iz_i^*\mathbb{A}_{ij}z\right] . \qquad (4.280)
\end{aligned}
$$

Since $\mathbb{A}_{ij}$ is Hermitian, there exists an unitary similarity transformation that diagonalizes $\mathbb{A}$, so that we can write [cf. Eq. (4.279)]

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\prod_{i=1}^{N}\left[\frac{1}{2\pi i}dz_i^* dz_i\right]\exp\left[iz_i^*\mathbb{A}_{ij}z_j\right] \\
&= \underbrace{\int_{-\infty}^{\infty}\prod_{i=1}^{N}\left[\frac{1}{2\pi i}dc_i^* dc_i\right]\underbrace{J}_{=1}\exp\left(i\lambda_i|c_i|^2\right)}_{(\det\mathbb{A})^{-1}} . \qquad (4.281)
\end{aligned}
$$

Here we have used the fact that the Jacobian of any unitary matrix is 1. Thus we obtain that Eq. (4.280) is equal to

$$
(\det\mathbb{A})^{-1}e^{-ib^*\mathbb{A}^{-1}b} . \qquad (4.282)
$$

As an exercise, following the same route as for Hermitian scalar fields, show that

Here N contains all constant factors, Fresnel measure and determinant.

$$
\begin{aligned}
&N\int\mathcal{D}\phi\mathcal{D}\phi^*\exp\left[iS_0\left[\phi,\phi^*\right] + i\int d^4x\phi(x)J^*(x) + i\int d^4x\phi^*(x)J(x)\right] \\
&= \left[\det\left(\Box + m^2\right)\right]^{-1}\exp\left[-\int d^4x d^4y J^*(x)G(x,y)J(y)\right] , \qquad (4.283)
\end{aligned}
$$

where $G(x,y) = \langle 0|T\left[\hat{\phi}(x),\hat{\phi}^\dagger(y)\right]|0\rangle = i\Delta_F(x,y)$, and thus finally

$$
\frac{Z\left[J,J^*\right]}{Z\left[0\right]} = \frac{\int\mathcal{D}\phi\mathcal{D}\phi^* e^{iS[\phi,\phi^*]+i\int d^4x\phi J^* + i\int d^4x\phi^* J}}{\int\mathcal{D}\phi\mathcal{D}\phi^* e^{iS[\phi,\phi^*]}} , \qquad (4.284)
$$

which implies

$$
\langle x_1,\ldots,x_n\rangle = \frac{\int\mathcal{D}\phi\mathcal{D}\phi^*\phi(x_1)\ldots\phi(x_n)\, e^{iS[\phi,\phi^*]}}{\int\mathcal{D}\phi\mathcal{D}\phi^* e^{iS[\phi,\phi^*]}} . \qquad (4.285)
$$

Similar identity holds for the correlator of ϕ^* fields or mixed correlator of ϕ and ϕ^* fields.

Note — The Feynman-Matthews-Salam formula

Previous relations can be generalized to any functional or function of fields, e.g.

$$\langle\Omega| T\left[F\left[\hat{\phi}_H\right]\right]|\Omega\rangle = N\int \mathcal{D}\phi F[\phi]e^{iS[\phi]}, \tag{4.286}$$

and similarly for

$$\langle\Omega| T\left[G\left[\hat{\phi}^*_H, \hat{\phi}_H\right]\right]|\Omega\rangle = N\int \mathcal{D}\phi\mathcal{D}\phi^* G[\phi^*,\phi]e^{iS[\phi^*,\phi]}. \tag{4.287}$$

4.9 Perturbative calculus

As a toy model we will discuss the case with

$$\mathcal{L}_I = -\frac{\lambda}{4!}\phi^4. \tag{4.288}$$

So, we consider a single real scalar field at this stage.

We have seen that in order to compute $\langle x_1,\dots,x_n\rangle$ we need to know the normalized generating functional $Z[J]/Z[0]$. Indeed

$$\langle x_1,\dots,x_n\rangle = \frac{1}{Z[0]}\frac{(-i)^n\delta^n Z[J]}{\delta J(x_1)\dots\delta J(x_n)}\bigg|_{J=0}. \tag{4.289}$$

Let us call the normalized generating functional as $\tilde{Z}[J]$, then

$$\tilde{Z}[j] = \frac{\exp\left[i\int d^4z\mathcal{L}_I\left(-i\frac{\delta}{\delta J(z)}\right)\right]\exp\left[-\frac{i}{2}\int d^4xd^4yJ(x)J(y)\Delta_F(x,y)\right]}{(\text{ ditto })|_{J=0}}.$$

The only way how to treat $\exp\left(i\int d^4x\mathcal{L}_I\right)$ is via power series expansion in the coupling constant λ, i.e. via *perturbation theory*. In particular, for the numerator we can write

$$\left[1-i\frac{\lambda}{4!}\int\left(-i\frac{\delta}{\delta J(z)}\right)^4 d^4z + O(\lambda^2)\right]\exp\left[-\frac{i}{2}\int d^4xd^4y\,J(x)\Delta_F(x,y)J(y)\right].$$

To order λ^0, we have just the free-particle generating functional $Z_0[J]$. To order λ, we proceed as follows. We compute first the single functional derivative

$$\begin{aligned}&(-i)\frac{\delta}{\delta J(z)}\exp\left[-\frac{i}{2}\int d^4xd^4y\,J(x)\Delta_F(x-y)J(y)\right]\\ &= -\int d^4x\,\Delta(z-x)J(x)\exp\left[-\frac{i}{2}\int d^4xd^4yJ(x)\Delta_F(x-y)J(y)\right].\end{aligned}$$

Similarly we continue further with higher functional derivatives. For the second functional derivative we have

$$\begin{aligned}&\left(-i\frac{\delta}{\delta J(z)}\right)^2\exp\left[-\frac{i}{2}\int d^4xd^4y\,J(x)\Delta_F(x-y)J(y)\right]\\ &= \left\{i\Delta_F(0)+\left[\int d^4x\Delta_F(z-x)J(x)\right]^2\right\}\exp\left[-\frac{i}{2}\int J_x\Delta_F(x-y)J_y\right].\end{aligned}$$

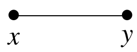
For third derivative

$$\left(-i\frac{\delta}{\delta J(z)}\right)^3 \exp\left[-\frac{i}{2}\int d^4x d^4y\, J_x \Delta_F(x-y) J_y\right]$$
$$= \left\{3\left[-i\Delta_F(0)\right]\int d^4x \Delta_F(z-x)J(x) - \left[\int d^4x \Delta_F(z-x)J(x)\right]^3\right\}$$
$$\times \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right], \qquad (4.290)$$

and finally for the fourth derivative

$$\left(-i\frac{\delta}{\delta J(z)}\right)^4 \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right]$$
$$= \left\{-3\left[\Delta_F(0)\right]^2 + 3i\Delta_F(0)\left[\int d^4x \Delta_F(z-x)J(x)\right]^2\right.$$
$$\left. + \; 3i\Delta_F(0)\left[\int d^4x \Delta_F(z-x)J(x)\right]^2 + \left[\int d^4x \Delta_F(z-x)J(x)\right]^4\right\}$$
$$\times \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right]. \qquad (4.291)$$

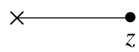
We may write this last expression diagrammatically. Let

$\Delta_F(x-y) \quad \sim$  (4.292)

represents the free propagator (often is instead of $\Delta_F(x-y)$ taken $i\Delta_F(x-y)$). In particular, $\Delta_F(0) = \Delta_F(z,z) = \Delta_F(z-z)$ is then represented by a closed loop (bubble diagram)

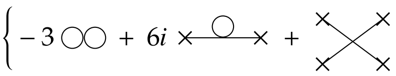
$$\Delta_F(z,z) = \Delta_F(0) \quad \sim \quad \bigcirc. \qquad (4.293)$$

We also introduce the notation

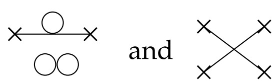
$\int d^4x\, J(x)\Delta_F(x-z) \quad \sim$  (4.294)

With these we can write (4.291) as

$$\left(-i\frac{\delta}{\delta J(z)}\right)^4 \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right]$$

$= \Big\{$  $\Big\} \exp\left(-\frac{i}{2}\int J\Delta_F J\right)$. (4.295)

The meeting of four lines at a point in diagrams



is clearly a consequence of the fact that $\mathcal{L}_I$ contains the ϕ^4 term. Moreover, the coefficients 3, 6 and 1 in Eq. (4.295) follow from rather simple symmetry considerations:

- Factor 3 results from joining up the 2 pairs of lines in the $\times$ diagram. In particular, pick up any line, there are 3 ways how to connect it with remaining 3 lines. This will give us one closed loop diagram. The second loop in the "double" bubble diagram is obtained by connecting the remaining two lines (there is only one way how this can be done). Altogether there are 3 ways how to generate the $\bigcirc\!\!\bigcirc$ diagram.

- Factor 6 results from joining any two lines in the $\times$ diagram (3 ways). This gives one bubble. The remaining two legs have two ways how to orient themselves (which one goes left and which one right). Altogether there are 6 ways how to generate the $\times\!\!-\!\!\overset{\bigcirc}{-}\!\!-\!\!\times$ diagram.

These numerical factors (or better their inverses) are known as *symmetry factors*. Diagram $\bigcirc\!\!\bigcirc$ is known as *vacuum graph* or *bubble diagram* or *vacuum bubble diagram* because it has no external lines. The meaning of this terminology will become clearer shortly.

It is easy to write down the denominator of $\tilde{Z}[J]$. In particular

$$\left[\exp\left(i\int d^4x\mathcal{L}_I\right)\exp\left(-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right)\right]\bigg|_{J=0} = 1 - i\frac{\lambda}{4!}\int(-3\,\bigcirc\!\!\bigcirc)\,d^4z\,, \tag{4.296}$$

and the complete generating functional $\tilde{Z}[J]$ to order λ is equal to

$$\frac{\left[1 - i\frac{\lambda}{4!}\int\left(-3\,\bigcirc\!\!\bigcirc + 6i\, \times\!\!-\!\!\overset{\bigcirc}{-}\!\!-\!\!\times + \;\times\!\!\!\!\times\right)d^4z\right]e^{-\frac{i}{2}\int J\Delta_F J}}{1 - i\frac{\lambda}{4!}\int(-3\,\bigcirc\!\!\bigcirc)\,dz}\,. \tag{4.297}$$

By employing the binomial expansion we finally obtain (again to order λ)

$$\tilde{Z}[J] = \left[1 - i\frac{\lambda}{4!}\int\left(6i\, \times\!\!-\!\!\overset{\bigcirc}{-}\!\!-\!\!\times + \;\times\!\!\!\!\times\right)d^4z\right]e^{-\frac{i}{2}\int J\Delta_F J}\,. \tag{4.298}$$

Clearly, the order of the perturbation is given by the considered order of $\exp\left[i\int \mathrm{d}^4x\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right]$ in the Taylor expansion, while the order n of the correlation function $\langle x_1, \ldots, x_n\rangle$ follows from the number of $J(x)$'s we keep in the expansion of $\tilde{Z}[J]$ (or $Z[J]$).

Let us now consider a toy model with the self-interaction given by

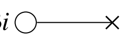
$$\mathcal{L}_I = -\frac{g}{3!}\phi^3\,. \tag{4.299}$$

We will be interested in the second perturbation order in g. To this end we expand $\tilde{Z}[J]$ to order g^2, i.e.

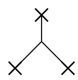
$$\begin{aligned}\tilde{Z}[J] &= \frac{\exp\left[-i\int \mathrm{d}^4x \frac{g}{3!}\left(-i\frac{\delta}{\delta J(x)}\right)^3\right] e^{-\frac{i}{2}\int J\Delta_F J}}{(\text{ ditto })|_{J=0}} \\ &= \frac{\left\{1 - i\frac{g}{3!}\int \mathrm{d}^4x\left(-i\frac{\delta}{\delta J(x)}\right)^3 - \frac{g^2}{2(3!)^2}\left[\int \mathrm{d}^4x\left(-i\frac{\delta}{\delta J(x)}\right)^3\right]^2\right\}}{(\text{ditto})|_{J=0}} \\ &\quad\times \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right]. \end{aligned} \tag{4.300}$$

We will first consider the numerator, i.e., $Z[J]$

$$\begin{aligned} Z[J] &= \left\{1 - i\frac{g}{3!}\int d^4z\left(-3i\right.\right. \end{aligned}$$



$$-$$



$$\begin{aligned} &\left.\left.\right) - \frac{g^2}{2(3!)^2}\left[\int d^4x\left(-i\frac{\delta}{\delta J(x)}\right)^3\right]^2\right\} \\ &\quad\times \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right]. \end{aligned} \tag{4.301}$$

On the first line of (4.301) we have used our result from $\lambda\phi^4$ theory, namely the fact that we know what is

$$\left(-i\frac{\delta}{\delta J(x)}\right)^3 \exp\left[-\frac{i}{2}\int d^4x d^4y\, J(x)\Delta_F(x-y)J(y)\right], \tag{4.302}$$

see (4.290). Let us now proceed with the remaining 3 functional derivatives $\left(-i\frac{\delta}{\delta J(z)}\right)$. In particular, we get

$$\begin{aligned} &\left(-i\frac{\delta}{\delta J(z)}\right)\left(-i\frac{\delta}{\delta J(x)}\right)^3 Z_0[J] \\ &= \left(-i\frac{\delta}{\delta J(z)}\right)\left\{-3i\Delta_F(0)\int d^4y\Delta_F(x-y)J(y)\right. \\ &\quad - \left.\left[\int d^4y\Delta_F(x-y)J(y)\right]^3\right\} e^{-\frac{i}{2}\int J\Delta_F J} \\ &= \left\{-3\Delta_F(0)\Delta_F(x-z) + 3i\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\Delta_F(x-z)\right. \\ &\quad + 3i\,\Delta_F(0)\int d^4y\Delta_F(x-y)J(y)\int d^4y\Delta_F(z-y)J(y) \\ &\quad + \left.\left[\int d^4y\Delta_F(x-y)J(y)\right]^3\int d^4y\Delta_F(z-y)J(y)\right\} e^{-\frac{i}{2}\int J\Delta_F J}. \end{aligned} \tag{4.303}$$

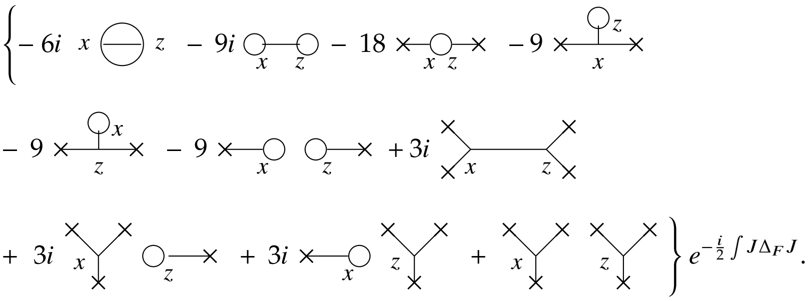
We now proceed with the second variation. This gives

$$
\begin{aligned}
&\left(-i\frac{\delta}{\delta J(z)}\right)^2\left(-i\frac{\delta}{\delta J(x)}\right)^3 Z_0\,[J]\\
&= \left\{6\left[\int d^4y\Delta_F(x-y)J(y)\right][\Delta_F(x-z)]^2\right.\\
&+\ 3\Delta_F(0)\Delta_F(x-z)\int d^4y\Delta_F(z-y)J(y)\\
&+\ 3\,[\Delta_F(0)]^2\int d^4y\Delta_F(x-y)J(y)\\
&-\ 3i\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\left[\int d^4y\Delta_F(z-y)J(y)\right]\Delta_F(x-z)\\
&-\ i\left[\int d^4y\Delta_F(x-y)J(y)\right]^3\Delta_F(0)\\
&+\ 3\Delta_F(0)\Delta_F(x-z)\int d^4y\Delta_F(z-y)J(y)\\
&-\ 3i\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\left[\int d^4y\Delta_F(z-y)J(y)\right]\Delta_F(x-z)\\
&-\ 3i\Delta_F(0)\int d^4y\Delta_F(x-y)J(y)\left[\int dy\Delta_F(z-y)J(y)\right]^2\\
&-\ \left.\left[\int d^4y\Delta_F(x-y)J(y)\right]^3\left[\int d^4y\Delta_F(z-y)J(y)\right]^2\right\}\\
&\times\ \exp\left[-\frac{i}{2}\int d^4x d^4y\,J(x)\Delta_F(x-y)J(y)\right]\\
&= \left\{6\left[\int d^4y\Delta_F(x-y)J(y)\right][\Delta_F(x-z)]^2\right.\\
&+\ 6\Delta_F(0)\Delta_F(x-z)\int d^4y\Delta_F(z-y)J(y)\\
&+\ 3\,[\Delta_F(0)]^2\int d^4y\Delta_F(x-y)J(y)\\
&-\ 6i\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\left[\int d^4y\Delta_F(z-y)J(y)\right]\Delta_F(x-z)\\
&-\ i\left(\int dy\Delta_F(x-y)J(y)\right)^3\Delta_F(0)\\
&-\ -3i\Delta_F(0)\int dy\Delta_F(x-y)J(y)\left(\int dy\Delta_F(z-y)J(y)\right)^2\\
&-\ \left.\left(\int dy\Delta_F(x-y)J(y)\right)^3\left(\int dy\Delta_F(z-y)J(y)\right)^2\right\}e^{-\frac{i}{2}\int J\Delta_F J}.
\end{aligned}
$$

Finally, the third functional derivatives gives

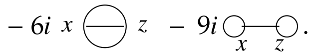
$$
\begin{aligned}
&\left(-i\frac{\delta}{\delta J(z)}\right)^3\left(-i\frac{\delta}{\delta J(x)}\right)^3 Z_0[J] \\
&= \Big\{-6i\left[\Delta_F(x-z)\right]^3 - 6i\Delta_F(0)\Delta_F(x-z)\Delta_F(0) \\
&- 3i\Delta_F(0)\Delta_F(0)\Delta_F(x-z) \\
&- 12\int d^4y\Delta_F(x-y)J(y)\int d^4y\Delta_F(z-y)J(y)\left[\Delta_F(x-z)\right]^2 \\
&- 6\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\Delta_F(x-z)\Delta_F(0) \\
&- 3\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\Delta_F(x-z)\Delta_F(0) \\
&- 3\Delta_F(0)\Delta_F(x-z)\left[\int d^4y\Delta_F(z-y)J(y)\right]^2 \\
&- 6\Delta_F(0)\int d^4y\Delta_F(x-y)J(y)\int d^4y\Delta_F(z-y)J(y)\Delta_F(0) \\
&+ 3i\Delta_F(x-z)\left[\int d^4y\Delta_F(x-y)J(y)\right]^2\left[\int d^4y\Delta_F(z-y)J(y)\right]^2 \\
&+ 2i\Delta_F(0)\left[\int d^4y\Delta_F(x-y)J(y)\right]^3\int d^4y\Delta_F(z-y)J(y) \\
&- 6\int d^4y\Delta_F(x-y)J(y)\int d^4y\Delta_F(z-y)J(y)\left[\Delta_F(x-z)\right]^2 \\
&- 6\Delta_F(0)\Delta_F(x-z)\left[\int d^4y\Delta_F(z-y)J(y)\right]^2 \\
&- 3\Delta_F(0)\Delta_F(0)\int d^4y\Delta_F(x-y)J(y)\int d^4y\Delta_F(z-y)J(y) \\
&+ i\left[\int d^4y\Delta_F(x-y)J(y)\right]^3\left[\int d^4y\Delta_F(z-y)J(y)\right]\Delta_F(0) \\
&+ 3i\Delta_F(0)\int d^4y\Delta_F(x-y)J(y)\left[\int d^4y\Delta_F(z-y)J(y)\right]^3 \\
&+ \left[\int d^4y\Delta_F(x-y)J(y)\right]^3\left[\int d^4y\Delta_F(z-y)J(y)\right]^3\Big\}\, e^{-\frac{i}{2}\int J\Delta_F J}.
\end{aligned}
$$

This rather lengthy expression has quite simple diagrammatic representation, namely



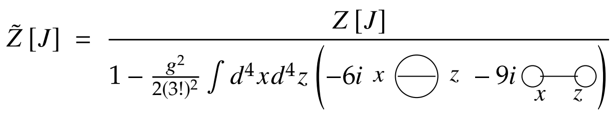
Here the integrations over the vertex positions x and z are implicitly understood.

Thus, at the order g^2 we find the following contribution to $Z[0]$



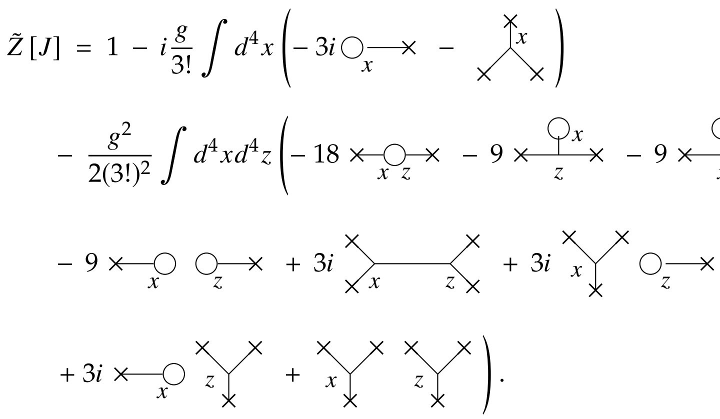
(4.304)

Consequently we can write for $\tilde{Z}[J]$ [cf. Eq. (4.301)]



$$= \frac{\left[1 - i\frac{g}{3!}\,\big(\text{only current diagr.}\big) - \frac{g^2}{2(3!)^2}\,\big(\text{vacuum + current diagr.}\big)\right]}{\left[1 - \frac{g^2}{2(3!)^2}\,\big(\text{only vacuum diagr.}\big)\right]} Z_0[J].$$

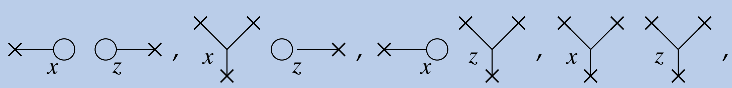
Again, by expanding the denominator the vacuum diagrams will cancel:



(4.305)

Types of diagrams

Diagrams of the following types



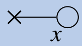
are called **disconnected**. On the other hand, diagrams of the type

are called **vacuum diagrams**, since they have no external lines present.

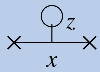
Check that for all connected (and planar) diagrams holds the formula

$$L = I - V + 1, \tag{4.306}$$

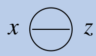
where L is the number of closed loops, I is the number of internal lines and V is the number of vertices. This is the famous **Euler formula for planar graphs**. So, for instance



$\Rightarrow\ I = 1,\ V = 1,\ L = 1\,,$



$\Rightarrow\ I = 2,\ V = 2,\ L = 1\,,$



$\Rightarrow\ I = 3,\ V = 2,\ L = 2\,.$

4.10 More complicated interactions

This section is slightly more technical and can be omitted on a first reading. In some cases (e.g., lower dimensional QFT systems, condense-matter systems or exactly solvable statistical systems) the interacting Lagrangian is complicated (not a simple polynomial), then in order to compute $Z[J]$ (or $\tilde{Z}[J]$) one can use the following identity

$$\begin{aligned} Z[J] &= \exp\left[-i\int d^4x\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right] e^{-\frac{i}{2}\int d^4x d^4y J(x)\Delta_F(x-y)J(y)} \\ &= \exp\left[\frac{i}{2}\int d^4x d^4y \frac{\delta}{\delta\phi(x)}\Delta_F(x-y)\frac{\delta}{\delta\phi(y)}\right] \\ &\quad \times \exp\left\{i\int d^4x\,[-\mathcal{L}_I(\phi(x)) + J(x)\phi(x)]\right\}\Bigg|_{\phi=0}. \end{aligned} \tag{4.307}$$

The passage from the first line to second comes from the simple observation that

$$G\left(-i\frac{\delta}{\delta J}\right)F[iJ] = F\left[\frac{\delta}{\delta\phi}\right]G[\phi]\,e^{i\int d^4x\phi(x)J(x)}\Big|_{\phi=0}, \tag{4.308}$$

which is an infinite-dimensional form of the equation

$$G\left(\frac{\partial}{\partial \boldsymbol{b}}\right)F(\boldsymbol{b}) = F\left(\frac{\partial}{\partial \boldsymbol{x}}\right)G(\boldsymbol{x})e^{\boldsymbol{x}\cdot\boldsymbol{b}}\Big|_{\boldsymbol{x}=0}. \tag{4.309}$$

Here $\partial/\partial\boldsymbol{b}$ is a shorthand notation for a vector $\{\partial/\partial b_i\}_{i=1}^N$ and similarly for $\partial/\partial\boldsymbol{x}$.

The proof is as follows. First we prove (4.309) for a special case $G(\boldsymbol{x}) =$

$e^{\boldsymbol{x}\cdot\boldsymbol{\alpha}}$ and $F(\boldsymbol{b}) = e^{\boldsymbol{\beta}\cdot\boldsymbol{b}}$. The left-hand side then reads

$$G\left(\frac{\partial}{\partial \boldsymbol{b}}\right) F(\boldsymbol{b}) \;=\; e^{\boldsymbol{\alpha}\cdot\frac{\partial}{\partial \boldsymbol{b}}} F(\boldsymbol{b}) \;=\; F(\boldsymbol{b}+\boldsymbol{\alpha}) \;=\; e^{\boldsymbol{\beta}(\boldsymbol{b}+\boldsymbol{\alpha})}, \tag{4.310}$$

and for the right-hand side we get

$$\begin{aligned} F\left(\frac{\partial}{\partial \boldsymbol{x}}\right) G(\boldsymbol{x}) e^{\boldsymbol{x}\cdot\boldsymbol{b}}\Big|_{\boldsymbol{x}=0} &= e^{\boldsymbol{\beta}\cdot\frac{\partial}{\partial \boldsymbol{x}}} e^{\boldsymbol{x}(\boldsymbol{\alpha}+\boldsymbol{b})}\Big|_{\boldsymbol{x}=0} = e^{(\boldsymbol{x}+\boldsymbol{\beta})(\boldsymbol{\alpha}+\boldsymbol{b})}\Big|_{\boldsymbol{x}=0} \\ &= e^{\boldsymbol{\beta}(\boldsymbol{\alpha}+\boldsymbol{b})}, \end{aligned} \tag{4.311}$$

which clearly coincide with the left-hand-side result. The result is then true for any F and G as one may express F and G as a Fourier series, which the preserves the result by term.

To provide a simple illustration of (4.307), we consider

$$Z \;=\; \exp\left(\frac{1}{2}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}\right)\exp\left[-V(x)+\boldsymbol{b}\boldsymbol{x}\right]\Big|_{\boldsymbol{x}=0}. \tag{4.312}$$

We get a perturbative expansion by expanding both exponentials. Let us begin with the case where $\boldsymbol{b} = 0$ and use the notation

$$V_{i_1,i_2,i_3\ldots i_k} \;=\; \frac{\partial}{\partial x_{i_1}}\frac{\partial}{\partial x_{i_2}}\frac{\partial}{\partial x_{i_3}}\cdots\frac{\partial}{\partial x_{i_k}}V(\boldsymbol{x})\Big|_{\boldsymbol{x}=0}. \tag{4.313}$$

Assume further that $V(\mathbf{0}) = 0$, $V_i(\mathbf{0}) = 0$, so that $V(x)$ is at least quadratic. Then we get to the second order in V

$$\begin{aligned} Z \;&=\; \left(1 + \frac{1}{2}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}} + \frac{1}{8}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}} + \ldots\right) \\ &\times \left[1 - V(x) + \frac{1}{2}V^2(x) + \ldots\right]\Big|_{\boldsymbol{x}=0} \\ &= 1 - \frac{1}{2}\mathbb{A}^{-1}_{ij}V_{ij} - \frac{1}{8}\mathbb{A}^{-1}_{ij}\mathbb{A}^{-1}_{kl}V_{ijkl} + \frac{1}{4}\mathbb{A}^{-1}_{ij}\left(\partial_{x_i}\partial_{x_j}V^2\right)\Big|_{\boldsymbol{x}=0} \\ &+ \frac{1}{16}\partial_{x_i}\mathbb{A}^{-1}_{ij}\partial_{x_j}\partial_{x_k}\mathbb{A}^{-1}_{kl}\partial_{x_l}V^2\Big|_{\boldsymbol{x}=0} + \ldots . \end{aligned} \tag{4.314}$$

The fourth terms in (4.314) can further be written as

$$\begin{aligned} \mathbb{A}^{-1}_{ij}\left(\partial_{x_i}\partial_{x_j}V^2\right)\Big|_{\boldsymbol{x}=0} \;&=\; \mathbb{A}^{-1}_{ij}2\partial_{x_i}\left(VV_j\right)\Big|_{\boldsymbol{x}=0} \\ &=\; \mathbb{A}^{-1}_{ij}\left(2V_iV_j + 2VV_{ij}\right)\Big|_{\boldsymbol{x}=0} \;=\; 0. \end{aligned} \tag{4.315}$$

In the fifth term

$$\partial_{x_i}\mathbb{A}^{-1}_{ij}\partial_{x_j}\partial_{x_k}\mathbb{A}^{-1}_{kl}\partial_{x_l}V(\boldsymbol{x})V(\boldsymbol{x})\Big|_{\boldsymbol{x}=0}, \tag{4.316}$$

the contributions V_iV_{jkl} or VV_{jklm} are zero due to conditions $V_i|_{\boldsymbol{x}=0} = V|_{\boldsymbol{x}=0} = 0$. The only non-tivial contributions are from two derivatives acting on each V separately. There are 3 possible pairings $V_{ij}V_{kl}$, $V_{ik}V_{jl}$

and $V_{il}V_{jk}$ which result in

$$2\mathbb{A}^{-1}_{ij}V_{ij}\mathbb{A}^{-1}_{kl}V_{kl} \;+\; 2\mathbb{A}^{-1}_{ij}V_{ik}\mathbb{A}^{-1}_{kl}V_{jl} \;+\; 2\mathbb{A}^{-1}_{ij}V_{il}\mathbb{A}^{-1}_{ij}V_{jk}\,. \tag{4.317}$$

Factor 2 results from symmetry of VV coming from the first derivative. In addition, the *second* and the *third* term are identical after re-indexing.

Note that both $\mathbb{A}^{-1}_{ij}$ and V_{ij} are symmetric in the two indices.

The corresponding contribution to Z is thus

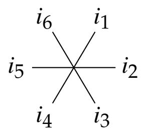
$$\frac{1}{8}\mathbb{A}^{-1}_{ij}V_{ij}\mathbb{A}^{-1}_{kl}V_{kl} \;+\; \frac{1}{4}\mathbb{A}^{-1}_{ij}V_{ik}\mathbb{A}^{-1}_{kl}V_{jl}\,. \tag{4.318}$$

As an exercise, show that should we have expanded (4.314) to the 3rd order in propagator then the corresponding contribution (still to 2nd order in V) would be

$$\frac{1}{3!}\frac{1}{2^3}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}V^2(\boldsymbol{x})\bigg|_{\boldsymbol{x}=0}$$

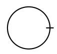
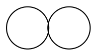
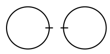
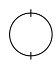
$$= \frac{1}{8}V_{ijk}\mathbb{A}^{-1}_{ij}\mathbb{A}^{-1}_{kl}\mathbb{A}^{-1}_{mn}V_{lmn} \;+\; \frac{1}{12}V_{ijk}\mathbb{A}^{-1}_{il}\mathbb{A}^{-1}_{jm}\mathbb{A}^{-1}_{kn}V_{lmn} \;+\; \ldots\,. \tag{4.319}$$

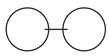
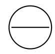
This can be diagrammatically represented as follows: $\mathbb{A}^{-1}_{ij}$ joins points i and j and $V_{i_1 i_2 \ldots i_n}$ represents a vertex with n lines, for instance for $n = 6$ we would have



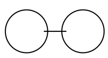
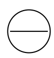
. (4.320)

Then

$Z \;=\; 1 - \frac{1}{2}$  $- \frac{1}{8}$  $+ \frac{1}{8}$  $+ \frac{1}{4}$ 

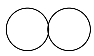
$+ \frac{1}{8}$  $+ \frac{1}{12}$  $+ \ldots\,.$ (4.321)

These are vacuum diagrams (the third one is disconnected). If specially $V(\boldsymbol{x}) = \sum_i x_i^3$, only diagrams with $V_{ijk} \neq 0$ survive, i.e.

 and .

Both are of the second order and up to a different symmetry factor they coincide with vacuum diagrams in $\frac{g}{3!}\varphi^3$ theory.

Similarly, for $V(\boldsymbol{x}) = \sum_i x_i^4$, only diagrams with $V_{ijkl} \neq 0$ survive, which are represented by

.

This is a first order vacuum diagram for $\frac{\lambda}{4!}\varphi^4$ theory (again modulo different symmetry factor).

For the case $\boldsymbol{b} \neq 0$ (i.e., by including also external legs) we still assume that $V(\mathbf{0}) = 0$ and $V_i(\mathbf{0}) = 0$ for $\forall i$. Then

$$
\begin{aligned}
Z[\boldsymbol{b}] \;&=\; \left(1 + \frac{1}{2}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}} + \frac{1}{8}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}} + \ldots\right) \\
&\times\; \left\{1 - [V(x) + \boldsymbol{b}\boldsymbol{x}] + \frac{1}{2}[V(x) + \boldsymbol{b}\boldsymbol{x}]^2 + \ldots\right\}\Big|_{\boldsymbol{x}=0} . \qquad (4.322)
\end{aligned}
$$

Now the following new terms appear

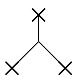
$$
\text{Term 1:} \quad \frac{1}{2}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{b}\boldsymbol{x}\Big|_{\boldsymbol{x}=0} \;=\; 0 ,
$$

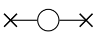
$$
\begin{aligned}
\text{Term 2:} \quad & \frac{1}{2}\cdot\frac{1}{2}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}2\boldsymbol{b}\boldsymbol{x}V(x)\Big|_{\boldsymbol{x}=0} \;=\; 2\frac{1}{4}\mathbb{A}^{-1}_{ij}\left(V b_k x_k\right)_{ij}\Big|_{\boldsymbol{x}=0} , \\
&= 2\frac{1}{4}\mathbb{A}^{-1}_{ij}\Big(V_{ij}\boldsymbol{b}\boldsymbol{x} + \underbrace{V_i b_j + V_j b_i}_{=\,0}\Big)\Big|_{\boldsymbol{x}=0} \;=\; 0
\end{aligned}
$$

$$
\begin{aligned}
\text{Term 3:} \quad & \frac{1}{2}\cdot\frac{1}{2}\frac{\partial}{\partial \boldsymbol{x}}\mathbb{A}^{-1}\frac{\partial}{\partial \boldsymbol{x}}b_k x_k b_l x_l\Big|_{\boldsymbol{x}=0} \\
&= \frac{1}{4}\mathbb{A}^{-1}_{ij}\partial_i\left(b_j b_l x_l + b_k x_k b_j\right)\Big|_{\boldsymbol{x}=0} \;=\; \frac{1}{2}b_i\mathbb{A}^{-1}_{ij}b_j ,
\end{aligned}
$$

where the last term is the first non-trivial contribution (apart from already computed vacuum diagrams).

One can show that there are other higher-order terms like

$-\frac{1}{6}$  $\sim\; -\frac{1}{6}b_i b_j b_k\mathbb{A}^{-1}_{il}\mathbb{A}^{-1}_{kn}\mathbb{A}^{-1}_{jm}V_{lmn}$,

$\frac{1}{4}$  $\sim\; \frac{1}{4}b_i b_j\mathbb{A}^{-1}_{ik}\mathbb{A}^{-1}_{jl}V_{kmn}V_{lpq}\mathbb{A}^{-1}_{mp}\mathbb{A}^{-1}_{nq}$, (4.323)

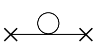
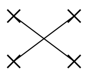
which we have already seen in the $\frac{g}{3!}\phi^3$ theory.

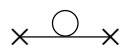
Full two-point Green Function

Let us now come back and proceed with the $\frac{\lambda}{4!}\phi^4$ system. Important quantity of interest is the full two-point Green function, i.e

$$
\langle x_1, x_2\rangle \;\equiv\; \tau(x_1, x_2) \;=\; (-i)^2\frac{\delta^2\tilde{Z}[J]}{\delta J(x_2)\delta J(x_1)}\Big|_{J=0} . \qquad (4.324)
$$

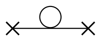
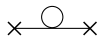
Let us remind that to the leading order in λ we have [cf. Eq. (4.298)]

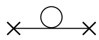
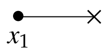
$\tilde{Z}[J] = \Big[1 - \frac{i\lambda}{4!}\int\Big(6i$  $+$  $\Big)d^4x\Big]\, e^{-\frac{i}{2}\int J\Delta_F J}$. (4.325)

So, the first term in $\langle x_1, x_2\rangle$ is $i\Delta_F(x_1 - x_2)$, which is the free particle propagator. Term  contains 4 J's and so gives no contribution to the two-point Green function. The term  equals to [recall (4.291) and (4.295)]

$$\underbrace{\frac{-i\lambda 6i}{4!}}_{\lambda/4}\Delta_F(0)\int d^4x d^4y d^4z \Delta_F(z-x)J(x)\Delta_F(z-y)J(y)\,. \quad (4.326)$$

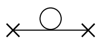
On differentiation we get

$(-i)\dfrac{\delta}{\delta J(x_1)}\Big($  $e^{-\frac{i}{2}J\Delta_F J}\Big) = (-i)\Big[\dfrac{\delta}{\delta J(x_1)}\Big($  $\Big)\Big]e^{-\frac{i}{2}J\Delta_F J}$

$+\ (-i)$  $\Big(-$  $\Big)e^{-\frac{i}{2}J\Delta_F J}$

$$= \frac{-i\lambda}{2}\Delta_F(0)\int d^4y d^4z \Delta_F(z-x_1)\Delta_F(z-y)J(y)e^{-\frac{i}{2}J\Delta_F J} + \dots, \quad (4.327)$$

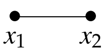
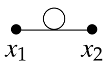
where the remaining terms are not important in the $J \to 0$ limit. The second derivative then reads

$(-i)^2\dfrac{\delta^2}{\delta J(x_2)\delta J(x_1)}\Big($  $e^{-\frac{i}{2}J\Delta_F J}\Big)$

$$= -\frac{\lambda}{2}\Delta_F(0)\int d^4z \Delta_F(z-x_1)\Delta(z-x_2)e^{-\frac{i}{2}J\Delta_F J} + \dots, \quad (4.328)$$

where "..." denotes the terms that do not contribute in the limit $J \to 0$. Finally, we can write the two-point Green function as

$$\langle x_1, x_2\rangle = i\Delta_F(x_1,x_2) - \frac{\lambda}{2}\Delta_F(0)\int d^4z \Delta_F(z-x_1)\Delta_F(z-x_2) + O(\lambda^2)$$

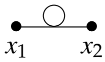
$= i$  $-\ \dfrac{\lambda}{2}$  $+\ O(\lambda^2)\,.$ (4.329)

To order λ, this represents the effect of interaction on the free-particle propagation.

Let us remind that the free propagator is given as

$$\Delta_F(x-y) = \frac{1}{(2\pi)^4}\int \frac{e^{-ik(x-y)}}{k^2-m^2+i\varepsilon}d^4k\,, \quad (4.330)$$

and its Fourier transform contains a pole at $k^2 = m^2$. This identifies mass of the particle as m. We will see that this is not a coincidence but a consequence of the structure of S-matrix. Let us now see that the effect of the interaction is to change the value of the physical mass away from m. Indeed, the second term in (4.374) is



$$\begin{aligned} &= \Delta_F(0)\int d^4z\Delta_F(x_1-z)\Delta_F(x_2-z) \\ &= \frac{\Delta_F(0)}{(2\pi)^8}\int \frac{e^{-ip(x_1-z)}}{p^2-m^2+i\varepsilon}\frac{e^{-iq(x_2-z)}}{q^2-m^2+i\varepsilon}\,d^4pd^4qd^4z \\ &= \frac{\Delta_F(0)}{(2\pi)^4}\int \frac{e^{-ip(x_1-x_2)}}{(p^2-m^2+i\varepsilon)^2}\delta^{(4)}(p+q)\;d^4pd^4q \\ &= \frac{\Delta_F(0)}{(2\pi)^4}\int \frac{e^{-ip(x_1-x_2)}}{(p^2-m^2+i\varepsilon)^2}\,d^4p\,. \end{aligned} \tag{4.331}$$

So, to the leading order in λ we have for z-point Green function

$$\langle x_1, x_2\rangle = \frac{i}{(2\pi)^4}\int \frac{e^{-ip(x_1-x_2)}}{p^2-m^2+i\varepsilon}\left[1-\frac{\frac{i}{2}\lambda\Delta_f(0)}{p^2-m^2+i\varepsilon}\right]d^4p\,. \tag{4.332}$$

Technical note

$$\frac{1}{(A+\lambda B)} = \{\lambda << 1\} = \frac{1}{A(1+\lambda A^{-1}B)} = \frac{1}{A}(1-\lambda A^{-1}B)\,.$$

With this we can rewrite (4.332)

$$\langle x_1, x_2\rangle = \frac{i}{(2\pi)^4}\int \frac{e^{-ip(x_1-x_2)}}{p^2-m^2-\frac{i}{2}\Delta_F(0)\lambda+i\varepsilon}\,d^4p\,. \tag{4.333}$$

The Fourier transform of $\langle x_1, x_2\rangle$ will now possess a pole at p^2 equal to

$$m^2+\frac{i}{2}\lambda\Delta_F(0) = m^2+\delta m^2 = m_R^2\,, \tag{4.334}$$

where $\delta m^2 = \frac{i}{2}\lambda\Delta_F(0)$. The mass m_R is now identified with the *physical mass* and for reasons to be explained in the chapter on renormalization is known also as *renormalized mass*.

Note I.

$\Delta_F(0)$ is divergent. One says that $\Delta_F(0)$ is quadratically divergent. This is because for large p the integrand behaves as $\frac{d^4p}{p^2} = d\Omega dp\frac{p^3}{p^2} = d\Omega dpp$. Integral over p behaves as $\frac{1}{2}p^2|_0^{+\infty}$, which diverges quadratically. We will discuss this point more in the part dedicated to renormalization.

Note II.

Important observation is that the *renormalized mass* is not the same as the parameter m in the Lagrangian. The same will be true also for *renormalized couplings*.

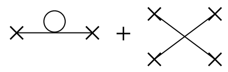
4-point Green function

Let us now compute 4-point Green function to the first order in λ. We start we the defining relation

$$\langle x_1, x_2, x_3, x_4\rangle_0 = (-i)^4 \frac{\delta^4 \tilde{Z}[J]}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\bigg|_{J=0}, \qquad (4.335)$$

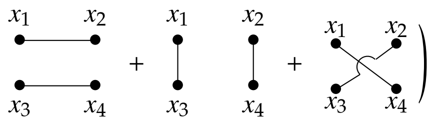
where to the leading order in λ we know [cf. Eq. (4.325)] that

$$\tilde{Z}[J] = \left[1 - \frac{i\lambda}{4!}\int \left(6i \; \text{[diagram]} \; + \; \text{[diagram]}\right) d^4x\right] e^{-\frac{i}{2}\int J\Delta_F J}, \qquad (4.336)$$



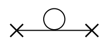
The first (i.e., order λ^0) term in $\langle x_1, \ldots, x_4\rangle$ is

$$\begin{aligned}\langle x_1, \ldots, x_4\rangle_0 &= -[\Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4) + \Delta_F(x_1 - x_3)\Delta_F(x_2 - x_4) \\ &\quad + \Delta_F(x_1 - x_4)\Delta_F(x_2 - x_3)] \\ &= -\left(\text{[diagram]} + \text{[diagram]} + \text{[diagram]}\right). \qquad (4.337)\end{aligned}$$

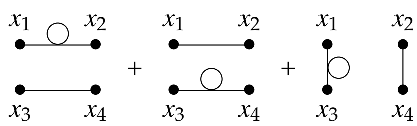


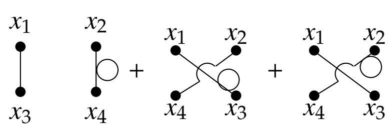
The next term in $\tilde{Z}[J]$ of order λ is given by

$$\frac{\lambda}{4}(-i)^4 \frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\left[\text{[diagram]} \; e^{-\frac{i}{2}\int J\Delta_F J}\right]\bigg|_{J=0}$$



$$\begin{aligned}&= \frac{\lambda}{4}\frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\left[\Delta_F(0)\int d^4x d^4y d^4z \Delta_F(x-z)\Delta_F(y-z)\right. \\ &\qquad\qquad \left.\times \; J(y)J(x)\, e^{-\frac{i}{2}\int J\Delta_F J}\right]\bigg|_{J=0} \\ &= \frac{-i\lambda}{8}\Delta_F(0)\int d^4x d^y d^4z d^4z_1 d^4z_2 \; \Delta_F(x-z)\Delta_F(y-z) \\ &\quad \times \; \Delta_F(z_1 - z_2)\; \frac{\delta^4 J_x J_y J_{z_1} J_{z_2}}{\delta J_{x_1}\delta J_{x_2}\delta J_{x_3}\delta J_{x_4}} \\ &= \frac{-i\lambda}{8}\int d^4z \left[\text{[diagram]} + \text{[diagram]} + \text{[diagram]} + \text{[diagram]} + \text{[diagram]} + \text{[diagram]}\right]. \qquad (4.338)\end{aligned}$$

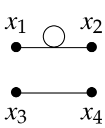




So, we have 24 terms — each diagram represents 4 equivalent terms.

Here, for instance

$$\int d^4z \left[\text{[diagram]}\right] = \int d^4z \Delta_F(0)\Delta_F(x_1 - z)\Delta_F(x_2 - z)\Delta(x_3 - x_4), \qquad (4.339)$$

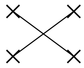


etc.

Note

$$\Delta_F(x) = \frac{1}{(2\pi)^4}\int \frac{e^{-ipx}d^4p}{p^2-m^2+i\varepsilon} \Rightarrow \Delta_F(x-y) = \Delta_F(y-x) \quad (4.340)$$

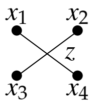
The final term in $\tilde{Z}[J]$ of order λ is given by

$$\frac{-i\lambda}{4!}\frac{\delta^4}{\delta J_{x_1}\delta J_{x_2}\delta J_{x_3}\delta J_{x_4}}\Bigg[$$

$$e^{-\frac{i}{2}\int J\Delta_F J}\Bigg]\Bigg|_{J=0}$$

$$= \frac{-i\lambda}{4!}\cdot\frac{\delta^4}{\delta J_{x_1}\delta J_{x_2}\delta J_{x_3}\delta J_{x_4}}\int d^4z\left[\int d^4x\Delta_F(z-x)J(x)\right]^4$$

$$= \frac{-i\lambda}{4!}\int d^4zd^4y_1d^4y_2d^4y_3d^4y_4\ \Delta_F(z-y_1)\Delta_F(z-y_2)$$

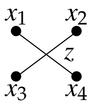
$$\times\ \Delta_F(z-y_3)\Delta_F(z-y_4)\frac{\delta^4 J(y_1)J(y_2)J(y_3)J(y_4)}{\delta J_{x_1}\delta J_{x_2}\delta J_{x_3}\delta J_{x_4}}$$

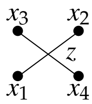
$$= \frac{-i\lambda}{4!}\int d^4z\Bigg[$$


$$+\ \text{all permutations of } x_1,\ldots,x_4.\Bigg]. \quad (4.341)$$

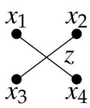
Note

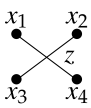
$\langle x_1,\ldots,x_4\rangle$ is by its very formulation given via time ordered product symmetric under permutation of positions $x_1,\ldots,x_4$ (this can also be directly seen from the functional integral representation of $\langle x_1,\ldots,x_4\rangle$, there $\phi(x_1),\ldots,\phi(x_4)$ enter as c-numbered functions, which clearly commute.

From this point of view diagrams


and


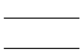
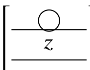
are the same and we can count them as 2. This is true for all 24 copies. So, the previous result can be written as

$$\frac{-i\lambda}{4!}\int d^4z\Bigg[$$


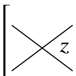
$$\Bigg]\times 24 = -i\lambda\int d^4z\Bigg[$$


$$\Bigg]. \quad (4.342)$$

The same is true also for previous 3 types of diagrams - each with multiplicity 4. So, finally we can schematically wite

$$\langle x_1,\ldots,x_4\rangle = -3\Bigg[$$

$$\Bigg] - 3i\lambda\int d^4z\Bigg[$$


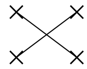
$$\Bigg]$$

$$-\ i\lambda\int d^4z\Bigg[$$


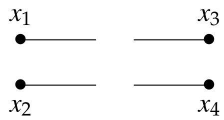
$$\Bigg] \quad (4.343)$$

The first term of order λ^0 does not contribute to the scattering since propagating particles are not disturbed in these evolution (no interaction is present). The numerical coefficients are easily derived by simple combinatorics.

For instance, if we want to find the contribution to order λ^n, we need to consider n-vertices. In short

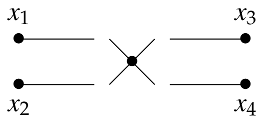
n vertices of the type  contribute to order λ^n . (4.344)

For 4-point function we draw four external legs



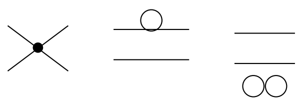
$$\Leftarrow \langle x_1 \ldots x_4 \rangle \,. \tag{4.345}$$

In particular, the 4-point function in $\lambda\phi^4$ theory to order λ is constructed from following diagrammatic building blocks (so-called Feynman prediagrams)



. (4.346)

Now we should join all lines (keeping external legs) and create all **topologically distinct** types of diagrams. Corresponding diagrams, the so-called **Feynman diagrams** are:



. (4.347)

Let us see how to deal with combinatorial factors (also known as *multiplicity of diagram*). The general idea is the following. If we want to build first diagram from (4.347), we start with prediagram (4.346), where we can connect one of the legs with vertex in 4 different ways. After that we have 3 legs remaining and 3 free legs in vertex, etc. [cf. Fig. 4.8]

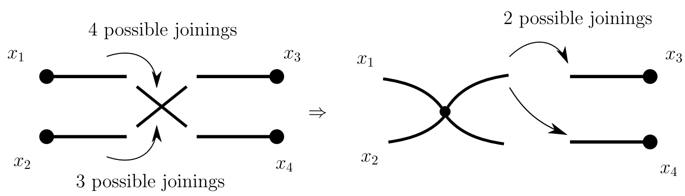


Figure 4.8: Construction of the multiplicity for the first diagram in (4.347).

Hence, we can see that multiplicity of this diagram is

$$4 \cdot 3 \cdot 2 \cdot 1 \;=\; 4! \;=\; 24 \,. \tag{4.348}$$

This precisely cancel the factor 4! in the definition of $\lambda\phi^4/4!$. This, in turn provides precisely the corresponding coefficient in $\langle x_1 \ldots x_4 \rangle$ in Eq. (4.343).

For the middle diagram in (4.347), we first can connect one external line with another external line in 3 possible ways. After that there are 2 ways to connect one leg in middle vertex with remaining external legs and there are 4 legs in the vertex. Then again, there are 3 ways to connect remaining legs to vertex. Finally, we have 1 way to connect remaining two lines in the middle prediagram, which builds loop, see Fig 4.10.

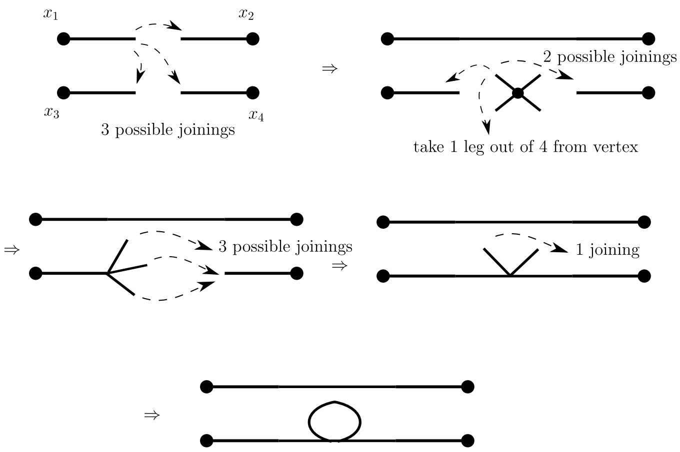


Figure 4.9: Construction of the multiplicity for the second diagram in (4.347).

Thus, the multiplicity is

$$3 \cdot 2 \cdot 4 \cdot 3 \;=\; 24 \cdot 3\,, \tag{4.349}$$

which when taken together with $\frac{1}{4!}$ gives precisely the factor of 3 in Eq. (4.343).

As for last diagram in (4.347), we can again start with prediagram (4.346) and connect external legs in 3 different ways. Then, the vertex can be connected into double bubble diagram in 3 different ways, see Fig **??**.

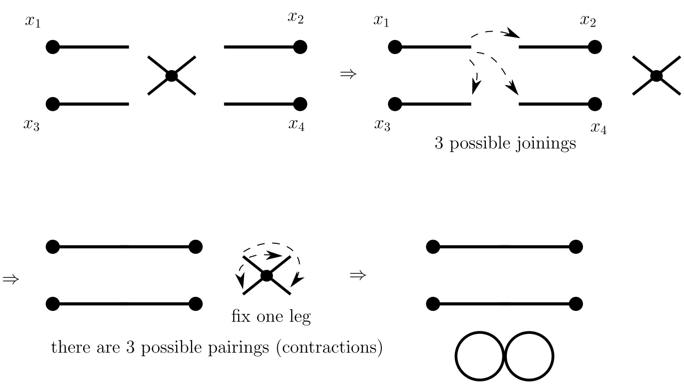


Figure 4.10: Construction of the multiplicity for the third diagram in (4.347).

Thus this diagram has the multiplicity

$$3 \cdot 3 \;=\; 9\,. \tag{4.350}$$

One can check that we have all diagrams by realizing that

$$\begin{aligned}\langle x_1 \dots x_4 \rangle &= \frac{\int \mathcal{D}\varphi \varphi(x_1)\dots\varphi(x_4) \exp\left\{i \int d^4x \left(\mathcal{L}_0 - \mathcal{V}(\varphi)\right)\right\}}{\int \mathcal{D}\varphi \exp\left\{i \int d^4x \left(\mathcal{L}_0 - \mathcal{V}(\varphi)\right)\right\}} \\ &= \frac{\int \mathcal{D}\varphi \varphi(x_1)\dots\varphi(x_4) e^{-i\int d^4x \mathcal{V}(\varphi)} \exp\left\{i \int d^4x \left(\mathcal{L}_0\right)\right\}}{\int \mathcal{D}\varphi \exp\left\{i \int d^4x \left(\mathcal{L}_0\right)\right\}} \\ &\quad \times \frac{\int \mathcal{D}\varphi \exp\left\{-i \int d^4x \mathcal{L}_0\right\}}{\int \mathcal{D}\varphi \exp\left\{i \int d^4x \mathcal{V}(\varphi)\right\} \exp\left\{i \int d^4x \mathcal{L}_0\right\}} \\ &= \frac{\left\langle \varphi(x_1)\dots\varphi(x_4) e^{-i\int d^4x V}\right\rangle_0}{\left\langle e^{-i\int d^4x V}\right\rangle_0} . \end{aligned} \tag{4.351}$$

In particular, in $\lambda\frac{\varphi^4}{4!}$ theory we have

$$\begin{aligned}&\langle x_1 \dots x_4 \rangle \\ &= \frac{\langle 0| T\left[\varphi(x_1)\dots\varphi(x_4)\left(1 - i\frac{\lambda}{4!}\int d^4x \varphi^4(x) + O(\lambda^2)\right)\right]|0\rangle}{\langle 0| T\left[\left(1 - i\frac{\lambda}{4!}\int d^4x \varphi^4(x) + O(\lambda^2)\right)\right]|0\rangle} . \end{aligned} \tag{4.352}$$

So, to order λ contribute all contractions from

$$\begin{aligned}&\langle 0| T\left[\varphi(x_1)\dots\varphi(x_4)\int d^4x \varphi^4(x)\right]|0\rangle \\ &\qquad = \int d^4x \langle 0| T\left[\varphi(x_1)\dots\varphi(x_4)\varphi^4(x)\right]|0\rangle . \end{aligned} \tag{4.353}$$

We know that there are in total $\frac{(2M)!}{2^M M!}$ contractions [cf. Eq. (4.222)]. Since in our case $M = 4$, we have $\frac{8!}{2^4 4!} = 105$ contractions. On the other hand, our 3 contributing Feynman diagrams have the multiplicity $24 + 24 \cdot 3 + 9 = 105$. So, we have correct number of diagrams and respective multiplicities.

The reason why the vacuum diagram does not appear in $\langle x_1 \dots x_4 \rangle$ in Eq. (4.343) is because it is precisely cancelled by the very same diagram in denominator. This result is completely general and it is known as **linked cluster property**. We will derive this result shortly.

In summary, the **Feynman rules** for $\lambda\frac{\phi^4}{4!}$ scalar field theory **in coordinate space** are

- Draw all topologically distinct diagrams. For given n-point Green function with n external legs. For order λ^m use m vertices.
- A line between points x and y represents propagator $i\Delta_F(x - y)$.
- A vertex with 4 lines represents a factor $-i\lambda$.
- Integrate over z for all vertices.
- Introduce combinatorial factor, where necessary.

Symmetry factor

Inverse of the overall pre-factor in front of each diagram is known as a **symmetry factor**. For a simple monomial interaction (such those considered so far)

$$s = \frac{n!(\eta)^n}{r}, \tag{4.354}$$

where n is a number of vertices, η is a coupling constant factor (e.h. 4! or 3!) and r is the multiplicity factor (i.e., the combinatorial factor). E.g. s for left diagram in (4.347) is $\frac{1!(4!)^1}{4!} = 1$ and for middle diagram is $\frac{1!(4!)^1}{3\cdot 4!} = \frac{1}{3}$.

Linked Cluster Theorem

We have seen that vacuum diagrams will cancel when $\langle x_1 \dots x_n \rangle$ is perturbatively computed to the order λ for $\lambda\frac{\varphi^4}{4!}$ theory. This result is, in fact, general and true to all orders in λ and also true for quite general potentials. This result is known as **linked cluster theorem**.

Proof: Let us illustrate the situation on the monomial interaction of the type $g\frac{\varphi^k}{k!}$. If we concentrate on n-th perturbative order of general m-point Green function we get

$$\langle x_1 \dots x_m \rangle^{(n)} \equiv \langle F[x] \rangle^{(n)} = \frac{(-ig)^n}{n!(k!)^n} \left\langle F[x] \left(\int d^4z \varphi^k(z) \right)^n \right\rangle_0 . \tag{4.355}$$

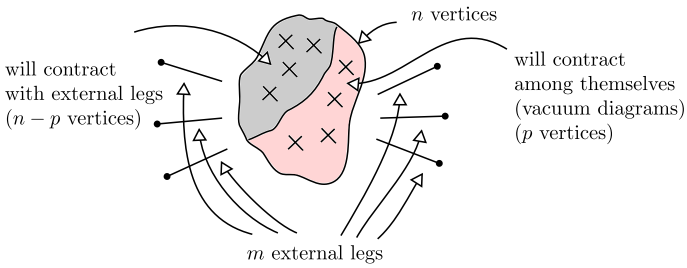


Figure 4.11: Illustration of the Linked Cluster Theorem.

Fig. 4.11 implies that the contribution to (4.355) from vacuum diagrams of p-th order (in coupling g) is

$$\frac{(-ig)^n}{n!(k!)^n} \binom{n}{p} \left\langle F[x] \left(\int d^4z \varphi^k(z) \right)^{n-p} \right\rangle_0^{nv} \left\langle \left(\int d^4z \varphi^k(z) \right)^p \right\rangle_0^{v} . \tag{4.356}$$

Here the combinatorial factor counts how many times one can select p vertices out of n vertices. Acronym "nv" denotes non-vacuum diagrams while v vacuum diagrams (i.e., diagrams without external legs).

By summing over p we get perturbation expansion of the n-th order with all possible vacuum diagrams included. The entire perturbation expansion thus reads

$$\sum_{n=0}^{\infty}\sum_{p=0}^{n}\frac{1}{(k!)^n}\frac{(-ig)^n}{(n-p)!p!}\left\langle F[x]\left(\int d^4z\varphi^k(z)\right)^{n-p}\right\rangle_0^{nv}\left\langle\left(\int d^4z\varphi^k(z)\right)^{p}\right\rangle_0^{v}$$

$$=\sum_{p=0}^{\infty}\sum_{n=p}^{\infty}\frac{1}{(k!)^n}\frac{(-ig)^n}{(n-p)!p!}\left\langle F[x]\left(\int d^4z\varphi^k(z)\right)^{n-p}\right\rangle_0^{nv}\left\langle\left(\int d^4z\varphi^k(z)\right)^{p}\right\rangle_0^{v}.$$

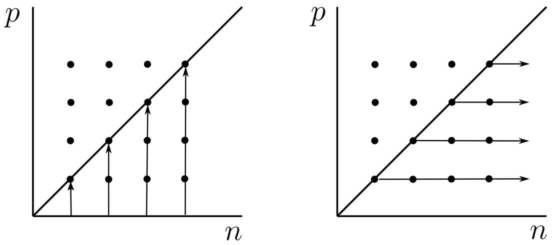


Figure 4.12: Equivalence between $\sum_{n=0}^{\infty}\sum_{p=0}^{n}$ and $\sum_{p=0}^{\infty}\sum_{n=p}^{\infty}$.

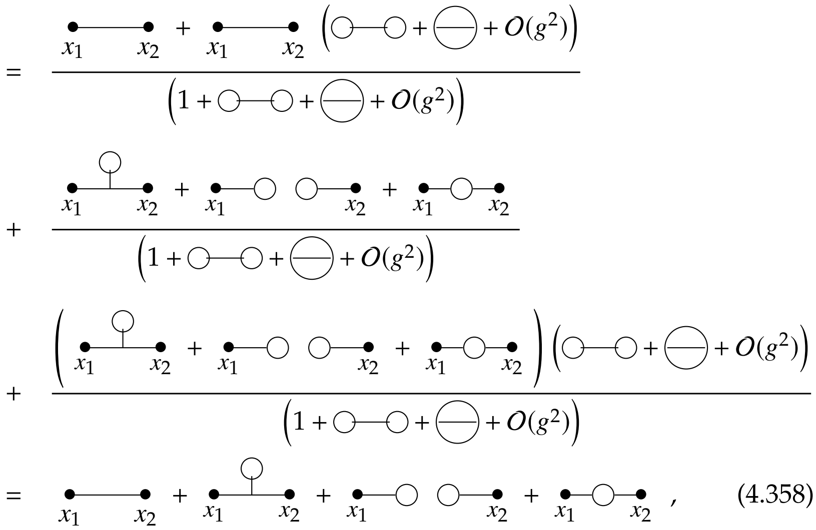
By denoting $n' = n - p$ we can further write

$$\sum_{p=0}^{\infty}\sum_{n'=0}^{\infty}\frac{(-ig)^{n'}}{n'!(k!)^{n'}}\frac{(-ig)^p}{p!(k!)^p}\left\langle F[x]\left(\int d^4z\varphi^k(z)\right)^{n'}\right\rangle_0^{nv}\left\langle\left(\int d^4z\varphi^k(z)\right)^{p}\right\rangle_0^{v} \tag{4.357}$$

Note that the summation over p precisely gives the denominator in $\langle x_1 \dots x_m\rangle$, see, e.g., Eq. (4.351).

Let us now illustrate the Linked Cluster Theorem on a simple example of the $g\varphi^3/3!$ theory to second order in g.

$$\langle x_1x_2\rangle = \frac{\left\langle\phi(x_1)\phi(x_2)e^{-i\int d^4z\mathcal{V}}\right\rangle}{\left\langle e^{-i\int d^4z\mathcal{V}}\right\rangle} = \left\{\mathcal{V} = \frac{g}{3!}\varphi^3\right\}$$



(4.358)

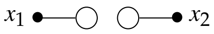
As an exercise, try to fill in the correct symmetry factors.

4.11 Generating Functional for Connected Diagrams

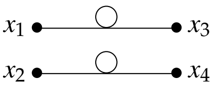
Before attempting to evaluate sums of Feynman diagrams we shall perform some further formal manipulations to simplify the task of organizing them.

As they stand, the Green functions $\langle x_1, \ldots, x_m \rangle$ (say, for example, for scalar theory) are cumbersome quantities to use. In fact, even when vacuum diagrams are removed, there are many diagrams in $\langle x_1, \ldots, x_m \rangle$ that are *disconnected*.

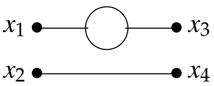
We note here that by a disconnected diagram we mean a diagram composed of two or more subdiagrams that are not linked by propagators, e.g.



$\leftrightarrow \quad \langle x_1 x_2 \rangle_{\phi^3}$,



$\leftrightarrow \quad \langle x_1 x_2 x_3 x_4 \rangle_{\phi^4}$,



$\leftrightarrow \quad \langle x_1 x_2 x_3 x_4 \rangle_{\phi^3}$,   etc.

(Here the sub-index of Green function denotes the type of considered potential.)

On one hand side, the connected diagrams are more elementary building blogs from which one can systematically generate more complicated perturbative diagrams. On other side, we will see that for scattering purposes in particle physics the connected diagrams are very important tools. To this end we will in this section isolate connected parts of Greens functions — the so-called *connected Green functions*. The first step is to break down the Feynman diagrams (and hence Green functions) into their connected parts.

Previously in Eq. (4.243) we have seen that the generating functional for Green functions $Z[J]$ can be expanded as

$$\tilde{Z}[J] = \frac{Z[J]}{Z[0]} = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int_{\mathbb{R}^{4n}} \prod_{i=1}^{n} d^4 x_i \, J(x_1) \ldots J(x_n) \langle x_1 x_2 \ldots x_n \rangle \,, \quad (4.359)$$

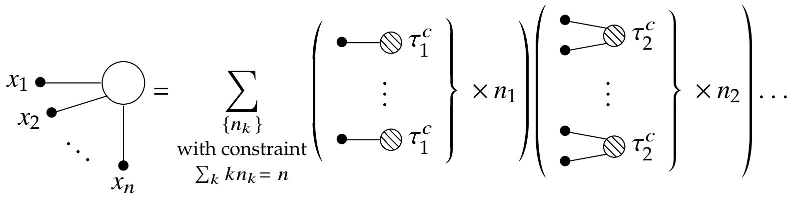
where

$$\langle x_1 x_2 \ldots x_n \rangle = \frac{\int \mathcal{D}\phi \, \phi(x_1) \ldots \phi(x_n) e^{iS[\phi]}}{\int \mathcal{D}\phi \, e^{iS[\phi]}} \,. \quad (4.360)$$

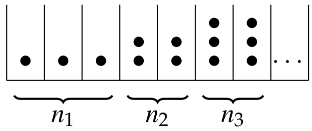
Disconnected Green functions arise when $\langle x_1 x_2 \ldots x_n \rangle$ factorises into (typically sum) of products of Green functions.

Let $\langle x_1 x_2 \ldots x_n \rangle^c \equiv \tau_n^c$ denote connected Green function with n fields. A general Green function can be written as the sum of products of connected Green functions. Let $\langle x_1 x_2 \ldots x_n \rangle \equiv \tau_n$ have n_1 factors of

$\langle x_1\rangle^L \equiv \tau_1^c$, n_2 factors of $\langle x_1 x_2\rangle^c \equiv \tau_2^c$, etc.



This number of ways of factorizing $\langle x_1 x_2 \ldots x_n\rangle$ in this fashion is the same as the number of ways of partitioning n particles with n_1 boxes with one particle each, n_2 boxes with two particles each etc.



$$= \frac{n!}{n_1!n_2!n_3!\ldots(1!)^{n_1}(2!)^{n_2}(3!)^{n_3}\ldots}$$

$$= \frac{n!}{n_1!(1!)^{n_1}n_2!(2!)^{n_2}n_3!(3!)^{n_3}\ldots}\,.$$

To understand this result, it is illuminating to go through the following two examples.

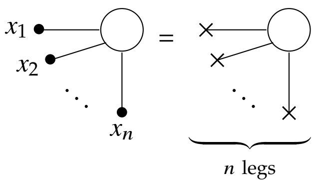
Example 1: In how many ways can we arrange r_1 balls of color 1, r_2 balls of color 2, ..., r_k balls of color k in a sequence of length $n := r_1 + r_2 + \ldots r_k$? If we number the balls 1 to n, then there are $n!$ arrangements. Since we ignore the numbering, any permutation of the set of r_i balls of color i, $1 \le i \le k$, produces the same arrangement. So the answer to the question is the multinominal coefficient $\binom{n}{r_1,\ldots,r_k}$.

Example 2: We wish to split $\{1, 2, \ldots, n\}$ into b_1 subsets of size 1, b_2 subsets of size 2, ..., b_k subsets of size k. Here $\sum_{i=1}^k ib_i = n$. The same argument as used in the previous example applies. Furthermore, the subsets of the same cardinality can be permuted among themselves without changing the configuration. So the solution is

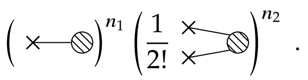
$$\frac{n!}{b_1!b_2!\ldots b_k!(1!)^{b_1}(2!)^{b_2}\ldots(k!)^{b_k}}\,. \tag{4.361}$$

We now multiply by $J(x_1)\ldots J(x_n)$ and integrate to get

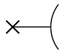
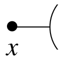
$$\int d^4x_1\ldots d^4x_n\, J(x_1)\ldots J(x_n)$$

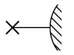
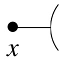


$$= \sum_{\substack{\{n_k\}\\ \text{with constraint}\\ \sum_k kn_k = n}} \frac{n!}{n_1!n_2!\ldots} \tag{4.362}$$



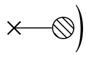
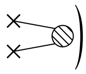
Here we use the obvious symbolic notation

 $= \int d^4x\, J(x)$  (4.363)

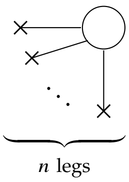
 $= \int d^4x\, J(x)$  (4.364)

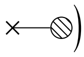
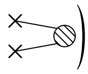
and similarly for more legs.

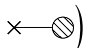
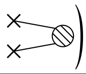
Eq. (4.362) can be further rewritten as

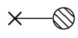
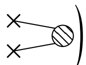
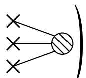
$\sum_{\{n_k\}} \delta\left(n - \sum_k k n_k\right) \frac{n!}{n_1! n_2! n_3! \ldots} \Big($  $\Big)^{n_1} \Big(\frac{1}{2!}$  $\Big)^{n_2} \times \ldots$ (4.365)

where the symbol δ-function stands for corresponding Kronecker's δ function. With this we can write

$\tilde{Z}[J] = \frac{Z[J]}{Z[0]} = \sum_n \frac{i^n}{n!}$ 

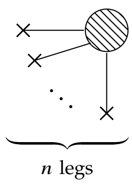
$= \sum_n \frac{i^n}{n!} \sum_{\{n_k\}} \frac{n!}{n_1! n_2! n_3! \ldots} \delta\left(n - \sum_k k n_k\right) \Big($  $\Big)^{n_1} \Big(\frac{1}{2!}$  $\Big)^{n_2} \ldots$

$= \sum_{\{n_k\}} \frac{1}{n_1!} \Big(i$  $\Big)^{n_1} \frac{1}{n_2!} \Big(i^2 \frac{1}{2!}$  $\Big)^{n_2} \ldots$

$= \exp\Big(i$  $\Big) \cdot \exp\Big(i^2 \frac{1}{2!}$  $\Big) \cdot \exp\Big(i^3 \frac{1}{3!}$  $\Big) \ldots$

$$= \exp\left(\tilde{W}[J]\right)\,, \tag{4.366}$$

where

$\tilde{W}[J] = \sum_{n=1} \frac{i^n}{n!}$ 

$$= \sum_{n=1} \frac{i^n}{n!} \int d^4x_1 \ldots d^4x_n \, \langle x_1, \ldots, x_n \rangle^c \, J(x_1) \ldots J(x_n)\,. \tag{4.367}$$

We should note here that it is often convenient to rescale $\tilde{W}[J]$ as $\tilde{W}[J] = iW[J]$ so that

$$\tilde{Z}[J] \;=\; \exp\left(iW[J]\right)\,. \tag{4.368}$$

Since $\tilde{Z}[J] = Z[J]/Z[0]$, we work with normalized generating functional and hence no vacuum diagrams are present in $\tilde{Z}[J]$ (nor in $\langle x_1, \ldots, x_n \rangle \; \forall n$)). This implies that W generates connected diagrams of non-vacuum type —- they have at least one external leg.

Relation to Characteristic Functions

In probability theory one introduces *characteristic function* for any given (multinomial) probability density function $p(\boldsymbol{x})$.

Characteristic function is defined as the Fourier transform of $p(\boldsymbol{x})$, i.e.

$$\phi(\boldsymbol{t}) = \int_{\mathbb{R}^n} d^n\boldsymbol{x}\, e^{i\boldsymbol{t}\boldsymbol{x}}\, p(\boldsymbol{x}) \quad \leftrightarrow \quad \tilde{Z}[J] = \int \mathcal{D}\varphi\, e^{i\int J\varphi}\, \frac{e^{iS[\varphi]}}{\int \mathcal{D}\varphi\, e^{iS[\varphi]}}\,,$$

where the last fraction is analogous to probability density function $p(\boldsymbol{x})$.

Characteristic function carries all information on moments and correlations (if they exist), e.g.,

$$\langle x_i, x_j\rangle = \frac{\partial^2}{\partial(ix_i)\partial(ix_j)}\phi(\boldsymbol{t})\Big|_{\boldsymbol{t}=0} \quad \leftrightarrow \quad \langle 0|\, T[\varphi(x_i)\varphi(x_j)]\, |0\rangle$$

$$= \frac{\delta^2}{\delta(iJ_{x_i})\delta(iJ_{x_j})}\tilde{Z}[J]\Big|_{J=0}\,.$$

Generating function of *cumulants* is defined as

$$H(\boldsymbol{t}) = \log\phi(\boldsymbol{t}) = \log\mathbb{E}\left(e^{i\boldsymbol{t}\boldsymbol{x}}\right) \quad \leftrightarrow \quad \tilde{W}[J] = \log\tilde{Z}[J]\,.$$

Analogy with probability theory will become even stronger when we perform the so-called *Euclidezation* of the functional integral.

Some examples ($W[J]$ in action)

Let us now show how the prescription

$$\tilde{Z}[J] \;=\; e^{iW[J]} \quad \Leftrightarrow \quad W[J] \;=\; -i\log\tilde{Z}[J]\,, \tag{4.369}$$

allows to generate connected diagrams in perturbative analysis. To this end we will consider the 2-point and 4-point Green functions in $\lambda\phi^4$ theory.

We have, firstly

$$\begin{aligned}\frac{\delta^2 W}{\delta J(x_1)\delta J(x_2)} &= \frac{\delta}{\delta J(x_1)}\left(-\frac{i}{\tilde{Z}[J]}\frac{\delta\tilde{Z}[J]}{\delta J(x_2)}\right) = \frac{\delta}{\delta J(x_1)}\left(-\frac{i}{Z[J]}\frac{\delta Z[J]}{\delta J(x_2)}\right)\\ &= \frac{i}{Z^2[J]}\frac{\delta Z[J]}{\delta J(x_1)}\frac{\delta Z[J]}{\delta J(x_2)} - \frac{i}{Z[J]}\frac{\delta^2 Z[J]}{\delta J(x_1)\delta J(x_2)}\,.\end{aligned} \tag{4.370}$$

When $J = 0$,

$$\frac{-i}{Z[J]} \frac{\delta Z[J]}{\delta J(x)}\bigg|_{J=0} = \frac{-i\delta \tilde{Z}[J]}{\delta J(x)}\bigg|_{J=0} = \frac{\int \mathcal{D}\varphi \, \varphi(x) e^{iS[\varphi]}}{\int \mathcal{D}\varphi \, e^{iS[\varphi]}}$$

$$= \frac{\langle \varphi(x) e^{-iV[\varphi]}\rangle_0}{\langle e^{-iV[\varphi]}\rangle_0} = \left\{ V[\varphi] = \lambda \int d^4x \frac{\varphi^4}{4!} \right\} = 0 \,,$$

This would not be true e.g. for $g\varphi^3$ theory.

since $\langle \varphi(x) e^{-iV[\varphi]}\rangle_0$ is (when $e^{-iV[\varphi]}$ is expanded) the vacuum expectation of the time ordered product of *odd* number of free fields. Alternatively, this can be seen by changing $\varphi(x)$ to $-\varphi(x)$ in the Feynman functional integral.

Furthermore, because $\tilde{Z}[J = 0] = 1$, one gets

$$\frac{\delta^2 W}{\delta J(x_1)\delta J(x_2)}\bigg|_{J=0} = -\frac{i\delta^2 Z}{\delta J(x_1)\delta J(x_2)}\bigg|_{J=0} = i\langle x_1 x_2\rangle \,, \tag{4.371}$$

which shows that W generates the propagator (2-point Green function) to any order in λ. This could be expected, since the propagator has no disconnected parts.

As already suggested, this would not be the case, e.g., for $g\varphi^3$ theory where we have diagrams of the type

(4.372)

The expansion, however, becomes less trivial when we consider the 4-point connected Green function. To this end, we differentiate Eq. (4.370) twice more and set $J = 0$ at the end. This gives

$$\frac{\delta^4 W}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\bigg|_{J=0}$$

$$= \frac{\delta^2}{\delta J(x_4)\delta J(x_3)} \left[\frac{i}{Z^2[J]} \frac{\delta Z[J]}{\delta J(x_1)} \frac{\delta Z[J]}{\delta J(x_2)} - \frac{i}{Z} \frac{\delta^2 Z[J]}{\delta J(x_1)\delta J(x_2)} \right]_{J=0}$$

$$= \frac{\delta}{\delta J(x_4)} \left[-\frac{2i}{Z^3} \frac{\delta Z}{\delta J(x_3)} \frac{\delta Z}{\delta J(x_1)} \frac{\delta Z}{\delta J(x_2)} + \frac{i}{Z^2} \frac{\delta^2 Z}{\delta J(x_3)\delta J(x_1)} \frac{\delta Z}{\delta J(x_2)} \right.$$

$$+ \frac{i}{Z^2} \frac{\delta Z}{\delta J(x_1)} \frac{\delta^2 Z}{\delta J(x_2)\delta J(x_3)} + \frac{i}{Z^2} \frac{\delta Z}{\delta J(x_3)} \frac{\delta^2 Z}{\delta J(x_1)\delta J(x_2)}$$

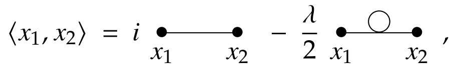
$$\left. -\frac{i}{Z} \frac{\delta^3 Z}{\delta J(x_3)\delta J(x_1)\delta J(x_2)} \right]_{J=0}$$

$$= \left[\frac{i}{Z^2} \frac{\delta^2 Z}{\delta J(x_4)\delta J(x_1)} \frac{\delta^2 Z}{\delta J(x_2)\delta J(x_3)} + \frac{i}{Z^2} \frac{\delta^2 Z}{\delta J(x_4)\delta J(x_3)} \frac{\delta^2 Z}{\delta J(x_1)\delta J(x_2)} \right.$$

$$\left. + \frac{i}{Z^2} \frac{\delta^2 Z}{\delta J(x_3)\delta J(x_1)} \frac{\delta^2 Z}{\delta J(x_4)\delta J(x_2)} - \frac{i}{Z} \frac{\delta^4 Z}{\delta J(x_4)\delta J(x_3)\delta J(x_2)\delta J(x_1)} \right]_{J=0}$$

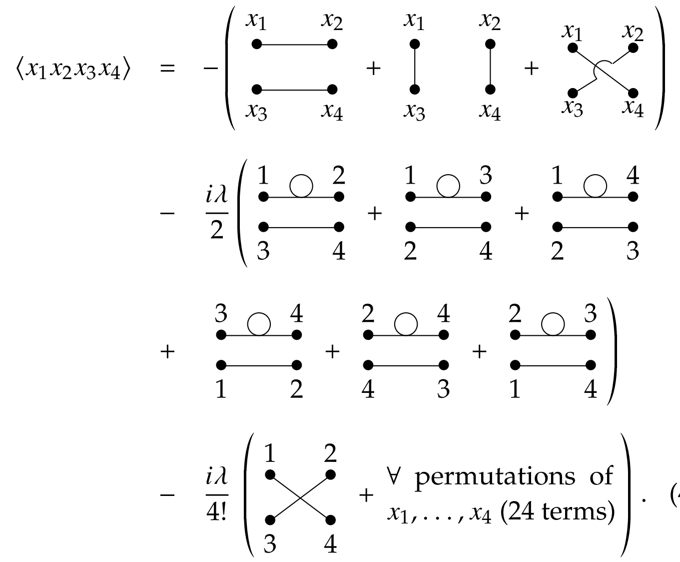
$$= i\langle x_4 x_1\rangle\langle x_2 x_3\rangle + i\langle x_4 x_3\rangle\langle x_1 x_2\rangle + i\langle x_3 x_1\rangle\langle x_4 x_2\rangle - i\langle x_4 x_3 x_2 x_1\rangle \,. \tag{4.373}$$

To see that this expression contains no disconnected diagrams, let us check it to order λ. From Eq. (4.374) we know that



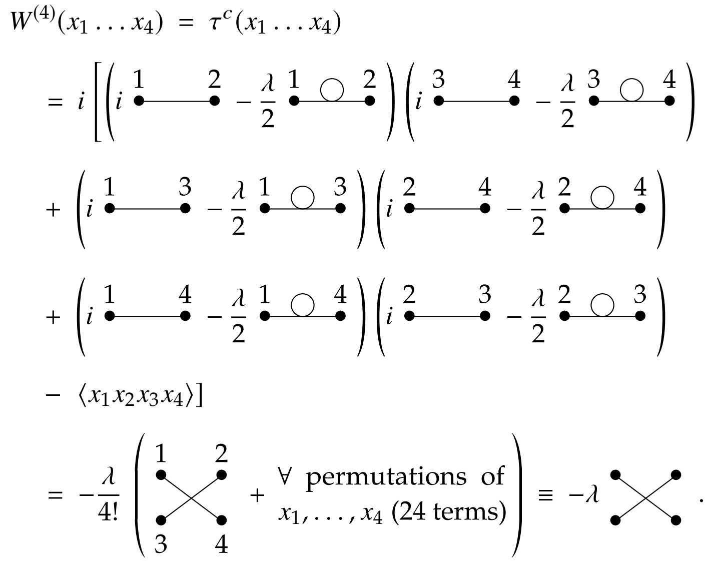
(4.374)

while



(4.375)

Thus,



(4.376)

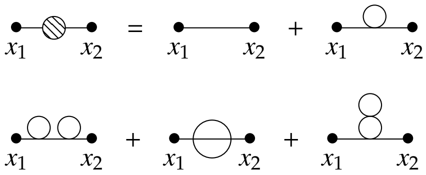
In terms of $\tilde{W}$ we would have the multiplicative coefficient $-i\lambda$ instead of $-\lambda$.

So, the disconnected pieces cancelled, and the only terms which survived are connected pieces, which form a topology of a cross.

We will discuss diagrammatics of the perturbation expansion more in the chapter dedicated to renormalization. Let us now make one more observation that will be relevant at the later stage.

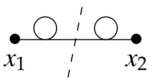
Let us write the 2-point connected Green function up to order λ^2 in the $\lambda\varphi^4$ theory. This reads (symmetry factors omitted)

$\langle x_1 x_2 \rangle^c =$



(4.377)

The organization of the perturbation series is quite straightforward. While $W[J]$ generates connected diagrams, it still contains diagrams that are reducible to two connected diagrams upon cutting an internal line, e.g.



(4.378)

is irreducible upon cutting the internal propagator. Such diagrams are called 1P (one particle) reducible. It is clear that 1P irreducible diagrams are more fundamental building blocks of generic diagrams since we can construct all the connected diagrams from them. We will study the 1PI (one particle irreducible) diagrams in connection with effective action and renormalization. We will see their role also when discussing the Källén–Lehmann spectral representation of 2-point Green functions.

Loop Expansion

The loop (or loopwise) perturbation expansion, i.e., the expansion according to the increasing number of independent loops of connected Green functions, may be identified with an expansion in powers of $\hbar$. To show this, let us reinsert $\hbar$. The best starting point is the functional-integral representation of the generating functional $Z[J]$. In particular, when $\hbar = 1$ we know that

$$
\begin{aligned}
Z[J] &\stackrel{\hbar=1}{=} N \int \mathcal{D}\varphi \exp\left\{ i \int d^4x \, [\mathcal{L}(\varphi, \partial\varphi) + J\varphi] \right\} \\
&\stackrel{\hbar\neq 1}{=} N \int \mathcal{D}\varphi \exp\left\{ \frac{i}{\hbar} S[\varphi] + i \int d^4x \, J\varphi \right\} , \qquad (4.379)
\end{aligned}
$$

where N is normalization constant. On the dimensional ground we had to divide the action $S[\varphi]$ by $\hbar$. Note also that there is no $\hbar$ factor in front of the $\int d^4x \, J\varphi$ term. This is because we require that

$$
\begin{aligned}
\langle \Omega | T[\varphi_H(x_1) \ldots \varphi_H(x_n)] | \Omega \rangle &= \frac{\int \mathcal{D}\varphi \, \varphi(x_1) \ldots \varphi(x_n) \exp\left\{ \frac{i}{\hbar} S[\varphi] \right\}}{\int \mathcal{D}\varphi \exp\left\{ \frac{i}{\hbar} S[\varphi] \right\}} \\
= \frac{(-i)^n \delta^n}{\delta J(x_1) \ldots \delta J(x_n)} & \left. \frac{\int \mathcal{D}\varphi \exp\left\{ \frac{i}{\hbar} S[\varphi] + i \int d^4x \, J\varphi \right\}}{\int \mathcal{D}\varphi \exp\left\{ \frac{i}{\hbar} S[\varphi] \right\}} \right|_{J=0} . \qquad (4.380)
\end{aligned}
$$

Eq. (4.379) implies [see also Eq. (4.271)] that

$$
Z[J] = \exp\left[\frac{i}{\hbar} \int d^4x \mathcal{L}_I \left(-i \frac{\delta}{\delta J(x)} \right) \right] Z_0[J] . \qquad (4.381)
$$

Here we have divided the Lagrangian into free (quadratic) and interaction parts, $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I$,

$$Z_0[J] = \exp\left(-\frac{i}{2}\int d^4x d^4y J(x)\mathbb{A}(x,y)J(y)\right), \tag{4.382}$$

where the operator $\mathbb{A}(x,y)$ deduces from $\mathcal{L}_0 = \int d^4y d^4x \frac{1}{2}\phi_x \mathbb{A}^{-1}(x,y)\phi_y$. Since $\mathbb{A}^{-1}(x,y) = -\frac{i}{\hbar}\delta(x-y)(\Box + m^2)_x$, we immediately get that $\mathbb{A}(x,y) = \hbar\Delta_F(x,y)$, and consequently

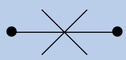
$$Z_0[J] = \exp\left(-\frac{i}{2}\int d^4x d^4y J(x)\hbar\Delta_F(x,y)J(y)\right). \tag{4.383}$$

Let us now count $\hbar$ in a typical Feynman graph. For any Feynman diagram, each propagator comes from $Z_0[J]$ and is multiplied by $\hbar$. However, each vertex, because it appears in the combination $\mathcal{L}_I/\hbar$, is multiplied by a factor $\hbar^{-1}$. So, for an arbitraty Feynman graph, the total $\hbar$ has power $\hbar^{E+I-V}$ (E – external line, I – internal line, V – vertex), which is (by Eulers theorem that states $L = I - (V-1)$) equal to $\hbar^{E-1+L}$.

In particular, for a fixed number of external legs (lines), i.e. for a given Green function each loop contributes with one $\hbar$.

Note

The minimal values of $\hbar$ occures (with fixed E) for $I = 0$ and $V = 1$, e.g.

 , (4.384)

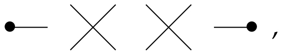
resulting in $\hbar^{E-1}$. Simple propagator does not count as it does not have clear concept external-internal lines.

For vacuum bubble diagrams the minimum is reached with 1 vertex and $I = 2$ (with one it does not provide vacuum diagram). Then $\hbar^{I-V} = \hbar^{2-1} = \hbar$ meaning minimal contributions disappear in the limit $\hbar \to 0$. This implies that vacuum diagrams are entirely of quantum origin.

It is interesting to observe that in theories with a single coupling constant (e.g., $g\frac{\varphi^3}{3!}$, $\lambda\frac{\varphi^4}{4!}$ etc.) the loopwise expansion coincides with expansion according to powers of coupling constants. This is because there exist in these cases auxiliary relations between V, i.e. number of vertices (i.e., power of λ, g, ...) and L. Indeed, e.g. for $\lambda\frac{\varphi^4}{4!}$ we have

$$\begin{aligned} 4V &= E + 2I \\ &= \{\text{Euler form}\} = E + 2(L + (V-1)) \\ &\Rightarrow 2V = E + 2L - 2 \\ &\Rightarrow V = \frac{E}{2} + L - 1 \quad \left(E \text{ is event for } \varphi^4\right). \end{aligned} \tag{4.385}$$

This holds because each vertex has 4 legs, thus the total number of legs from vertices in graph is $4V$, e.g.

 (4.386)

where each I connects 2 legs from $4V$ and $V = 2$, $E = 2$. For fixed E, power of λ is dictated by number of loops.

Similarly for $g\frac{\varphi^3}{3!}$ where one has

$$\begin{aligned} 3V &= E + 2I = E + 2(L + (V-1)) \\ &\Rightarrow V = E + 2L - 2\,. \end{aligned} \tag{4.387}$$

Again for fixed E, the total power of g is equal to $g^V \sim g^{2L} = (g^2)^L$. Hence, expansion in number of L.

Note

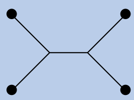
Because our conclusions are valid due to Euler formula (which in turn holds only for planar graphs), they are valid only for connected Feynman diagrams. In particular, they are not valid e.g. for

 (4.388)

diagram, where $E = 2$, $I = 2$, $V = 2$ and $L = 2$, thus $2 = L \neq I - (V-1) = 1$).

Note

All loops disappear in the limit $\hbar \to 0$. Diagrams that do not disappear in this limit are known as *tree* or *Born diagrams*, e.g.

 (4.389)

Beyond simple scalar fields

For a scalar field multiplet we have the generating functional

$$\tilde{Z}[J_1, J_2, \ldots, J_n] = \frac{\int \left[\prod_{r=1}^n \mathcal{D}\phi_r\right] e^{iS[\phi_1,\ldots,\phi_n] + i\int d^4x \sum_r J_r(x)\phi_r(x)}}{\int \left[\prod_{r=1}^n \mathcal{D}\phi_r\right] e^{iS[\phi_1,\ldots,\phi_n]}}\,. \tag{4.390}$$

In particular, for $n = 2$ it is conventional to introduce a complex field

$$\varphi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \quad \text{and} \quad \varphi^* = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)\,, \tag{4.391}$$

which allows us to write

$$\begin{aligned} S[\phi_1, \phi_2] &\rightarrow S[\varphi, \varphi^*] , \\ \tilde{Z}[J, J^*] &= \frac{\int \mathcal{D}\varphi\mathcal{D}\varphi^* e^{iS[\varphi,\varphi^*]+i\int d^4x(J^*\varphi+J\varphi^*)}}{\int \mathcal{D}\varphi\mathcal{D}\varphi^* e^{iS[\varphi,\varphi^*]}} . \end{aligned} \tag{4.392}$$

Objects of interest are again Green functions. In order to do perturbative calculus we need to identify first the corresponding propagators. As before, propagators follow from the term in the action that is quadratic in the fields.

$$\begin{aligned} S[\phi_1, \ldots, \phi_n] &= \sum_{r=1}^{n} \frac{1}{2} \int d^4x \left\{ \phi_r(x) \left[-\left(\Box + m_r^2 \right) \right] \phi_r + \mathcal{L}_I(\phi_1, \ldots, \phi_n) \right\} , \\ S[\varphi, \varphi^*] &= \int d^4x \left\{ \varphi^*(x) \left[-\left(\Box + m^2 \right) \right] \varphi(x) + \mathcal{L}_I(\varphi, \varphi^*) \right\} . \end{aligned} \tag{4.393}$$

Thus,

$$\begin{aligned} Z[J] &= \exp\left[i \int_{R^4} d^4x \mathcal{L}_I \left(-i\frac{\delta}{\delta J_1(x)}, \ldots, -i\frac{\delta}{\delta J_n(x)} \right) \right] \\ &\times \exp\left(-\frac{1}{2} \int d^4x d^4y \sum_{r=1}^{n} J_r(x) i\Delta_F^r(x,y) J_r(x) \right) , \end{aligned} \tag{4.394}$$

where

$$\left[\Delta_F^r(x,z) \right]^{-1} = -\delta(x-y) \left(\Box + m_r^2 \right)_x . \tag{4.395}$$

Note

We recall that $$\begin{aligned} \int \left[\Delta_F^r(x,z) \right]^{-1} \Delta_F(z,y) d^4z &= \delta(x-y) , \\ \Leftrightarrow (\Box + m_r^2)_x \Delta_F(x,y) &= -\delta(x,y) . \end{aligned}$$

So, perturbation expansion of, e.g.,

$$\langle \Omega | T[\phi_{H1}(x_1)\phi_{H2}(x_2) \ldots \phi_{Hn}(x_n)] | \Omega \rangle , \tag{4.396}$$

consist of diagrams constructed from various typologically distinct vertices.

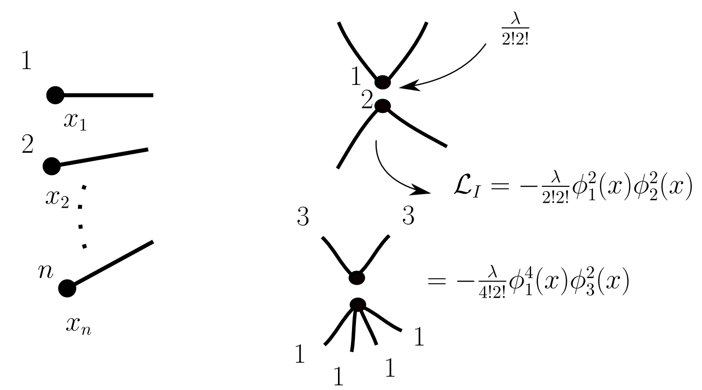


Figure 4.13: n external legs should be joined with vertices prescribed by the interaction Lagrangian $\mathcal{L}_I$. For illustrative purposes we sub-divided the vertex point (should be a single point!) into sub-vertices of the same field type.

Note

There are no propagators of the type $\langle 0| T[\phi_1(x)\phi_2(y)] |0\rangle$ in the Feynman diagrams. In fact these would-be propagators are zero, because

$$\begin{aligned}\langle 0| T[\phi_1(x)\phi_2(y)] |0\rangle \quad &\propto \quad \langle 0| a_1(p)a_2^\dagger(q) |0\rangle \; = \; 0 \\ \text{or} \quad &\langle 0| a_2(p)a_1^\dagger(q) |0\rangle \; = \; 0 \,. \end{aligned} \tag{4.397}$$

Situation with complex fields is a bit more complicated. Let us recall that complex fields are convenient if the theory is invariant under phase transformations. For instance, the Lagrangian

$$\mathcal{L} = |\partial^\mu \varphi|^2 - m^2|\varphi|^2 - \frac{\lambda}{4}|\varphi|^4 \tag{4.398}$$

is invariant under $\varphi \longrightarrow \varphi' = e^{i\alpha}\varphi$ with α being an arbitrary parameter. In these cases

$$\begin{aligned} Z[J,J^*] \;\; &= \;\; \exp\left[i\int_{R^4} d^4x \mathcal{L}_I\left(-i\frac{\delta}{\delta J^*(x)}, -i\frac{\delta}{\delta J(x)}\right)\right] \\ &\times \exp\left(-\int d^4x d^4y J^*(x) i\Delta_F(x,y) J(y)\right) , \end{aligned} \tag{4.399}$$

where $\Delta_F(x, y)$ now corresponds to

$$\begin{aligned} \langle 0| T[\varphi(x)\varphi^*(y)] |0\rangle \;\; &= \;\; \frac{1}{2}\langle 0| T[\phi_1(x)\phi_1(y)] |0\rangle + \frac{1}{2}\langle 0| T[\phi_2(x)\phi_2(y)] |0\rangle \\ &= \;\; i\int \frac{d^4p}{(2\pi)^4}\frac{e^{-ip(x-y)}}{p^2 - m^2 + i\varepsilon} \,. \end{aligned} \tag{4.400}$$

From this follows that

$$\begin{aligned} \langle 0| T[\varphi(x)\varphi^*(y)] |0\rangle \;\; &= \;\; \langle 0| T[\varphi(y)^*\varphi(x)] |0\rangle = \langle 0| T[\varphi(y)\varphi^*(x)] |0\rangle \\ &= \;\; \langle 0| T[\varphi^*(x)\varphi(y)] |0\rangle \,, \end{aligned} \tag{4.401}$$

We note here that

$\langle 0| T[\varphi(x)\varphi(y)] |0\rangle$

$= \langle 0| T[\varphi^*(x)\varphi^*(y)] |0\rangle = 0\,.$

is the same propagator as for real scalar field. Corresponding full Green functions are obtained from $Z[J, J^*]$ as before, e.g.

$$\langle \Omega| T[\varphi_H(x_1)\varphi_H^*(x_2)] |\Omega\rangle \; = \; (-i)^2\frac{\delta^2}{\delta J^*(x_1)\delta J(x_2)}Z[J,J^*]\Big|_{J,J^*=0} \,. \tag{4.402}$$

First, to order λ we have

$$Z[J,J^*] = \left[1 - i\frac{\lambda}{4}\int d^4x(-i)^4\frac{\delta^2}{\delta J^*(x)^2}\frac{\delta^2}{\delta J(x)^2}\right] e^{-\int J^* i\Delta_F J} \,. \tag{4.403}$$

To order of λ^0 we have just the free particle generating functional. To order λ we can write

$$\frac{\delta}{\delta J(z)} \exp\left[-i\int d^4x d^4y J^*(x)\Delta_F(x,y)J(y)\right]$$
$$= -i\int d^4x J^*(x)\Delta_F(x,z)\exp\left[-i\int d^4x d^4y \Delta_F(x,y)J^*(x)J(y)\right]. \quad (4.404)$$

Then the corresponding second variation $\delta^2/\delta J(z)^2$ needed in (4.403) is

$$\frac{\delta}{\delta J(z)}[\text{Eq. (4.404)}]$$
$$= \left[-i\int d^4x J^*(x)\Delta_F(x,z)\right]^2 e^{-i\int d^4x d^4y J^*(x)\Delta_F(x,y)J(y)}. \quad (4.405)$$

Analogously, by taking variation wrt. $J^*(z)$ we get

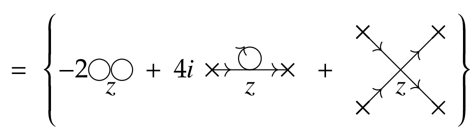
$$\frac{\delta}{\delta J^*(z)}[\text{Eq. (4.405)}]$$
$$= 2\left[-i\int d^4x J^*(x)\Delta_F(x,z)\right][-i\Delta_F(0)]\, e^{-i\int J^*\Delta_F J}$$
$$+ \left[-i\int d^4x J^*(x)\Delta_F(x,z)\right]^2 \left[-i\int d^4y \Delta_F(z,y)J(y)\right] e^{-i\int J^*\Delta_F J}. \quad (4.406)$$

And finally, the last variation wrt. $J^*(z)$ provides

$$\frac{\delta^2}{\delta J^*(z)^2}[\text{Eq. (4.405)}]$$
$$= -2\,[\Delta_F(0)]^2\, e^{-i\int J^*\Delta_F J}$$
$$- 2\int d^4x J^*(x)\Delta_F(x,z)\Delta_F(0)\int(-i)d^4y J(y)\Delta_F(z,y)e^{iJ^*\Delta_F J}$$
$$-2\left[\int d^4x J^*(x)\Delta_F(x,z)\right]\Delta_F(0)(-i)\int d^4y \Delta_F(z,y)J(y)e^{-iJ^*\Delta_F J}$$
$$+\left[-i\int d^4x J^*(x)\Delta_F(x,z)\right]^2\left[-i\int d^4y \Delta_F(z,y)J(y)\right]^2 e^{-iJ^*\Delta_F J}. \quad (4.407)$$

This result can be graphically represented as

$$\left(\frac{\delta}{\delta J^*(z)}\right)^2\left(\frac{\delta}{\delta J(z)}\right)^2 \exp\left(-i\int J^*\Delta_F J\right)$$

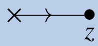


$$e^{-i\int J^*\Delta_F J}. \quad (4.408)$$

Arrow in the propagator indicates that the propagator is oriented in the sense that the endpoints of the propagator lines refer to independent (different) fields φ and φ^*. Combinatorial factors 2 and 4 are simple result of symmetry considerations.

Standard convention

The standard convention is that incoming arrows refer to φ and outgoing ones to φ^*. I.e.



$$
\begin{aligned}
&= \int d^4x J(x) \Delta_F(x,z) \\
&= -i \int d^4x J(x) \langle 0| T[\varphi^*(x)\varphi(z)] |0\rangle \,. \qquad (4.409)
\end{aligned}
$$

As we said, a formulation in terms of complex fields (rather than real ones) is useful if the theory is invariant under phase transformation, i.e. $\varphi \to \varphi' = e^{i\alpha}\varphi$. In that case, every Green function must contain an equal number of φ and φ^* fields (otherwise it is zero) since in $\mathcal{L}_I(\varphi, \varphi^*)$ must be for each field φ also field φ^*. So, each vertex has an equal number of incoming and outgoing lines. In forming diagrams, the lines can be only joined if their orientation arrows match. The orientation often corresponds to the flow of electric charge, obviously, charge will be conserved if the number of incoming and outgoing arrows is the same at each vertex.

4.12 Functional Integral for Fermions

In canonical quantization $[,] \to \{,\}$ for fermions. This will bring various signs modifications into Wick's theorem. As in Bose case, we can derive a generating relation for Wick's theorem that will serve as a basis for corresponding functional integral treatment. The simplest passage to a generator for fermionic Wick's theorem and ensuing Feynman functional integral is via *Grassman variables* and *Berezin calculus*.

4.13 Grassmann variables

Grassmann variables are a set of anticommuting symbols. Name "variable" is really misnomer, as Grassman variables are not really variables. Nevertheless, terminology "Grassmann variables" is standardly used, and so we stick to it also in this lecture. Suppose there are n Grassmann variables. We denote them as θ_i. The only properties we require is that they are linearly independent and that

$$
\theta_i\theta_j \;+\; \theta_j\theta_i \;=\; 0 \quad\Rightarrow\quad \theta_i^2 \;=\; 0\,. \qquad (4.410)
$$

So, θ_i are *nilpotent*. We combine θ_i with a coefficient field (either $\mathbb{R}$ or $\mathbb{C}$) and form the algebra $\mathcal{A}_n$ consisting of all sums of products of θ_i. A typical element of $\mathcal{A}_n$ would have form

$$
p(\theta_1, \theta_2, \ldots, \theta_n) \;=\; p_0 + p_i\theta_i + \frac{1}{2}p_{ij}\theta_i\theta_j + \frac{1}{3!}p_{ijk}\theta_i\theta_j\theta_k + \cdots\,, \qquad (4.411)
$$

The combinatorial factor 1/2!, 1/3!, etc. are only conventional and often are omitted.

where $p_{\dots}$ are elements of the coefficient field. We assume that they are antisymmetrical under exchange of pairs of indices. The expansion in (4.411) clearly terminates at the $(n+1)$-th term due to $\theta_i^2 = 0$.

Elements containing only terms with an even number of θ_i factors commute with all elements of the algebra and are called *even* or *bosonic* elements. Those with odd numbers of θ_i anticommute with one another, and are known as *odd* or *fermionic*. In physics context, we will find ourselves only adding even elements to even elements and odd elements to odd elements, but this is not a mathematical requirement.

One can also mix Grassmann variables with usual variables (say x) within one function. In such as cases, a generic element of the algebra (say algebra $\mathcal{A}_1$) will be of the form

$$p(x,\theta) \;=\; p_0(x) \;+\; p_1(x)\theta\,. \tag{4.412}$$

Integrals over Grassmann variables were introduced by Berezin in 1966.

Motivation: One of the features we would like to incorporate is analogue of the fact, that the integral over space is transitionally invariant, i.e.

$$\int_{\mathbb{R}} dx\,\phi(x) \;=\; \int_{\mathbb{R}} dx\,\phi(x+c)\,. \tag{4.413}$$

We define the "integral" as a linear functional taking the elements of the algebra to elements of the coefficient field and satisfying $(n = 1)$

$$\begin{aligned}\int d\theta p(x,\theta) \;&=\; \int d\theta\,[p_0(x) + p_1(x)\theta] \\ &=\; \int d\theta\,[p_0(x) + p_1(x)(\theta+\alpha)]\,.\end{aligned} \tag{4.414}$$

Let us define $I_0 = \int d\theta$, $I_1 = \int d\theta\theta$, then

$$\begin{aligned}\int d\theta p(x,\theta) \;&=\; \int d\theta p_0(x) + \int d\theta\theta p_1(x) \;=\; I_0 p_0(x) + I_1 p_1(x) \\ &=\; (p_0 + \alpha p_1)I_0 + I_1 p_1\,.\end{aligned} \tag{4.415}$$

Thus we see that $I_0 = 0$ and $I_1 = 1$. This, in turn, provides the following unorthodox definition:

$$\int d\theta \;=\; 0, \int d\theta\theta \;=\; 1\,. \tag{4.416}$$

The choice $I_1 = 1$ is only conventional (could by, in principle, any number). Number 1 is chose so that the integral over Grassmann variables behaves as a derivative

$$\int d\theta \;=\; \frac{d}{d\theta}\,. \tag{4.417}$$

For more variables we can use the prescription (4.411). By again requiring

$$\int d\theta_i p(x,\theta_1,\ldots,\theta_n) = \int d\theta_i p(x,\theta_1,\ldots,\theta_i+\alpha,\ldots,\theta_n), \quad (4.418)$$

we get

$$\int d\theta_i = 0, \int d\theta_i\theta_i = 1, \int d\theta_i\theta_j = 0, \text{ for } i \neq j. \quad (4.419)$$

Again, we might notice that our convention implies

$$\int d\theta_i = \frac{\partial}{\partial\theta_i}. \quad (4.420)$$

For example

$$\begin{aligned}\int d\theta_1 d\theta_2(\theta_1\theta_2) &= \frac{\partial}{\partial\theta_1}\left(\frac{\partial}{\partial\theta_2}\theta_1\theta_2\right)\\ &= \frac{\partial}{\partial\theta_1}\frac{\partial}{\partial\theta_2}(-\theta_2\theta_1) \quad (4.421)\\ &= -\frac{\partial}{\partial\theta_1}\theta_1 = -1. \quad (4.422)\end{aligned}$$

The same result can be achieved if we prescribe that

$$\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j} = -\frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_i}, \quad (4.423)$$

and

$$d\theta_1 d\theta_2 = -d\theta_2 d\theta_1. \quad (4.424)$$

Another example would be

$$\begin{aligned}\int d\theta_1 d\theta_2 f(x;\theta_1,\theta_2) &= \int d\theta_1 d\theta_2\left(a + b_i\theta_i + \frac{1}{2}\varepsilon_{ij}c\theta_i\theta_j\right)\\ &= c\int d\theta_1 d\theta_2\left(\frac{1}{2}\theta_1\theta_2 - \frac{1}{2}\theta_2\theta_1\right)\\ &= -c\int d\theta_2 d\theta_1\theta_2\theta_1 = -c(x). \quad (4.425)\end{aligned}$$

One can also define analogue of Dirac's δ-function. In fact, by analogy with classical calculus we want

$$\begin{aligned}\int d\theta\,\delta(\theta)f(x,\theta) &= f(x,0),\\ \int d\theta\,\delta(\theta)\left[f_0(x)+f_1(x)\theta\right] &= f_0(x). \quad (4.426)\end{aligned}$$

This implies that we can choose $\delta(\theta) = \theta$. Note that for this representation of δ-function holds also other consistency conditions, e.g. $\theta\delta(\theta) = \theta^2 = 0$ (analogue of $x\delta(x) = 0$) or more generally $f(\theta)\delta(\theta) = (f_0 + f_1\theta)\delta(\theta) = f_0\theta = f(0)\delta(\theta)$.

Also, note that the only terms that will survive in the integral $\int d\theta_1 \ldots d\theta_n$

will be the one with n θ's (all other terms will not have enough θ's or too much θ's to survive integration). So, that

$$\begin{aligned}
\int d\theta_n \dots d\theta_1 p(x, \theta_1, \dots, \theta_n) \\
&= \int d\theta_n \dots d\theta_1 \left[p_0 + p_i\theta_i + p_{ij}\theta_i\theta_j + \dots + p_{i_1,\dots,i_n}\theta_{i_1} \cdots \theta_{i_n}\right] \\
&= \varepsilon_{i_1,i_2,\cdots i_n} p_{i_1,i_2,\cdots i_n} \\
&= n! p_{1,2,\cdots n} . \qquad (4.427)
\end{aligned}$$

If we now change variable of integration according to

$$\hat{\theta}_i = a_{ij}\theta_j , \qquad (4.428)$$

then

$$\int d\hat{\theta}_1 \dots d\hat{\theta}_n p(x, \hat{\theta}) = \int d\theta_1 \dots d\theta_n p(x, \hat{\theta})(?) . \qquad (4.429)$$

We can see that

$$\begin{aligned}
p(x, \hat{\theta}(\theta)) &= p_0(x) + \dots + n! p_{1\dots n} a_{1i_1} \dots a_{ni_n} \theta_{1i_1} \dots \theta_{ni_n} \\
&= p_0(x) + \dots + n! p_{1\dots n} \det(a) \theta_1 \cdots \theta_n . \qquad (4.430)
\end{aligned}$$

Consequently, we can write

$$\int d\theta_n \dots d\theta_1 p(x, \hat{\theta}(\theta)) = n! p_{1\dots n} \det(a) . \qquad (4.431)$$

This leads to

$$\int d\hat{\theta}_n \dots d\hat{\theta}_1 p(x, \hat{\theta}) = \int d\theta_1 \dots d\theta_n p(x, \hat{\theta}(\theta)) \left[\det(a)\right]^{-1} . \qquad (4.432)$$

So, we have the following equation for differentials

$$\begin{aligned}
d\hat{\theta}_n \dots d\hat{\theta}_1 &= [\det(a)]^{-1} d\theta_n \dots d\theta_1 \\
&= \det\left[\frac{\partial(\hat{\theta}_1, \dots, \hat{\theta}_n)}{\partial(\theta_1, \dots, \theta_n)}\right]^{-1} d\theta_n \dots d\theta_1 , \qquad (4.433)
\end{aligned}$$

which is different than expected form of Jacobian — it is inverse of the Jacobian.

4.14 Gaussian Integrals over Grassmann Variables

We wish to compute

$$\int d\theta_n \dots d\theta_1 \exp\left(\frac{1}{2}\boldsymbol{\theta}^T \boldsymbol{A} \boldsymbol{\theta}\right) . \qquad (4.434)$$

Here we consider that n is even and that the matrix $\boldsymbol{A}$ is antisymmetric. Specifically for the case $n = 2$ we have

$$\begin{aligned}\int d\theta_2 d\theta_1 e^{\frac{1}{2}\boldsymbol{\theta}^T \boldsymbol{A}\boldsymbol{\theta}} &= \int d\theta_2 d\theta_1 e^{\frac{1}{2}\theta_1 A_{12}\theta_2 + \frac{1}{2}\theta_2 A_{12}\theta_1} \\ &= \int d\theta_2 d\theta_1 e^{A_{12}\theta_1\theta_2} \\ &= \int d\theta_2 d\theta_1 \left[1 + A_{12}\theta_1\theta_2\right] \\ &= A_{12} = \sqrt{\det \boldsymbol{A}} = \operatorname{Pf}\boldsymbol{A}\,. \end{aligned} \tag{4.435}$$

$\sqrt{\det \boldsymbol{A}}$ is well defined for both $A_{12} > 0$ and $A_{21} < 0$. $\operatorname{Pf}\boldsymbol{A}$ is called *Pfaffian*. For any antisymmetric (skew symmetric) matrix we have

$$(\operatorname{Pf}\boldsymbol{A})^2 = \det \boldsymbol{A}\,. \tag{4.436}$$

For general n we first recall that for each real antisymmetric matrix $\boldsymbol{A}$ there exists unitary transformation $\boldsymbol{U}$ such that

$$\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger = \boldsymbol{A}_s\,. \tag{4.437}$$

Here $\boldsymbol{A}_s$ is matrix in a block diagonal *Jacobi form*

$$\boldsymbol{A}_s = \begin{pmatrix} a\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix} & 0 & \dots & 0 \\ \vdots & b\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix} & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix} \tag{4.438}$$

If n is even (our case) $a, b, \dots$ are real and positive definite. Define now matrix

$$T = \begin{pmatrix} \pm a^{-1/2} & \dots & \dots & 0 \\ 0 & \pm a^{-1/2} & \dots & 0 \\ \vdots & & \pm b^{-1/2} & \\ \vdots & & & \ddots \\ 0 & & & \end{pmatrix}. \tag{4.439}$$

Note that $\det\left(\boldsymbol{T}^{-1}\right) = \sqrt{\det \boldsymbol{A}}$ and, in addition

$$\begin{aligned}\boldsymbol{T}\left(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\right)\boldsymbol{T} &= \boldsymbol{T}\boldsymbol{A}_s\boldsymbol{T} \\ &= \tilde{\boldsymbol{A}}_s = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix} & 0 & \dots & 0 \\ \vdots & \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix} & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix}. \end{aligned} \tag{4.440}$$

Let us now introduce new Grassmann variable

$$\tilde{\theta} = \left(\boldsymbol{T}^{-1}U\right)\theta. \tag{4.441}$$

Then

$$\begin{aligned}\exp\left(\frac{1}{2}\boldsymbol{\theta}^T\boldsymbol{A}\boldsymbol{\theta}\right) &= \exp\left(\frac{1}{2}\boldsymbol{\theta}^T\boldsymbol{U}^{\dagger}\boldsymbol{T}^{-1}\tilde{\boldsymbol{A}}\boldsymbol{T}^{-1}\boldsymbol{U}\boldsymbol{\theta}\right) \\ &= \exp\left(\frac{1}{2}\tilde{\boldsymbol{\theta}}^T\tilde{\boldsymbol{A}}_s\tilde{\boldsymbol{\theta}}\right). \end{aligned}\tag{4.442}$$

This implies that

$$\begin{aligned}&\int d\theta_n\cdots d\theta_1 \exp\left(\frac{1}{2}\boldsymbol{\theta}^T\boldsymbol{A}\boldsymbol{\theta}\right) \\ &\quad= \int d\tilde{\theta}_n\cdots d\tilde{\theta}_1 \det\left(\boldsymbol{T}^{-1}\boldsymbol{U}\right)\exp\left(\frac{1}{2}\tilde{\boldsymbol{\theta}}^T\tilde{\boldsymbol{A}}_s\tilde{\boldsymbol{\theta}}\right). \end{aligned}\tag{4.443}$$

Note that

$$\det\left(\boldsymbol{T}^{-1}\boldsymbol{U}\right) = \det\left(\boldsymbol{T}^{-1}\right)\det\boldsymbol{U} = \sqrt{\det\boldsymbol{A}} = \mathrm{Pf}\,\boldsymbol{A}. \tag{4.444}$$

Because

$$\begin{aligned}&\int d\tilde{\theta}_1 d\tilde{\theta}_2 \exp\left[\frac{1}{2}\begin{pmatrix}\tilde{\theta}_1 & \tilde{\theta}_2\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\tilde{\theta}_1\\ \tilde{\theta}_2\end{pmatrix}\right] \\ &\quad= \int d\tilde{\theta}_2 d\tilde{\theta}_1\left(1+\tilde{\theta}_1\tilde{\theta}_2\right) = 1. \end{aligned}\tag{4.445}$$

Generally

$$\begin{aligned}&\int d\tilde{\theta}_n\cdots d\tilde{\theta}_1 \exp\left(\frac{1}{2}\tilde{\boldsymbol{\theta}}^T\tilde{\boldsymbol{A}}_s\tilde{\boldsymbol{\theta}}\right) \\ &= \int d\tilde{\theta}_n\cdots d\tilde{\theta}_1 \exp\left(\tilde{\theta}_1\tilde{\theta}_2+\tilde{\theta}_3\tilde{\theta}_4+\cdots+\tilde{\theta}_{n-1}\tilde{\theta}_n\right) \\ &= \int d\tilde{\theta}_n\cdots d\tilde{\theta}_1 \exp\left(\tilde{\theta}_1\tilde{\theta}_2\right)\exp\left(\tilde{\theta}_3\tilde{\theta}_4\right)\cdots\exp\left(\tilde{\theta}_{n-1}\tilde{\theta}_n\right) \\ &= 1. \end{aligned}\tag{4.446}$$

So, when we finally collect all our results we get

$$\int d\theta_n\ldots d\theta_1 \exp\left(\boldsymbol{\theta}^T\boldsymbol{A}\boldsymbol{\theta}\right) = \sqrt{\det\boldsymbol{A}} = \mathrm{Pf}\,\boldsymbol{A}. \tag{4.447}$$

To be able to treat Dirac (charged) fermions we double the number of generators in the algebra and define an involution that takes an element θ_i to an associated element θ_i^*, inverts the orders of products, and takes the complex conjugation of coefficients. The term "involution" means that if we perform the mapping twice, we get back the original element, i.e. $(\theta_i^*)^* = \theta_i$. Despite the similarity of this procedure to the operation of Hermitian conjugation, the variable θ_i^* should be regarded as being an object quite independent of θ_i. This means that $\{\theta_i\}$ and $\{\theta_i^*\}$ are distinct sets of Grassmann variables.

Involution θ_i^* is often also denoted as $\bar{\theta}_i$.

Following the rules above, we have

$$\int d\theta d\theta^* e^{\theta^* a\theta} \;=\; \int d\theta d\theta^* \,(1 + \theta^* a\theta) \;=\; a\,. \qquad (4.448)$$

The exponential series terminated after the second term because $\theta^2 = (\theta^*)^2 = 0$.

Gaussian Integrals with Complex Grassmann Variables

Let us use the notation

$$[d\theta]\,[d\theta^*] \;=\; \prod_{i=1}^{N} d\theta_i d\theta_i^*\,. \qquad (4.449)$$

We wish to compute

$$\begin{aligned}
&\int [d\theta]\,[d\theta^*]\, e^{\theta_i^* A_{ij}\theta_j}\\
&= \int [d\theta]\,[d\theta^*] \left[\frac{1}{N!}(\theta_i^* A_{ij}\theta_j)^N\right]\\
&= \int [d\theta]\,[d\theta^*] \left[\frac{1}{N!} N!\theta_1^* A_{1i_1}\theta_{i_1}\theta_2^* A_{2i_2}\theta_{i_2}\cdots\theta_N^* A_{Ni_N}\theta_{i_N}\right]\\
&= \int [d\theta]\,[d\theta^*]\, A_{1i_1}A_{2i_2}\cdots A_{Ni_N}\theta_1^*\theta_{i_1}\theta_2^*\theta_{i_2}\cdots\theta_N^*\theta_{i_N}\\
&= \int [d\theta]\,[d\theta^*]\, A_{1i_1}A_{2i_2}\cdots A_{Ni_N}\theta_1^*\theta_2^*\cdots\theta_N^*\theta_{i_1}\cdots\theta_{i_N}(-1)^{1+3+5+\cdots}\\
&= \int [d\theta]\,[d\theta^*]\, A_{1i_1}A_{2i_2}\cdots A_{Ni_N}\theta_1^*\theta_2^*\cdots\theta_N^*\theta_1\cdots\theta_N\varepsilon_{i_1,\cdots i_N}(-1)^{1+3+5+\cdots}\\
&= \int [d\theta]\,[d\theta^*]\, A_{1i_1}A_{2i_2}\cdots A_{Ni_N}\varepsilon_{i_1,\cdots i_N}\theta_1^*\theta_1\theta_2^*\theta_2\cdots\theta_N^*\theta_N\\
&= \underbrace{\int \left(\prod_{i=1}^{N} d\theta_i d\theta_i^*\theta_i^*\theta_i\right)}_{=1}\underbrace{A_{1i_1}A_{2i_2}\cdots A_{Ni_N}\varepsilon_{i_1,\cdots i_N}}_{\det \boldsymbol{A}}\\
&= \det \boldsymbol{A}\,. \qquad (4.450)
\end{aligned}$$

An essential point is that the determinant appears in the numerator (!!), rather than in the denominator (as one could naively expect).

To complete the analogy with Gaussian/Fresnelian integration, we should define and evaluate integrals of the form

η_i here is an analogue of Schwinger's source.

$$\int [d\theta]\, e^{\frac{1}{2}\theta_i A_{ij}\theta_j + \eta_i\theta_i}\,. \qquad (4.451)$$

To do this, we must embed the original Grassmann algebra in a larger one, where the vectors $\boldsymbol{\eta}$ form a set of elements that anticommute with each other and with θ_i. They serve as "constants" that will not be integrated over (i.e., no integral of type $\int d\eta_i$, but $\int d\theta_i = 0 \forall i, j$). To evaluate the integral we must complete the square in the exponent and shift the variable of integration. Despite of no "domain of integration" in Berezin integration we know that

$$\int d\theta\theta \;=\; 1 \;=\; \int d\theta(\theta+\eta)\,. \tag{4.452}$$

So, the integral is by construction invariant under shifts. The same holds true for the general shift $\theta_i \to \theta_i + \eta_i$. By using the fact that the inverse of an antisymmetric matrix is also antisymmetric matrix one can write

$$\begin{aligned}
&\int [d\theta]\, e^{\frac{1}{2}\theta_i A_{ij}\theta_j + \eta_i\theta_i} \\
&\quad = \int [d\theta]\, e^{\frac{1}{2}(\theta_i - A^{-1}_{ik}\eta_k)A_{ij}\left(\theta_j - A^{-1}_{jl}\eta_l\right)+\frac{1}{2}\eta_k A^{-1}_{kl}\eta_l} \\
&\quad = \int [d\theta]\, e^{\frac{1}{2}\theta_i A_{ij}\theta_j}\, e^{\frac{1}{2}\eta_k A^{-1}_{kl}\eta_l} \\
&\quad = (\mathrm{Pf}\,\boldsymbol{A})\, e^{\frac{1}{2}\eta_k A^{-1}_{kl}\eta_l} \;=\; \sqrt{\det \boldsymbol{A}}\, e^{\frac{1}{2}\eta_k A^{-1}_{kl}\eta_l}\,.
\end{aligned} \tag{4.453}$$

This is exactly the same result we obtained for real Gaussian integral — except that the determinant is now in the numerator rather than denominator.

As an exercise, prove that

$$\int [d\theta]\,[d\theta^*]\, e^{\theta_i^* A_{ij}\theta_j + \theta_j^*\eta_j + \eta_j^*\theta_j} \;=\; (\det \boldsymbol{A})\, e^{-\eta_i^* A^{-1}_{ij}\eta_j}\,. \tag{4.454}$$

Measure definitions

Sometimes the measure (4.449) is defined differently, e.g.

$$\int [d\theta]\,[d\theta^*] = \int \prod d\theta_i \prod d\theta_i^* = \int d\theta_1 \ldots d\theta_N d\theta_1^* \ldots d\theta_N^*.$$

This brings about an extra sign $(-1)^{n(n-1)/2}$ in comparison with our definition of the measure given by Eq. (4.449).

We agian stress that $*$ denotes here involution and not complex conjugation.

Now, we make transition from discrete-index Grasmann variables θ_i and θ_i^* ($\equiv \bar{\theta}_i$) to two sets of continuous-index Grassmann variables $\psi_\alpha(x)$ and $\bar{\psi}_\beta(x) \equiv \psi_\beta^*(x)$. We further introduce two Grassmann sources $\eta_\alpha(x)$ and $\bar{\eta}_\beta(x)$ and define the generating functional

$$Z[\eta,\bar{\eta}] \;=\; N\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, e^{i\int d^4x[\mathcal{L}_0(\psi,\bar{\psi})+\bar{\eta}\psi+\bar{\psi}\eta]}\,, \tag{4.455}$$

where

$$\mathcal{L}_0 \;=\; \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi\,, \tag{4.456}$$

and

$$\mathcal{D}\psi\mathcal{D}\bar{\psi} \;=\; \lim_{N\to\infty}\left(\prod_{i=1}^{N}\prod_{\alpha=0}^{3} d\psi_\alpha(x_i)d\bar{\psi}_\alpha(x_i)\right). \tag{4.457}$$

Let us now compute $Z[\eta, \bar{\eta}]$ by using the same analysis as for discrete Grassmann variables θ_i and $\bar{\theta}_j$, i.e.

$$\begin{aligned} Z[\eta,\bar{\eta}] &= N \int \mathcal{D}\psi \mathcal{D}\bar{\psi}\, e^{i\int d^4x d^4y \left[\bar{\psi}(x)\delta(x-y)(i\gamma^\mu\partial_\mu - m)\psi(y)\right] + i\int d^4x (\bar{\eta}\psi + \bar{\psi}\eta)} \\ &= N \int \mathcal{D}\psi \mathcal{D}\bar{\psi}\, e^{i\int d^4x d^4y \left(\bar{\psi}(x) + \bar{\eta}(z)A^{-1}(z,x)\right)A(x,y)\left(\psi(y)+A^{-1}(y,z)\eta(z)\right)} \\ &\quad \times e^{-i\int d^4x d^4y \left[\bar{\eta}(x)A^{-1}(x,y)\eta(y)\right]} . \end{aligned} \tag{4.458}$$

In expressions $\bar{\eta}(z)A^{-1}(z,x)$ and $A^{-1}(y,z)\eta(z)$ the integration over z is tacitly assumed.

Recalling that $(i\partial\!\!\!/ - m)S_F(x,y) = \delta(x-y)$ we get (cf. Eq. (4.454))

$$Z[\eta,\bar{\eta}] = \underbrace{\tilde{N}\,\det\left(i\partial\!\!\!/ - m\right)}_{Z[0,0]} e^{-i\int d^4x d^4y \bar{\eta}(x)S_F(x,y)\eta(y)} . \tag{4.459}$$

To obtain a free particle Green function we should compute

$$\begin{aligned} &\frac{\partial}{\partial\bar{\eta}(x)}\frac{\partial}{\partial\eta(y)}\frac{Z[\eta,\bar{\eta}]}{Z[0,0]}\bigg|_{\eta=\bar{\eta}=0} \\ &= \frac{\partial}{\partial\bar{\eta}(x)}\frac{\partial}{\partial\eta(y)}\left(-i\int d^4x_1 d^4x_2 \bar{\eta}(x_1)S_F(x_1,x_2)\eta(x_2)\right)\bigg|_{\eta=\bar{\eta}=0} \\ &= \frac{\partial}{\partial\bar{\eta}(x)}\frac{\partial}{\partial\eta(y)}\left(-i\int d^4x_1 d^4x_2 \eta(x_2)S_F(x_2,x_1)\bar{\eta}(x_1)\right)\bigg|_{\eta=\bar{\eta}=0} \\ &= -iS_F(y,x) \;=\; iS_F(x,y) . \end{aligned} \tag{4.460}$$

Antisymmetry of Fermion Propagator

From definitions of time-ordered products:

$$\begin{aligned} T\left[\psi_\alpha(x)\bar{\psi}_\beta(y)\right] &= \theta(t_x - t_y)\psi_\alpha(x)\bar{\psi}_\beta(y) - \theta(t_y - t_x)\bar{\psi}_\beta(y)\psi_\alpha(x) , \\ T\left[\bar{\psi}_\beta(y)\psi_\alpha(x)\right] &= \theta(t_y - t_x)\bar{\psi}_\beta(y)\psi_\alpha(x) - \theta(t_x - t_y)\psi_\alpha(x)\bar{\psi}_\beta(y) . \end{aligned}$$

This implies that

$$T\left[\psi_\alpha(x)\bar{\psi}_\beta(y)\right] = -T\left[\bar{\psi}_\beta(y)\psi_\alpha(x)\right] , \tag{4.461}$$

and hence for propagator we get that

$$\{S_F(x,y)\}_{\alpha\beta} = -\{S_F(y,x)\}_{\beta\alpha} . \tag{4.462}$$

4.15 Wick Theorem for Dirac Fermions

In order to formulate Wick's theorem for Fermion fields we introduce anticommuting sources $\bar{\eta}$ and η for ψ and $\bar{\psi}$. These sources anticommute among themselves as well as with ψ and $\bar{\psi}$ (so that $\bar{\eta}\psi + \bar{\psi}\eta$ are bosonic quantities that can enter in action). With the help of η and $\bar{\eta}$ we can write

Note that the sources are Grassmann variables while ψ and $\bar{\psi}$ are operators here.

$$
\begin{aligned}
&[\bar{\eta}(x)\psi(x), \bar{\psi}(y)\eta(y)] \\
&\quad = \bar{\psi}(y)\,[\bar{\eta}(x)\psi(x), \eta(y)] \;+\; [\bar{\eta}(x)\psi(x), \bar{\psi}(y)]\,\eta(y) \\
&\quad = \bar{\psi}(y)\bar{\eta}(x)\,\underbrace{\{\psi(x), \eta(y)\}}_{=0} \;-\; \bar{\psi}(y)\,\underbrace{\{\bar{\eta}(x), \eta(y)\}}_{=0}\,\psi(x) \\
&\quad + \bar{\eta}(x)\,\{\psi(x), \bar{\psi}(y)\}\,\eta(y) \;-\; \underbrace{\{\bar{\eta}(x), \bar{\psi}(y)\}}_{=0}\,\psi(x)\eta(y)\,. \qquad (4.463)
\end{aligned}
$$

So, $[\bar{\eta}(x)\psi(x), \bar{\psi}(y)\eta(y)] = \bar{\eta}(x)\,\{\psi(x), \bar{\psi}(y)\}\,\eta(y)$. We now introduce the source Lagrangian

$$
\mathcal{L}_S(x) \;=\; \bar{\eta}(x)\psi(x) \;+\; \bar{\psi}(x)\eta(x)\,. \qquad (4.464)
$$

Note that $[\mathcal{L}_S(x), \mathcal{L}_S(y)]$ is a c-number and hence it commutes with $\mathcal{L}_S(z)$, indeed

$$
\begin{aligned}
[\mathcal{L}_S(x), \mathcal{L}_S(y)] &= [\bar{\eta}(x)\psi(x), \bar{\eta}(y)\psi(y)] + [\bar{\eta}(x)\psi(x), \bar{\psi}(y)\eta(y)] \\
&\quad + [\bar{\psi}(x)\eta(x), \bar{\eta}(y)\psi(y)] + [\bar{\psi}(x)\eta(x), \bar{\psi}(y)\eta(y)] \\
&= -\bar{\eta}(x)\,\underbrace{\{\psi(x), \psi(y)\}}_{=0}\,\bar{\eta}(y) + \bar{\eta}(x)\,\{\psi(x), \bar{\psi}(y)\}\,\eta(y) \\
&\quad + \eta(x)\,\{\bar{\psi}(x), \psi(y)\}\,\bar{\eta}(y) - \eta(x)\,\underbrace{\{\bar{\psi}(x), \bar{\psi}(y)\}}_{=0}\,\eta(y) \\
&= c - \text{number}\,. \qquad (4.465)
\end{aligned}
$$

Thus indeed $[\mathcal{L}_S(z), [\mathcal{L}_S(x), \mathcal{L}_S(y)]] = 0$.

To prove Wick's theorem for fermion field we will follow the same strategy we employed when dealing with scalar field. In particular, we will show that

$$
\begin{aligned}
&T\left[\exp\left(i\int d^4x\,[\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)]\right)\right] \\
&\quad = :\exp\left(i\int d^4x\,[\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)]\right): \\
&\quad \times\, \exp\left(-\int d^4x d^4y\;\bar{\eta}(x)\,\langle 0|\,T[\psi(x)\bar{\psi}(y)]\,|0\rangle\,\eta(y)\right). \qquad (4.466)
\end{aligned}
$$

Since $[\mathcal{L}_S(z), [\mathcal{L}_S(x), \mathcal{L}_S(y)]] = 0$, we can use our strategy from Chapter 4.6 and substitute instead of $-J(x)\phi(x)$ the source term $\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)$. By employing the Baker–Campbell–Hausdorff formula we can write

$$
\begin{aligned}
&T\left[\exp\left(i\int d^4x\,[\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)]\right)\right] \;=\; e^{i\int d^4x\,\mathcal{L}_S(x)} \\
&\times\, \exp\left(-\frac{1}{2}\int d^4x d^4y\,[\bar{\psi}(x)\eta(x) + \bar{\eta}(x)\psi(x), \bar{\psi}(y)\eta(y) + \bar{\eta}(y)\psi(y)]\right).
\end{aligned}
$$

In the second step we split ψ and $\bar{\psi}$ to positive and negative frequency parts and write

$$e^{i\int d^4x \mathcal{L}_S(x)} = e^{i\int d^4x \left[(\bar{\eta}(x)\psi^{(-)}(x) + \bar{\psi}^{(-)}(x)\eta(x)) + (-)\to(+)\right]} . \qquad (4.467)$$

With the help of the BCH formula $e^{A+B+\frac{1}{2}[A,B]} = e^A e^B \Leftrightarrow e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}$ we can rewrite (4.467) as

$$\begin{aligned} e^{i\int d^4x \mathcal{L}_S(x)} & \\ &= e^{i\int d^4x \left[\bar{\eta}(x)\psi^{(-)}(x) + \bar{\psi}^{(-)}(x)\eta(x)\right]} \\ &\times e^{i\int d^4x \left[\bar{\eta}(x)\psi^{(+)}(x) + \bar{\psi}^{(+)}(x)\eta(x)\right]} \\ &\times e^{\frac{1}{2}\iint d^4x d^4y \left[\bar{\eta}(x)\psi^{(-)}(x) + \bar{\psi}^{(-)}(x)\eta(x),\, x\to y \text{ and } (-)\to(+)\right]} . \end{aligned} \qquad (4.468)$$

Plugging this results back we obtain

$$T\left[\exp\left(i\int d^4x \mathcal{L}_S(x)\right)\right] = :\exp\left(i\int d^4x \mathcal{L}_S(x)\right): e^A, \qquad (4.469)$$

where

$$A = \frac{1}{2}\iint d^4x d^4y \left\{\left[\mathcal{L}_S^{(-)}(x), \mathcal{L}_S^{(+)}(y)\right] - \theta(x_0-y_0)\left[\mathcal{L}_S(x), \mathcal{L}_S(y)\right]\right\} .$$

Since the integrand is a c-number, it can be evaluated via its vacuum expectation value, i.e.

$$\begin{aligned} \langle 0|&\left[\mathcal{L}_S^{(-)}(x), \mathcal{L}_S^{(+)}(y)\right] - \theta(x_0-y_0)\left[\mathcal{L}_S(x), \mathcal{L}_S(y)\right]|0\rangle \\ &= -\langle 0|\mathcal{L}_S^{(+)}(x)\mathcal{L}_S^{(-)}(y)|0\rangle - \theta(x_0-y_0)\langle 0|\mathcal{L}_S(x)\mathcal{L}_S(y)|0\rangle \\ &\quad + \theta(x_0-y_0)\langle 0|\mathcal{L}_S(y)\mathcal{L}_S(x)|0\rangle \\ &= \underbrace{-\mathbf{1}}_{\theta(x_0-y_0)\,+\,\theta(y_0-x_0)} \cdot \langle 0|\mathcal{L}_S(y)\mathcal{L}_S(x)|0\rangle - \theta(x_0-y_0)\langle 0|\mathcal{L}_S(x)\mathcal{L}_S(y)|0\rangle \\ &\quad + \theta(x_0-y_0)\langle 0|\mathcal{L}_S(y)\mathcal{L}_S(x)|0\rangle \\ &= -\theta(y_0-x_0)\langle 0|\mathcal{L}_S(y)\mathcal{L}_S(x)|0\rangle - \theta(x_0-y_0)\langle 0|\mathcal{L}_S(x)\mathcal{L}_S(y)|0\rangle \\ &= -\langle 0|T[\mathcal{L}_S(x)\mathcal{L}_S(y)]|0\rangle . \end{aligned} \qquad (4.470)$$

Note that terms of the type $\langle 0|T[\bar{\psi}(x)\eta(x)\bar{\psi}(y)\eta(y)]|0\rangle = 0$, since $\bar{\psi}$ contain only $a^\dagger$ and b and there is no way how the product $\bar{\psi}\bar{\psi}$ could survive vacuum expectation value. So, the only surviving parts are

$$\begin{aligned} \langle 0|T[\mathcal{L}_S(x)\mathcal{L}_S(y)]|0\rangle &= \langle 0|T[\bar{\psi}(x)\eta(x)\bar{\eta}(y)\psi(y)]|0\rangle \\ &\quad + \langle 0|T[\bar{\eta}(x)\psi(x)\bar{\psi}(y)\eta(y)]|0\rangle \\ &= \eta(x)\langle 0|T[\bar{\psi}(x)\psi(y)]|0\rangle\,\bar{\eta}(y) \\ &\quad + \bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle\,\eta(y) . \end{aligned} \qquad (4.471)$$

By noting that $T\left[\psi_\alpha(x)\bar{\psi}_\beta(y)\right] = -T\left[\bar{\psi}_\beta(y)\psi_\alpha(x)\right]$ we have that

$$\begin{aligned}\langle 0|\,T[\mathcal{L}_S(x)\mathcal{L}_S(y)]\,|0\rangle &= \bar{\eta}(x)\,\langle 0|\,T[\psi(x)\bar{\psi}(y)]\,|0\rangle\,\eta(y)\\ &\quad + \bar{\eta}(y)\,\langle 0|\,T[\psi(y)\bar{\psi}(x)]\,|0\rangle\,\eta(x)\,. \end{aligned}\quad (4.472)$$

This consequently implies that

$$A = -\int d^4x d^4y\ \bar{\eta}(x)\underbrace{\langle 0|\,T[\psi(x)\bar{\psi}(y)]\,|0\rangle}_{=iS_F(x,y)}\eta(y)\,. \quad (4.473)$$

If we now take the vacuum expectation value of the Wick theorem generating identity (4.466), we obtain

$$\begin{aligned}&\langle 0|\,T\left[\exp\left(i\int d^4x\left[\bar{\eta}(x)\psi(x)+\bar{\psi}(x)\eta(x)\right]\right)\right]|0\rangle\\ &\quad = \exp\left(-\iint d^4xd^4y\bar{\eta}(x)\,\langle 0|\,T[\psi(x)\bar{\psi}(y)]\,|0\rangle\,\eta(y)\right). \end{aligned}\quad (4.474)$$

This again allows to obtain relation between $\langle 0|\,T[\psi\psi\ldots\psi\bar{\psi}\bar{\psi}\ldots\bar{\psi}]\,|0\rangle$ and two point Green functions $\langle 0|\,T[\psi\bar{\psi}]\,|0\rangle$. In particular, we see that the n-point (free field) Green function must have equal number of ψ's and $\bar{\psi}$'s, so that n must be even.

Working with the fermionic Wick's theorem is analogous to the situation with scalar fields. Let us, for instance, twice variationally differentiate the LHS of (4.474). This yields

$$\begin{aligned}&\left(-i\frac{\delta}{\delta\bar{\eta}(x)}\right)\left(i\frac{\delta}{\delta\eta(y)}\right)\langle 0|\,T\left[e^{i\int d^4z[\bar{\eta}(z)\psi(z)+\bar{\psi}(z)\eta(z)]}\right]|0\rangle\Big|_{\eta,\bar{\eta}=0}\\ &\quad = \frac{\delta}{\delta\bar{\eta}(x)}\,\langle 0|\,T\left[-i\bar{\psi}(y)e^{i\int d^4z[\bar{\eta}(z)\psi(z)+\bar{\psi}(z)\eta(z)]}\right]|0\rangle\Big|_{\eta,\bar{\eta}=0}\\ &\quad = \langle 0|\,T\left[(-i\bar{\psi}(y))(-i\psi(x))\right]|0\rangle = -\langle 0|\,T\left[\bar{\psi}(y)\psi(x)\right]|0\rangle\\ &\quad = \langle 0|\,T\left[\psi(x)\bar{\psi}(y)\right]|0\rangle\ , \end{aligned}\quad (4.475)$$

When the same differentiation is performed on the RHS of (4.474) we get

$$\begin{aligned}&\left(-i\frac{\delta}{\delta\bar{\eta}(x)}\right)\left(i\frac{\delta}{\delta\eta(y)}\right)e^{-\int d^4z_1d^4z_2\ \bar{\eta}(z_1)\langle 0|T[\psi(z_1)\bar{\psi}(z_2)]|0\rangle\eta(z_2)}\Big|_{\eta,\bar{\eta}=0}\\ &\quad = \frac{\delta}{\delta\bar{\eta}(x)}\frac{\delta}{\delta\eta(y)}\left[-\int d^4z_1d^4z_2\ \bar{\eta}(z_1)\,\langle 0|\,T[\psi(z_1)\bar{\psi}(z_2)]\,|0\rangle\,\eta(z_2)\right]\\ &\quad = \frac{\delta}{\delta\bar{\eta}(x)}\int d^4z_1\bar{\eta}(z_1)\,\langle 0|\,T[\psi(z_1)\bar{\psi}(y)]\,|0\rangle\\ &\quad = \langle 0|\,T[\psi(x)\bar{\psi}(y)]\,|0\rangle\ . \end{aligned}\quad (4.476)$$

Let us also notice that successive differentiation over the source fields on the LHS gives

$$\prod_{i=1}^{n}\left(-i\frac{\delta}{\delta\bar{\eta}(x_i)}\right)\prod_{j=1}^{n}\left(i\frac{\delta}{\delta\eta(y_j)}\right)\langle 0|\,T\left[e^{i\int d^4z[\bar{\eta}\psi+\bar{\psi}\eta]}\right]|0\rangle\Big|_{\eta,\bar{\eta}=0}$$

$$= \langle 0|\,T\left[\psi(x_1)\ldots\psi(x_n)\bar{\psi}(y_1)\ldots\bar{\psi}(y_n)\right]|0\rangle\ . \qquad (4.477)$$

As an exercise, we now show that

$$\begin{aligned}&\langle 0|\,T\left[\psi(x_1)\psi(x_2)\bar{\psi}(x_3)\bar{\psi}(x_4)\right]|0\rangle\\ &\quad= -\ \langle 0|\,T\left[\psi(x_1)\bar{\psi}(x_3)\right]|0\rangle\,\langle 0|\,T\left[\psi(x_2)\bar{\psi}(x_4)\right]|0\rangle\\ &\qquad+\ \langle 0|\,T\left[\psi(x_1)\bar{\psi}(x_4)\right]|0\rangle\,\langle 0|\,T\left[\psi(x_2)\bar{\psi}(x_3)\right]|0\rangle\ . \end{aligned} \qquad (4.478)$$

To see this, we 4 times functionally differentiate the RHS of (4.474), namely

$$\begin{aligned}&\left[\frac{(-i)\delta}{\delta\bar{\eta}(x_1)}\frac{(-i)\delta}{\delta\bar{\eta}(x_2)}\right]\left[\frac{i\delta}{\delta\eta(x_3)}\frac{i\delta}{\delta\eta(x_4)}\right]e^{-\int d^4x d^4y \bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle\eta(y)}\Big|_{\eta,\bar{\eta}=0}\\
&= \frac{\delta}{\delta\bar{\eta}(x_1)}\frac{\delta}{\delta\bar{\eta}(x_2)}\frac{\delta}{\delta\eta(x_3)}\int d^4x\,\bar{\eta}(x)\,\langle 0|\,T\left[\psi(x)\bar{\psi}(x_4)\right]|0\rangle\\
&\quad\times e^{-\int d^4x d^4y\,\bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle\eta(y)}\Big|_{\eta,\bar{\eta}=0}\\
&= \frac{\delta}{\delta\bar{\eta}(x_1)}\frac{\delta}{\delta\bar{\eta}(x_2)}\left[-\int d^4x\,\bar{\eta}(x)\,\langle 0|\,T\left[\psi(x)\bar{\psi}(x_4)\right]|0\rangle\right.\\
&\quad\times\int d^4x\,\bar{\eta}(x)\,\langle 0|\,T\left[\psi(x)\bar{\psi}(x_3)\right]|0\rangle\\
&\quad\left.\times\ e^{-\int d^4x d^4y \bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle\eta(y)}\right]\Big|_{\eta,\bar{\eta}=0}\\
&= \frac{\delta}{\delta\bar{\eta}(x_1)}\left[-\langle 0|\,T\left[\psi(x_2)\bar{\psi}(x_4)\right]|0\rangle\int d^4x\,\bar{\eta}(x)\,\langle 0|\,T\left[\psi(x)\bar{\psi}(x_3)\right]|0\rangle\right.\\
&\quad\times e^{-\int d^4x d^4y \bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle\eta(y)}\\
&\quad+\int d^4x\,\bar{\eta}(x)\,\langle 0|\,T\left[\psi(x)\bar{\psi}(x_4)\right]|0\rangle\,\langle 0|\,T\left[\psi(x_2)\bar{\psi}(x_3)\right]|0\rangle\\
&\quad\left.\times e^{-\int d^4x d^4y \bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle\eta(y)}+\cdots\right]\Big|_{\eta,\bar{\eta}=0}\\
&= -\ \langle 0|\,T\left[\psi(x_1)\bar{\psi}(x_3)\right]|0\rangle\,\langle 0|\,T\left[\psi(x_2)\bar{\psi}(x_4)\right]|0\rangle\\
&\quad+\ \langle 0|\,T\left[\psi(x_1)\bar{\psi}(x_4)\right]|0\rangle\,\langle 0|\,T\left[\psi(x_2)\bar{\psi}(x_3)\right]|0\rangle\ . \end{aligned} \qquad (4.479)$$

This coincides with the assertion (4.478). The minus sign is clearly associated with the odd permutation 1234 → 1324 while the plus sign with the even permutation 1234 → 1423. Analogous statement holds also for higher-order Green functions.

More generally, one can write

$$\langle 0|\,T[[F[\psi,\bar{\psi}]\,|0\rangle = F\left[-i\frac{\delta}{\delta\bar{\eta}}, i\frac{\delta}{\delta\eta}\right] e^{-\int d^4x d^4y \bar{\eta}(x)\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle \eta(y)}\Big|_{\eta,\bar{\eta}=0}, \quad (4.480)$$

where F is a function (or functional) of Dirac field operators.

In order to compute the vacuum expectation value of the time ordered product of Dirac fields in Heisenberg picture, we can follow the same strategy as for scalar fields. It is not difficult to see we obtain

$$\langle\Omega|\,T[\psi_H(x_1)\ldots\psi_H(x_n)\bar{\psi}_H(y_1)\ldots\bar{\psi}_H(y_n)]\,|\Omega\rangle = \frac{\langle 0|\,T[\psi(x_1)\ldots\psi(x_n)\bar{\psi}(y_1)\ldots\bar{\psi}(y_n)e^{i\int d^4x\mathcal{L}_I}]\,|0\rangle}{\langle 0|\,T[e^{i\int d^4x\mathcal{L}_I}]\,|0\rangle}. \quad (4.481)$$

This is the *Gell-Mann–Low formula* for Dirac fields. The actual reason why the aforestated form emulates the form (4.240) for scalar fields is because the basic logical steps that went into the derivation of (4.240) are not dependent on spin.

Now, we note that for free fields we can write

$$\langle 0|\,T[\psi(x_1)\ldots\bar{\psi}(y_n)e^{i\int d^4x\mathcal{L}_I}]\,|0\rangle = \prod_{i=1}^{n}\left(-i\frac{\delta}{\delta\bar{\eta}(x_i)}\right)\prod_{j=1}^{n}\left(i\frac{\delta}{\delta\eta(y_j)}\right)e^{i\int d^4x\mathcal{L}_I\left(-i\frac{\delta}{\delta\bar{\eta}(x_i)}, i\frac{\delta}{\delta\eta(y_j)}\right)} \times e^{-i\iint d^4xd^4y\bar{\eta}(x)S_F(x,y)\eta(y)}\Big|_{\eta,\bar{\eta}=0}. \quad (4.482)$$

This can be equivalently rewritten as

$$\begin{aligned}\langle 0|\,T[\psi(x_1)\ldots\bar{\psi}(y_n)e^{i\int d^4x\mathcal{L}_I}]\,|0\rangle &= \prod_{i=1}^{n}\left(-i\frac{\delta}{\delta\bar{\eta}(x_i)}\right)\prod_{j=1}^{n}\left(i\frac{\delta}{\delta\eta(y_j)}\right)Z^{-1}[0,0]e^{i\int d^4x\mathcal{L}_I\left(-i\frac{\delta}{\delta\bar{\eta}(x_i)}, i\frac{\delta}{\delta\eta(y_j)}\right)} \\ &\quad\times\int\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{iS_0[\psi,\bar{\psi}]+\int\bar{\eta}\psi+\int\bar{\psi}\eta}\Big|_{\eta,\bar{\eta}=0} \\ &= Z^{-1}[0,0]\int\mathcal{D}\psi\mathcal{D}\bar{\psi}\psi(x_1)\ldots\bar{\psi}(y_n)e^{iS_0[\psi,\bar{\psi}]+iS_I[\psi,\bar{\psi}]} \\ &= Z^{-1}[0,0]\int\mathcal{D}\psi\mathcal{D}\bar{\psi}\psi(x_1)\ldots\bar{\psi}(y_n)e^{iS[\psi,\bar{\psi}]}. \end{aligned} \quad (4.483)$$

Here $S[\psi,\bar{\psi}]$ is a full action. With this we can write

$$\langle\Omega|\,T[\psi_H(x_1)\ldots\bar{\psi}_H(y_n)]\,|\Omega\rangle = \frac{\int\mathcal{D}\psi\mathcal{D}\bar{\psi}\,\psi(x_1)\ldots\bar{\psi}(y_n)e^{iS[\psi,\bar{\psi}]}}{Z[0,0]\,\langle 0|\,T\left[e^{iS_I[\psi,\bar{\psi}]}\right]|0\rangle}. \quad (4.484)$$

In particular, for $n = 0$ we get

$$1 = \langle\Omega|\Omega\rangle = \frac{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, e^{iS[\psi,\bar{\psi}]}}{Z[0,0]\,\langle 0|\, T\left[e^{iS_I[\psi,\bar{\psi}]}\right]|0\rangle}\,. \qquad (4.485)$$

With this we finally arrive at functional-integral representation of the $2n$-point Green function for Dirac fields

$$\begin{aligned}&\langle\Omega|\, T[\psi_H(x_1)\ldots\bar{\psi}_H(y_n)]\,|\Omega\rangle \\ &\quad = \frac{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\,\psi(x_1)\ldots\bar{\psi}(y_n)\, e^{iS[\psi,\bar{\psi}]}}{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, e^{iS[\psi,\bar{\psi}]}}\,. \qquad (4.486)\end{aligned}$$

Yukawa Interaction

In cases when only scalars and spin-1/2 fermions are present in the theory, or when we are interested only in scalar-fermion sector of a theory (e.g., when describing pion-nucleon scattering) the Lagrangian has the generic form

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_{int} \\ &= \bar{\psi}\gamma^\mu\partial_\mu\psi + M\bar{\psi}\psi + \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 + \mathcal{L}_{int}(\psi,\bar{\psi},\phi)\,. \qquad (4.487)\end{aligned}$$

Here both ψ and ϕ might be generally field multiplets. A particular form of $\mathcal{L}_{int}$ is the so-called *Yukawa interaction*, which appears in two versions:

a) when ϕ is a *parity even scalar* then

$$\mathcal{L}_{Y,int} = -g\bar{\psi}\phi\psi\,, \qquad (4.488)$$

b) when ϕ is a *parity odd scalar* (i.e., *pseudoscalar*) then

$$\mathcal{L}_{Y,int} = -ig\bar{\psi}\gamma^5\phi\psi\,. \qquad (4.489)$$

(Here i ensures that $\mathcal{L}_{Y,int}$ is Hermitian).

Note

For real pion-nucleon interaction the Yukawa interaction term is a bit more complicated, because both nucleons and pions are field multiplets:

$$\psi = \begin{pmatrix}\psi_p \\ \psi_n\end{pmatrix}, \quad \text{and} \quad \phi \longrightarrow \boldsymbol{\phi} = \begin{pmatrix}\pi^+ \\ \pi^- \\ \pi^0\end{pmatrix}.$$

Besides, Higgs scalar field (in Standar Model of particle physics it is a *complex scalar doublet*) is also coupled to quarks and leptons via Yukawa interaction.

In order to quantize such systems via functional integrals, we first write generating functional (we consider for simplicity only one, parity even scalar field and one Dirac fermion)

$$
\begin{aligned}
Z[\eta,\bar{\eta},J] & \\
&= \exp\left(i\int_{\mathbb{R}^4}\mathcal{L}_{int}\left[-i\frac{\delta}{\delta\bar{\eta}(x)}, i\frac{\delta}{\delta\eta(x)}, -i\frac{\delta}{\delta J(x)}\right]\right) \\
&\times \langle 0|\, T\left[e^{i\int_{\mathbb{R}^4} d^4x[\bar{\psi}(x)\eta(x)+\bar{\eta}(x)\psi(x)+\phi(x)J(x)]}\right]|0\rangle \\
&= \exp\left(i\int_{\mathbb{R}^4}\mathcal{L}_{int}\left[-i\frac{\delta}{\delta\bar{\eta}(x)}, i\frac{\delta}{\delta\eta(x)}, -i\frac{\delta}{\delta J(x)}\right]\right) \\
&\times N\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, e^{iS_0[\psi,\bar{\psi}]\, +\, i\int d^4x\bar{\eta}\psi\, +\, i\int d^4x\bar{\psi}\eta} \\
&\times \int \mathcal{D}\phi\, e^{iS_0[\phi]\, +\, i\int d^4x J\phi} \\
&= N\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}\phi\, e^{iS[\psi,\bar{\psi},\phi]\, +\, i\int d^4x\bar{\eta}\psi\, +\, i\int d^4x\bar{\psi}\eta\, +\, i\int d^4xJ\phi}\,. \qquad (4.490)
\end{aligned}
$$

Here

$$S[\psi,\bar{\psi},\phi] \;=\; \int d^4x\left[\bar{\psi}\left(i\partial\!\!\!/ - M\right)\psi + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - g\phi\bar{\psi}\psi\right].$$

Here M and m are masses of fermion and scalar particle, respectively. We can get rid of the factor N by working directly with the generating functional for Green functions

$$\tilde{Z}[\eta,\bar{\eta},J] \;=\; \frac{Z[\eta,\bar{\eta},J]}{Z[0,0,0]}\,. \qquad (4.491)$$

For this we get

$$\tilde{Z}[\eta,\bar{\eta},J] = \frac{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}\phi\, e^{iS[\psi,\bar{\psi},\phi]+i\int d^4x\bar{\eta}\psi+i\int d^4x\bar{\psi}\eta+i\int d^4xJ\phi}}{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}\phi\, e^{iS[\psi,\bar{\psi},\phi]}}\,. \qquad (4.492)$$

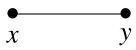
By having $\tilde{Z}[\eta,\bar{\eta},J]$, we can generate the mixed full Green function in a standard way, for instance

$$
\begin{aligned}
&\langle\Omega|\, T\left[\phi_H(x_1)\ldots\phi_H(x_n)\psi_H(y_1)\ldots\bar{\psi}_H(z_m)\right]|\Omega\rangle \\
&= \prod_{i=1}^{n}\left(-i\frac{\delta}{\delta J(x_i)}\right)\prod_{l=1}^{m}\left(-i\frac{\delta}{\delta\bar{\eta}(y_l)}\right)\prod_{j=1}^{m}\left(i\frac{\delta}{\delta\eta(z_j)}\right)\tilde{Z}[\eta,\bar{\eta},J]\Big|_{\eta,\bar{\eta},J=0} \\
&= \frac{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}\phi\,\phi(x_1)\ldots\phi(x_n)\psi(y_1)\ldots\bar{\psi}(z_m)\, e^{iS[\psi,\bar{\psi},\phi]}}{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}\phi\, e^{iS[\psi,\bar{\psi},\phi]}}\,. \qquad (4.493)
\end{aligned}
$$

Feynman Rules for Yukawa Interaction

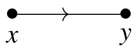
In the position space we can formulate Feynman rules as we did for scalar fields. *Lines* (i.e. propagators) are deduced from quadratic parts of the action (propagators correspond to the inverse of the integral kernel), while the *vertices* are implied by the interaction term. So, Feynman rules read:

- Draw all topologically distinct diagrams for given n-point Green function with n external legs for order g^m use m vertices.
- A line between points x and y can be either bosonic



$$= \int \frac{d^4p}{(2\pi)^4} \frac{ie^{-ip(x-y)}}{p^2 - m^2 + i\varepsilon}, \tag{4.494}$$

or fermionic



$$= \int \frac{d^4p}{(2\pi)^4} \frac{ie^{-ip(x-y)}}{\not{p} - m + i\varepsilon}. \tag{4.495}$$

Arrow Orientation

The orientation of the arrow is just a convention. Since

$$\psi = \sum_{p,\lambda} \left[a_{p,\lambda} u_{p,\lambda} e^{-ipx} + b^\dagger_{p,\lambda} v_{p,\lambda} e^{ipx} \right],$$

$$\bar{\psi} = \sum_{p,\lambda} \left[b_{p,\lambda} \bar{v}_{p,\lambda} e^{-ipx} + a^\dagger_{p,\lambda} \bar{u}_{p,\lambda} e^{ipx} \right].$$

For $x_0 > y_0$ we have that

$$\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle = \sum_{p,\lambda} \langle 0| \underbrace{a_{p,\lambda}}_{\nwarrow x} \underbrace{a^\dagger_{p,\lambda}}_{\nwarrow y} |0\rangle \cdots,$$

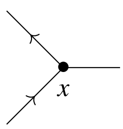
which describes a particle created at y and annihilated at x and particle, hence charge of a *particle* flows from x to y.

For $x_0 < y_0$ we have

$$\langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle = \sum_{p,\lambda} \langle 0| \underbrace{b_{p,\lambda}}_{\nwarrow y} \underbrace{b^\dagger_{p,\lambda}}_{\nwarrow x} |0\rangle \cdots,$$

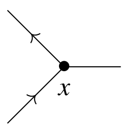
terms which describe antiparticle created at x and annihilated at y, hence charge of an *antiparticle* flows from y to x, which is equivalent to saying that charge of a *particle* flows from x to y.

- A vertex with 3 lines is represented by



$$= -ig \int d^4x \cdots, \tag{4.496}$$

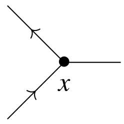
or



$$= +g\gamma^5 \int d^4x \cdots. \tag{4.497}$$

- Introduce symmetry factors where necessary.

Show explicitly that the rule for pseudoscalar



$$= +g\gamma^5 \int d^4x \cdots , \tag{4.498}$$

is correct correctly (factor i absent and the overall sign is +).

This can be directly seen from the fact that in the functional integral we have the action multiplied by i and the interaction term in the action comes with the factor $-ig$, so the overall factor in the vertex in Feynman diagram should be $i \times (-i)g = +g$. This can also be independently checked by computing an appropriate 3-point Green function in a tree approximation. In particular, let us consider the generating functional

$$Z[\eta, \bar\eta, J] = e^{iS_{Y,int}\left[-i\frac{\delta}{\delta\bar\eta(x)}, i\frac{\delta}{\delta\eta(x)}, -i\frac{\delta}{\delta J(x)}\right]} Z_0[\eta, \bar\eta] Z_0[J] . \tag{4.499}$$

The connected 3-point function at tree level is given by the term

$$i(-ig)\int d^4x \left(i\frac{\delta}{\delta\eta(x)}\right)\gamma^5\left(-i\frac{\delta}{\delta\bar\eta(x)}\right)\left(-i\frac{\delta}{\delta J(x)}\right) Z_0[\eta,\bar\eta] Z_0[J] . \tag{4.500}$$

Given that

$$\begin{aligned} Z_0[J] &= e^{-\frac{i}{2}\int d^4x_1 d^4x_2 J(x_1)\Delta_F(x_1-x_2)J(x_2)} , \\ Z_0[\eta,\bar\eta] &= e^{-i\int d^4x_1 d^4x_2 \bar\eta(x_1) S_F(x_1-x_2)\eta(x_2)} . \end{aligned} \tag{4.501}$$

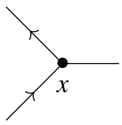
we find from Eq. (4.499) that

$$\begin{aligned} & i(-ig)\int d^4x \left(i\frac{\delta}{\delta\eta(x)}\right)\gamma^5\left(-i\frac{\delta}{\delta\bar\eta(x)}\right) Z_0[\eta,\bar\eta]\left(-i\frac{\delta}{\delta J(x)}\right) Z_0[J] \\ & = \{\text{relevant part only, i.e., we want to end up with connected diagram} \\ & \quad \text{with two Grassmann sources and one } J \text{ source}\} \\ & = \frac{g}{2}\int d^4x \left(i\frac{\delta}{\delta\eta_\alpha(x)}\right)\gamma^5_{\alpha\beta}\left(-i\frac{\delta}{\delta\bar\eta_\beta(x)}\right)\int d^4x_1 d^4x_2\, \bar\eta^a_{x_1} iS_F^{ab}(x_1-x_2)\eta^b_{x_2} \\ & \times \int d^4y_1 d^4y_2\, \bar\eta^c_{y_1} iS_F^{cd}(y_1-y_2)\eta^d_{y_2}\int d^4z[-\Delta_F(z-x)]J(x) \\ & = -g\int d^4x d^4x_1 d^4x_2 d^4z\, \bar\eta(x_1) S_F(x_1-x)\gamma^5 S_F(x-x_2)\bar\eta(x_2) \\ & \times\ \Delta_F(z-x)J(z). \end{aligned} \tag{4.502}$$

The corresponding 3-point Green function is obtained by taking 3 functional derivatives of $\tilde Z[\eta, \bar\eta, J]$ (which on the tree level is the same as $Z[\eta, \bar\eta, J]$), in particular

$$
\begin{aligned}
&\langle 0|\, T\,[\psi(x_1)\bar{\psi}(x_2)\phi(x_3)]\,|0\rangle \\
&= \left(-i\frac{\delta}{\delta\bar{\eta}(x_1)}\right)\left(i\frac{\delta}{\delta\eta(x_2)}\right)\left(i\frac{\delta}{\delta J(x_3)}\right)\tilde{Z}[\eta,\bar{\eta},J]\Big|_{\bar{\eta},\eta,J} \\
&= -(-i)g\int d^4x\,(-1)S_F(x_1-x)\gamma^5 S_F(x-x_2)\Delta_F(x-x_3)+O(g^2) \\
&= (-i)g\int d^4x S_F(x_1-x)\gamma^5 S_F(x-x_2)\Delta_F(x-x_3)+O(g^2) \\
&= \int d^4x\,[iS_F(x_1-x)]\,g\gamma^5\,[iS_F(x-x_2)]\,[i\Delta_F(x-x_3)]+O(g^2)\,. \quad (4.503)
\end{aligned}
$$

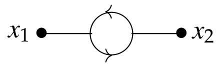
This is precisely the result that we would have obtained should we have used Feynman rules with the vertex prescription



$$\sim\ g\gamma^5\int d^4x\cdots\,. \qquad (4.504)$$

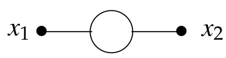
As an exercise, compare relative signs of loops in:

A) Yukawa theory with $\mathcal{L}_{Y,int}=-g\bar{\psi}\psi\phi$



$$\leftrightarrow\quad \langle x_1x_2\rangle_{\bar{\psi}\psi\phi}\,.$$

B) scalar theory with $\mathcal{L}_{int}=-\frac{g}{3!}\phi^3$ theory



$$\leftrightarrow\quad \langle x_1x_2\rangle_{\phi^3}\,.$$

To make this comparison, we can employ the generating functionals $Z[\bar{\eta},\eta,J]$ and $Z[J]$ for respective Green functions. We should use only that parts of $Z[\bar{\eta},\eta,J]$ and $Z[J]$ that contributes to the second order in the coupling constant and use only as many external source terms J that are relevant to the above Feynman diagrams. In particular, for the diagram A) we can thus write

$$\left(-i\frac{\delta}{\delta J(x_1)}\right)\left(-i\frac{\delta}{\delta J(x_2)}\right)\frac{1}{2}(-i)g\int d^4x\left(i\frac{\delta}{\delta\eta(x)}\right)\left(-i\frac{\delta}{\delta\bar{\eta}(x)}\right)\left(-i\frac{\delta}{\delta J(x)}\right)$$
$$\times(-i)g\int d^4y\left(i\frac{\delta}{\delta\eta(y)}\right)\left(-i\frac{\delta}{\delta\bar{\eta}(y)}\right)\left(-i\frac{\delta}{\delta J(y)}\right)$$
$$\times\left[-\frac{1}{2!}\int d^4x'd^4y'\,\bar{\eta}(x')S_F(x'-y')\eta(y')\right.$$
$$\left.\times\int d^4xd^4y\,\bar{\eta}(x)S_F(x-y)\eta(y)\right]$$
$$\times\left[-\frac{1}{2!2^2}\int d^4x'd^4y'\,J(x')\Delta_F(x'-y')J(y')\right.$$
$$\left.\times\int d^4xd^4y\,J(x)\Delta(x-y)J(y)\right]. \quad (4.505)$$

Similar formula would hold for diagram B) but with only J sources and different combinatorial factors.

Now, due to *anticommuting* property of Grassmann derivatives we obtain from the functional differentiation of Grassmann sources overall -1 sign. On the other hand, in the diagram B) the structure of computations would be analogous, but the *commuting* nature of functional derivatives $\frac{\delta}{\delta J}$ brings an overall sign of $+1$.

Notes on Fermionic loops

i) Above result is quite generic. Fermionic loops appear with opposite sign than analogous bosonic loops. In fact, one should add to Feynman rules for Yukawa theory that *each Fermionic loop carries extra* -1 *factor.*

ii) In exactly supersymmetric theories bosonic loop diagrams are cancelled by fermionic loop diagrams.

4.16 Feynman Rules in Momentum Space

It is often technically simpler and conceptually more convenient to give Feynman rules in *momentum space*, i.e., to consider the Fourier transform of $\tau(x_1,\ldots,x_n)$

$$\tilde{\tau}(p_1,\ldots,p_n)=\int d^4x_1e^{-ip_1x_1}\ldots\int d^4x_ne^{+ip_nx_n}\,\tau(x_1,\ldots,x_n), \quad (4.506)$$

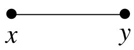
where

$$\tau(x_1,\ldots,x_n)\;\equiv\;\langle\Omega|\,T\left[\phi_H(x_1)\ldots\phi_H(x_n)\right]|\Omega\rangle\;. \quad (4.507)$$

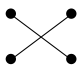
"Sign convention"

The minus sign in the exponent in (4.506) is associated to incoming particles (momenta flow to interaction zone) and *the plus sign* in the exponent is associated to outgoing particles (momenta flow from interaction zone). We will say more about this convention when discussing LSZ formula.

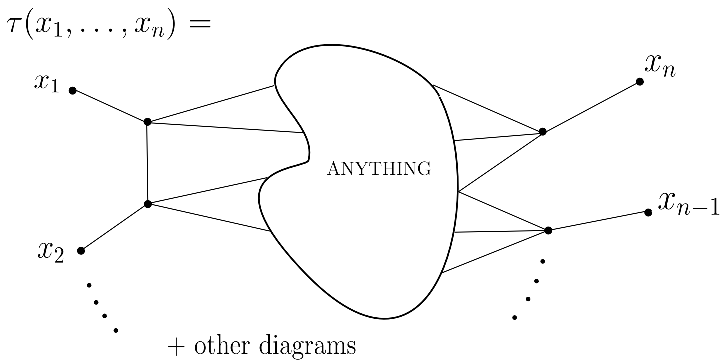
Recall that in the *position space* we had following rules (e.g., for $\lambda\phi^4$ theory)



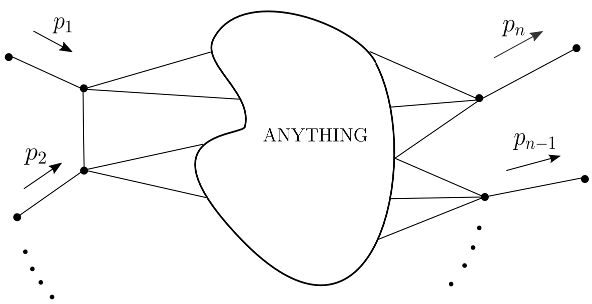
$$= \Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\varepsilon}e^{-ip(x-y)},$$



$$= -i\lambda\int d^4x\cdots. \tag{4.508}$$



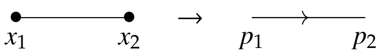
Clearly, in $\tilde{\tau}$ the momentum of every external line is affiliated to the external momentum (appropriate argument of $\tilde{\tau}$). This is due to the fact that d^4x_i integration that is followed by d^4p_i integration. So, that



where $p_1,\ldots,p_n$ appear as arguments of $\tilde{\tau}$.

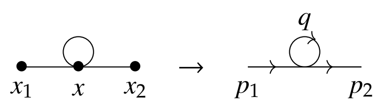
In order to better understand the situation let us discuss some examples.

Example 1



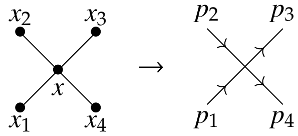
$$
\begin{aligned}
&= \int d^4x_1 d^4x_2\, e^{-ip_1x_1} e^{ip_2x_2} \Delta_F(x_2 - x_1) \\
&= \int d^4x_1 d^4x_2\, e^{-ip_1x_1} e^{ip_2x_2} i \int \frac{d^4p}{(2\pi)^4} \frac{e^{-ip(x_1-x_2)}}{p^2 - m^2 + i\varepsilon} \\
&= \int \frac{d^4p}{(2\pi)^4} d^4x_1 d^4x_2\, e^{-ix_1(p_1+p)} e^{ix_2(p_2+p)} \frac{i}{p^2 - m^2 + i\varepsilon} \\
&= (2\pi)^4 \delta(p_1 - p_2) \frac{i}{p_1^2 - m^2 + i\varepsilon} \,. \qquad (4.509)
\end{aligned}
$$

Example 2



$$
\begin{aligned}
&= \int d^4x_1 d^4x_2\, e^{-ip_1x_1} e^{ip_2x_2} (-i\lambda) \int d^4x\, \Delta_F(x_1 - x) \Delta_F(x - x) \Delta_F(x - x_2) \\
&= (-i\lambda) \int d^4x_1 d^4x_2 d^4x\, e^{-ip_1x_1} e^{ip_2x_2} \int \frac{d^4q_1}{(2\pi)^4} \frac{i e^{-iq_1(x_1-x)}}{q_1^2 - m^2 + i\varepsilon} \\
&\quad \times \int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon} \int \frac{d^4q_2}{(2\pi)^4} \frac{i e^{-iq_2(x-x_2)}}{q_2^2 - m^2 + i\varepsilon} \\
&= (-i\lambda) \int d^4x_1 d^4x_2 d^4x \frac{d^4q_1}{(2\pi)^4} \frac{d^4q}{(2\pi)^4} \frac{d^4q_2}{(2\pi)^4} \\
&\quad \times e^{-ix_1(p_1+q_1)} e^{ix_2(p_2+q_2)} e^{ix(q_1-q_2)} \prod_{i=1}^{3} \frac{i}{q_i^2 - m^2 + i\varepsilon} \\
&= (-i\lambda) \frac{1}{(2\pi)^{12}} \int d^4q_1 d^4q d^4q_2 \delta(p_1 + q_1) \delta(p_2 + q_2) \delta(q_1 - q_2) \\
&\quad \times (2\pi)^{12} \prod_{i=1}^{3} \frac{i}{q_i^2 - m^2 + i\varepsilon} \\
&= (-i\lambda)(2\pi)^4 \delta(p_1 - p_2) \left(\frac{i}{p_1^2 - m^2 + i\varepsilon} \right)^2 \int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon} \,. \qquad (4.510)
\end{aligned}
$$

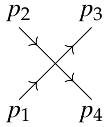
Example 3



$$
\begin{aligned}
&= (-i\lambda)\int d^4x_1\, e^{-ip_1x_1}\int d^4x_2\, e^{-ip_2x_2}\int d^4x\int d^4x_3\, e^{ip_3x_3}\int d^4x_4\, e^{ip_4x_4} \\
&\quad\times\ \Delta_F(x_1-x)\Delta_F(x_2-x)\Delta_F(x-x_4)\Delta_F(x-x_3) \\
&= (-i\lambda)\int d^4x_1 d^4x_2 d^4x_3 d^4x_4 d^4x\, e^{-ip_1x_1}e^{-ip_2x_2}e^{ip_3x_3}e^{ip_4x_4} \\
&\quad\times \int\frac{d^4q_1}{(2\pi)^4}\frac{i}{q_1^2-m^2+i\varepsilon}e^{-iq_1(x_1-x)}\int\frac{d^4q_2}{(2\pi)^4}\frac{i}{q_2^2-m^2+i\varepsilon}e^{-iq_2(x_2-x)} \\
&\quad\times \int\frac{d^4q_3}{(2\pi)^4}\frac{i}{q_3^2-m^2+i\varepsilon}e^{-iq_3(x-x+3)}\int\frac{d^4q_4}{(2\pi)^4}\frac{i}{q_4^2-m^2+i\varepsilon}e^{-iq_4(x-x_4)} \\
&= (-i\lambda)\int d^4x_1 d^4x_2 d^4x_3 d^4x_4 d^4x\,\frac{d^4q_1 d^4q_2 d^4q_3 d^4q_4}{(2\pi)^{16}} \\
&\quad\times\ e^{-ix_1(p_1+q_1)}e^{-ix_2(p_2+q_2)}e^{ix_3(p_3+q_3)}e^{ix_4(p_4+q_4)}e^{ix(q_1+q_2-q_3-q_4)} \\
&\quad\times\prod_{i=1}^{4}\frac{i}{q_i^2-m^2+i\varepsilon} \\
&= (-i\lambda)\int d^4q_1 d^4q_2 d^4q_3 d^4q_4 (2\pi)^4\delta(p_1+q_1)\delta(p_2+q_2)\delta(p_3+q_3) \quad (4.511) \\
&\quad\times\ \delta(p_4+q_4)\delta(q_1+q_2-q_3-q_4)\prod_{i=1}^{4}\frac{i}{q_i^2-m^2+i\varepsilon} \\
&= (-i\lambda)(2\pi)^4\delta(p_1+p_2-p_3-p_4)\prod_{i=1}^{4}\frac{i}{p_i^2-m^2+i\varepsilon}\,. \quad (4.512)
\end{aligned}
$$

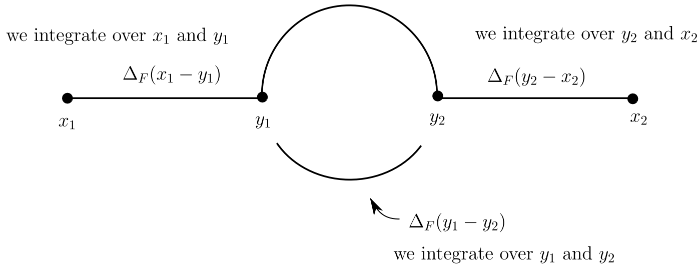
What we have learned from the foregoing 3 examples:

- Momentum of every external line is affiliated with the external momentum (the argument of $\tilde{\tau}$). This is because each d^4x_i integration is followed by d^4p_i integration.
- Each d^4x integration of a vertex enforces momentum conservation at that vertex. For instance, for $\lambda\phi^4$ theory we have



$$
\begin{aligned}
&\propto\quad (-i\lambda)\int d^4x\, e^{ix(p_4+p_3-p_1-p_2)} \\
&=\quad (-i\lambda)(2\pi)^4\,\delta(p_1+p_2-p_3-p_4)\,.
\end{aligned}
$$

- Since each propagator has argument either at a vertex or on external point, all $e^{\pm ipx}$ factors of propagators Δ_F's are used up.



In particular, all $e^{\pm ipx}$ disappear and are turned into δ-functions.

- Momentum conservation at all vertices is enforced via δ-function and "kills" many of the d^4p integrations.
- All those propagator moments that are not fixed by the δ-functions (originating from integrating d^4x over external or vertex points) are still integrated over with $\int \frac{d^4p}{(2\pi)^4}$. These are the so-called "loop momentum" integrals. In fact, the following statement holds:

Note

There are as many remaining momentum integrations (in a given Feynman diagram) as there are loops.

Proof: External lines do not have any integration (as we said, d^4x_i and d^4p_i integration follow each other and $e^{\pm ipx}$ in the propagator produces δ-functions that cancel integration and set the momenta in propagator to corresponding external momenta.

- What remains are integrations for each *internal propagatot/line* $(= I)$.
- Each vertex produces one δ-function representing momentum conservation. Hence the number of δ-functions is equal to the number of *vertices* $(= V)$.
- All δ-functions are not independent, they provide overal momentum conservation δ-function.
- Hence we end-up with $V - 1$ independent δ-functions.

Total number of integrations is then $I - (V - 1) = I - V + 1 = L$ by *Euler formula*.

Note

The fact that the δ-function corresponding to *total momentum conservation* must always be factored out from $\tilde{\tau}$ is a consequence of translational invariance of τ.

Indeed, we first note that

$$\begin{aligned}
&\langle\Omega| T\left[\phi_H(x_1)\ldots\phi_H(x_n)\right]|\Omega\rangle \\
&\quad = \langle\Omega| e^{iP^\nu a_\nu} T\left[e^{-iP^\nu a_\nu}\phi_H(x_1+a)e^{iP^\nu a_\nu}\times\cdots\right. \\
&\qquad \left.\cdots\times\, e^{-iP^\nu a_\nu}\phi_H(x_n+a)e^{iP^\nu a_\nu}\right] e^{-iP^\nu a_\nu}|\Omega\rangle\,. \qquad (4.513)
\end{aligned}$$

By employing assumption of the translational invariance of the vacuum state $|\Omega\rangle$ (i.e. $e^{iP^\nu a_\nu}|\Omega\rangle = |\Omega\rangle$) we get

$$\tau(x_1,\ldots,x_n) = \tau(x_1+a,\ldots,x_n+a)\,. \qquad (4.514)$$

Now,

$$\tilde{\tau}(p_1,\ldots,p_n) = \int d^4x_1\ldots d^4x_n\, e^{-i\sum_i p_i x_i}\,\tau(x_1,\ldots,x_n)\,. \qquad (4.515)$$

Here p_i's appear with appropriate signs. Then

$$\begin{aligned}
&\tilde{\tau}(p_1,\ldots,p_n) \\
&= \int d^4(x_1+a)\ldots d^4(x_n+a)\, e^{-i\sum_i p_i(x_i+a)}\tau(x_1+a,\ldots,x_n+a) \\
&= e^{-i\sum_i p_i a}\int d^4x_1\ldots d^4x_n\, e^{-i\sum_i p_i x_i}\,\tau(x_1,\ldots,x_n) \\
&= e^{-i\sum_i p_i a}\,\tilde{\tau}(p_1,\ldots,p_n)\,. \qquad (4.516)
\end{aligned}$$

This gives an equation

$$\left(e^{-i\sum_i p_i a} - 1\right)\tilde{\tau}(p_1,\ldots,p_n) = 0\,. \qquad (4.517)$$

That must be satisfied for all a. Particularly for small a we have up to the first order in a that

$$\left(\sum_i p_i\right)\tilde{\tau}(p_1,\ldots,p_n) = 0. \qquad (4.518)$$

This has a *general solution*

$$\tilde{\tau}(p_1,\ldots,p_n) = \delta\left(\sum_i p_i\right)(2\pi)^4\tau(p_1,\ldots,p_n)\,, \qquad (4.519)$$

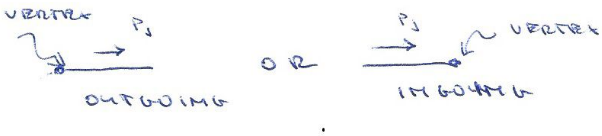
where the residual Green function $\tau(p_1,\ldots,p_n)$ is (for simplicity) denotes with the same symbol "τ" as the position-space Green function. Factor $(2\pi)^4$ is mere convention.

Consequently, the total momentum conservation is always factorized out from momentum-space Green functions. Since this is true for any full Green's function, it must be true also order by order and diagram by diagram.

Summary of Feynman rules for momentum-space Green functions $\tau(p_1, \ldots, p_n)$

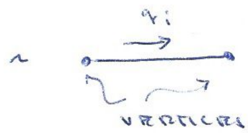
1. Draw all topologically distinct diagrams with n external lines with ensuing momenta $p_1, \ldots, p_n$. *Incoming momenta* are considered to be *positive*, while *outgoing momenta* are *negative*. For each diagram denote by $q_1, \ldots, q_I$ the momenta of internal lines. (In scalar theory without derivative coupling the choice of an orientation of internal lines is irrelevant.)
2. To the j-th external line assign the factor

$$\frac{i}{p_j^2 - m^2 + i\varepsilon}$$

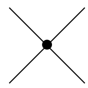
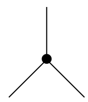


3. To the i-th internal line assign the factor

$$\frac{i}{q_i^2 - m^2 + i\varepsilon}$$

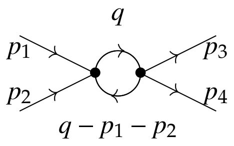


4. To each vertex assign vertex factor, i.e. $(-i\lambda)$ for $\lambda\phi^4/4!$ theory and $(-ig)$ for $g\phi^3/3!$ theory, i.e.

 $\sim -i\lambda\,,$  $\sim -ig\,.$

5. Additionally the following rules apply:
   - Assign momenta at each vertex so that the momentum conservation is ensured.
   - Multiply by $\int \frac{d^4q}{(2\pi)^4}$ for each closed loop, here q is the free (unconstrained by momentum conservation) momenta propagation along the loop.
   - Factor out total factor $(2\pi)^4\delta(p_f - p_i)$ representing total momentum conservation.
6. Divide by the *symmetry factor*.
7. Sum the contributions of all topologically distinct diagrams to a given order in λ or g, etc.

Example 4



$$= (-i\lambda)^2 \frac{i}{p_1^2 - m^2 + i\varepsilon} \frac{i}{p_2^2 - m^2 + i\varepsilon} \frac{i}{p_3^2 - m^2 + i\varepsilon} \frac{i}{p_4^2 - m^2 + i\varepsilon}$$
$$\times \int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m^2 + i\varepsilon} \frac{i}{(q - p_1 - p_2)^2 - m^2 + i\varepsilon} .$$

This is a second order contribution to the 4-point Green's function $\tau(p_1, p_2, p_3, p_4)$.

By convention momenta entering vertex have positive sign and outgoing momenta have negative sign. So, the momentum z propagating on the lower half of the loop satisfies $p_1 + p_2 - q + z = 0$ or equivalently $z = q - p_1 - p_2$.

4.17 LSZ Formalism

Particle physicists and phenomenologists are mostly interested in S-matrix elements

$$_{out}\langle p'_1, \dots, p'_n | p_1, \dots, p_m \rangle_{in} ,$$

that are directly relevant, e.g., for cross-section computations. On the other hand, quantum field theorists are mostly interested in Green functions

$$\langle \Omega | T\left[\phi_H(x_1) \dots \phi_H(x_{n+m})\right] | \Omega \rangle \equiv \tau(x_1, \dots, x_{n+m}) ,$$

because they are easily calculable in perturbation theory and they also provide basic building blocks in applications that go beyond simple scattering theory.

We will now demonstrate that it is possible to compute the S-matrix elements (i.e., scattering amplitudes) directly in terms of $\tau(x_1, \dots, x_{n+m})$, so that all our labor with a perturbation computations of $\tau(x_1, \dots)$ can be employed, e.g., in cross-section computations. Let us, however, first start with two important concepts.

Spectral density and Z-factors

For simplicity's sake we will carry out our following argumentation in terms of scalar fields, even though the results obtained will be more general and with small modifications valid also for Dirac fermions and gauge fields.

In the following, we will use the Heisenberg picture (hats over operators are omitted)

$$\phi_H(x) = e^{iHt} \phi_S(\boldsymbol{x}) e^{-iHt} , \quad H = H_0 + H_I ,$$
$$|\psi\rangle \equiv |\psi_H\rangle = |\psi_S(t = 0)\rangle . \tag{4.520}$$

Here H is the full Hamiltonian in Schrödinger picture.

Consider now the correlation function

$$i\Delta_+(x-y) \;=\; \langle\Omega|\,\phi_H(x)\phi_H(y)\,|\Omega\rangle\;. \tag{4.521}$$

For a free field (let us denote it here as ϕ_0) with mass m_0 we have

$$\begin{aligned}
\langle 0|\,\phi_0(x)\phi_0(y)\,|0\rangle &= \sum_p\sum_q \langle 0|\,a(p)a^\dagger(q)\,|0\rangle\, e^{-ipx+iqy}\\
&= \sum_p\sum_q \langle 0|\,[a(p),a^\dagger(q)]\,|0\rangle\, e^{-ipx+iqy}\\
&= \int_{\mathbb{R}^3}\frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_p}e^{-ip(x-y)}\\
&= \int_{\mathbb{R}^4}\frac{d^4p}{(2\pi)^3}e^{-ip(x-y)}\delta(p^2-m_0^2)\theta(p_0)\\
&\equiv iD_+(x-y,m_0)\,.
\end{aligned} \tag{4.522}$$

Let us now turn to the general case (i.e., situation when interaction is included). Then we can write

$$\langle\Omega|\,\phi_H(x)\phi_H(y)\,|\Omega\rangle = \sum_\alpha \langle\Omega|\,\phi_H(x)\,|\alpha\rangle\,\langle\alpha|\,\phi_H(y)\,|\Omega\rangle\;, \tag{4.523}$$

where the sum runs over some complete set of states in the Heisenberg picture. We chose these base states to be eigenstates of the full Hamiltonian ($H_H = H$). Since the momentum $\hat{\boldsymbol{P}}$ operator commutes with H, we can chose $|\alpha_0\rangle$ to be eigenstates of H with momentum zero (i.e., $\hat{\boldsymbol{P}}\,|\alpha_0\rangle = 0$), then all the boosts of $|\alpha_0\rangle$ are also eigenstates of H. The eigenvalues of the 4-momentum operator $p^\mu = (H,\boldsymbol{p})$ are organized in sets of hyperboloids.

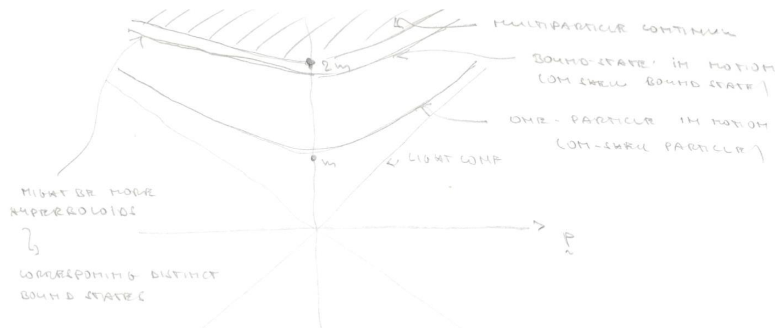


Figure 4.14: Schematic picture of the energy spectrum for scalar field theory.

Multiparticle continuum is bounded by a hyperboloid with $H = \sqrt{(2m)^2+(\boldsymbol{p})^2}$. Consider two-particle state with

$$H \;=\; \sqrt{m^2+\boldsymbol{p}_1^2}\;+\;\sqrt{m^2+\boldsymbol{p}_2^2}\,. \tag{4.524}$$

Taking into account that $\sqrt{m^2+\boldsymbol{p}^2}$ is a convex function in $\boldsymbol{p}$, one can then use the *Jensen inequality*

$$\sqrt{m^2 + (s\boldsymbol{p}_1 + (1-s)\boldsymbol{p}_2)^2} \;\leq\; s\sqrt{m^2 + \boldsymbol{p}_1^2} + (1-s)\sqrt{m^2 + \boldsymbol{p}_2^2}\,, \tag{4.525}$$

which is valid for any $s \in [0, 1]$. For $s = 1/2$ we get

$$\frac{1}{2}\sqrt{(2m)^2 + (\boldsymbol{p}_1 + \boldsymbol{p}_2)^2} \;\leq\; \frac{1}{2}\sqrt{m^2 + \boldsymbol{p}_1^2} \;+\; \frac{1}{2}\sqrt{m^2 + \boldsymbol{p}_2^2}\,, \tag{4.526}$$

which implies that (we set the total 3-momenta $\boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2$)

$$\sqrt{(2m)^2 + \boldsymbol{P}^2} \;\leq\; \sqrt{m^2 + \boldsymbol{p}_1^2} + \sqrt{m^2 + \boldsymbol{p}_2^2}\,. \tag{4.527}$$

Thus, the lower bound for two-particle state is hyperboloid $H = \sqrt{(2m)^2 + \boldsymbol{P}^2}$. For more than two particle states the Jensen inequality implies

$$\sqrt{(Nm)^2 + \left(\sum_i^N \boldsymbol{p}_i\right)^2} \;\leq\; \sum_{i=1}^{N}\sqrt{m^2 + \boldsymbol{p}_i^2}\,, \tag{4.528}$$

which is still bounded from below by the hyperboloid $H = \sqrt{(2m)^2 + \boldsymbol{P}^2}$.

Bound states have lower energy than is a sum of free particles (due to negative binding energy). So, for instance, energy for bound state of two particles must appear in the graph H vs. $\boldsymbol{P}$ below the hyperboloid $H = \sqrt{(2m)^2 + \boldsymbol{P}^2}$, see, Fig. 4.14.

Let $|\alpha_{\boldsymbol{p}}\rangle$ be the boost of $|\alpha_{\boldsymbol{0}}\rangle$ with momentum $\boldsymbol{p}$. The resolution of unity can be formally written as

$$\begin{aligned}
\mathbf{1} \;&=\; \sum_{\alpha} |\alpha\rangle\,\langle\alpha| \;\equiv\; |\Omega\rangle\,\langle\Omega| \;+\; \sum_{\alpha}\int \frac{d^3\boldsymbol{P}}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{P}}(\alpha)}\,|\alpha_{\boldsymbol{P}}\rangle\,\langle\alpha_{\boldsymbol{P}}| \\
&\equiv\; |\Omega\rangle\,\langle\Omega| \;+\; \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\frac{1}{2\omega_{\boldsymbol{p}}}\,|\boldsymbol{p}\rangle\,\langle\boldsymbol{p}| \\
&+\; \cdots \text{ multiparticle contributions}\,.
\end{aligned} \tag{4.529}$$

Here $\omega_{\boldsymbol{P}}(\alpha) = \sqrt{m_\alpha^2 + \boldsymbol{P}^2}$, m_α is the mass of the state $|\alpha_{\boldsymbol{P}}\rangle$, i.e. the energy of the state $|\alpha_{\boldsymbol{0}}\rangle$. $\sum_\alpha$ in the first sum in α is meant both over discrete and continuous indices and in the second sum we sum over all zero-momentum states $|\alpha_{\boldsymbol{0}}\rangle$.

Let us now use two already known relations, namely

$$\phi(x + a) \;=\; e^{i\hat{P}^\nu a_\nu}\phi(x)e^{-i\hat{P}^\nu a_\nu} \;\Rightarrow\; \phi(x) \;=\; e^{i\hat{P}^\nu x_\nu}\phi(0)e^{-i\hat{P}^\nu x_\nu}\,, \tag{4.530}$$

and

$$e^{-i\hat{P}^\nu x_\nu}\,|\alpha_{\boldsymbol{P}}\rangle = e^{-iP^\nu_\alpha x_\nu}\,|\alpha_{\boldsymbol{P}}\rangle\,, \tag{4.531}$$

(in short $e^{-i\hat{P}^\nu x_\nu}|\alpha\rangle = e^{-P_\alpha^\nu x_\nu}|\alpha\rangle$). With these we get

$$\begin{aligned}\langle\Omega|\,\phi_H(x)\phi_H(y)\,|\Omega\rangle &= \sum_\alpha e^{-iP_\alpha^\nu(x-y)_\nu}|\,\langle\Omega|\,\phi_H(0)\,|\alpha\rangle\,|^2 \\ &= \int d^4q \sum_\alpha e^{-iq(x-y)}|\,\langle\Omega|\,\phi_H(0)\,|\alpha\rangle\,|^2\delta(P_\alpha - q) \\ &= \int \frac{d^4q}{(2\pi)^3} e^{-iq(x-y)}\rho(q)\,. \end{aligned} \tag{4.532}$$

Here we have used the fact that $e^{-i\hat{P}^\nu x_\nu}|\Omega\rangle = |\Omega\rangle$ and defined

$$\rho(q) \equiv (2\pi)^3 \sum_\alpha \delta(p_\alpha - q)\,|\langle\Omega|\,\phi_H(0)\,|\alpha\rangle|^2\,. \tag{4.533}$$

Note that $\rho(q)$ is obviously positive and vanishes for $q^0 < 0$ (due to positivity of the energy of physical states $|\alpha\rangle$). Furthermore, it is invariant under a Lorentz transformation as required by the corresponding property of the field ϕ_H. To see this we use the fact that under the Lorentz transformation

$$U(\Lambda)\phi_H(x)U^\dagger(\Lambda) = \phi(\Lambda x)\,. \tag{4.534}$$

and $U(\Lambda)\,|\Omega\rangle = |\Omega\rangle$. With this we see that the LHS of (4.532) equals to

$$\begin{aligned}&\langle\Omega|\,\phi_H(x)\phi_H(y)\,|\Omega\rangle \\ &\quad= \langle\Omega|\,U^\dagger(\Lambda)U(\Lambda)\phi_H(x)U^\dagger(\Lambda)U(\Lambda)\phi_H(y)U^\dagger(\Lambda)U(\Lambda)\,|\Omega\rangle \\ &\quad= \langle\Omega|\,\phi_H(\Lambda x)\phi_H(\Lambda y)\,|\Omega\rangle\,. \end{aligned} \tag{4.535}$$

This implies that the LHS of (4.532) is a Lorentz scalar and hence also ρ is a Lorentz scalar.

At this point we can then introduce the *spectral density* $\sigma(q^2)$

$$\rho(q) \equiv \theta(q_0)\sigma(q^2)\,. \tag{4.536}$$

σ thus quantifies the contribution of the intermediate states $|\alpha\rangle$ with $p_\alpha^2 = q^2$.

Further rewriting yields

$$\begin{aligned}&\langle\Omega|\,\phi_H(x)\phi_H(y)\,|\Omega\rangle \\ &\quad= \int_0^\infty d(M^2)\int \frac{d^4q}{(2\pi)^3}e^{-iq(x-y)}\delta(q^2 - M^2)\theta(q^0)\sigma(M^2) \\ &\quad= \int_0^\infty d(M^2)iD_+(x-y, M^2)\sigma(M^2)\,. \end{aligned} \tag{4.537}$$

It can be easily seen that by choosing $\sigma(q^2) = \delta(q^2 - m_0^2)$ we obtain the result for free field. At this point we note that

$$\begin{aligned}\rho(q) = \theta(q_0)\sigma(q^2) \quad &= (2\pi)^3 \sum_\alpha \delta(p_\alpha - q)| \langle\Omega| \phi_H(0) |\alpha\rangle |^2 \\ &= (2\pi)^3 \delta(q)| \langle\Omega| \phi_H(0) |\Omega\rangle |^2 \\ &+ (2\pi)^3 \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}} \delta(\boldsymbol{p} - \boldsymbol{q})\delta(q_0 - \omega_{\boldsymbol{p}})| \langle\Omega| \phi_H(0) |p\rangle |^2 \\ &+ (2\pi)^3 \sum_{\alpha_0} \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2\omega_{\boldsymbol{p}}(\alpha)} \delta(\boldsymbol{p} - \boldsymbol{q})\delta(q_0 - \omega_{\boldsymbol{p}}(\alpha))| \langle\Omega| \phi_H(0) |\alpha_{\boldsymbol{p}}\rangle |^2.\end{aligned} \tag{4.538}$$

The vacuum term $\langle\Omega| \phi_H(0) |\Omega\rangle$ is typically zero by symmetry (cf. $\lambda\phi^4$ theory, but not $g\phi^3$) and for higher-spin fields, it is zero by Lorentz invariance. If the vacuum term is non-zero, we can appropriately shift the field $\phi_H \to \phi_H + \text{const}$. So, in the following we neglect $\langle\Omega| \phi_H(0) |\Omega\rangle$. We can further manipulate matrix elements $\langle\Omega| \phi_H(0) |\alpha_{\boldsymbol{p}}\rangle$ as follows

$$\langle\Omega| \phi_H(0) |\alpha_{\boldsymbol{p}}\rangle = \langle\Omega| \phi_H(0) U_{\boldsymbol{p}}^{-1} U_{\boldsymbol{p}} |\alpha_{\boldsymbol{p}}\rangle = \langle\Omega| \phi_H(0) |\alpha_{\boldsymbol{0}}\rangle. \tag{4.539}$$

Here we used that $U_{\boldsymbol{p}}\phi_H(0)U_{\boldsymbol{p}}^{-1} = \phi_H(0)$ which implies that $\phi_H(0)U_{\boldsymbol{p}}^{-1} = U_{\boldsymbol{p}}^{-1}\phi_H(0)$ and $|\Omega\rangle\, U_{\boldsymbol{p}}^{-1} = |\Omega\rangle$ so, $\langle\Omega| \phi_H(0) |\alpha_{\boldsymbol{p}}\rangle$ is momentum independent due to Lorentz invariance. For fermions it is more difficult to show, but it works as well. Consequently we can write

$$\begin{aligned}\rho(q) \quad &= \frac{\delta(q_0 - \sqrt{m^2 + \boldsymbol{q}^2})}{\sqrt{m^2 + \boldsymbol{q}^2}} | \langle\Omega| \phi_H(0) |1_{\boldsymbol{p}=0}\rangle |^2 \\ &+ \sum_{\lambda_0} \frac{\delta(q_0 - \sqrt{m_\lambda^2 + \boldsymbol{q}^2})}{\sqrt{m_\lambda^2 + \boldsymbol{q}^2}} | \langle\Omega| \phi_H(0) |\lambda_0\rangle |^2 \\ &= \delta(q^2 - m^2)\theta(q^0) z + \sum_{\lambda_0} \delta(q^2 - m_\lambda^2)\theta(q^0)| \langle\Omega| \phi_H(0) |\lambda_0\rangle |^2.\end{aligned} \tag{4.540}$$

This means that

$$\sigma(q^2) = \delta(q^2 - m^2) z + \sum_{\lambda_0} \delta(q^2 - m_\lambda^2)| \langle\Omega| \phi_H(0) |\lambda_0\rangle |^2. \tag{4.541}$$

Here $z = | \langle\Omega| \phi_H(0) |1_{\boldsymbol{p}=0}\rangle |^2$ and it is known as **field-strength renormalization** or **wave-function renormalization**. For free fields $z = 1$ (try to show).

The quantity m is the **exact mass of a single particle** - exact energy eigenvalue at rest. m will in general differ from the value of mass parameter that appears in the Lagrangian.

Note on Mass

It is customary to refer to parameter in the Lagrangian as m_0 and call it **bare mass**. m is called the **physical mass** (or **renormalized mass**).

Using all of that we can write

$$\begin{aligned}\langle\Omega|\,\phi_H(x)\phi_H(y)\,|\Omega\rangle &= i\int_0^\infty d(M^2)D_t(x-y,M^2)\sigma(M^2)\\ &= izD_t(x-y,m^2)+i\int_{M_t^2}^\infty d(M^2)\sigma(M^2)D_t(x-y,M^2).\end{aligned} \tag{4.542}$$

M_t^2 is known as multiparticle threshold and $M_t^2 \approx 4m^2$.

From this we can get the **full Pauli-Jordan function** as

$$\langle\Omega|\,[\phi_H(x),\phi_H(y)]\,|\Omega\rangle = iz\Delta(x-y,m^2)+i\int_{M_t^2}^\infty d(M^2)\sigma(M^2)\Delta(x-y,M^2). \tag{4.543}$$

With $i\Delta(x-y,M^2) = \langle 0|\,[\phi(x),\phi(y)]\,|0\rangle$. Here $\phi(x)$ is a **free field** with mass M.

To understand some further properties of z we apply $\frac{\partial}{\partial y_0}\Big|_{x_0=y_0}$ to the full Pauli-Jordan function. For scalar field we use the fact that $\dot{\phi}=\pi$ and obtain

$$[\phi(x^0,\boldsymbol{x}),\pi(x^0,\boldsymbol{y})] = i\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y}). \tag{4.544}$$

Both for the interacting and free fields. Thus we get

$$i\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y}) = zi\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y})+i\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y})\int_{M_t^2}^\infty d(M^2)\sigma(M^2). \tag{4.545}$$

$z-1$ accounts for the overlap of $\phi(0)\,|\Omega\rangle$ with multiparticle states.

This implies that

$$1 = z+\int_{M_t^2} d(M^2)\sigma(M^2) \tag{4.546}$$

which means that $z \le 1$ (and particularly $z=1$ for free theory).

Finally, in complete analogy, we can derive spectral expansion for

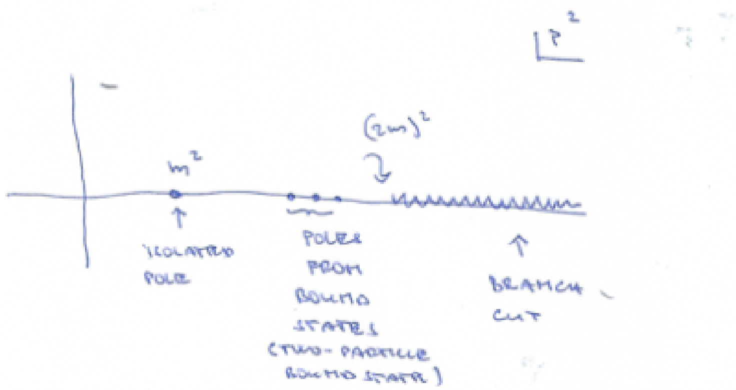
$$\langle\Omega|\,T\,[\phi(x)\phi(y)]\,|\Omega\rangle = iz\Delta_F(x-y,m^2)+i\int_{M_t^2}^\infty d(M^2)\sigma(M^2)\Delta_F(x-y,M^2). \tag{4.547}$$

Spectral representations for $\langle\Omega|\,\phi(x)\phi(y)\,|\Omega\rangle$,$\langle\Omega|\,[\phi(x),\phi(y)]\,|\Omega\rangle$ and $\langle\Omega|\,T\,[\phi(x)\phi(y)]\,|\Omega\rangle$ are known as Källen-Lehmann representations.

In momentum space we can thus write

$$\tau(p) = \frac{iz}{p^2-m^2+i\varepsilon}+i\int_{M_t^2} d(M^2)\sigma(M^2)\frac{1}{p^2-M^2+i\varepsilon}. \tag{4.548}$$

The analytic structure of this function can be seen on Fig. **??**



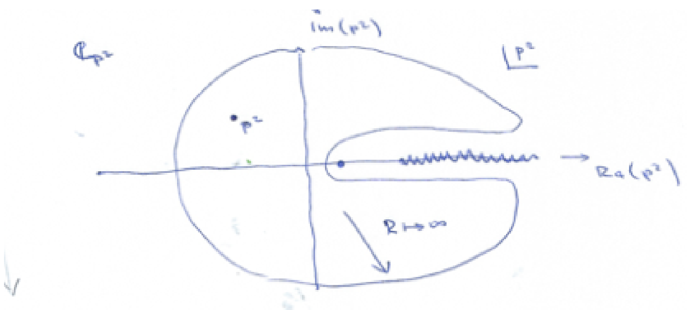
The Källen-Lehmann dispersion relation has also representation in terms of contour integral. For arbitrary analytic function $f(p^2)$ Cauchy's theorem states that

$$f(p^2) = \frac{1}{2\pi i} \oint_\gamma df \frac{f(s)}{s - p^2}. \tag{4.549}$$

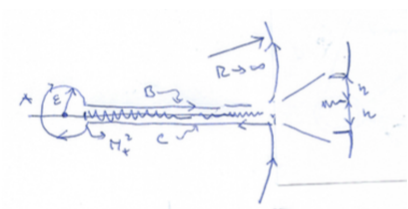
Provided that

- ▶ $p^2 \in \mathbb{C}$ is inside the contour γ.
- ▶ Contour γ does not cross any singularity.

Let us apply this to the z-point function $\tau(p)$ and use the knowledge of the analytic structure. We choose the contour γ as in Fig. ??.



Consider first that $f(p^2)$ has only branch cut but no poles. We further assume that $f(p^2)$ falls of rapidly enough so that the contribution from the large radius circle can be neglected moreover the contribution "A" from the figure be llow goes to zero as $\varepsilon \to 0$. Indeed



$$\int_{\gamma_A} ds \frac{f(s)}{s-z} = \lim_{\varepsilon \to 0} i\varepsilon \int_0^{2\pi} d\varphi \frac{f(M_t^2 + \varepsilon + \varepsilon e^{i\varphi})}{M_t^2 - \varepsilon + \varepsilon e^{i\varphi} - z} \tag{4.550}$$

Note that $\lim_{\varepsilon \to 0} |\cdots| \le \lim_{\varepsilon \to 0} \varepsilon \int_0^{2\pi} d\varphi \frac{|f(\cdots)|}{|M_t^2 - \varepsilon + \varepsilon^{i\varphi} - z|} = 0$. Here $|M_t^2 - \varepsilon + \varepsilon^{i\varphi} - z| = M_t^2 - z + O(\varepsilon)$.

Thus after the double limit $\varepsilon \to 0$ and $R \to \infty$ we are left with γ_C and γ_B.

$$\begin{aligned} f(z) &= \lim_{\eta \to 0} \frac{1}{2\pi i} \left\{ \int_{M_t^2 + i\eta}^{\infty + i\eta} ds \frac{f(s)}{s-z} - \int_{M_t^2 - i\eta}^{\infty - i\eta} ds \frac{f(s)}{s-z} \right\} \\ &= \lim_{\eta \to 0} \frac{1}{2\pi i} \left\{ \int_{M_t^2}^{\infty} ds \frac{f(s+i\eta)}{s+i\eta-z} - \int_{M_t^2}^{\infty} ds \frac{f(s-i\eta)}{s-i\eta-z} \right\}. \end{aligned} \tag{4.551}$$

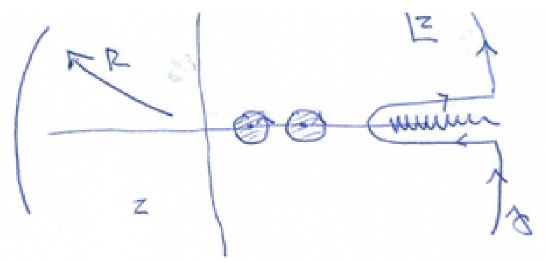
Since z is not on the cut, we can neglect $\pm i\eta$ in the denominators and write

$$f(z) = \frac{1}{2\pi i} \int_{M_t^2}^{\infty} ds \frac{f(s+i\eta) - f(si\eta)}{s-z}. \tag{4.552}$$

The numerator of the integrand is the discontinuity of $f(z)$ across the cut (denotes as disc $f(s)$). When $f(z)$ is real on the real axis except for a cut, then $f^*(z) = f(z^*)$ for $z \in \mathbb{R} \backslash$ branch cut. This relation is known as the **Schwartz reflection principle**.

$$f(z) = \frac{1}{\pi} \int_{M_t^2}^{\infty} ds \frac{\text{Im}\, f_+(s)}{s-z}. \tag{4.553}$$

If $f(z)$ has simple poles z_k, $k = 1, \ldots$ (for us are relevant poles $z_k \in \mathbb{R}^+$) then $f(z)$ is analytic inside of the curve.



$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{\gamma} ds \frac{f(s)}{s-z} \\ &= \frac{1}{2\pi i} \int_{\Leftrightarrow} - \frac{1}{2\pi i} \int_{\rightarrow} + \frac{1}{2\pi i} \sum_k \int_{\circ\gamma_k} ds \frac{f(s)}{s-z} \\ &= \frac{1}{\pi} \int_{M_t^2}^{\infty} ds \frac{\text{Im}\, f_+(s)}{s-z} + \frac{1}{2\pi i} \sum_k \text{Res}\, f(z_k) \int_{\circ\gamma_k} \frac{ds}{(s-z)(s-z_k)} \\ &= \frac{1}{\pi} \int_{M_t^2}^{\infty} ds \frac{\text{Im}\, f_+(s)}{s-z} + \sum_k \frac{\text{Res}\, f(z_k)}{z_k - z}. \end{aligned} \tag{4.554}$$

Now by setting $z = p^2 + i\eta$ ($\eta \to 0$) we get

$$f_+(p^2) = \frac{1}{\pi}\int_{M_t^2}^{\infty} ds \frac{\operatorname{Im} f_+(s)}{s - p^2 - i\eta} + \sum_k \frac{\operatorname{Res} f(z_k)}{z_k - p^2 - i\eta}. \tag{4.555}$$

Comparing this with formula

$$\tau(p) = \frac{iz}{p^2 - m^2 + i\varepsilon} + i\int_{M_t^2} d(M^2)\sigma(M^2)\frac{1}{p^2 - M^2 + i\varepsilon} \tag{4.556}$$

we see that $i\tau(p) \equiv f_+(p^2)$, $s \equiv M^2$ and $z_k \equiv m_k$, where m_k is a mass of single particle state and bound states.

Thus we arrive at the following relations

$$\sigma(M^2) = \frac{\operatorname{Im}(i\tau(p^2 = M^2))}{\pi}, \tag{4.557}$$

$$Z \;=\; \operatorname{Res}(i\tau(p^2 = m^2)). \tag{4.558}$$

LSZ Reduction Formulas

Now we will relate time ordered correlation functions $\langle x_1 \dots x_n\rangle \equiv \langle\Omega|\,T\,[\varphi_H(x_1)\dots\varphi_H(x_n)]\,|\Omega\rangle$ to scattering amplitudes.

Let us denote $\alpha = \{p_1, \dots, p_n\}$ to be set of initial state momenta and $\beta = \{q_1, \dots, q_m\}$ to be set of momenta of outgoing particles. In scattering processes we are interested in scattering amplitudes

$$\langle\beta, \text{in}|\hat{S}|\alpha, \text{in}\rangle = \langle\beta, \text{out}|\alpha, \text{in}\rangle\ . \tag{4.559}$$

This can be rewritten as

$$\begin{aligned}\langle b, \text{out}|a^\dagger_{in}(p_1)|\alpha', \text{in}\rangle \;&=\; \langle b, \text{out}|a^\dagger_{in}(p_1)|\alpha', \text{in}\rangle\\ &-\; \langle b, \text{out}|a^\dagger_{out}(p_1)|\alpha', \text{in}\rangle\\ &+\; \langle b, \text{out}|a^\dagger_{out}(p_1)|\alpha', \text{in}\rangle\ .\end{aligned} \tag{4.560}$$

Here $\alpha' = \{p_2, \dots, p_n\}$ denotes set of momenta. But

$$\begin{aligned}a_{out}(p_1)\,|\beta, \text{out}\rangle \quad &= a_{out}(p_1)\,|q_1, \dots, q_m, \text{out}\rangle\\ &= a_{out}(p_1)a^\dagger_{out}(q_1)\dots a^\dagger_{out}(q_m)\,|0\rangle_{out}\\ &= \left[a_out(p_1), a^\dagger_{out}(q_1)\dots a^\dagger_{out}(q_m)\right]|0\rangle_{out}\\ &= \sum_{j=1}^{m}(2\pi)^3 2\omega_{p_j}\delta(\boldsymbol{q}_j - \boldsymbol{p}_j)\,|\beta'_j, \text{out}\rangle\,,\end{aligned} \tag{4.561}$$

where $\beta'_j \;=\; \{q_1, q_2, \dots, q_{j-1}, q_{j+1}, \dots, q_m\}$. From the mode expansion for free scalar field, i.e.

$$\begin{aligned}\phi(\boldsymbol{x},t) &= \frac{1}{(2\pi)^3}\int \frac{d^3\boldsymbol{p}}{2\omega_{\boldsymbol{p}}}\left[a(p)e^{-ipx}+a^\dagger(p)e^{ipx}\right] \\ &= \sum_{\boldsymbol{p}}\left[a_{\boldsymbol{p}}e^{-ipx}+a^\dagger_{\boldsymbol{p}}e^{ipx}\right] \\ &= \sum_{\boldsymbol{p}}\left[a_{\boldsymbol{p}}e^{-i\omega_p t}+a^\dagger_{-\boldsymbol{p}}e^{i\omega_p t}\right]e^{i\boldsymbol{px}}. \end{aligned} \tag{4.562}$$

$$\begin{aligned}\pi(\boldsymbol{x},t)=\partial_0\phi(\boldsymbol{x},t) &= \sum_{\boldsymbol{p}}\left[-i\omega_{\boldsymbol{p}}a_{\boldsymbol{p}}e^{-ipx}+i\omega_{\boldsymbol{p}}a^\dagger_{\boldsymbol{p}}e^{ipx}\right] \\ &= \sum_{\boldsymbol{p}}\left[-i\omega_{\boldsymbol{p}}a_{\boldsymbol{p}}e^{-i\omega_p t}+i\omega_{\boldsymbol{p}}a^\dagger_{-\boldsymbol{p}}e^{i\omega_p t}\right]e^{ixp} \end{aligned} \tag{4.563}$$

This might be inverted in favour of $a_{\boldsymbol{p}}$ and $a^\dagger_{\boldsymbol{p}}$. In fact, the inverse Fourier transform directly gives

$$\begin{aligned}a(\boldsymbol{p}) &= \int d^3\boldsymbol{x}e^{ipx}\left[\omega_{\boldsymbol{p}}\phi(\boldsymbol{x},t)+i\pi(\boldsymbol{x},t)\right] \\ &+i\int d^3\boldsymbol{x}e^{ipx}\overset{\leftrightarrow}{\partial}_0\phi(\boldsymbol{x},t). \end{aligned} \tag{4.564}$$

Here $u\overset{\leftrightarrow}{\partial}v = u(\partial v)-(\partial u)v$. By Hermitan conjugation we also get

$$a^\dagger(\boldsymbol{p}) = -i\int d^3\boldsymbol{x}e^{-ipx}\overset{\leftrightarrow}{\partial}_0\phi(\boldsymbol{x},t). \tag{4.565}$$

With this, we can write

$$\begin{aligned}&\langle\beta,\text{out}|\alpha,\text{in}\rangle-\langle\beta,\text{out}|a^\dagger_{out}(\boldsymbol{p_1})|\alpha',\text{in}\rangle \\ &= \langle\beta,\text{out}|a^\dagger_{in}(\boldsymbol{p}_1)|\alpha',\text{in}\rangle-\langle\beta,\text{out}|a^\dagger_{out}(\boldsymbol{p}_1)|\alpha',\text{in}\rangle \\ &= -i\lim_{t\to-\infty}\int d^3\boldsymbol{x}_1e^{-ip_1x_1}\overset{\leftrightarrow}{\partial}_0\langle\beta,\text{out}|\phi_{in}(\boldsymbol{x},t)|\alpha',\text{in}\rangle \\ &+i\lim_{t\to\infty}\int d^3\boldsymbol{x}_1e^{-ip_1x_1}\overset{\leftrightarrow}{\partial}_0\langle\beta,\text{out}|\phi_{in}(\boldsymbol{x},t)|\alpha',\text{in}\rangle\,. \end{aligned} \tag{4.566}$$

Now we use the identity

$$\left(\lim_{t\to\infty}-\lim_{t\to-\infty}\right)\int d^3\boldsymbol{x}f(x)\overset{\leftrightarrow}{\partial}_0g(x)=\int d^4x\partial_0(f(x)\overset{\leftrightarrow}{\partial}_0g(x)=\int d^4x(f\partial_0^2g-g\partial_0^2f). \tag{4.567}$$

By setting $f(x) = e^{-ipx}$ we have

$$\int d^4x(f\partial_0^2 g - g\partial_0^2 f) = \{\partial_0^2 f = -p_0^2 f = (-\boldsymbol{p}^2 - m^2)f = (\nabla^2 - m^2)f\}$$
$$= \int d^4x(f\partial_0^2 g - g\nabla^2 f + gm^2 f) = \int d^4x f(\partial_0^2 - \nabla^2 + m^2)g$$
$$= \int d^4x f(\Box + m^2)g. \tag{4.568}$$

So, this allows us to write

$$\langle\beta, \text{out}||\alpha, \text{in}\rangle - \langle\beta, \text{out}|a^\dagger_{out}(\boldsymbol{p}_1)|\alpha', \text{in}\rangle$$
$$= \lim_{t\to\infty} \frac{i}{\sqrt{z}} \int d^3\boldsymbol{x}_1 e^{-ip_1x_1} \overleftrightarrow{\partial_0} \langle\beta, \text{out}|\phi_H(\boldsymbol{x}, t)|\alpha', \text{in}\rangle$$
$$- \lim_{t\to-\infty} \frac{i}{\sqrt{z}} \int d^3\boldsymbol{x}_1 e^{-ip_1x_1} \overleftrightarrow{\partial_0} \langle\beta, \text{out}|\phi_H(\boldsymbol{x}, t)|\alpha', \text{in}\rangle$$
$$= \frac{i}{\sqrt{z}} \int d^4x_1 e^{-ip_1x_1}(\Box + m^2) \langle\beta, \text{out}|\phi_H(\boldsymbol{x}, t)|\alpha', \text{in}\rangle\,. \tag{4.569}$$

Mass m is on-shell asymptotic (measured - hence physical) mass. It enters through in/out-states in S-matrix. Originally we have introduced the limits

$$\phi_{in}(x) \rightarrow_{t\to-\infty} \phi_H(x) \Leftrightarrow_{t\to\infty} \phi_{out}(x). \tag{4.570}$$

However, this naive assumption is actually incorrect. If we take this "strong" operatorial assumption, then it can be shown that the S-matrix becomes trivial and no scattering takes place. For this reason, Lehmann-Symanzik-Zimmerman (LSZ) proposed as the form the asymptotic condition

$$\lim_{t\to\infty} \langle\psi_1|\phi_H(x)|\psi_2\rangle \quad =_{in} \langle\psi_1|\phi_{out}(\boldsymbol{x}, \infty)|\psi_2\rangle_{out}$$
$$\lim_{t\to-\infty} \langle\psi_1|\phi_h(x)|\psi_2\rangle \quad =_{in} \langle\psi_1|\phi_{in}(\boldsymbol{x}, -\infty)|\psi_2\rangle_{in}\,, \tag{4.571}$$

for all states ψ_1 and ψ_2.

Convergence Issues

Convergence only in this **weak operatorial** sense is not strong enough to ensure that the limit of a product is the product of the limits consequentially. It is generally not true that the limit of a commutator of the commutator of the limits, i.e.

$$\lim_{t\to\infty} \langle\psi_1|[\phi_H(x), \phi_H(y)]|\psi_2\rangle \neq_{out} \langle\psi_1|[\phi_{out}(\boldsymbol{x}, \infty), \phi_{out}(y, \infty)|\psi_2\rangle_{out}\,. \tag{4.572}$$

On the other hand we know that the Heisenberg field $\phi_H(x)$ that we used in deriving Lehman-Kälen representation and $\phi_{in/out}$ fields obey canonical commutation relations. At the same time

$$\begin{aligned}\langle\Omega|\phi_H(x)|p\rangle &= e^{-ipx}\,\langle\Omega|\phi_H(0)|p\rangle \\ &= \sum_q \langle 0|(a(q)e^{-iqx}+a^\dagger(q)e^{iqx})a^\dagger(p)|0\rangle \\ &= \sum_q e^{-iqx}\,\langle 0|a(q)a^\dagger(p)|0\rangle \\ &= \sum_q e^{-iqx}\,\langle 0|[a(q),a^\dagger(p)]|0\rangle \\ &= e^{iqx}. \end{aligned} \tag{4.573}$$

So, Lehmann-Källen representation implies that

$$\begin{aligned}\lim_{t\to\infty}\langle\Omega|\phi_H(x)|p\rangle &=_{in}\langle 0|\phi_{out}(\boldsymbol{x},\infty)|p\rangle_{out}\, z^{1/2} \\ \lim_{t\to-\infty}\langle\Omega|\phi_h(x)|p\rangle &=_{in}\langle 0|\phi_{in}(\boldsymbol{x},-\infty)|p\rangle_{in}\, z^{1/2}. \end{aligned} \tag{4.574}$$

By denoting the interpolation Heisenberg field as $\tilde{\phi}_H$ and the Lehman-Källen Heisenberg field as $\phi_H(x)$, we see that the weak relation must hold

$$\begin{aligned}\phi_H(x)=\sqrt{z}\tilde{\phi}_H(x) &\to_{t\to+\infty}\sqrt{z}\phi_{out}(x,\infty) \\ &\to_{t\to-\infty}\sqrt{z}\phi_{in}(x,-\infty). \end{aligned} \tag{4.575}$$

So, we might write that

$$\langle\beta,\mathrm{out}|\alpha,\mathrm{in}\rangle = \frac{i}{\sqrt{z}}\int d^4x_1 e^{-ip_1x_1}(\Box+m^2)\,\langle\beta,\mathrm{out}|\phi_H(x_1,t_1)|\alpha,\mathrm{in}\rangle + \text{disconnected term}. \tag{4.576}$$

We have neglected the term $\langle\beta,\mathrm{out}|a^\dagger_{out}(p_1)|\alpha',\mathrm{in}\rangle$. In fact, this term can be written as

$$\begin{aligned}&\langle q_1,\ldots,q_n,\mathrm{out}|a^\dagger_{out}(p_1)|p_1,\ldots,p_m,\mathrm{in}\rangle \\ &= \langle 0|a_{out}(q_1)\ldots a_{out}(q_n)a^\dagger_{out}(p_1)|p_1,\ldots,p_m,\mathrm{in}\rangle \\ &= \langle 0|[a_{out}(q_1)\ldots a_{out}(q_n),a^\dagger_{out}(p_1)]|p_1,\ldots,p_n,\mathrm{in}\rangle \\ &= \sum_{i=1}^n 2\omega_{q_i}(2\pi)^3\delta(\boldsymbol{q}_i-\boldsymbol{p}_i)\,\langle q_1,\ldots,q_{i-1},q_{i+1}\ldots,q_n,\mathrm{out}|p_2,\ldots,p_m,\mathrm{in}\rangle\,. \end{aligned} \tag{4.577}$$

This is **disconnected term** because one particular energy unaffected by the scattering process. We have now completed the **first step** in the LSZ program. We can now proceed further, by defining $\beta' = \{q_2,q_2,\ldots,q_m\}$.

We get

$$\begin{aligned}
&\langle\beta,\mathrm{out}|\hat{\phi}_H(x_1)|\alpha',\mathrm{in}\rangle - \langle\beta',\mathrm{out}|\phi_H(x_1)a_{in}(q_1)|\alpha',\mathrm{in}\rangle \\
&= \langle\beta',\mathrm{out}|a_{out}\phi_H(x_1)|\alpha',\mathrm{in}\rangle - \langle\beta',\mathrm{out}|\phi_H(x_1)a_{in}(q_1)|\alpha',\mathrm{in}\rangle \\
&= i\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|\phi_{out}(y_1)\phi_H(x_1)|\alpha',\mathrm{in}\rangle \\
&-i\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|\phi_H(x_1)\phi_{in}(y_1)|\alpha',\mathrm{in}\rangle\,. \qquad (4.578)
\end{aligned}$$

Since the relation between $a_{in/out}$ and $\phi_{in/out}$ is true for any time argument, we can again rewrite the later identity as

$$\begin{aligned}
&i\lim_{t_y\to\infty}\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|\phi_{out}(y_1)\phi_H(x_1)|\alpha',\mathrm{in}\rangle \\
&-i\lim_{t_y\to-\infty}\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|\phi_H(x_1)\phi_{in}(y_1)|\alpha',\mathrm{in}\rangle \\
&= \frac{i}{\sqrt{z}}\lim_{y_0\to\infty}\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|\phi_H(y_1)\phi_H(x_1)|\alpha',\mathrm{in}\rangle \\
&-\frac{i}{\sqrt{z}}\lim_{y_0\to-\infty}\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|\phi_H(x_1)\phi_H(y_1)|\alpha',\mathrm{in}\rangle\,.
\end{aligned} \qquad (4.579)$$

Clearly, some tricks is needed in order to rewrite the integral as the four-dimensional integral. Time ordering is what does the job. Note, that previous identity can be written as

$$\begin{aligned}
&\frac{i}{\sqrt{z}}\lim_{y_0\to\infty}\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|T[\phi_H(y_1)\phi_H(x_1)]|\alpha',\mathrm{in}\rangle \\
&-\frac{i}{\sqrt{z}}\lim_{y_0\to-\infty}\int d^3\boldsymbol{y}_1 e^{iq_1y_1}\overleftrightarrow{\partial_0}\,\langle\beta',\mathrm{out}|T[\phi_H(y_1)\phi_H(x_1)]|\alpha',\mathrm{in}\rangle \\
&= \frac{i}{\sqrt{z}}\left(\lim_{y_0\to\infty} - \lim_{y_0\to-\infty}\right)\int d^3\boldsymbol{y}_1\cdots \\
&= \frac{i}{\sqrt{z}}\int d^4y_1 e^{iq_1y_1}\left(\Box_{y_1}+m^2\right)\langle\beta',\mathrm{out}|T[\phi_H(y_1)\phi_H(x_1)]|\alpha',\mathrm{in}\rangle\,.
\end{aligned} \qquad (4.580)$$

So, once two particles have been reduced the element of the S-matrix looks like

$$\begin{aligned}
&\langle\beta,\mathrm{out}|\alpha,\mathrm{in}\rangle = \langle\beta,\mathrm{in}|\S|\alpha,\mathrm{in}\rangle \\
&= \text{Disconnected terms} \\
&+\left(\frac{i}{\sqrt{z}}\right)^2\int d^4x_1 d^4y_1 e^{iq_1y_1-ip_1x_1}\left(\Box_{y_1}+m^2\right)\left(\Box_{x_1}+m^2\right) \\
&\times\langle\beta',\mathrm{out}|T[\phi_H(y_1)\phi_H(x_1)]|\alpha',\mathrm{in}\rangle\,. \qquad (4.581)
\end{aligned}$$

Disconnected part here involve one or two $\delta^{(3)}$ functionals. The same reasoning can be now carried further until all incoming and outgoing particles have been reduced

$$
\begin{aligned}
&\langle \beta, \text{out}|\alpha, \text{in}\rangle = \langle q_1, \dots, q_m, \text{out}|p_1, \dots, p_n, \text{in}\rangle \\
&= \text{Disconnected terms} \\
&+ \left(\frac{i}{\sqrt{z}}\right)^{n+m} \int d^4y_1 \dots d^4y_m d^4x_1 \dots d^4x_n \exp\left(i\sum_{k=1}^{m} q_k y_k - i\sum_{k=1}^{n} p_k x_k\right) \\
&\times \left(\Box_{y_1} + m^2\right) \cdots \left(\Box_{y_m} + m^2\right) \left(\Box_{x_1} + m^2\right) \cdots \left(\Box_{x_n} + m^2\right) \\
&\times \langle \Omega| T[\phi_H(y_1) \dots \phi_H(x_n)|\Omega\rangle \, .
\end{aligned}
\tag{4.582}
$$

In the last line we passed from $|0\rangle_{in}$ to $|\Omega\rangle$ by using the weak limit, namely by denoting

$$
\langle \psi| \equiv \langle \cdots, \text{out}|\, T[\phi_H(y_1) \dots \phi_H(x_{n-1})] \tag{4.583}
$$

we have

$$
\lim_{t_{x_n} \to -\infty} \langle \phi|\phi_{in}(x_1)|0\rangle_{in} = \lim_{\tau_{x_n} \to -\infty} \langle \psi|\phi_H(x_n)|\Omega\rangle \, \frac{1}{\sqrt{z}}. \tag{4.584}
$$

Expression (4.582) provides the relation between the on-shell transition amplitudes and the general Green functions. This relation is known as **LSZ Reduction Formula** and it implies, in particular, that in the momentum space the Green functions have poles in the variables p_i^2 (p_i are conjugates to x_i) so, up to a normalization constant the S**-matrix elements are nothing but the residue of the multi-pole structure of full Green function**. For general $n + m$ scattering process we still use the same $n + m$ point full Green function. Green functions are more elementary than scattering amplitudes.

LSZ for Dirac Spinors

For Dirac spinors one can derive LSZ Reduction Formula along the same lines as for bosons. Due to extra indices and anticommutativity, the derivation is more involved.

4.18 Cross Section

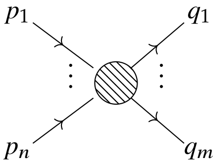
Define

$$
\tilde{\tau}(p_1, \dots, p_n, -q_1, \dots, -q_m) = \left(\prod_i \int d^4y_i\right)\left(\prod_j \int d^4x_j\right) e^{i\sum_i p_i y_j - \sum_j q_j x_j} \tau(x_1, \dots, x_n, y_1, \dots \tag{4.585}
$$

LSZ formula in the momentum space reads

$$\begin{aligned} S_{fi} &= \langle \{f\}, \text{in}|S|\{i\}, \text{in}\rangle = \langle \{f\}, \text{out}|\{i\}, \text{in}\rangle \\ &= \lim_{p^2, q^2 \to m_P^2} \left(\prod_l \frac{1}{\sqrt{z}} \frac{(p_l^2 - m_P^2)}{i} \right) \left(\prod_j \frac{1}{\sqrt{z}} \frac{(q_l^2 - m_P^2)}{i} \right) \\ &\times \underbrace{\tilde{\tau}(p1, \ldots, p_n, -q_1, \ldots, -q_m)}_{(2\pi)^4 \delta\left(\sum_{i=1}^n p_i - \sum_{j=1}^m q_j\right) \tau(p_1, \ldots, p_n, -q_1, \ldots, q_m)} \quad + \text{disconnected term.} \end{aligned} \tag{4.586}$$

More explicitly by $\langle \{f\}, \text{out}|\{i\}, \text{in}\rangle$ we mean $\langle q_1, \ldots, q_m, \text{out}|p_1, \ldots, p_n \text{in}\rangle$



(4.587)

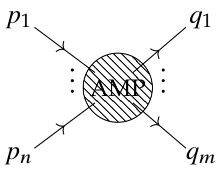
Presence of the terms $(p_l^2 - m_P^2), \ldots, (q_l^2 - m_P^2)$ causes that the external lines in the diagrams contribution to the τ **are amputated**. We speak about **amputated Green function** in the LSZ formula. So, for instance, for a scattering of 2 particles to 2 particles

$$\langle q_1, q_2, \text{out}|p_1, p_2, \text{in}\rangle = \frac{(-i)^4}{(\sqrt{z})^4} \lim_{p_i^2, q_j^2 \to m_P^2} \prod_{i=1}^{2} \left(p_i^2 - m_P^2\right) \prod_{i=1}^{2} \left(q_i^2 - m_P^2\right) \tilde{\tau}(p_1, p_2, -q_1, -q_2). \tag{4.588}$$

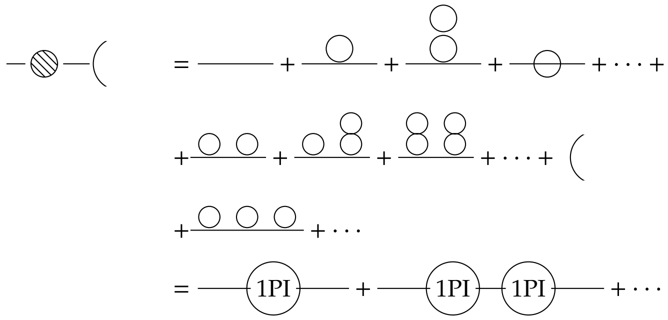
For a general diagram with external legs, we define **amputation** in the following way: start from the tip of each external leg, find the last point at which the diagram can be cut by removing a single propagator, such that this operation separates the leg from the rest of the diagram. Cut there. Consequently we can graphically write that

$$\langle q_1, \ldots, q_m, \text{out}|p_1, \ldots, p_n, \text{in}\rangle$$

$$= \left(\sqrt{z})\right)^{n+m} \tag{4.589}$$

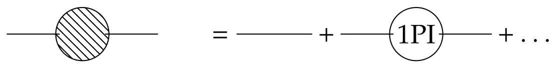


Here the circle AMP is the sum of amputated $n + m$-point diagrams and z is the field strength renormalization factor. The fact that we have factor $(\sqrt{z})^{n+m}$ and not $\left(\frac{1}{\sqrt{z}}\right)^{m+n}$ (as one could expect) requires some explanation. Before truncation the external propagator has the structure:



(4.590)

Where 1PI stands for all **1-particle irreducible diagrams**, i.e. diagrams that cannot be split to two by removing/cutting one line. By denoting the value of 1PI as $-i\Sigma(p^2)$ (also known as 1PI self-energy), we see that



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

$$= \text{Geometric Series} = \frac{i}{p^2 - m_0^2 - \Sigma(p^2) + i\varepsilon}. \tag{4.591}$$

Here m_0 is a mass parameter in Lagrangian, not a physical mass m_p. If we now power expand the self-energy correction around $p^2 = m_p^2$, we get

$$\Sigma(p^2) = \Sigma(m_p^2) + (p^2 - m_p^2)\Sigma'(m_p^2) + \underbrace{\tilde{\Sigma}(p^2)}_{\sim O((p^2 - m_p^2)^2)} . \tag{4.592}$$

This implies

$$\frac{i}{p^2 - m_0^2 - \Sigma(p^2) + i\varepsilon} = \frac{i}{p^2 - m_0^2 - \Sigma(p_m^2) - (p^2 - m_p^2)\Sigma'(m_p^2) - \tilde{\Sigma} + i\varepsilon}$$

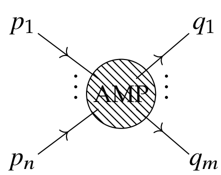
$$= \frac{i}{p^2 - \left[m_0^2 + \Sigma(m_p^2)\right] - (p^2 - m_p^2)\Sigma'(m_p^2) - \tilde{\Sigma}(p^2) + i\varepsilon}$$

$$= \frac{1}{(p^2 - m_p^2)(1 - \Sigma'(m_p^2)) - \tilde{\Sigma}(p^2) + i\varepsilon}$$

$$= \frac{z}{p^2 - m_p^2 + i\varepsilon} + \text{regular terms}. \tag{4.593}$$

So by cutting external leg it will survive factor iz (from denominator of Green's function). On the other hand, LSZ reduction formula multiplies each external leg with the factor $\frac{-i}{\sqrt{z}}$. So, consequently

$$\langle q_1, \dots, q_m, \text{out}|p_1, \dots, p_n, \text{in}\rangle = \left(\sqrt{z}\right)^{n+m}$$



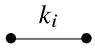
. (4.594)

In fact, all amputated Green functions should be only connected Green functions, because in LSZ we discard disconnected scattering.

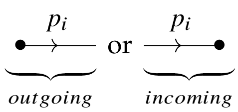
By knowing Feynman rules for Green functions, we can now directly write Feynman rules (in momentum space) for the elements of the S-matrix $\langle q_1, \ldots, q_m, \text{out}|p_1, \ldots, p_n, \text{in}\rangle$.

- Draw all topologically distinct connected diagrams with $n+m$ external lines with incoming momenta considered as positive and outgoing momenta considered as negative.
- For internal propagator assign

$$\frac{i}{k_i^2 - m^2 + i\varepsilon} \sim \tag{4.595}$$



- To each vertex assign vertex factor (i.e. $(-i\lambda)$ for ϕ^4 theory and $(-ig)$ for ϕ^3 theory).
- To each external propagator

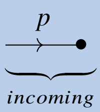


(4.596)

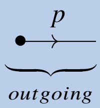
affiliate the factor $\sqrt{z}$.
- Impose momenta conservation at each vertex.
- Integrate over undetermined loop momenta $\int \frac{d^4p}{(2\pi)^4}$.
- Divide by the symmetry factors.

Fermionic propagators and LSZ formula

For Fermions the LSZ reduction formula prescribes that to each external line we affiliate, for particles

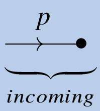


$$\sim \sqrt{z}u(p) \tag{4.597}$$

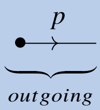


$$\sim \sqrt{z}\bar{u}(p) \tag{4.598}$$

and for antiparticles



$$\sim -\sqrt{z}\bar{v}(p) \tag{4.599}$$



$$\sim -\sqrt{z}v(p) \tag{4.600}$$

The factor z is irrelevant for calculations of the leading order of perturbation theory, but are important in the calculation of higher-order corrections.

4.19 Cross-Section - Practical Part

The S-matrix has generically the structure

$$S = 1 + iT. \tag{4.601}$$

T here is the so called T-matrix and it contains the information on the interaction

$$\langle f|S|i\rangle = \delta_{fi} + i\,\langle f|T|i\rangle\,. \tag{4.602}$$

Here δ_{fi} symbolically represents the particles not interacting at all and $\langle f|T|i\rangle$ is represented by LSZ. Thus

$$\begin{aligned} i\,\langle f|T|i\rangle \quad &= i\left(\sqrt{z}\right)^{n+m}\tilde{\tau}(p_1,\ldots,p_n,-q_1,\ldots,-q_m)_{amp} \\ &= i(2\pi)^4\left(\sqrt{z}\right)^{n+m}\delta\left(\sum p_i - \sum q_i\right)\tau(p_1,\ldots,p_m,-q_1,\ldots,-q_n)_{amp} \\ &= i(2\pi)^4\delta(p_i - q_f)T_{fi}. \end{aligned} \tag{4.603}$$

In the following we will consider only $f \neq i$ which implies that $\delta_{fi} = 0$. Then the probability of making the transition $i \to f$ is

$$|\,\langle f|T|i\rangle\,|^2 = \left((2\pi)^4\left[\delta(p_i - p_f)\right]\right)^2 |T_{fi}|^2. \tag{4.604}$$

What sense we should make of the square of a δ-function? We can proceed heuristically and use the fact that $\delta(x)f(x) = delta(x)f(0)$ for any function $f(x)$ and take $\delta^2(x) = \delta(0)\delta(x)$. By using the Fourier representation of δ-function we have

$$\delta^{(4)}(0) = \delta^{(4)}(p=0) = \int \frac{d^4x}{(2\pi)^4}e^{ix\cdot 0} = \frac{VT}{(2\pi)^4}. \tag{4.605}$$

Here V is the volume of the universe and T is the time of the universe duration. Thus we can proceed as

$$|\,\langle f|T|i\rangle\,|^2 = V\cdot T\cdot(2\pi)^4\,\delta^{(4)}(p_f - p_i)|T_{if}|^2. \tag{4.606}$$

This is nothing but Fermi's "Golden Rule" known from quantum mechanics.

In other words, the transition rate $i \to f$ per unit volume is

$$\Gamma_{if} = (2\pi)^4\,\delta^{(4)}(p_f - p_i)|T_{if}|^2. \tag{4.607}$$

If we restrict the rate of transition to some range of final states (i.e. final momenta of particles are not sharp but belong to some allowed range of values)

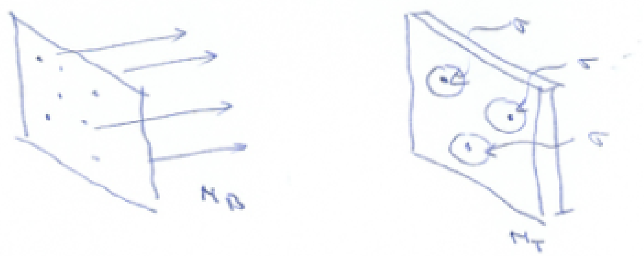
$$\Gamma_{if} = \sum_f (2\pi)^4\delta^{(4)}(p_f - p_i)|T_{if}|^2. \tag{4.608}$$

Now the connection between Γ_{if} and scattering cross-sections can be obtained in the following way.

There are many ways in which to define the **cross section**. The simplest and most intuitive is to define it as the effective size of each particle in the target.

Consider a thin target with N_T particles in it. Each particle has the effective area σ (cross-section). As seen from an incoming beam, the total amount of area taken by these particles is $N_T\sigma$. If we aim a beam of particles at the target with area A.

$$\text{Probability of hitting particle} = \frac{N_T\sigma}{A}. \tag{4.609}$$



Let the beam has N_B particles. Number of events is

$$N_B \times \text{Probability of hitting} = \frac{N_B N_T \sigma}{A}. \tag{4.610}$$

Thus

$$\sigma = \left(\frac{Number of Events}{N_B N_T}\right) A. \tag{4.611}$$

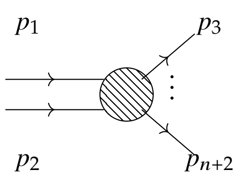
This can be rewritten in more expedient form. If beam is moving at velocity v towards a stationary target. The number of particles is $\rho_B V$. If the beam is a pulse that is turned fro t seconds $V = vtA$ and hence $N_B = \rho_b vtA$, which implies that

$$\begin{aligned} \sigma &= \frac{(\text{Number of Events})/t}{(\rho_b vtA)N_T/t} \\ &= \frac{(\text{Number of events})/t}{\rho_B v \rho_T V_T} \\ &= \frac{\text{Transition Rate}}{\rho_B v \rho_T V_T}. \end{aligned} \tag{4.612}$$

If N_B and $N_T = 1$, then transition rate $/V_T$ is Γ_{if} = Probability transition rate per unit volume. Thus

$$\sigma = \frac{1}{\rho_B \rho_T v}\Gamma_{if} = \frac{1}{\rho_B \rho_T v}\sum_f (2\pi)^4 \delta(p_f - p_i)|T_{if}|^2. \tag{4.613}$$

If we consider a final state of n distinct spinless particels



(4.614)

The cross-section is given by

$$\sigma_{2\to n} = \frac{1}{\rho_T \rho_T V}\int_\Delta \frac{d^3p_3}{(2\pi)^3 2\omega_{p_3}}\cdots\frac{d^3p_{n+2}}{(2\pi)^3 2\omega_{p_{n+2}}}(2\pi)^4\delta^{(4)}\left(p_1+p_2-p_3-\cdots p_{n+2}\right)|T_{if}|^2. \tag{4.615}$$

Here Δ is restricted range of observed momenta. In case we will have scattering of particle that are indistinguishable, we need to add $1/n!$ factor before integral.

We now use the fact that the velocity of particle $v = \frac{|\boldsymbol{p}|}{E_{\boldsymbol{p}}}$ and that with our relativistic normalization for the plane-wave states $|p\rangle$ the number of particles per unit volume (i.e. particle density in state with momenta p) is $2E_{\boldsymbol{p}} = 2\omega_{\boldsymbol{p}}$. Indeed, average number of particles in state $|p\rangle$ is

$$\begin{aligned}\langle p|\hat{N}|p\rangle &= \rho_p V = \langle p|p\rangle = \langle 0|a_p a_p^\dagger|0\rangle \\ &= \langle 0|[a_p, a_p^\dagger]|0\rangle = (2\pi)^3 2\omega_p \delta(0).\end{aligned} \tag{4.616}$$

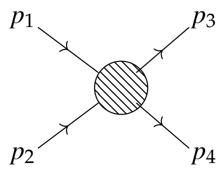
Thus $\rho_p = 2\omega_p = 2E_p$. If we take particle 1 to be at rest (target particle) we have $p_1 = (E_1, \mathbf{0})$, $p_2 = (E_2, \boldsymbol{p}_2)$, $v = v_2 = \frac{|\boldsymbol{p}_2|}{E_2}$, hence

$$\rho_T \rho_B v = 4m_1 E_2 \frac{|\boldsymbol{p}_2|}{E_2} = 4m_1|\boldsymbol{p}_2|. \tag{4.617}$$

Now we can finally rewrite cross section as

$$\sigma_{2\to n} = \frac{1}{4m_1|\boldsymbol{p}_2|}\frac{1}{n!}\int_\Delta \frac{d^3p_3}{(2\pi)^3 2\omega_{p_3}}\cdots\frac{d^3p_{n+2}}{(2\pi)^3 2\omega_{p_{n+2}}}(2\pi)^4\delta^{(4)}\left(p_1+p_2-p_3-\cdots p_{n+2}\right)|T_{if}|^2. \tag{4.618}$$

Important is the situation with **elastic scattering**



(4.619)

Since $(p_1+p_2)^2 = (E_1+E_2)^2 - (\boldsymbol{p}_1+\boldsymbol{p}_2)^2$ and since we are in CM system, the second term is zero.

In this context one introduces kinematic Lorentz invariant

$$s = (p_1+p_2)^2 = (p_3+p_4)^2, \tag{4.620}$$

which is nothing but the squared center-of-mass energy (i.e. $s > 0$).

Mandelstam Variables

s invariant is called **Mandelstam variables**. One introduces also another two Mandelstam variables

$$t = (p_1-p_3)^2 = (p_2-p_4)^2, \tag{4.621}$$

$$u = (p_1-p_4)^2 = (p_2-p_3)^2. \tag{4.622}$$

One can show that $s^2+t^2+u^2 = m_1^2+m_2^2+m_3^2+m_4^2$.

Since σ represents effective area of the scatterer, perpendicular to incident beam, it remains invariant under a Lorentz transformation to any other collinear frame (Lorentz contractions do not affect the size of any area provided boosts are in directions perpendicular to that area).

Cross section is not a true Lorentz invariant, since it transforms line area under arbitraty Lorentz transformation. Common collinear frames are the **laboratory frame** (one particle is in rest) and **CM frame**. Both frames are used in σ analysis.

Let us now analyze the elastic scattering $(m_1 + m_2 = m_3 + m_4)$ $p_1 + p_2 \to p_3 + p_4$ in laboratory frame. In this case

$$p_1 = (E_L, 0, 0, p_L), \quad p_2 = (m, 0, 0, 0) \tag{4.623}$$

and then

$$s = (E_L + m)^2 - p_L^2 = E_L^2 + 2mE_L + m^2 - p_L^2 = 2m^2 + 2mE_L = 2m(m + E_L). \tag{4.624}$$

And thus

$$E_L = \frac{s - 2m^2}{2m}. \tag{4.625}$$

Similarly $p_L^2 = E_L^2 - m^2 = \left(\frac{s-2m^2}{2m}\right)^2 - m^2 = \frac{s^2-4m^2s+4m^4-4m^4}{4m^2} = \frac{s^2-4sm^2}{4m^2}$.

So $p_L = \frac{\sqrt{s(s-4m^2)}}{2m}$ and so $s \geq 4m^2$ for the process to occur. $4m^2$ is the threshold value of s. This gives us

$$\sigma_{2\to n} = \frac{1}{\sqrt{s(s-4m^2)}} \frac{1}{2!} \int_\Delta \frac{d^3p_3}{(2\pi)^3 2\omega_{p_3}} \frac{d^3p_4}{(2\pi)^3 2\omega_{p_4}} (2\pi)^4 \delta^{(4)}\left(p_1 + p_2 - p_3 - p_4\right) |T_{if}|^2. \tag{4.626}$$

We can evaluate the cross-section in any collinear frame. We now shift to CM frame. The initial particles have 3-momenta $\boldsymbol{p}$ and $-\boldsymbol{p}$ and energy $\frac{\sqrt{s}}{2}$ and $\frac{\sqrt{s}}{2}$. Thus

$$p_1 = (\frac{\sqrt{s}}{2}, 0, 0, p), \quad p_2 = (\frac{\sqrt{s}}{2}, 0, 0, -p). \tag{4.627}$$

Since

$$p_1^2 = m^2 \implies \frac{s}{4} - p^2 = m^2 \implies p = \sqrt{\frac{s}{4} - m^2}. \tag{4.628}$$

So we get that $p_1 + p_2 = (\sqrt{s}, 0, 0, 0)$.

In computing $\sigma_{2\to 2}$ we need to evaluate

$$\int_\Delta \frac{d^3p_3}{(2\pi)^3 2E_{\boldsymbol{p}_3}} \frac{d^3p_4}{(2\pi)^3 2E_{\boldsymbol{p}_4}} (2\pi)^4 \delta\left(E_{\boldsymbol{p}_4} + E_{\boldsymbol{p}_3} - \sqrt{s}\right) \delta^{(3)}(\boldsymbol{p}_3 + \boldsymbol{p}_4) |T_{if}|^2. \tag{4.629}$$

We now extend $\int_\Delta$ to $\int_{\mathbb{R}^3}$. This provides the so-called **total cross-section**

$\sigma_{tot,2\to 2}$. The integral then simplifies to

$$\int_{\mathbb{R}^3} \frac{d^3p_3}{(2\pi)^6 4E_{\boldsymbol{p}_3}E_{\boldsymbol{p}_4}} (2\pi)^4 \, \delta\left(E_{\boldsymbol{p}_4} + E_{\boldsymbol{p}_3} - \sqrt{s}\right) = \int_{\mathbb{R}^3} \frac{d^3p_3}{(2\pi)^2 4E^2_{\mathbf{p}_3}} \delta\left(E_{\boldsymbol{p}_4} + E_{\boldsymbol{p}_3} - \sqrt{s}\right). \tag{4.630}$$

At this stage we use that $\boldsymbol{p}_4 = -\boldsymbol{p}_3$, which means that $E_{\boldsymbol{p}_3} = \sqrt{m^2 + |\boldsymbol{p}_3|^2}$ and $E_{\boldsymbol{p}_4} = \sqrt{m^2 + |\boldsymbol{p}_3|^2}$. Next, we denote $W \equiv E_{\boldsymbol{p}_3} + E_{\boldsymbol{p}_4} = 2\sqrt{m^2 + |\boldsymbol{p}_3|^2}$. Then

$$\begin{aligned}
\int_{\mathbb{R}^3} \frac{d^3p_3}{(2\pi)^2 2E^2_{\boldsymbol{p}_3}} \delta(W - \sqrt{s}) &= \int d\Omega(\boldsymbol{p}_3) d|\boldsymbol{p}_3| \frac{|\boldsymbol{p}_3|}{(2\pi)^2 2E^2_{\boldsymbol{p}_3}} \delta(W - \sqrt{s}) = \left\{ E_{\boldsymbol{p}_3} = E_{\boldsymbol{p}_4} = \frac{\sqrt{s}}{2} \right\} \\
&= \int \frac{d\Omega(\boldsymbol{p}_3)}{(2\pi)^2 s} \frac{|\boldsymbol{p}_3|^2}{\left|\frac{dW}{d|\boldsymbol{p}_3|}\right|} = \left\{ \frac{dW}{d|\boldsymbol{p}_3|} = \frac{2|\boldsymbol{p}_3|}{\sqrt{m^2 + |\boldsymbol{p}_3|^2}} = \frac{2|\boldsymbol{p}_3|}{E_{\boldsymbol{p}_3}} \right\} \\
&= \int \frac{d\Omega(\boldsymbol{p}_3)}{(2\pi)^2 s} \frac{|\boldsymbol{p}_3|}{2} \frac{\sqrt{s}}{2} = \int \frac{d\Omega(\boldsymbol{p}_3)}{(2\pi)^2 4} \frac{\sqrt{\frac{s}{4} - m^2}}{\sqrt{s}} \\
&= \int \frac{d\Omega(\boldsymbol{p}_3)}{16\pi^2} \frac{\sqrt{\frac{s}{4} - m^2}}{\sqrt{s}}
\end{aligned} \tag{4.631}$$

Thus total cross-section has form

$$\begin{aligned}
\sigma_{tot,2\to 2} &= \frac{1}{2\sqrt{s(s - 4m^2)}} \frac{\sqrt{s - 4m^2}}{16\pi^2 2! \sqrt{s}} \int d\Omega(\boldsymbol{p}_3) |T_{if}|^2 \\
&= \frac{1}{64\pi^2 s} \int d\Omega(\boldsymbol{p}_3) |T_{if}|^2.
\end{aligned} \tag{4.632}$$

And corresponding differential cross-section is

$$\frac{d\sigma_{tot,2\to 2}}{d\Omega(\boldsymbol{p}_3)} = \frac{1}{64\pi^2 s} |T_{if}|^2. \tag{4.633}$$

This is so-called differential cross-section of particles with identical (so also equal mass) particles. Formula works also for more general $2 \to 2$ elastic scatterings such as, e.h. $e^- e^+ \to \mu^- \mu^+$.

Note on Frames

$$\underbrace{\rho_1 \rho_2 |V_{12}|}_{\text{Lorentz invariant under collinear boosts}} = 4m_1 |\boldsymbol{p}_2| = 2\sqrt{s(s - 4m^2)}. \tag{4.634}$$

Generically it can be written in the form

$$4\sqrt{(p_1 p_2)^2 - p_1^2 p_2^2} = 4\sqrt{(p_1 p_2)^2 - m_1^2 m_2^2}. \tag{4.635}$$

Indeed in the Laboratory frame $p_1 = (E_1, \mathbf{0})$, $p_2 = (E_2, \boldsymbol{p}_2)$.

This then implies that

$$4\sqrt{E_1^2E_2^2 - E_1^2(E_2^2 - \boldsymbol{p}_2^2)} = 4\sqrt{E_1^2\boldsymbol{p}_2^2} = 4|\boldsymbol{p}_2|E_1 = 4\frac{|\boldsymbol{p}_2|}{E_2}E_2E_1 = V\rho_1\rho_2. \tag{4.636}$$

Recap

- $$\sigma = \frac{(\text{Number of Events})/t}{\rho_R V N_T} = \frac{\text{\# of events per unit time per unit target}}{\text{incident flux}}.$$

  (Here incident flux $= \rho_R V = (\rho_R V t A)/(tA) =$ # of incoming particles per unit time and per unit area).
- $$\sigma_{2\to n} = \sum_f \frac{(2\pi)^4|T_{fi}|^2\delta^{(4)}(p_f - p_i)}{4\sqrt{(p_1p_2)^2 - m_1^2m_2^2}}$$

  Which in differential form reads as

  $$d\sigma_{2\to n} = \frac{1}{4\sqrt{(p_1p_2)^2 - m_1^2m_2^2}}\frac{(2\pi)^4}{S}|T_{if}|^2\delta(p_i - p_f)d\tilde{p}_3\cdots d\tilde{p}_{n+2} \tag{4.637}$$

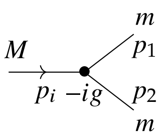
  Here $d\tilde{p} = \frac{d^3\boldsymbol{p}_i}{(2\pi)^3 2\omega_{\boldsymbol{p}_i}}$.
- $$\frac{d\sigma_{tot,2\to 2}}{d\Omega(\boldsymbol{p}_3)} = \frac{1}{64\pi^2 s}|T_{if}|^2. \text{ For identical particles.}$$

  Here $T_{fi} = \tau(p_1, \cdots, p_n, -q_1, \cdots, -q_m)_{amp}$.

As an exercise, let us use S-matrix elements in computation that describes **decay of unstable particle**. Assume

$$\mathcal{L}_I = -\frac{g}{2}\Phi(x)\varphi^2(x). \tag{4.638}$$

The corresponding vertex is



(4.639)

The initial state is a single unstable particle state $P_i = (M, \mathbf{0})$ in the rest frame of the unstable particle.

$$S_{fi} = \langle p_1, p_2|S|P_i\rangle = i(2\pi)^4\delta^{(4)}(p_1 + p_2 - P_i)T_{fi}. \tag{4.640}$$

In the lowest order in g we get $T_{fi} = (-i)^3(-ig) = g$.

Probability of transition is

$$P_{fi} = |S_{fi}|^2 = \sum_{p_1,p_2} \underbrace{(2\pi)^4\delta^{(4)}(0)}_{=V\cdot T}(2\pi)^4\delta^{(4)}(p_1 + p_2 - P_i)|T_{fi}|^2. \tag{4.641}$$

Rate of transition per unit volume and unit time is

$$\omega = \frac{P_{fi}}{VT} = (2\pi)^4 \sum_{p_1,p_2} \delta^{(4)}(p_1 + p_2 - P_i)|T_{fi}|^2. \tag{4.642}$$

Then rate per particle per unit of time is

$$\Gamma = \frac{\omega}{\rho} = \frac{P_{fi}}{\rho VT} = \frac{(2\pi)^4}{2M} \sum_{p_1,p_2} \delta^{(4)}(p_1 + p_2 - P_i)|T_{fi}|^2 . \tag{4.643}$$

Here $\rho = 2E$ which is $2M$ in the rest frame. Consequently, to the lowest order we have

$$\begin{aligned} \Gamma &= \frac{(2\pi)^4}{2M} g^2 \frac{1}{2!} \int \frac{d^3\boldsymbol{p}_1}{(2\pi)^3 2E_{\boldsymbol{p}_1}} \frac{d^3\boldsymbol{p}_2}{(2\pi)^3 2E_{\boldsymbol{p}_2}} \delta^{(4)}(p_1 + p_2 - P_i) \\ &= \frac{g^2}{(2\pi)^2} \frac{1}{4M} \int \frac{d^3\boldsymbol{p}_1}{4E_{\boldsymbol{p}_1}E_{\boldsymbol{p}_2}} \delta(E_{\boldsymbol{p}_1} + E_{\boldsymbol{p}_2} - M) . \end{aligned} \tag{4.644}$$

By using the fact that $\boldsymbol{p}_1 + \boldsymbol{p}_2 = 0$ (), which implies $E_{\boldsymbol{p}_1} = E_{\boldsymbol{p}_1} = \sqrt{m^2 + |\boldsymbol{p}_1|^2}$ we get

$$\begin{aligned} \int d^3\boldsymbol{p}_1 \delta(W - M) &= \left\{ W \equiv E_{\boldsymbol{p}_1} + E_{\boldsymbol{p}_2} = 2\sqrt{m^2 + |\boldsymbol{p}_1|^2} \right\} \\ &= \int d|\boldsymbol{p}_1||\boldsymbol{p}_1|^2 d\Omega(\boldsymbol{p}_1)\delta(W - M) \\ &= \int d\Omega(\boldsymbol{p}_1) \frac{dW}{\left|\frac{dW}{d|\boldsymbol{p}_1|}\right|} |\boldsymbol{p}_1|^2 \delta(W - M) \\ &= \left\{ \frac{dW}{d|\boldsymbol{p}_1|} = \frac{2|\boldsymbol{p}_1|}{E_{\boldsymbol{p}}} = \frac{2|\boldsymbol{p}_1|}{\frac{M}{2}} = \frac{4|\boldsymbol{p}_1|}{M} \right. \\ &\qquad \left. |\boldsymbol{p}_1| = \sqrt{E_{\boldsymbol{p}_1}^2 - m^2} = \sqrt{\frac{M^2}{4} - m^2} = \sqrt{\frac{M^2 - 4m^2}{2}} \right\} . \end{aligned} \tag{4.645}$$

Thus we get

$$\begin{aligned} \Gamma &= \frac{g^2}{4\pi^2} \frac{1}{4M} \int \frac{d\Omega}{\frac{4|\boldsymbol{p}_1|}{M}} \frac{|\boldsymbol{p}_1|^2}{4E_{\boldsymbol{p}_1}^2} = \frac{g^2}{4\pi^2} \frac{1}{4M} \int \frac{d\Omega}{8M^2} \sqrt{M^2 - 4m^2} \\ &= \frac{g^2}{32\pi M^2} \sqrt{M^2 - 4m^2}. \end{aligned} \tag{4.646}$$

Since density of unstable particle decays as $e^{-\Gamma t}$, the $\frac{1}{\Gamma}$ is the mean lifetime of a particle.