Der Beginn aller Wissenschaften
ist das Erstaunen, dass die Dinge
sind, wie sie sind.

Aristoteles

Electroweak Theory and LEP Physics

F. Jegerlehner *
DESY Zeuthen, Platanenallee 6, D-15738 Zeuthen, Germany

*Lectures given in the "Troisième Cycle de la Physique en Suisse Romande" at the Ecole Polytechnique Fédérale Lausanne (see `http://www-com.physik.hu-berlin.de/~fjeger/books.html`)

Contents

1 Natural Units

Standard physical units are the units of length (cm), mass (gr), time (sec) and charge (Cb). In particle physics it is customary and convenient to choose natural units such that the fundamental constants, the speed of light c, the reduced Planck constant $\hbar$ and the permittivity (electric field constant) ϵ_0 are unity : $c = \hbar = \epsilon_0 = 1$. This convention eliminates three of the standard units once one fundamental unit is chosen. The most commonly used unit is an energy, e.g., $[E] = 1$ GeV $= 10^9$ eV. Mass (m), momentum (mc) and energy (mc^2) are given in GeV. We note that

$$1 \text{ eV} \quad = \quad 1.6022 \cdot 10^{-12} \text{ erg} \,. \tag{1.1}$$

Using the experimental values for c, $\hbar$ and ϵ_0 in standard units one finds the **conversion factors**:

$$\begin{aligned} c &= 2.9979 \cdot 10^{10} \text{ cm s}^{-1} \\ \hbar &= 6.5821 \cdot 10^{-25} \text{ GeV s} \\ \hbar c &= 1.9733 \cdot 10^{-14} \text{ GeV cm} \\ (\hbar c)^2 &= 0.38938 \text{ GeV}^2 \text{ mbarn} \\ \epsilon_0 \hbar c &= 2.798 \cdot 10^{-37} \text{ Cb}^2 \end{aligned} \tag{1.2}$$

with 1 barn$=10^{-24}$ cm^2. Standard units are then given in energy units by:

$$\begin{aligned} 1 \text{ fermi} \ \ (\text{fm}) &= 10^{-13}\text{cm} = 5.0677 \text{ GeV}^{-1} \\ 1 \text{ s} &= 1.5193 \cdot 10^{24} \text{ GeV}^{-1} \\ 1 \text{ g} &= 5.6096 \cdot 10^{23} \text{ GeV} \\ 1 \text{ mbarn} \ \ (\text{mb}) &= 2.5682 \text{ GeV}^{-2} \\ 1 \text{ Cb} &= 1.89005 \cdot 10^{18} \end{aligned} \tag{1.3}$$

The charge magnitude of the electron is

$$e = 1.602177 \cdot 10^{-19} \text{ Cb} = 0.302822$$

and the fine structure constant

$$\alpha = \frac{e^2}{4\pi} = (137.036)^{-1}$$

Cross sections are usually obtained in GeV^{-2} and then converted to millibarns using

$$1 \text{ GeV}^{-2} \quad = \quad 0.38938 \text{ mbarn}$$

Hint: For updated precise values of **physical constants** contact the World-Wide-Web (WWW) page *http://www-pdg.lbl.gov/1996/contents_sports.html#constantsetc*

1.1 Exercises: Section 1

① Calculate the conversion factors using the values for c, $\hbar$ and ϵ_0 in standard units.

② The width of the Z boson has been measured at LEP (October 2001) to be

$$\Gamma_Z = (2.4952 \pm 0.0023) \text{ GeV}.$$

Calculate the Z lifetime $\tau_Z = \Gamma_Z^{-1}$ in seconds. What distance (in mm) does a Z particle travel before it decays (length of track in the detector) ? Hint: Use the velocity v (in units of c), which is determined by the magnitude of the momentum $|\vec{p}\,| = \frac{vM_Z}{\sqrt{1-v^2}} = \sqrt{E_{CM}^2 - M_Z^2}$. The distance traveled in the laboratory frame is then given by (Lorentz contraction!)

$$\ell_Z(E_b) = \frac{v}{\sqrt{1-v^2}} c\tau_Z \simeq \sqrt{\frac{E_{CM}^2}{M_Z^2} - 1} \ \times\ 7.9 \times 10^{-14} \ \text{ mm} \ .$$

The experimental value for the Z mass is

$$M_Z = (91.1875 \pm 0.0021) \text{ GeV}.$$

The Z is produced as a real (though unstable) particle provided $E_{CM} > M_Z$. Consider typical LEP energies $E_{CM} = M_Z + n\Gamma_Z$ for $n = 1, 2, 5$.

Use the Boltzmann constant $k = 8.6173 \cdot 10^{-5}$ eV °K^{-1} to evaluate the equivalent "temperature of a Z event" at LEP.

In nature such temperatures must have existed in our universe shortly after the big bang. In the early universe the time–temperature relationship (in the radiation dominated era) is given by

$$t = \frac{2.42}{\sqrt{N(T)}} \left(\frac{1\text{MeV}}{kT}\right)^2 \ sec \ .$$

where

$$N(T) = \sum_B g_B + \frac{7}{8} \sum_F g_F$$

counts the effectively massless ($m_i \ll kT$) degrees of freedom of bosons and fermions. Calculate at what time in the history of the universe the temperature of the universe was equivalent to the mass of the Z boson.

③ The QED cross section for μ–pair production in e^+e^- annihilation at high energies ($s \gg m_\mu^2$) is given by

$$\frac{d\sigma}{d\Omega}\left(e^+e^- \to \mu^+\mu^-\right) = \frac{\alpha^2}{8E_b^2}\frac{1+\cos^2\theta}{2},$$

where E_b is the e^-–beam energy and θ the μ^- production angle $\sphericalangle(e^-, \mu^-)$ in the center of mass frame. Calculate σ_{total} in cm^2 for E$_b$=1 GeV. What is the event rate if the beam luminosity is $L = 10^{32}$ cm^{-2}sec^{-1}?. The luminosity measures the incoming flux of particles per cm^2 and per second.

④ Range of interactions : the range of a field and the mass m of the corresponding field quantum are related by the Compton wave length

$$r_0 = \frac{\hbar}{mc},$$

where r_0 appears in the static potential (Yukawa)

$$\Phi(r) \propto \frac{e^{-r/r_0}}{r}.$$

Calculate the range of the strong, the weak and the electromagnetic interaction in cm under the assumption that the interactions are mediated by exchange of a pion ($m_\pi = 135$ MeV), a W boson ($M_W \simeq 80.45 \pm 0.04$ GeV) and a photon ($m_\gamma < 2 \cdot 10^{-16}$ eV experimental bound), respectively. In QED, $m_\gamma = 0$. Discuss this limiting case and the role played by Gauss's law. Hint: Look for static, spherically symmetric solutions of the Klein-Gordon equation.

Hint: For updated precise values of **electroweak parameters** contact the World-Wide-Web (WWW) page *http://lepewwg.web.cern.ch/LEPEWWG/Welcome.html*

2 Basic Principles, Particles and Fields

The basic theoretical framework for a theory of elementary particle interactions is **quantum field theory** (the prototype is quantum electrodynamics QED) which derives from the following principles:

1. Quantum theory : superposition principle, unitarity, probability interpretation, Hilbert space structure.
2. Special relativity : Lorentz invariance, translation invariance
3. Causality (locality) : no signals propagate at speed $v > c$.
4. Existence of a unique normalizable ground state called vacuum.
5. Spectral condition : positivity of the physical state spectrum
6. Symmetry principles : Gauge invariance, global symmetries (quantum numbers, selection rules, multiplets).

All known fundamental interactions (strong, electromagnetic, weak, Gravity?) fit into this scheme. Important properties of relativistic quantum fields, which incorporate wave–particle duality, follow in a straight forward manner from the above principles. A lot can be learned by considering free relativistic particles, which show up in Nature as asymptotic scattering states (at times $t \to \pm\infty$). The aim of the following discussion is to set up notation and to sketch the basics of quantum field theory.

2.1 Minkowski space, Lorentz transformations

The known simple form of Maxwell's equations of classical electrodynamics holds for a restricted class of coordinate frames only, the so called **inertial frames**. The space–time transformations which leave the Maxwell equations invariant form the **Lorentz group**. Accordingly, **Lorentz transformations** are transformations between different inertial frames. The invariance group of Maxwell's equations in this way singles out a particular space-time structure, the **Minkowski space**. **Lorentz invariance** is a basic principle which applies for the other fundamental interactions. We briefly sketch the elements which we will need for a discussion of relativistic field theory.

A space-time event is described by a point (**contravariant** vector)

$$x^\mu = \left(x^0, x^1, x^2, x^3\right) = \left(x^0, \vec{x}\right) \,;\, \mathrm{x}^0 = \mathrm{t}\ (= \mathrm{time})$$

in Minkowski space with metric

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

The metric defines a scalar product [1]

$$x \cdot y = x^0 y^0 - \vec{x} \cdot \vec{y} = g_{\mu\nu} x^\mu y^\nu$$

invariant under Lorentz transformations (L-invariance) which include

1. rotations
2. special Lorentz transformations (boosts)

The set of linear transformations

$$x^\mu \to x^{\mu'} = \Lambda^\mu{}_\nu x^\nu + a^\mu$$

which leave invariant the **distance**

$$(x-y)^2 = g_{\mu\nu}(x^\mu - y^\mu)(x^\nu - y^\nu)$$

between two events x and y form the **Poincaré group** $\mathcal{P}$. $\mathcal{P}$ includes the Lorentz transformations and the translations. We denote the group elements by (Λ, a). To two transformations (Λ_1, a_1) and (Λ_2, a_2), applied successively, there corresponds a transformation $(\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2)$ (multiplication law of the group). The Lorentz transformations $(\Lambda, 0)$ by themselves form the Lorentz group. L-invariance of the scalar products implies the invariance condition

$$\Lambda^\mu{}_\nu \Lambda^\rho{}_\sigma g_{\mu\rho} = g_{\nu\sigma}$$

for the metric. This condition on the matrices $\Lambda^\mu{}_\nu$ implies $\det\Lambda = \pm 1$ and $\mid \Lambda^0{}_0 \mid \geq 1$.Transformations with $\det\,\Lambda = +1$ are called **proper** $(+)$. Such transformations do not change the orientation of frames. Transformations with the property $\Lambda^0{}_0 \geq 1$ are called **orthochronous**, since they exclude time inversions.

Special relativity requires physical laws to be invariant under proper orthochronous Poincaré transformations $\mathcal{P}_+^\uparrow$. Thus $\mathcal{P}_+^\uparrow$ exhibits the general transformation law between **inertial frames**.

We denote by

$$\partial_\mu = \frac{\partial}{\partial_\mu} = \left(\frac{\partial}{\partial_0}, \vec{\nabla}\right)$$

the derivative with respect to $x^\mu = (x^0, \vec{x})$. ∂_μ transforms as a **covariant** vector i.e. it has the same transformation property as $x_\mu = g_{\mu\nu} x^\nu = (x^0, -\vec{x})$. The invariant D'Alembert operator (four-dimensional Laplace operator) is given by

$$\Box = \partial_\mu \partial^\mu = g^{\mu\nu} \partial_\mu \partial_\nu = \frac{\partial^2}{\partial x^{0\,2}} - \triangle.$$

A contravariant **tensor** $T^{\mu_1\mu_2\cdots\mu_n}$ of rank n is an object which has the same transformation property as the products of n contravariant vectors $x_1^{\mu_1}\, x_2^{\mu_2}\, \cdots\, x_n^{\mu_n}$. Covariant or mixed tensors are defined correspondingly.

[1] As usual we adopt the summation convention: Repeated indices are summed over unless stated otherwise. For Lorentz indices $\mu, \cdots = 0, 1, 2, 3$ summation only makes sense (i.e. respects L-invariance) between upper (contravariant) and lower (covariant) indices and is called **contraction**.

The Kronecker symbol

$$\delta^{\mu}{}_{\nu} = g^{\mu\rho}\, g_{\rho\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is a $2nd$ rank mixed tensor. With its help, contracting the invariance condition of the metric with $g^{\sigma\lambda}$, we may write

$$\Lambda^{\mu}{}_{\nu}\Lambda_{\mu}{}^{\lambda} = \delta_{\nu}{}^{\lambda}$$

which shows that

$$\Lambda^{\mu}{}_{\nu} = (\Lambda^{-1})_{\nu}{}^{\mu}$$

is the **transpose of the inverse** of Λ. Covariant vectors transform like

$$x'_{\mu} = \Lambda_{\mu}{}^{\nu} x_{\nu} = x_{\nu}(\Lambda^{-1})^{\nu}{}_{\mu}$$

and the L-invariance of $x^2 = x_\mu x^\mu$ follows immediately.

Finally we will need the totally antisymmetric pseudo-tensor[2]

$$\varepsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & (\mu\nu\rho\sigma) \text{ even permutation of } (0123) \\ -1 & (\mu\nu\rho\sigma) \text{ odd permutation of } (0123) \\ 0 & \text{otherwise} \end{cases} .$$

With the help of this tensor the determinant of any 4x4 matrix A is given by

$$\varepsilon^{\mu\nu\rho\sigma} A^{\alpha}{}_{\mu} A^{\beta}{}_{\nu} A^{\gamma}{}_{\rho} A^{\delta}{}_{\sigma} = detA\ \varepsilon^{\alpha\beta\gamma\delta} .$$

The 4–dimensional volume element is invariant under proper Lorentz transformations which satisfy $\det\Lambda = 1$:

$$d^4x \to \det\Lambda\, d^4x = d^4x$$

A 3–dimensional hyper-surface element is defined by a covariant vector

$$dS_\mu = \varepsilon_{\mu\nu\rho\sigma}\, dx^\nu dx^\rho dx^\sigma$$

and by partial integration we obtain

$$\int_V d^4x\, \partial_\mu f(x) = \int_\Sigma dS_\mu f(x)$$

where $\Sigma = \partial V$ is the boundary of V. The Gauss law takes the form

$$\int_V d^4x\, \partial_\mu f^\mu(x) = \int_\Sigma dS_\mu f^\mu(x) \quad .$$

[2]One easily checks that it transforms as a rank 4 tensor and that it is numerically invariant (identically the same in any inertial frame). Useful relations are

$$\begin{aligned} \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu\rho\sigma} &= -24 \\ \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu\rho\sigma'} &= -6\delta^{\sigma}_{\sigma'} \\ \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu\rho'\sigma'} &= -2\delta^{\rho}_{\rho'}\delta^{\sigma}_{\sigma'} + 2\delta^{\rho}_{\sigma'}\delta^{\sigma}_{\rho'} \\ \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu'\rho'\sigma'} &= -\delta^{\nu}_{\nu'}\delta^{\rho}_{\rho'}\delta^{\sigma}_{\sigma'} + \delta^{\nu}_{\nu'}\delta^{\rho}_{\sigma'}\delta^{\sigma}_{\rho'} + \delta^{\nu}_{\rho'}\delta^{\rho}_{\nu'}\delta^{\sigma}_{\sigma'} - \delta^{\nu}_{\rho'}\delta^{\rho}_{\sigma'}\delta^{\sigma}_{\nu'} - \delta^{\nu}_{\sigma'}\delta^{\rho}_{\nu'}\delta^{\sigma}_{\rho'} + \delta^{\nu}_{\sigma'}\delta^{\rho}_{\rho'}\delta^{\sigma}_{\nu'} \end{aligned}$$

For an infinite volume

$$\int_V \ldots \to \int d^4x \ \ldots \ = \int_{-\infty}^{+\infty} dx^0 \int_{-\infty}^{+\infty} dx^1 \int_{-\infty}^{+\infty} dx^2 \int_{-\infty}^{+\infty} dx^3 \ \ldots$$

the surface terms vanish if the function falls off sufficiently fast in all directions. In this case a component–wise partial integration yields

$$\int d^4x \, g(x) \, \partial_\mu f(x) \ = - \int d^4x \, (\partial_\mu \, g(x)) \, f(x)$$

and the integral of a divergence is vanishing

$$\int d^4x \, \partial_\mu f(x) \ = 0 \ \ .$$

2.2 States of free relativistic particles

The concepts we will develop in the following apply to any isolated relativistic quantum system, irrespective of whether it is elementary or composite. On the one hand, elementary particles like the photon γ, the leptons e, μ, τ, the neutrinos ν_e, ν_μ, ν_τ etc., according to present day knowledge, are elementary and not composite. On the other hand, we know that all hadrons like the proton p, the neutron n, the hyperon Λ or the pions π^+, π^0, π^-, the kaons $K^+, K^0, \bar{K}^0, K^-$ etc. are composite. As described by Quantum Chromodynamics (QCD), hadrons are made up from "colored" quarks which are permanently confined inside the hadrons by strong forces mediated by "colored" gluons. Hadrons thus are **elementary composite particles**, elementary in the sense that they are not breakable into their constituents. In spite of their complex internal structure, with valence quarks, sea quarks and a cloud of gluons, a proton, for example, has precisely spin 1/2, i.e. it is quantized according to the rules of isolated elementary objects. How this dynamically happens in the proton is still a mystery and far from being understood theoretically. But this does not limit the validity of the discussion to follow.

A relativistic quantum mechanical system is described by a state vector $|\psi\rangle \ \in \ \mathcal{H}$ in Hilbert space, which transforms in a specific way under $\mathcal{P}_+^\uparrow$. We denote by $|\psi'\rangle$ the state transformed by $(\Lambda, a) \ \in \ \mathcal{P}_+^\uparrow$. Since the system is required to be invariant transition probabilities must be conserved

$$| < \phi' | \psi' > |^2 = | < \phi | \psi > |^2 \ \ .$$

Therefore there must exist a unitary operator $U(\Lambda, a)$ such that

$$|\psi\rangle \to |\psi'\rangle = U \left(\Lambda, a\right) |\psi\rangle \ \in \ \mathcal{H}$$

and $U(\Lambda, a)$ must satisfy the group law:

$$U \left(\Lambda_2, a_2\right) U \left(\Lambda_1, a_1\right) = \omega U \left(\Lambda_2 \Lambda_1, \Lambda_2 a_1 + a_2\right) \ .$$

This means that $U(\Lambda, a)$ is a **representation up to a phase** ω (ray representation) of $\mathcal{P}_+^\uparrow$. Without loss of generality one can choose $\omega = \pm 1$ (Wigner 1939).

The generators of $\mathcal{P}_+^\uparrow$ are the relativistic energy–momentum operator P_μ

$$U \left(a\right) \equiv U \left(1, a\right) = e^{i P_\mu a^\mu} = 1 + i P_\mu a^\mu + \ldots$$

and the relativistic angular momentum operator $M_{\mu\nu}$

$$U(\Lambda) \equiv U(\Lambda, 0) = e^{\frac{i}{2}\omega^{\mu\nu}M_{\mu\nu}} = 1 + \frac{i}{2}\omega^{\mu\nu}M_{\mu\nu} + \dots$$

Since for infinitesimal transformations we have

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} \quad \text{with} \quad \omega_{\mu\nu} = -\omega_{\nu\mu},$$

the generators $M_{\mu\nu}$ are antisymmetric:

$$M_{\mu\nu} = -M_{\nu\mu} \ .$$

By unitarity of $U(\Lambda, a)$, P_μ and $M_{\mu\nu}$ are Hermitean operators on the Hilbert space. The generator of the time translations P_0 represents the Hamiltonian H of the system ($H \equiv P_0$) and determines the **time evolution**. If $|\psi\rangle = |\psi\rangle_H$ is a Heisenberg state, which coincides with the Schrödinger state $|\psi(0) >_S$ at $t = 0$, then $|\psi(t) >_S = e^{-iHt} \ |\psi(0) >_S$ represents the state of the system at time t.

We usually work in the Heisenberg picture: The state of a system is represented by a time independent vector $|\psi\rangle$ (Heisenberg state) while the physical observables are represented by time dependent Hermitean operators

$$\mathcal{O}(t) = e^{-iHt}\, \mathcal{O}(0)\, e^{iHt}$$

which satisfy the Heisenberg equation of motion

$$[\mathcal{O}(x), H] = i\partial_0 \mathcal{O}(x) \ .$$

The latter is the time component of

$$[\mathcal{O}(x), P_\mu] = i\partial_\mu \mathcal{O}(x)$$

valid for relativistic systems.

The components of $M_{\mu\nu}$ can be expressed in terms two vectors (by antisymmetry it has 6 independent elements):

$$M_{ik} := (M_{23}, M_{31}, M_{12}) = \vec{J}$$

are the operators of the total angular momentum of the system (generators of space rotations), and

$$M_{i0} := (M_{10}, M_{20}, M_{30}) = \vec{K}$$

are the generators of the **Lorentz boost**.

A finite rotation of magnitude $|\vec{\omega}|$ about the direction of $\vec{\omega}$ is represented by

$$U(R(\vec{\omega}), 0) = e^{-i\vec{\omega}\vec{J}} \ .$$

Similarly, a special Lorentz transformation by "velocity" $\vec{\beta}$ is represented by

$$U(L(\vec{\beta}), 0) = e^{i\vec{\beta}\vec{K}} \ .$$

In order to construct a complete set (basis) of state vectors we have to find a complete set of commuting observables, which are represented by Hermitean operators. The set has to include the Hamiltonian P_0 in order have conserved (time independent) quantities. A basis of states is then given by the simultaneous eigenvectors and is labeled by the corresponding eigenvalues.

The properties of states which derive from relativistic behavior are determined by the Lie-algebra of $\mathcal{P}_+^\uparrow$. This Lie algebra, the commutation relations for the P_μ and $M_{\mu\nu}$ $(\vec{J}, \vec{K})$, can be obtained by insertion of infinitesimal transformations into the group law for the $U(\Lambda,a)'s$ (see Exercises 2.8 and E.2). One finds

$$\begin{aligned}
[P_\mu, P_\nu] &= 0 \\
[\,\vec{J}, P_0] &= 0 \quad \text{(conserved)} \\
[\vec{K}, P_0] &= -i\vec{P} \neq 0 \quad \text{(not conserved)}
\end{aligned}$$

and the angular momentum algebra $[J_i, J_k] = i\epsilon_{ikl}J_l$. Furthermore, the generators of the boosts satisfy

$$[J_i, K_k] = i\epsilon_{ikl}K_l \;, \quad [K_i, K_k] = -i\epsilon_{ikl}J_l \;, \quad [K_i, P_k] = iP_0 g_{ik} \;.$$

We notice that the generators of the special Lorentz transformations are not conserved. Therefore no new conserved quantum numbers associated with the special Lorentz transformations are obtained.

We are now able to define the concept of a **free relativistic particle**. It is an isolated system corresponding to a quantum mechanical one particle state which transforms unitary and simple, i.e. as an irreducible unitary representation, under $\mathcal{P}_+^\uparrow$. The irreducible representations are those representations which cannot be decomposed into simpler ones. They are the basic building blocks from which all representations can be obtained by forming products. We briefly discuss the simplest one particle states in the following.

As in ordinary (non relativistic) quantum mechanics $[J_i, P_k] = i\epsilon_{ikl}P_l \neq 0$. Thus simultaneous eigenstates of momentum and angular momentum can only be constructed in the rest frame, where angular momentum determines the **spin** of the particle. Here the usual angular momentum quantization applies: Simultaneous eigenstates to $P^0 = m$, $\vec{J}^{\,2} = j(j+1)$ and $J_3 = j_3$ exist ($m \neq 0$ assumed):

$$|m, j, j_3; \alpha\rangle \qquad j \text{ integer or half integer, } j_3 = -j, -j+1, \ldots, j-1, j$$

By α we denote other possible quantum numbers like charge, lepton number, baryon number etc. .

Since $[P_\mu, P_\nu] = 0$ we may construct simultaneous eigenstates with four-momentum p^μ. Given p^μ, the Lorentz transformation $L(p)$ transforming $(m, \vec{0})$ to $p^\mu : p^\mu = L^\mu{}_0 m$ is uniquely determined by $L^0{}_0 = \frac{p^0}{m}$, $L^i{}_0 = \frac{p_i}{m}$, $L^i{}_k = 0$. Then the boosted state[3]

$$|p, j, j_3; \alpha\rangle = U\left(L(p), 0\right)|m, j, j_3; \alpha\rangle \tag{2.1}$$

is an eigenstate of P_μ

$$P_\mu|p, j, j_3; \alpha\rangle = p_\mu|p, j, j_3; \alpha\rangle$$

which transforms according to

$$U\left(\Lambda, a\right)|p, j, j_3; \alpha\rangle = e^{ia\Lambda p}|\Lambda p, j, j_3'; \alpha\rangle D^{(j)}\left(R_{\Lambda,p}\right)_{j_3 j_3'} \tag{2.2}$$

[3]The boost operator $U(L(p),0)$ may always be represented by a combination of rotations by an angle $\pm\phi$ about the z-axis, rotations by an angle $\pm\theta$ about the y-axis and a boost in z-direction. The angles are determined by writing the momentum $\vec{p} = |\vec{p}\,|(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ in polar coordinates. If we rotate $\vec{p}$ into the direction of the z-axis then perform a boost along the z-axis and rotate back to the original direction of $\vec{p}$ we obtain

$$U(L(p),0) = U(R_{\phi,\theta},0)\,U(L_z(|\vec{p}\,|),0)\,U(R^{-1}_{\phi,\theta},0)$$

with $U(L_z(|\vec{p}\,|),0) = e^{i\beta K_3}$, $\tanh\beta = |\vec{p}\,|/p^0 = v$ the velocity and $U(R_{\phi,\theta},0) = e^{-i\phi J_3}e^{-i\theta J_2}e^{i\phi J_3}$.

where $R_{\Lambda,p} = L^{-1}(\Lambda p)\,\Lambda L(p)$ is a pure rotation (so called Wigner rotation) and $D^{(j)}(R)$ is a $2j+1$–dimensional representation of the rotation. There exist two **Casimir operators** (invariant operators commuting with all generators of $\mathcal{P}_+^\uparrow$). One is the mass operator

$$M^2 = P^2 = g_{\mu\nu}P^\mu P^\nu \tag{2.3}$$

the other is

$$L^2 = g_{\mu\nu}L^\mu L^\nu \;;\quad L^\mu \doteq \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}P_\nu M_{\rho\sigma} \tag{2.4}$$

where L^μ is the Pauli-Lubansky operator. These operators characterize mass m and spin j of the states in an invariant way:

$$M^2|p,j,j_3;\alpha\rangle = p^2|p,j,j_3;\alpha\rangle;\; p^2 = m^2 \tag{2.5}$$

yields the mass of the particle and

$$L^2|p,j,j_3;\alpha\rangle = -m^2 j(j+1)|p,j,j_3;\alpha\rangle \tag{2.6}$$

tells us that the spin j has an invariant meaning. The **on–(mass)–shell condition** $p^2 = m^2$ means that

$$p^\mu = (E,\vec{p}\,) \quad \text{with} \quad E = E(p) = \sqrt{\vec{p}^{\,2}+m^2} \;\; .$$

Physical representations must fulfill the spectral condition

$$p^2 \geq 0,\; p^0 \geq 0 \;,$$

namely, **no particle can travel at speed faster that light** and **a particle must have positive energy**.

Lowest lying states are:

1. $p^2 = 0,\quad p^0 = 0$: vacuum
2. $p^2 = 0,\quad p^0 > 0$: photon, neutrino, $\cdots$: massless particles
   1 state for fixed p
3. $p^2 = m^2,\quad p^0 > 0$: electron, pion, $\cdots$: massive particles
   2j+1 states for fixed p

States of two and more particles of masses m_1, m_2, ... have a **continuous spectrum**:

$$(p_1+p_2+\ldots)^2 \geq (m_1+m_2+\ldots)^2 \quad .$$

For massless particles which must travel at the speed of light, quantization in the rest frame is not possible. Instead of the quantization relative to the z–axis in the rest frame a quantization relative to the direction of the momentum $\vec{p}$ can be done. As a standard vector one conveniently chooses a light-like momentum vector in the z-direction $q^\mu = (q,0,0,q)$ which afterwards may be boosted and rotated to an arbitrary light-like four-momentum p^μ. Massless states are always

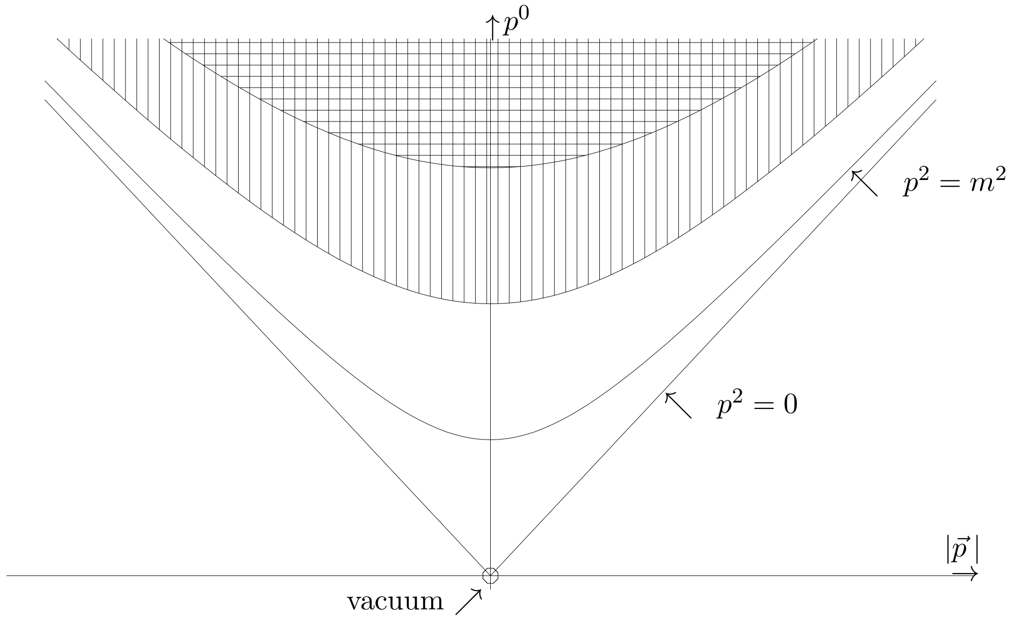


Figure 2.1:
Momentum spectrum of a particle of mass m. continuum of two particle states. continuum of three particle states.

eigenstates of **helicity** $h = \vec{J} \cdot \vec{P} / |\vec{P}\,|$ defined as the projection of angular momentum in direction of $\vec{p}\,$:

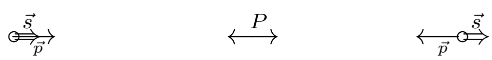
$$h|p, \lambda; \alpha\rangle = \lambda|p, \lambda; \alpha\rangle, \text{ where } \lambda = \pm j,\ j \text{ spin of the particle.}$$

One proves the transformation law (see Appendix C.1)

$$U(\Lambda, a)\,|p, \lambda; \alpha\rangle = e^{ia\Lambda p} e^{-i\phi\lambda} |\Lambda p, \lambda; \alpha\rangle \quad , \tag{2.7}$$

which tells us that massless states always transform diagonal in λ[4]. This result is not very surprising because in order to flip the helicity one would have to perform a Lorentz–boost at speed exceeding the speed of the particle which is not possible for particles traveling at the speed of light.

Thus, to a given spin j there exist exactly two states can not be mixed by $\mathcal{P}_+^\uparrow$ transformations.



h:+ (right–handed = R) h:– (left–handed = L))

The states transform into each other under parity transformations.

Examples:

[4]The phase ϕ depends on the momentum $p^\mu = (|\vec{p}\,|, p^1, p^2, p^3)$ and is determined by

$$e^{i\phi} = \frac{p^1 + ip^2}{\sqrt{(p^1)^2 + (p^2)^2}} \quad .$$

Neutrino: Spin $\frac{1}{2}$, helicity $-$

$|p,\frac{1}{2},-\frac{1}{2}>$ $\quad \overset{\vec{s}}{\Longleftarrow}\underset{\nu}{\circ}\underset{\vec{p}}{\longrightarrow}$ one irreducible representation

$\downarrow\ P$

$|p,\frac{1}{2},\frac{1}{2}>$ $\quad \underset{\nu}{\circ}\overset{\vec{s}}{\Longrightarrow}\underset{\vec{p}}{\rightarrow}$ not existent in Nature

Antineutrino: Spin $\frac{1}{2}$, helicity $+$

$|p,\frac{1}{2},\frac{1}{2}>$ $\quad \underset{\bar{\nu}}{\circ}\overset{\vec{s}}{\Longrightarrow}\underset{\vec{p}}{\rightarrow}$ one irreducible representation

$\downarrow\ P$

$|p,\frac{1}{2},-\frac{1}{2}>$ $\quad \overset{\vec{s}}{\Longleftarrow}\underset{\bar{\nu}}{\circ}\underset{\vec{p}}{\longrightarrow}$ not existent in Nature

Photon: Spin 1, helicity ± 1

$|p,1,\ \ 1>$ right circularly polarized photon

$\updownarrow\ P$ two irreducible representations

$|p,1,-1>$ left circularly polarized photon

In general a 1 photon state is a reducible mixture of the two helicity states.

In high energy physics very often masses of light particles are negligible. It is then natural to describe also the massive particles in a helicity basis with $\vec{p}/|\vec{p}\,|$ as a quantization axis. The canonical basis is related to the helicity basis by a rotation

$$|p,j,j_3;\alpha>=|p,j,\lambda;\alpha>\ D^{(j)}\left(R^{-1}_{\phi,\theta}\right)_{j_3\lambda} \tag{2.8}$$

where $R_{\phi,\theta}$ is the rotation which rotates the z-axis into the direction of $\vec{p}$. Of course for $m\neq 0$ the label λ takes the same $2j+1$ values as j_3 .

Space and time reflections

The **discrete** transformations **parity** P and **time reversal** T defined by

$$Px^\mu=\left\{\begin{array}{lcl} Px^0 & = & x^0 \\ P\vec{x} & = & -\vec{x}\end{array}\right. \quad ; \quad Tx^\mu=\left\{\begin{array}{lcl} Tx^0 & = & -x^0 \\ T\vec{x} & = & \vec{x}\end{array}\right. \tag{2.9}$$

represented on Hilbert space by unitary (P) and anti-unitary (T) operators $U(P)$ and $\bar{U}(T)$, respectively. Notice that $P^2=T^2=1$ and one can show that the phases may be chosen such that $U(P)^2=1$ and $\bar{U}(T)^2=\pm 1$.

The proper orthochronous Poincaré group $\mathcal{P}_+^\uparrow$ together with the reflections P, T and PT=TP generate the full Poincaré group $\mathcal{P} = \mathcal{P}_+^\uparrow \cup \mathcal{P}_-^\uparrow \cup \mathcal{P}_-^\downarrow \cup \mathcal{P}_+^\downarrow$ where the different pieces are obtained as shown in the following diagram:

$$\begin{array}{ccc} \mathcal{P}_+^\uparrow & \stackrel{P}{\rightarrow} & \mathcal{P}_-^\uparrow \\ \downarrow T & & \downarrow T \\ \mathcal{P}_-^\downarrow & \stackrel{P}{\rightarrow} & \mathcal{P}_+^\downarrow \end{array} \quad .$$

Combining a translation $U(a)$ with a reflection I the group law yields

$$U(I)U(a)U^{-1}(I) = U(Ia)\ ; \quad I = P, T$$

or for an infinitesimal transformation

$$U(I)\,(1 + ia^\mu P_\mu)\,U^{-1}(I) = 1 + i(Ia)^\mu P_\mu \quad .$$

Therefore

$$U(I)\,P_\mu U^{-1}(I) = \epsilon_I (IP)_\mu\ ; \quad \epsilon_I = \left\{ \begin{array}{ccc} 1 & U & \text{unitary} \\ -1 & U & \text{anti} - \text{unitary} \end{array} \right.$$

since i changes sign if we pull through an anti-unitary operator U. For the Hamiltonian this reads

$$\begin{array}{rcl} U(P)\,P_0 U^{-1}(P) & = & \epsilon_P P_0 \\ U(T)\,P_0 U^{-1}(T) & = & -\epsilon_T P_0 \end{array}$$

and if we require the energy to remain positive we must have $\epsilon_P = +1$ (U unitary) and $\epsilon_T = -1$ (U anti-unitary). We thus impose the condition

$$[U(I), H] = 0$$

as usual if $U(I)$ is to be a symmetry. In order to indicate the anti-unitarity of $U(T)$ we will denote it by $\bar{U}(T)$ in the sequel. Because of the anti-unitarity of $\bar{U}(T)$ all signs of the generators of $\mathcal{P}_+^\uparrow$ change sign relative to the classical expectations.

Anti-unitarity is defined by the properties

$$\bar{U}(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha^* \bar{U}|\psi\rangle + \beta^* \bar{U}|\phi\rangle = \alpha^*|\psi'\rangle + \beta^*|\phi'\rangle$$

and

$$< \psi'\phi' > = < \psi\phi >^* \quad .$$

The complex conjugation of matrix elements is admitted by the fact that it also preserves the probability $|< \psi|\phi >|^2$. Because of the complex conjugation of matrix elements an anti-unitary transformation implies a **Hermitean transposition** of states and operators. Like any anti-unitary operator we may write $\bar{U}(T)$ in the form $\bar{U}(T) = U(T)\,K$ with $U(T)$ unitary and K an anti-unitary "unit–operator", satisfying $K^+ = K$ and $K^{-1} = K$. Then, while $U(T)$ just changes signs of the momenta and spins, K interchanges "incoming" and "outgoing" states and implies a Hermitean conjugation of the operators:

$$\begin{array}{ccc} K|\phi> = <\phi| & , & <\psi|K = |\psi> \\ KABK & = & B^+A^+ \end{array}$$

such that

$$< \psi|AB\ldots|\phi > \stackrel{K}{\rightarrow} < \phi|\ldots B^+A^+|\psi > \quad .$$

These properties show that an anti-unitary operator cannot have eigenstates and eigenvalues and hence T–invariance cannot be a symmetry in the usual sense of the word. It merely is a substitution rule which preserves probabilities. The interchange of incoming and outgoing states is of course natural for the time reversal operator, and for scattering states and scattering matrix–elements has a literal meaning. After this digression on anti-unitary transformations we proceed with the discussion of P and T.

By experiment : P and T are not universal symmetries . They are conserved by strong and electromagnetic interactions but broken by weak interactions.

The generators of $\mathcal{P}_+^\uparrow$ transform as

$$\begin{array}{llllll} P: & P_0 \to P_0, & \vec{P} \to -\vec{P}, & \vec{J} \to \vec{J}, & \vec{K} \to -\vec{K}, & h \to -h \\ T: & P_0 \to P_0, & \vec{P} \to -\vec{P}, & \vec{J} \to -\vec{J}, & \vec{K} \to \vec{K}, & h \to h \end{array} \tag{2.10}$$

The physical states transform like:

$$\begin{array}{rlcl} m \neq 0: & U(P)|\vec{p}, j, j_3; \alpha\rangle & = & \eta_P |-\vec{p}, j, j_3; \alpha\rangle \\ & \bar{U}(T)|\vec{p}, j, j_3; \alpha\rangle & = & \eta_T \, (-1)^{j+j_3} \, \langle -\vec{p}, j, -j_3; \alpha| \\ m = 0: & U(P)|\vec{p}, \lambda; \alpha\rangle & = & \eta_P (-1)^{j-\lambda} e^{2i\phi\lambda} |-\vec{p}, -\lambda; \alpha\rangle \\ & \bar{U}(T)|\vec{p}, \lambda; \alpha\rangle & = & \eta_T e^{-2i\phi\lambda} \langle -\vec{p}, \lambda; \alpha| \ . \end{array} \tag{2.11}$$

η_P and η_T are possible phase factors called **inner P- and T-parities**, respectively. Spin dependent factors have been split off such that the parities are spin–independent. The term $T\!-\!parity$ is somewhat misleading because as we mentioned above an anti-unitary operator has no eigenvalues (see the discussion of the CPT-theorem below). The phase ϕ is the same as in the transformation law of the states under $\mathcal{P}_+^\uparrow$ transformations.

<u>Result:</u> **Free relativistic particles are described by irreducible unitary representations of $\mathcal{P}_+^\uparrow$. They are characterized by $(m, j; \alpha)$ i.e. mass, spin and charge-like quantum numbers α and by the coordinate–dependent three momentum $\vec{p}$ and the 3^{rd} component j_3 of spin in the rest frame.**

In the following we frequently omit the invariant label $(m, j; \alpha)$ and denote canonical states by $|\vec{p}, j_3\rangle$, helicity states by $|\vec{p}, \lambda\rangle$. The free four-momentum eigenstates (plane wave states) are normalized in a relativistically invariant way by

$$< \vec{p}\,', j_3' | \vec{p}, j_3 >= \delta_{j_3, j_3'} \, (2\pi)^3 \, 2\omega_p \, \delta^{(3)}(\vec{p}\,' - \vec{p}\,) \tag{2.12}$$

with $\omega_p = \sqrt{\vec{p}^{\,2} + m^2}$. The completeness relation reads:

$$\sum_{j_3} \frac{1}{(2\pi)^3} \int \frac{d^3p}{2\omega_p} |\vec{p}, j_3\rangle\langle\vec{p}, j_3| = 1 \ . \tag{2.13}$$

Here $\frac{d^3p}{2\omega_p}$ is the relativistically invariant volume element. By

$$d\mu(p) = \frac{1}{h^3} \frac{d^3p}{2\omega_p}$$

we denote the number of states in this volume element. Possible spin multiplicities are not included here. Since $\hbar = \frac{h}{2\pi} = 1$ we have $h = 2\pi$ for Planck's constant and hence

$$d\mu(p) = \frac{1}{(2\pi)^3} \frac{d^3p}{2\omega_p} = (2\pi)^{-3} \, \Theta(p^0) \, \delta(p^2 - m^2) \, d^4p \ . \tag{2.14}$$

A basis for the **multiple particle states** of free particles can be constructed by forming tensor products of single particle states, which simply means that the state is described by the product of the wave functions of the single particle states. The n–particles states

$$\begin{array}{rcl} |0\rangle & & \text{vacuum} \\ |\vec{p}, j_3\rangle & & \text{one particle state} \\ |\vec{p}_1, \vec{p}_2, j_{3_1}, j_{3_2}\rangle & = & |\vec{p}_1, j_{3_1}\rangle \otimes |\vec{p}_2, j_{3_2}\rangle \\ \vdots & & \\ |\vec{p}_1, \vec{p}_2 \ldots \vec{p}_n, j_{3_1}, j_{3_2} \ldots j_{3_n}\rangle & = & |\vec{p}_1, j_{3_1}\rangle \otimes |\vec{p}_2, j_{3_2}\rangle \ldots \otimes |\vec{p}_n, j_{3_n}\rangle \\ \vdots & & \end{array}$$

span the so called **Fock space** of the particle species $(m, j; \alpha)$. The vacuum is normalized by $< 0|0 >= 1$.

According to the **Pauli principle** states of integer spin particles are symmetric under permutations (Bose statistics), states of half-odd-integer spin particles are antisymmetric under permutations (Fermi statistics). The spin–statistics connection is a consequence of the requirement of causality of quantum fields, as we shall see below. Particles obeying the Bose statistics are called **bosons**, those obeying Fermi statistics **fermions**.

2.3 Creation and annihilation operators

In relativistic processes particles can be created or annihilated. Creation and annihilation operators for relativistic states can be immediately associated with the states discussed above.

We define a creation operator $a^+(\vec{p}, j_3)$ by the requirement that $a^+(\vec{p}, j_3)$ creates, from the vacuum, a particle of species (m, j, α) with momentum $\vec{p}$ and $3rd$ component j_3 of spin in the rest frame

$$| \vec{p}, j_3 > \doteq a^+ (\vec{p}, j_3) \mid 0 > .$$

The corresponding annihilation operator $a(\vec{p}, j_3)$ is defined by the adjoint (Hermitean conjugate) of the creation operator

$$a(\vec{p}, j_3) \doteq \left(a^+ (\vec{p}, j_3)\right)^+ .$$

$a(\vec{p}, j_3)$ annihilates a particle of species (m, j, α) of momentum $\vec{p}$ and 3^{rd} component of spin j_3 in the rest frame. Obviously, $a(\vec{p}, j_3)$ annihilates the vacuum, because

$$\begin{array}{rcl} < 0 \mid \vec{p}, j_3 > & = & 0 \ \text{implies} \\ < 0 \mid a^+(\vec{p}, j_3) \mid 0 > & = & 0 \quad \text{or} \\ < 0 \mid a^+(\vec{p}, j_3) & = & 0 \end{array}$$

hence, taking the Hermitean conjugate,

$$a(\vec{p}, j_3) \mid 0 >= 0 .$$

Furthermore, $a(\vec{p}, j_3)$ acts on one particle states by

$$\sum_{j_3'} \frac{1}{(2\pi)^3} \int \frac{d^3p'}{2\omega_{p'}}\, a(\vec{p}\,', j_3') \mid \vec{p}, j_3 >=\mid 0 > .$$

The operators a and a^+ cannot commute or anticommute since

$$\begin{aligned} < \vec{p}\,', j_3' \mid \vec{p}, j_3 > &= < 0 \mid a\,(\vec{p}\,', j_3')\, a^+(\vec{p}, j_3) \mid 0 > \\ &= \delta_{j_3 j_3'}\, (2\pi)^3\, 2\omega_p\, \delta^{(3)}\, (\vec{p} - \vec{p}\,')\,. \end{aligned}$$

The commutator $[A, B]_- = [A, B] \equiv AB - BA$ (for bosons) or the anticommutator $[A, B]_+ = \{A, B\} \equiv AB + BA$ (for fermions) obviously must read

$$\begin{aligned} \left[a\,(\vec{p}\,', j_3')\,, a^+(\vec{p}, j_3)\right]_\pm &= a\,(\vec{p}\,', j_3')\, a^+\,(\vec{p}, j_3)) \pm a^+\,(\vec{p}, j_3)\, a\,(\vec{p}\,', j_3') \\ &= \delta_{j_3 j_3'}\, (2\pi)^3\, 2\omega_p\, \delta^{(3)}\, (\vec{p} - \vec{p}\,')\,. \end{aligned} \tag{2.15}$$

One easily derives that in addition we must have

$$\left[a\,(\vec{p}\,', j_3')\,, a(\vec{p}, j_3)\right]_\pm = \left[a^+\,(\vec{p}\,', j_3')\,, a^+(\vec{p}, j_3)\right]_\pm = 0 \tag{2.16}$$

These are either the **canonical commutation relations** valid for **bosons** or the **canonical anticommutation relations** valid for **fermions**. Note that at this stage the spin of the particles has not yet been fixed.

These relations hold on any free multiparticle state

$$|\vec{p}_1, \vec{p}_2 \ldots \vec{p}_n, j_{3_1}, j_{3_2} \ldots j_{3_n}\rangle \;=\; \frac{1}{N} a^+(\vec{p}_1, j_{3_1})\, a^+(\vec{p}_2, j_{3_2})\, \ldots a^+(\vec{p}_n, j_{3_n})\, |0 >$$

which is generated by application of products of creation operators to the vacuum. N is the normalization factor. The (anti-) commutation relations are valid on the whole Fock space which is spanned by the above multi-particle states.

Obviously: The wave functions

$$< \Psi|\vec{p}_1, \vec{p}_2 \ldots \vec{p}_n, j_{3_1}, j_{3_2} \ldots j_{3_n}\rangle$$

are symmetric under permutations of the particles for the commutator algebra in which case the particles are bosons and they are antisymmetric under permutations of the particles for the anticommutator algebra in which case the particles are fermions.

Notice that, besides the relativistic transformation laws for the states, which directly carry over to the operators a and a^+, we are dealing with ordinary quantum mechanics of free multiparticle states.

2.4 Fields associated with particles

The implementation of Einstein causality which requires physical signals to propagate at speed $v \leq c$, makes it necessary to consider states and operators in configuration space. Momentum space and configuration space are related by Fourier transformation.

The **quantum field** $\Psi_\alpha(x)$ associated with a particle of mass m and spin j is the relativistically invariant Fourier transform of a linear combination of an annihilation operator $a(\vec{p}, j_3)$ and a creation operator $b^+(\vec{p}, j_3)$:[5]

$$\Psi_\alpha(x) = \sum_{j_3} \int d\mu(p) \left\{ u_\alpha(\vec{p}, j_3)\, a(\vec{p}, j_3)\, e^{-ipx} + v_\alpha(\vec{p}, j_3)\, b^+(\vec{p}, j_3)\, e^{ipx} \right\} \quad . \tag{2.17}$$

This equation expresses the field-particle duality: **particles are the quanta of fields**.

The quanta a and b may have different charge quantum number ($\alpha_a \neq \alpha_b$). They must have the same mass and spin, however. In case of the Dirac field, for example, a describes the electron and b the positron, the anti–particle of the electron (see below). If $a = b$ we call a field **neutral**, when $a \neq b$ we call it **charged**.

The amplitudes $u_\alpha(\vec{p}, j_3)$ and $v_\alpha(\vec{p}, j_3)$ are classical free particle wave functions. In the Fourier transformation we cannot use simply d^4p because p^μ must satisfy $p^2 = m^2$ and $p^0 > 0$. Therefore d^4p is replaced by $d\mu(p) = (2\pi)^{-3}\, \theta(p^0)\, \delta(p^2 - m^2)\, d^4p$. As a consequence $\Psi_\alpha(x)$ satisfies the Klein-Gordon equation

$$(\Box + m^2)\Psi_\alpha(x) = 0 \tag{2.18}$$

which is the Fourier transform of the on–shell condition $\left(-p^2 + m^2\right) \tilde{\Psi}_\alpha(p) = 0$. So far the introduction of the field operator $\Psi_\alpha(x)$ looks quite trivial. Deep consequences follow from the properties which are required to be satisfied by the fields and which determine very specific wave functions u_α and v_α.

The fields $\Psi_\alpha(x)$ are required to satisfy the following properties:

(a) $\Psi_\alpha(x)$ transforms **simple and local** under $\mathcal{P}_+^\uparrow$, specifically,

$$U(\Lambda, a)\Psi_\alpha(x)U^{-1}(\Lambda, a) = D_{\alpha\beta}(\Lambda^{-1})\Psi_\beta(\Lambda x + a)\,, \tag{2.19}$$

where $D(\Lambda)$ is a finite–dimensional (non–unitary) representation of the group $SL(2, C)$ which, in contrast to $\mathcal{P}_+^\uparrow$ itself, exhibits true **spinor representations** (see Appendix B)[6]. The above transformation law is **local** in the sense that the Lorentz transformed field is determined by its value at the Lorentz transformed point only.

[5] Remark for the reader not yet well familiar with free quantum fields: One aspect is that we are looking for the Fourier-transform of operators which live on the mass hyperboloid $p^2 = m^2$ in momentum space. This hyperboloid has a positive energy (**positive frequency part**) mass shell with $p^0 > 0$ ($\to$ "annihilation part") as well as a negative energy (**negative frequency part**) mass shell with $p^0 < 0$ ($\to$ "creation part") which both are separately invariant under orthochronous Lorentz transformation. Therefore the Fourier transform we are looking for exhibits two parts which go with different signs of p^0 in $\exp -ipx$ and $\exp +ipx$ (by relativistic invariance p^0 can enter only via a scalar product, px in our case, i.e., with p^0 the entire scalar product px has to change sign): the first term "living" on the upper hyperboloid ($p^0 > 0$) corresponds to the **annihilation part** which annihilates a particle of mass m and spin j of positive energy $p^0 = \sqrt{\vec{p}^2 + m^2}$ while the second term "living" on the lower hyperboloid ($p^0 < 0$) seems to be destroying a particle of negative energy $p^0 = -\sqrt{\vec{p}^2 + m^2}$ which actually translates into the **creation part** which creates a particle of mass m and spin j of positive energy $p^0 = \sqrt{\vec{p}^2 + m^2}$. For the mathematics associated with the Fourier transform on a hyperboloid we refer to Sec. 2.8 Exercise ⑤ (see also Appendix E.2). As we shall see both terms **must** be present for a causal relativistic field. They must correspond to particles of the same mass and spin but need not describe the same particle. The need for both terms in a relativistic quantum field gives raise to the famous particle–antiparticle pairing discovered by P. Dirac in 1928. Only four years later in 1932 C.D. Anderson dicovered the first antiparticle. The positron as a positively charged electron. For more details and the basic derivations we refer to the Appendices A, B and C.

[6] $SL(2, C)$ is the group of complex 2 x 2 matrices with determinant $\det U = 1$. $\mathcal{P}_+^\uparrow$ and $SL(2, C)$ have the same Lie-Algebra (infinitesimal transformations) but different global transformation properties. $SL(2, C)$ is the simply connected so called "covering group" of the full Poincaré group $\mathcal{P}$, which is multiply connected.

For pure translations the transformation law has the simple form

$$e^{ia_\mu P^\mu}\Psi_\alpha(x)\, e^{-ia_\mu P^\mu} = \Psi_\alpha(x+a)$$

the infinitesimal form of which is

$$[P_\mu, \Psi_\alpha(x)] = -i\partial_\mu \Psi_\alpha(x) \tag{2.20}$$

Similarly, for the generators of the Lorentz transformations one obtains

$$[M_{\mu\nu}, \Psi_\alpha(x)] = -\left(L_{\mu\nu}\,\delta_{\alpha\beta} + (\Sigma_{\mu\nu})_{\alpha\beta}\right)\ \Psi_\beta(x) \tag{2.21}$$

with

$$L_{\mu\nu} = i\,(x_\mu\partial_\nu - x_\nu\partial_\mu)$$

the four-dimensional analog of the orbital angular momentum operator and

$$(\Sigma_{\mu\nu})_{\alpha\beta} = \begin{cases} 0 & \text{spin}\ \ 0 \\ \frac{1}{2}(\sigma_{\mu\nu})_{\alpha\beta} & \text{spin}\ \ \frac{1}{2} \\ i\,(g_{\mu\alpha}g_{\nu\beta} - g_{\nu\alpha}g_{\mu\beta}) & \text{spin}\ \ 1 \end{cases}$$

are the appropriate spin matrices for spin 0,1/2 and 1, respectively. The 4×4 matrix $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ is the antisymmetric tensor constructed from the Dirac matrices.

(b) Causality: Two fields at space-likely separated points x and x' must be independent. This means that the fields must commute (bosons) or anticommute (fermions) at space-like separations (see Fig. 2.2):

$$\left[\Psi_\alpha(x), \Psi^+_\beta(x')\right]_\mp = 0 \quad \text{for} \quad (x-x')^2 < 0 \quad . \tag{2.22}$$

For free fields the (anti-) commutator can easily be calculated by using the representation of the fields in terms of the creation and annihilation operators which satisfy the canonical (anti-) commutation relations. The latter yield c-number distributions involving a delta function which allows to perform one integration trivially. The result is an integral over a c-number distribution.

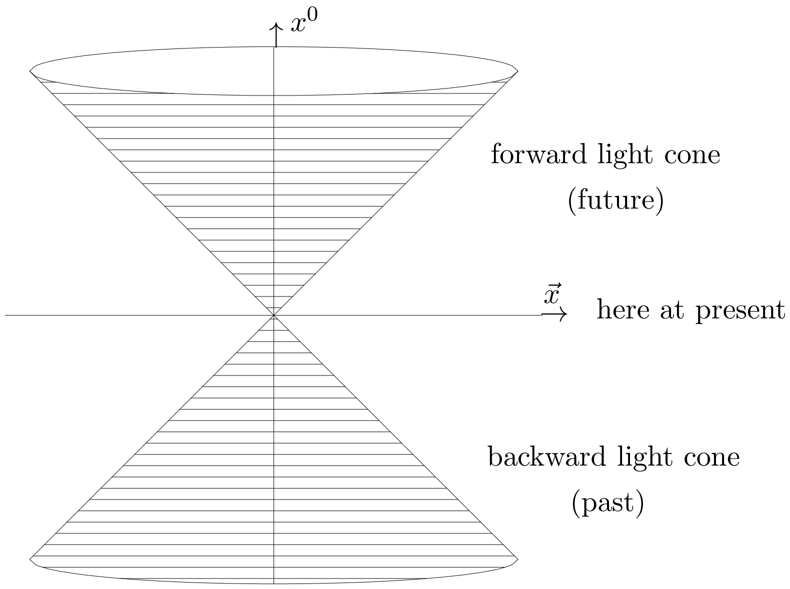


Figure 2.2: Causally connected regions in space–time.

Since the fields always satisfy the Klein-Gordon equation, the commutator (anticommutator) must also satisfy the Klein-Gordon equation. The Klein-Gordon equation has exactly one scalar causal solution

$$(\Box + m^2)\Delta(x) = 0 \quad \text{with } \Delta(x) = 0 \quad \text{when } x^2 < 0 \quad .$$

Up to normalization which is fixed by convention one finds

$$\begin{aligned} i\Delta(x) &= \frac{1}{(2\pi)^3} \int d^4p\; \epsilon(p^0)\; \delta(p^2 - m^2)\; e^{-ipx} \\ &= \frac{1}{(2\pi)^3} \int \frac{d^3p}{2\omega_p} \left(e^{-ipx} - e^{ipx} \right) \Big|_{p^0=\omega_p=\sqrt{\vec{p}^{\,2}+m^2}} \\ &= \frac{1}{2\pi}\; \epsilon(x^0) \left\{ \delta(x^2) - \frac{m^2}{2} \theta(x^2) \frac{J_1(m\sqrt{x^2})}{m\sqrt{x^2}} \right\} \end{aligned} \tag{2.23}$$

where $J_1(z)$ is a Bessel function. The singular function (distribution) $\Delta(x)$ satisfies the singular initial conditions

$$\Delta(0, \vec{x}\,) = 0\,, \;\; \partial^0 \Delta(x^0, \vec{x}\,)|_{x^0=0} = -\delta^{(3)}(\vec{x}\,)$$

and is antisymmetric

$$\Delta(x) = -\Delta(-x) \quad .$$

Fields describing spin j particles must have at least $2j + 1$ components and are characterized by different transformation laws, namely, finite dimensional representations of $SL(2, C)$ which necessarily are non-unitary. More familiar is the characterization of fields by specific field equations (see Appendices B and C):

$$\textbf{Scalar field} \text{ (mass } m\text{, spin 0)} \qquad \varphi(x): \qquad (\,\Box + m^2\,)\,\varphi(x) = 0$$

$$\textbf{Dirac field} \text{ (mass } m\text{, spin 1/2)} \qquad \psi_\alpha(x): \qquad (\,i\gamma^\mu\partial_\mu - m\,)\,\psi_\alpha(x) = 0$$

$$\textbf{Proca field} \text{ (mass } m\text{, spin 1)} \qquad V_\mu(x): \qquad (\,\Box + m^2\,)V_\mu(x) - \partial_\mu(\,\partial_\nu V^\nu\,) = 0$$

$$\text{implies } \partial^\mu V_\mu(x) = 0$$

$$\textbf{Photon field} \text{ (mass 0, spin 1)} \qquad A_\mu(x): \qquad (\,\Box g^{\mu\nu} - (1-\xi^{-1})\,\partial^\mu\partial^\nu\,)\,A_\nu = 0\ .$$

In the field equation for the photon field ξ is a **gauge parameter**, which is necessary to determine the gauge potential $A_\mu(x)$. Physical observables like cross–sections, the electromagnetic field strength tensor or the electromagnetic current are **gauge invariant**, i.e., they are independent of the gauge parameter ξ (see Sec. 4).

One easily checks that the one–particle wave functions u_α and v_α must be classical (i.e. c–number) solutions of the momentum–space versions of the corresponding field equations. They can be determined easily in momentum space. For a scalar field $u = v = 1$, for a Dirac field u and v are four-component spinors and for a spin 1 field u, $v \to \epsilon_\mu$, ϵ^*_μ are the polarization vectors.

Explicit formulae and properties of these fields are collected in Appendix A at the end of this section. A more detailed discussion of spinors and fields may be found in Appendix B. Peculiarities of massless particles and fields are considered in Appendix C.

For the free field (anti-) commutators one obtains:

$$\begin{array}{llcl}
\text{Neutral scalar field :} & [\,\phi\,(x),\ \phi\,(0)] & = & i\Delta(x) \\[2ex]
\text{Charged scalar field :} & [\,\phi\,(x), \phi^+(0)] & = & i\Delta(x) \\
 & [\,\phi\,(x),\ \phi\,(0)] & = & 0 \\
 & [\phi^+(x), \phi^+(0)] & = & 0 \\[2ex]
\text{Dirac field :} & \{\psi_\alpha(x), \bar\psi_\beta(0)\} & = & i\,(i\gamma^\mu\partial_\mu + m)_{\alpha\beta}\,\Delta(x) \\
 & \{\psi_\alpha(x), \psi_\beta(0)\} & = & 0 \\
 & \{\bar\psi_\alpha(x), \bar\psi_\beta(0)\} & = & 0 \\[2ex]
\text{Neutral Proca field :} & [V_\mu\,(x), V_\nu\,(0)] & = & i\left(g_{\mu\nu} + \frac{\partial_\mu\partial_\nu}{m^2}\right)\Delta(x) \\[2ex]
\text{Charged Proca field :} & [V_\mu\,(x), V^+_\nu(0)] & = & i\left(g_{\mu\nu} + \frac{\partial_\mu\partial_\nu}{m^2}\right)\Delta(x) \\
 & [V_\mu\,(x), V_\nu\,(0)] & = & 0 \\
 & [V^+_\mu(x), V^+_\nu(0)] & = & 0\ .
\end{array}$$

These are the configuration space versions of the basic canonical (anti-) commutation relations satisfied by the creation and annihilation operators. By $\bar{\psi}_\beta \doteq \psi^+_\alpha \gamma^0_{\alpha\beta}$, as usual, we denoted the **adjoint spinor** which is defined such that $\bar{\psi}\psi$ is a Lorentz scalar (see the Appendix).

The crucial consequences of the requirements of locality and causality are:

1. **Spin - statistics theorem:**
   Bosons quantized with commutation relations must have integer spin.
   Fermions quantized with anticommutation relation must have half odd–integer spin.

2. **Particle - antiparticle crossing theorem:**

   Each particle of mass m and spin j must have associated an antiparticle with the same mass and spin, which transform under identical representations of $\mathcal{P}^\uparrow_+$. A particle can be its own antiparticle e.g. π^0, γ. In general, a particle and its antiparticle have different quantum numbers like charge, baryon number, lepton number and flavor. Examples are: (e^-, e^+), $(p, \bar{p})$, $(n, \bar{n})$, (π^+, π^-), $(K^0, \bar{K}^0)$ etc. .

   Since particle - antiparticle pairs can annihilate electromagnetically into a photon and the photon is neutral, $Q_\gamma = B_\gamma = L_{\ell\gamma} = \cdots = 0$, particles and antiparticles must have opposite quantum numbers.

   For massless particles we get the stronger result: To each left–handed (right–handed) particle there must exist a right–handed (left–handed) antiparticle.

Examples of massless fields:

a) Antiparticle = particle, e.g. γ, requires that both helicities <u>must</u> exist in Nature and

$$A^\mu(x) = \sum_\pm \int d\mu(p)\, \varepsilon^\mu_\pm(p) \left(a(\vec{p},\pm)\, e^{-ipx} - a^+(\vec{p},\pm)\, e^{ipx} \right)$$

is a Hermitean field with natural inner parity $\eta_P = (-1)^j = -1$.

b) Antiparticle $\neq$ particle, e.g. $\nu_L, \bar{\nu}_R$ exist in Nature

$$\psi_{L\alpha}(x) = \int d\mu(p) \left(u_{L\alpha}(\vec{p},-)\, a(\vec{p},-)\, e^{-ipx} + v_{L\alpha}(\vec{p},+)\, b^+(\vec{p},+)\, e^{ipx} \right)$$

where $\psi_{L\alpha} = \frac{1}{2}(1-\gamma_5)\psi_\alpha$ is the left–handed neutrino field and ψ_α a massless Dirac field. a^+ creates a left–handed neutrino, b^+ creates a right–handed antineutrino. Notice that ν_L and $\bar{\nu}_L$ can not be combined into a local and causal field, because they transform in a different way under $\mathcal{P}^\uparrow_+$. Given the existence of ν_L, causality requires the existence of $\bar{\nu}_R$. It does not require $\bar{\nu}_L$ to exist, however.

In order to understand better how the above theorems come about let us consider a modified field

$$\Psi_\alpha(x) = \sum_{j_3} \int d\mu(p) \left\{ \xi\, u_\alpha(\vec{p},j_3)\, a(\vec{p},j_3)\, e^{-ipx} + \eta\, v_\alpha(\vec{p},j_3)\, b^+(\vec{p},j_3)\, e^{ipx} \right\} \tag{2.24}$$

where we have multiplied the annihilation part by ξ and the creation part by η. Now we calculate the commutator and obtain a result of the form (see Appendix B.6)

$$\begin{aligned} \left[\Psi_\alpha(x), \Psi^+_\beta(x')\right]_\mp &= \left(\tfrac{i}{m}\right)^{2j}\; t^{\mu_1\ldots\mu_{2j}}_{\alpha\beta}\, \partial_{\mu_1} \ldots \partial_{\mu_{2j}} \times \\ \tfrac{1}{(2\pi)^3} \int \tfrac{d^3p}{2\omega_p} &\left(|\xi|^2 e^{-ip(x-x')} \mp (-1)^{2j} |\eta|^2 e^{ip(x-x')} \right) \Big|_{p^0=\omega_p=\sqrt{\vec{p}^{\,2}+m^2}} \end{aligned} \tag{2.25}$$

with appropriate tensor coefficients $t^{\mu_1 \dots \mu_{2j}}_{\alpha\beta}$ the precise form of which is not important here. Now, locality only holds if the integral is proportional to $\Delta(x-x')$ and this requires

$$|\xi|^2 = \pm(-1)^{2j}|\eta|^2 \quad .$$

This condition indeed implies two things:

$$\pm(-1)^{2j} = 1 \tag{2.26}$$

which is the spin-statistics theorem and

$$|\xi| = |\eta| \tag{2.27}$$

which is the crossing theorem. Normalization and phases can then be chosen without loss of generality such that

$$\xi = \eta = m^j \qquad \text{by convention !} \tag{2.28}$$

This normalization is chosen such that the fields have smooth limits as $m \to 0$.

Under very general conditions a field theory which involves different species of causal and local fields can be arranged always to satisfy **"normal commutation relations"** which means that for space–like separations different Bose fields commute, Fermi fields commute with Bose fields and different Fermi fields anti–commute:

$$\begin{aligned}
[\ \phi_i(x),\ \phi_j(y)] &= 0 \\
\{\psi_i(x), \psi_j(y)\} &= 0 \quad (x-y)^2 < 0 \\
[\ \phi_i(x),\ \psi_j(y)] &= 0 \qquad\qquad .
\end{aligned}$$

The validity of normal commutation relations is a condition for the validity of the PCT–theorem which will be discussed later.

Important remark: The fields have been constructed as a tool to control Lorentz invariance and causality of relativistic particles in Minkowski space. The fields associated with particles of given mass, spin and "charge" are determined in an unambiguous manner if we discard some ambiguities with the massless spin 1 fields. These fields in general are **non-Hermitean** and hence cannot themselves be observables. They merely serve as auxiliary fields in terms of which observables may be represented. Thereby the general properties required for the fields carry over to the observables in a simple way. **Observable fields** not only must be **Hermitean, covariant and gauge-invariant** but they also must commute for space-like separations by causality. Examples of observable fields are the electromagnetic field strength tensor $F^{\mu\nu}$ and the electromagnetic current $j^\mu = \bar{\psi}\gamma^\mu\psi$. The Fermi fields themselves cannot have a direct physical meaning since they never commute outside the light-cone by the fact that the spin-statistics theorem requires them to anti-commute at space-like separations. Only the Hermitean local bilinear forms $\bar{\psi}_\alpha(x)\Gamma_{(r)\alpha\beta}\psi_\beta(x)$ (Γ an appropriate 4 by 4 matrix (see the Appendix)) or monomials of them are also causal. Let us consider as a typical example the commutator

$$\left[\bar{\psi}(x)\,\Gamma_{(r)}\psi(x)\,,\ \bar{\psi}(y)\,\Gamma_{(r')}\psi(y)\,\right] = \Gamma_{(r)\alpha\beta}\Gamma_{(r')\gamma\delta}\,[\bar{\psi}_\alpha(x)\,\psi_\beta(x)\,,\ \bar{\psi}_\gamma(y)\,\psi_\delta(y)\,] \ \ .$$

One easily checks that the bilinear commutator

$$\begin{aligned}
& \bar{\psi}_\alpha(x)\,\psi_\beta(x)\,\bar{\psi}_\gamma(y)\,\psi_\delta(y)\ - \bar{\psi}_\gamma(y)\,\psi_\delta(y)\,\bar{\psi}_\alpha(x)\,\psi_\beta(x) \\
=\ & \bar{\psi}_\alpha(x)\,\{\psi_\beta(x)\,,\ \bar{\psi}_\gamma(y)\,\}\psi_\delta(y)\ - \bar{\psi}_\gamma(y)\,\{\psi_\delta(y)\,,\ \bar{\psi}_\alpha(x)\,\}\psi_\beta(x) \\
-\ & \{\bar{\psi}_\alpha(x)\,,\ \bar{\psi}_\gamma(y)\,\}\psi_\beta(x)\,\psi_\delta(y)\ + \bar{\psi}_\gamma(y)\,\bar{\psi}_\alpha(x)\,\{\psi_\beta(x)\,,\ \psi_\delta(y)\,\}
\end{aligned}$$

vanishes for $(x-y)^2 < 0$, where the anti-commutators of the Dirac fields vanish.

There are important consequences of this. Fermion fields must always enter in a physical problem as adjoint pairs $\bar{\psi}(x) \ldots \psi(y)$. This gives raise to **fermion number conservation** if we assign opposite fermion number to ψ and $\bar{\psi}$. Transition matrix elements between states of different fermion number all vanish and the **phases of individual fermion fields are never observable**.

2.5 Chiral fields

As we have mentioned earlier massless particles with non-vanishing spin are described in the helicity basis. For any spin j there are exactly two states of helicity $\pm j$ which do not mix under Poincaré transformations. This decomposition into two invariant subspaces carries over to the local fields as we will briefly discuss now. We consider a massless Dirac field. In the **helicity representation** appropriate to describe a massless field the Dirac matrix γ_5 and the so called **chiral projectors**

$$\Pi_\pm = \frac{1}{2}\,(1 \pm \gamma_5)$$

are diagonal (see Appendix):

$$\gamma_5 = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}$$

$$\Pi_+ = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} , \quad \Pi_- = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$$

where **1** and **0** are the 2×2 unit and zero matrices, respectively. The four-spinors $\psi_\alpha, u_\alpha, v_\alpha$ decompose into right–handed and left–handed two-component spinors or **Weyl spinors**

$$\begin{aligned} \psi_\alpha &= \psi_{R\alpha} + \psi_{L\alpha} \\ \psi_{R\alpha} &= (\Pi_+\psi)_\alpha \ , \quad \psi_{L\alpha} = (\Pi_-\psi)_\alpha \\ \psi_{R\alpha} &= 0 \text{ for } \alpha = 3,4 \ ; \quad \psi_{L\alpha} = 0 \text{ for } \alpha = 1,2 \end{aligned}$$

It is very simple to check that $\Pi_\pm$ are Hermitean projection operators, meaning that they have the properties:

$$\Pi_+ + \Pi_- = 1 \ , \ \ \Pi_+^2 = \Pi_+ \ , \ \ \Pi_-^2 = \Pi_- \ , \ \ \Pi_+\Pi_- = \Pi_-\Pi_+ = 0 \ . \tag{2.29}$$

We point out once more that a **left–handed chiral field** describes a particle with negative helicity (left–handed particle) and at the same time an antiparticle with positive helicity (right–handed antiparticle) which are the two states with identical transformation properties under $\mathcal{P}_+^\uparrow$. A corresponding statement holds for a **right–handed chiral field**. Note that the "handedness" of a field and of the anti–particle it describes are opposite.

May be not so well known is the corresponding representation for spin 1 particles like the photon. From the field strength tensor $F^{\mu\nu}$ and its dual $\tilde{F}^{\mu\nu} \doteq \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$ we may form the tensors

$$\mathcal{F}_\pm^{\mu\nu} = \frac{1}{2}\left(F^{\mu\nu} \pm \tilde{F}^{\mu\nu}\right)$$

which have the following properties:

- $\partial_\mu \mathcal{F}^{\mu\nu}_\pm = 0$
  since the free field satisfies the homogeneous Maxwell equations: $\partial_\mu F^{\mu\nu} = 0$ and $\partial_\mu \tilde{F}^{\mu\nu} = 0$.
- $\tilde{\mathcal{F}}^{\mu\nu}_\pm \doteq \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\mathcal{F}_{\pm\,\rho\sigma} = \pm\mathcal{F}^{\mu\nu}_\pm$
  such that $\mathcal{F}^{\mu\nu}_+$ **self-dual** and $\mathcal{F}^{\mu\nu}_-$ **anti-self-dual**. This implies that the two fields $\mathcal{F}^{\mu\nu}_\pm$ have each exactly 3 independent components and the two fields do not mix under Lorentz transformations. They describe each a pure state of fixed helicity.
- $(\mathcal{F}^{\mu\nu}_\pm)^* = \mathcal{F}^{\mu\nu}_\mp$

As usual the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is given by the curl of the photon field which is represented by a vector potential A_μ.

2.6 Addendum: Finite dimensional representations of $SL(2, C)$

The Lie algebra of $SL(2, C)$, the group of complex unimodular 2×2 matrices, is identical to the one of Lorentz group $L^\uparrow_+$. The 6 generators are given by the generators of the rotations $\vec{J}$ plus the generators of the Lorentz boosts $\vec{K}$. The Lie algebra given by

$$[J_i, J_k] = i\epsilon_{ikl}J_l \ , \ [J_i, K_k] = i\epsilon_{ikl}K_l \ , \ [K_i, K_k] = -i\epsilon_{ikl}J_l \quad .$$

is a coupled algebra of the J_i's and K_i's. Since these generators are Hermitean $\vec{J} = \vec{J}^+$ and $\vec{K} = \vec{K}^+$ the group elements $e^{-i\vec{\omega}\vec{J}}$ and $e^{i\vec{\beta}\vec{K}}$ are unitary[7] . This algebra can be decoupled by the linear transformation

$$\vec{A} = \frac{1}{2}\left(\vec{J} + i\vec{K}\right) \ , \ \vec{B} = \frac{1}{2}\left(\vec{J} - i\vec{K}\right)$$

under which the Lie algebra takes the form

$$\vec{A} \times \vec{A} = i\vec{A} \ , \ \vec{B} \times \vec{B} = i\vec{B} \ , \ [A_i, B_j] = 0$$

of two decoupled angular momentum algebras. Since $\vec{A}^+ = \vec{B}$ and $\vec{B}^+ = \vec{A}$, the new generators are not Hermitean any more and hence give raise to **non–unitary** irreducible representations. These are **finite dimensional** and evidently characterized by a pair (A, B), with $2A$ and $2B$ integers. The dimension of the representation (A, B) is $(2A + 1) \cdot (2B + 1)$. Note that this classification exhausts all finite dimensional irreducible representations which are necessarily non-unitary. The unitary representations, under which physical states transform, on the other hand are necessarily infinite dimensional.

One may introduce "states" in the representation space:

$$|a, b\rangle \ , \quad a = -A, -A + 1, \ldots, A - 1, A \ , \ \ b = -B, -B + 1, \ldots, B - 1, B$$

for which the matrix elements of the generators read:

$$\begin{aligned} \langle a, b|\vec{A}|a', b'\rangle &= \delta_{bb'}\vec{J}^{(A)}_{aa'} \\ \langle a, b|\vec{B}|a', b'\rangle &= \delta_{aa'}\vec{J}^{(B)}_{bb'} \end{aligned}$$

[7]In $SL(2, C)$ the Lie algebra obviously has the 2×2 matrix representation $J_i = \sigma_i/2 \ \ K_i = -I\ \sigma_i/2$ in terms of the Pauli matrices, however, $\vec{K}^+ = -\vec{K}$ is non-Hermitean and the corresponding finite dimensional representation non-unitary. Unitary representations of the Lorentzgroup, required to implement relativistic covariance on the Hilbert space of physical states, are necessarily infinite dimensional

with $\vec{J}^{(j)}$ the usual $(2j+1)$ dimensional representation of the rotation group:

$$\begin{aligned}\left(J_x^{(j)} \pm iJ_y^{(j)}\right)_{j_3 j_3'} &= \delta_{j_3 j_3' \pm 1} \left\{(j \mp j_3)(j \pm j_3 + 1)\right\}^{1/2} \\ \left(J_z^{(j)}\right)_{j_3 j_3'} &= \delta_{j_3 j_3'} \cdot j_3 \quad .\end{aligned}$$

The simplest irreducible representations are of type $(j,0)$, $(0,j)$ and (j,j) :

$$\begin{array}{lll}(j,0) \quad : & \left.\begin{array}{l}A \to J^{(j)} \\ B \to 0\end{array}\right\} \Rightarrow & \begin{array}{ll}J \to & J^{(j)} \\ K \to & -iJ^{(j)}\end{array} \\ \Rightarrow \text{ Representation :} & D^{(j)}(\Lambda) = e^{(\vec{\beta} - i\vec{\omega})\,\vec{J}^{(j)}} & \\ (0,j) \quad : & \left.\begin{array}{l}A \to 0 \\ B \to J^{(j)}\end{array}\right\} \Rightarrow & \begin{array}{ll}J \to & J^{(j)} \\ K \to & iJ^{(j)}\end{array} \\ \Rightarrow \text{ Representation :} & \bar{D}^{(j)}(\Lambda) = e^{(-\vec{\beta} - i\vec{\omega})\,\vec{J}^{(j)}} & \end{array}$$

Note that $\bar{D}^{(j)}(\Lambda)$ is the adjoint of $D^{(j)}(\Lambda)$: $\bar{D}^{(j)}(\Lambda) = D^{(j)+}(\Lambda^{-1})$.

We are now ready to classify the fields in terms of the simplest finite dimensional (non-unitary) representations. We denote group element as follows: $A \in SL(2,C)$ and $\Lambda \in \mathcal{P}_+^{\uparrow}$. The simplest representations the are:

$$\begin{array}{ccl}1 & D(0,0) & \text{scalar field} \\ A & D(1/2,0) & \text{neutrino field} \\ \bar{A} & D(0,1/2) & \text{antineutrino field} \\ \begin{pmatrix} A & 0 \\ 0 & \bar{A}\end{pmatrix} & D(1/2,0) \oplus D(0,1/2) & \text{Dirac field} \\ \begin{pmatrix} \mathrm{Re}A & \mathrm{Im}A \\ -\mathrm{Im}A & \mathrm{Re}A\end{pmatrix} & D(1/2,0) \oplus D(0,1/2) & \text{Majorana field} \\ \Lambda & D(1/2,1/2) & \text{vector field (contravariant)} \\ (\Lambda^{-1})^T & D(1/2,1/2) & \text{vector field (covariant)} \\ \Lambda \otimes \Lambda & D(1/2,1/2) \otimes D(1/2,1/2) & \\ & = D(1,1) & \text{graviton field} \\ & \multicolumn{2}{c}{\text{(symmetric, traceless 2nd rank tensor)}} \\ & \oplus D(1,0) \oplus D(0,1) & \text{electromagnetic field} \\ & \multicolumn{2}{c}{\text{(self-dual and anti-self-dual antisymmetric 2nd rank tensor)}} \\ & = D(0,0) & \text{scalar field} \\ & \multicolumn{2}{c}{\text{(scalar)}}\end{array}$$

the general rule for the reduction of direct product representations is the following:

$$D(\frac{i}{2},\frac{j}{2}) \otimes D(\frac{m}{2},\frac{n}{2}) = \sum_{a=0}^{min(i,m)} \sum_{b=0}^{min(j,n)} \oplus D(\frac{i+m-2a}{2},\frac{j+n-2b}{2}) \quad .$$

The representations $D(j,k)$, while irreducible with respect to $SL(2,C)$, are not irreducible with respect to the subgroup $SU(2) \subset SL(2,C)$ of rotations, in general. This means that the representations $D(j,k)$, in general, do not have a fixed spin. The spin content may be obtained by performing the angular momentum decomposition:

$$D(j,k) \supset D_R^{(j,k)} = D_R^{(j+k)} \oplus D_R^{(j+k-1)} \oplus \cdots \oplus D_R^{(|j-k|)} \quad . \tag{2.30}$$

2.7 Charge conjugation, parity, time reversal and the CPT-theorem

Particle - antiparticle crossing gives rise to the definition of **charge conjugation C** (Kramer 1937), represented by a unitary operator $U(C)$ on Hilbert space (representations up to a phase η_C). Charge conjugation, parity and time reversal act as follows on the annihilation operators:

$$\begin{aligned}
C \quad &: \quad U(C)\, a(\vec{p},j_3)\, U^{-1}(C) = \eta_C\; b(\vec{p},j_3) \\
&\qquad U(C)\, b(\vec{p},j_3)\, U^{-1}(C) = \bar{\eta}_C\; a(\vec{p},j_3) \\
P \quad &: \quad U(P)\, a(\vec{p},j_3)\, U^{-1}(P) = \eta_P\; a(-\vec{p},j_3) \\
&\qquad U(P)\, b(\vec{p},j_3)\, U^{-1}(P) = \bar{\eta}_P\; b(-\vec{p},j_3) \\
T \quad &: \quad \bar{U}(T)\, a(\vec{p},j_3)\, \bar{U}^{-1}(T) = \eta_T\; C_{j_3 j_3'}\; a^+(-\vec{p},j_3') \\
&\qquad \bar{U}(T)\, b(\vec{p},j_3)\, \bar{U}^{-1}(T) = \bar{\eta}_T\; C_{j_3 j_3'}\; b^+(-\vec{p},j_3')
\end{aligned} \tag{2.31}$$

The matrix $C_{j_3 j_3'}$ shows up due to the anti-unitarity of $\bar{U}(T)$ which includes a complex conjugation. C is a $2j+1 \times 2j+1$ matrix which transforms a rotation matrix into its complex conjugate representation: $CD^{(j)}(R)C^{-1} = D^{(j)}(R)^*$ with $C^*C = (-1)^{2j}, C^+C = 1$. Explicitly, we have

$$C_{j_3 j_3'} = \left(e^{i\pi J_2^{(j)}}\right)_{j_3 j_3'} = (-1)^{j+j_3}\delta_{j_3,-j_3'} \quad .$$

As we mentioned earlier the intrinsic parities are defined to be spin–independent by explicitly taking out the spin–dependent factors. The requirement of simple local transformation laws of the fields under C,P,T transformations has interesting consequences. Simple means that the creation part proportional to b^+ and the annihilation part proportional to a transform with the same phase such that the field is multiplied by the corresponding common phase. As we shall derive below this implies:

$$\bar{\eta}_P = \eta_P^*(-1)^{2j}\;,\quad \bar{\eta}_T = \eta_T^*\;,\quad \bar{\eta}_C = \eta_C^*$$

or, equivalently,

$$\underline{\eta_P\bar{\eta}_P = (-1)^{2j}}\;,\quad \eta_T\bar{\eta}_T = 1\;,\quad \eta_C\bar{\eta}_C = 1 \quad . \tag{2.32}$$

Consequently:

A particle - antiparticle pair must have inner parity:

$$\eta_P\bar{\eta}_P = \begin{cases} \;\;1 & \text{for bosons} \\ -1 & \text{for fermions .} \end{cases} \tag{2.33}$$

For <u>massive</u> neutral particles (particle $\equiv$ antiparticle) we then must have

$$\eta_P = \begin{cases} \pm 1 & \text{for neutral bosons} \\ \pm i & \text{for neutral fermions} \end{cases} \tag{2.34}$$

and $\eta_T = \pm 1$, $\eta_C = \pm 1$ for neutral particles.

The crucial question in any case is whether a particular type of interaction preserves the corresponding symmetry. If a symmetry is good **the interaction fixes the relative phases of different fields** as we shall see later.

For neutral (Hermitean) fields like the photon field or the π^0 and the ρ^0 fields the phases are observable and have to be determined by experiment. Thus neutral particles have absolute intrinsic parities. Depending on whether $\eta_P = +1$ or $\eta_P = -1$ one distinguishes between **scalar** or **pseudo-scalar** spin 0 bosons and between **vector** or **axial vector** spin 1 bosons, respectively. For charged (non-Hermitean) fields only relative phases may have a physical meaning. The individual phases of charged particles are not observable and may be chosen by convention as far as they are not fixed relative to each other.

In order to establish the results just given, let us now consider the C, P and T transformation properties of free fields in some more detail. We first notice, that a field Ψ contains a and b^+ while its Hermitean conjugate Ψ^+ contains b and a^+ as annihilation and creation operators. The transformed field must be a non-singular linear combination of the components of either the field (P) or its Hermitean adjoint (C,T). This follows from the transformation properties of the annihilation and creation operators and from the fact that the transformed field must be local again and satisfies the appropriate field equation. We then find:

A) Scalar fields

$$\begin{aligned} U(C)\,\phi(x)\,U^{-1}(C) &= \eta_C\,\phi^+(x) \\ U(P)\,\phi(x)\,U^{-1}(P) &= \eta_P\,\phi\,(Px) \\ \bar{U}(T)\,\phi(x)\,\bar{U}^{-1}(T) &= \eta_T\,\phi^+(Tx)\ . \end{aligned} \tag{2.35}$$

For real fields $\phi = \phi^+$ in particular we have

$$\phi(x) \overset{P}{\to} \begin{cases} \phi(Px) & \text{scalar field} \\ -\phi(Px) & \text{pseudoscalar field} \end{cases}$$

B) Vector fields

For the assignment of inner parities an important reference quantity is the Hermitean, gauge-invariant and conserved electromagnetic current $j^\mu(x)$. As an observable for which also a classical limit exists it has no free phase. As suggested by classical correspondence, the natural convention is to require it to transform like an ordinary four-vector:

$$\begin{aligned} U(P)\,j^\mu(x)\,U^{-1}(P) &= (Pj)^\mu(Px) &= (j^0, -\vec{j}\,)(Px) \\ \bar{U}(T)\,j^\mu(x)\,\bar{U}^{-1}(T) &= -(Tj)^\mu(Tx) &= (j^0, -\vec{j}\,)(Tx) \end{aligned}$$

where the "phase" has to be chosen such that the charge keeps its sign. Correspondingly, j^0, which represents the charge density, must keep sign. On the other hand under charge conjugation the charge changes sign which implies

$$U(C)\,j^\mu(x)\,U^{-1}(C) = -j^\mu(x)\ .$$

Notice that for any contravariant four-vector j^μ we may write the parity transformed vector $(j^0, -\vec{j}\,) \equiv j_\mu$ as a covariant vector. We will use this notation in the following.

Since the electromagnetic interaction $\mathcal{L}_{int}^{QED} = ej^\mu(x)A_\mu(x)$ respects C-, P- an T-invariance separately we immediately get the following transformation properties for the photon field:

$$\begin{array}{rclcl} U(C)\,A^\mu(x)\,U^{-1}(C) & = & -A^\mu(x) & & \\ U(P)\,A^\mu(x)\,U^{-1}(P) & = & (PA)^\mu(Px) & = & A_\mu(Px) \\ \bar{U}(T)\,A^\mu(x)\,\bar{U}^{-1}(T) & = & -(TA)^\mu(Tx) & = & A_\mu(Tx)\ . \end{array} \tag{2.36}$$

Notice that the charge parity for the photon is $\eta_C^\gamma = -1$.

In general we have

$$\begin{array}{rcl} U(C)\,V^\mu(x)\,U^{-1}(C) & = & \eta_C\,V^{+\mu}(x) \\ U(P)\,V^\mu(x)\,U^{-1}(P) & = & \eta_P\,V_\mu(Px) \\ \bar{U}(T)\,V^\mu(x)\,\bar{U}^{-1}(T) & = & -\eta_T\,V_\mu^+(Tx)\ . \end{array} \tag{2.37}$$

For real fields $V^\mu = V^{+\mu}$ in particular we have

$$V^\mu(x) \stackrel{P}{\to} \left\{ \begin{array}{rl} V_\mu(Px) & \text{vector field} \\ -V_\mu(Px) & \text{axial vector field}\ . \end{array} \right.$$

C) Dirac fields

$$\begin{array}{rcl} U(C)\,\psi_\alpha(x)\,U^{-1}(C) & = & i\,(\gamma^2\gamma^0)_{\alpha\beta}\,\bar{\psi}_\beta^T(x) \\ U(P)\,\psi_\alpha(x)\,U^{-1}(P) & = & (\gamma^0)_{\alpha\beta}\,\psi_\beta(Px) \\ \bar{U}(T)\,\psi_\alpha(x)\,\bar{U}^{-1}(T) & = & i\,(\gamma^2\gamma_5)_{\alpha\beta}\,\bar{\psi}_\beta^T(Tx) \end{array} \tag{2.38}$$

where the phases have been chosen conveniently. We observe that, in contrast to the boson fields, the transformation properties of the Dirac fields is by no means obvious. A short digression may help to show how these results come about:

By the arguments given above, under C, P and T the Dirac field must transform as $\psi_\alpha(x) \stackrel{C}{\to} (X_C)_{\alpha\beta}\bar{\psi}_\beta^T(x)$, $\psi_\alpha(x) \stackrel{P}{\to} (X_P)_{\alpha\beta}\psi_\beta(Px)$ and $\psi_\alpha(x) \stackrel{T}{\to} (X_T)_{\alpha\beta}\bar{\psi}_\beta^T(Tx)$. The transpose $\bar{\psi}^T$ appears just because we have to get a column spinor rather than a row spinor $\bar{\psi}$. The easiest way to determine the 4 x 4 matrices X_I is to use the Dirac equation.

Using the Hermitecity property $\gamma^{\mu+} = (\gamma^0, -\vec{\gamma}\,)$ of the $\gamma-$matrices we may write $\gamma^\mu\partial_\mu \stackrel{P}{\to} \gamma^{\mu+}\partial_\mu$ and $\gamma^\mu\partial_\mu \stackrel{T}{\to} -\gamma^{\mu+}\partial_\mu$. Therefore $(i\gamma^{\mu+}\partial_\mu - m)\,\psi(Px) = 0$ and $(i\gamma^{\mu+}\partial_\mu + m)\,\psi(Tx) = 0$. For the transpose of the adjoint spinor we have $(i\gamma^{\mu T}\partial_\mu + m)\,\bar{\psi}^T(x) = 0$ and $(i\gamma^{\mu *}\partial_\mu - m)\,\bar{\psi}^T(Tx) = 0$.

Now, since $U(I)\psi_\alpha(x)U^{-1}(I)$ satisfies the Dirac equation, we obtain:

C: $(i\gamma^\mu\partial_\mu - m)X_C\bar{\psi}^T(x) = 0$ and this requires $X_C^{-1}\gamma^\mu X_C = -\gamma^{\mu T}$. If we choose the arbitrary phase such that X_C is real and antisymmetric $X_C^T = -X_C$ we find $X_C = -i\gamma_5 B = i\gamma^2\gamma^0$. For specific representations of B see the Appendix.

P: $(i\gamma^\mu\partial_\mu - m)X_P\psi(Px) = 0$ which requires $X_P^{-1}\gamma^\mu X_P = \gamma^{\mu+}$ and hence $X_P = \gamma^0$ is the obvious choice.

T: $(i\gamma^\mu\partial_\mu - m)X_T\bar{\psi}^T(Tx) = 0$ and therefore $X_T^{-1}\gamma^\mu X_T = \gamma^{\mu *}$. Again we may choose the phase such that X_T is real and we find $X_T = iC = i\gamma^2\gamma_5$. For specific representations of C see the Appendix.

For the case of spin 1/2, we are ready now to verify the non-trivial result that for half odd integer spin particle and antiparticle have opposite inner P–parity. Equivalently, this means that a particle-antiparticle pair has negative inner parity.

To see this we compare the transformations of a Dirac field $\psi(x)$ and its charge conjugate $\psi^c(x) \equiv i\gamma^2\gamma^0\bar{\psi}^T(x)$ under space inversion: First we have

$$\begin{aligned} \psi(x) &\xrightarrow{P} \gamma^0\psi(Px) \\ \bar{\psi}^T(x) &\xrightarrow{P} \gamma^0\bar{\psi}^T(Px) \end{aligned}$$

where the second transformation law follows from the first one by taking the Hermitean conjugate, multiplication with γ^0 from the right and taking the transpose. When taking the Hermitean conjugate the unitarity of $U(P)$ must be used. For the charge conjugate field we then obtain

$$\psi^c(x) \xrightarrow{P} i\gamma^2\gamma^0 U(P)\,\bar{\psi}^T(x)\,U(P)^{-1} = i\gamma^2\gamma^0\gamma^0\bar{\psi}^T(Px)$$

and anticommuting one γ^0 to the left of γ^2 we find

$$\psi^c(x) \xrightarrow{P} -\gamma^0\psi^c(Px)$$

which proves that particle and antiparticle transform with opposite sign under parity (see Eq. 2.32). Consequences of this result will be considered in an example at the end of this section.

We finally note the following transformation properties of the **chiral fields** (**Weyl fields**) defined in Sec. 2.5

$$\begin{aligned} U(C)\,\psi_{L\alpha}(x)\,U^{-1}(C) &= i\,(\gamma^2\gamma^0)_{\alpha\beta}\,\bar{\psi}^T_{R\beta}(x) \\ U(P)\,\psi_{L\alpha}(x)\,U^{-1}(P) &= (\gamma^0)_{\alpha\beta}\,\psi_{R\beta}(Px) \\ \bar{U}(T)\,\psi_{L\alpha}(x)\,\bar{U}^{-1}(T) &= i\,(\gamma^2\gamma_5)_{\alpha\beta}\,\bar{\psi}^T_{L\beta}(Tx) \end{aligned} \tag{2.39}$$

and the same relations with L and R interchanged. Notice that parity (P) and charge conjugation (C) interchange left–handed (ψ_L) and right–handed (ψ_R) (Weyl-) fields, which means that a field which transforms under a given irreducible representation of $\mathcal{P}^\uparrow_+$ is transformed into a field which transforms under a different representation under $\mathcal{P}^\uparrow_+$. In order to avoid this, one has to describes fermions by the reducible Dirac representation, and not by the irreducible two component Weyl representation. The latter, at first sight, seems to be more natural for a description of spin 1/2 particles which have two independent degrees of freedom and not four. The reason for the preference given to the Dirac representation in particle physics are the parity conserving interactions QED and QCD, which must include both the L and the R fields in a symmetric manner. This is simply achieved by using the Dirac field $\psi = \psi_L + \psi_R$, as we know.

A very important consequence of local field theory is the **CPT-theorem:** Any $\mathcal{P}^\uparrow_+$ invariant field theory with normal commutation relations is CPT invariant. Let $\Theta = CPT$ where C, P and T may be applied in any order, there exists an anti-unitary operator $\bar{U}(\Theta)$ such that with appropriate phase conventions the transformation

$$\begin{aligned} &\text{Scalar field :} & \bar{U}(\Theta)\,\phi(x)\,\bar{U}^{-1}(\Theta) &= \phi^+(-x) \\ &\text{Dirac field :} & \bar{U}(\Theta)\,\psi(x)\,\bar{U}^{-1}(\Theta) &= i\gamma_5\psi(-x) \\ &\text{Photon field :} & \bar{U}(\Theta)\,A_\mu(x)\,\bar{U}^{-1}(\Theta) &= -A_\mu(-x) \end{aligned} \tag{2.40}$$

etc. of the fields is a symmetry of the theory (Lüders 1954, Pauli 1955, Jost 1957).

The basic reason for the validity of the CPT-theorem is the following: If we consider a Lorentz transformation $\Lambda \in L^\uparrow_+$ represented by a unitary operator $U(\vec{\chi}, \vec{\omega} = \vec{n}\,\theta)$, then the operator $U(\vec{\chi}, \vec{n}\,(\theta + 2\pi)) = -U(\vec{\chi}, \vec{n}\,\theta)$ is representing the same L-transformation. In a quantum field theory the mapping $\Lambda \to -\Lambda$ for $\Lambda \in L_+^\uparrow$, which is equivalent to $\theta : x \to -x$, should be a symmetry: the **invariance under four-dimensional reflections** .

The physical consequences are the following: CPT-symmetry implies that the masses (more generally the energy spectrum) of particles and antiparticles are equal

$$m = \bar{m}$$

for interactive theories. The best experimental test of this property is provided by the $K_L - K_S$ mass difference:

$$(M_K - M_{\bar{K}})/M_K \leq (M_L - M_S)/M_K \simeq 0.7 \times 10^{-14}$$

as an upper limit of the CPT-violation of the strong interactions. Furthermore the partial and total widths of particles and antiparticles are identical

$$\Gamma_{A \to B} = \Gamma_{\bar{A} \to \bar{B}}$$

if the initial and final states do not mix:

$$< \bar{A}|A > = < \bar{B}|B > = 0 \ \ .$$

The latter condition is not satisfied for the $(K^0, \bar{K}^0)$ and the $(B^0, \bar{B}^0)$ systems. Of course charged particles never mix. Here the best limits for the identity of particle and antiparticle widths are

$$(\tau^+ - \tau^-)/(\tau^+ + \tau^-) = \begin{cases} (1.5 \pm 3.4) \times 10^{-5} & \text{for} \quad \mu^\pm \\ (2.6 \pm 3.4) \times 10^{-4} & \text{for} \quad \pi^\pm \\ (5.7 \pm 4.6) \times 10^{-4} & \text{for} \quad K^\pm \end{cases}$$

where $\tau^\pm = \Gamma_\pm^{-1}$ is the life-time of the positively and negatively charged partners, respectively.

Another static property which must be equal for particles and their antiparticles is the magnetic moment (g-factors). The best limits are

$$\begin{aligned} g(e^+) - g(e-) &= (\ \ 1.0 \pm 4.2) \times 10^{-12} \\ g(\mu^+) - g(\mu^-) &= (-5.2 \pm 3.2) \times 10^{-8} \ \ . \end{aligned}$$

They are testing the CPT-invariance of the electromagnetic interaction[8].

The above properties also follow from charge conjugation invariance (C), which holds for strong and electromagnetic interactions only. Weak interactions violate parity P maximally ($V - A$ interaction) but conserve CP to good accuracy. Thus together with P also C is strongly violated in weak processes. Most strikingly this manifests itself such that a right–handed neutrino and a left–handed anti-neutrino seem not to exist in Nature[9]. Here the CPT-theorem yields a very nontrivial result. CP is weakly violated by weak interactions. Experimentally it was observed so

[8]When a measurement of some quantity is performed in reality, all effects of all possible kind of interactions in principle contribute. Since among the known interactions the weak ones violate P, C as well as CP, only CPT can be strictly conserved. The only condition then is that nature is described by a quantum field theory. When the C–violating weak contributions are smaller than the experimental accuracy, the equality of particle and anti-particle properties derive from C invariance alone. A corresponding statement can be made with respect to CP violating effects which in many cases are much more suppressed then the genuine weak interactions.

[9]Today we know from the observed neutrino oscillations, that also the latter neutrinos species must exist. In contrast to “ordinary helicity” neutrinos, they seem not to couple directly to ordinary matter, however. Such “sterile” wrong helicity neutrinos in first place seem to participate only via helicity flip transitions mediated by a tiny mass term.

far only in $(K^0, \bar{K}^0)$ decays. A crucial experimental goal is the observation of CP violation in the $(B^0, \bar{B}^0)$ system! In fact, according to present day knowledge, CP is an exact symmetry in the lepton sector as well as in the (u, d) quark sector which describes nucleons, pions, the $\rho-$ meson etc. Of course the CPT-theorem implies that always at least two of the tree discrete symmetries have to be violated simultaneously. In particular, CP **violation implies the violation of time reversal invariance** T.

If a symmetry $I = C$, P or CP is preserved by an interaction, which is universally true for the strong and electromagnetic interactions, the intrinsic parities η_I represent **multiplicatively conserved quantum numbers** in multiparticle reactions. This is obvious since multiparticle states are created (annihilated) from the vacuum by corresponding products of creation (annihilation) operators, with each operator carrying a corresponding I-parity factor.

Let us discuss this in some more detail for the P-parity as an example. Considering space-reflections the parity property of the spatial wave–function must be taken into account. If we consider a two-body reaction

$$A + B \to C + D$$

in the center of mass (CM) frame we may separate the radial and the angular part of the spatial wave–function

$$\psi_{LM}(x) = R(r)\, Y_L^M(\theta, \phi)$$

if the system is in a state of fixed angular momentum L. For a space reflection $r \to r\ ,\ \ \theta \to \pi - \theta\ ,\ \ \phi \to \phi + \pi$ the spherical harmonics change sign according to

$$Y_L^M(\pi - \theta, \phi + \pi) = (-1)^L\, Y_L^M(\theta, \phi)$$

which means that an angular momentum eigenstate has positive or negative parity depending on whether L is even or odd. We thus obtain

$$\eta_P^A \eta_P^B (-1)^L = \eta_P^C \eta_P^D (-1)^{L'}$$

if the initial an final states have angular momenta L and L', respectively. Generally, **the parity of a multiparticle (composite) system, in a state of angular momentum L, is the product of the inner parities of the individual particles (constituents) times the orbital momentum parity** $(-1)^L$.

Suppose now that a particle C can be **singly produced** in a reaction

$$A + B \to A + B + C\ .$$

Such a particle cannot carry any conserved charge-like quantum number $Q, B, L_\ell ...$ or fermion number. Notice that strong interactions conserve all flavors and thus also flavor must be conserved in this case. Thus C has to be a neutral boson and from

$$\eta_P^A \eta_P^B\, (-1)^L = \eta_P^A \eta_P^B \eta_P^C\, (-1)^{L'}$$

we conclude that

$$\eta_P^C = (-1)^{L+L'}$$

is the intrinsic parity of the particle C. Such an absolute intrinsic parity is assigned to the photon, the π^0 and the ρ^0. On the other hand if C is a fermion or carries any charge-like quantum numbers it can be **pair produced** only

$$A + B \to A + B + C + D\ .$$

In this case an unambiguous intrinsic parity $\eta_P^C \eta_P^D$ is obtained for the pair only. It is still meaningful to assign an intrinsic parity to such a particle **by convention**. In the parity conserving strong interaction physics one chooses $\eta_P = 1$ **for the normal baryons** $p, n, \Lambda, \ldots$. What we said here about the P-parity of course is true for the C-parity and the CP-parity.

We end this section by a discussion of some interesting examples which illustrate the interplay of discrete symmetries.

Examples

(a) P and C violation and CP conservation in weak leptonic decays

The leptonic decay $\pi^+ \to \mu^+ + \nu_\mu$ of the charged pion is a beautiful example to observe P and C violation under the condition of strict CP conservation in a weak interaction process. If we apply C and P to the given process we obtain processes which include neutrinos of the "wrong" helicity the couplings of which appear to be absent in Nature. However, the combined transformation CP=PC yields the observed decay $\pi^- \to \mu^- + \bar{\nu}_\mu$.

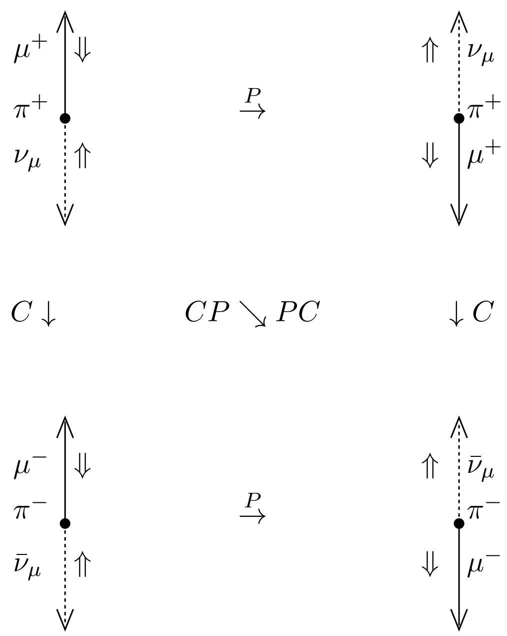


Fig. 2.3: Starting from the observed π^+ decay P and C map this decay into decays which do not exist. CP or PC however relate the observed π^+ and π^- decays.

(b) C parities of π^0 and η^0

Since the photon has C–parity $\eta_C^\gamma = -1$ as mentioned earlier, the observed C–conserving electromagnetic decays

$$\pi^0 \to \gamma\gamma \;, \quad \eta^0 \to \gamma\gamma$$

tell us that

$$\eta_C^{\pi^0} = \eta_C^{\eta^0} = +1 \quad .$$

The conservation of C–parity on the other hand implies that

$$\pi^0 \not\to \gamma\gamma\gamma \ , \ \ \eta^0 \not\to \pi^0\gamma \text{ or } \gamma\gamma\gamma$$

are forbidden decays.

(c) Spin, isospin and parity of the pions

The pions are the lightest of all strongly interacting particles. They are known to have integer spin and baryon number zero as can be seen from the production and absorption reactions (with d=(pn) denoting the deuteron)

$$\begin{aligned} &p + p \to \pi^+ + d \\ &\pi^+ + d \to p + p \qquad , \end{aligned}$$

which also shows that pions, like photons, may be created or annihilated singly in particle reactions.

① *Spin of the charged pion:*

A simultaneous measurement of production and absorption allows one to determine the spin of the charged pion. The idea is that the unpolarized cross-section depends in a particular way on the spin, since one has to average over the initial state polarizations and sum over to final state polarizations. For a reaction $A = a + b \to B = c + d$ we have (see Sec. (3.5) below for details on the notation)

$$\left(\frac{d\sigma}{d\Omega}\right)^{\text{unpol}}_{CM} = \frac{1}{(8\pi)^2 s}\frac{1}{2S_a+1}\frac{1}{2S_b+1}\sqrt{\frac{\lambda(s,m_c^2,m_d^2)}{\lambda(s,m_a^2,m_b^2)}}\sum_{\text{spins}} |< B|T|A >|^2 \quad .$$

By time reversal invariance the spin sum of the absolute square of the matrix element is the same for a reaction and its inverse. Therefore we obtain

$$\begin{aligned} &\left(\frac{d\sigma_{AB}}{d\Omega}\right)^{\text{unpol}}_{CM} (2S_a+1)\,(2S_b+1)\,\lambda(s,m_a^2,m_b^2) \quad = \\ &\left(\frac{d\sigma_{BA}}{d\Omega}\right)^{\text{unpol}}_{CM} (2S_c+1)\,(2S_d+1)\,\lambda(s,m_c^2,m_d^2) \end{aligned}$$

a relationship which is completely independent of the details of the interaction and usually is called "**principle of detailed balance**". For the above reaction using the fact that the spin of the proton is 1/2 the one of the deuteron 1 we find

$$\left(\frac{d\sigma_{\text{production}}}{d\Omega}\right)^{\text{unpol}}_{CM} \Bigg/ \left(\frac{d\sigma_{\text{absorption}}}{d\Omega}\right)^{\text{unpol}}_{CM} = \frac{3}{4}\,(2S_\pi+1)\,\frac{\lambda(s,m_\pi^2,M_d^2)}{\lambda(s,m_p^2,m_p^2)}$$

The experimental results require $S_{\pi^\pm} = 0$.

② *Spin of the neutral pion:*

The neutral pion predominantly decays electromagnetically into two photons $\pi^0 \to 2\gamma$. If we consider the decay in the rest frame of the pion, the two gammas are emitted mono-energetically, with energy $E_\gamma = m_{\pi^0}/2$, in opposite directions. The total angular momentum of the two spin 1 photons can only be even. If the state has orbital angular momentum zero we can have a spin 0 or a spin 2 state only (see Sec. 5). Since the π^0 is the isospin partner of the charged pions its spin must be zero.

③ *Isospin of the pions:*

A comparison of proton-proton (pp) and proton-neutron (pn) scattering experiments reveals a fundamental symmetry of the strong interactions, the (strong) **isospin symmetry** (charge independence of the nuclear force). The internal symmetry group is $SU(2)$ and the nucleons p and n form a doublet $\binom{p}{n}$. The isospin of the pions is then obtained by comparing the production of charged and neutral pions in the reactions

$$\begin{aligned} p + p &\to d + \pi^+ \\ p + n &\to d + \pi^0 \end{aligned} \qquad .$$

The pp initial state must be in an $I_i = 1$ state. The deuteron has $I_d = 0$. By conservation of isospin in strong interaction we must have $I_{\pi^+} = 1$. The pn state can be in singlet $I_i = 0$ or in a triplet $I_i = 1$ state, with equal probability. But, if $I_{\pi^0} = 1$ the final state must have $I_f = 1$ and thus only half of the initial states create a π^0. Therefore we expect

$$\sigma(p + n \to d + \pi^0)/\sigma(p + p \to d + \pi^+) = 1/2$$

which is confirmed by experiment. The pions form a isospin triplet where π^+, π^0, π^- correspond to the $I_3 = +1, 0, -1$ components.

④ *Parity of the pions:*

The occurrence of the reaction

$$\pi^- + d \to 2\,n \;\; ,$$

at rest, is possible only if the pion is a pseudo-scalar particle i.e. it has inner parity $\eta_P^{\pi^-} = -1$. In order to show this we apply the relation $\eta_P^{\pi^-}\, \eta_P^d\, (-1)^L \;=\; (\eta_P^n)^2\, (-1)^{L'}$. In the reaction under consideration the pion is trapped in an S-orbit of the deuteron, thus $L = 0$, and the total angular momentum of the initial state is equal to the spin of the deuteron which is 1. Since the orbital angular momentum is zero, the parity of the deuteron is $\eta_P^d = +1$. The final state exhibiting two neutrons is restricted by the Pauli exclusion principle. The triplet state $S = 1$ (parallel spins) is symmetric, the singlet state $S = 0$ (anti-parallel spins) is antisymmetric. Under the exchange of the two fermions the total wave function changes by the factor $(-1)^{(L+S+1)}$. Since only an antisymmetric state is admitted, $L + S$ must be even. Thus, in spectroscopic notation ${}^{2S+1}L_J$, possible states are ${}^1S_0$, ${}^3P_{2,1,0}$, ${}^1D_2$, ... , not however a ${}^3S_1$ state, for example. By angular momentum conservation only the ${}^3P_1$ state is possible, which has negative parity. Therefore the parity of the charged pion must be negative. Note that the parity of the proton is set to $\eta_P^p = +1$ by convention; the neutron, as the isospin partner of the proton, then also has parity $\eta_P^n = +1$.

While the reaction

$$\pi^- + d \to 2\,n + \gamma$$

occurs at a rate comparable to the reaction considered so far, the production of a single additional π^0

$$\pi^- + d \not\to 2\,n + \pi^0$$

seems to be strictly forbidden by experiment. Both processes are allowed by charge and baryon number conservation, but, the conservation of parity P in strong interactions implies that the second reaction is forbidden if $\eta_P^{\pi^0} = -1$.

The parity of the neutral pion may be determined again from the decay $\pi^0 \to 2\gamma$ under the assumption that the spin of the pion is 0, by analyzing the polarizations of the emitted photons. This is possible by studying the angular correlations of the two Dalitz pairs (e^+e^-–pairs) obtained by the conversion of the two virtual photons in $\pi^0 \to \overset{*}{\gamma} + \overset{*}{\gamma} \to e^+e^-e^+e^-$.

(d) Positronium and the intrinsic parity of a fermion–antifermion pair

It is perhaps unexpected that a fermion–antifermion system <u>must</u> have negative inner parity. This has important consequences for example for the properties of positronium, a hydrogen–like (e^+e^-) bound state. The wave function of the two fermions must be antisymmetric if we interchange at the same time the positions the spins and the charges[10]. The interchange of the positions is equivalent to a spatial reflection about the center of mass of the system which implies that the wave function multiplies with $(-1)^L$, where L is the orbital angular momentum. Under the exchange of the two spins, the spin wave–function is antisymmetric for S=0 and symmetric for S=1 (see Sec. 5), hence, the wave function gets multiplied by $(-1)^{1+S}$. Together with the charge conjugation C we obtain: $(-1)^L(-1)^{1+S}C = -1$ and thus $C = (-1)^{L+S}$. Because of the negative inner parity the system has parity $P = (-1)^{1+L}$ and thus $CP = (-1)^{1+S}$. The CP–invariance of the electromagnetic interaction thus implies that the spin S is an exact quantum number. In spectroscopic notation $^{2S+1}L_J$, the ground states are

State	Spin	Name	Charge-parity	Decays
$^1S_0$	$S = 0$	parapositronium	$C = \ \ 1$	$2\gamma, 4\gamma, \cdots$
$^3S_1$	$S = 1$	orthopositronium	$C = -1$	$3\gamma, 5\gamma, \cdots$

Positronium decays electromagnetically into photons. By energy–momentum conservation at least two photons must be in the final state. The $^3S_1$ state cannot decay into two photons because a two photon system cannot have total angular momentum 1. By its negative C-parity it cannot decay into any even number of photons (**Furry's theorem**). The $^1S_0$ state has even C-parity and cannot decay into an odd number of photons. The inner P-parity of positronium can be observed directly in the angular correlation of the emitted photons.

2.8 Exercises: Section 2

① Determine the matrix Λ^μ_ν for

a) a rotation by an angle φ about the z-axis

b) a special Lorentz transformation of velocity $\vec{v}$ in z-direction.

Write down the operators $U(\Lambda)$ for the above transformation.

② Show that for a general Poincaré transformation $U(\Lambda, a)$ the generators of the Poincaré group satisfy

$$U(\Lambda, a)P_\mu U(\Lambda, a)^{-1} = \Lambda^\nu{}_\mu \, P_\nu$$

[10]Formally, we have

$$\psi_\alpha(x)\psi^c_\beta(y) \to \psi^c_\beta(y)\psi_\alpha(x) = -\psi_\alpha(x)\psi^c_\beta(y)$$

by the anticommutation relations for the fermions fields.

and

$$U(\Lambda, a) M_{\mu\nu} U(\Lambda, a)^{-1} = \Lambda^\rho{}_\mu \, \Lambda^\sigma{}_\nu \, \left(M_{\rho\sigma} - P_\rho a_\sigma + P_\sigma a_\rho \right) \quad .$$

While the first equation tells us that P_μ transforms as a covariant four-vector, the second proves that $M_{\mu\nu}$ is a 2nd rank tensor only with respect to <u>homogeneous</u> Poincaré transformation.

③ Use the previous result to derive the Lie algebra of $\mathcal{P}_+^\uparrow$ by expanding $U(\Lambda, a)$ to first order. The result should be

$$[P_\mu, P_\nu] = 0 \quad ,$$

$$[P_\rho, M_{\mu\nu}] = -i \left(g_{\rho\mu} P_\nu - g_{\rho\nu} P_\mu \right) \quad ,$$

$$[M_{\mu\nu}, M_{\rho\sigma}] = i \left(g_{\mu\rho} M_{\nu\sigma} - g_{\mu\sigma} M_{\nu\rho} + g_{\nu\sigma} M_{\mu\rho} - g_{\nu\rho} M_{\mu\sigma} \right) \quad .$$

Give a physical interpretation.

④ Prove that

$$\frac{d^3p}{\omega_p} \quad \text{and} \quad \omega_p \delta^{(3)}(\vec{p}\,' - \vec{p}\,) \quad \text{with} \quad \omega_p = \sqrt{\vec{p}^{\,2} + m^2}$$

are relativistically invariant. Hint: Use

$$\Theta(p^0) \, \delta(p^2 - m^2) \, d^4p = \frac{d^3p}{2\omega_p} \quad .$$

⑤ Show that the Fourier transform of a distribution $\tilde{\phi}(p) = 2\pi\delta(p^2 - m^2)\tilde{\xi}(p)$ with support on the hyperbola $p^2 = m^2$ has the form

$$\phi(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2\omega_p} \left(e^{-ipx} \tilde{\xi}(\omega_p, \vec{p}\,) + e^{ipx} \tilde{\xi}(-\omega_p, -\vec{p}\,) \right)$$

which is a decomposition into a positive and a negative frequency part. Hint: Use

$$\int_{-\infty}^{+\infty} dp^0 \, \tilde{f}(p^0, \vec{p}\,) = \int_0^\infty dp^0 \, \left(\tilde{f}(p^0, \vec{p}\,) + \tilde{f}(-p^0, \vec{p}\,) \right)$$

and

$$\int_{-\infty}^{+\infty} d^3p \, e^{i\vec{p}\vec{x}} \, \tilde{g}(\vec{p}\,) = \int_{-\infty}^{+\infty} d^3p \, e^{-i\vec{p}\vec{x}} \, \tilde{g}(-\vec{p}\,) \quad .$$

Compare the form obtained with the representation of a free field in terms of creation and annihilation operators.

⑥ Show that

$$\Delta(x) = \frac{i}{(2\pi)^3} \int \frac{d^3p}{2\omega_p} \left(e^{-ipx} - e^{ipx} \right) \Big|_{p_0 = \omega_p = \sqrt{\vec{p}^{\,2} + m^2}} = 0 \quad \text{if} \quad x^2 < 0$$

⑦ For a system of free particles and antiparticles the four-momentum operator can be expressed in the simple form

$$P^\mu = \sum_r \int d\mu(p) p^\mu \left\{ a^+(\vec{p},r)a(\vec{p},r) + b^+(\vec{p},r)b(\vec{p},r) \right\} \quad .$$

Show that P^μ has the properties

$$\begin{aligned} P^\mu \mid 0 > &= 0 \\ P^\mu \mid \vec{p}, r, \alpha > &= p^\mu \mid \vec{p}, r, \alpha > \end{aligned}$$

and satisfies the commutation relations

$$\begin{aligned} [P^\mu, a \ \ (\vec{p},r)] &= -p^\mu a(\vec{p},r) \\ [P^\mu, a^+(\vec{p},r)] &= +p^\mu a^+(\vec{p},r) \quad \text{etc.} \end{aligned}$$

Give a physical interpretation of these properties.

⑧ For a Dirac field the charge operator is given by

$$Q = \int d^3x \, j^0(x) = \int d^3x \ : \psi_\alpha^+(x) \, \psi_\alpha(x) :$$

where

$$j^\mu = \bar{\psi}\gamma^\mu\psi \ , \quad \partial_\mu j^\mu = 0$$

is the conserved electromagnetic current operator. Show that

$$Q = \sum_r \int d\mu(p) \ \left\{ a^+(\vec{p},r)a(\vec{p},r) - b^+(\vec{p},r)b(\vec{p},r) \right\}$$

is time independent. Give a physical interpretation of Q by means of the commutation relations

$$[Q, a] = -a \ , \quad [Q, b^+] = -b^+ \ , \quad [Q, \psi] = -\psi \quad .$$

⑨ Prove that for a Dirac particle

$$\begin{aligned} (\Lambda_+)_{\alpha\beta} &= \frac{1}{2m} \sum_r u_\alpha(p,r)\bar{u}_\beta(p,r) = \frac{1}{2m}(\not{p} + m)_{\alpha\beta} \\ (\Lambda_-)_{\alpha\beta} &= -\frac{1}{2m} \sum_r v_\alpha(p,r)\bar{v}_\beta(p,r) = -\frac{1}{2m}(\not{p} - m)_{\alpha\beta} \end{aligned}$$

are projection operators with property

$$\begin{aligned} &\Lambda_+ u(p,s) = u(p,s) \quad , \quad \Lambda_+ v(p,s) = 0 \\ &\Lambda_- u(p,s) = 0 \qquad\qquad , \quad \Lambda_- v(p,s) = v(p,s) \quad . \end{aligned}$$

Give a physical interpretation of this result.

Note: In the space of four-spinors the usual Hermitean conjugation is replaced by going to the adjoint

$$\Gamma \rightarrow \Gamma^\dagger = \gamma^0 \Gamma^+ \gamma^0 \quad .$$

Thus, the usual Hermitecity $\Gamma = \Gamma^+$ requirement is replaced be self–adjointness requirement $\Gamma = \Gamma^\dagger$, because the L–invariant scalar product between two spinors u and v is $\bar{u}v \equiv u^+\gamma^0 v$, and not u^+v. The latter is not L–invariant.

⑩ Prove that

$$\Pi_{\pm} = \frac{1}{2}\left(1 \pm \gamma_5 \not{n}\right)$$

for n a space-like vector orthogonal to the momentum p of a Dirac particle

$$n^2 = -1 \;\; ; \quad n \cdot p = 0$$

are projection operators with the property

$$\begin{aligned} \Pi_{\pm} u(p,s) &= u(p,s)\, \delta_{s,\pm} \\ \Pi_{\pm} v(p,s) &= v(p,s)\, \delta_{s,\pm} \quad . \end{aligned}$$

Give a physical interpretation of the latter properties.

3 Interactions, S–matrix, Perturbation Expansion and Cross–sections

3.1 Interacting fields

The free fields are characterized by free homogeneous field equations

$$\begin{aligned} (\Box + m^2)\phi(x) &= 0 \\ (i\gamma^\mu \partial_\mu - m)\psi(x) &= 0 \\ (\Box + m^2)V_\mu(x) - \partial_\mu \left(\partial_\nu V^\nu(x)\right) &= 0 \ . \end{aligned} \tag{3.1}$$

These equations can be derived from an invariant Lagrangian density

$$\mathcal{L} = \mathcal{L}(\varphi, \partial\varphi)(x) \ ; \quad \varphi = \phi, \psi, V, \ldots$$

by means of Hamilton's principle of stationary action which states that for physical trajectories [11]

$$\delta \int d^4x \, \mathcal{L}(\varphi, \partial\varphi)(x) \ = 0 \quad \text{under variations} \quad \varphi \to \varphi + \delta\varphi \ .$$

The stationarity condition yields the Euler-Lagrange equations

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi} = \frac{\partial \mathcal{L}}{\partial \varphi} \ ; \quad \varphi = \phi, \psi, V, \ldots \tag{3.2}$$

which coincide with the field equations given above for the Lagrangian densities

$$\begin{aligned} \mathcal{L}_0 &= \frac{1}{2}\left(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\right) \\ \mathcal{L}_0 &= \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi \\ \mathcal{L}_0 &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{m^2}{2}V_\mu V^\mu \ ; \ F_{\mu\nu} \doteq \partial_\mu V_\nu - \partial_\nu V_\mu \ . \end{aligned} \tag{3.3}$$

Interactions are simply included now by adding **local invariant products of fields** to the free Lagrangian densities:

$$\mathcal{L}(x) = \mathcal{L}_0(x) + \mathcal{L}_{int}(x) \tag{3.4}$$

where $\mathcal{L}_{int}(x)$ is a sum of local monomials of fields, called (interaction-) **vertices**. The field products also may include derivatives of fields. By local we mean that all fields must be taken at the same space-time point x. L-invariance requires each monomial to be a Lorentz scalar. $\mathcal{L}_{int}(x)$ determines the **dynamics** of the system. Notice that adding a total derivative $\partial_\mu K^\mu$ with $K^\mu(x)$ some nonsingular vector field to the Lagrangian does not change the action $i\int d^4x \, \mathcal{L}(x)$ and hence leaves unaffected the dynamics of the system.

Since the bilinear terms are included in $\mathcal{L}_0(x)$ already, the interaction term in general includes products of at least three fields. The requirement of renormalizability (see later) in fact only

[11] Notice that we always have to subtract the vacuum expectation value from the Lagrangian

$$\mathcal{L}(x) \to \mathcal{L}(x) - <0|\mathcal{L}(x)\,|0>$$

in order that the action $\int d^4x\mathcal{L}(x)$ exists in the limit of an infinite space-time volume.

allows for products of at most four boson fields, for fermion fields only a pair $\bar{\psi} \cdots \psi$ times a boson field is admitted. The most familiar example is the electromagnetic interaction (see Sec. 4)

$$\mathcal{L}_{int}^{QED}(x) \;=\; -e j_{em}^{\mu}(x) A_{\mu}(x)$$

where $j_{em}^{\mu}(x) = \bar{\psi}(x)\gamma^{\mu}\psi(x)$ is the conserved electromagnetic current and $A_{\mu}(x)$ is the photon field.

Local interactions of local and causal fields allow to control the causality of the interacting theory. Of course the basic principles listed at the beginning of this section are required to hold for the interacting theory. In addition we assume that the vacuum is a **cyclic** vector, which means that a dense set of physical states may be generated by applying arbitrary products of field operators to the vacuum vector $|0>$.

The equations of motion for the interacting theory are given by the Euler-Lagrange equations for the full Lagrangian. For example, for a scalar field theory with self-interaction

$$\mathcal{L}_{int} = -\frac{\lambda}{4!}\phi^4\,(x) \tag{3.5}$$

the free homogeneous field equation is modified to

$$(\Box + m^2)\,\phi(x) = -\frac{\lambda}{3!}\phi^3\,(x) \doteq -\lambda \hat{J}\{\phi\}(x)\;. \tag{3.6}$$

Typically, an inhomogeneous non-linear source term which is proportional to a coupling constant, λ in our case, is added.

Interacting "particles" in general do not have definite mass, spin, charge etc. because they are no longer isolated systems. However, if the interaction is sufficiently short-ranged the effective interaction time may be considered to be finite and after a sufficiently long time ($t \to +\infty$) the system will be in a state of freely moving outgoing particles. Similarly we assume that before interaction sets in at early enough times ($t \to -\infty$) we start with a state of freely moving incoming particles. This <u>scattering theory picture</u> gives rise to the **asymptotic condition** for interacting fields: The assumption is that the interacting fields $\varphi(x)$ may be chosen such that for any states $|A>$ and $|B>$

$$\lim_{t\to -\infty/+\infty} < B|\varphi(t,\vec{x}\,)|A> = < B|\varphi_{in/out}(t,\vec{x}\,)|A> \tag{3.7}$$

where $\varphi_{in}(x)$ and $\varphi_{out}(x)$ describe a free incoming and a free outgoing particle, respectively. A field which satisfies the asymptotic condition is called an **interpolating field** (Heisenberg field). It is the asymptotic condition which provides the key for the particle interpretation of the interacting theory and exhibits particles as asymptotic states. It is via the asymptotic condition that the spin-statistics and particle–anti-particle crossing theorems and other properties we have discussed for free fields carry over to the interacting theory.

Although we shall assume that the asymptotic condition holds in the following we should mention here that there are serious limitations to its validity. The reason is that the simple scattering theory picture only works if no massless particles and if no bound states exist in a theory. In many important cases, therefore, it does not apply. A well known example is QED. Due to the masslessness of the photon charged particles are always accompanied by a cloud of an arbitrary number of soft photons such that, strictly speaking, a "one-electron" state is never (zero probability) produced in QED. In perturbation theory the unphysical nature of charged one-particle states manifests itself by the occurrence of **infrared (IR) singularities** in the charged particle wave-functions. This problem is called the **infrared problem** of QED and can be solved by a suitable modification of the definition of charged particle states (dressed states or infra-particle states) and a generalization of the asymptotic condition. In QCD, which formally looks like a non-Abelian version of QED, the situation is dramatically different. Quarks and gluons do not form asymptotically free particle states. They are confined in hadronic bound states. Only mesons, bound states of a quark–anti-quark pair, and (anti-) baryons, bound states of three (anti-) quarks, appear as scattering states. In this case the relationship between the fields in the Lagrangian (quarks and gluons) and the interpolating fields of the asymptotic states (hadrons) is complicated and related problems presently are unsolved. In any case, where bound states are involved we are hampered by our rather limited ability to handle the relativistic bound state problem.

Renormalization and the asymptotic condition.

So far we have not mentioned the serious problem of **ultraviolet (UV) divergences** which show up due to the fact that the interaction terms are ill-defined local products of singular fields. In mathematical terms, quantum fields are **operator–valued distributions**, which means that their matrix elements are singular functions, so–called **distributions**, which require smearing with suitable test–functions (wave packets). The reason for the problem of course is that we are working, for simplicity, with plane wave fields and states rather than with well behaved wave packets. Products of free fields already exhibit singularities

$$\lim_{x \to y} \varphi(x)\varphi(y) = ?$$

in the limit of coinciding arguments. By inspecting the two point functions $< 0|\varphi(x)\varphi(y)\,|0>$ of free fields in configuration space one finds terms like $\delta((x-y)^2)$, $1/(x-y)^2$, $\Theta((x-y)^2)$ and $\ln((x-y)^2)$ which are singular on the light cone and at the tip of the light cone. These **light–cone** and **short–distance singularities** in configuration space are "represented" in momentum-space as integrals which diverge at large momenta and therefore usually are called ultraviolet singularities. It is relatively easy to prove that vacuum expectation values of free fields are distributions which are regular functions if no arguments coincide. In order to obtain a well–defined starting point, one has to use the fact that any distribution may be defined as a limit of a sequence of ordinary functions

$$\Delta(x-y) = \lim_{r \to 0} \Delta^r(x-y) \ .$$

One thus starts with a **regularization** of the theory, parameterized by some parameter r. In the regularized version of the theory local products of fields are well-defined. The problem is that a regularization in general violates one or several of our basic principles like unitarity, causality, Poincaré invariance etc., in particular, symmetries may be violated and we have to make sure that at the end, in the limit $r \to 0$, all the required properties are recovered. At this stage we cannot go into these technical problems further. The essence of renormalization theory, which deals with these questions, is that after a re-parameterization of parameters and fields (renormalization) the regularization can be removed in a non-singular way. Renormalization has to do with imposing appropriate boundary conditions for the field equations. This is what we are going to discuss now. Thereby it is assumed that the theory is suitably regularized.

If the fields in the Lagrangian correspond to particles that can be detected in a scattering experiment, the asymptotic condition has to be used for the physical interpretation of the field theory model. For the moment we ignore the subtleties associated with the infrared problem of QED. In quantum field theory parameters like the mass or the coupling constant get modified by the interaction, as we shall see in examples later on. Therefore, if we want the parameters to have some prescribed values we have to fine tune the corresponding parameters in the Lagrangian in an appropriate way. This tuning is called **parameter renormalization** . The parameters in the original Lagrangian are called **bare parameters** the numerically prescribed parameters, determined usually by some experiment, are called **renormalized** or **physical** parameters. Similarly, nothing guarantees the fields in the Lagrangian to converge at large times to properly normalized free in- and out-fields. We thus have to renormalize the bare fields in the Lagrangian in order to obtain the proper interpolating fields. The renormalization of the fields is called **field renormalization** or **wave-function renormalization** .

On a formal level renormalization is needed to implement the asymptotic condition as a physical boundary condition. Let us illustrate this for the ϕ^4-model described by the bare Lagrangian

$$\mathcal{L} \; = \; \frac{1}{2}\,(\partial\phi)^2 - \frac{m^2}{2}\phi^2 - \frac{\lambda}{4!}\phi^4 \,.$$

The desired boundary conditions may be achieved by performing the following substitutions on the bare Lagrangian:

1. Parameter renormalization:

$$\begin{aligned} m^2 \;\; &\to \;\; m_0^2 \;\; = \;\; m^2 + \delta m^2 \\ \lambda \;\; &\to \;\; \lambda_0 \;\; = \;\; \lambda \; + \delta\lambda \end{aligned}$$

2. Wave-function renormalization:

$$\phi(x) \to \phi_0(x) \; = \sqrt{Z}\,\phi(x)$$

where the bare quantities appearing in the original Lagrangian are indexed by 0 now and the quantities without an index are the physical ones. As an interpolating field $\phi(x)$ has to satisfy

$$< 0|\phi(x)|p > \stackrel{t\to -/+\infty}{\longrightarrow} e^{-ipx} \quad p^0 = \sqrt{\vec{p}^{\,2} + m^2}$$

for a scalar one-particle state $|p>$ of physical mass m. As a first step we **re-parameterize** the bare Lagrangian in terms of the renormalized parameters and fields as follows:

$$\begin{aligned} \mathcal{L} \;\; &= \;\; \frac{1}{2}\,(\partial\phi_0)^2 - \frac{m_0^2}{2}\phi_0^2 - \frac{\lambda_0}{4!}\phi_0^4 \\ &= \;\; \frac{Z}{2}\,(\partial\phi)^2 - \frac{(m^2+\delta m^2)\,Z}{2}\phi^2 - \frac{(\lambda+\delta\lambda)\,Z^2}{4!}\phi^4 \,. \end{aligned}$$

In a second step we now split the Lagrangian into a free part and an interaction part where the free part is chosen such that it describes the asymptotic free field of mass m:

$$\begin{aligned} \mathcal{L} \;\; &= \;\; \mathcal{L}_0 + \mathcal{L}_{int} \\ \mathcal{L}_0 \;\; &= \;\; \frac{1}{2}\,(\partial\phi)^2 - \frac{m^2}{2}\phi^2 \\ \mathcal{L}_{int} \;\; &= \;\; -\frac{\lambda}{4!}\phi^4 + \Delta\mathcal{L} \\ \Delta\mathcal{L} \;\; &= \;\; -\frac{\lambda_0 Z^2 - \lambda}{4!}\phi^4 + \frac{Z-1}{2}\,(\partial\phi)^2 - \frac{m_0^2 Z - m^2}{2}\phi^2 \\ &= \;\; \frac{c}{4!}\phi^4 + \frac{b}{2}\,(\partial\phi)^2 + \frac{a}{2}\phi^2 \end{aligned} \tag{3.8}$$

The terms of $\Delta\mathcal{L}$ are called **counter terms** and the counter term coefficients a, b and c have to be adjusted by the physical boundary conditions. The number of counter terms is equal to the number of conditions we have to satisfy, namely, fixing the mass, the coupling constant and the normalization of the field. A theory is called **renormalizable** if after re-normalizing the parameters and fields of the original (bare) Lagrangian the regularization can be removed in an unambiguous manner. While the relationship between the bare and the renormalized parameters and fields become singular as the regularization is removed the matrix elements of the renormalized fields expressed as functions of the renormalized parameters remain finite.

Formal solution of the field equations.

We briefly illustrate how we may obtain a solution of the field equations for the interacting theory. In order to keep notation as simple as possible we consider the scalar self-interacting ϕ^4-model introduced above. First we notice that a field equation may always be written in an equivalent way as an integral equation, the so-called **Yang-Feldman equation**. This is possible with the help of the **Green function** determined by the **point-source** equation

$$(\Box + m^2)\,\Delta_0(z) = -\delta^{(4)}(z)$$

plus suitable boundary conditions, to be specified in a moment. The minus sign for the source term is a pure matter of convention. Let us denote by $\phi_0(x)$ a solution of the free field equation, then the interacting field

$$\hat{\phi}(x) = \phi_0(x) + \lambda \int d^4y\, \Delta_0(x-y)\; \hat{J}\{\hat{\phi}\}(y)$$

is a solution of the full field equation. This can be verified easily by applying the Klein-Gordon operator to $\hat{\phi}(x)$. If λ is small enough we expect that the interaction term can be treated as a perturbation. In this case we may try to solve the integral equation by iteration starting from the free field solution $\hat{\phi}^{(0)} = \phi_0$:

$$\hat{\phi}^{(n)}(x) = \phi_0(x) + \lambda \int d^4y\, \Delta_0(x-y)\; \hat{J}\{\hat{\phi}^{(n-1)}\}(y)$$

such that

$$\hat{\phi}(x) = \lim_{n\to\infty} \hat{\phi}^{(n)}(x) = \mathcal{F}\{\phi_0\}(x)$$

where $\mathcal{F}\{\phi_0\}$ is a functional of the free field $\phi_0(x)$. This iteration defines a **perturbation expansion** and yields the solution as a formal power series in the coupling constant λ.

We call it formal since we do not expect the series to converge. We merely expect it to be an **asymptotic expansion** meaning that the partial sums $s_N(\lambda) = \sum_{n=0}^{N} c_n \lambda^n$ (result to finite order in the perturbation expansion) for finite N and for sufficiently small λ give a good approximation of the generally unknown exact answer $s(\lambda)$, while the sequence s_N does not converge for $N \to \infty$. One can expect that for a given value of $\lambda (\ll 1)$ there exists an optimal choice for N such that $|s_N(\lambda) - s(\lambda)|$ reaches a minimum at $N = N_0$. N_0 may be estimated in some cases. In general it is unknown, however.

Now we have to impose the appropriate boundary condition which for scattering problems is the asymptotic condition. We thus have to work with the interpolating field and the renormalized parameters. The renormalized field equation reads:

$$\begin{aligned}
(\Box + m^2)\,\phi(x) &= \frac{\partial \mathcal{L}_{int}}{\partial \phi(x)} - \partial_\mu \frac{\partial \mathcal{L}_{int}}{\partial \partial_\mu \phi(x)} = -\lambda J\{\phi\}(x) \\
-\lambda J\{\phi\}(x) &= -\frac{\lambda}{3!}\phi^3(x) + \frac{c}{3!}\phi^3(x) - b\,\Box\,\phi(x) + a\,\phi(x)
\end{aligned}$$

and the renormalized Yang-Feldman equation is given by

$$\begin{aligned}\phi(x) &= \phi_{in}\,(x) + \lambda \int d^4y\, \Delta_{ret}(x-y)\, J\{\phi\}(y)\\ \phi(x) &= \phi_{out}(x) + \lambda \int d^4y\, \Delta_{adv}(x-y)\, J\{\phi\}(y)\end{aligned}$$

in terms of the in- and out-fields, respectively. The Green functions are determined by

$$(\Box + m^2)\, \Delta_{ret,adv}(z) = -\delta^{(4)}(z)$$

and must satisfy the boundary conditions

$$\begin{aligned}\Delta_{ret}(z) &= 0 \quad \text{for} \quad z^0 < 0\\ \Delta_{adv}(z) &= 0 \quad \text{for} \quad z^0 > 0\ .\end{aligned}$$

We easily convince ourselves that these boundary conditions guarantee that $\phi(x)$ fulfills the asymptotic condition. The two Green functions may be represented in terms of the causal free field commutator function $\Delta(z)$ (defined in Sec. 2.4) as

$$\begin{aligned}\Delta_{ret}(z) &= \Theta(z^0) \quad \Delta(z)\\ \Delta_{adv}(z) &= -\Theta(-z^0) \quad \Delta(z)\end{aligned}$$

and have the property

$$\Delta_{ret}(z) \quad - \quad \Delta_{adv}(z) = \Delta(z)\ .$$

For obvious reasons $\Delta_{ret(adv)}$ is called retarded (advanced) commutator function. Since $\Delta(z) = 0$ for $z^2 < 0$ the theta-function

$$\Theta(x) = \begin{cases} 1 & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x < 0 \end{cases}$$

just is a prescription to pick the forward (retarded) or the backward (advanced) light cone. Therefore Δ_{ret} and Δ_{adv} are Lorentz invariant Green functions.

Taking the difference of the two Yang-Feldman equations we find

$$\phi_{out}(x) \quad = \quad \phi_{in}(x) + \lambda \int d^4y\, \Delta(x-y)\, J\{\phi\}(y)$$

which proves that the in- and out-fields are related in a causal manner also for the interacting theory. As explained before we could solve this equation by iteration as a perturbation series in the renormalized coupling constant and would obtain as a result a functional $\phi_{out} = \mathcal{F}\{\phi_{in}\}$. Rather than going this direct way, we shall proceed in a somewhat different more elegant manner to find $\mathcal{F}$ in the following.

If scattering takes place

$$\phi_{out} \neq \phi_{in} \quad ,$$

however, we expect that the in-states and the out-states span the same physical Hilbert space in which case there exists a unitary transformation, the **scattering operator** S which maps the in-field into the out-field

$$\phi_{out}(x) = S^{-1}\phi_{in}(x)S = \phi_{in}(x) + S^{-1}[\phi_{in}(x), S]\ . \tag{3.9}$$

This formula directly compares to a Yang-Feldman equation. It is clear that the determination of $\mathcal{F}$ is equivalent to the determination of S as a functional of the in-fields.

Since the S–operator must have a representation as a functional of the in-fields we may write

$$S = 1 + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^4x_1 \cdots d^4x_n \, C_n(x_1, \cdots, x_n) \, W_n(x_1, \cdots, x_n)$$

where

$$W_n(x_1, \cdots, x_n) = : \phi_{in}(x_1) \cdots \phi_{in}(x_n) :$$

is a normal–ordered[12] product of n in-fields. This expression is just a definition of the symmetric coefficient functions C_n the explicit form of which is not important here. We have to calculate the commutator $[\phi_{in}(x), S]$. The commutator $[\phi_{in}(x), W_n]$ may be easily calculated. In the first term we commute $\phi_{in}(x)$ n-times to the right. Each time we are commuting $\phi_{in}(x)$ with one of the fields, $\phi_{in}(x_j)$ say, from the product W_n we obtain the c-number free field commutator $i\Delta(x - x_j)$ times the product W_n but with $\phi_{in}(x_j)$ missing now. Thus

$$[\phi_{in}(x), W_n] = i \sum_{j=1}^{n} \Delta(x - x_j) \frac{\partial W_n}{\partial \phi_{in}(x_j)} .$$

Using this and the symmetry of the C_n's we obtain after suitable relabeling of the summation indices

$$[\phi_{in}(x), S] = - \int d^4y \, \Delta(x - y) \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^4x_1 \cdots d^4x_n \, C_{n+1}(y, x_1, \cdots, x_n) \, W_n(x_1, \cdots, x_n) .$$

The sum is in fact what we get as a result of a functional differentiation [13]

$$-i \frac{\delta S}{\delta \phi_{in}(y)}$$

of the S–operator. Thus we obtain

$$\phi_{out}(x) = \phi_{in}(x) + \int d^4y \, \Delta(x - y) \, iS^{-1} \frac{\delta S}{\delta \phi_{in}(y)} .$$

The comparison with the Yang-Feldman equation tells us that

$$\begin{aligned} iS^{-1} \frac{\delta S}{\delta \phi_{in}(y)} &= \lambda J\{\phi\}(y) \\ &= -\frac{\partial \mathcal{L}_{int}}{\partial \phi(y)} \\ &= -\frac{\delta}{\delta \phi(y)} \int d^4x \, \mathcal{L}_{int}(x) . \end{aligned}$$

[12] this notion is defined and discussed in Sec. 3.4

[13] A **functional** F{J} is a quantity F which depends on a space-time function $J(x)$. We define the **functional derivative** by

$$\frac{\delta F\{J\}}{\delta J(x)} = \lim_{\epsilon \to 0} \frac{F\{J'\} - F\{J\}}{\epsilon}$$

with $J'(y) = J(y) + \epsilon\delta(y - x)$ or, equivalently,

$$\int d^4x \, f(x) \frac{\delta F\{J\}}{\delta J(x)} = \lim_{\epsilon \to 0} \frac{F\{J + \epsilon \, f\} - F\{J\}}{\epsilon}$$

for smooth test functions $f(x)$.

This is a functional differential equation. We are not yet able to solve this equation because the l.h.s is represented in terms of the in-fields while the r.h.s. is given in terms of the interpolating fields. However, with the help of the S–operator it is not difficult to find the interpolating field as a functional of the in-field. The interpolating field must satisfy the asymptotic conditions:

$$\begin{aligned}\lim_{t\to-\infty}\phi(x) &= \phi_{in}\,(x) = S^{-1}S\,\phi_{in}(x)\\ \lim_{t\to+\infty}\phi(x) &= \phi_{out}(x) = S^{-1}\phi_{in}(x)\,S\,.\end{aligned}$$

Thus besides the common factor S^{-1} we need a function which yields

$$S\,\phi_{in}(x)\quad\text{for}\quad t\to-\infty$$

on the one hand and

$$\phi_{in}(x)\,S\quad\text{for}\quad t\to+\infty$$

on the other hand. This can be achieved be using the so called **time-ordered product** (T–product), defined by

$$T\{A(x),B(y)\}\doteq\Theta(x^0-y^0)A(x)\,B(y)\pm\Theta(y^0-x^0)B(y)\,A(x)\tag{3.10}$$

for the simplest case of two fields $A(x)$ and $B(y)$. The "–"-sign in the second term applies for fermion-fields which must be anti-commuting. For a product of n fields this generalizes to

$$T\{A_1(x_1),A_2(x_2),\cdots,A_n(x_n)\}\doteq(-1)^F\;A_{i_1}(x_{i_1}),A_{i_2}(x_{i_2}),\cdots,A_{i_n}(x_{i_n})\tag{3.11}$$

where $(i_1,i_2,\cdots,i_n)$ is the permutation of $(1,2,\cdots,n)$ for which $x^0_{i_1}>x^0_{i_2}>\cdots>x^0_{i_n}$. The sign $(-1)^F$ accounts for the Fermi statistics of fermion-fields, hence F denotes the number of fermion-fields which have to be anti-commuted in the permutation $(1,2,\cdots,n)\to(i_1,i_2,\cdots,i_n)$.

When a field is integrated over, the time-ordering prescription must be performed before integration, for example, let $\mathcal{B}=\int d^4y\,B(y)$ then

$$T\{A(x),\mathcal{B}\}\doteq A(x)\int_{-\infty}^{x^0}d^4y\,B(y)+\int_{x^0}^{+\infty}d^4y\,B(y)\,A(x)\,.\tag{3.12}$$

This generalizes in an obvious way to functionals where multiple integrals are involved like the expression for the S–operator in terms of the in-fields. Notice that under the T–prescription all operators commute (boson-fields) or anti-commute (fermion-fields).

Using these definitions we obtain

$$\Big\|\;\phi(x)\;=S^{-1}T\{\phi_{in}(x),\;S\}\tag{3.13}$$

for the interpolating field. This generalizes to

$$F\{\phi\}(x)\;=S^{-1}T\{F\{\phi_{in}\}(x),\;S\}$$

for any regular functional F of the fields.

With the help of this important result we obtain

$$\frac{\delta S}{\delta\phi_{in}(y)}\quad=\quad T\{\frac{\delta\mathcal{A}_{int}}{\delta\phi_{in}(y)},S\}$$

where

$$\mathcal{A}_{int} = i \int d^4x \, \mathcal{L}_{int}\{\phi_{in}\}(x)$$

denotes the interaction part of the action expressed in terms of the in-fields and we have skipped the common factor iS^{-1}. Now, using the fact that the operators commute under the T–product we easily verify that the solution of this functional differential equation reads

$$\boxed{\begin{array}{rcl} S &=& T\left(e^{\mathcal{A}_{int}}\right)_{\otimes} \\ &=& T\left(e^{i\int_{-\infty}^{+\infty} d^4x \, \mathcal{L}_{int}\{\phi_{in}\}(x)}\right)_{\otimes} \\ &=& \sum_{n=0}^{\infty} \frac{i^n}{n!} \int d^4x_1 \cdots d^4x_n \, T\{\mathcal{L}_{int}^{(in)}(x_1), \cdots, \mathcal{L}_{int}^{(in)}(x_n)\}_{\otimes} \end{array}} \tag{3.14}$$

where $\mathcal{L}_{int}^{(in)}(x) = \mathcal{L}_{int}\{\phi_{in}\}(x)$ and the prescription $\otimes$ indicates that the solution has to be normalized such that

$$_{in}< 0|S|0 >_{in}= 1 \;\Leftrightarrow\; S = \frac{S_{formal}}{_{in}< 0|S_{formal}|0 >_{in}} \tag{3.15}$$

which means to skip all vacuum-to-vacuum transition amplitudes [14]. This is our key result. The S–operator is given in terms of the in-fields and can be evaluated order-by-order in a perturbative expansion.

Dimensional counting and renormalizability.

Let us add here a brief consideration of the dimensions (in energy units) of different quantities. Dimensional counting will provide a simple rule (necessary condition) to check the renormalizability of a field theory. Let us denote the dimension of space-time by d. Of course we will set $d = 4$ at the end.

We start from the observation that the integrated Lagrangian density

$$\mathcal{A} = i \int d^dx \, \mathcal{L}(x) \; ,$$

called **action** and measured in units of $\hbar = 1$ is a dimensionless quantity. Therefore the Lagrangian must have dimension $dim \; \mathcal{L}(x) = d$ in mass units. By inspection of the individual terms of the bilinear (free) part of the Lagrangian we find that the fields must carry a dimension. Taking into account that a derivative has dimension 1 we find the following dimensions for the fields:

$$\begin{array}{rcrcl} \partial_\mu\phi\partial^\mu\phi &:& dim \; \phi &=& \frac{d-2}{2} \\ \bar{\psi}\gamma^\mu\partial_\mu\psi &:& dim \; \psi &=& \frac{d-1}{2} \\ (\partial_\mu V_\nu - \partial_\nu V_\mu)^2 &:& dim \; V_\mu &=& \frac{d-2}{2} \; . \end{array}$$

Hence, in $d = 4$ dimensions, boson fields have dimension 1 while spin 1/2 fermion fields carry dimension 3/2. The dimension of a vertex is given by the sum of the dimensions of the fields plus

[14] The functional differential equation obviously determines S up to a multiplicative c-number factor only. This factor is fixed by the physical normalization condition $S|0 >_{in}= |0 >_{in}$ which is not automatically satisfied. On a purely formal level $S_{formal} = T(\exp \mathcal{A}_{int})$ seems to solve the functional differential equation. However, closer inspection shows that this expression is well defined only for a system in a finite space-time box of volume VT. One then obtains for large finite volumes $_{in} < 0|S_{formal}|0 >_{in}= e^{ifVT}$ and the "thermodynamic limit" $VT \to \infty$ cannot exist, because $f \neq 0$ for any interacting theory (Haag theorem, Haag 1955). The limit only exists **after** the proper normalization has been imposed.

the number of derivatives present at the vertex. Thus a vertex $\mathcal{V}$ given by a product of n_B boson fields and n_F Fermi fields and which includes n_D derivatives has dimension

$$dim\mathcal{V} = n_B\,\frac{d-2}{2} + \,n_F\,\frac{d-1}{2} + n_D\;.$$

One can show that a field theory can be renormalizable only if all vertices in the Lagrangian have dimensions $\leq d$. In $d = 4$ dimensions this yields the very severe restriction $dim\mathcal{V} = n_B + \frac{3}{2}n_F + n_D \leq 4$ for the possible interaction vertices. Notice that Lorentz invariance requires fermion fields to enter in pairs $\bar{\psi}\cdots\psi$. Since we require at least 3 fields: $dim\mathcal{V} \geq 3$. Allowed are vertices of the type

$$\begin{array}{cccc}
\phi^4\;, & V_\mu V^\mu V_\nu V^\nu\;, & V_\mu V^\mu \partial_\nu V^\nu\;, & V_\mu V^\nu \partial_\nu V^\mu\;,\\
\phi^2\partial_\nu V^\nu\;, & \phi V^\nu \partial_\nu \phi\;, & \phi^3\;, & V_\mu V^\mu \phi\;,\\
\bar{\psi}\gamma^\mu\psi V_\mu\;, & \bar{\psi}\gamma^\mu\gamma_5\psi V_\mu\;, & \bar{\psi}\psi\phi\;, & \bar{\psi}\gamma_5\psi\phi\;\;.
\end{array}$$

Notice that repeated fields may stand for different fields of the given type.

This dimensional counting rule provides a <u>necessary condition</u> for a model to be renormalizable. If no spin 1 fields are involved it is also a <u>sufficient condition</u>. However, for models which involve spin 1 bosons, like the models of strong, weak and electromagnetic interactions, the question of renormalizability is much more involved mainly because a viable theory not only has to be renormalizable but at the same time must be unitary. It turns out that only locally gauge invariant vector boson models, the Abelian and non-Abelian gauge theories, are renormalizable and unitary (see Sec. 6 and 8).

Models with vertices of the type

$$\bar{\psi}\gamma^\mu\psi\bar{\psi}\gamma_\mu\psi\;,\quad \bar{\psi}\sigma^{\mu\nu}\psi F_{\mu\nu}\quad \text{or}\quad \bar{\psi}\gamma^\mu\gamma_5\psi\partial_\mu\phi$$

can not be renormalizable. Nevertheless, models exhibiting such interaction terms often appear as low energy effective models as we shall see later. A corresponding remark applies to models which include fields of higher spin like spin 3/2, spin 2 etc.

3.2 Scattering processes, the S–matrix

In the last subsection we introduced the S–operator as a unitary operator which provides a transformation between the free in-fields and free out-fields

$$\phi_{out}(x) = S^{-1}\phi_{in}(x)S\;\;.$$

We obtained a formal expression Eq. (3.14) for S which allows to compute the properties of the interacting theory in terms of free fields. Here we first consider some general properties of scattering matrix elements.

We are mainly interested in scattering, production and decays of particles. For simplicity we shall assume that the interaction is short-ranged (no massless particles), such that particles at times $t \to \pm\infty$ are described by free asymptotic multiparticle states. We denote them as follows:

$$\begin{array}{cccc}
\text{initial states:} & \mid\alpha,\,\text{in}> & \text{in-states} & (t\to-\infty)\\
\text{final states:} & \mid\beta,\,\text{out}> & \text{out-states} & (t\to+\infty)
\end{array}$$

α and β are the labels of the states and include mass, spin, momentum and other quantum numbers.

The quantum mechanical transition probability is given by

$$w(\beta,\alpha) = |<\beta,\mathrm{out} \mid \alpha,\mathrm{in}\ >|^2 \quad .$$

If scattering takes place

$$\mid \alpha,\mathrm{out}> \neq \mid \alpha,\mathrm{in}> \quad .$$

However, we expect that the in-states and the out-states span the same physical Hilbert space

$$\mathcal{H}_{\mathrm{in}} \simeq \mathcal{H}_{\mathrm{out}} \simeq \mathcal{H}_{\mathrm{phys}}$$

i.e. we assume **asymptotic completeness** of the states. In this case there exists a **unitary** transformation, the **scattering matrix** or S-matrix, which maps the in-state onto the out-states

$$\mid \alpha,\mathrm{out}> = S^+ \mid \alpha,\mathrm{in}> \quad .$$

For the vacuum state

$$\mid 0,\mathrm{out}> = \mid 0,\mathrm{in}> = \mid 0> \quad .$$

Similarly, for one particle states of **stable particles**

$$\mid \alpha,\mathrm{out}> = \mid \alpha,\mathrm{in}> = \mid \alpha>\ ,$$

since a single particle cannot scatter. The conservation of probability requires the unitarity

$$SS^+ = S^+S = 1 \ \Leftrightarrow\ S^+ = S^{-1} \tag{3.16}$$

of the S-matrix.

Using the completeness relation $\sum_\beta \mid \beta\ \mathrm{out}><\beta\ \mathrm{out}\mid = 1$ of the out-states we obtain

$$\begin{aligned}
\mid \alpha,\mathrm{in}> &= S \mid \alpha,\mathrm{out}> \\
&= \sum_\beta \mid \beta,\mathrm{out}><\beta,\mathrm{out} \mid S \mid \alpha,\mathrm{out}> \\
&= \sum_\beta \mid \beta,\mathrm{out}> S_{\beta\alpha}
\end{aligned}$$

where $S_{\beta\alpha}$ are the **S-matrix elements** . Using $\mid \alpha,\mathrm{out}>= S^+ \mid \alpha,\mathrm{in}>$ and $SS^+ = 1$ we observe that $S_{\beta\alpha}$ can be evaluated either in the out-basis or in the in-basis:

$$\begin{aligned}
S_{\beta\alpha} &= <\beta,\ \mathrm{out} \mid S \mid \alpha,\ \mathrm{out}> \\
&= <\beta,\ \mathrm{in} \mid SSS^+ \mid \alpha,\ \mathrm{in}> \\
&= <\beta,\mathrm{in} \mid S \mid \alpha,\mathrm{in}>
\end{aligned} \tag{3.17}$$

In the following we briefly discuss some symmetry properties of S-matrix elements.

Symmetries of the S-matrix

A symmetry of S is represented by unitary transformations $U = e^{i\sum_i \omega_i G_i}$ with ω_i a set of real parameters and G_i the Hermitean generators of the symmetry transformations (see Sec. 5 below). Symmetries of the scattering matrix not only imply relation between matrix elements but also **selection rules**:

Let

i) G be a generator of a symmetry transformation such that $[G, S] = 0$

ii) and $\mid \alpha > and \mid \beta >$ be eigenstates of G

$$\begin{aligned} G \mid \alpha > &= g_\alpha \mid \alpha > \\ G \mid \beta > &= g_\beta \mid \beta > \end{aligned}$$

then

$$\begin{aligned} < \beta \mid SG \mid \alpha > &= g_\alpha < \beta \mid S \mid \alpha > \\ < \beta \mid GS \mid \alpha > &= g_\beta < \beta \mid S \mid \alpha > \end{aligned}$$

and hence

$$(g_\alpha - g_\beta) < \beta \mid S \mid \alpha > = 0 \, .$$

This tells us that $S_{\beta\alpha}$ can be different from zero only if $\mid \alpha >$ and $\mid \beta >$ are eigenstates of G with the same eigenvalue. We thus have a **conservation law**, $g_\alpha = g_\beta$, or a **selection rule** , if $g_\alpha \neq g_\beta$ then $S_{\beta\alpha} = 0$, associated with a quantum number g . For absolutely conserved quantities like charge, baryon number and the lepton numbers the selection rules are called super selection rule.

We now list some properties of S-matrix elements which derive from the space-time symmetries. We consider multiparticle states where each particle is characterized by its momentum p, spin j and helicity λ.

1. Translation invariance

Translation invariance

$$\begin{aligned} & < \text{out } p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots \text{in} > \\ = \; & < \text{out } p_1', j_1', \lambda_1', \ldots \mid U^{+}(1,a) \, U(1,a) \mid p_1, j_1, \lambda_1, \ldots \text{in} > \\ = \; & e^{iP_i a - iP_f a} < \text{out } p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots \text{in} > \end{aligned}$$

implies total four-momentum conservation

$$\begin{aligned} & < \text{out } p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots \text{in} > \\ = \; & \delta^{(4)}(P_f - P_i) \, \tilde{S}(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots) \, . \end{aligned}$$

By $P_f = \sum p_i'$ we denoted the total outgoing momentum by $P_i = \sum p_i$ the total incoming momentum.

If no interaction takes place, like for free fields, we have a trivial S-matrix $S = S_0 \equiv 1$ and hence $\mid \alpha, \text{ out} >= \mid \alpha, \text{ in} >$. Consequently

$$\begin{aligned} & < \text{out } p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots \text{in} >_0 = I(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots) \\ = \; & (2\pi)^3 \, 2\omega_{p_1} \, \delta^{(3)}(\vec{p}_1 - \vec{p}_1{}') \, \delta_{j_1 j_1'} \, \delta_{\lambda_1 \lambda_1'} \, (2\pi)^3 \, 2\omega_{p_2} \, \delta^{(3)}(\vec{p}_2 - \vec{p}_2{}') \, \delta_{j_2 j_2'} \, \delta_{\lambda_2 \lambda_2'} \cdots \end{aligned}$$

It is convenient and customary to split off the identity from the S–matrix and to define the T**–matrix** by

$$\Big\| \; S = 1 + i\,(2\pi)^4\,\delta^{(4)}(P_f - P_i)\,T \quad . \tag{3.18}$$

For the matrix elements this reads

$$\begin{aligned} S\,(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots) \;&=\; I(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots) \\ &\quad + i\,(2\pi)^4 \delta^{(4)}(P_f - P_i)\,T\,(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots) \end{aligned}$$

2. Lorentz invariance

   Lorentz invariance implies

$$\sum_{\lambda_1', \ldots, \lambda_1, \ldots} \mid T\,(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots)\mid^2 = \sum_{\lambda_1', \ldots, \lambda_1, \ldots} \mid T\,(\Lambda p_1', j_1', \lambda_1', \ldots \mid \Lambda p_1, j_1, \lambda_1, \ldots)\mid^2$$

   Summation over the helicities (or 3^{rd} components of spin) is necessary here because these are coordinate dependent state labels.

3. Parity invariance (in case parity is a symmetry)

$$\mid T\,(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots)\mid^2 = \mid T\,(P\,p_1', j_1', -\lambda_1', \ldots \mid P\,p_1, j_1, -\lambda_1, \ldots)\mid^2$$

   with $P\,p = (p^0, -\vec{p}\,)$.

4. Time reversal invariance (in case CP is a symmetry)

$$\mid T\,(p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots)\mid^2 = \mid T\,(T\,p_1', j_1', \lambda_1', \ldots \mid T\,p_1, j_1, \lambda_1, \ldots)\mid^2$$

   with $T\,p = (p^0, -\vec{p}\,)$.

Unitarity and the T–matrix

Let us denote the in-state by $|i>$ (initial state) and the out-state by $|f>$ (final state) and correspondingly by $S_{fi} = < f|S|i >$ the S–matrix elements and by $T_{fi} = < f|T|i >$ the T–matrix elements. Consider the unitarity relation of S

$$< f|S^+ S|i > = < f|I|i > = \delta_{fi}$$

and insert a complete set of intermediate states. We obtain

$$\sum_n < f|S^+|n >< n|S|i > = \sum_n < n|S|f >^* < n|S|i > = \delta_{fi}\ .$$

Inserting the definition of the T–matrix this yields

$$\begin{aligned} 0 = \textstyle\sum_n \{ \quad & i\,(2\pi)^4\,\delta^{(4)}(P_n - P_i) \quad < n|I|f >^* < n|T|i > \\ & -i\,(2\pi)^4\,\delta^{(4)}(P_f - P_n) \quad < n|T|f >^* < n|I|i > \\ & +(2\pi)^8\,\delta^{(4)}(P_f - P_n)\,\delta^{(4)}(P_n - P_i) \quad < n|T|f >^* < n|T|i > \} \end{aligned}$$

or

$$\Big\| \; i\left\{T^*_{if} - T_{fi}\right\} = \sum_n (2\pi)^4\,\delta^{(4)}(P_n - P_i)\,T^*_{nf}T_{ni} \tag{3.19}$$

where $P_f = P_i$. In more explicit form this reads

$$\begin{gathered} i\,\{T^*\,(p_1,j_1,\lambda_1,\ldots \mid p'_1,j'_1,\lambda'_1,\ldots) - T\,(p'_1,j'_1,\lambda'_1,\ldots \mid p_1,j_1,\lambda_1,\ldots)\} \\ = \sum_{\lambda''_1\ldots} \int d\mu(p''_1)\ldots(2\pi)^4\,\delta^{(4)}(\textstyle\sum_i p''_i - \sum_i p_i) \\ T^*\,(p''_1,j''_1,\lambda''_1,\ldots \mid p'_1,j'_1,\lambda'_1,\ldots)\cdot T\,(p''_1,j''_1,\lambda''_1,\ldots \mid p_1,j_1,\lambda_1,\ldots) \end{gathered}$$

and for elastic forward scattering where $p'_1 = p_1,\ j'_1 = j_1,\ \lambda'_1 = \lambda_1,\ \ldots$ we obtain the **optical theorem**

$$\begin{gathered} 2\,\mathrm{Im}\,T\,(p_1,j_1,\lambda_1,\ldots \mid p_1,j_1,\lambda_1,\ldots) \\ = \sum_{\lambda''_1\ldots} \int d\mu(p''_1)\ldots(2\pi)^4\,\delta^{(4)}(\textstyle\sum_i p''_i - \sum_i p_i)\mid T\,(p''_1,j''_1,\lambda''_1,\ldots \mid p_1,j_1,\lambda_1,\ldots)\mid^2 \end{gathered}$$

which tells us that the imaginary part of the forward scattering amplitude is related to the total probability for transitions $i \to n$ "summed" over all possible states n. Physical applications of the optical theorem we will encounter later on.

Let us now discuss, as a next step, how to calculate scattering matrix elements.

Calculating S–and T–matrix elements

In principle we are able to calculate S- and T–matrix elements once a theory is specified by a given Lagrangian. The S–matrix elements are obtained using

$$\begin{aligned} &< \mathrm{out}\ p'_1,j'_1,\lambda'_1,\ldots \mid p_1,j_1,\lambda_1,\ldots\,\mathrm{in} > \\ = \;&< \mathrm{in}\ p'_1,j'_1,\lambda'_1,\ldots \mid S \mid p_1,j_1,\lambda_1,\ldots\,\mathrm{in} > \end{aligned}$$

where the S–operator is represented in terms of the free in–fields (see Eq. (3.14))

$$S = T(e^{i\int d^4x\,\mathcal{L}_{int}\{\varphi_{in}\}(x)})_\otimes$$

by expanding the exponential into a formal power series. Since both the states and the fields (appearing in $\mathcal{L}_{int}$) are associated with the same free in–states, and hence may be represented in terms of free creation and annihilation operators, we are able to actually calculate the S–matrix elements

$$\begin{gathered} < \mathrm{out}\ p'_1,j'_1,\lambda'_1,\ldots \mid p_1,j_1,\lambda_1,\ldots\,\mathrm{in} >= \sum_{n=0}^{\infty} \frac{i^n}{n!}\int d^4x_1\cdots d^4x_n \\ {}_{in}\!< 0|a^+_{in}(p'_1,j'_1,\lambda'_1)\ldots T\{\mathcal{L}^{(in)}_{int}(x_1),\cdots,\mathcal{L}^{(in)}_{int}(x_n)\}\,a_{in}(p_1,j_1,\lambda_1)\ldots|0>_{in_\otimes} \end{gathered}$$

to any order n of the perturbation expansion. Besides the integrations the problem is reduced to the problem of calculating vacuum expectation values of products of free creation and annihilation operators which can be evaluated by just using the canonical (anti-) commutation relations together with the vacuum annihilation property of the annihilation operators.

Thus, once we know the particles and the interaction Lagrangian we can evaluate S matrix elements in a straight forward way. It is useful, however, to develop some simple rules, known as **Feynman rules**, which allow for an efficient calculation of perturbation series. The reader who is not interested in a detailed derivation of the Feynman rules may skip reading the next two chapters which are somewhat technical.

Efficient techniques are important in perturbation theory, since, even in simple applications, one easily encounters a large number of terms. Feynman rules will be given later when we will discuss specific models of interactions. Before we are doing so we discuss some useful tools in the following: The calculation of vacuum expectation values of free fields (Wick's theorem), which is the configuration space version of the vacuum expectation values of products of free creation and annihilation operators, and the relationship between vacuum expectation values of fields (Green functions) and S–matrix elements (LSZ–formulas), which we consider first in the next subsection.

3.3 The LSZ Reduction Formulas

The purpose of this chapter is to show that each external particle which takes part in a scattering process contributes a one–particle wave function as a factor to the scattering amplitude. This is summarized in Tab. 3.1.

The dynamics of an interacting system is characterized by particle fields and their field equations which derive from a Lagrangian by the principle of stationary action. How can we calculate S matrix elements in terms of these fields? A conceptually clear answer to this question was given by Lehmann, Symanzik and Zimmermann (LSZ) in 1957. We will discuss it in following. We start with a short digression on wave packets:

As usual we will work with plane waves in the following discussion. Several points could be made more precise if we would work with wave packets instead of plane waves. All arguments of our derivation remain valid if we replace $\exp(-ipx)$ and $\exp(ipx)$ by normalized positive and negative frequency solutions, respectively, of the Klein-Gordon equation. Let us consider a complete set, labeled by α, of positive frequency solutions

$$\left(\Box + m^2\right) f_\alpha(x) = 0$$

orthonormalized by

$$i \int d^3x f^*_{\alpha'}(x) \stackrel{\leftrightarrow}{\partial}_0 f_\alpha(x) = \delta_{\alpha'\alpha} \ .$$

As a positive frequency solution of the Klein-Gordon equation we may write $f_\alpha(x)$ as a Fourier transform of a function with support on the positive mass hyperboloid

$$f_\alpha(x) = \int \frac{d^4p}{(2\pi)^4} 2\pi\Theta(p^0)\delta(p^2 - m^2)\tilde{f}_\alpha(p)\, e^{-ipx}$$

whence

$$\int d\mu(p)\tilde{f}^*_{\alpha'}(p)\tilde{f}_\alpha(p) = \delta_{\alpha'\alpha} \ .$$

Smeared states are defined by

$$\mid f_\alpha > = \int d\mu(p)\tilde{f}_\alpha(p) \mid \vec{p} >$$

and one may introduce creation and annihilation operators by

$$\mid f_\alpha > = a^+_\alpha \mid 0 > \quad , \qquad a_\alpha = \left(a^+_\alpha\right)^+$$

which satisfy

$$[a_{\alpha'}, a^+_\alpha] = \delta_{\alpha'\alpha}$$

and the field is represented by

$$\varphi(x) = \sum_\alpha \left\{a_\alpha f_\alpha(x) + a^+_\alpha f^*_\alpha(x)\right\}$$

and

$$a_\alpha = i \int d^3x\, f^*_\alpha(x) \overleftrightarrow{\partial}_0\, \varphi(x) \quad .$$

The wave packet solutions we may think of as being smooth functions extending far enough in space to approximate very closely to plane waves but are fastly decreasing at spatial infinity such that surface terms obviously vanish. In momentum space $\tilde{f}_\alpha(p)$ may be taken as a smooth function concentrated around p with some finite spread.

Before we are going to consider scattering matrix elements we present a few concepts which will be needed:

3.3.1 Creation and annihilation operators in terms of fields

We are familiar with the representation of the free fields in terms of annihilation and creation operators which destroy or create free particle states as they appear in scattering states. What we need is the inverse relationship, the creation and annihilation operators in terms of the fields. These can be found easily by inverting the known formulas. In the following free field formulas one has to identify $p^0 = \omega_p = \sqrt{\vec{p}^{\,2} + m^2}$ where m is the particle mass.

- **Charged scalar field**

$$\varphi(x) = \int d\mu(p) \left\{ a(\vec{p}\,)\, e^{-ipx} + b^+(\vec{p}\,)\, e^{ipx} \right\}$$

$$a(\vec{p}\,) = \; i \int d^3x\, e^{ipx} \overleftrightarrow{\partial}_0\, \varphi(x) \quad , \quad b^+(\vec{p}\,) = -i \int d^3x\, e^{-ipx} \overleftrightarrow{\partial}_0\, \varphi(x)$$

- **Dirac field**

$$\psi_\alpha(x) = \sum_{r=\pm 1/2} \int d\mu(p) \left\{ u_\alpha(\vec{p},r)\, a(\vec{p},r)\, e^{-ipx} + v_\alpha(\vec{p},r)\, b^+(\vec{p},r)\, e^{ipx} \right\}$$

$$a(\vec{p},r) = \int d^3x\, e^{ipx} \bar{u}(\vec{p},r) \gamma^0 \psi(x) \quad , \quad b^+(\vec{p},r) = \int d^3x\, e^{-ipx} \bar{v}(\vec{p},r) \gamma^0 \psi(x)$$

- **Charged vector field**

$$V^\mu(x) \;=\; \sum_{r=\pm 1,0} \int d\mu(p) \left\{ \epsilon^\mu(\vec{p},r)\, a(\vec{p},r)\, e^{-ipx} + \epsilon^{\mu *}(\vec{p},r)\, b^+(\vec{p},r)\, e^{ipx} \right\}$$

$$a(\vec{p},r) = -i \int d^3x\, e^{ipx} \overleftrightarrow{\partial}_0\, V_\mu(x)\, \epsilon^{\mu *}(\vec{p},r) \quad , \quad b^+(\vec{p},r) = \; i \int d^3x\, e^{-ipx} \overleftrightarrow{\partial}_0\, V_\mu(x)\, \epsilon^\mu(\vec{p},r)$$

For neutral particles one has to identify $b = a$, and neutral Bose fields are real.

3.3.2 The asymptotic condition

For free fields, by construction, the spatial integrals which represent the creation and annihilation operators are **time independent**. We now consider these operators to be in- or out-operators which create incoming or outgoing free scattering states, respectively. We then can use the asymptotic condition

$$\phi_{in/out}(x^0, \vec{x}\,) = \lim_{x^0 \to -\infty/+\infty} \phi(x^0, \vec{x}\,) \ ,$$

where $\phi = \varphi,\ \psi,\ V_\mu$, to express these operators as limits in terms of the interpolating (interacting) fields as follows:

$$\begin{aligned} a_{in}(\vec{p}\,) &= i \int d^3x \, e^{ipx} \overset{\leftrightarrow}{\partial}_0 \, \varphi_{in}(x) \\ &= i \lim_{x^0 \to -\infty} \int d^3x \, e^{ipx} \overset{\leftrightarrow}{\partial}_0 \, \varphi(x) \end{aligned}$$

for the in-operators,

$$\begin{aligned} a_{out}(\vec{p}\,) &= i \int d^3x \, e^{ipx} \overset{\leftrightarrow}{\partial}_0 \, \varphi_{out}(x) \\ &= i \lim_{x^0 \to +\infty} \int d^3x \, e^{ipx} \overset{\leftrightarrow}{\partial}_0 \, \varphi(x) \end{aligned}$$

for the out-operators, and analogous expressions for the other operators and fields.

3.3.3 Partial integration

The in- and out- operators are interrelated by a partial integration. Using

$$\int d^3x \int_{-\infty}^{+\infty} dx^0 \frac{\partial}{\partial x^0} F(x^0, \vec{x}\,) = \lim_{x^0 \to +\infty} \int d^3x F(x^0, \vec{x}\,) - \lim_{x^0 \to -\infty} \int d^3x F(x^0, \vec{x}\,)$$

and the representations of the in- and out- operators in terms of the interpolating fields it is easy to find the explicit form of these interrelations as follows:

- **Scalar fields**

For the scalar field we have

$$F(x) = e^{ipx} \overset{\leftrightarrow}{\partial}_0 \, \varphi(x)$$

and

$$\partial_0 F(x^0, \vec{x}\,) = e^{ipx} \partial_0^2 \varphi(x) - \partial_0^2 \left(e^{ipx} \right) \varphi(x) \ .$$

Since e^{ipx} satisfies the Klein-Gordon equation we have

$$\partial_0^2 e^{ipx} = (\vec{\partial}^{\,2} - m^2) \, e^{ipx} \ .$$

By two partial integrations, in which the surface terms can be dropped, we can bring $\vec{\partial}^{\,2}$ to operate on $\varphi(x)$ and obtain

$$\int d^4x \, \partial_0 F(x) = \int d^4x \, e^{ipx} \left(\Box + m^2 \right) \varphi(x) \ .$$

The relations we are looking for therefore read

$$\begin{aligned} a_{out}(\vec{p}\,) &= a_{in}(\vec{p}\,) + i\int d^4x\, e^{ipx}\left(\Box + m^2\right)\varphi(x) \\ b^+_{in}(\vec{p}\,) &= b^+_{out}(\vec{p}\,) + i\int d^4x\, e^{-ipx}\left(\Box + m^2\right)\varphi(x) \quad . \end{aligned}$$

This is a very interesting relationship since it directly relates the evolution of in- into out- states to the interaction of the fields, given by the source term of the field equations. For example, we have

$$\left(\Box + m^2\right)\varphi(x) = \begin{cases} 0 & \text{free field} \\ -\lambda J\{\varphi\}(x) & \lambda\varphi^4 \text{ -theory} \quad . \end{cases}$$

In fact the above relation is not new, it is just the momentum space version of the Yang-Feldman equation in the form

$$\varphi_{out}(x) = \varphi_{in}(x) + \lambda\int d^4y\,\Delta(x-y)\,J\{\varphi\}(y)$$

which we discussed in Sec. 3.1 for the $\lambda\varphi^4$-theory.

- **Dirac fields**

For the Dirac field we have

$$F(x) = e^{ipx}\bar{u}(\vec{p},r)\gamma^0\psi(x)$$

and

$$\partial_0 F(x^0,\vec{x}\,) = e^{ipx}\bar{u}(\vec{p},r)\gamma^0\partial_0\psi(x) + \partial_0\left(e^{ipx}\right)\bar{u}(\vec{p},r)\gamma^0\psi(x)$$

Taking the derivative of the exponential and using the Dirac equation $\not{p}u = mu$ for the spinor

$$\partial_\mu\left(e^{ipx}\right)\bar{u}(\vec{p},r)\gamma^\mu = ie^{ipx}\bar{u}(\vec{p},r)\not{p} = ie^{ipx}\bar{u}(\vec{p},r)m$$

we obtain

$$\begin{aligned} \partial_0\left(e^{ipx}\right)\bar{u}(\vec{p},r)\gamma^0\psi(x) &= -\vec{\partial}\left(e^{ipx}\right)\bar{u}(\vec{p},r)\vec{\gamma}\psi(x) + ie^{ipx}\bar{u}(\vec{p},r)m\psi(x) \\ &\overset{\text{p.i.}}{=} -ie^{ipx}\bar{u}(\vec{p},r)\left(i\vec{\gamma}\vec{\partial} - m\right)\psi(x) \quad . \end{aligned}$$

In the last line we have performed a partial integration (dropping surface terms as usual) already as the relation needs holds under the integral only. We then arrive at

$$\int d^4x\,\partial_0 F(x) = -i\int d^4x\, e^{ipx}\bar{u}(\vec{p},r)\left(i\gamma^\mu\partial_\mu - m\right)\psi(x)$$

and thus

$$\begin{aligned} a_{out}(\vec{p},r) &= a_{in}(\vec{p},r) - i\int d^4x\, e^{ipx}\bar{u}(\vec{p},r)\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) \\ b^+_{in}(\vec{p},r) &= b^+_{out}(\vec{p},r) + i\int d^4x\, e^{-ipx}\bar{v}(\vec{p},r)\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) \quad . \end{aligned}$$

Again we find that particle creation and annihilation is in direct correspondence to the **source of the field**, the interaction term of the field equation , for example,

$$\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = \begin{cases} 0 & \text{free field} \\ -eA_\mu\gamma^\mu\psi(x) & \text{in QED} \quad . \end{cases}$$

- **Vector fields**

For the vector fields we may directly use the scalar formulas if we identify $\varphi(x) = -V_\mu(x)\epsilon^{\mu(*)}(\vec{p}, r)$. Thus

$$\begin{aligned} a_{out}(\vec{p}, r) &= a_{in}(\vec{p}, r) - i \int d^4x\, e^{ipx} \left(\Box + m^2\right) V_\mu(x)\, \epsilon^{\mu *}(\vec{p}, r) \\ b^+_{in}(\vec{p}, r) &= b^+_{out}(\vec{p}, r) - i \int d^4x\, e^{-ipx} \left(\Box + m^2\right) V_\mu(x)\, \epsilon^{\mu}(\vec{p}, r) \end{aligned}$$

with source terms, for example, ($V_\mu = A_\mu$ real and $m_A = 0$)

$$\Box A_\mu(x) = \begin{cases} (1 - \xi^{-1})\, \partial_\mu \partial_\nu A^\nu(x) & \text{free photon field} \\ -e j_{\mu\, em}(x) + (1 - \xi^{-1})\, \partial_\mu \partial_\nu A^\nu(x) & \text{in QED} \end{cases}$$

A crucial point here is that the gauge dependent term drops out in the "reduction formula". In fact, a partial integration of ∂_μ for the term proportional to $(1 - \xi^{-1})$ yields

$$\begin{aligned} \int d^4x\, e^{\pm ipx} \partial_\mu \partial_\nu A^\nu(x)\, \epsilon^\mu(\vec{p}, r) &= -\int d^4x\, \partial_\mu \left(e^{\pm ipx}\right) \partial_\nu A^\nu(x)\, \epsilon^\mu(\vec{p}, r) \\ &= \mp i \int d^4x\, e^{\pm ipx} \partial_\nu A^\nu(x)\, p_\mu \epsilon^\mu(\vec{p}, r) \\ &= 0\ , \end{aligned}$$

since the polarization vectors satisfy $p_\mu \epsilon^\mu(\vec{p}, r) = 0$. Thus for physical states, as described by the creation and annihilation operators, gauge dependent terms drop out i.e. they couple to **physical sources**. For QED this directly proves the physical form of Maxwell's equation to be valid when acting on the physical state space.

Taking advantage of the fact that the derivative term does not contribute one customarily writes the reduction formula for the photon like

$$\begin{aligned} a_{out}(\vec{p}, r) &= a_{in}(\vec{p}, r) - i \int d^4x\, e^{ipx} \left(\Box g_{\mu\nu} - (1 - \xi^{-1})\, \partial_\mu \partial_\nu\right) A^\nu(x)\, \epsilon^{\mu *}(\vec{p}, r) \\ b^+_{in}(\vec{p}, r) &= b^+_{out}(\vec{p}, r) - i \int d^4x\, e^{-ipx} \left(\Box g_{\mu\nu} - (1 - \xi^{-1})\, \partial_\mu \partial_\nu\right) A^\nu(x)\, \epsilon^{\mu}(\vec{p}, r) \end{aligned}$$

such that with

$$\left(\Box g_{\mu\nu} - (1 - \xi^{-1})\, \partial_\mu \partial_\nu\right) A^\nu(x) = \begin{cases} 0 & \text{free field} \\ -e j_{\mu\, em}(x) & \text{in QED} \end{cases}$$

the relationship between the physical in and out operators takes a more physical appearance. As a rule one always may take the inverse free propagator, as given by the kernel of the bilinear part of the Lagrangian, to act on the interpolating field.

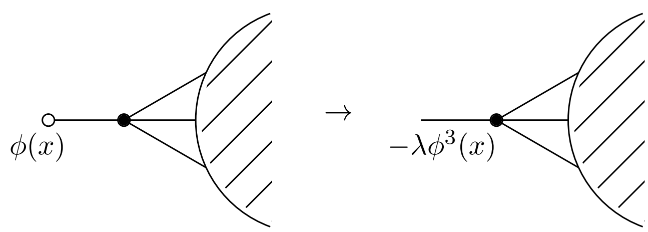


Fig. 3.1: Relationship between field $\phi(x)$ and state (source ($\bullet$): $(\Box + m^2)\phi(x) = -\lambda\phi^3(x)$)

3.3.4 Reduction of states

In order to keep notation simple let us consider scattering with an in-state $|i>$ and an out-state $|f>$ each containing a neutral spin zero particle of mass m, besides some other particles labeled by α and β, respectively:

$$\begin{aligned} |i> &= |p,\alpha \text{ in}> &= a^+_{in}(p)\,|\alpha \text{ in}> \\ \\ |f> &= |q,\beta \text{ out}> &= a^+_{out}(q)\,|\beta \text{ out}> \quad . \end{aligned}$$

The S–matrix element is given by

$$\begin{aligned} & S(q,\beta|p,\alpha) =< \text{out } \beta, q \mid p,\alpha \text{ in} >=< \text{out } \beta, q \mid a^+_{in}(p) \mid \alpha \text{ in}> \\ = \;& < \text{out } \beta, q \mid a^+_{out}(p) \mid \alpha \text{ in}> + i\int d^4x\, e^{-ipx}\left(\Box_x + m^2\right) < \text{out } \beta, q \mid \varphi(x) \mid \alpha \text{ in}> \end{aligned}$$

We say that we have "reduced" or "contracted" a particle from the in-state. The first term is a matrix element of the unit operator, unity for forward scattering and zero otherwise: For example, if $|f>$ contains only one spin zero boson of mass m:

$$< \text{out } \beta, q \mid a^+_{out}(p) \mid \alpha \text{ in}>=< \text{out } \beta \mid a_{out}(q)\, a^+_{out}(p) \mid \alpha \text{ in}>= 2\omega_p\,(2\pi)^3\delta^{(3)}(\vec{q}-\vec{p}\,)\; < \text{out } \beta \mid \alpha \text{ in}>$$

It thus corresponds to a freely moving, unscattered particle and therefore is uninteresting as it does not contribute to the T–matrix proportional to $S-1$.

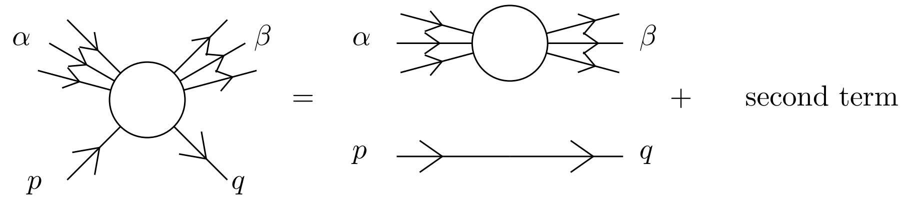


Fig. 3.2: Disconnected contribution from an unscattered particle

The second term is the contribution to the T–matrix element. If we denote the source of the field $\varphi(x)$ by $j(x)$

$$j(x) \;=\left(\Box_x + m^2\right)\varphi(x)$$

and use translation invariance to write

$$< \text{out } \beta, q \mid j(x) \mid \alpha \text{ in}>= e^{-i(p_\alpha - p_\beta - q)x} < \text{out } \beta, q \mid j(0) \mid \alpha \text{ in}>$$

we find

$$\begin{aligned} & i\int d^4x\, e^{-ipx}\left(\Box_x + m^2\right) < \text{out } \beta, q \mid \varphi(x) \mid \alpha \text{ in}> \\ =\; & i\int d^4x\, e^{-ipx} e^{-i(p_\alpha - p_\beta - q)x} < \text{out } \beta, q \mid j(0) \mid \alpha \text{ in}> \\ =\; & i(2\pi)^4\delta^{(4)}(p_\beta + q - p_\alpha - p)\; < \text{out } \beta, q \mid j(0) \mid \alpha \text{ in}> \end{aligned}$$

Thus the result of the reduction is

$$S(q,\beta|p,\alpha) = I(q,\beta|p,\alpha) + i(2\pi)^4\delta^{(4)}(p_f - p_i)\; < \text{out } \beta, q \mid j(0) \mid \alpha \text{ in}>$$

which tells us that the T–matrix element is the matrix element of the source of the interpolating field of the particle we have reduced:

$$T(q,\beta|p,\alpha) =< \text{out }\beta, q \mid j(0) \mid \alpha\text{ in} > \quad .$$

We only deal with this second term of the reduction in the following.

We proceed by reducing an outgoing particle from the remaining matrix element:

$$\begin{aligned} < \text{out }\beta, q \mid \varphi(x) \mid \alpha\text{ in} > \;&=\; < \text{out }\beta \mid a_{out}(q)\, \varphi(x) \mid \alpha\text{ in} > \\ = i \lim_{y^0\to+\infty} \textstyle\int d^3y\, e^{iqy} \overset{\leftrightarrow}{\partial}_{y0} &< \text{out }\beta \mid \varphi(y)\, \varphi(x) \mid \alpha\text{ in} > \end{aligned}$$

At this stage we have to be ingenious and notice that the product $\varphi(y)\,\varphi(x)$ can be rewritten as a "product" $\tilde{T}\{\varphi(y)\,\varphi(x)\}$ which satisfies the boundary condition

$$\lim_{y^0\to+\infty} \tilde{T}\{\varphi(y)\,\varphi(x)\} = \lim_{y^0\to+\infty} \varphi(y)\,\varphi(x) = \varphi_{out}(y)\,\varphi(x)$$

We again perform a partial integration

$$\begin{aligned} & i \lim_{y^0\to+\infty} \int d^3y\, e^{iqy} \overset{\leftrightarrow}{\partial}_{y0} < \text{out }\beta \mid \tilde{T}\{\varphi(y)\,\varphi(x)\} \mid \alpha\text{ in} > \\ = \;& i \lim_{y^0\to-\infty} \int d^3y\, e^{iqy} \overset{\leftrightarrow}{\partial}_{y0} < \text{out }\beta \mid \tilde{T}\{\varphi(y)\,\varphi(x)\} \mid \alpha\text{ in} > \\ + \;& i \int d^4y\, \partial_{y0} \left(e^{iqy} \overset{\leftrightarrow}{\partial}_{y0} < \text{out }\beta \mid \tilde{T}\{\varphi(y)\,\varphi(x)\} \mid \alpha\text{ in} > \right) \\ = \;& i \int d^3y\, e^{iqy} \overset{\leftrightarrow}{\partial}_{y0} < \text{out }\beta \mid \tilde{T}\{\varphi_{in}(y)\,\varphi(x)\} \mid \alpha\text{ in} > \\ + \;& i \int d^4y\, e^{iqy} \left(\Box_y + m^2\right) < \text{out }\beta \mid \tilde{T}\{\varphi(y)\,\varphi(x)\} \mid \alpha\text{ in} > \end{aligned}$$

The prescription $\tilde{T}$ has to have the property that the first term can be evaluated. This requires $\varphi_{in}(y)$ to appear not to the left of $\varphi(x)$. And it must be Lorentz invariant. Since the operation of $\varphi_{in}(y)$ on $\mid \alpha$ in $>$ is known it is natural to require as a second boundary condition

$$\lim_{y^0\to-\infty} \tilde{T}\{\varphi(y)\,\varphi(x)\} = \lim_{y^0\to-\infty} \varphi(x)\,\varphi(y) = \varphi(x)\,\varphi_{in}(y)$$

The only L-invariant solution satisfying both boundary conditions is the time ordered product

$$\begin{aligned} \tilde{T}\{\varphi(y)\,\varphi(x)\} = T\{\varphi(y)\,\varphi(x)\} &= \Theta(y^0-x^0)\,\varphi(y)\,\varphi(x) + \Theta(x^0-y^0)\,\varphi(x)\,\varphi(y) \\ &= \varphi(y)\,\varphi(x) - \Theta(x^0-y^0)\;[\varphi(y),\,\varphi(x)] \end{aligned}$$

Notice that the T–product differs from the ordinary product by the retarded commutator. Since locality requires $[\varphi(y),\,\varphi(x)] = 0$ for $(y-x)^2 < 0$ multiplication with the Θ-function $\Theta(x^0-y^0)$ is not in conflict with L-invariance since it just cuts the forward cone. The only problem can show up for $y = x$ i.e. at the tip of the light cone and is dealt with in renormalization theory.

As a result of the reduction of a second particle from the out-state we find

$$\begin{aligned} < \text{out }\beta, q \mid \varphi(x) \mid \alpha\text{ in} > \;&=\; < \text{out }\beta \mid a_{out}(q)\,\varphi(x) \mid \alpha\text{ in} > \\ =< \text{out }\beta \mid \varphi(x)\, a_{in}(q) \mid \alpha\text{ in} > &+ i \textstyle\int d^4y\, e^{iqy}\,(\Box_y + m^2) < \text{out }\beta \mid T\{\varphi(y)\,\varphi(x)\} \mid \alpha\text{ in} > \end{aligned}$$

Again the first term is uninteresting. It vanishes if $|\alpha$ in $>$ does not contain a spin zero boson of mass m and momentum $p' = q$, otherwise the term corresponds to matrix element of the unit operator. The T–matrix element is given by

$$(S-I)\,(q,\beta|p,\alpha) = i^2 \int d^4y\, d^4x\, e^{iqy} e^{-ipx} \left(\Box_y + m^2\right)\left(\Box_x + m^2\right) < \text{out }\beta \mid T\{\varphi(y)\,\varphi(x)\} \mid \alpha\text{ in} >$$

In this way we can reduce all particles from the states. The rule for obtaining the T–matrix element is very simple:

- Write the states defining the S–matrix element in terms of a product of the appropriate free in- and out-creation operators acting onto the vacuum state.
- Replace the creation operators from the in-state and the annihilation operators from the out-state by their interpolating fields (source terms).
- As a result the T–matrix element is represented in terms of the connected part of the vacuum expectation value of the time ordered product of the interpolating Heisenberg fields.

For scattering of neutral spin zero particles of mass m, with n particles incoming and m particles outgoing, this reads

$$\begin{gathered}(S-I)\,(q_1,\cdots,q_m|p_1,\cdots,p_n) = i^{n+m}\int d^4y_1\,\cdots d^4y_m\,d^4x_1\,\cdots d^4x_n \\ e^{iq_1y_1}\cdots e^{iq_my_m}e^{-ip_1x_1}\cdots e^{-ip_nx_n}\left(\Box_{y_1}+m^2\right)\cdots\left(\Box_{y_m}+m^2\right)\left(\Box_{x_1}+m^2\right)\cdots\left(\Box_{x_n}+m^2\right) \\ <0\mid T\left\{\varphi(y_1)\,\cdots\varphi(y_m)\,\varphi(x_1)\cdots\varphi(x_n)\right\}\mid 0>^{conn}\quad .\end{gathered}$$

As indicated we are interested only in the connected part given by diagrams which do not factorize into two or more independent parts corresponding to independent subprocesses. Note that in momentum space the Klein-Gordon operators are the inverse free propagators which cancel the external one-particle propagators of the time-ordered amplitude. The main result is that in momentum space the **S–matrix elements are the residues of the one-particle poles of the time-ordered Green functions**.

This generalizes to charged and higher spin particles in an obvious way. For non integer spin the T–product must be defined in accordance with the anti-commutativity of the Fermi fields:

$$\begin{aligned}T\left\{\psi(y)\,\psi(x)\right\} &= \Theta(y^0-x^0)\,\psi(y)\,\psi(x)-\Theta(x^0-y^0)\,\psi(x)\,\psi(y) \\ &= \psi(y)\,\psi(x)-\Theta(x^0-y^0)\,\left\{\psi(y),\,\psi(x)\right\}\end{aligned}$$

The vacuum expectation value of the time ordered products of fields are called time ordered Green functions or simply τ-functions:

$$\tau(x_1,x_2,\cdots)\equiv<0\mid T\left\{\phi(x_1)\,\phi(x_2)\,\cdots\right\}\mid 0>$$

Fields in a time ordered product commute (bosons) or anti-commute (fermions).

Usually it is more convenient to work directly in momentum space. We use the convention

$$\phi(x) = (2\pi)^{-4}\int d^4q\,e^{-iqx}\tilde{\phi}(q)\;\equiv\int_{q_i} e^{-iqx}\tilde{\phi}(q) \tag{3.20}$$

for the Fourier transformation of the fields.

We summarize the reduction rules as follows: p is the on–shell particle momentum $p^2=m^2$ and q is an off–shell momentum which at the end will be taken in the on–shell limit $q^2\to p^2=m^2$

Spin 0 bosons:		
Incoming particle	$i\int d^4x\, \varphi^+(x)\,\left(\overleftarrow{\Box}+m^2\right) e^{-ipx}$	$-i\,(q^2-m^2)\,\tilde{\varphi}^+(q)$
Incoming antiparticle	$i\int d^4x\, e^{-ipx}\left(\overrightarrow{\Box}+m^2\right)\varphi(x)$	$-i\,(q^2-m^2)\,\tilde{\varphi}(-q)$
Outgoing particle	$i\int d^4x\, e^{ipx}\left(\overrightarrow{\Box}+m^2\right)\varphi(x)$	$-i\,(q^2-m^2)\,\tilde{\varphi}(q)$
Outgoing antiparticle	$i\int d^4x\, \varphi^+(x)\,\left(\overleftarrow{\Box}+m^2\right) e^{ipx}$	$-i\,(q^2-m^2)\,\tilde{\varphi}^+(-q)$
Spin 1/2 fermions:		
Incoming particle	$i\int d^4x\, \bar{\psi}(x)\,\left(i\gamma^\mu\overleftarrow{\partial}_\mu+m\right)u(p,r)e^{-ipx}$	$-i\tilde{\bar{\psi}}(q)\,(\not{q}-m)\,u(p,r)$
Incoming antiparticle	$i\int d^4x\, e^{-ipx}\bar{v}(p,r)\left(i\gamma^\mu\overrightarrow{\partial}_\mu-m\right)\psi(x)$	$-i\bar{v}(p,r)\,(\not{q}+m)\,\tilde{\psi}(-q)$
Outgoing particle	$-i\int d^4x\, e^{ipx}\bar{u}(p,r)\left(i\gamma^\mu\overrightarrow{\partial}_\mu-m\right)\psi(x)$	$-i\bar{u}(p,r)\,(\not{q}-m)\,\tilde{\psi}(q)$
Outgoing antiparticle	$-i\int d^4x\, \bar{\psi}(x)\,\left(i\gamma^\mu\overleftarrow{\partial}_\mu+m\right)v(p,r)e^{ipx}$	$-i\tilde{\bar{\psi}}(-q)\,(\not{q}+m)\,v(p,r)$
Spin 1 bosons:		
Incoming particle	$-i\int d^4x\, V_\mu^+(x)\,\left(\overleftarrow{\Box}+m^2\right)\epsilon^\mu(p,r)e^{-ipx}$	$i\tilde{V}_\mu^+(q)\,(q^2-m^2)\,\epsilon^\mu(p,r)$
Incoming antiparticle	$-i\int d^4x\, e^{-ipx}\epsilon^\mu(p,r)\left(\overrightarrow{\Box}+m^2\right)V_\mu(x)$	$i\epsilon^\mu(p,r)\,(q^2-m^2)\,\tilde{V}_\mu(-q)$
Outgoing particle	$-i\int d^4x\, e^{ipx}\epsilon^{\mu*}(p,r)\left(\overrightarrow{\Box}+m^2\right)V_\mu(x)$	$i\epsilon^{\mu*}(p,r)\,(q^2-m^2)\,\tilde{V}_\mu(q)$
Outgoing antiparticle	$-i\int d^4x\, V_\mu^+(x)\,\left(\overleftarrow{\Box}+m^2\right)\epsilon^{\mu*}(p,r)e^{ipx}$	$i\tilde{V}_\mu^+(-q)\,(q^2-m^2)\,\epsilon^{\mu*}(p,r)$

The **external** field $\tilde{\phi}(q)$ appearing in the vacuum expectation value of the time ordered product only gives a contribution if it contracts with a field of the same kind from a vertex. This yields an external propagator for that field times an **amputated amplitude** with the external propagator omitted. Graphically:

$$\lim_{q^2\to m^2} -i\,(q^2-m^2)$$

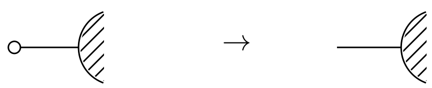
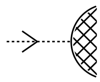
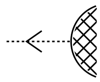
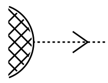
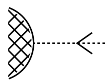
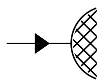
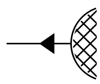
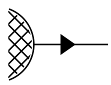
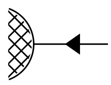
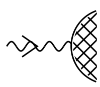
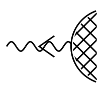
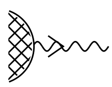
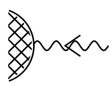


Fig. 3.3:: Amputation for a scalar field

The external propagators of the time ordered Green functions multiply to unity with the inverse propagators obtained by the reduction of states. Hence the T–matrix elements are given by the **on–shell limits** of the amputated τ-functions multiplied with the classical wave functions listed in Tab. 3.1 describing the state. For one particle states which can emit massless quanta (bremsstrahlung like process) the on–shell limit does not exist, which tells us that the notion of a free one particle state in this case is not physical because the state is always **dressed** by an arbitrary number of soft massless quanta. This is a well known problem, the **infrared problem**, in QED, where an electron is always dressed by soft photons. The problem can be solved by a physically more realistic definition of these states. Here we assume all particles to be massive in which case, after cancelation of the propagator poles, one can prove that the on–shell limits exist to all orders in perturbation theory (Steinmann 1971).

Table 3.1: Rules for the treatment of external legs in the evaluation of T–matrix elements.

scattering state	graphical representation	wave function
Scalar particles:		
incoming particle		1
incoming antiparticle		1
outgoing particle		1
outgoing antiparticle		1
Dirac particles:		
incoming particle		$u(p,r)$
incoming antiparticle		$\bar{v}(p,r)$
outgoing particle		$\bar{u}(p,r)$
outgoing antiparticle		$v(p,r)$
Vector particles:		
incoming particle		$\epsilon^{\mu}(p,r)$
incoming antiparticle		$\epsilon^{\mu}(p,r)$
outgoing particle		$\epsilon^{\mu *}(p,r)$
outgoing antiparticle		$\epsilon^{\mu *}(p,r)$

3.4 Perturbation theory.

3.4.1 The Gell-Mann–Low formula.

In the last chapter the calculation of S–matrix elements was related to the calculation of the time–ordered Green functions. For interacting fields the time–ordered Green functions are given by

$$\begin{aligned} & < 0|T\{\varphi(x_1)\dots\varphi(x_N)\}|0> \\ = \; & {}_{in}< 0|T\{\varphi_{in}(x_1)\dots\varphi_{in}(x_N)\, S\}|0>_{in\otimes} \end{aligned}$$

which is the **Gell-Mann–Low formula** (Gell-Mann and Low 1954).

A formal proof of the Gell-Mann–Low formula may proceeds as follows: We first remind the reader of the definition of time ordering for a functional. For an exponential functional exp $\{i\int_{-\infty}^{\infty} d^4x'\, F(x')\}$ by definition

$$T\left(e^{i\int_{-\infty}^{\infty} d^4x'\, F(x')}\right) = \sum_{n=0}^{\infty}\frac{i^n}{n!}\int d^4x_1\cdots d^4x_n\, T\left(F(x_1),\cdots,F(x_n)\right)$$

such that for any function $f(x)$ we have

$$\begin{aligned} & T\{f(x)\, T(e^{i\int_{-\infty}^{\infty} d^4x'\, F(x')})\} = T\{f(x)\, e^{i\int_{-\infty}^{\infty} d^4x'\, F(x')}\} \\ = \; & T(e^{i\int_{x^0}^{\infty} d^4x'\, F(x')}) f(x^0,\vec{x}\,)\, T(e^{i\int_{-\infty}^{x^0} d^4x'\, F(x')}) \quad . \end{aligned}$$

Now we use the representation of the interpolating fields in terms of free in–fields which we derived in Sec. 3.1

$$\varphi(x) = S^{-1}T\{\varphi_{in}(x)\, S\} = T(e^{i\int_{-\infty}^{x^0} d^4x\, \mathcal{L}_{int}^{(in)}})^{-1}\varphi_{in}(x) \quad T(e^{i\int_{-\infty}^{x^0} d^4x'\, \mathcal{L}_{int}^{(in)}(x')})$$

where we have used

$$S = T(e^{i\int_{-\infty}^{+\infty} d^4x'\, \mathcal{L}_{int}^{(in)}(x')}) = T(e^{i\int_{x^0}^{+\infty} d^4x'\, \mathcal{L}_{int}^{(in)}(x')})T(e^{i\int_{-\infty}^{x^0} d^4x'\, \mathcal{L}_{int}^{(in)}(x')})$$

and assume S to be normalized properly (see Eq. (3.15)) such that

$$|0>=|0>_{in}=|0>_{out}= S^{+}|0>_{in} \quad .$$

For the time ordered Green functions we then obtain

$$\begin{aligned} & < 0|T\{\varphi(x_1)\dots\varphi(x_N)\}|0>= \textstyle\sum_i \Theta_i < 0|\varphi(x_{i_1})\dots\varphi(x_{i_N})|0> \\ = \; & \textstyle\sum_i \Theta_i\, {}_{in}< 0|S^{-1}T\{\varphi_{in}(x_{i_1})S\}\dots T\{\varphi_{in}(x_{i_N})S\}|0>_{in\otimes} \\ = \; & \textstyle\sum_i \Theta_i\, {}_{in}< 0|T\{\varphi(x_{i_1})S\}\dots T\{\varphi(x_{i_N})S\}|0>_{in\otimes} \\ = \; & \textstyle\sum_i \Theta_i\, {}_{in}< 0|U(\infty,x_{i_1}^0)\varphi_{in}(x_{i_1})U(x_{i_1}^0,x_{i_2}^0)\varphi_{in}(x_{i_2})U(x_{i_2}^0,x_{i_3}^0)\varphi_{in}(x_{i_3})\dots\varphi_{in}(x_{i_N})U(x_{i_N}^0,-\infty)|0>_{in\otimes} \\ = \; & {}_{in}< 0|T\{\varphi_{in}(x_1)\dots\varphi_{in}(x_N)\, S\}|0>_{in\otimes} \end{aligned}$$

where

$$U(a,b) \equiv T\left(e^{i\int_{x'^0=b}^{x'^0=a} d^4x'\, \mathcal{L}_{int}^{(in)}(x')}\right) \quad .$$

The sum extends over all permutations $i:\ (1,2,\dots,N)\to(i_1,i_2,\dots,i_N)$ and Θ_i is a shorthand for the time–ordering product of Θ–functions for a given permutation i .

Thus, in perturbation theory we have

$$\begin{array}{c} < 0|T\{\varphi(x_1)\ldots\varphi(x_N)\}|0>= \sum_{n=0}^{\infty}\frac{i^n}{n!}\int d^4y_1\cdots d^4y_n \\ {}_{in}< 0|T\{\varphi_{in}(x_1)\ldots\varphi_{in}(x_N)\,\mathcal{L}_{int}^{(in)}(y_1),\cdots,\mathcal{L}_{int}^{(in)}(y_n)\}|0>_{in\otimes} \end{array} \tag{3.21}$$

which compares to the S–matrix element Eq. (3.20). The two expressions are related by the LSZ reduction formulas.

Eq. (3.21) is the analog of the expression for the S–matrix element with the replacement of the creation and annihilation operators by fields subjected to the time ordering prescription. One thus has to calculate vacuum expectation values of free fields. This we consider in the following paragraph.

3.4.2 Normal-ordering, Wick's theorem.

For notational convenience, in this paragraph, we use the convention that all states and fields describe free particles. In the perturbation expansion we usually have to evaluate products of monomials of free fields. We consider free fields $\varphi = \phi, \psi, V, \ldots$ which have vanishing vacuum expectation values

$$< 0|\varphi(x)|0 >= 0\ . \tag{3.22}$$

The vacuum expectation values of products of fields

$$< 0|\prod_i \varphi_i(x_i)|0 >$$

are called **Wightman functions**, those of time-ordered products

$$< 0|T\{\prod_i \varphi_i(x_i)\ \}|0 >$$

time-ordered Green functions or **τ–functions**. In order to evaluate vacuum-expectation values of field products we have to represent the fields in terms of the creation and annihilation operators and, using the canonical (anti-) commutation relations, to (anti-) commute all annihilation operators to the right of the creation operators until we can use the vacuum annihilation property

$$a(\vec{p}\,)|0 >= 0 \quad \text{or} \quad < 0|a^+(\vec{p}\,) = 0$$

(Dyson 1949). This procedure gives raise to the definition of normal–ordered products of fields. In configuration space we may split a field

$$\varphi(x) = \varphi^{(+)}(x) + \varphi^{(-)}(x)$$

into a **positive frequency part** $\varphi^{(+)}(x)$, which contains the annihilation operator, and a **negative frequency part** $\varphi^{(-)}(x)$, which contains the creation operator. In products the annihilation parts $\varphi^{(+)}(x)$ and the creation parts $\varphi^{(-)}(x)$ each (anti-) commute among themselves as a immediate consequence of the canonical (anti-) commutation relations.

Definition: A **normal–product** $:\varphi_1(x_1)\cdots\varphi_n(x_n):$ is defined by the prescription: Write all fields as a sum of the annihilation part and the creation part and write all annihilation parts to the right of the creation parts. For fermion-fields a factor (-1) has to be taken into account for each interchange of two fields. The field arguments may partially or all coincide. By definition all boson-fields are commuting and all fermion fields are anti-commuting under the $: \cdots :$-prescription and the normal products have vanishing vacuum expectation values

$$< 0| : \prod_i \varphi_i(x_i) : |0>= 0 \; . \tag{3.23}$$

The formal definition of a normal–product is

$$: \varphi_1(x_1) \cdots \varphi_n(x_n) : = \sum_{m=0}^{n} \sum_i (-1)^{F_i} \varphi_{i_1}^{(-)}(x_{i_1}) \cdots \varphi_{i_m}^{(-)}(x_{i_m}) \varphi_{i_{m+1}}^{(+)}(x_{i_{m+1}}) \cdots \varphi_{i_n}^{(+)}(x_{i_n})$$

where $\sum_i$ extends over all possibilities to chose an ordered subset of m elements $(i_1, \cdots, i_m)$ out of the ordered set $(1, \cdots, n)$. $(i_{m+1}, \cdots, i_n)$ is the ordered complement of $(i_1, \cdots, i_m)$ in $(1, \cdots, n)$. Furthermore, for Fermi fields, we have a sign factor $(-1)^{F_i}$ where $F_i = 0(1)$ if $(i_1, \cdots, i_m)$ exhibits an even(odd) permutation of the Fermi fields relative to the set $(1, \cdots, n)$. The number of terms is given by the binomial coefficients $\binom{n}{m}$ defined by

$$(x+y)^n = \sum_{m=0}^{n} \left(\begin{array}{c} n \\ m \end{array} \right) x^m y^{n-m} \; ; \quad \left(\begin{array}{c} n \\ m \end{array} \right) = \frac{n!}{m!(n-m)!} \; .$$

Correspondingly one finds for coinciding Bose fields

$$: \phi^n(x) \; := \sum_{m=0}^{n} \left(\begin{array}{c} n \\ m \end{array} \right) \left(\phi^{(-)}(x) \right)^m \left(\phi^{(+)}(x) \right)^{n-m} \; .$$

The normal products represent a basis of field-products and any ordinary product as well as time-ordered products of fields may be represented as a linear combination of normal products.

Consider a pair of fields. The ordinary product

$$\varphi(x)\varphi(y) = \varphi^{(-)}(x)\varphi^{(-)}(y) + \varphi^{(-)}(x)\varphi^{(+)}(y) + \varphi^{(+)}(x)\varphi^{(-)}(y) + \varphi^{(+)}(x)\varphi^{(+)}(y)$$

differs from the corresponding normal–product

$$: \; \varphi(x)\varphi(y) \; := \varphi^{(-)}(x)\varphi^{(-)}(y) + \varphi^{(-)}(x)\varphi^{(+)}(y) \pm \varphi^{(-)}(y)\varphi^{(+)}(x) + \varphi^{(+)}(x)\varphi^{(+)}(y)$$

by a free field (anti-) commutator

$$\varphi(x)\varphi(y) =: \; \varphi(x)\varphi(y) \; : + [\varphi^{(+)}(x), \varphi^{(-)}(y) \,]_{\mp} \tag{3.24}$$

and since $< 0| : \; \varphi(x)\varphi(y) \; : \; |0>= 0$ we have

$$[\varphi^{(+)}(x), \varphi^{(-)}(y) \,]_{\mp} =< 0|\varphi(x)\varphi(y) \,|0> \quad . \tag{3.25}$$

We say that $\varphi(x)\varphi(y)$ and $: \; \varphi(x)\varphi(y) \; :$ differ by a **contraction** which is represented by a two-point function $< 0|\varphi(x)\varphi(y) \,|0>$. For a scalar field we find

$$\begin{aligned}
[\phi^{(+)}(x), \phi^{(-)}(y) \,] &= < 0|\phi(x)\phi(y) \,|0> \\
&= \int d\mu(p) \, e^{-ipx} \int d\mu(q) \, e^{iqy} \, [a(\vec{p}\,), \; a^+(\vec{q}\,)] \\
&= \int d\mu(p) \, e^{-ipx} \int d\mu(q) \, e^{iqy} \, (2\pi)^3 \, 2\omega_p \, \delta^{(3)}(\vec{p} - \vec{q}\,) \\
&= \int d\mu(p) \, e^{-ip(x-y)} \\
&= (2\pi)^{-3} \int d^4p \, \Theta(p^0) \, \delta(p^2 - m^2) \, e^{-ip(x-y)} \\
&\doteq i\Delta^+(x-y; m^2)
\end{aligned} \tag{3.26}$$

This is the so called positive frequency part of the commutator. The latter is the sum of a positive and a negative frequency part

$$\Delta(x;m^2) \, = \Delta^+(x;m^2) \, + \Delta^-(x;m^2) \tag{3.27}$$

where

$$\Delta^-(x;m^2) \, = -\Delta^+(-x;m^2) \, = i(2\pi)^{-3} \int d^4p \, \Theta(-p^0) \, \delta(p^2-m^2) \, e^{-ipx} \quad . \tag{3.28}$$

For an adjoint pair of fermion-fields, similarly, we find

$$\begin{aligned} \{\psi^{(+)}(x), \bar{\psi}^{(-)}(y)\, \} &= \; < 0|\psi(x)\bar{\psi}(y)\,|0> \\ &= \; iS_F^+(x-y;m) \\ &= \; i\,(i\gamma^\mu\partial_\mu + m)\;\Delta^+(x-y;m^2) \end{aligned}$$

and for a vector field

$$\begin{aligned} [V_\mu^{(+)}(x), V_\nu^{(-)}(y)\,] &= \; < 0|V_\mu(x)V_\nu(y)\,|0> \\ &= \; iD_{\mu\nu}^+(x-y;m^2) \\ &= \; -i\left(g_{\mu\nu} + \frac{\partial_\mu\partial_\nu}{m^2}\right)\;\Delta^+(x-y;m^2) \quad . \end{aligned}$$

Notice that only fields describing identical particle–antiparticle species yield non-vanishing two–point functions.

If we normal-order several fields then for each pair of fields there is <u>one</u> term which must be reordered and yields besides the normal-ordered term an additional term where the corresponding pair is replaced by its vacuum-expectation value. With these observations one easily verifies the following statements:

Statement 1: The normal products have the orthogonality property

$$< 0| \; :\varphi_1(x_1)\cdots\varphi_n(x_n) :: \varphi_1'(y_1)\cdots\varphi_m'(y_m) \; : \; |0>= \delta_{nm}\sum_i\prod_{j=1}^n < 0|\varphi_j(x_j)\,\varphi_{i_j}'(y_{i_j})\,|0>$$

where the sum extends over all possible **pairings** (contractions) between the two sets of points $(x_1\cdots x_n)$ and $(y_1\cdots y_m)$.

A problem in calculating vacuum expectation values of products of free fields is that there are many identical contributions for which one only needs the multiplicities called usually **combinatorial factors**. It is sufficient then to concentrate in evaluating differing contributions. For this purpose the following statements about contractions between two normal–ordered monomials are useful:

Statement 2:

$$\frac{1}{n!\,m!} : \varphi^n(x) :: \varphi^m(y) := \sum_{k=1}^{min(n,m)} \frac{1}{k!}(< 0|\varphi(x)\varphi(y)|0>)^k \frac{1}{(n-k)!\,(m-k)!} : \varphi^{n-k}(x) :: \varphi^{m-k}(y) : \tag{3.29}$$

because for picking k fields out of n we have $\binom{n}{k}$ possibilities, a corresponding factor applies for picking k fields out of m, and there are $k!$ equivalent possibilities to contract the k fields at x with the k fields at y. This yields

$$\frac{n!\,m!}{(n-k)!\,k!\,(m-k)!\,k!}\,k!$$

and distributing the factorials appropriately we obtain the stated result. By normal ordering we avoided self–contractions at a point. For the latter we have the rule:

Statement 3:

$$\frac{1}{n!}\varphi^n(x) = \sum_{k=1}^{n/2} \frac{1}{2^k\, k!}\,(< 0|\varphi(x)\,\varphi(x)|0 >)^k \frac{1}{(n-2k)!} : \varphi^{n-2k}(x) : \tag{3.30}$$

Here we have $\binom{n}{2k}$ possibilities to pick $2k$ fields out of n. The $2k$ fields can be contracted in $(2k-1)!! = (2k-1)(2k-3)\cdots 3\cdot 1$ different ways, namely, the first one with one of the $(2k-1)$ others, after two are used up, the next one with one of the $(2k-3)$ remaining ones and so forth. The result then follows.

These rules will be useful for evaluating the combinatorial factors in Feynman rules, as we shall see below.

Simple examples illustrating the relationship between normal ordered and normal field monomials are

$$\begin{aligned}
:\phi(x): &= \phi(x) \\
:\phi^2(x): &= \phi^2(x) - < 0|\phi^2(x)|0 > \\
:\phi^4(x): &= \phi^4(x) - 3!\ < 0|\phi^2(x)|0 >: \phi^2(x) : - < 0|\phi^4(x)|0 > \\
:\partial_\mu\phi\partial^\mu\phi(x): &= \partial_\mu\phi\partial^\mu\phi(x) \quad \text{since} \quad < 0|\partial_\mu\phi\partial^\mu\phi(x)|0 >= 0 \\
& \quad etc.
\end{aligned}$$

By translation invariance

$$< 0|\phi^n(x)|0 >=< 0|\phi^n(0)|0 >=< 0|\phi^n|0 >$$

are constants. The transition from ordinary field products to normal-products amounts to a change of basis and inhibits a reparametrization. This is illustrated for the ϕ^4 -Lagrangian

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2}\,(\partial\phi)^2 - \frac{m^2}{2}\phi^2 \\
& \quad -\frac{\lambda}{4!}\phi^4 + \frac{c}{4!}\phi^4 + \frac{b}{2}\,(\partial\phi)^2 + \frac{a}{2}\phi^2 \\
& \quad + < 0|\mathcal{L}|0 > \\
&= \frac{1}{2}\ :(\partial\phi)^2: - \frac{m^2}{2} : \phi^2 : \\
& \quad -\frac{\lambda}{4!} : \phi^4 : + \frac{c}{4!} : \phi^4 : + \frac{b}{2}\ :(\partial\phi)^2: + \frac{a'}{2} : \phi^2 : \\
& \quad + < 0|\mathcal{L}|0 >'
\end{aligned}$$

where

$$\begin{aligned}
a' &= a + (c-\lambda)/2\ \ < 0|\phi^2|0 > \\
< 0|\mathcal{L}|0 >' &= < 0|\mathcal{L}|0 > + (a-m^2)/2\ < 0|\phi^2|0 > + (c-\lambda)/4!\ < 0|\phi^4|0 >
\end{aligned}$$

Since vacuum expectation values of $\mathcal{L}$ have to be subtracted in any case the only difference is a reparametrization of the mass term.

It should be noted that normal ordering in general leads to a reparametrization which spoils the classical form of the equations of motion. For example, field theories exhibiting a spontaneously broken symmetry (B. W. Lee 1969) or non–Abelian gauge theories broken by a Higgs mechanism are described more concisely if one avoids reparametrization by normal ordering .

The considerations on normal–ordering hold for time–ordered products as well if we apply the time-ordering prescription on both sides of the equations and notice that

$$T\{:\cdots:\} =: \cdots :$$

One then easily proves the following theorem (Wick 1950):

Wick's theorem:

$$\begin{aligned}
& T\left(\varphi_1(x_1)\cdots\varphi_n(x_n)\right) =: \varphi_1(x_1)\cdots\varphi_n(x_n) \ : \\
& +\textstyle\sum_i <0|T\left(\varphi_{i_1}(x_{i_1})\varphi_{i_2}(x_{i_2})\right)|0> : \varphi_{i_3}(x_{i_3})\cdots\varphi_{i_n}(x_{i_n}) \ : \\
& +\textstyle\sum_i <0|T\left(\varphi_{i_1}(x_{i_1})\varphi_{i_2}(x_{i_2})\right)|0> <0|T\left(\varphi_{i_3}(x_{i_3})\varphi_{i_4}(x_{i_4})\right)|0> : \varphi_{i_5}(x_{i_5})\cdots\varphi_{i_n}(x_{i_n}) \ : \\
& +\cdots \\
& +\textstyle\sum_i \begin{cases} <0|T\left(\varphi_{i_1}(x_{i_1})\varphi_{i_2}(x_{i_2})\right)|0> \cdots <0|T\left(\varphi_{i_{n-1}}(x_{i_{n-1}})\varphi_{i_n}(x_{i_n})\right)|0> & \text{(n even)} \\ <0|T\left(\varphi_{i_1}(x_{i_1})\varphi_{i_2}(x_{i_2})\right)|0> \ \cdots <0|T\left(\varphi_{i_{n-2}}(x_{i_{n-2}})\varphi_{i_{n-1}}(x_{i_{n-1}})\right)|0> : \varphi_{i_n}(x_{i_n}) \ : & \text{(n odd)} \end{cases}
\end{aligned}$$

which may be proved by induction. The sums extend over all permutations i . A simple consequence of Wick's theorem is that vacuum expectation values of time ordered products of fields are given by

$$\begin{aligned}
& <0|T\left(\varphi_1(x_1)\cdots\varphi_{2n}(x_{2n})\right)|0> \\
&= \textstyle\sum_i(-1)^{F_i} <0|T\left(\varphi_{i_1}(x_{i_1})\varphi_{i_2}(x_{i_2})\right)|0> \cdots <0|T\left(\varphi_{i_{2n-1}}(x_{i_{2n-1}})\varphi_{i_{2n}}(x_{i_{2n}})\right)|0> \\
& <0|T\left(\varphi_1(x_1)\cdots\varphi_{2n-1}(x_{2n-1})\right)|0> = 0
\end{aligned} \tag{3.31}$$

Note that non-vanishing contributions are obtained for **complete contractions**, only, where all fields are contracted pairwise. Next we look for the explicit form of a time–ordered contraction.

3.4.3 Stückelberg–Feynman Propagators.

In perturbation theory the main objects are the time–ordered products of free fields. We consider a scalar field first. In evaluating

$$\begin{aligned}
& T\{\cdots\varphi(x)\cdots\varphi(y)\cdots\} \\
= \ & T\{\cdots \quad \varphi(x)\varphi(y) \quad \cdots\}
\end{aligned}$$

we have to consider

$$\Theta(x^0-y^0)\varphi(x)\varphi(y) \ +\Theta(y^0-x^0)\varphi(y)\varphi(x)$$

and for the first term we have

$$\begin{aligned}
x^0 > y^0 \quad : \quad \varphi(x)\varphi(y) \ &= \ \left(\varphi^{(+)}(x) \ +\varphi^{(-)}(x)\right)\left(\varphi^{(+)}(y) \ +\varphi^{(-)}(y)\right) \\
&= \ \varphi^{(+)}(x)\varphi^{(+)}(y) \ +\varphi^{(-)}(x)\varphi^{(+)}(y) \ +\varphi^{(+)}(x)\varphi^{(-)}(y) \ +\varphi^{(-)}(x)\varphi^{(-)}(y) \\
&= \ :\varphi(x)\varphi(y): \ +\left[\varphi^{(+)}(x),\varphi^{(-)}(y)\right]
\end{aligned}$$

where the third term was reordered. Since the vacuum expectation value of the normal ordered term vanishes we obtain as a total time–ordered contribution

$$\Theta(x^0-y^0)\left[\varphi^{(+)}(x),\varphi^{(-)}(y)\right] \ +\Theta(y^0-x^0)\left[\varphi^{(+)}(y),\varphi^{(-)}(x)\right]$$

where

$$\left[\varphi^{(+)}(x), \varphi^{(-)}(y)\right] \equiv i\Delta^{+}(x-y)$$

is a c–number, the positive frequency part of the free field commutator. The time–ordered two–point Green function

$$\Theta(x^0 - y^0)\, i\Delta^{+}(x-y) \ + \Theta(y^0 - x^0)\, i\Delta^{+}(y-x) \equiv i\Delta_F(x-y) \tag{3.32}$$

describes the propagation of a free relativistic particle as it appears in relativistic scattering theory and is called **Feynman propagator** (Stückelberg 1941, Feynman 1949).

The explicit form of the Feynman propagator $\Delta_F(z)$ may be obtained easily in momentum space. Using Eqs. (3.26), (3.28) and (3.32) we can calculate $\tilde{\Delta}_F(q) = \int d^4z\, e^{iqz}\, \Delta_F(z)$. We obtain

$$\begin{aligned}
\tilde{\Delta}_F(q) &= -i(2\pi)^{-3} \int d^4p\, \delta(p^2 - m^2)\, (2\pi)^3\, \delta^{(3)}(\vec{p} - \vec{q}\,) \times \\
& \quad \left\{ \Theta(p^0) \int_0^\infty dz^0\, e^{-i(p^0 - q^0)\, z^0} + \Theta(-p^0) \int_{-\infty}^0 dz^0\, e^{-i(p^0 - q^0)\, z^0} \right\} \\
&= \frac{-i}{2\omega_q} \left\{ \int_0^\infty dz^0\, e^{-i(\omega_q - q^0)\, z^0} + \int_{-\infty}^0 dz^0\, e^{+i(\omega_q + q^0)\, z^0} \right\} \\
&= \frac{1}{q^2 - m^2 + i\epsilon} \quad \text{with} \quad \epsilon \to +0
\end{aligned}$$

For the last step we have used the following oscillatory integrals, which are well defined only as **boundary values** of an appropriate **analytic function** (so called $i\epsilon$–**prescription**): The integral

$$\int_0^\infty dz^0\, e^{-i\omega_q z^0}\, e^{iq^0 z^0} = ?$$

is obviously well defined in the complex q^0–plane for Im $(q^0) > 0$. We write $q^0 \to q^0 + i\epsilon$, $\epsilon > 0$ and obtain

$$\int_0^\infty dz^0\, e^{i(q^0 - \omega_q)\, z^0} = \left. \frac{e^{i(q^0 + i\epsilon - \omega_q)\, z^0}}{i(q^0 + i\epsilon - \omega_q)} \right|_0^\infty = \frac{1}{i(\omega_q - q^0 - i\epsilon)} \ .$$

In a similar way, with $q^0 \to q^0 - i\epsilon$, $\epsilon > 0$ now,

$$\int_{-\infty}^0 dz^0\, e^{i(q^0 + \omega_q)\, z^0} = \left. \frac{e^{i(q^0 - i\epsilon + \omega_q)\, z^0}}{i(q^0 - i\epsilon + \omega_q)} \right|_{-\infty}^0 = \frac{1}{i(\omega_q + q^0 - i\epsilon)}$$

and the sum of the two integrals is

$$\frac{1}{i} \left\{ \frac{1}{\omega_q + q^0 - i\epsilon} + \frac{1}{\omega_q - q^0 - i\epsilon} \right\} = \frac{1}{i} \frac{2\omega_q}{m^2 - q^2 - i\epsilon} \ .$$

Here we have replaced $i\epsilon\omega_q$ by the equivalent $i\epsilon$ and neglected terms of order $O(\epsilon^2)$. This is justified because we are interested in the limit $\lim_{\epsilon \to 0}$ only and because $\omega_q > 0$ in any case. The $i\epsilon$–prescription means that the limit $\epsilon \to +0$ is to be taken at the end. This will be understood always in the following.

In Fig. 3.4 we show how the poles of the Feynman propagator must approach the real q^0–axis starting from the finite ϵ "regularization".

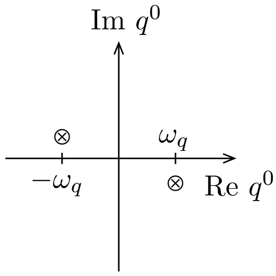


Fig. 3.4: Location of the two poles of the Feynman propagator in the complex q^0–plane.

As a result we find in momentum space

$$\tilde{\Delta}_F(q) = \frac{1}{q^2 - m^2 + i\epsilon} \qquad (\epsilon \to +0)$$

and

$$\Delta_F(z) = (2\pi)^{-4} \int d^4q \, \frac{e^{-iqz}}{q^2 - m^2 + i\epsilon} \qquad (\epsilon \to +0) \quad . \tag{3.33}$$

3.4.4 Feynman rules for the ϕ^4-model.

We are prepared now to formulate the Feynman rules. As a simple example we consider the ϕ^4-model, which we introduced in Sec. 3.1 as an example of an interacting theory. We thus consider a single scalar field of mass m with self-interaction

$$\mathcal{L}_{int} = -\frac{\lambda}{4!}\phi^4\,(x) \quad .$$

We first consider the τ–functions Eq. (3.21). For a given order n we have to find all possible contributions. We already know that vacuum expectation values only are non-vanishing if all fields are pairwisely contracted. We thus arrive at the result:

$$\begin{aligned} {}_{in}\!< 0|T\{\varphi_{in}(x_1)\ldots\varphi_{in}(x_N) \int d^4y_1\, \mathcal{L}^{(in)}_{int}(y_1), \cdots, \int d^4y_n\, \mathcal{L}^{(in)}_{int}(y_n)\}|0>_{in\otimes} \\ = \sum_{\Gamma} \int d^4y_1 \cdots d^4y_n \prod_{z_i, z_k \in x_1,\ldots,y_n} i\Delta_F(z_i - z_k) \end{aligned} \tag{3.34}$$

where the sum is over all possible complete contractions of free fields and each complete contraction corresponds to a product of $(N+2n)/2$ Feynman propagators. Each possible contraction scheme may be characterized by a Feynman graph Γ which is obtained as follows:

- Each contribution at n-th order perturbation theory is characterized by a Feynman diagram with N external $(x_1, x_2 \ldots, x_N)$ and n internal $(y_1, y_2, \ldots, y_n)$ vertices (drawn as points in a plane) which are completely contracted i.e. connected by lines (propagators).

- In configuration space the contribution to the τ–function characterized by a particular Feynman graph is given by the products of propagators which are represented by the lines of the graph. The internal vertices are integrated out as follows from Eq. (3.34).

For actual calculations it is much simpler to work in momentum space. We use the convention Eq. (3.20) for the Fourier transformation of the fields. We thus take the Fourier transform

$$\tilde{\tau}^{(N)}(p_1,p_2,\ldots,p_N) = \int d^4x_1\, d^4x_2 \ldots d^4x_N\, e^{i\sum_{i=1}^N p_i x_i} < 0|T\{\varphi(x_1)\,\varphi(x_2)\ldots\varphi(x_N)\}|0 >$$

with the convention that all momenta are incoming. By translation invariance the τ–functions

$$\tau^{(N)}(x_1,x_2,\ldots,x_N) =< 0|T\{\varphi(x_1)\,\varphi(x_2)\ldots\varphi(x_N)\}|0 >$$

depend on the coordinate differences $\xi_i = x_i - x_{i+1}\ \ (i=1,\ldots,N-1)$ only

$$\tau^{(N)}(x_1,x_2,\ldots,x_N) = \mathcal{T}^{(N)}(\xi_1,\xi_2,\ldots,\xi_{N-1})$$

and this implies four-momentum conservation, $\sum_{i=1}^N p_i = 0$, such that

$$\tilde{\tau}^{(N)}(p_1,p_2,\ldots,p_N) = (2\pi)^4\delta^{(4)}(\sum_{i=1}^N p_i)\,\tilde{\mathcal{T}}^{(N)}(p_1,p_1+p_2,\ldots,p_1+p_2+\ldots+p_{N-1})$$

where

$$\tilde{\mathcal{T}}^{(N)}(q_1,q_2,\ldots,q_{N-1}) = \int d^4\xi_1\, d^4\xi_2 \ldots d^4\xi_{N-1}\, e^{i\sum_{i=1}^{N-1} q_i\xi_i}\mathcal{T}^{(N)}(\xi_1,\xi_2,\ldots,\xi_{N-1})$$

depends on the $N-1$ independent momenta $q_1 = p_1,\ q_2 = p_1+p_2,\ \ldots\ ,q_{N-1} = p_1+p_2+\ldots+p_{N-1}$. In the perturbation expansion an internal vertex is proportional to

$$\int d^4x\,\phi^4(x) \;=\; \int d^4x\,\prod_{i=1}^4\int_{q_i} e^{-iq_ix}\tilde{\phi}(q_i) \;=\; \prod_{i=1}^4\int_{q_i}(2\pi)^4\,\delta^{(4)}(\sum_{i=1}^4 q_i)\,\tilde{\phi}(q_i)$$

and one four-momentum integration can be performed trivially with the help of the four-momentum conservation delta–function at each vertex. We obtain the following **Feynman rules** in momentum space:

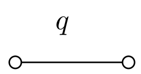
Expand the Fourier transform of the τ–function in terms of Feynman diagrams

$$< 0|T\{\tilde{\varphi}(p_1)\,\tilde{\varphi}(p_2)\ldots\tilde{\varphi}(p_N)\}|0 >= \sum_\Gamma < 0|T\{\tilde{\varphi}(p_1)\,\tilde{\varphi}(p_2)\ldots\tilde{\varphi}(p_N)\}|0 >^\Gamma \quad .$$

The sum extends over all topologically distinct Feynman diagrams Γ which we can draw with N external lines and the interaction vertices. Vacuum sub-diagrams (disconnected components of Γ without external lines) are not permitted ($\otimes$–prescription$\equiv$ vacuum normalization). We label external lines by momenta $p_1,p_2,\ldots,p_N$. These are chosen incoming by convention. Internal lines we label by $k_1,k_2,\ldots,k_I$. The contribution from the diagram Γ is then obtained via the following correspondence:

Lines and vertices represent the factors: (see Eq. (3.8))

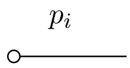
❶ Lines:



$\frac{1}{q^2-m^2+i\epsilon}$; $\quad q \in \{p_i, k_j\}$

❷ Vertices:

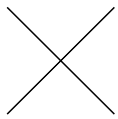
a) external:



$$1$$

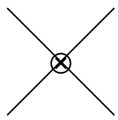
b) internal:

(bare interaction vertex)



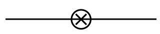
$$-\lambda$$

(vertex renormalization)



$$c(\lambda)$$

(mass and wave–function renormalizations)



$$a(\lambda) + b(\lambda)\, q^2$$

❸ Integration: At each vertex we have four-momentum conservation. After taking into account this, those internal momenta k_i which are not determined by the external momenta are called **loop momenta** and remain to be integrated over

$$\frac{1}{(2\pi)^d} \int d^d k_i \cdots$$

where d is the space–time dimension.

❹ Factors: Multiply the integrals over the products of propagators and coupling matrices obtained so far by the following additional factors:

total four-momentum conservation	:	$(2\pi)^d \delta^{(d)} \left(\sum p_{i\, ext}\right)$
each interaction vertex	:	i
each propagator	:	i
combinatorial factors	:	$\frac{1}{c(\Gamma)}$ (see below)

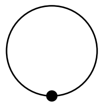
A diagrams is **connected** if it does not consist of disconnected parts. Note that four-momentum conservation holds for each connected component, such that in general one has a product of δ–functions in place of just one overall one.

The combinatorial factor is given by the symmetry number

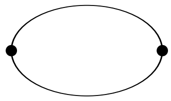
$$c(\Gamma) = s 2^{n_1} 2^{n_3} (3!)^{n_3}$$

of a diagram. n_1 is the number of lines connecting a vertex with itself, n_2 is the number of double lines (two lines connecting a pair of vertices) and n_3 the number of triple lines (three lines connecting a pair of vertices). s is the number of permutations of vertices which leave the diagram unchanged (external vertices fixed).

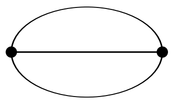
The rules follow from Eqs. (3.29) and (3.30): Take out a permutation symmetry factor $\frac{1}{n!}$ for each product of n identical fields at a vertex. These symmetry factors are omitted in the Feynman rules. Then the multiple lines have the weight factors



$: \ \frac{1}{2!}$



$: \ \frac{1}{2!}$



$: \ \frac{1}{3!}$

as given above.

The contribution from a diagram Γ has the form

$$< 0|T\{\tilde{\varphi}(p_1)\,\tilde{\varphi}(p_2)\ldots\tilde{\varphi}(p_N)\}|0 >^\Gamma = \frac{1}{c(\Gamma)}\delta_\Gamma \int \frac{d^dk_1}{(2\pi)^d}\cdots\frac{d^dk_L}{(2\pi)^d}\, I_\Gamma(p_i,k_j)$$

where $I_\Gamma(p_i,k_j)$ is the **Feynman integrand** and

$$\delta_\Gamma = \prod_\gamma (2\pi)^d \delta^{(d)}\left(\sum p^\gamma_{i\;ext}\right)$$

is the product δ–functions expressing energy–momentum conservation, one factor for each connected component γ of Γ. The number of **loops** (nontrivial integrations) L is at most

$$L \leq I - (V-1)$$

where I is the number of internal lines and V the number of internal vertices. This is just counting the number of δ–functions at the vertices minus one for the overall energy–momentum conservation. This holds for each connected component of a diagram. Diagrams which have no parts which are connected just by a single particle line (the diagram cannot be cut into two parts by cutting a single line) we call **one particle irreducible** ($1pi$). For connected $1pi$ diagrams we have

$$L = I - (V-1)\ \ .$$

The S–matrix elements for the scattering of N particles are obtained from the τ–functions by replacing the external propagators (attached at the external vertices) by the one–particle wave functions, as required by the LSZ reduction formulas (see Tab. 3.1). The latter are just 1 for scalar particles.

When calculating a certain quantity in perturbation theory one, obviously, proceeds order by order, starting with the **Born or tree approximation**, which is represented by tree type diagrams which do not exhibit loops. The expansion in the coupling constant corresponds to a loop expansion, i.e., corrections are represented by the one-loop, two-loop, ... diagrams contributing to a given process, which is fixed by specifying the external legs. In a renormalizable theory interaction vertices can be field monomials with three or four fields only. Since all connected diagrams are trees $1pi$ diagrams we need consider the $1pi$ diagrams only. Let the diagram have N external legs, V_3 three-vertices and V_4 four-vertices. Then we have

$$3V_3 + 4V_4 - N = 2I \;\; ; \quad V_3 + V_4 = V$$

where, again, I is the number of internal lines and V the number of internal vertices. Now there are several possibilities:

1. $\lambda\varphi^3$ theory: $V_4 = 0$ and hence

$$3V - N = 2I = 2(L + V - 1) \qquad \text{or} \;\; V = N - 2 + 2L$$

which shows that the expansion in powers V of λ is in fact a loop expansion in powers of λ^2, since L can take integer values only.

2. $\lambda\varphi^4$ theory: $V_3 = 0$ and hence

$$4V - N = 2I = 2(L + V - 1) \qquad \text{or} \;\; 2V = N - 2 + 2L$$

which shows that the expansion in powers V of λ is in fact a loop expansion in powers of λ, since L can take integer values only.

3. Non–Abelian gauge theory: Here, as we shall see later, three-vertices are of order g and four-vertices of order g^2 where g is the gauge coupling constant. The order in g is now

$$O = V_3 + 2V_4$$

and hence, eliminating $V_3 = 2V - O$ and $V_4 = O - V$,

$$2V + O - N = 2I = 2(L + V - 1) \qquad \text{or} \;\; O = N - 2 + 2L$$

which shows that the expansion in powers O of g is in fact a loop expansion in powers of g^2. V has dropped from the relationship, here.

For an amplitude $A^{(N)}$ with N external legs we have

$$A^{(N)} = g^{-(N-2)} \left(1 + a_1 \left(\frac{\alpha}{\pi}\right) + \cdots + a_L \left(\frac{\alpha}{\pi}\right)^L + \cdots\right) \;\; ; \;\; \alpha = \frac{g^2}{4\pi}$$

where the factor π is included because it naturally comes in via the loop integrals.

4. In theories like the electroweak Standard Model one has, in general, many coupling constants and it is not obvious that there is a simple relationship between the number of loops and the multi coupling constant expansion. In spite of the many couplings in the SM there is however a natural way to set up a perturbation expansion, which relates to the loop expansion. The key for this possibility is the fact that masses in the SM are generated by the Higgs mechanism (spontaneously broken gauge theory). Typically masses are generated via the vacuum expectation value v of a scalar field, i. e. a mass m_i is proportional to a coupling constant g_i times the vacuum expectation value: $m_i \propto g_i v$. If, as in the case of the SM, the theory in the unbroken phase $v = 0$ is massless (the scalar field itself may be an exception) and only has dimensionless couplings then after the breaking one has the following situation: The quadrilinear couplings are ratios of the form M^2/v^2 whereas the trilinear couplings are of the form M^2/v or m/v (M denotes a boson mass, m a fermion mass). We immediately see that counting powers of v^{-1} is the same as counting powers of g in the previous example. Therefore the expansion in powers of v^{-2}: For an amplitude $A^{(N)}$ with N external legs we have

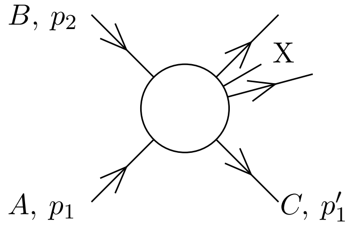
$$A^{(N)} = v^{-(N-2)} \left(1 + a_1 x + \cdots + a_L x^L + \cdots\right) \ ; \ \ x = \frac{1}{(4\pi v)^2}$$

where the factor 4π is included because it naturally comes in via the loop integrals.

3.5 Cross sections and decay rates

The **differential cross section** for a two particle collision

$$A(p_1) \ + B(p_2) \ \rightarrow \ C(p_1') \ + D(p_2') \ \cdots$$



$$P_i = p_1 + p_2$$
$$P_f = p_1' + p_2' + \cdots$$

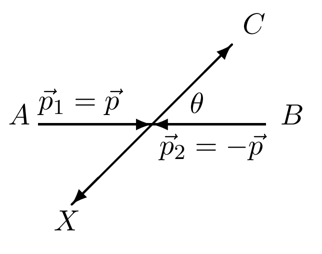
is given by

$$d\sigma = \frac{(2\pi)^4 \delta^{(4)}\left(P_f - P_i\right)}{2\sqrt{\lambda\left(s, m_1^2, m_2^2\right)}} \mid T_{fi} \mid^2 \, d\mu(p_1') d\mu(p_2') \cdots \tag{3.35}$$

$s = (p_1 + p_2)^2$ is the square of the total CM energy and $\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz$ is a two body phase-space function. In the CM frame (see the figure):

$$\sqrt{\lambda} = \sqrt{\lambda\left(s, m_1^2, m_2^2\right)} = 2 \mid \vec{p} \mid \sqrt{s}$$

where $\vec{p} = \vec{p}_i$ is the three-momentum of the initial state particle A.



The **total cross section** follows from the differential one by "summation" over all final states

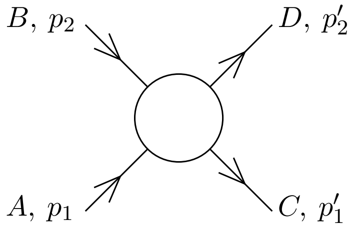
$$\sigma_{tot} = \int d\sigma = \frac{(2\pi)^4}{2\sqrt{\lambda\left(s, m_1^2, m_2^2\right)}} \sum_{spins'} \int \prod_i d\mu(p_i') \, \delta^{(4)} \left(\sum_i p_i' - p_1 - p_2 \right) \mid T_{fi} \mid^2 \ .$$

In particular, for a two-body final state $A+B \to C+D$ one defines the Lorentz invariant kinematic variables

$$\begin{aligned} s &= (p_1+p_2)^2 = (p_1'+p_2')^2 \\ t &= (p_1-p_1')^2 = (p_2-p_2')^2 \\ u &= (p_1-p_2')^2 = (p_2-p_1')^2 \end{aligned}$$

the so–called **Mandelstam variables**. They satisfy

$$s+t+u = m_1^2+m_2^2+m_1'^2+m_2'^2 \doteq \Sigma$$



$$\begin{aligned} P_i &= p_1+p_2 \\ P_f &= p_1'+p_2' \end{aligned}$$

The final state two-body phase-space function

$$\sqrt{\lambda'} = \sqrt{\lambda(s, m_1'^2, m_2'^2)} = 2 \mid \vec{p}_f \mid \sqrt{s} = s\sqrt{y_+ y_-}$$

determines the modulus of final state CM momentum $\vec{p}_f = \vec{p}_1\,' = -\vec{p}_2\,'$. The variables $y_\pm$ are given by

$$y_\pm = 1 - \frac{(m_1' \pm m_2')^2}{s} \ .$$

The invariant **momentum transfer** t is given by

$$t = (p_1-p_1')^2 = m_1^2+m_1'^2 - 2p_1p_1' = m_1^2+m_1'^2 - 2\left(E_1E_1' - |\vec{p}_i\,||\vec{p}_f\,|\cos\theta\right)$$

with $E_j = \sqrt{m_j^2+\vec{p}_j\;^2}$, or, written in a manifestly invariant way,

$$t = \frac{1}{2}\left(\Sigma - s + \frac{1}{s}\sqrt{\lambda}\sqrt{\lambda'}\cos\theta\right)$$

and u follows from t by a substitution $\cos\theta \to -\cos\theta$. For a final state particle the three-momentum volume element may be written as $d^3p = \vec{p}\,^2\, d|\vec{p}\,|d\Omega$ where $d\Omega$ is the element of solid angle

$$d\Omega = \sin\theta\, d\theta\, d\phi = -d(\cos\theta)\, d\phi \ .$$

Notice that $|T|^2$ cannot depend on the azimuthal angle ϕ because the process is symmetric with respect to rotations about the beam-axis. Formally this follows because $|T|^2$ is L-invariant and hence can depend only on the scalar products of the momenta which all are determined by s, t and the four particle masses. Then, performing all trivial integrations (using four-momentum conservation), we obtain

$$\left(\frac{d\sigma}{d\Omega}\right)_{CM} = \frac{1}{(8\pi)^2 s}\sqrt{\frac{\lambda(s,m_1'^2,m_2'^2)}{\lambda(s,m_1^2,m_2^2)}} \mid T_{fi}\mid^2$$

or

$$\frac{d\sigma}{dt} = \frac{\mid T_{fi} \mid^2}{16\pi\lambda(s, m_1^2, m_2^2)}$$

where

$$dt = -2|\vec{p}_i \,||\vec{p}_f \,|\frac{d\Omega}{2\pi} \; .$$

If initial and final states are unpolarized we have to average over the initial state polarizations and to sum over the final state polarizations such that

$$\frac{d\sigma}{dt}^{unpol} = \frac{1}{2j_1+1}\frac{1}{2j_2+1}\sum_{spins}\frac{d\sigma}{dt} \; . \tag{3.36}$$

We may reconsider at this point the optical theorem. The imaginary part or **absorptive part** of the forward scattering amplitude of an elastic process $A+B \to A+B$ is proportional to the sum over all possible final states $A+B \to$ "anything" which defines the total **inclusive cross section**

$$\sigma_{tot}\,(A+B \to X) = \frac{(2\pi)^4}{2\sqrt{\lambda(s,m_1^2,m_2^2)}} \; \cdot$$
$$\sum_n \sum_{spins'} \int \prod_{i=1}^n \frac{d^3p'_i}{(2\pi)^3\, 2\omega_{p'_i}} \delta^{(4)}(P'_n - P_i) \;\mid T\,(p'_1, j'_1, \lambda'_1, \ldots, p'_n, j'_n, \lambda'_n \mid p_1, j_1, \lambda_1, p_2, j_2, \lambda_2) \mid^2 \; .$$

The state X is anything allowed by conservation laws. Thus X must have the same conserved quantum numbers as the initial state. We know are able to write down the final version of the optical theorem

$$\mathrm{Im}\; T_{\mathrm{forward}}\;(A+B \to A+B) = \sqrt{\lambda\,(s, m_1^2, m_2^2)}\;\sigma_{tot}\,(A+B \to \;\text{anything}) \tag{3.37}$$

For unpolarized particles in the initial state one has to average over spins on both sides of the equation. The relation is depicted graphically in Fig. 3.3.

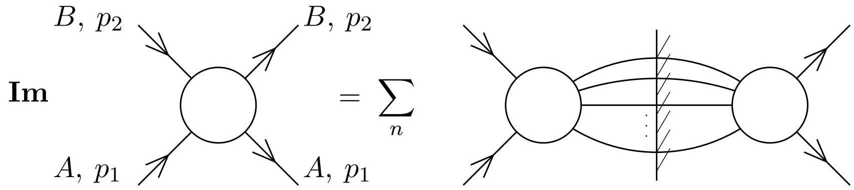


Figure 3.3: The optical theorem yields the inelasticity (3.37) in terms of the forward scattering amplitude.

This relationship can be tested experimentally: On the one hand one measures the differential cross-section $\frac{d\sigma}{d\Omega}(\theta)$ as a function of the scattering angle θ for elastic scattering $A+B \to A+B$ and then extrapolates to $\theta \to 0$. On the other hand one measures the total inclusive cross-section for $A+B \to$ anything.

Since

$$\left(\frac{d\sigma}{d\Omega}\right)^{\mathrm{elastic}}_{\theta=0} = \frac{1}{(8\pi)^2 s} \mid T_{\mathrm{forward}}\,(A+B \to A+B) \mid^2$$

and

$$\mid T_{\mathrm{forward}} \mid^2 = (\mathrm{Re}\; T_{\mathrm{forward}})^2 + (\mathrm{Im}\; T_{\mathrm{forward}})^2$$

one obtains the unitarity bound

$$\left(\frac{d\sigma}{d\Omega}\right)^{\text{elastic}}_{\theta=0} \geq \frac{1}{(8\pi)^2 s}\lambda\left(s, m_1^2, m_2^2\right)\ \sigma^2_{tot}\left(A + B \to \text{ anything}\right)\ .$$

Finally we consider the decay of unstable particles. The differential **decay rate** for $A \to B + C + \cdots$ is given by

$$\Big\| \ d\Gamma = \tfrac{(2\pi)^4 \delta^{(4)}\left(P_f - P_i\right)}{2m_1} \mid T_{fi} \mid^2 d\mu(p_1')d\mu(p_2')\cdots \tag{3.38}$$

By "summing" over all possible decay channels we find the total width

$$\Gamma = \not{\sum}\!\!\!\!\!\int d\Gamma = \frac{1}{\tau}$$

where τ is the **lifetime** of the particle, which decays via the exponential **decay law**

$$N(t) = N_0\, e^{-t/\tau}$$

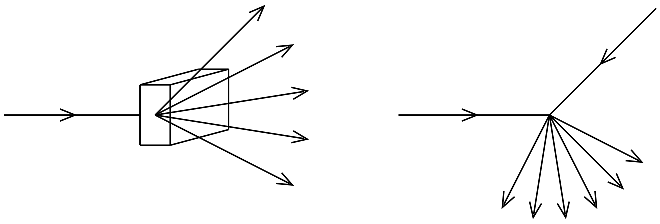
For a two body decay the momenta of the decay particles are fixed by the masses $A(M) \to B(m_1) + C(m_2)$ $(M > m_1 + m_2)$ and one obtains:

$$\boxed{\Gamma = \tfrac{M}{16\pi}\sqrt{1 - \tfrac{(m_1 - m_2)^2}{M^2}}\sqrt{1 - \tfrac{(m_1 + m_2)^2}{M^2}}\,\tfrac{1}{2s_V + 1}\,\tfrac{\sum \mid T_{12} \mid^2}{M^2}} \tag{3.39}$$

Here s_V is the spin of the decaying particle and $\frac{1}{2s_V+1}\sum \mid T_{12} \mid^2$ the trensition matrixelement averaged over the spin of the decaying particle. This formula (see (14.4) will be derived later in Sec. 14.

Appendix to Section 3:
Derivation of the differential cross section

In a typical scattering experiment a beam of practically free particles hits either a target, in a fixed target experiment, or another beam, in a collider experiment.



After the interaction the scattered and/or newly produced particles move practically freely before they are kinematically analyzed and identified in a detector.

Given a suitably prepared initial state $|i>$, we are interested to calculate the probabilities for the system to undergo a transition into a possible final state $|f>$. The result of a collision is a superposition

$$\sum_f |f><f|S|i> \tag{3.40}$$

of all possible final states. S is the scattering operator introduced in Sec. 3.2 (see also Sec. 3.1) and the amplitudes $S_{fi} =< f|S|i>$ are the S–matrix elements. The probability for observing a particular final state $|f>$ is given by $|S_{fi}|^2$.

It is convenient to split off the trivial part of the S–matrix, which describes the free passing of particles, and to factor out the total four-momentum conservation, which leads us to the transition matrix T. The T–matrix is defined by

$$< f|S|i >= S_{fi} = \delta_{fi} + i(2\pi)^4\, \delta^{(4)}(P_f - P_i)\ \cdot T_{fi}\ \ . \tag{3.41}$$

For $\underline{f \neq i}$ we have

$$S_{fi} = i(2\pi)^4\, \delta^{(4)}(P_f - P_i)\ \cdot T_{fi}\ \ . \tag{3.42}$$

At this point a serious problem is encountered if we attempt to calculate the **transition probability** $|S_{fi}|^2$. Obviously, with the above expression, we get an infinite answer, a non-sense result proportional to $\left(\delta^{(4)}(P_f - P_i)\right)^2$. The reason for the problem is that, so far, we have calculated S_{fi} using plane wave scattering states, which are not localized in space-time and hence are not normalizable. The problem may be circumvented either by going to "box quantization" or by working with normalizable states.

Heuristic derivation of the differential cross section

We first present a simple heuristic derivation of the differential cross section formula Eq. (3.35) and of the differential width formula (3.38). For this purpose we consider the experiment to be confined in a box of spatial volume V, where V is assumed to be sufficiently large as compared to the effective range of the interaction. The number of states in a box of volume V with momenta in a momentum space element d^3p is:

$$\frac{V}{(2\pi)^3}\, d^3p \ \ .$$

Once we have the four-momentum conservation

$$(2\pi)^4\,\delta^{(4)}(P_f - P_i) = \int d^4x e^{-i(P_f-P_i)\,x}$$

as we go to infinity with the volume, a second such factor for $P_f = P_i$ becomes

$$\text{“}\ \ (2\pi)^4\,\delta^{(4)}(P_f - P_i)\ \ \text{”} = \int d^4x e^{-i(P_f-P_i)\,x}\Big|_{P_f=P_i} = \int d^4x = V\cdot t$$

where t is a time interval, sufficiently large as compared to the interaction time. We thus may write

$$\begin{aligned}\text{“}\Big((2\pi)^4\,\delta^{(4)}(P_f - P_i)\Big)^2\text{”} &= \text{“}\ \ (2\pi)^4\,\delta^{(4)}(P_f - P_i)\int d^4x e^{-i(P_f-P_i)\,x}\ \ \text{”}\\ &= (2\pi)^4\,\delta^{(4)}(P_f - P_i)\cdot V\cdot t\ \ ,\end{aligned}$$

such that

$$|S_{fi}|^2 = (2\pi)^4\,\delta^{(4)}(P_f - P_i)\,|T_{fi}|^2\,Vt\ \ .$$

We then obtain the transition probability per unit time

$$\mathcal{P}_{fi} = (2\pi)^4\,\delta^{(4)}(P_f - P_i)\,|T_{fi}|^2\,V\ \ .$$

In order to proceed, we have to match the normalization of the states in the box to the ones we previously used to calculate T_{fi}. The normalization of the states in the box is

$$< p_{n'}|p_n >= \delta_{n',n}\ \ ,$$

where $|p_n>$ is a state of <u>one</u> particle in the box in the state n. This has to be compared with our plane wave normalization

$$< \vec{p}\,', \alpha' \mid \vec{p}, \alpha >= \delta_{\alpha\alpha'}\,(2\pi)^3\,2\omega_p\,\delta^{(3)}\left(\vec{p} - \vec{p}\,'\right)\ \ ,$$

which can be obtained as a limit from the box states as

$$\int_V \left(e^{i\vec{p}_{n'}\,\vec{x}}\right)^* e^{i\vec{p}_n\,\vec{x}} = V\cdot\delta_{n',n} \to (2\pi)^3\,\delta^{(3)}\left(\vec{p} - \vec{p}\,'\right)\ \ .$$

Here we used that

$$\int_V d^3x\, e^{i(\vec{p}-\vec{p}\,')\,\vec{x}} \ \to\ \begin{cases}\int d^3x = V & \text{for } \vec{p}\,' = \vec{p}\\ \quad 0 & \text{otherwise}\end{cases}$$

on the one hand, and

$$\int_V d^3x\, e^{i(\vec{p}-\vec{p}\,')\,\vec{x}} \ \to\ (2\pi)^3\,\delta^{(3)}\left(\vec{p} - \vec{p}\,'\right)$$

on the other hand, for $V \to \infty$. Accordingly, we must identify

$$|p_n> \to \frac{1}{\sqrt{2\omega_p V}} \; |\vec{p}, \alpha>$$

and thus

$$\begin{array}{ccc} T_{fi} & \to \frac{1}{\sqrt{2\omega_{p_1}V\cdots}\sqrt{2\omega_{p'_1}V\cdots}} & T_{fi} \\ \uparrow & & \uparrow \\ \text{box normalization} & & \text{conventional normalization} \end{array} \quad .$$

The transition probability for final states with particles with momentum in the element $d^3p'_i$ around $\vec{p}\,'_i$ is

$$\begin{aligned} dP &= \mathcal{P}_{fi} \prod_{i=1}^{n} \frac{V}{(2\pi)^3}\, d^3p'_i \\ &= (2\pi)^4\, \delta^{(4)}(P_f - P_i)\, V\, \frac{|T_{fi}|^2}{2\omega_{p_1}V\cdots 2\omega_{p'_1}V\cdots} \prod_{i=1}^{n} \frac{V}{(2\pi)^3}\, d^3p'_i \\ &= (2\pi)^4\, \delta^{(4)}(P_f - P_i) \frac{V|T_{fi}|^2}{2\omega_{p_1}V\cdots} \prod_{i=1}^{n} d\mu(p'_i) \quad . \end{aligned}$$

We may now consider particular initial states.

For one particle in the initial state we directly obtain the differential **decay width**. In the rest frame of the decaying particle we have

$$d\Gamma = \frac{(2\pi)^4\, \delta^{(4)}(P_f - P_i)}{2m_1} \, |T_{fi}|^2\, d\mu(p'_1)\cdots d\mu(p'_n) \quad .$$

For two particles in the initial state we have the transition probability

$$dP = \frac{(2\pi)^4\, \delta^{(4)}(P_f - P_i)}{2E_1\, 2E_2\, V} \, |T_{fi}|^2\, d\mu(p'_1)\cdots d\mu(p'_n) \quad .$$

The cross section is defined by the ratio of the transition probability dP per unit time and the current density j of the incoming particles:

$$d\sigma = \frac{dP}{j}$$

with

$$j = \frac{I}{VE_1E_2} \; ; \; I = \left((p_1p_2)^2 - m_1^2m_2^2\right)^{1/2} \quad .$$

Note that I is Lorentz invariant, as it should be. In the c.m. system we have $\vec{p}_1 = -\vec{p}_2 = \vec{p}$ and hence $I = |\vec{p}|\,(E_1 + E_2)$ and thus

$$j = \frac{|\vec{p}|}{V} \left(\frac{1}{E_1} + \frac{1}{E_2}\right) = \frac{v_1 + v_2}{V} \quad .$$

Since $v_1 + v_2 = v_{in}$ is the relative velocity of the incoming particles, this is the conventional definition of the current density for colliding particles, normalized to one particle. For an arbitrary frame we have

$$j = \frac{1}{V} \left((\vec{v}_1 - \vec{v}_2\,)^2 - (\vec{v}_1 \times \vec{v}_2\,)^2\right)^{1/2}$$

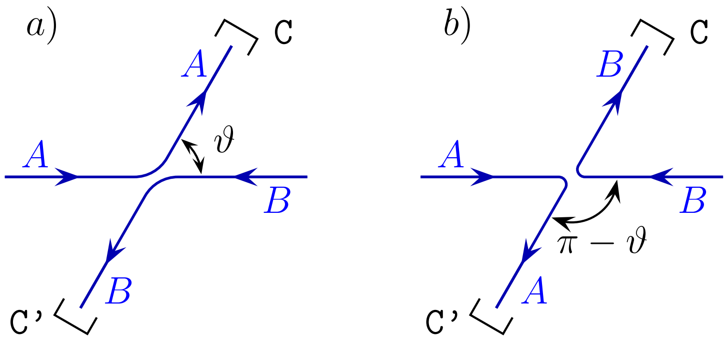


Figure 3.4: Distinguishable versus indistinguishable particle scattering

which reduces to the conventional current density $j = |\vec{v}_1 - \vec{v}_2|/V$ in the case $\vec{v}_1 \,||\, \vec{v}_2$.

As a final result we have

$$d\sigma = \frac{(2\pi)^4\,\delta^{(4)}(P_f - P_i)}{2\sqrt{\lambda(s, m_1^2, m_2^2)}}\,|T_{fi}|^2\,d\mu(p_1') \cdots d\mu(p_n')$$

with $\sqrt{\lambda(s, m_1^2, m_2^2)} = 2I$.

Scattering and decay involving identical particles

So far we assumed all scattering particles to be distinguishable. In many cases actually we have to deal with identical particles, for example, if two photons or two Z bosons or two or more π^0's are produced. The problem is illustrated in Fig. 3.4. If A and B are different kinds of particles the probability that a particle (A or B) is observed at the counter C is[15]

$$|A(\vartheta)|^2 + |A(\pi - \vartheta)|^2\;. \tag{3.43}$$

We add the two probabilities since the two possible processes are distinguishable, but we have chosen not to distinguish between them. In contrast, if the particles are indistinguishable, the two processes are indistinguishable, they do not represent different processes. Thus it appears that for identical particles one has to integrate only over one hemisphere, which may be converted into a full angular integration provided we divide by 2. As a rule in the cross section and decay rate formulae (3.35) and (3.38), respectively, one always has to divide by the number of indistinguishable states (degeneracy factor) and for n identical bosons there are $n!$ indistinguishable permutations (averaging over indistinguishable possibilities).

The problem addressed here has its formal solution remembering that in quantum mechanics multi-particle states with identical particles have to be symmetrized (Bosons) or anti-symmetrized (Fermions) and normalized appropriately. The proper QM two-particle sates must have the form

$$|\Phi\rangle_{\rm in,out} = \frac{1}{\sqrt{2}}\,(|\phi_A, \phi_B\rangle \pm |\phi_B, \phi_A\rangle)_{\rm in,out} \tag{3.44}$$

where the + sign holds for bosons and − sign for fermions. On the right-hand side of these equations it is supposed that we can label the indistinguishable (non-interacting) particles by subscripts A and B where the particle A and B are in states described by a set of eigenvalues

[15] the azimuthal angle φ is redundant here because of the azimuthal symmetry of the problem considered

ϕ_A and ϕ_B, respectively. The "state" $|\phi_A, \phi_B\rangle_{\rm out}$ corresponds to the one detected by the counter of Fig. 3.4a alone while $|\phi_B, \phi_A\rangle_{\rm out}$ corresponds to the one detected by the counter of Fig. 3.4b. However, quantum mechanically (physically) only the state described by the superposition $|\Phi\rangle_{\rm out}$ is detectable (an apparatus which would be able to distinguish the two "states" does not exist). Taking the transition matrix element between the states (3.44) we have

$$\begin{aligned} {}_{\rm out}\langle\Phi'|\Phi\rangle_{\rm in} &= \frac{1}{2}\Big({}_{\rm out}\langle\phi'_A,\phi'_B|\phi_A,\phi_B\rangle_{\rm in} \pm {}_{\rm out}\langle\phi'_B,\phi'_A|\phi_A,\phi_B\rangle_{\rm in} + (A\leftrightarrow B)\Big) \\ &= \quad {}_{\rm out}\langle\phi'_A,\phi'_B|\phi_A,\phi_B\rangle_{\rm in} \pm {}_{\rm out}\langle\phi'_B,\phi'_A|\phi_A,\phi_B\rangle_{\rm in}\;. \end{aligned}$$

The second equality holds because A and B are dummy labels when the particles are indistinguishable. This translates into

$$|A(\vartheta) \pm A(\pi-\vartheta)|^2\;. \tag{3.45}$$

This result has to be compared with the one obtained for the case of distinguishable particles (3.43).

An additional complication arises for identical particles because of the role of the spin. Two otherwise identical particles (same species) may be distinguishable if they have different spin. Hence, in general, in a mixed state we have to distinguish the component in which the particles have identical spin and thus are indistinguishable and a component where the particles have different spin and thus are distinguishable. For two spin 1/2 fermions, two electrons say , and for unpolarized electrons then the probability of detecting an electron at counter `C` reads

$$\frac{1}{2}\,\left(|A(\vartheta)|^2+|A(\pi-\vartheta)|^2\right)+\frac{1}{2}\,\left(|A(\vartheta)-A(\pi-\vartheta)|^2\right)$$

as there is an equal probability that the electrons can be distinguished or not.
These considerations show that one has to be careful in the formal definition of the transition probabilities, because the result in general depends on the experimental set up and the precise definition of the observable of interest. Here so called *maximal observations* play a special role, where a system is analyzed with respect to all possible simultaneous measurements. The latter are related to a complete system of commuting observables, which posses a complete system of simultaneous eigenvectors and eigenvalues. Note that quantum mechanically such complete information does not remove the identical particle degeneracy as we know.

Exercise: Show that $\lambda(s,m_1^2,m_2^2)$ can be written as

$$\lambda(s,m_1^2,m_2^2) = \left(s-(m_1+m_2)^2\right)\left(s-(m_1-m_2)^2\right)\;;\quad s=(p_1+p_2)^2\quad .$$

λ is the two–body phase space function, sometimes also called Källén–function.

Units: The total cross section

$$\sigma_{\rm tot} = \int d\sigma$$

has the dimension of an area. Typical cross section units are:

$$\begin{aligned} 1 \text{ barn} &= 10^{-24}\ \mathrm{cm}^2 \\ 1 \text{ millibarn (mb)} &= 10^{-27}\ \mathrm{cm}^2 \end{aligned}$$

and the subunits micro- (μb), nano- (nb), pico- (pb) and femto- (fb) barns. An orders of magnitude comparison we consider the Compton wave length of a particle of mass m: $\lambda_C = \hbar/(mc)$

as a measure for the uncertainty principle in quantum mechanical localization. A natural area to compare with accordingly is λ_C^2.

Examples: strong interaction (short ranged) considered as mediated by one pion exchange:

$$\begin{aligned}
\frac{\hbar}{m_\pi c} &\simeq 1.41 \text{ fermi } = 1.41 \times 10^{-13} \text{ cm} \\
\left(\frac{\hbar}{m_\pi c}\right)^2 &\simeq 2(\text{ fermi })^2 = 20 \text{ mb}
\end{aligned}$$

the total cross section for π^+p–scattering at energies 5 to 10 GeV, well above the proton mass, is

$$\sigma_{\text{tot}} \simeq 25mb \simeq 1.25 \times \left(\frac{\hbar}{m_\pi c}\right)^2$$

A proper derivation of the differential cross section

Particle beams and target in reality have a finite momentum resolution with momenta concentrated around average momenta $\bar{p}_1, \bar{p}_2 \ldots$. They must be described by appropriate wave packets. The improper translationally invariant momentum eigenstates

$$|\,\vec{p}, \alpha> = |\,\vec{p}, \lambda, \bar{\alpha}>$$

of helicity λ and other quantum numbers $\bar{\alpha}$ are normalized by

$$<\vec{p}\,', \alpha' \,|\, \vec{p}, \alpha> = \delta_{\alpha\alpha'}\,(2\pi)^3\,2\omega_p\,\delta^{(3)}\,(\vec{p} - \vec{p}\,')\,.$$

A state of finite momentum resolution is described by a distribution function $\tilde{\varphi}(p)$ in Fourier space:

$$\begin{aligned}
|\,\varphi, \lambda, \bar{\alpha}> &\doteq \int \frac{d^4p}{(2\pi)^4}\,2\pi\,\Theta(p^0)\,\delta(p^2 - m^2)\,\tilde{\varphi}(p)\,|\,\vec{p}, \lambda, \bar{\alpha}> \\
&= \int d\mu(p)\tilde{\varphi}(p)\,|, \vec{p}\lambda\bar{\alpha}>
\end{aligned}$$

with normalization

$$\begin{aligned}
<\varphi, \lambda, \alpha \,|\, \varphi, \lambda, \alpha> &= \int d\mu(p')\,d\mu(p)\,\tilde{\varphi}^*(p')\,\tilde{\varphi}(p) \cdot <\vec{p}\,', \lambda, \alpha \,|\, \vec{p}, \lambda, \alpha> \\
&= \int d\mu(p)|\tilde{\varphi}(p)|^2 = N
\end{aligned}$$

as a number of particles in a beam pulse or in a target. In configuration space we then have

$$\begin{aligned}
\varphi(x) &= \int \frac{d^4p}{(2\pi)^4}\,2\pi\,\Theta(p^0)\,\delta(p^2 - m^2)\,\tilde{\varphi}(p)\,e^{-ipx} \\
&= \int \frac{d^3p}{(2\pi)^3\,2\omega_p}\,\tilde{\varphi}(\omega_p, \vec{p}\,)\,e^{-i\omega_p x^0}\,e^{i\vec{p}\vec{x}}
\end{aligned}$$

with

$$|\varphi(x)|^2 = \int d\mu(p')\,d\mu(p)\,\tilde{\varphi}^*(p')\,\tilde{\varphi}(p)\,e^{-i(p-p')\,x}$$

and

$$\begin{aligned}
\int d^3x|\varphi(x)|^2 &= \int d\mu(p')\, d\mu(p)\, \tilde{\varphi}^*(p')\, \tilde{\varphi}(p)\, e^{-i(\omega_p-\omega_{p'})\, x^0}\, (2\pi)^3\, \delta^{(3)}\, (\vec{p}-\vec{p}\,') \\
&= \int d\mu(p) \frac{|\tilde{\varphi}(p)|^2}{2\omega_p} \\
&= \int d\mu(p) \frac{|\tilde{\varphi}(p)|^2}{2\omega_p} \\
&\simeq \frac{1}{2\bar{\omega}_p} \int d\mu(p) |\tilde{\varphi}(p)|^2 = \frac{N}{2\bar{\omega}_p} \quad .
\end{aligned}$$

Here, we have made an approximation. We have assumed that the momentum resolution is sufficiently narrow in momentum space, such that the wave packet approximates to some extent a plane wave. This assumption will be made also in the following discussion. Thus we have the

Result:

$$|\varphi(x)|^2 = \frac{\rho(x)}{2\bar{\omega}_p} \tag{3.46}$$

where

$\bar{\omega}_p$ is the average energy of the beam particles or of the target particles, and

$\rho(x)$ is the particle density function: the probability per unit volume to find a particle at time $t = x^0$ at the point $\vec{x}$.

For later use we note that

$$\int d^4x\, e^{ipx}\, \varphi(x) = 2\pi\, \Theta(p^0)\, \delta(p^2 - m^2)\, \tilde{\varphi}(p) \quad .$$

We now consider a particle collision process

$$\begin{array}{ccc}
A + B & \to & n - \text{particles} \\
q_1, q_2 & & p_1, p_2, \ldots, p_n \\
\tilde{\varphi}_1(q_1), \tilde{\varphi}_2(q_2) & &
\end{array}$$

where $\tilde{\varphi}_1(q_1), \tilde{\varphi}_2(q_2)$ are the momentum distributions of the two incoming particles.

The transition matrix element is given by

$$\int d\mu(q_1) d\mu(q_2) \tilde{\varphi}_1(q_1) \tilde{\varphi}_2(q_2) < p_1, p_2, \ldots, p_n, \beta|\, (S - I)\, |q_1, q_2, \alpha >$$

where α and β describe the collection of quantum numbers of the in and out states, respectively. The transition probability for the transition into a state with momenta $p_1, p_2, \ldots, p_n$ in the continuum part of the spectrum reads

$$\begin{aligned}
P(\varphi_1, \varphi_2; p_1, p_2, \ldots, p_n) = \textstyle\int d\mu(q_1') d\mu(q_2') d\mu(q_1) d\mu(q_2) \tilde{\varphi}_1^*(q_1') \tilde{\varphi}_2^*(q_2') \tilde{\varphi}_1(q_1) \tilde{\varphi}_2(q_2) \cdot \\
< q_1', q_2', \alpha|\, (S - I)^+\, |p_1, p_2, \ldots, p_n, \beta >< p_1, p_2, \ldots, p_n, \beta|\, (S - I)\, |q_1, q_2, \alpha >
\end{aligned}$$

with

$$< p_1, p_2, \ldots, p_n, \beta|\, (S - I)\, |q_1, q_2, \alpha >= i(2\pi)^4\, \delta^{(4)}(\textstyle\sum p_i - q_1 - q_2)\, T(p_1, p_2, \ldots, p_n, \beta; q_1, q_2, \alpha) \quad .$$

The transition matrix element T is a continuous function of the momenta, to all orders in perturbation theory. As before we assume that the wave packets $\tilde{\varphi}_1(q_1), \tilde{\varphi}_2(q_2)$ are concentrated to a narrow range of momenta, such that $T(\ldots)$ is well approximated by a constant on the support of the wave functions.

Therefore we obtain

$$\begin{aligned}&\int d\mu(q_1)d\mu(q_2)\tilde{\varphi}_1(q_1)\tilde{\varphi}_2(q_2)\cdot(2\pi)^4\,\delta^{(4)}(\textstyle\sum p_i - q_1 - q_2)\,T(p_1,p_2,\ldots,p_n,\beta;q_1,q_2,\alpha)\\ \simeq\, & T(p_1,p_2,\ldots,p_n,\beta;\bar{q}_1,\bar{q}_2,\alpha)\times\int d\mu(q_1)d\mu(q_2)\tilde{\varphi}_1(q_1)\tilde{\varphi}_2(q_2)\cdot(2\pi)^4\,\delta^{(4)}(\textstyle\sum p_i - q_1 - q_2)\ .\end{aligned}$$

We denote

$$T_{ni} = T(p_1,p_2,\ldots,p_n,\beta;\bar{q}_1,\bar{q}_2,\alpha)$$

and get

$$\begin{aligned}P(\varphi_1,\varphi_2;p_1,p_2,\ldots,p_n) \;=\; |T_{ni}|^2 \;\;&\times\;\int d\mu(q_1')d\mu(q_2')\tilde{\varphi}_1^*(q_1')\tilde{\varphi}_2^*(q_2')(2\pi)^4\,\delta^{(4)}(\textstyle\sum p_i - q_1' - q_2')\\ &\times\;\int d\mu(q_1)d\mu(q_2)\tilde{\varphi}_1(q_1)\tilde{\varphi}_2(q_2)(2\pi)^4\,\delta^{(4)}(\textstyle\sum p_i - q_1 - q_2)\end{aligned}$$

The distribution factors may be evaluated

$$\begin{aligned}&\int d\mu(q_1)d\mu(q_2)\tilde{\varphi}_1(q_1)\tilde{\varphi}_2(q_2)(2\pi)^4\,\delta^{(4)}(\textstyle\sum p_i - q_1 - q_2)\\ =\;&\int \frac{d^4q_1}{(2\pi)^4}\,2\pi\,\Theta(q_1^0)\,\delta(q_1^2-m_1^2)\,2\pi\,\Theta(q_2^0)\,\delta(q_2^2-m_2^2)\,\tilde{\varphi}_1(q_1)\,\tilde{\varphi}_2(q_2)\Big|_{q_2=\sum p_i - q_1}\\ =\;&\int \frac{d^4q_1}{(2\pi)^4}\int d^4x_1\int d^4x_2\,e^{iq_1x_1}e^{i(\sum p_i - q_1)\,x_2}\,\varphi_1(x_1)\,\varphi_2(x_2)\\ =\;&\int d^4x\,e^{i\sum p_i\,x}\varphi_1(x)\varphi_2(x)\end{aligned}$$

and hence

$$P(\varphi_1,\varphi_2;p_1,p_2,\ldots,p_n) = |T_{ni}|^2\times\int d^4x\,d^4y\,e^{i\sum p_i\,(x-y)}\varphi_1(x)\varphi_2(x)\varphi_1^*(y)\varphi_2^*(y) \tag{3.47}$$

as a probability density.

Our final states are normalized such that the covariant projection

$$P_{n,\beta} = \int\prod_{i=1}^{n} d\mu(p_i)\,|p_1,\ldots,p_n,\beta><\beta,p_1,\ldots,p_n|$$

onto the n–particle states with quantum numbers β and

$$\sum_\beta P_{n,\beta} = 1$$

on the n–particle states. Therefore, we have

$$dP \propto P(\varphi_1,\varphi_2,p_1,\cdots,p_n)\,d\mu(p_1)\cdots d\mu(p_n)\ . \tag{3.48}$$

Note that

$$d\mu(p_i) = \frac{d^3p_i}{(2\pi)^3 2\omega_{p_i}}$$

is just the density of states in the relativistically invariant volume element.

So far particles in the beam and in the target have been prepared with a *finite resolution*. We therefore have to "sum" over all final states which are compatible with the initial distribution. The finite spread manifests itself in a distribution in the total c.m. four-momentum

$$P = q_1 + q_2$$

which we have to integrate over. The support properties of our distribution functions thereby take automatically care that just those final states are integrated over which are compatible with the given distribution. We first note the identity

$$\prod_{i=1}^{n} d\mu(p_i) \equiv \int \frac{d^4P}{(2\pi)^4} \prod_{i=1}^{n} d\mu(p_i)\, (2\pi)^4\, \delta^{(4)}(\sum p_i - P)$$

If this is inserted in Eq. (3.48), because of the support properties of the distributions $\tilde{\varphi}_i$ we have $P \simeq \bar{P} = \bar{q}_1 + \bar{q}_2$ such that the δ–function can be take out from the integral:

$$\begin{aligned} dP &= \int \frac{d^4P}{(2\pi)^4}\, (2\pi)^4\, \delta^{(4)}(\sum p_i - P)\, P(\varphi_1, \varphi_2, p_1, \cdots, p_n)\, d\mu(p_1) \cdots d\mu(p_n) \\ &= \int \frac{d^4P}{(2\pi)^4}\, (2\pi)^4\, \delta^{(4)}(\sum p_i - P)\, |T_{ni}|^2\, d\mu(p_1) \cdots d\mu(p_n) \\ &\quad \times \int d^4x\, d^4y\, e^{iP\,(x-y)} \varphi_1(x)\varphi_2(x)\varphi_1^*(y)\varphi_2^*(y) \\ &\simeq (2\pi)^4\, \delta^{(4)}(\sum p_i - \bar{P})\, |T_{ni}|^2\, d\mu(p_1) \cdots d\mu(p_n) \\ &\quad \times \int \frac{d^4P}{(2\pi)^4} \int d^4x\, d^4y\, e^{iP\,(x-y)} \varphi_1(x)\varphi_2(x)\varphi_1^*(y)\varphi_2^*(y) \;\; . \end{aligned}$$

The integral over P yields a δ–function in configuration space and we may use Eq. (3.46) to obtain

$$\begin{aligned} &\int \frac{d^4P}{(2\pi)^4} \int d^4x\, d^4y\, e^{iP\,(x-y)} \varphi_1(x)\varphi_2(x)\varphi_1^*(y)\varphi_2^*(y) \\ &= \int d^4x\, |\varphi_1(x)|^2\, |\varphi_2(x)|^2 = \frac{1}{2\bar{\omega}_{q_1} 2\bar{\omega}_{q_2}} \int d^4x\, \rho_1(x)\, \rho_2(x) \end{aligned}$$

in terms of the particle density functions $rho(x)$. As a final result we have

$$dP \;\simeq\; \frac{(2\pi)^4\, \delta^{(4)}(\sum p_i - \bar{P})}{2\bar{\omega}_{q_1} 2\bar{\omega}_{q_2}}\, |T_{fi}|^2 \prod_{i=1}^{n} d\mu(p_i)\; \times \int d^4x\, \rho_1(x)\, \rho_2(x) \tag{3.49}$$

This is the differential probability for the scattering into n–particle final states with momenta p_i in the momentum space elements d^3p_i which are allowed by the resolution of the beam and the target, integrated over time. The **counting rate** is in fact given by

$$\frac{dP}{dt} \simeq dn \;=\; \frac{(2\pi)^4\, \delta^{(4)}(\sum p_i - \bar{P})}{2\bar{\omega}_{q_1} 2\bar{\omega}_{q_2}}\, |T_{fi}|^2 \prod_{i=1}^{n} d\mu(p_i)\; \times \int d^3x\, \rho_1(t, \vec{x}\,)\, \rho_2(t, \vec{x}\,) \;\; .$$

The relativistically invariant differential cross section Eq. (3.35)

$$d\sigma = \frac{(2\pi)^4 \delta^{(4)}\,(P_f - P_i)}{2\sqrt{\lambda\,(s, m_1^2, m_2^2)}}\, \mid T_{fi} \mid^2\, d\mu(p_1') \cdots d\mu(p_n')$$

is defined as the ratio of the counting rate and the particle current density j. The latter is given by

$$j = \frac{2\sqrt{\lambda\left(s, m_1^2, m_2^2\right)}}{2\bar{\omega}_{q_1} 2\bar{\omega}_{q2}} \cdot \int d^3x \, \rho_1(x) \, \rho_2(x)$$

where $\rho_1(x)$ is the beam particle density and $\rho_2(x)$ the target particle density.

In the c.m. frame $\vec{q}_1 = -\vec{q}_2 = \vec{q}$, $\sqrt{s} = E_1 + E_2$ and $\lambda = 2|\vec{q}|\sqrt{s}$. Thus the factor $2\sqrt{\lambda}/(2\bar{\omega}_{q_1} 2\bar{\omega}_{q_2})$, with $\bar{\omega}_{q_i} = E_i$, reads $|\vec{q}\,|\,(\frac{1}{E_1} + \frac{1}{E_2}) = v_{in}$, which is the relative velocity of the incident particles. Thus

$$j = v_{in} \int d^3x \, \rho_1(x) \, \rho_2(x) \quad .$$

Interpretation: In the laboratory system $\Phi(t) = v_{in} \; \rho_1(t, \vec{x}\,)$ is the incoming flux of particles, the number of particles which hit the target per unit area and per unit time, at position $\vec{x}$ at time t. In other words, $v_{in}\, \rho_1(x)\, dA\, dt$ is the number of beam particles, which hit within the time interval dt the area element dA.

$\rho_2(t, \vec{x}\,)$ is the target particle density. In other words, $\rho_2(x)\, dA \; d\ell$ is the number of target particles in the volume element $dV = dA \, d\ell$, at position $\vec{x}$ at time t.

Let σ_{fi} be the effective area of a target particle for scattering of a beam particle in the process $|i> \to \; |f>$ then the number of beam particles per unit time which are scattered by the target into the state $|f>$ is

$$\frac{dN_{fi}(t)}{dt} = n_{fi}(t) = \sigma_{fi} \; v_{in} \int d^3x \, \rho_1(t, \vec{x}\,) \, \rho_2(t, \vec{x}\,)$$

or

$$n_{fi}(t) = \sigma_{fi} \; j(t)$$

Let

$$N_{fi}(t) = \int_{t_0}^{t} dt' \; n_{fi}(t')$$

be the total number of events during the period (t, t_0) into the state $|f>$, then

$$\sigma_{fi} = \frac{N_{fi}(t)}{v_{in} \; \int_{t_0}^{t} \; dt' \; d^3x \; \rho_1(t', \vec{x}\,) \; \rho_2(t', \vec{x}\,)}$$

or for the differential cross section

$$d\sigma_{fi} = \frac{dN_{fi}(t)}{v_{in} \; \int d^4x \; \rho_1(x) \; \rho_2(x)} = \frac{dn_{fi}}{j}$$

which is independent on t, because

$$\int d^4x \; \cdots = \int_{t_0}^{t} \; dx^0 \; \int_V d^3x \; \cdots$$

extends over the duration of the experiment and spatially over the whole target volume. The only important condition is that

$$\begin{aligned} \Delta x &\gg \hbar/\Delta p \\ \Delta t &\gg \hbar/\Delta E \qquad . \end{aligned}$$

In an idealized situation, where the particle densities would be constant in space and time we would obtain

$$n_{fi} = \sigma_{fi}\, v_{in} \cdot V \cdot n_1 n_2 = \sigma_{fi}\, \frac{v_{in}}{V}\, N_1 N_2 = \sigma_{fi}\, j$$

with

$$j = \frac{v_{in}}{V}\, N_1 N_2$$

Summary of the relevant assumptions which go into the derivation of the cross-section formulae:

- All particles in the beam have practically the same momentum $\bar{q}_1$ and similarly the target particles have momenta strongly peaked near $\bar{q}_2$. Consequently, the relative velocity v_{in} is practically the same for all collisions.
- The total number of events N_{fi} must be sufficiently large, such that the result is statistically significant.
- The volume and the duration of the experiment must be sufficiently large, relative to the quantum fluctuations and range of the interactions.

Note: There are additional conditions, even for short ranged interactions as we have always assumed in the above derivations. As an example we mention that the target should be very thin such that multiple scattering is excluded. In colliders beam–beam interaction may lead to additional complications etc. The theoretician usually leaves these problems to the experimenters and do not think about them.

3.6 Dispersion relations, spectral representation

Here we consider some general properties of two–point functions of field operators. By "general properties" we mean properties which rely on basic features of any QFT, independent of a specific form of the interaction Lagrangian and which do not require to resort to perturbation theory. We first derive some results for the simplest case of a scalar field $\varphi(x)$ and its two–point function $W(x-y) =< 0|\varphi(x)\varphi(y)|0 >$, which is often called Wightman function. We shall assume the theory to be regularized in some appropriate way, such that it is well defined. The Hilbert space associated with the fields may considered to be spanned by a complete set of eigenstates of the four–momentum operators P_μ

$$P_\mu|p,\alpha\rangle = p_\mu|p,\alpha\rangle$$

with α the additional quantum numbers. The basic properties (P) the following considerations rely on are the following:

P1) completeness: $\sum_\alpha \frac{1}{(2\pi)^3}\int d^4p|p,\alpha\rangle\langle p,\alpha| = 1$[16]

P2) spectral condition: $p^2 \geq 0$; $p^0 \geq 0$

P3) translation invariance of the vacuum: $P_\mu|0>= 0$

P4) translation invariance

P5) L–invariance

P6) parity invariance (optional QED, QCD)

We now consider the above mentioned two–point function and insert a complete set of intermediate states:

$$\begin{aligned} < 0|\varphi(x)\varphi(y)|0 > &= \sum_\alpha \frac{1}{(2\pi)^3}\int d^4p\ < 0|\varphi(x)|p,\alpha\rangle\langle p,\alpha|\varphi(y)|0 > \\ &= \sum_\alpha \frac{1}{(2\pi)^3}\int d^4p\,|< 0|\varphi(0)|p,\alpha\rangle|^2\, e^{-ip(x-y)} \end{aligned}$$

where we used translational invariance (2.19) $U(a) = e^{ia_\mu P^\mu}$ the translation operator

$$e^{ia_\mu P^\mu}\varphi(x)\, e^{-ia_\mu P^\mu} = \varphi(x+a)$$

and the transformation law (2.2)

$$U(a)|p,\alpha\rangle = e^{iap}|p,\alpha\rangle\ ;\ \ U(a)|0>= |0>$$

of the eigenstates of P_μ.

$$\begin{aligned} < 0|\varphi(x)|\vec{p},\alpha\rangle &= < 0|U(x)\,\varphi(0)U^{-1}(x)|\vec{p},\alpha\rangle \\ &= < 0|\varphi(0)|\vec{p},\alpha\rangle e^{-ixp} \end{aligned}$$

etc.
We define the **spectral function**

$$\tilde{\rho}(p) \doteq \sum_\alpha |< 0|\varphi(0)|p,\alpha\rangle|^2 \tag{3.50}$$

which has the properties

[16] Since the states in general are not free one particle states the completeness relation (2.13) has to be generalized. The completeness integral also includes the integration over p^0 since we do not have $p^0 = \sqrt{m^2+\vec{p}^{\,2}}$ any longer.

1) positivity: $\tilde{\rho}(p) \geq 0$

2) spectral condition: $\tilde{\rho}(p) \neq 0 \Leftrightarrow p^2 \geq 0 \;, \quad p^0 \geq 0$

3) L–invariance: $\tilde{\rho}(p) = \tilde{\rho}(\Lambda p)$

The latter property implies with $U = U(\Lambda)$ see (2.2)

$$\begin{aligned} & < 0|\varphi(0)|p,\alpha\rangle =< 0|U^+U\varphi(0)U^+U|p,\alpha\rangle \\ = \;& < 0|\varphi(0)|\Lambda p,\alpha'\rangle D_{\alpha'\alpha} =< 0|\varphi(0)|\bar{p},\bar{\alpha}\rangle \end{aligned}$$

where $|\bar{p},\bar{\alpha}\rangle = |\Lambda p,\alpha'\rangle D_{\alpha'\alpha}$ is a new basis, in which we again have completeness since we performed a unitary transformation (we have used $U\varphi(0)U^+ = \varphi(0)$ (scalar field) and $U|0>=|0>$ etc.).

As a consequence $\tilde{\rho}(p)$ must be a function of p^2 only:

$$\tilde{\rho}(p) = \Theta(p^0)\,\Theta(p^2)\,\rho(p^2) \tag{3.51}$$

and we may write

$$\Theta(p^2) = \int_0^\infty dm^2 \delta(p^2 - m^2)\;.$$

Thus for the Wightman function

$$W(x-y) = \int_0^\infty dm^2 \int d^4p\,\rho(p^2)\,\delta(p^2-m^2)\,\Theta(p^0)\,e^{-ip(x-y)} \tag{3.52}$$

or utilizing the invariant function (3.26)

$$i\Delta^+(x-y;m^2) = (2\pi)^{-3}\int d^4p\,\Theta(p^0)\,\delta(p^2-m^2)\,e^{-ip(x-y)}$$

we finally have

$$W(x-y) = i\int_0^\infty dm^2 \rho(m^2)\Delta^+(x-y;m^2) \tag{3.53}$$

which is the wanted **spectral representation** for the scalar Wightman function with spectral density $\rho(m^2)$. It generalizes the $i\Delta^+(x-y;m^2)$ invariant function of a free field to the corresponding invariant function for an arbitrary interacting field.

Using the following relations between the invariant functions

$$\begin{aligned} \Delta^-(x;m^2) &= -\Delta^+(-x;m^2) \\ \Delta(x;m^2) &= \Delta^+(x;m^2) + \Delta^-(x;m^2) \\ \Delta^F(x;m^2) &= \Theta(x^0)\,\Delta^+(x;m^2) - \Theta(-x^0)\,\Delta^-(x;m^2) \end{aligned}$$

we easily find

$$
\begin{aligned}
<0|\varphi(x)\varphi(y)|0> &= i\int\limits_0^\infty dm^2\rho(m^2)\Delta^+(x-y;m^2) \\
<0|\left[\varphi(x)\varphi(y)\right]|0> &= i\int\limits_0^\infty dm^2\rho(m^2)\Delta(x-y;m^2) \\
<0|\,T\left(\varphi(x)\varphi(y)\right)|0> &= i\int\limits_0^\infty dm^2\rho(m^2)\Delta^F(x-y;m^2)\;.
\end{aligned}
\tag{3.54}
$$

The vacuum expectation value of the commutator $<0|\left[\varphi(x)\varphi(y)\right]|0>$ for interacting field vanishes for space-like separation $(x-y)^2<0$. This directly follows from the spectral representation and by the fact that $\Delta(x-y;m^2)$ is L–invariant and vanishes at $x^0=y^0$: $\Delta(x-y;m^2)|_{x^0=y^0}=0$. The mass spectrum represented by the eigenvalues of P_μ is displayed in 2.1. Here we assume $\varphi(x)$ to be a neutral pseudo-scalar field which describes the neutral pion of mass μ. Then automatically $<0|\varphi(0)|2n>=0$ for $n=0,1,2,\cdots$ due to parity mismatch between field and states, i.e., parity only allows states with an odd number of pions. The continuous part of the spectrum thus starts at $p^2\geq(3\mu)^2$. The one particle intermediate states which are on the mass shell $p^2=\mu^2$ may be taken into account separately as a contribution to the spectral function ρ which is proportional to $\delta(p^2-\mu^2)$, thus

$$
\begin{aligned}
\rho(p^2) &= |<0|\varphi(0)|p;\;p^2=\mu^2\rangle|^2\,\delta(p^2-\mu^2)+\sum\!\!\!\!\!\!\int_{\alpha,\;\mathrm{continuum}}|<0|\varphi(0)|p,\alpha\rangle|^2 \\
&= Z_3\,\delta(p^2-\mu^2)+\sigma(p^2)
\end{aligned}
$$

Z_3 is the square of the φ wave function renormalization constant. The full (interacting) Feynman propagator thus takes the form

$$
\boxed{<0|\,T\left(\varphi(x)\varphi(y)\right)|0>=i\,Z_3\,\Delta^F(x-y;\mu^2)\;+\;i\int\limits_{(3\mu)^2}^\infty dm^2\sigma(m^2)\Delta^F(x-y;m^2)\;.}
\tag{3.55}
$$

If we require the canonical commutation relations to hold

$$
\begin{aligned}
\left[\dot\varphi(x),\varphi(y)\right]|_{x^0=y^0} &= -i\,\delta^3(\vec{x}-\vec{y}) \\
<0|\left[\varphi(x)\varphi(y)\right]|0> &= i\int\limits_0^\infty dm^2\rho(m^2)\Delta(x-y;m^2) \\
\partial_0\,\Delta(x-y;m^2)\Big|_{x^0=y^0} &= -\delta^3(\vec{x}-\vec{y})
\end{aligned}
$$

implies

$$
\int\limits_0^\infty dm^2\rho(m^2)=Z_3+\int\limits_{(3\mu)^2}^\infty dm^2\sigma(m^2)=1
\tag{3.56}
$$

which implies

$$
\boxed{0\leq Z_3\leq 1}
\tag{3.57}
$$

3.6.1 Bare and physical mass

Here the dynamics comes into play. The equations of motion have the form

$$\left(\Box + \mu_0^2\right) \varphi(x) = j(x)$$

with a scalar current $j(x)$ describing the interaction. When parity is conserved and φ is a pseudo-scalar field the corresponding Lagrangian term would read $\mathcal{L}_{int} = -\frac{g_0}{4!}\varphi^4$ and hence $j(x) = \frac{\partial \mathcal{L}_{int}}{\partial \phi(x)} = -\frac{g_0}{3!}\varphi^3$. In the presence of interactions the parameters are bare parameters in first palace (see Sec. 3.1). On the other hand, the mass in the spectral representation is the physical mass, i.e.,

$$\left(\Box + m^2\right) \Delta(x; m^2) = 0$$

Applying the equation of motion to

$$< 0| \left[\varphi(x)\varphi(y)\right] |0 >= i \int_0^\infty dm^2 \rho(m^2) \Delta(x - y; m^2)$$

we obtain

$$< 0| \left[j(x)\varphi(y)\right] |0 >= i \int_0^\infty dm^2 \rho(m^2) \left(\mu_0^2 - m^2\right) \Delta(x - y; m^2)$$

which does not vanish because of the mismatch between the bare and the physical mass. If we take the time derivative at $x^0 = y^0$ using the property of $\Delta(x - y; m^2)$ we find

$$\int_0^\infty dm^2 \rho(m^2) \left(\mu_0^2 - m^2\right) = 0 \tag{3.58}$$

Together with $\rho(m^2) = Z_3\, \delta(m^2 - \mu^2) + \sigma(m^2)$ and $\int_0^\infty dm^2 \rho(m^2) = 1$ we obtain

$$\mu_0^2 = Z_3 \mu^2 + \int_{(3\mu)^2}^\infty dm^2 m^2\, \sigma(m^2)$$

$$\mu_0^2 - \mu^2 = \int_{(3\mu)^2}^\infty dm^2 \left(m^2 - \mu^2\right) \sigma(m^2)$$

which implies

$$\boxed{\mu_0^2 > \mu^2} \tag{3.59}$$

We may give the following interpretation to the above result:

$$\mu = \mu_0 + M + \varepsilon$$

μ the physical mass

μ_0 the bare mass

M the contribution of the cloud of virtual excitations

ε the binding energy

By consideration of the full propagator im Fourier space

$$\tilde{G}(k^2) = \int_0^\infty dm^2 \frac{\rho(m^2)}{k^2 - m^2 + i\varepsilon} = \frac{Z_3}{k^2 - \mu^2 + i\varepsilon} + \int_{(3\mu)^2}^\infty dm^2 \frac{\rho(m^2)}{k^2 - m^2 + i\varepsilon}$$

we learn that in the region of

- small k: main contributions from $k^2 \sim \mu^2$; μ **dressed** (i.e. physical) **mass** cloud has time to follow the motion of the particle.
- large k: using (3.58) and (3.56) we find $\int_0^\infty dm^2 \frac{\rho(m^2)}{k^2-m^2+i\varepsilon} = \int_0^\infty dm^2 \frac{\rho(m^2)}{k^2} \left\{1 + \frac{m^2}{k^2} + \cdots\right\} = \frac{1}{k^2}\left\{1 + \frac{\mu_0^2}{k^2} + \cdots\right\} = \frac{1}{k^2-\mu_0^2+i\varepsilon}$ which shows that the **bare mass** is relevant.

This is similar to the charge screening in QED: at long distances the vacuum polarization causes the screening of the charge due to virtual pairs (dipoles) adjusting in the field of the charge. If one pores closer and closer with higher energies one is penetrating the cloud of virtual pairs and more and more sees the bare charge.

3.6.2 Analyticity and dispersion relations

The analytic behavior of free fields has been discussed earlier in Sec. 3.4.3. Here with the help of the spectral decomposition (3.55) we may extend the consideration to interacting fields. We consider the time–ordered product

$$G(x-y) =< 0|\, T\left(\varphi(x)\varphi(y)\right) |0 >= i \int_0^\infty dm^2 \rho(m^2) \Delta^F(x-y;m^2)$$

and

$$\Delta_F(z) = (2\pi)^{-4} \int d^4q \, \frac{e^{-iqz}}{q^2 - m^2 + i\epsilon} \qquad (\epsilon \to +0) \quad .$$

we get

$$\tilde{G}(k^2) = \frac{Z_3}{k^2 - \mu^2 + i\varepsilon} + \int_{(3\mu)^2}^\infty dm^2 \frac{\sigma(m^2)}{k^2 - m^2 + i\varepsilon}$$

which we consider now for arbitrary complex $\zeta = k^2$

$$\tilde{G}(\zeta) = \frac{Z_3}{\zeta - \mu^2} + \int_{(3\mu)^2}^\infty dm^2 \frac{\sigma(m^2)}{\zeta - m^2}$$

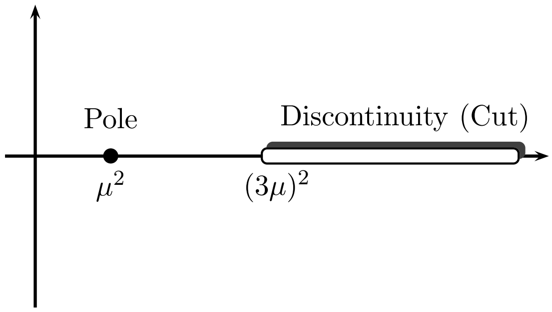


Figure 3.5: Analyticity domain of the full Feynman propagator of a pseudo scalar field

The domain of analyticity is displayed in Fig. 3.5 [17]

Apart from a pole at μ^2 and a cut staring at $(3\mu)^2$ along to positive real axis the function is analytic in the entire ζ–plane.

Let us look at the decomposition of $\tilde{G}(\zeta)$ into real and imaginary part: we write $\zeta = \xi + i\eta$ and use $\frac{1}{a \pm ib} = \frac{a \mp ib}{a^2+b^2}$ to obtain

$$\begin{aligned} \tilde{G}(\xi + i\eta) \quad &= \quad \frac{Z_3\,(\xi - \mu^2)}{(\xi - \mu^2)^2 + \eta^2} + \int\limits_{(3\mu)^2}^{\infty} dm^2 \frac{\sigma(m^2)\,(\xi - m^2)}{(\xi - m^2)^2 + \eta^2} \\ & \qquad -i\eta \left\{ \frac{Z_3}{(\xi - \mu^2)^2 + \eta^2} + \int\limits_{(3\mu)^2}^{\infty} dm^2 \frac{\sigma(m^2)}{(\xi - m^2)^2 + \eta^2} \right\} \end{aligned} \tag{3.60}$$

applying

$$\lim_{\varepsilon \to 0} \frac{\varepsilon}{x^2 + \varepsilon^2} = \pi \delta(x)$$

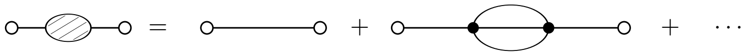
this explicitely shows that zeros only show up on the real positive axis. The discontinuity (jump) on the cut is

$$\tilde{G}(\xi + i\eta) - \tilde{G}(\xi - i\eta) = -2\pi i \left(Z_3\, \delta(\xi - \mu^2) + \Theta(\xi - (3\mu)^2)\, \sigma(\xi) \right) \tag{3.61}$$

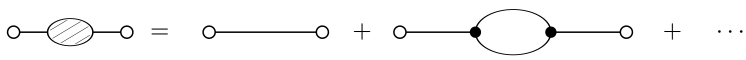
which means

$$\lim_{\eta \to +0} \operatorname{Im} \tilde{G}(\xi + i\eta) = -\pi \left(Z_3\, \delta(\xi - \mu^2) + \Theta(\xi - (3\mu)^2)\, \sigma(\xi) \right)$$

[17] The perturbation expansion with the P conserving interaction Lagrangian $\mathcal{L}_{int} = -\frac{g}{4!}\varphi^4$ invariant under $\varphi \to -\varphi$ starts as

$G^{(2)}$ = 

which exhibits the free $O(1)$ one–particle line and as an $O(g^2)$ contribution a diagram exhibiting a 3–particle cut starting at 3μ. In case a P–violating term $\Delta\mathcal{L}_{int} = -\frac{gv}{3!}\varphi^3$ is present which breaks the Z_2–symmetry $\varphi \to -\varphi$ the expected 2–particle cut diagram also is present with a cut starting at 2μ:

$G^{(2)}$ = 

and therefore

$$-\frac{1}{\pi}\int d\xi' \frac{\mathrm{Im}\,\tilde{G}(\xi')}{\xi-\xi'+i\varepsilon} = \frac{Z_3}{\xi-\mu^2+i\varepsilon} + \int\limits_{(3\mu)^2}^{\infty} d\xi' \frac{\sigma(\xi')}{\xi-\xi'+i\varepsilon} = \tilde{G}(\xi+i\varepsilon)\ .$$

Thus, utilizing (P = principal value)

$$\frac{1}{x\pm i\varepsilon} = \mathrm{P}\frac{1}{x} \mp i\pi\delta(x) \tag{3.62}$$

we arrive at

$$\begin{aligned}\tilde{G}(k^2) &= -\frac{1}{\pi}\int dk'^2 \frac{\mathrm{Im}\,\tilde{G}(k'^2)}{k^2-k'^2+i\varepsilon} \\ &= \frac{\mathrm{P}}{\pi}\int dk'^2 \frac{\mathrm{Im}\,\tilde{G}(k'^2)}{k'^2-k^2} + i\mathrm{Im}\,\tilde{G}(k^2)\end{aligned}$$

which leads us to the final result, the wanted **dispersion relation** (DR)

$$\boxed{\mathrm{Re}\,\tilde{G}(k^2) = \frac{1}{\pi}\mathrm{P}\!\!\int dk'^2 \frac{\mathrm{Im}\,\tilde{G}(k'^2)}{k'^2-k^2}} \tag{3.63}$$

which in mathematics is called a **Hilbert transformation**. It tells us that the analytic function $\tilde{G}(\zeta)$ is uniquely determined in terms of its imaginary part along the positive real axis, provided the dispersion integral converges. Often the integral in fact does not converge because the behavior of the integral at $k'^2 \to \infty$ does not fall off fast enough. We then have a typical ultraviolet problem and hence one ore more subtractions may be required. Suppose, e.g., the dispersion integral is logarithmically divergent. Then one subtraction cures the problem:

$$\begin{aligned}\mathrm{Re}\,\tilde{G}(k^2)-\tilde{G}(0) &= \frac{1}{\pi}\mathrm{P}\!\!\int dk'^2 \frac{\mathrm{Im}\,\tilde{G}(k'^2)}{k'^2-k^2} - \frac{1}{\pi}\mathrm{P}\!\!\int dk'^2 \frac{\mathrm{Im}\,\tilde{G}(k'^2)}{k'^2} \\ &= \frac{k^2}{\pi}\mathrm{P}\!\!\int dk'^2 \frac{\mathrm{Im}\,\tilde{G}(k'^2)}{k'^2\,(k'^2-k^2)}\end{aligned}$$

where the new integral is convergent now. The prize we have to pay is the free subtraction constant $\tilde{G}(0)$ which remains undetermined by the DR, it rather must be fixed by a renormalization condition.

3.6.3 Causality and analyticity

In the above derivation of the DR one important aspect, namely, the physics origin of analyticity has not been made very explicite. Here we would like to show that it actually derives from the classical **principle of cause and effect**. We consider a physical system

The principle of cause and effect may be put into precise mathematical form by the following requirements (and the consequences thereof):

[1) internal properties of system constant in time

[2) $g(t)$ causally dependent of $f(t')$

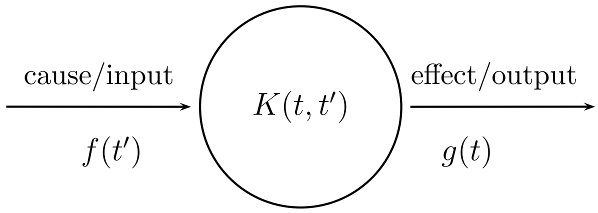


Figure 3.6: The principle of cause and effect.

[3) $g(t)$ linear functional of $f(t')$

which translates into:

- 1) $\mapsto K(t,t') = K(t-t')$
- 2) $\mapsto K(t-t') = 0$ for $t' > t$
- 3) $\mapsto g(t) = \int dt' K(t-t')\, f(t')$

The last assumption seems natural, as it allows for a general consideration. There is no theory of non-linear systems and only specific examples could be considered otherwise.

The consequences in Fourier space may be easily worked out: [1)] implies that K only depends on one variable $\tau = t - t'$ such that its Fourier transform reads $\tilde{K}(\omega) = \int K(\tau)\, e^{i\omega\tau}$. The convolution integral [3)] then translates into $\tilde{g}(\omega) = \tilde{K}(\omega)\, \tilde{f}(\omega)$. The crucial property in our context is [2)] which implies that ω may be analytically continued to complex values $\omega = \xi + i\eta$! The reason is that

$$\tilde{K}(\omega) = \int_{-\infty}^{+\infty} d\tau K(\tau)\, e^{i\omega\tau} = \int_{0}^{+\infty} d\tau K(\tau)\, e^{-\eta\tau} e^{i\xi\tau}$$

such that $\tilde{K}(\omega)$ is a regular analytic function in the upper half ω–plane.

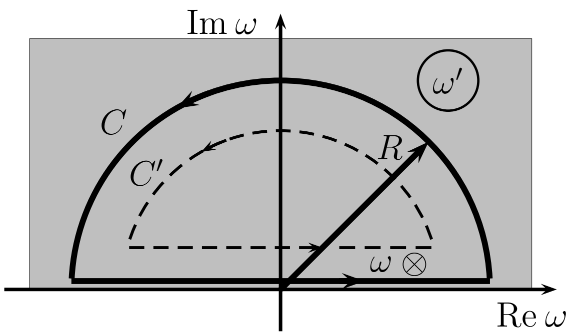


Figure 3.7: Analyticity domain and Cauchy contour for the causal transmission function $\tilde{K}(\omega)$

We then may apply Cauchy's theorem and write

$$\tilde{K}(\omega) = \frac{1}{2\pi i} \int_C d\omega' \frac{\tilde{K}(\omega')}{\omega' - \omega} + \frac{1}{2\pi i} \int_{C'} d\omega' \frac{\tilde{K}(\omega')}{\omega' - \omega}$$

with ω a point inside the contour C but outside the second contour C'. Thus the second integral does not contribute but it will play a role when we are going to take a limit $R,\ R' \to \infty$ with the paths parallel to the real axis at the same time approaching the real axis and thus squeezing ω to become real. Thats what we are going to do now. Suppose $\tilde{K}(\omega)$ falls off sufficiently fast at infinity[18] in the upper half plane $\mathrm{Im}\,\omega' > 0$, such that we may take the limit $R(R') \to \infty$ and the contributions from the half–circles are vanishing: then we arrive at

$$\tilde{K}(\omega) = \frac{1}{2\pi i}\int\limits_{-\infty}^{+\infty} d\omega' \frac{\tilde{K}(\omega')}{\omega' - \omega - i\varepsilon} + \frac{1}{2\pi i}\int\limits_{-\infty}^{+\infty} d\omega' \frac{\tilde{K}(\omega')}{\omega' - \omega + i\varepsilon}$$

with the $i\varepsilon$–prescription such that the path does not cross the ω–pole. As a result the physical (real ω) transfer function $\tilde{K}(\omega)$ is a boundary value

$$\boxed{\tilde{K}(\xi) = \lim_{\varepsilon \to +0} \tilde{K}(\xi + i\varepsilon)} \tag{3.64}$$

of a regular analytic function which is analytic in $\mathrm{Im}\,\omega > 0$. Applying now (3.62), which implies

$$\frac{1}{x - i\varepsilon} + \frac{1}{x + i\varepsilon} = 2\,\frac{P}{x}$$

we arrive at (ω real)

$$\tilde{K}(\omega) = -\frac{i}{\pi}\mathrm{P}\!\!\int\limits_{-\infty}^{+\infty} d\omega' \frac{\tilde{K}(\omega')}{\omega' - \omega}$$

which we may write as our main result as follows:

$$\boxed{\begin{aligned} \mathrm{Re}\,\tilde{K}(\omega) &= \frac{1}{\pi}\mathrm{P}\!\!\int\limits_{-\infty}^{+\infty} d\omega' \frac{\mathrm{Im}\,\tilde{K}(\omega')}{\omega' - \omega} \\ \mathrm{Im}\,\tilde{K}(\omega) &= -\frac{1}{\pi}\mathrm{P}\!\!\int\limits_{-\infty}^{+\infty} d\omega' \frac{\mathrm{Re}\,\tilde{K}(\omega')}{\omega' - \omega} \end{aligned}} \tag{3.65}$$

and hence the validity of the dispersion relations reflects the causality of a physical system. The relations (3.65) say that $\mathrm{Re}\,\tilde{K}(\omega)$ and $\mathrm{Im}\,\tilde{K}(\omega)$ are Hilbert transforms of each other.

The heuristic discussion presented above may be made mathematically rigurous and is known as Titcharsh's theorem. Losely speaking it states that the following three properties of $\tilde{K}(\omega)$ are equivalent:

(a) $\tilde{K}(\omega)$ obeys the DR (3.65)

(b) $\tilde{K}(\omega)$ has a Fourier transform $K(\tau)$ vanishing for $\tau < 0$

(c) $\tilde{K}(\omega)$ is holomorphic in $\mathrm{Re}\,\omega > 0$.

[18] i.e., we assume $\tilde{K}(\omega)$ to obey an un-subtracted DR, otherwise the appropriate number of subtractions have to be applied.

3.6.4 An example from classical optics

Let us consider a polarizable medium in a electromagnetic field. The function f here is the electric field, the transmission function K is the susceptibility and g is the polarization. We have the relation

$$P(\omega) = \varepsilon_0\,(\varepsilon(\omega) - 1)\,E(\omega) \equiv \tilde{K}(\omega)\,E(\omega)$$

which identifies the transmission function K. ε_0 is the permittivity of the vacuum, $\varepsilon(\omega) = n^2(\omega)$ the dielectric constant of the medium and n is the refraction index. Also here causality implies the validity of DR's. They read

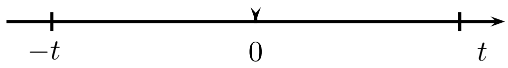
$$\begin{aligned}
\mathrm{Re}\, n_1^2(\omega) - 1 &= \frac{1}{\pi}\mathrm{P}\!\int_{-\infty}^{\infty} d\omega' \frac{\mathrm{Im}\, n_2^2(\omega')}{\omega' - \omega} \\
\mathrm{Im}\, n_2^2(\omega) &= -\frac{\omega}{\pi}\mathrm{P}\!\int_{-\infty}^{\infty} d\omega' \frac{\mathrm{Re}\, n_1^2(\omega') - 1}{\omega' - \omega}
\end{aligned}$$

which exhibits the causality of the polarizability of matter in an electric field.

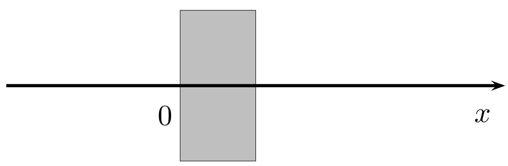
What causality means may be illustrated by the following reasoning about a light-flash in a dark room and the possibility to "see in the dark" before the light-flash was on:

Let us consider a light-flash at time $t = 0$

$$\delta(\tau) = \frac{1}{2\pi}\int d\omega e^{i\omega t}$$



all frequencies are contained as ever lasing waves in such a way that they cancel each other except at the instant $t = 0$. All the Fourier components are present before $t = 0$. With the help of colored glasses it should be possible to see the Fourier components at time -t, say. As a color filter we consider



A wave traveling along the x–axis has the form

$$\begin{aligned}
f(x,t) &= \frac{1}{2\pi}\int d\omega e^{-i\omega\,(t - n(\omega)\frac{x}{c})} \\
&= \frac{1}{2\pi}\int d\omega \left\{ e^{-\omega n_2(\omega)\frac{x}{c}} e^{-i\omega\,(t - n_1(\omega)\frac{x}{c})} \right\}
\end{aligned}$$

which is analytic in the upper half ω–plane. What we have to show:

$$f(x,t) = 0 \;\;\forall\;\; t < 0\,.$$

We may write

$$f(x,t) = \lim_{R\to\infty} \frac{1}{2\pi} \int_C d\omega e^{g(x,t;\omega)}$$

with a contour as displayed in Fig. 3.7 and

$$g(x,t;\omega) = -i\omega\,(t - n_1(\omega)\frac{x}{c}) - \omega n_2(\omega)\frac{x}{c}\ .$$

We note that $g_1 \equiv \mathrm{Re}\, g \to -\infty$ for $t < 0$. Writing $\omega = R\cos\phi + iR\sin\phi$ $(0 \leq \phi \leq \pi)$for the arc of the integration path C we have

$$g_1 = R\sin\phi\left(t - n_1(\omega)\frac{x}{c}\right) - R\cos\phi n_2(\omega)\frac{x}{c}\ .$$

For the DR the behavior of n_1 and n_2 for large R is crucial:

$$n_1(\omega) \sim 1 + O(R^{-2})\ \ ;\quad n_2(\omega) \sim O(R^{-1})$$

such that

$$g_1 \sim R\sin\phi\left(t - \frac{x}{c}\right) - \cos\phi\frac{x}{c}$$

and hence we have

$$\lim_{R\to\infty} g_1(x,t;\omega) = -\infty\ \ \forall\ \ t < \frac{x}{c}\ .$$

This result is even stronger than required. No signal can be obtained at speed faster than light. Because of causality glasses must have absorption $(n_2(\omega))$ in addition to dispersion (transmission) $(n_1(\omega))$.

3.6.5 Vector fields, vector currents

Let $A^\mu(x)$ be a contravariant vector field. We first want to find a representation for the two–point Wightman function $< 0|A_\mu(x)A_\nu(y)|0 >$. We may proceed as in the scalar case. The basic assumptions are the same.

We now consider the above mentioned two–point function and insert a complete set of intermediate states:

$$\begin{aligned} < 0|A^\mu(x)A^\nu(y)|0 > &= \sum_\alpha \frac{1}{(2\pi)^3}\int d^4p\ < 0|A^\mu(x)|p,\alpha\rangle\langle p,\alpha|A^\nu(y)|0 > \\ &= \sum_\alpha \frac{1}{(2\pi)^3}\int d^4p\ < 0|A^\mu(0)|p,\alpha\rangle\langle p,\alpha|A^\nu(0)|0 >\ e^{-ip(x-y)} \end{aligned}$$

where we used translational invariance.

We define the **spectral function**

$$\tilde{\rho}^{\mu\nu}(p) \doteq \sum_\alpha < 0|A^\mu(0)|p,\alpha\rangle\langle p,\alpha|A^\nu(0)|0 > \tag{3.66}$$

which has the properties

1) L–transformations: $\tilde{\rho}^{\rho\sigma}(p)$ is a second rank tensor

$$\tilde{\rho}^{\mu\nu\prime}(p) = \Lambda^{\mu}{}_{\rho}\, \Lambda^{\nu}{}_{\sigma}\, \tilde{\rho}^{\rho\sigma}(p)$$

and thus must have the form

$$\tilde{\rho}^{\mu\nu}(p) = -\tilde{\rho}_1\, g^{\mu\nu} \;+\; \tilde{\rho}_2\, p^\mu p^\nu$$

where $\tilde{\rho}_1$ and $\tilde{\rho}_2$ are scalar functions depending on p^2 only

2) spectral condition: $\tilde{\rho}(p) \neq 0 \Leftrightarrow p^2 \geq 0$, $\quad p^0 \geq 0$

3) current conservation/transversality: choose tensor coefficient functions $\tilde{\rho}_1$ and $\tilde{\rho}_2 = \tilde{\rho}_0 + \tilde{\rho}_1/p^2$ such that $\partial_\mu A^\mu = 0$ implies $\tilde{\rho}_0 = 0$

$$\boxed{\tilde{\rho}^{\mu\nu}(p) = \Theta(p^0) \left\{ \left[\frac{p^\mu p^\nu}{p^2} - g^{\mu\nu} \right] \tilde{\rho}_1(p^2) \;+\; p^\mu p^\nu \tilde{\rho}_0(p^2) \right\}} \tag{3.67}$$

As before we may write

$$\tilde{\rho}_i(p^2) = \int\limits_0^\infty dm^2\, \tilde{\rho}_i(m^2)\, \delta(p^2 - m^2) \;,$$

furthermore $p^\mu = i\partial_x^\mu$ when acting on $e^{-ip(x-y)}$ and we obtain (see 3.26)

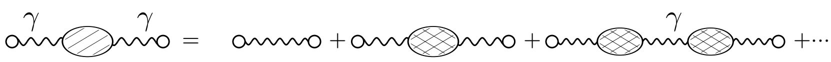
$$\begin{aligned} < 0|A^\mu(x)A^\nu(y)|0 > \;&=\; \frac{1}{(2\pi)^3} \int d^4p\, \Theta(p^0) \left\{ \left[\frac{p^\mu p^\nu}{p^2} - g^{\mu\nu} \right] \tilde{\rho}_1(p^2) \;+\; p^\mu p^\nu \tilde{\rho}_0(p^2) \right\}\, e^{-ip(x-y)} \\ &=\; -i \int\limits_0^\infty dm^2 \left\{ \tilde{\rho}_1(m^2) \left(g^{\mu\nu} + \frac{\partial^\mu \partial^\nu}{m^2} \right) \;+\; \tilde{\rho}_0(m^2)\, (\partial^\mu \partial^\nu) \right\} \Delta^+(x-y;m^2) \end{aligned}$$

and corresponding expressions follow for the commutator ($\Delta^+(x-y;m^2) \to \Delta(x-y;m^2)$) as well as for the Feynman propagator ($\Delta^+(x-y;m^2) \to \Delta^F(x-y;m^2)$).

3.6.6 Vacuum polarization, correlator of two electromagnetic currents

The fundamenta electromagnetic fine structure constant α in fact is a function of the energy scale $\alpha \to \alpha(E)$ of a process due to charge screening. The latter is a result of the fact that a naked charge is surrounded by a cloud of virtual particle–antiparticle pairs (dipoles) which line up in the field of the central charge and such lead to a vacuum polarization which screens the central charge. From long distances (classical charge) one thus sees less charge than if one comes closer, such that one seen an increasing charge with energy.

The vacuum polarization mainly affects the photon propagator. The full or dressed propagator is given by the geometrical progression of self–energy insertions $-i\Pi_\gamma(q^2)$ (Dyson summation) (for simplicity we consider the Feynman gauge and omit an overall metric tensor $g^{\mu\nu}$)



$$\begin{aligned} -i\, D_\gamma(q^2) \;&\equiv\; \frac{-i}{q^2} + \frac{-i}{q^2}\,(-i\Pi_\gamma)\,\frac{-i}{q^2} + \frac{-i}{q^2}\,(-i\Pi_\gamma)\,\frac{-i}{q^2}\,(-i\Pi_\gamma)\,\frac{-i}{q^2} + \cdots \\ &=\; \frac{-i}{q^2} \left\{ 1 + \left(\frac{-\Pi_\gamma}{q^2} \right) + \left(\frac{-\Pi_\gamma}{q^2} \right)^2 + \cdots \right\} = \frac{-i}{q^2} \left\{ \frac{1}{1 + \frac{\Pi_\gamma}{q^2}} \right\} = \frac{-i}{q^2 + \Pi_\gamma(q^2)} \end{aligned}$$

By $U(1)_{\rm em}$ gauge invariance the photon remains massless and hence we have $\Pi_\gamma(q^2) = \Pi_\gamma(0) + q^2\,\Pi'_\gamma(q^2)$ with $\Pi_\gamma(0) \equiv 0$. As a result we obtain

$$-i\,D^{\mu\nu}_\gamma(q^2) = -ig^{\mu\nu}\;D_\gamma(q^2) + \text{ gauge terms } = \frac{-ig^{\mu\nu}}{q^2\,(1+\Pi'_\gamma(q^2))} + \text{ gauge terms } .$$

Including a factor e^2 and considering the renormalized propagator (wave function renormalization factor Z) we have

$$e^2 D_{\mu\nu}(q) = \frac{g_{\mu\nu}\,e^2\,Z}{q^2\,\left(1+\Pi'_\gamma(q^2)\right)} + \text{ gauge terms}$$

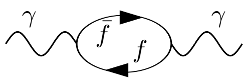
which in effect means that the charge has to be replaced by a **running charge**

$$e^2 \to e^2(q^2) = \frac{e^2 Z}{1+\Pi'_\gamma(q^2)} .$$

The wave function renormalization factor Z is fixed by the condition that at $q^2 \to 0$ one ontains the classical charge (charge renormalization in the Thomson limit). Thus the renormalized charge is

$$e^2 \to e^2(q^2) = \frac{e^2}{1+(\Pi'_\gamma(q^2)-\Pi'_\gamma(0))} \tag{3.68}$$

where the lowest order diagram in perturbation theory which contributes to $\Pi'_\gamma(q^2)$ is the following:



which describes virtual creation and reabsorption of fermion pairs $\gamma^* \to e^+e^-, \mu^+\mu^-, \tau^+\tau^-, u\bar{u}, d\bar{d}, \cdots \to\; \gamma^*$

In terms of the fine structure constant $\alpha = \frac{e^2}{4\pi}$ (3.68) reads

$$\alpha(q^2) = \frac{\alpha}{1-\Delta\alpha} \quad ; \quad \Delta\alpha = -\mathrm{Re}\,\left(\Pi'_\gamma(q^2) - \Pi'_\gamma(0)\right) .$$

The various contributions to the shift in the fine structure constant come from the leptons (lep $= e$, μ and τ) the 5 light quarks (u, b, s, c, and b and the corresponding hadrons = had) and from the top quark:

$$\Delta\alpha = \Delta\alpha_{\rm lep} + \Delta^{(5)}\alpha_{\rm had} + \Delta\alpha_{\rm top} + \cdots$$

Also W–pairs contribute at $q^2 > M_W^2$. While the other contributions can be calculated order by order in perturbation theory the hadronic contribution $\Delta^{(5)}\alpha_{\rm had}$ exhibits low energy strong interaction effects and hence cannot be calculated by perturbative means. Here the dispersion relations play a key role. We thus consider in the following the hadronic contribution

$\mathbf{\Pi'_{\rm had}(q^2)}$ =

1pi blob

The one particle irreducible (1pi) blob (diagrams which cannot be cut into two disconnected parts by cutting a single photon line) at low energies exhibits intermediate states like $\pi^0\gamma, \rho, \pi\pi, \pi\pi\gamma, \pi\pi Z$,

$\cdots, \pi\pi H, \cdots$ (at least one hadron plus any strong, electromagnetic or weak interaction contribution)

The vacuum expectation value of the time ordered product of two electromagnetic currents has the form given above for general vector fields

$$< 0|Tj^\mu_{\mathrm{em}}(x)j^\nu_{\mathrm{em}}(0)|0 > \;=\; -i\int\limits_0^\infty dm^2\,\tilde{\rho}(m^2)\left(m^2\,g^{\mu\nu}+\partial^\mu\partial^\nu\right)\,\Delta^F(x-y;m^2)\;, \tag{3.69}$$

however, due to vector current conservation $\partial_\mu j^\mu_{\mathrm{em}}(x) = 0$ only the transversal amplitude is present: thus $\tilde{\rho}_0 \equiv 0$ and we denote $\tilde{\rho}_1$ by $\tilde{\rho}$, simply. In Fourier space

$$\begin{aligned} i\int d^4x e^{iqx} < 0|Tj^\mu_{\mathrm{em}}(x)j^\nu_{\mathrm{em}}(0)|0 > &= -\left(q^2g^{\mu\nu}-q^\mu q^\nu\right)\,\Pi'_\gamma(q^2) \\ &= \int\limits_0^\infty dm^2\,\tilde{\rho}(m^2)\left(m^2\,g^{\mu\nu}-q^\mu q^\nu\right)\,\frac{1}{q^2-m^2+i\varepsilon} \end{aligned} \tag{3.70}$$

where $\Pi'_\gamma(q^2)$ defines the **photon vacuum polarization function**. By taking imaginary parts we have

$$\boxed{\mathrm{Im}\,\Pi'_\gamma(q^2) = \pi\,\tilde{\rho}(q^2)} \tag{3.71}$$

Again causality implies analyticity and the validity of a dispersion relation. In fact the electromagnetic current correlator exhibits a logarithmic UV singularity and thus requires one subtraction such that from (3.71) we find

$$\Pi'_\gamma(q^2)-\Pi'_\gamma(0) = \frac{q^2}{\pi}\int\limits_0^\infty ds\frac{\mathrm{Im}\,\Pi'_\gamma(s)}{s\,(s-q^2-i\varepsilon)} \tag{3.72}$$

Unitarity (3.19) implies the optical theorem (see Fig. 3.8), which tells us that the imaginary part of the photon propagator is proportional to the total cross section $\sigma_{\mathrm{tot}}(e^+e^- \to \gamma^* \to \text{anything})$ ("anything" means any possible state). The precise relationship reads

$$\mathrm{Im}\,\Pi'_\gamma(s) = \frac{1}{12\pi}R(s) \tag{3.73}$$

where

$$R(s) = \frac{\sigma_{\mathrm{tot}}}{\frac{4\pi\alpha^2}{3s}} \tag{3.74}$$

The shift of the fine structure constant α due to the vacuum polarization effects is thus given by

$$\Delta^{(5)}\alpha_{\mathrm{had}} = -\frac{\alpha s}{3\pi}\left(\,\mathrm{P}\!\!\int_{4m_\pi^2}^{E^2_{\mathrm{cut}}} ds'\frac{R^{\mathrm{data}}_\gamma(s')}{s'(s'-s)} + \mathrm{P}\!\!\int_{E^2_{\mathrm{cut}}}^\infty ds'\frac{R^{\mathrm{QCD}}_\gamma(s')}{s'(s'-s)}\right)$$

where

$$R_\gamma(s) \equiv \frac{\sigma(e^+e^- \to \gamma^* \to \text{hadrons})}{\sigma(e^+e^- \to \gamma^* \to \mu^+\mu^-)} = 12\pi\mathrm{Im}\Pi'_\gamma(s)\;.$$

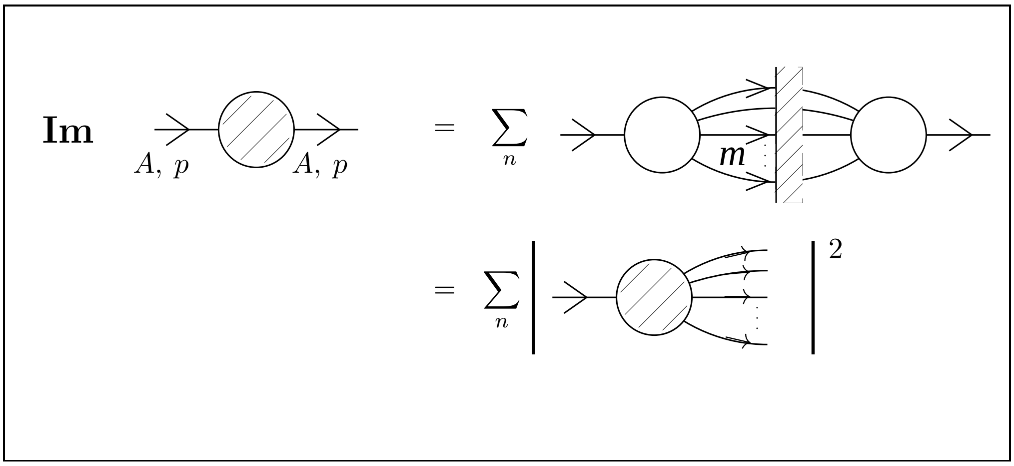


Figure 3.8: Optical theorem (3.19) for the photon propagator. The cut lines are on-shell particles $\frac{1}{q^2-m^2+i\varepsilon} \to -i\pi\delta(q^2-m^2)$ integrated out over phase space. It hence is proportional to $|T|^2$ and thus the the total cross–sextion $\sigma_{\text{tot}}(s)$

The optical theorem for the propagator is a special case of the more general version valid for scattering amplitudes depicted in Fig. 3.3.

4 Quantum Electrodynamics

Quarks and leptons, the constituents of matter, interact via gauge fields. Quantum electrodynamics (QED) describes the interaction of charged particles with the photon, which is described by an **Abelian gauge field** $A_\mu(x)$. As we shall see the form of the interaction may be understood as a consequence of a local symmetry: **local gauge invariance** (Weyl 1929). QED has been tested with extreme accuracy at its quantum effects level (Lamb shift, anomalous magnetic moments) and is the prototype of a very successful quantum field theory.

We assume the reader to be familiar with QED. Here we give a short account of its basic features only. The aim is to remind the reader of some basic problems which one encounters with massless spin 1 fields and how they are solved. Similar problems will show up in non-Abelian gauge theories which we will considered at a later stage.

We first consider the free photon field $A_\mu(x)$ and a free electron field $\psi_\alpha(x)$. The independent free Lagrangian densities read:

$$\begin{array}{rclcl} \mathcal{L}_{0A} & = & \mathcal{L}_{\text{free photon}} & = & -\frac{1}{4} : F_{\mu\nu}F^{\mu\nu} : \\ \mathcal{L}_{0\psi} & = & \mathcal{L}_{\text{free electron}} & = & : \bar{\psi}_\alpha \left(i\gamma^\mu \partial_\mu - m\right)_{\alpha\beta} \psi_\beta : \end{array}$$

The ": $\cdots$:" prescription means <u>normal ordering</u> (see Sec. 3.4.2) i.e. represent the fields in terms of annihilation and creation operators and commute (anticommute) all creation operators to the left of the annihilation operators. The c-number commutator (anticommutator) terms are omitted (subtracted)[19]. By this prescription we have subtracted the vacuum density such that now

$$< 0 \mid \mathcal{L}_0^{A,\psi}(x) \mid 0 >= 0 \ \ .$$

Notice that the action $i \int d^4x \mathcal{L}(x)$, for the infinite space-time volume, only may exist after subtraction of the vacuum density.

The photon field determines the antisymmetric electromagnetic field strength tensor

$$F_{\mu\nu}(x) = \partial_\mu A_\nu - \partial_\nu A_\mu \ \ , \tag{4.1}$$

which is gauge invariant – i.e. an Abelian gauge transformation

$$A_\mu \to A_\mu - \partial_\mu \alpha(x) \ \ , \tag{4.2}$$

where $\alpha(x)$ is an arbitrary scalar function, leaves $F_{\mu\nu}(x)$ invariant. As we know, if we represent $F_{\mu\nu}(x)$ as a curl of a vector-potential A_μ, Eq. (4.1), the homogeneous Maxwell equation

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad \text{ with } \ \tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \tag{4.3}$$

is automatically satisfied. The pseudo-tensor (parity odd) $\tilde{F}^{\mu\nu}$ is the dual of $F^{\mu\nu}$.

We now consider the free field equations.

- **Field equation for** $A^\mu(x)$

[19] For operators which are bilinear in the (free) fields, the normal ordering prescription is equivalent to the subtraction of the vacuum expectation value. For higher powers in the fields the relationship between ordinary products and normal products is discussed in Sec. 3.4.2.

The free photon is a solution of the free Maxwell equation, which is the Euler Lagrange equation for the free photon Lagrangian:

$$\partial_\mu \frac{\partial \mathcal{L}_{0A}}{\partial \partial_\mu A_\nu} = \frac{\partial \mathcal{L}_{0A}}{\partial A_\nu} \quad \Rightarrow \quad \partial_\mu F^{\mu\nu}(x) = 0$$

or, in terms of the vector-potential,

$$\Box A^\mu(x) - \partial^\mu \left(\partial_\nu A^\nu(x)\right) = \left(\Box g^{\mu\nu} - \partial^\mu \partial^\nu\right) \, A_\nu = 0$$

If we try to solve this latter equation a fundamental problem shows up. This equation of motion does not determine the field A_μ. The differential operator $\Box g^{\mu\nu} - \partial^\mu \partial^\nu$ has no inverse, since $\phi_\mu = \partial_\mu \alpha(x)$ is a solution with eigenvalue zero.

There is a simple reason for the problem. The field A^μ is not an observable and therefore has unphysical properties. A^μ is supposed to describe a massless spin 1 particle, which has two physical degrees of freedom only, the two transversely polarized states. Therefore two components of A^μ must be redundant. In particular, A^μ has a scalar component $\partial_\mu A^\mu = \phi$ which cannot be physical and must be required to vanish or to decouple from the physical degrees of freedom.

An idea of how to cure the problem we get if we notice that

$$\mathcal{L}_{0A} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

is a degenerate quadratic form in A^μ, which means that a change of A^μ does not necessarily change $\mathcal{L}_{0A}$. In particular we know that a gauge transformation of A^μ leaves $\mathcal{L}_{0A}$ unchanged. Obviously, in order to obtain an equation of motion which determines A^μ uniquely we have to break this degeneracy. This forces us to **add a gauge dependent term** to the invariant Lagrangian. Doing so, we fix a particular gauge and loose manifest gauge invariance.

In order to get an idea of what kind of term we may add in order to break the gauge symmetry of the Lagrangian without affecting the physics, let us consider the the problem on the level of the field equation. This is a second order linear partial differential equation. Since it is not sufficient to determine A^μ uniquely we need some supplementary condition $C(A) = 0$. The latter should be linear, in order to keep the problem linear, and covariant, both requirements are not mandatory, however. A possibility, actually the only covariant and linear choice, is the use of the Lorentz gauge condition

$$\partial_\mu A^\mu(x) = 0$$

as a subsidiary condition. Strictly speaking this condition does not determine uniquely a gauge, because we still can perform a gauge transformation

$$\partial_\mu A^\mu(x) = 0 \to \partial_\mu A^\mu(x) - \Box \alpha(x) = 0 \quad \text{if} \quad \Box \alpha(x) = 0$$

which respects the Lorentz condition if we choose a restricted class of gauge functions $\alpha(x)$ which are solutions of $\Box \alpha(x) = 0$. In practice we need not bother about this problem further because we will see that we get a well defined perturbation expansion if we use the Lorentz condition for gauge fixing.

Geometrically the gauge condition picks a hyper-surface in the A^μ field space. Each point on the hyper-surface corresponds to a physically distinct field. Gauge transformations move the field orthogonal to the surface. Fields connected by gauge transformations are called **gauge copies** of each other. They are physically equivalent and form so called **gauge orbits**. The gauge condition selects the cut point of the gauge orbit with the hyper-surface as one particular representative field from each gauge orbit.

The gauge condition $C(A) = 0$ may be imposed by adding a Lagrange-multiplier term $\lambda \frac{1}{2} C(A)^2$ to $\mathcal{L}_{0A}$ where λ is an arbitrary constant. We obtain the gauge dependent Lagrangian

$$\mathcal{L}_{0A}^{\xi} = \mathcal{L}_{0A} + \mathcal{L}_{GF} \;\; ; \quad \mathcal{L}_{GF} = -\frac{1}{2\xi} \left(\partial_\mu A^\mu(x) \right)^2 \;\; ,$$

where $\mathcal{L}_{GF}$ is the **gauge fixing term** which lifts the degeneracy of $\mathcal{L}_{0A}$. ξ is called the **gauge parameter** and ξ^{-1} corresponds to the Lagrange multiplier. If we can show that physical predictions, like scattering matrix-elements, for example, are independent of ξ we also have shown that $\partial_\mu A^\mu = 0$ for what concerns physics. The gauge fixed Lagrangian $\mathcal{L}_{0A}^{\xi}$ is now suitable as a starting point for the quantization of the vector-potential while the original invariant Lagrangian was not.

The modified equation of motion following from the Lagrangian

$$\mathcal{L}_{0A}^{\xi} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} \xi^{-1} \left(\partial_\mu A^\mu(x) \right)^2$$

is

$$\left(\Box g^{\mu\nu} - (1 - \xi^{-1}) \partial^\mu \partial^\nu \right) A_\nu(x) = 0$$

obtained by adding the extra term

$$\partial_\mu \frac{\partial \mathcal{L}_{GF}}{\partial \partial_\mu A_\nu} = -\xi^{-1} \partial^\nu \left(\partial_\mu A^\mu(x) \right)$$

to the previous form of the equation of motion.

Now the free Maxwell equation

$$\partial_\mu F^{\mu\nu} = \xi^{-1} \partial^\nu \left(\partial_\rho A^\rho(x) \right) \neq 0 \quad !$$

has no longer its classical form unless $\partial_\mu A^\mu(x) = 0$ in some sense. This is in contradistinction to the Proca field (massive spin 1 field) for which the Proca equation

$$\left((\Box + m^2) g^{\mu\nu} - \partial_\mu \partial_\nu \right) A_\nu = 0$$

automatically implies $\partial_\mu A^\mu(x) \equiv 0$.

Notice that the "vacuum" $|0>$ also makes troubles. Under a gauge transformation

$$\begin{aligned} <0 \mid A_\mu(x) \mid 0> \;\; \to \;\; & <0 \mid A_\mu(x) \mid 0> - \partial_\mu \alpha(x) <0 \mid 0> \\ & = <0 \mid A_\mu(x) \mid 0> - \partial_\mu \alpha(x) \end{aligned}$$

which would be a contradiction if $A_\mu(x)$ and $\mid 0>$ are supposed to have the naive properties, we except them to have. It turns out that the "vacuum" of the photon states must be considered as an equivalence class $\{\mid 0>\}_{A^\mu}$ and each gauge representative of a gauge orbit has a different formal vacuum.

The only clean covariant way to treat the problem is the Gupta-Bleuler formalism. In this formalism one can show that the physical state space $\mathcal{H}_{\text{phys}}$ is characterized by

$$\partial_\mu A^{\mu(+)}(x)\mathcal{H}_{\text{phys}} = 0$$

where $A^{\mu(+)}(x)$ denotes the positive frequency part (annihilation term) of the covariant photon field. The formal Hilbert space, before imposing this Gupta-Bleuler condition, includes unphysical states. Only on the physical subspace the physical laws have the "classical" form. For example, the Maxwell equation

$$\partial_\mu F^{\mu\nu} = -e j^\nu_{em} \tag{4.4}$$

is only true in the sense

$$\partial_\mu F^{\mu\nu(+)}\mathcal{H}_{\text{phys}} = -e j^{\nu(+)}_{em}\mathcal{H}_{\text{phys}} \ .$$

The reason why we need not worry to much about these problems is the fact that we have a simple check of whether or not the physics is gauge invariant: Physical matrix elements must turn out to be independent of the gauge parameters ξ ! In this case $\mathcal{L}_{GF}$ does not affect physical predictions and it looks as if $\partial_\mu A^\mu(x) = 0$. Thus gauge invariance is the "instrument" which allows to single out physics from technical artifacts.

- **Field equation for** $\psi_\alpha(x)$

The Euler-Lagrange equation for the free electron Lagrangian is the Dirac equation:

$$\partial_\mu \frac{\partial \mathcal{L}_{0\psi}}{\partial \partial_\mu \psi} = \frac{\partial \mathcal{L}_{0\psi}}{\partial \psi} \quad \Rightarrow \quad (i\gamma^\mu \partial_\mu - m)\,\psi(x) = 0 \ .$$

The electromagnetic current of the electron can be constructed from $\mathcal{L}_{0\psi}$ as follows:

$\mathcal{L}_{0\psi}$ has a **global** $U(1)$ symmetry: $\psi \to e^{-i\alpha}\psi$ where α is an arbitrary **constant**. Then by the **Noether theorem** there exists a conserved current.

$$\delta \mathcal{L}_{0\psi} \;=\; 0 \quad \text{under} \quad \begin{cases} \psi \to \psi + \delta\psi & ; \quad \delta\psi = -i\alpha\psi \\ \bar{\psi} \to \bar{\psi} + \delta\bar{\psi} & ; \quad \delta\bar{\psi} = \ \ i\alpha\bar{\psi} \end{cases}$$

where

$$\delta \mathcal{L}_{0\psi} = \delta\bar{\psi}\,\frac{\partial \mathcal{L}}{\partial \bar{\psi}} + \delta(\partial_\mu \bar{\psi})\,\frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} + \frac{\partial \mathcal{L}}{\partial \psi}\,\delta\psi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\,\delta(\partial_\mu \psi) \ .$$

For **global** transformations we have $\delta\partial_\mu\psi = \partial_\mu\delta\psi$ and the equation of motion tells us that

$$\frac{\partial \mathcal{L}}{\partial \psi} = \partial_\mu \frac{\partial \mathcal{L}}{\partial_\mu \psi} \ ; \quad \frac{\partial \mathcal{L}}{\partial \bar{\psi}} = \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \bar{\psi}}$$

and hence

$$\begin{aligned} \delta \mathcal{L}_{0\psi} \;&=\; \partial_\mu \left(\delta\bar{\psi} \frac{\partial \mathcal{L}}{\partial \partial_\mu \bar{\psi}} + \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} \delta\psi \right) \\ &=\; -i\alpha \partial_\mu \left(\bar{\psi} i \gamma^\mu \psi \right) = \alpha \partial_\mu j^\mu_{em} = 0 \end{aligned}$$

with

$$j^\mu_{em} \;=\; :\bar{\psi}_\alpha(\gamma^\mu)_{\alpha\beta}\psi_\beta : \tag{4.5}$$

the conserved electromagnetic current.

Notice: For all interactions which do not exhibit derivations of ψ the form of j^μ_{em} is determined solely by $\mathcal{L}_{0\psi}$

- **Local gauge invariance and the electromagnetic interaction**

We observed that the problems with the photon field $A_\mu(x)$ requires local gauge invariance to hold. For the free electron we have another problem which at first seems not to be related to the problem of quantization of massless spin 1 particles: We know that in quantum mechanics the phase of a wave function is not observable. For a Dirac field we would expect therefore invariance under local phase transformations

$$\psi(x) \to e^{-ie\alpha(x)}\psi(x) \quad . \tag{4.6}$$

These transformations, which again are related to some redundancy in the description of a particle by a quantum field, correspond to the local gauge transformations discussed before for the photon field and therefore are also called gauge transformations. However, $\mathcal{L}_{0\psi}$ is not locally gauge invariant because

$$\partial_\mu\psi(x) \to e^{-ie\alpha(x)}\partial_\mu\psi(x) - ie\; e^{-ie\alpha(x)}\psi(x)\partial_\mu\alpha(x)$$

and thus

$$\begin{aligned}\delta\mathcal{L}_{0\psi} &= e\,\bar{\psi}(x)i\gamma^\mu\left(-i\partial_\mu\alpha(x)\right)\psi(x) \\ &= e\,\bar{\psi}(x)\gamma^\mu\psi\partial_\mu\alpha(x) = e\,j^\mu_{em}\partial_\mu\alpha(x) \quad .\end{aligned}$$

A free electron cannot be described in a locally gauge invariant way! The requirement of local gauge invariance implies that "electrons must couple to photons via minimal substitution". Which means that we have to replace the troublesome derivative

$$\partial_\mu\psi \to D_\mu\psi$$

by a **covariant derivative** $D_\mu\psi$ defined in such a way that it transforms in the same way as ψ:

$$D_\mu\psi \to e^{-ie\alpha(x)}D_\mu\psi \tag{4.7}$$

under a local gauge transformation of the electron-photon system[20]

$$\begin{aligned}\psi(x) &\to e^{-ie\alpha(x)}\psi(x) \\ A_\mu(x) &\to A_\mu(x) - \partial_\mu\alpha(x) \quad .\end{aligned} \tag{4.8}$$

[20] Note that (4.8) must be a **simultaneous** transformation of the electron and the photon field with identical local gauge function $\alpha(x)$. With separate local functions (4.7) [in conjunction with (4.9)] does not hold and the QED Lagrangian is obviously not manifestly invariant. However, the apparent non–invariance is not real, because once the photon couples to the electron the non-invariant terms can always be reabsorbed by a gauge transformation of the photon field which cannot (or should not) affect the physics.

The requirement (4.7) implies the form

$$D_\mu = \partial_\mu - ieA_\mu \tag{4.9}$$

for the covariant derivative, which thus may be obtained by the so called **minimal substitution**

$$\partial_\mu \to D_\mu = \partial_\mu - ieA_\mu \tag{4.10}$$

to be applied to the free electron Lagrangian. As a consequence

$$\mathcal{L}_{0\psi} \to \mathcal{L}_\psi = \mathcal{L}_{0\psi} + ej^\mu_{\mathrm{em}} A_\mu(x) = \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi$$

includes automatically a certain type of electron-photon interactions. Thus the principle of local gauge invariance implies the following specific form of the interaction:

$$\mathcal{L}_{\mathrm{int}} = ej^\mu_{\mathrm{em}} A_\mu(x). \tag{4.11}$$

Obviously $F^{\mu\nu}$ and j^μ are gauge invariant objects. Thus one easily checks:

$$\begin{aligned} & \delta\left(\mathcal{L}_{0\psi} + ej^\mu_{em} A_\mu(x)\right) \\ = \;& \delta\mathcal{L}_{0\psi} - ej^\mu_{em}\partial_\mu\alpha(x) \\ = \;& ej^\mu_{\mathrm{em}}\partial_\mu\alpha(x) - ej^\mu_{\mathrm{em}}\partial_\mu\alpha(x) = 0. \end{aligned}$$

We notice that the form (4.11) of the coupling of photons to the electromagnetic current, which is prescribed by local gauge invariance, is the crux why it is **mandatory** to describe the photon, interacting with the charged particles, by the gauge dependent **four–potential** A_μ (**gauge potential**). Note that the two physical photon states with the fixed helicities ± 1 transform separately as an irreducible representation of the Lorentz group (see Appendix C). The reducible representation combining the two photon states into one field would have two independent components. By the above construction, however, we are forced to describe the photon by a four component field, which necessarily has two superfluous components. This causes a lot of technical complications part of which have been addressed above when discussing the free photon field $A_\mu(x)$ (and the need for gauge fixing) and which will be discussed further below in Sec. 4.1. For a more detailed discussion of the properties of the photon field we refer to Appendices A and C.

The result of our discussion may be summarized as follows:

> Local $U(1)$ gauge invariance implies electron-photon interaction according to minimal coupling. The electromagnetic interaction is described by
>
> $$\begin{aligned} \mathcal{L}_{\mathrm{QED}} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\xi^{-1}\left(\partial_\mu A^\mu\right)^2 + \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi \\ &= \mathcal{L}^\xi_{0A} + \mathcal{L}_{0\psi} + ej^\mu_{em}(x)A_\mu(x) \\ \mathcal{L}_{\mathrm{int}} &= ej^\mu_{em}(x)A_\mu(x) \;. \end{aligned} \tag{4.12}$$
>
> Correspondingly, the field equations for QED read
>
> $$\left\| \begin{array}{l} \left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = -e : A_\mu(x)\gamma^\mu\psi(x) : \\ \left(\Box g^{\mu\nu} - \left(1-\xi^{-1}\right)\partial^\mu\partial^\nu\right)A_\nu(x) = -e : \bar{\psi}(x)\gamma^\mu\psi(x) : \end{array} \right. \tag{4.13}$$

The minimally coupled electron-photon system has more symmetry than the free electron system, namely, a local one instead of a global one only. Due to the particular form of the interaction, resulting from the minimal substitution, the global gauge symmetry is promoted to a local gauge symmetry.

Let us come back to the simultaneous local gauge transformation (4.8) under which the classical QED Lagrangian (i.e., discarding the gauge fixing term in first place) is manifestly invariant: it is important to note that we may relax from the manifest invariance requirement. If we transform the electron field only, for example, the non-invariant term which shows up may be always eliminated by a gauge transformation of the photon field. The latter is required to leave the physics unchanged, which has to be proven of course. Thus the formal non–invariance of the Lagrangian not necessarily implies the non–invariance of the physics. We may thus precise the meaning of (4.8) as follows: *to each local gauge transformation of the electron field* $\psi(x) \to e^{-ie\alpha(x)}\psi(x)$ *there exists a gauge transformation of the photon field, namely,* $A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x)$ *such that under the combined transformation the classical part of the Lagrangian is manifestly invariant.*

As we shall see nature frequently makes use of the possibility that particles (the electrons) conspire with other particles (the photon) in order to enhance the symmetry. In this sense local gauge symmetries are **conspirative symmetries** which are only possible by conspiracy of particles of different kind.

Empirical fact: Nature makes use of the principle of local gauge invariance which is similar to the equivalence principle known from general relativity. Known fundamental elementary particle interactions are minimal couplings with respect to an invariance principle:

interaction	gauge group	quantum numbers
QED	$U(1)_{em}$	electric charge
QFD	$SU(2)_L \otimes U(1)_Y \overset{\text{broken}}{\to} U(1)_{em}$	weak isospin and weak hypercharge
QCD	$SU(3)_c$	color

Quantum Flavordynamics QFD we call the combination of the **electroweak theory**, also called electroweak **Standard Model**, and QCD, the theory of strong interactions which describes the strong forces between nucleons and other hadrons. In the sections to follow the basic ideas behind the construction of these theories will be developed.

4.1 Covariant photon field, Gupta-Bleuler formalism

Let us summarize once more in brief the origin of the problems with non–physical properties of the photon field and the cure of them by gauge invariance. The photon field $A^\mu(x)$ (see Appendix A.7) and the related polarization vectors $\varepsilon_\pm(p)$ are four component quantities which, however, are assumed to encode two physical degrees of freedom only. Obviously one needs two subsidiary conditions or something equivalent. Since at the same time we require relativistic covariance and linearity we have a problem:

i) the only covariant and linear condition available is

$$\partial_\mu A^\mu(x) = 0 \;\text{ or equivalently }\; p_\mu \varepsilon^\mu_\pm(p) = 0$$

In fact the standard photon field given in Appendix A.7 is not covariant (see Appendix C.2) and one could choose as well a non–covariant gauge, e.g., the Coulomb gauge.
The part of the missing "second covariant subsidiary condition" is taken by

ii) the requirement of <u>gauge invariance</u> of physical observables

$$A^\mu(x) \to A^\mu(x) - \partial^\mu \alpha(x)$$

or equivalently

$$\varepsilon^\mu_\pm(p) \to \varepsilon^\mu_\pm(p) = \lambda p^\mu \quad .$$

must be satisfied in any case.

The missing covariance of the formalism in calculations beyond the tree level cannot be accepted for technical reasons. Not only for practical calculations of higher order corrections, also for formal proofs of properties to all orders in the perturbation expansion a manifestly covariant formalism is indispensable. Such a formalism has been developed by Gupta and Bleuler long time ago and we are going to outline its basic elements in the following.

In order to get a clearer idea about the problems mentioned above in connection with the photon field, we will elaborate here in more detail about the difficulties we have encountered. As already mentioned, the problem is due to the fact that the following catalog of requirements cannot be fulfilled simultaneously in a manifest manner:

1. the photon is described by a four-vector field, the vector-potential, as required to be able to write down a local coupling with the electromagnetic current
2. Lorentz covariance
3. locality
4. positivity of the state space (i.e. Fock space for the physical photons)
5. uniqueness of the vacuum

Since the vector potential is not a measurable physical quantity we may consider this to be a purely technical problem. In fact only "physics" i.e. the observables must satisfy the physical principles. The latter are simply not realized manifestly in the formalism. However, the technical feasibility of controlling physical properties a posteriori to any order in perturbation theory force us to utilize a formalism which is manifestly covariant and local. We thus insist in preserving manifestly the first three points listed above, and then have to accept to work in an unphysical state space $\mathcal{H}_{GB}$, the Gupta-Bleuler state space. The space $\mathcal{H}_{GB}$ necessarily has indefinite metric. One then can prove that $\mathcal{H}_{GB}$ has a subspace $\mathcal{H}_P$ of states with positive definite norm, which may be characterized in a explicitly L-invariant manner. The physical state space $\mathcal{H}_{\rm phys}$ of the physical photons will then be defined by the equivalence classes (gauge orbits) of $\mathcal{H}_P$.

Construction of $\mathcal{H}_{GB}$

The covariance problem related with the use of the vector-potential may be solved formally, when one considers $A^\mu(x)$ as a field of four degrees of freedom. This means that one has to introduce unphysical states, the Gupta-Bleuler ghosts (GB–ghosts), and corresponding creation and annihilation operators in addition to the two physical photons.

The polarization vectors may be chosen as follows: Let p be a positive light–like vector, $p^2 = 0$, $p^0 > 0$. Then the polarization vectors $\varepsilon_\pm(p)$ of the physical photons are given by two orthogonal space–like vectors orthogonal to p:

$$\varepsilon_r^\mu(p)\,\varepsilon^*_{\mu\,r'}(p) = -\delta_{rr'} \quad , \quad p_\mu \varepsilon_r^\mu(p) = 0 \ .$$

The ε_r's may be represented in the complex helicity base, $r = (+,-)$, or in a real Cartesian base, $r = (1,2)$. Given p and the ε_r's they determine uniquely a vector $\hat{p}$ with the properties:

$$\hat{p}^2 = 0 \quad , \quad \hat{p}p = 1 \quad , \quad \hat{p}_\mu \varepsilon_r^\mu(p) = 0 \ .$$

We now introduce two additional unphysical "polarization vectors"

$$\begin{aligned} \varepsilon_0^\mu(p) &= \frac{\alpha p^\mu + \alpha^{-1}\hat{p}^\mu}{\sqrt{2}} \\ \varepsilon_3^\mu(p) &= \frac{\alpha p^\mu - \alpha^{-1}\hat{p}^\mu}{\sqrt{2}} \end{aligned}$$

with $\alpha > 0$ arbitrary, which together with $\varepsilon_r^\mu(p)$ $(r = (+,-) \sim (1,2))$ form an orthogonal set of four-vectors:

$$\varepsilon_0^\mu\varepsilon_0^\nu - \varepsilon_1^\mu\varepsilon_1^\nu - \varepsilon_2^\mu\varepsilon_2^\nu - \varepsilon_3^\mu\varepsilon_3^\nu = g^{\mu\nu} = p^\mu\hat{p}^\nu + \hat{p}^\mu p^\nu - \varepsilon_+^\mu\varepsilon_+^\nu - \varepsilon_-^\mu\varepsilon_-^\nu$$

The polarization vectors given here stand for a particular representative out of an equivalence class of polarization vectors obtained by applying Lorentz transformations from G_p, the little group of p, to the specific ones chosen here.

A covariant photon field is then given by

$$A^\mu(x) = \int d\mu(p) \left\{ \left[\sum_\pm \varepsilon_\pm^\mu(p) a(\vec{p},\pm) + \hat{p}^\mu b(\vec{p}\,) + p^\mu c(\vec{p}\,) \right] e^{-ipx} + \text{h.c.} \right\} \tag{4.14}$$

which satisfies $\Box A^\mu(x) = 0$. The annihilation operators may be represented by

$$a(\vec{p},\pm) = -i\varepsilon_\pm^{\mu *}(p) \int d^3x\, e^{ipx} \overleftrightarrow{\partial}_0 A_\mu(x) \tag{4.15}$$

$$b(\vec{p}\,) = ip^\mu \int d^3x\, e^{ipx} \overleftrightarrow{\partial}_0 A_\mu(x) \tag{4.16}$$

$$c(\vec{p}\,) = i\hat{p}^\mu \int d^3x\, e^{ipx} \overleftrightarrow{\partial}_0 A_\mu(x) \ . \tag{4.17}$$

Now we require the photon field to be local and covariant, which means, in particular, that the field commutator must have the form

$$[A_\mu(x), A_\nu(y)] = -i\left(-g_{\mu\nu} + 2\beta\partial_\mu\partial_\nu\right)\, \Delta(x-y;0) \tag{4.18}$$

with β an arbitrary gauge parameter. This form determines a family of covariant gauges. If we insert the above ansatz for $A_\mu(x)$ into the last equation we obtain the following commutation relations for the creation and annihilation operators:

$$\begin{aligned} [a(\vec{p},\lambda), a^+(\vec{p}\,',\lambda')] &= (2\pi)^3\, 2p^0\delta^{(3)}(\vec{p}-\vec{p}\,')\,\delta_{\lambda\lambda'} \\ [b(\vec{p}\,), c^+(\vec{p}\,')] = [c(\vec{p}\,), b^+(\vec{p}\,')] &= -(2\pi)^3\, 2p^0\delta^{(3)}(\vec{p}-\vec{p}\,') \\ [c(\vec{p}\,), c^+(\vec{p}\,')] &= -(2\pi)^3\, 2p^0\delta^{(3)}(\vec{p}-\vec{p}\,')\, 2\beta \end{aligned} \tag{4.19}$$

with $p^0 = |\vec{p}|$ and all other commutators vanishing. Obviously this algebra is manifestly covariant.

Translation invariance of $A^\mu(x)$ means that

$$[P^\mu, A^\nu(x)] = -i\partial^\mu A^\nu(x) \ ,$$

such that

$$\begin{aligned}
[P^\mu, a^\pm(\vec{p}, \pm)] &= \pm p^\mu a^\pm(\vec{p}, \pm) \\
[P^\mu, b^\pm(\vec{p}\,)] &= \pm p^\mu b^\pm(\vec{p}\,) \\
[P^\mu, c^\pm(\vec{p}\,)] &= \pm p^\mu c^\pm(\vec{p}\,) \ ,
\end{aligned}$$

which tells us a, b and c destroy, a^+, b^+ and c^+ create quanta of four-momentum p^μ. With the help of the commutation relations one verifies that

$$P^\mu = \int d\mu(p) p^\mu \left\{ \sum_\pm a^+(\vec{p}, \pm)\, a(\vec{p}, \pm) - b^+(\vec{p}\,)\, c(\vec{p}\,) - c^+(\vec{p}\,)\, b(\vec{p}\,) + 2\beta\, b^+(\vec{p}\,)\, b(\vec{p}\,) \right\}$$

must hold if we require that for all $\vec{p}$:

$$\begin{aligned}
a(\vec{p}, \pm)|0> = b(\vec{p}\,)|0> = c(\vec{p}\,)|0> = 0 \\
< 0|0> = 1 \ \text{ and } \ P^\mu|0> = 0 \ .
\end{aligned}$$

The space $\mathcal{H}_{GB}$ is now defined as the linear space of states which may be created by repeated application of the creation operators $a^+(\vec{p}, \pm)$, $b^+(\vec{p}\,)$ and $c^+(\vec{p}\,)$ to the vacuum $|0>$.

One easily verifies the following properties of the states in $\mathcal{H}_{GB}$:

a) states which are obtained by application of $a^+(\vec{p}, \pm)$ to the vacuum have positive scalar products and contain transversal photons, but not necessarily exclusively such.

b) if a state contains in addition one or more b–quanta, then it has a vanishing scalar product with all states which do not contain c–quanta.

c) if a state contains in addition c–quanta, then is has indefinite scalar products.

These statements directly follow from the commutation relations and $< 0|0 >= 1$.

Restriction to $\mathcal{H}_P$

The condition

$$b(\vec{p}\,)\ \mathcal{H}_P = 0 \quad \forall \vec{p} \tag{4.20}$$

defines a linear subspace of $\mathcal{H}_{GB}$, which by virtue of the commutation relation $[b(\vec{p}\,), c^+(\vec{p}\,')] = -(2\pi)^3\ 2p^0\delta^{(3)}(\vec{p} - \vec{p}\,')$ is characterized as the subspace of states which do not contain any c–quanta. Therefore the scalar product on $\mathcal{H}_P$ is positive semi-definite. States from $\mathcal{H}_P$ consist of linear combinations of vectors which apart from photons exhibit none, one or more b–quanta. Only the part without b–quanta contributes to the scalar product.

The crucial point of this construction is that the operator $b(\vec{p})$ determined by Eq. (4.17) in terms of $A^\mu(x)$ for given $\vec{p}$ is determined uniquely, in contrast to $a(\vec{p}, \pm)$ and $c(\vec{p}\,)$, which represent equivalence classes.

We may rewrite Eq. (4.17) as follows:

$$\begin{aligned} b(\vec{p}\,) &= \int d^3x\, \left(\partial_\mu e^{ipx}\right) \overleftrightarrow{\partial_0}\, A^\mu(x) \\ &= \int d^3x\, \partial_\mu \left(e^{ipx} \overleftrightarrow{\partial_0}\, A^\mu(x)\right) \\ &- \int d^3x\, e^{ipx} \overleftrightarrow{\partial_0}\, (\partial_\mu A^\mu(x)) \ . \end{aligned}$$

The first term of the last equality vanishes since if we write it as

$$\int d^3x\, \partial_\mu \left(e^{ipx} \overleftrightarrow{\partial_0}\, A^\mu(x)\right) = \partial_0 \left(\int d^3x\, e^{ipx} \overleftrightarrow{\partial_0}\, A^0(x)\right) + \int d^3x\, \partial_i \left(e^{ipx} \overleftrightarrow{\partial_0}\, A^i(x)\right) = 0$$

we observe that both terms vanish. The first one due to the fact that the integral is time-independent. This can be inferred by partial integration and using the Klein-Gordon equations $\Box A^\mu = 0$ and $\Box e^{ipx} = 0$. The second term is zero because it is a spatial integral over a divergence. Note that the plane wave solution e^{ipx} of the Klein-Gordon equation is always thought to stand for a "wave packet" solution $f_p(x)$ i.e. a smooth function which is strongly decreasing towards spatial infinity.

As a result we have

$$b(\vec{p}\,) \;=\; -\int d^3x\, e^{ipx} \overleftrightarrow{\partial_0}\, \varphi(x) \tag{4.21}$$

with

$$\varphi(x) = \partial_\mu A^\mu(x) \quad \text{and} \quad \Box\varphi(x) = 0 \ .$$

We conclude that the b–quanta are massless spin 0 bosons, since φ describes the scalar part of $A^\mu(x)$. It will turn out that $\varphi \equiv \partial_\mu A^\mu$ remains a <u>free field</u> in the interacting theory i.e. in QED.

By the construction of $A^\mu(x)$ in terms of the creation and annihilation operators Eq. (4.14) we have

$$\partial_\mu A^\mu(x) = \int d\mu(p) \left(b(\vec{p}\,)e^{-ipx} - \text{h.c.}\right)$$

since $p_\mu \varepsilon^\mu_\pm = 0$, $p_\mu p^\mu = 0$ and $p_\mu \hat{p}^\mu = 1$. The positive frequency part is thus given by

$$\Big\| \; \partial_\mu A^{\mu(+)}(x) = \textstyle\int d\mu(p) b(\vec{p}\,)e^{-ipx} \tag{4.22}$$

and we may write the Gupta-Bleuler subsidiary condition Eq. (4.20) in the manifestly covariant form

$$\partial_\mu A^{\mu(+)}(x)\; \mathcal{H}_P = 0 \ . \tag{4.23}$$

We notice that for the definition of the scalar products in $\mathcal{H}_{GB}$ we only needed the covariant field commutator Eq. (4.18) or the equivalent algebra of the creation and destruction operators Eq. (4.19), respectively. Therefore the scalar products on $\mathcal{H}_{GB}$ are manifestly L-invariant. Because of the covariant projection, this is also true for the scalar products on $\mathcal{H}_P$. The particular choice of the polarization vectors, respectively, the equivalence classes of them, do not play a role anymore.

The physical state space $\mathcal{H}_{\text{phys}}$

The space $\mathcal{H}_P$ may be completed in an abstract sense to a Pre-Hilbert-space. This space still exhibits states of norm zero. A vector $\psi \in \mathcal{H}_P$ can be decomposed as $\psi = \psi_+ + \psi_0$ in a unique way into a vector $\psi_+ \in \mathcal{H}_+$ of positive norm and a vector $\psi_0 \in \mathcal{H}_0$ having zero norm. Unfortunately, however, this decomposition is not L-invariant! In fact, under a generic Lorentz transformation a vector $\psi_+ \in \mathcal{H}_+$ acquires a component in $\mathcal{H}_0$, while $\mathcal{H}_P$ and $\mathcal{H}_0$ both are L-invariant. In $\mathcal{H}_P$ Schwarz's inequality holds, which implies that $\mathcal{H}_0$ is a linear subspace of $\mathcal{H}_P$. We therefore can form the quotient space

$$\mathcal{H}_{phys} = \mathcal{H}_P / \mathcal{H}_0 \ .$$

This is the linear space of equivalence classes of vectors from $\mathcal{H}_P$, which differ by elements from $\mathcal{H}_0$. As a quotient space of two invariant spaces $\mathcal{H}_{phys}$ is L-invariant, too. This space is in one-to-one correspondence to the Fock space of the physical transversal photons:

$$\begin{array}{lccl} \mathcal{H}_{GB} & : & a^+(\vec{p},\pm), b^+(\vec{p}\,), c^+(\vec{p}\,) & \text{acting on } |0> \\ \mathcal{H}_P & : & a^+(\vec{p},\pm), b^+(\vec{p}\,) & \text{acting on } |0> \\ \mathcal{H}_{Fock} & : & \{a^+(\vec{p},\pm)\}_{G_p-\text{orbit}} & \text{acting on } |0> \end{array}$$

where $\mathcal{H}_{Fock} \simeq \mathcal{H}_{phys}$ is isomorphic to a Hilbert space.

What is the bottom line?

The operators $a^+(\vec{p},\pm)$ cannot be identified with our naive understanding of creation operators for physical transverse photons! The reason is that the relationship between the field $A^\mu(x)$ and the creation and annihilation operators involves the polarization vectors and hence cannot be covariant. Since we insisted to have the field $A^\mu(x)$ manifestly covariant the corresponding $a^+(\vec{p},\pm)$ cannot any longer transform according to a unitary representation of $SL(2,C)$. With other words, starting from a covariant $A^\mu(x)$, the formulas defining the annihilation and creation operators Eqs. (4.16), (4.17) and (4.17) in terms of $A^\mu(x)$ lead automatically to operators a and c which are determined in equivalence classes only, while b turns out to be unambiguous.

The equivalence classes of a and c operators are induced by the L-transformations from the little group G_p of the light-like vector p which are equivalent to substitutions

$$\varepsilon_\pm^{\mu\,*}(p) \to \varepsilon_\pm^{\mu\,*}(p) + \lambda\, p^\mu \ .$$

Accordingly, the vector $\hat{p}$ must transform as (note that $\hat{p}^* = \hat{p}$)

$$\hat{p} \to \hat{p} + \lambda^2 p + \lambda\, \varepsilon_+^*(p) + \lambda\, \varepsilon_-^*(p)$$

The Eqs. (4.16) and (4.17) then tell us that

$$\begin{array}{lcccl} a(\vec{p},\pm) & \to & \tilde{a}(\vec{p},\pm)_\lambda & = & a(\vec{p},\pm) - \lambda\, b(\vec{p}\,) \\ c(\vec{p}\,) & \to & \tilde{c}(\vec{p}\,)_\lambda & = & c(\vec{p}\,) + \lambda^2 b(\vec{p}\,) - \lambda\, a(\vec{p},+) - \lambda\, a(\vec{p},-) \end{array}$$

which means that a state which contains a c-quantum remains such a state, however a state which contains only a-quanta does not remain such a state. More precisely, we may summarize the effect of the application of the different creation operators as follows:

$a^+(\vec{p},\pm)$	:	generate transversal photons plus b–ghosts
$b^+(\vec{p}\,)$	:	generates b–ghosts, scalar massless unphysical bosons
$c^+(\vec{p}\,)$	:	generates unphysical longitudinal photons plus b–ghosts and transversal photons

Consequently the equivalence classes $\tilde{a}(\vec{p}, \pm)_\lambda$, respectively, the states $|\vec{p}, \pm >_\lambda$ the corresponding creation operators create from the vacuum are the objects to be identified with the unique quantities $a(\vec{p}, \pm)$, respectively, the states $|\vec{p}, \pm >$ constructed as unitary representations of the Poincaré group. We remind the reader that the problems are unavoidable, due to the fact that transversal polarization vectors for massless spin 1 (or higher) particles cannot be defined in a covariant way. The transversality refers to the transformation properties under rotations. In contrast to the massive case, where the rest frame is invariant with respect to rotations, there does not exist a rotationally invariant standard frame. In the massless case where we must start quantization using a light-like vector p we necessarily are lead to consider an equivalence class of standard frames, related to each other by the L-transformations which leave the chosen light-like vector invariant. This stability group also called little group G_p is actually equivalent to the group $E(2)$ of translations in a plane, which is not semi-simple and thus has an Abelian subgroup.

In short: physical basic principles in general cannot be realized in a manifest way, but only modulo Abelian gauge transformations. Physics must be independent of the gauge i.e. invariant under $A_\mu(x) \to A_\mu(x) + \partial_\mu \alpha(x)$ where $\alpha(x)$ is an arbitrary scalar field. This is quite similar to symmetries in quantum mechanics, which need not be realized by true representations of the symmetry group, it suffices to have representations up to a phase (ray representations).

Final remark: Relationship between the Gupta-Bleuler formulation and the formal argumentation with the covariant Lorentz condition:

1) Gupta-Bleuler: $A^\mu(x)$ covariant and $\Box A^\mu = 0$! and commutator given by Eq. (4.18).

2) Covariant gauge fixing with gauge condition $\partial_\mu A^\mu(x) = 0$ implemented by adding the gauge fixing term $\mathcal{L}_{GF} = -\frac{1}{2}\xi^{-1}\left(\partial_\mu A^\mu(x)\right)^2$. The corresponding equation of motion reads

$$\Box A^\mu - (1 - \xi^{-1})\partial^\mu \partial_\nu A^\nu(x) = 0$$

which yields 1) for $\xi = 1$. In this formalism the Lagrangian has a residual gauge invariance $A_\mu(x) \to A_\mu(x) + \partial_\mu \alpha(x)$ with $\alpha(x)$ an arbitrary differentiable function satisfying $\Box\alpha(x) = 0$.

3) Comparison: in a covariant gauge we still expect the commutator to be of the form Eq. (4.18) for an appropriate choice of β. Applying the equation of motion of 2) to the commutator yields

$$\begin{aligned} \Box\ [A^\mu(x), A^\nu(y)] &= -i\ (-g^{\mu\nu}\Box + 2\beta\Box\partial^\mu\partial^\nu)\ \Delta(x-y;0) \\ &\quad +i\left(1 - \frac{1}{\xi}\right)(-1 + 2\beta\Box)\ \partial^\mu\partial^\nu\Delta(x-y;0) \\ &= 0\ , \end{aligned}$$

which can only be satisfied for a "singular", non–c-number $\beta \propto \Box^{-1}$. Since $\Box\Delta(x-y;0) = 0$, but $''(\beta\Box)''\ \Delta(x-y;0) \neq 0$ the equation of motion requires $2\beta\Box = 1 - \xi$ or

$$\beta = \frac{1-\xi}{\Box}$$

a non–local operator! For $\xi = 1$ the non locality is absent and we are back in the Gupta-Bleuler formulation, where β is an arbitrary c-number constant.

The crucial point: ξ–dependent terms always decouple from the dynamics! In order to show this on has to extend the above consideretion to the ineracting theory.

4.2 Exercises: Section 4

① Show that the Maxwell equation $\partial_\mu F^{\mu\nu} = 0$ as a field equation for the vector potential takes the form

$$\Box A^\nu(x) - \partial^\nu \left(\partial_\mu A^\mu(x)\right) = 0 \ .$$

Show that $A^\nu(x)$ is not determined by this equation because the operator $\Box g^{\mu\nu} - \partial^\mu \partial^\nu$ has no inverse. Hint.: $\varphi_\mu = \partial_\mu \alpha(x)$, $\alpha(x)$ an arbitrary scalar function, is a solution of the above equation with eigenvalue 0.

② Show that for a massive spin 1 field the Proca equation

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

implies $\partial_\mu A^\mu(x) \equiv 0$ automatically. Comment on the number of degrees of freedom ! Show that the Proca equation is the Euler-Lagrange equation of the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{m^2}{2} A_\mu A^\mu \ \ ; \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ .$$

Discuss the invariance properties of $\mathcal{L}$ under gauge transformations.

③ Prove that under local gauge transformations

$$\psi \to e^{-ie\alpha(x)}\psi \ , \quad A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x)$$

the covariant derivative $D_\mu = \partial_\mu - ieA_\mu$ has the property: $D_\mu \psi$ transforms identical to ψ and $\bar{\psi}\Gamma D_\mu \psi$ is gauge invariant provided D_μ commutes with the 4 by 4 matrix Γ.

5 Internal symmetry groups

Empirically, physical states are known to often show up in multiplets of symmetry groups. Familiar examples are the flavor symmetries in strong interaction physics $SU(2)_{\text{flavor}}$ (= isospin), $SU(3)_{\text{flavor}}$ (= isospin plus hypercharge), etc. Internal symmetries are to be distinguished from the space-time symmetries, which are unified in the Poincaré group. The symmetry groups G of interest here are the groups $SU(n)$, defined by the set of complex $n \times n$-matrices U which are **unitary** ($U^{-1} = U^{+}$) and **unimodular** ($\det U = 1$) and with matrix multiplication as the group operation. The requirement of unitarity ensures that the transition probabilities between states are preserved:

$$|< \varphi \mid \psi >|^2 = |< \varphi' \mid \psi' >^2 = |< \varphi \mid U^{+}U \mid \psi >|^2$$

of course $\mid \psi > \to \mid \psi' >= U \mid \psi >$ is a symmetry if and only if all group elements U commute with the total Hamiltonian H of the system:

$$[U, H] = 0 \ \ \forall\, U \ \in \ G.$$

Since any unitary matrix $\tilde{U}$ can be written as a product $\tilde{U} = Ue^{i\varphi}$ of a matrix U with $\det U = 1$ and a phase factor $e^{i\varphi}$, a unitary group $U(n)$ is equivalent to a direct product $SU(n) \otimes U(1)$. Therefore we may restrict ourselves to a consideration of the simple groups $SU(n)$. Possible $U(1)$ factors may be discussed separately.

The groups $SU(n)$ have $r = n^2 - 1$ real continuous parameters ω_i ($i = 1, \ldots, r$). A complex $n \times n$ matrix has $2n^2$ real parameters, unitarity implies n^2 conditions and $\det U = 1$ yields one further condition. Therefore, $SU(n)$ is characterized by r infinitesimal **generators** T_i and a general $SU(n)$ transformation can be written as

$$U = U(\omega) = \exp\left(i \sum_{i=1}^{n^2-1} T_i \omega_i\right)$$

and r is called **order** of the group.

The generators are Hermitian $T_i = T_i^{+}$ (which guarantees that U is unitary), traceless $Tr\, T_i = 0$ (which implies $\det U = 1$) and may be normalized so that $Tr\ (T_i T_j) = \frac{1}{2}\delta_{ij}$.

A convenient (non unique) basis for the matrices T_i, written conventionally as $T_i = \lambda_i/2$, can be constructed as follows. For the $n-1$ possible diagonal traceless Hermitian λ_i choose

$$\begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & -2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}, \ldots, \sqrt{\frac{2}{n(n-1)}}\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & -(n-1) \end{pmatrix}.$$

Then form the $\frac{n(n-1)}{2}$ off-diagonal matrices λ_i with 1 in a given off-diagonal position above the diagonal, 1 in the transposed position and zeros elsewhere. Also form the $\frac{n(n-1)}{2}$ off-diagonal matrices λ_i with a $-i$ in a given off-diagonal position above the diagonal, $+i$ in the transposed position and zeros elsewhere.

The **Lie-algebra** (commutation rules)

$$[T_i, T_k] = ic_{ikl}T_l$$

determines the structure constants c_{ikl}, which are real, totally antisymmetric and satisfy the **Jacobi identity**

$$c_{ikn}c_{nlm} + \text{ cyclic terms in}(ikl) \equiv 0.$$

A Lie group G and its structure constants c_{ikl} uniquely determine each other in a neighborhood of the identity element of G. In a Lie-algebra there is a maximum number l of simultaneously commuting (i.e. diagonal) elements. l is called the **rank** of the group. The $SU(n)$ groups have rank $l = n - 1$, which is obvious in the basis given above. The states belonging to a $SU(n)$ multiplet may be labeled, as usual, by the eigenvalues of the simultaneous eigenstates of the l diagonal matrices which we denote by $H_1, \ldots, H_l$. The structure of a multiplet is characterized by a **weight diagram** which displays the eigenvalues of the states on a $(H_1, \ldots, H_l)$ plot.

The remaining generators may be combined into pairs of **ladder operators** (a raising and a lowering operator) $E_{\pm\alpha}(\alpha = 1, \ldots, \frac{r-l}{2})$ which map the different eigenstates of a multiplet into each other. The $E_{\pm\alpha}$'s are non-Hermitian matrices with 1 in a given off-diagonal position and zeros elsewhere.

For $SU(2)$ and $SU(3)$ we list some basic properties in the following.

a) $SU(2)$: Order $r = 3$, rank $l = 1$

Structure constants: $c_{ikl} = \epsilon_{ikl}$ the fully antisymmetric permutation tensor.

Generators: $T_i = \frac{\tau_i}{2}$; τ_i the Pauli matrices [21]

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Diagonal operators: $H_1 = \frac{\tau_3}{2} = I_3$: 3^{rd} component of isospin.

Eigenvectors:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Eigenvalues of I_3 : $\frac{1}{2}, -\frac{1}{2}$
Ladder operators: $E_{\pm 1} = \frac{1}{2}(\tau_1 \pm i\tau_2)$

$$E_{+1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{-1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

[21] Properties of the Pauli matrices:

$$\begin{aligned} [\tau_i, \tau_k] &= 2i\epsilon_{ikl}\tau_l , \quad \{\tau_i, \tau_k\} = 2\delta_{ik} \\ \tau_i^+ &= \tau_i , \quad \tau_i^2 = 1 , \quad Tr\, \tau_i = 0 \\ \tau_i\tau_k &= \frac{1}{2}\{\tau_i, \tau_k\} + \frac{1}{2}[\tau_i, \tau_k] = \delta_{ik} + i\epsilon_{ikl}\tau_l \end{aligned}$$

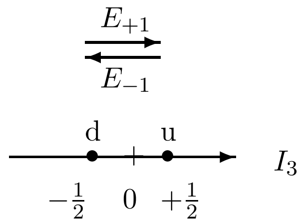


Fig. 5.1: Weight diagram for $\begin{pmatrix} u \\ d \end{pmatrix}$ quark doublet.

b) $SU(3)$: Order $r = 8$, rank $l = 2$

Structure constants : $c_{ikl} = f_{ikl}$, where the non-vanishing entries are permutations of the elements

$$f_{123} = 1$$
$$f_{147} = f_{165} = f_{246} = f_{257} = f_{345} = f_{376} = \tfrac{1}{2}$$
$$f_{458} = f_{678} = \sqrt{3}/2.$$

Generators: $T_i = \frac{\lambda_i}{2}$; λ_i the Gell-Mann matrices [22]

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Diagonal operators:

$$H_1 = \frac{\lambda_3}{2} = I_3 \quad 3^{rd} \text{ component of isospin}$$

$$H_2 = \frac{\lambda_8}{2} \doteq \frac{\sqrt{3}}{2} Y \ , \quad Y \text{ hypercharge}$$

[22] Properties of the Gell-Mann matrices:

$$\begin{aligned} [\lambda_i, \lambda_k] &= 2i\, f_{ikl} \lambda_l \ , \ \ \{\lambda_i, \lambda_k\} = \frac{4}{3}\delta_{ik} + 2d_{ikl}\lambda_l \\ Tr\,\lambda_i &= 0 \ , \ \ Tr\,\lambda_i\lambda_k = 2\delta_{ik} \\ Tr\,\lambda_i\,[\lambda_k, \lambda_l] &= 4i\, f_{ikl} \ , \ \ Tr\,\lambda_i\,\{\lambda_k, \lambda_l\} = 4i\, d_{ikl} \end{aligned}$$

Eigenvectors:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Eigenvalues of $(I_3, Y): \left(\frac{1}{2}, \frac{1}{3}\right)$, $\left(-\frac{1}{2}, \frac{1}{3}\right)$, $\left(0, -\frac{2}{3}\right)$

Ladder operators:

$$\begin{aligned}
E_{\pm 1} &= T_1 \pm iT_2 \;:\; E_{+1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \; E_{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
E_{\pm 2} &= T_4 \pm iT_5 \;:\; E_{+2} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \; E_{-2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\
E_{\pm 3} &= T_6 \pm iT_7 \;:\; E_{+3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} , \; E_{-3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}
\end{aligned}$$

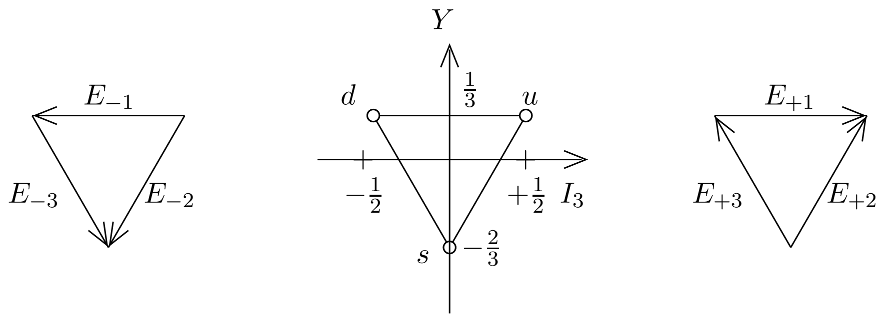


Fig. 5.2: Weight diagram for $\begin{pmatrix} u \\ d \\ s \end{pmatrix}$ quark triplet.

<u>Finite dimensional representations of $SU(n)$</u>

Given the structure constants c_{ikl} of $SU(n)$ any set of Hermite-an traceless $N \times N$ matrices $\tilde{T}_i (i = 1, \dots n^2 - 1)$ satisfying the Lie-algebra

$$\left[\tilde{T}_i, \tilde{T}_k\right] = ic_{ikl}\tilde{T}_l$$

is called a **representation** of the $SU(n)$ Lie-algebra. The unitary unimodular matrices $\tilde{U} = \exp(i \sum_{i=1}^{n^2-1} \tilde{T}_i \omega_i)$ then form a representation of $SU(n)$. The smallest non-trivial irreducible representation is the **fundamental representation** of dimension $N = n$. This is the representation which defines $SU(n)$. In gauge theories the fundamental spin 1/2 matter fields of quarks and leptons are in this representation. The Jacobi identity implies that there always exists the **adjoint representation** of dimension r with generators

$$(\tilde{T}_i)_{kl} = -ic_{ikl}$$

In gauge theories the spin 1 gauge fields in any case must be in this representation, as we shall see below.

The complex **conjugate** of a representation is also a representation since

$$(U_1U_2)^* = U_1^*U_2^*.$$

Two representations are **equivalent** if we can transform one into the other by a change of basis:

$$SD_1(U)S^{-1} = D_2(U) \ \ \forall\, U \ \in \ G.$$

The conjugate representation n^* of the fundamental representation n is not a new representation if it is equivalent to n. In fact for the fundamental representation 2 of $SU(2)$ 2^* is equivalent to 2. In contrast, the conjugate representations 3^* of the fundamental representation 3 of $SU(3)$ is a new (inequivalent) representation. In QCD this crucial property of $SU(3)$ allows to distinguish color triplets of quarks (which transform according to the 3 representation) from color triplets of antiquarks (which transform according to the 3^* representation).

A representation is called **irreducible** if it cannot be transformed by a change of basis to block-diagonal form:

$$D(U) = \begin{pmatrix} D_1(U) & 0 \\ 0 & D_2(U) \end{pmatrix} = D_1(U) \oplus D_2(U) \ \ \forall\, U \ \in \ G.$$

If such a transformation exist, the representation is **reducible**. The irreducible representations are the basic building blocks of any representation. Particle multiplets are classified by the irreducible representations of a symmetry group.

The possible irreducible representations can be constructed by decomposing products of the fundamental representation into irreducible blocks. In the following we briefly discuss how this can be done.

<u>Combining representations, reduction</u>

Let $\psi_i (i = 1, \ldots n)$ be a **vector** transforming under the fundamental representation n of $SU(n)$. A **tensor product** $\psi_{i_1} \ldots \psi_{i_m}$ forms a tensor $\psi_{i_1 \ldots i_m}$ which transforms according to

$$\psi_{i_1 \ldots i_m} \to \psi'_{i_1 \ldots i_m} = U_{i_1 i'_1} \ldots U_{i_m i'_m} \psi_{i'_1 \ldots i'_m} \, .$$

For $m > 1$ this product representation, denoted by $n \otimes n \otimes \ldots \otimes n$ (m factors), is reducible.

One can decompose $\psi_{i_1 \ldots i_m}$ into a sum of tensors of different **symmetry class** with respect to **permutations** of the indices $i_1 \ldots i_m$ as follows:

Choose a set of positive integers $n_1, n_2 \ldots, n_k$ with $n_1 \geq n_2 \geq \ldots \geq n_k$ which form a partition of m : $n_1 + n_2 + \ldots + n_k = m$. Then group the indices $i_1 \ldots i_m$ into k classes $(i_{11} \ldots i_{1n_1})$, $(i_{21} \ldots i_{2n_2}), \ldots, (i_{k1} \ldots i_{kn_k})$ and write them in form of a tableau of k stacked rows where the first row has n_1 boxes containing the indices $i_{11} \ldots i_{1n_1}$, the second row has n_2 boxes containing the indices $i_{21} \ldots i_{2n_2}$ and so on. The tableau obtained is called a **Young tableau** (often called

Young diagram).

1. symmetrize in rows →

2. antisymmetrize in columns ↓

$$\begin{array}{|c|c|c|c|c|c|c|c|}
\hline
i_{11} & i_{12} & \cdot & \cdot & \cdot & \cdot & \dots & i_{1n_1} \\
\hline
i_{21} & i_{22} & \cdot & \cdot & \dots & i_{2n_2} & \multicolumn{2}{c}{} \\
\cline{1-6}
\vdots & & & & \multicolumn{4}{c}{} \\
\cline{1-4}
i_{k1} & i_{k2} & \dots & i_{kn_k} & \multicolumn{4}{c}{} \\
\cline{1-4}
\end{array}$$

With a Young tableau we associate a tensor of a given symmetry class by the following construction. By **convention** first symmetrize $\psi_{i_1 \dots i_m}$ in each group of indices appearing in the rows. Afterwards anti-symmetrize in the indices appearing in each column. By this a tensor of a given symmetry class is defined. According to the convention (first symmetrize in rows then anti-symmetrize in columns) tableaus with indices permuted in columns represent the same tensor. Tableaus with indices permuted in rows represent the same tensor if and only if the indices are not anti-symmetrized with indices in a different row.

For $SU(n)$ a tensor index can take the values $i = 1, \dots, n$ only. Hence, there cannot be more than n rows for anti-symmetrization.

One easily verifies that group transformations do not mix tensors from different symmetry classes. The following central theorem holds:

a) Tensors in a given symmetry class form an invariant irreducible subspace. The group representation induced (by projection to the invariant subspace) in this subspace by the fundamental representation is irreducible.

b) The irreducible representations generated through all possible symmetry classes are exhaustive (i.e. there are no irreducible representations which cannot be obtained this way).

Symmetrization and anti-symmetrization obviously reduces the number of independent components of a tensor. The number of independent components of a tensor of a given symmetry class is equal to the **dimension** of the irreducible representation.

The irreducible representation of highest dimension is represented by the totally symmetric tensor.

$$\begin{array}{|c|c|c|c|}
\hline
i_1 & i_2 & \quad \dots \quad & i_m \\
\hline
\end{array}$$

There is only **one** such representation in $n \otimes \dots \otimes n$ (m factors)

$$\psi_{(i_1 \dots i_m)} = \frac{1}{m!} \sum_{\text{permutations}\, p} \psi_{ip(1) \cdots ip(m)} \ .$$

A column with n boxes represents a tensor of rank zero i.e. a singlet and corresponds to a 1-dimensional trivial representation:

$$\varphi = \epsilon^{i_1 \dots i_n} \psi_{i_1} \dots \psi_{i_n}$$

Therefore, if a column with n boxes is part of a larger Young tableau it can be omitted!

A $n-1$ fold antisymmetric product of ψ_i's transforms as a n^* (complex conjugate of the fundamental representation n):

$$\chi^i = \epsilon^{ii_1\ldots i_{n-1}}\psi_{i_1}\ldots\psi_{i_{n-1}}$$

since

$$\psi_i\chi^i = \varphi = \epsilon^{ii_1\ldots i_{n-1}}\psi_i\psi_{i_1}\ldots\psi_{n-1}$$

is a singlet.

We now present some specific properties of $SU(2)$ and $SU(3)$:

a) $SU(2)$:
The fundamental representation is 2 . 2^* is equivalent to 2 . Indices have 2 possible values $i = 1, 2$.

A tableau $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ is a singlet. $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ as a part of a larger tableau can be omitted i.e. $\begin{smallmatrix}\square\square\\\square\end{smallmatrix} \equiv \square$ etc. All nontrivial representations are characterized by a row:

Tableau:	$\square$,	$\square\square$,	$\square\square\square$,	etc.
Dimension:	2	3	4	

Product representations and their reduction follow by combining corresponding tableaus in all possible ways.

Examples: $SU(2)$ interpreted as spin

1.
$$2\otimes 2 \;=\; \square\times\square = \begin{smallmatrix}\square\\\square\end{smallmatrix} + \square\square = 1\oplus 3$$
i.e. two spin 1/2 particles can group into a singlet of spin 0 and a triplet of spin1

2.
$$\begin{aligned} 2\otimes 2\otimes 2 = (\begin{smallmatrix}\square\\\square\end{smallmatrix} + \square\square)\times\square \;&=\; \begin{smallmatrix}\square\square\\\square\end{smallmatrix} + \begin{smallmatrix}\square\square\\\square\end{smallmatrix} + \square\square\square \\ &=\; \square + \square + \square\square\square \\ &=\; 2\oplus 2\oplus 4 \end{aligned}$$
i.e. three spin 1/2 particles can group into two doublets of spin 1/2 and a quartet of spin 3/2.

b) $SU(3)$:
The fundamental representation is 3 . 3^* is inequivalent to 3. Indices have 3 possible values $i = 1, 2, 3$.

A tableau $\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}$ is a singlet. $\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}$ as part of a larger diagram can be omitted i.e. $\begin{smallmatrix}\square\square\\\square\\\square\end{smallmatrix} \equiv \square$. etc.

All non-trivial representations are characterized by tableaus with one or two columns:

$$\underbrace{\begin{smallmatrix}\square\square\square\\\square\square\square\end{smallmatrix}}_{q}\underbrace{\begin{smallmatrix}\square\square\square\\ \end{smallmatrix}}_{p}$$

Each $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ corresponds to a n^* i.e. an irreducible representation is characterized by two indices (p, q) and transforms as a tensor

$$\psi^{j_1\ldots j_q}_{i_1\ldots i_p} \text{ symmetrized in } (i_1\ldots i_p) \text{ and } (j_1\ldots j_q)$$

where $i_1\ldots i_p$ transform under 3 and $j_1\ldots j_q$ under 3^*.

We may write $\psi^{j_1\dots}_{i_1\dots}$ in product form

$$\psi^{j_1\dots j_q}_{i_1\dots i_p} = \chi^{j_1}\dots\chi^{j_q}\psi_{i_1}\dots\psi_{i_q}$$

with $\chi^i = \epsilon^{ikl}\psi_k\psi_l$. Together with the symmetrization it can be shown that the trace condition

$$\sum_{j=1}^{3}\psi^{jj_2\dots j_q}_{ji_2\dots i_p} = 0$$

must hold. This restricts the number of independent components of the tensor, which equals the dimension of the irreducible representation $D(p,q)$: One finds

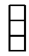
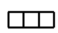
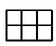
$$D(p,q) = \frac{1}{2}(p+1)(q+1)(p+q+2).$$

The generators $\tilde{T}_i$ of a given irreducible representation can be worked out from the transformation law

$$\psi'^{j_1\dots j_q}_{i_1\dots i_p} = U^*_{j_1j'_1}\dots U^*_{j_qj'_q}U_{i_1i'_1}\dots U_{j_pj'_p}\psi^{j'_1\dots j'_q}_{i'_1\dots i'_p}$$

for infinitesimal transformations.

The simplest irreducible representation are given in the following table:

(p,q)		$D(p,q)$	*tableau*	*tensor*	
$(0,0)$		1		1	singlet
$(1,0)$		3		ψ_i	triplet
$(0,1)$		3^*		ψ^i	antitriplet
$(2,0)$		6		ψ_{ik}	sextet
$(0,2)$		6^*		ψ^{ik}	antisextet
$(1,1)$	adjoint	$8=8^*$		ψ^i_k	octet
$(3,0)$		10		ψ_{ikl}	decaplet
$(0,3)$		10^*		ψ^{ikl}	antidecaplet

Application to $SU(3)_{\text{flavor}}$:

Low lying hadronic states may be classified in $SU(3)_{\text{flavor}}$ multiplets. The relevant quantum numbers are the baryon number B, isospin I and strangeness S. We can achieve that multiplets are centered on the origin if we replace strangeness S by hypercharge Y

$$Y = B + S.$$

Empirically, the electric charge of a hadron is given by

$$Q = I_3 + \frac{Y}{2}\ .$$

In the quark model of hadrons mesons ($B=0$) are quark – antiquark states

$$M = (q\bar{q})$$

baryons ($B = 1$) are three quark states

$$B = (qqq)$$

where $q = u, d, s$. The quarks (u, d, s) are in the fundamental representation 3, the antiquarks $(\bar{u}, \bar{d}, \bar{s})$ in the representation 3^*.

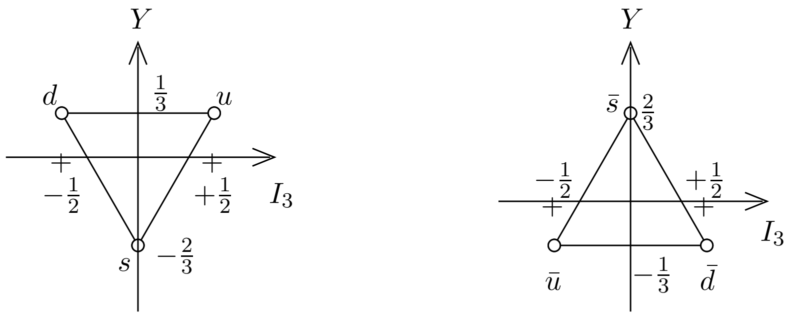


Fig. 5.3: Basic building blocks of the $SU(3)$ quark model.

Direct products of representations may be reduced (decomposed) into irreducible blocks by combining boxes of the corresponding Young tableaux in all possible ways with the restriction that antisymmetric pairs must be preserved. The latter condition is nontrivial but may be satisfied by the following construction:

In order to append to the first tableau the second one in all admissible ways which respect the (anti -) symmetrization, we place in each box of the second tableau letters (in lexicographic order) with identical letters in each given row (symmetrized). Thus we insert a's in the first row, b's in the second row, etc. All boxes of the second tableau are now appended to the right-hand ends of the rows of the first one (which represents the upper left-hand corner of the new diagram) in all possible ways. Thus we first append all a's to the first tableau (in all admissible ways) with no more than one a per column (anti-symmetrized). To the such obtained enlarged tableaux append all b's (in all admissible ways) with no more than one b per column, etc.

Some of the tableaux such obtained are not admissible because they do not take into account properly the (anti -) symmetrization of the original boxes and have to be thrown away (also in order to avoid double counting).

Here we need a definition: a sequence of letters a, b, c, $\cdots$ is *admissible* if at any point of the sequence at least as many a's have occurred as b's, at least as many b's have occurred as c's etc.

Examples: a) admissible: abcd, aabcb, ...

b) not admissible: abb, acb, ...

Now consider for each tableau constructed above the full sequence of letters formed by reading from *right to left* in the first row, then in the second row etc. The tableaux which we have to throw away are those which lead to sequences of letters which are not admissible.

The properties of the composed new tableaux may be summarized as follows:

1. Each tableau must be a Young tableau.

2. The number of boxes in the new tableau must be equal to the sum of the number of boxes in the original two tableaux.

3. If dealing with $SU(n)$, no tableau has more than n rows.
4. Making a journey through the tableau starting with the top row and entering each row from the right, at any point the number of b's encountered in any of the attached boxes must not exceed the number of previously encountered a's and the number of c's encountered in any of the attached boxes must not exceed the number of previously encountered b's, etc.
5. The letters must be in anti-lexicographical order when reading across a row from left to right.
6. The letters must differ and be in lexicographic order when reading a column from top to bottom.

The first three rules should be obvious. The purposes of the three rules 4) to 6) are to assure that states which were previously symmetrized are not anti-symmetrized in the product and vice versa, and to avoid double counting states.

Examples:

$$
\begin{array}{lll}
1.\ \ 3\otimes 3 & = & \square\times\square = \begin{smallmatrix}\square\\\square\end{smallmatrix} + \square\square = 3^*\oplus 6 \\
2.\ \ 3\otimes 3^* & = & \square\times\begin{smallmatrix}\square\\\square\end{smallmatrix} = \begin{smallmatrix}\square\\\square\\\square\end{smallmatrix} + \begin{array}{l}\square\square\\\square\end{array} = 1\oplus 8 \\
3.\ \ 3\otimes 3\otimes 3 & = & \left(\begin{smallmatrix}\square\\\square\end{smallmatrix} + \square\square\right)\times\square = \begin{smallmatrix}\square\\\square\\\square\end{smallmatrix} + \begin{array}{l}\square\square\\\square\end{array} + \begin{array}{l}\square\square\\\square\end{array} + \square\square\square \\
 & = & 1\oplus 8\oplus 8\oplus 10
\end{array}
$$

More than two tableaux may be combined by first combining the first two, then combining the result with the third one and so on.

Exercise: Show that $8\otimes 8 = 1\oplus 8\oplus 8\oplus 10\oplus 10\oplus 27$.

The quantum numbers of quarks are given by:

Quark	spin	B	Q	I_3	S	Y
u	1/2	1/3	2/3	1/2	0	1/3
d	1/2	1/3	$-1/3$	$-1/2$	0	1/3
s	1/2	1/3	$-1/3$	0	-1	$-2/3$

Exercise: Use the Young tableaux to construct the meson states in

$$3\otimes 3^*$$

and the baryon states in

$$3\otimes 3\otimes 3\ .$$

Notice that the indices of the tensors $\psi^{j_1\ldots j_q}_{i_1\ldots i_p}$, which characterize a irreducible representation (p,q), in the quark model have two different interpretations. Each upper index has associated either an antiquark or an anti-symmetrized pair of quarks. For lower indices antiquarks and quarks are interchanged: i.e.

$$\text{Upper index}: \quad \begin{pmatrix} \bar{u} \\ \bar{d} \\ \bar{s} \end{pmatrix} \quad or \quad \begin{pmatrix} (ds) \\ (su) \\ (ud) \end{pmatrix}$$
$$B = -1/3 \qquad\qquad B = 2/3$$

$$\text{Lower index}: \quad \begin{pmatrix} u \\ d \\ s \end{pmatrix} \quad or \quad \begin{pmatrix} (\bar{d}\bar{s}) \\ (\bar{s}\bar{u}) \\ (\bar{u}\bar{d}) \end{pmatrix}$$
$$B = 1/3 \qquad\qquad B = -2/3$$

where (ud) etc. denote anti-symmetrized pairs. Which interpretation is to be used is uniquely fixed if we specify the baryon number B of the state.

5.1 The spectrum of low lying hadrons:

Mesons: $\bar{q}q'$ bound states
A $\bar{q}q'$ with orbital angular momentum L has Parity $P = (-1)^{L+1}$. For $q' = q$ we have a $\bar{q}q$ bound state which is also an eigenstate of charge conjugation C with $C = (-1)^{L+S}$, where S is the spin 0 or 1. The $L = 0$ states are the pseudoscalar mesons, $J^P = 0^-$, and the vectors mesons, $J^P = 1^-$.

In the limit of exact $SU(3)$ the pure states would read

$$\begin{aligned} \pi^0 &= (\mathbf{\bar{u}u - \bar{d}d})/\sqrt{\mathbf{2}} \\ \eta_1 &= (\mathbf{\bar{u}u + \bar{d}d + \bar{s}s})/\sqrt{\mathbf{3}} \\ \eta_8 &= (\mathbf{\bar{u}u + \bar{d}d - 2\bar{s}s})/\sqrt{\mathbf{6}} \end{aligned}$$

$$\begin{aligned} \rho^0 &= (\mathbf{\bar{u}u - \bar{d}d})/\sqrt{\mathbf{2}} \\ \omega_1 &= (\mathbf{\bar{u}u + \bar{d}d + \bar{s}s})/\sqrt{\mathbf{3}} \\ \omega_8 &= (\mathbf{\bar{u}u + \bar{d}d - 2\bar{s}s})/\sqrt{\mathbf{6}} \end{aligned}$$

In fact $SU(2)_{\text{flavor}}$ breaking by the quark mass difference $m_d - m_u$ leads to $\rho - \omega$–mixing [mixing angle $\sim^\circ$]:

$$\begin{aligned} \rho^0 &= \cos\theta\,\rho' + \sin\theta\,\omega' \\ \omega &= -\sin\theta\,\rho' + \cos\theta\,\omega' \end{aligned}$$

Similarly, the substantially larger $SU(3)_{\text{flavor}}$ breaking by the quark masses, leads to large $\omega - \phi$–mixing [mixing angle $\sim 36^\circ$ close to so called ideal mixing where $\phi \sim$ is a pure $\bar{s}s$ state]:

$$\begin{aligned} \phi &= \cos\theta\,\omega_8 + \sin\theta\,\omega_1 \\ \omega &= -\sin\theta\,\omega_8 + \cos\theta\,\omega_1 \end{aligned}$$

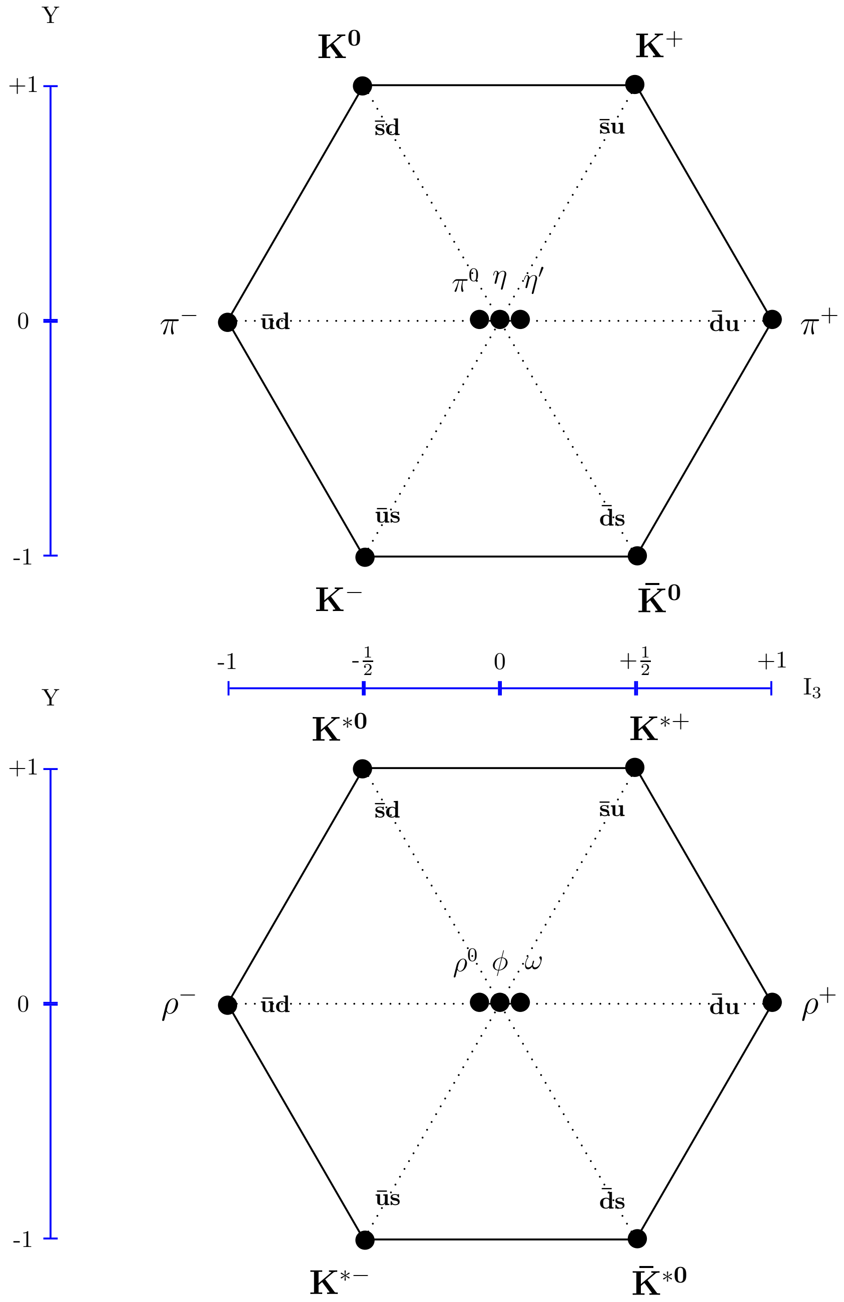
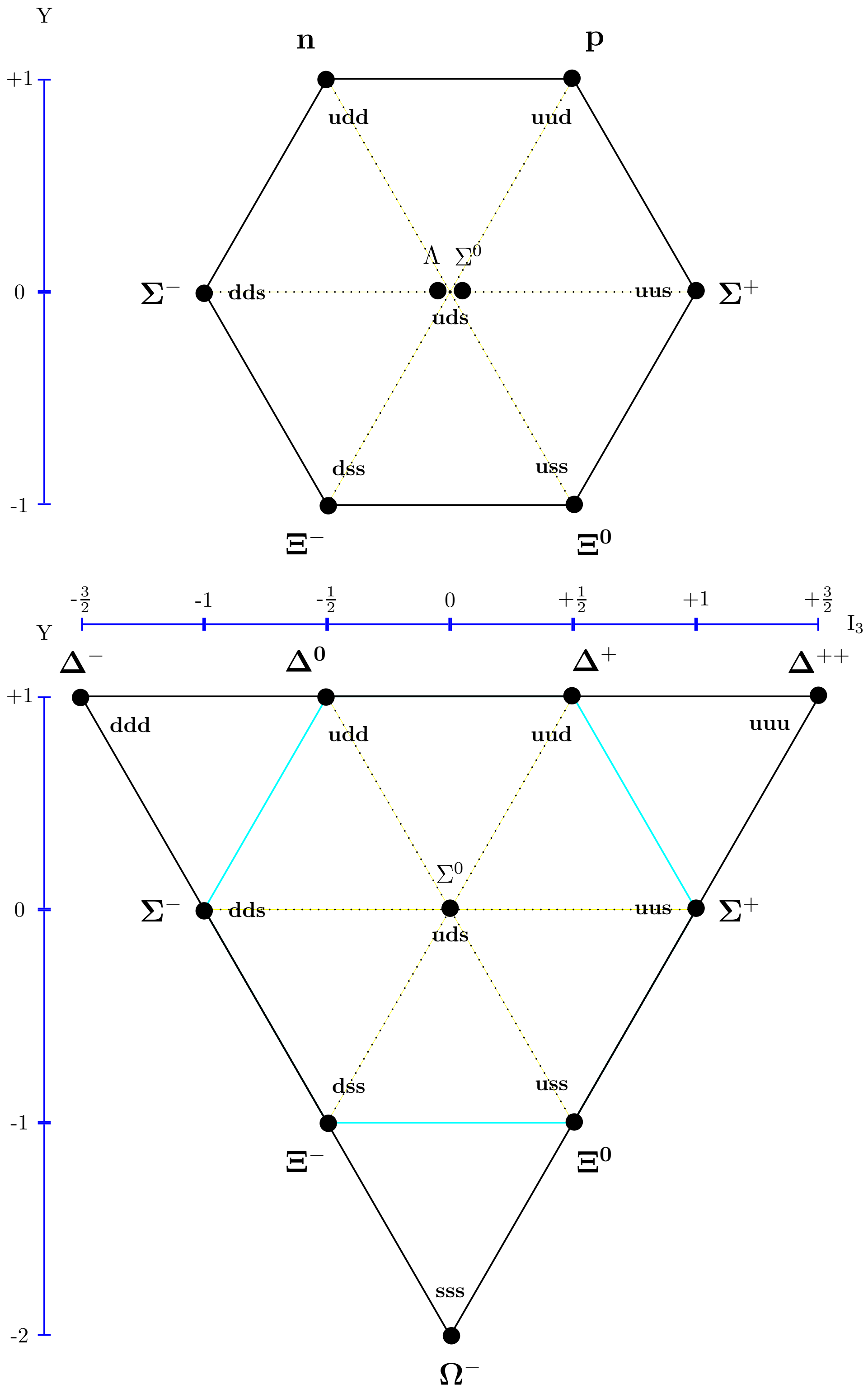

Y
+1
0
-1
$\mathbf{K^0}$
$\mathbf{K^+}$
$\mathbf{\bar{s}d}$
$\mathbf{\bar{s}u}$
π^0 η η'
π^-
$\mathbf{\bar{u}d}$
$\mathbf{\bar{d}u}$
π^+
$\mathbf{\bar{u}s}$
$\mathbf{\bar{d}s}$
$\mathbf{K^-}$
$\mathbf{\bar{K}^0}$
-1
$-\frac{1}{2}$
0
$+\frac{1}{2}$
+1
I_3
Y
+1
0
-1
$\mathbf{K^{*0}}$
$\mathbf{K^{*+}}$
$\mathbf{\bar{s}d}$
$\mathbf{\bar{s}u}$
ρ^0 ϕ ω
ρ^-
$\mathbf{\bar{u}d}$
$\mathbf{\bar{d}u}$
ρ^+
$\mathbf{\bar{u}s}$
$\mathbf{\bar{d}s}$
$\mathbf{K^{*-}}$
$\mathbf{\bar{K}^{*0}}$

Figure 5.9:



Y
+1
0
-1
n
p
udd
uud
Λ
Σ0
Σ−
dds
uds
uus
Σ+
dss
uss
Ξ−
Ξ0
-3/2
-1
-1/2
0
+1/2
+1
+3/2
I3
Y
Δ−
Δ0
Δ+
Δ++
+1
ddd
udd
uud
uuu
Σ0
0
Σ−
dds
uds
uus
Σ+
-1
dss
uss
Ξ−
Ξ0
sss
-2
Ω−

Figure 5.10:

Outlook

Nature plays all kind of games with symmetries. There are global and local space-time symmetries and global and local internal symmetries. The latter determine strong, weak and electromagnetic interactions by the local gauge group

$$G_{\mathrm{loc}} = SU(3)_c \otimes SU(2)_L \otimes U(1)_Y .$$

$SU(3)_c$ is the color gauge group of the strong interactions (QCD). $SU(2)_L \otimes U(1)_Y$ is the electroweak gauge group ($SU(2)_L$ the left-handed weak isospin and $U(1)_Y$ the weak hypercharge) which is broken to the Abelian electromagnetic $U(1)_{\mathrm{em}}$ of QED. Local space-time symmetry (general coordinate transformation invariance, general relativity) leads to classical gravity theory. Gravitational interactions break Poincaré invariance, which "in practice" appears as an absolute global symmetry due to the extreme weakness of the gravitational force. Global charge-like symmetries remain the most mysterious symmetries. The only known exact global symmetries are Abelian symmetries with the associated quantum numbers being **quantized**.

$$\begin{array}{lll} Q & \text{electric charge} & U(1)_Q \\ B & \text{baryon number} & U(1)_B \\ L_\ell (\ell = e, \mu, \tau) & \text{lepton numbers} & U(1)_{L_\ell} . \end{array}$$

The quantization of these charges is barely understood, notice that $U(1)$ invariance only implies the conservation of the corresponding charge not its quantization.

Baryon number B is an additive quantum number like the electric charge Q. It derives from a global $U(1)_B$ invariance of all standard model interactions. Empirically it is tested most accurately by the stability of the proton. Proton lifetime limits are

$$\begin{array}{ll} \tau_p^{geo} > & 1.6 \times 10^{25} \text{ years (geochemical estimate independent on decay modes)} \\ \tau_p^{lab} > & 10^{31} - 3 \times 10^{32} \text{ years (absence of "main" decay modes)} \end{array}$$

Possible decay modes which where searched for are:

$$\begin{array}{rcl} p & \to & e^+\gamma,\ e^+\pi^0,\ e^+\rho^0,\ e^+\omega,\ e^+K^0, \ldots, \\ & & \bar{\nu}_e\pi^+,\ \bar{\nu}_e\rho^+,\ \bar{\nu}_e K^+, \ldots, (e \to \mu) \\ p & \to & \pi^+ + \pi^0 \end{array}$$

By convention $B(p) = 1,\ B(e^-) = 0$. All observations support the assignments $B(B) = 1$ and $B(\bar{B}) = -1$ for baryons B and antibaryons $\bar{B}$, respectively. All other particles have $B = 0$.

□

Excursion on the *Baryon Asymmetry Of The Universe*: While we have no direct evidence that baryon number is not strictly conserved we know that at some level, presumably not far from current experimental limits, baryon number must be violated. Otherwise the observed baryon asymmetry, the asymmetry between matter and antimatter observed in our universe (galaxies, stars, planets,... but no anti–galaxies,anti–stars, anti–planets,...) could not be explained from properties of the fundamental interactions of nature. It would (and could) be just an accidental asymmetry in the initial condition at the moment of the creation of the universe at the big–bang.

In the early very hot universe matter and antimatter was created by the highly energetic photon collisions at (almost) equal rate. In fact with a tiny asymmetry: per unit density $\rho_{\bar{B}}$=1.000000000 of antimatter a portion

ρ_B=1.000000001 of matter must have been created. The expansion of the universe cooled down the radiation and matter and antimatter annihilated almost completely into photons with the relict $\rho_B - \rho_{\bar{B}}$=0.000000001 of matter. Thus the baryon number in units of the number of photons in the universe is $N_B/N_\gamma \sim 10^{-9}$.

A theory which is able of explaining the origin of the baryon asymmetry must satisfy three conditions (Sakharov 1967):

- It must violate B,
- it must violate CP and
- the universe must be out of thermal equilibrium.

The latter condition is satisfied since we know that the universe is expanding and not in a stationary equilibrium state.

Within the SM B is conserved to all orders in perturbation theory and only violated by extremely tiny non–perturbative effects. The latter are due to the existence of non–trivial classical solutions of the $SU(2)$ Yang–Mills equations, the so called instantons (Belavin et al. 1975). Quantum effects due to these four-dimensional pseudoparticles lead to symmetry breaking via Adler-Bell-Jackiw anomalies ('t Hooft 1976).

While B is practically conserved in the SM, CP violation is naturally incorporated in the SM with three (or more) families of quarks and leptons (Kobayashi and Maskawa 1973). In fact three families are known to exist, the last member of the third family the top quark with a mass of about 175 GeV has been found some time ago (CDF and D0 at Fermilab 1995). More recently CP violation has been established (Babar at SLAC and Belle at KEK 2001) in the B–meson system to be a large effect in accord with the SM prediction given the CP violation in the $K^0 - \bar{K}^0$ system, which has been discovered as a small effect $\varepsilon \simeq 2.3 \times 10^{-3}$ long time ago (Christenson, Cronin, Fitch and Turlay 1964).

The SM is very unlikely able to predict the correct size of the baryon asymmetry. The latter thus is a clear indication that the SM is only part of the full story.

□

The lepton numbers L_ℓ ($\ell = e, \mu, \tau$) are other additive quantum numbers which seem to be strictly conserved at first sight. By convention $L_\ell(\ell^-) = 1$. That L_μ is separately conserved follows from the non-observations of the decays

$$\begin{array}{cc}
\mu^+ \to e^+ + \gamma & \Gamma(\mu \to e\gamma)/\Gamma(\mu \to \mathrm{all}) < 1.2 \times 10^{-11} \\
\mu^+ \to e^+ + e^- + e^+ & \Gamma(\mu \to 3e)/\Gamma(\mu \to \mathrm{all}) < 1.0 \times 10^{-12} \\
K_L \to e + \mu & \Gamma(K_L \to e\mu)/\Gamma(K_L \to \mathrm{all}) < 4.7 \times 10^{-12} \\
K^+ \to \pi^+ + e + \mu & \Gamma(K^+ \to \pi^+ e\mu)/\Gamma(K^+ \to \mathrm{all}) < 2.1 \times 10^{-10} \\
\mu^- + (Z,A) \to e^- + (Z,A) & \Gamma(\mu^-\mathrm{Ti} \to e^-\mathrm{Ti})/\Gamma(\mu^-\mathrm{Ti} \to \mathrm{all}) < 4.0 \times 10^{-12} \\
\mu^- + (Z,A) \to e^+ + (Z-2,A) & \Gamma(\mu^-\mathrm{Ti} \to e^+\mathrm{Ca})/\Gamma(\mu^-\mathrm{Ti} \to \mathrm{all}) < 3.6 \times 10^{-11} \quad .
\end{array}$$

Tests of the separate conservation of L_τ are much less stringent: The best limits are:

$$\Gamma(\tau \to e\gamma)/\Gamma(\tau \to \mathrm{all}) < 2.7 \times 10^{-6} \;\; \text{and} \;\; \Gamma(\tau \to \mu\gamma)/\Gamma(\tau \to \mathrm{all}) < 1.1 \times 10^{-6} \quad .$$

Within the experimentally well established electroweak standard model strict lepton number conservation is only possible if the neutrinos are strictly massless. Non-vanishing neutrino masses lead to **neutrino-oscillations**. Neutrino mixing searches (ν-oscillations $\nu_\ell \leftrightarrow \nu_{\ell'}$) have confirmed the effect recently which implies the existence of non–vanishing neutrino masses. Present direct upper limits on the neutrino masses are:

$$\begin{array}{lcl}
m_{\nu_e} < & 3.0 \text{ eV} & (\text{from } {}^3H \to {}^3He\, e^- \,\bar{\nu}_e) \\
m_{\nu_\mu} < & 190 \text{ keV} & (\text{from } \pi \to \mu\, \nu_\mu) \\
m_{\nu_\tau} < & 18.2 \text{ MeV} & (\text{from } \tau^- \to 3\pi\, \nu_\tau)
\end{array}$$

Lower bounds are not yet so easy to establish at present but observed neutrino mixing phenomena indicate values of about two to three orders of magnitude lower than the above direct upper limits. In any case this implies corresponding lepton numbers L_ℓ ($\ell = e, \mu, \tau$)–violations.

Another important limit is the absence of $\Delta L_e = 2$ transitions. The limit from neutrino-less double beta decay $(Z, A) \to (Z+2, A) + e^+ + e^-$ is $t_{1/2} > 1.6 \times 10^{25}$ years for ^{76}Ge . The observation of such reactions would imply that the electron-neutrino is a massive **Majorana neutrino**, a self-conjugate fermion which is its own antiparticle.

Global non-Abelian symmetries are approximate (broken) only and correlated with the hierarchy of the fundamental interactions. The weaker the interaction the less symmetries it respects. Strong interaction symmetries are:

$$\left.\begin{array}{l}\left.\begin{array}{ll}\text{Isospin (charge independence of nucleon forces) } I & SU(2)_I \\ \text{Strangeness } S & \end{array}\right\} SU(3)_{\text{flavor}} \\ \text{Charm } C \end{array}\right\} SU(4)_{\text{flavor}}$$

$$\vdots$$

The larger the symmetry group the stronger it is broken by growing mass differences of the states in the multiplets. These symmetries are furthermore broken by electromagnetic and weak interactions. The latter in addition breaks parity P maximally and CP in accordance with the Cabibbo-Kobayashi-Maskawa (CKM) mixing scheme of the three quark–lepton family electroweak SM. CP violation was observed for the first time in K^0 decays in 1964 as a small effect at the 3 ppm level. The corresponding effect in B^0 decays has been established 2001 at dedicated B–factories.

□

Excursion on the *Quantum Chromodynamics* (QCD) and the *chiral group*: The modern theory of strong interaction is QCD. It views hadrons like the pions, the nucleons etc. as composite objects made out of quarks. Mesons are quark antiquark bound states, nucleons are three quark bound states etc. The quarks not only carry flavor quantum numbers like isospin, strangeness, charm, etc. but an additional one called *color*. More precisely quarks are color triplets antiquarks are color antitriplets of $SU(3)_c$. Color is a local symmetry (much like the local $U(1)_{\rm em}$ gauge invariance in QED) and requires the existence of eight colored gauge bosons the *gluons* which glue together the quarks in the hadrons. QCD is a unbroken non-Abelian gauge theory, which has the property of *asymptotic freedom*, the strength of the interaction becomes weaker and weaker as we look at shorter and shorter distances, i.e., inside the hadrons. Complementary it becomes stronger and stronger as we go to larger and larger distances. This means that we cannot separate the quarks in the hadrons to become free particles. If we try to separate a $q\bar{q}$–pair the color field between them get squeezed into flux tubes which at the end form strings which execute a linearly rising force. In this way quarks and gluons get in fact permanently confined inside the hadrons. This phenomenon is called *confinement*. Only objects not carrying net color can become free, these are the hadrons. They have typical sizes of about 1 fermi (=10^{-13} cm) and the color forces are screened at distances beyond the size of a hadron. The remnant forces are what we observe as nuclear forces in atomic nuclei or in low energy hadron scattering. Thus in spite of the fact the force carriers, the gluons, are massless the strong interaction forces are rather short ranged. Thus, interestingly, the spectrum of possible states of QCD are not the fields in the Lagrangian, the quarks and gluons, but the hadrons, and those must be color neutral which means they must be color singlets.

Thus from the point of view of the color $SU(3)_c$ the spectrum can be found by determining all possible singlets which we may form from quarks in 3 or 3^* and gluons in 8:

$$\begin{array}{lcl} 3 \otimes 3^* & = & 1 \oplus 8 \\ 3 \otimes 3 \otimes 3 & = & 1 \oplus 8 \oplus 8 \oplus 10 \\ 3^* \otimes 3^* \otimes 3^* & = & 1 \oplus 8 \oplus 8 \oplus 10^* \\ 8 \otimes 8 & = & 1 \oplus 8 \oplus 8 \oplus 10 \oplus 10^* \oplus 27 \end{array}$$

Only the singlets play a role here. They correspond to the mesons $\bar{\psi}_{i_c\alpha}(\Gamma)_{\alpha\beta}\psi_{i_c\beta}$ baryons, antibaryons and the so called *glueballs*, respectively. The glueballs, the singlet in $8 \otimes 8$ of gluons, do not contain any valence quarks. They are expected to show up as broad resonances between 1 and 2 GeV and have not yet been established by experiments.

A completely different kind of application concern the flavor symmetries, which are approximate global symmetries of the strong interactions and hence, in modern terminology, of the QCD Lagrangian. Since some of the quarks are rather light one may look at the approximation where quark masses are switched off. The QCD Lagrangian then has a huge global flavor symmetry, namely, the **chiral group**

$$\mathcal{G}_F = U(N_F)_V \otimes U(N_F)_A \simeq SU(N_F)_V \otimes SU(N_F)_A \otimes U(1)_V \otimes U(1)_A \tag{5.1}$$

with N_F the number of quark flavors. In the context of the SM, as we shall see later, this corresponds to the symmetric phase of the electroweak– or flavor– sector of the SM, i.e., before the local gauge symmetry $SU(2)_L \otimes U(1)_Y$ is broken by the Higgs mechanism to the residual exact electromagnetic $U(1)_{\mathrm{em}}$ local gauge symmetry. The unbroken phase may be understood as an *asymptotic symmetry* which is approached asymptotically at very high energies, when all masses which are small relative to a given energy scale are negligible. Since as we increase the number of flavors from $N_F = 2$ to 6, the above symmetry is broken more and more by increasingly heavy quark masses[23]

quark flavor	u	d	s	c	b	t
mass (MeV)	∼5	∼9	190	1650	4750	174200

the chiral symmetry is good only for the light flavors: for $N_F = 2$ we have with very good accuracy the isospin $SU(2)$ for $N_F = 3$ the slightly more broken $SU(3)$ of isospin plus strangeness, symmetries which manifest themselves in the hadronic spectrum. One might ask what is the precise sense of a *broken symmetry*, a symmetry which is not truly a symmetry? Associated with a symmetry there are currents and generalized charges, the generators of the symmetry transformations (see next Sec.) of particle multiplets. The point is that in spite of the fact that the symmetry is not perfect the states may be classified or labeled in terms of corresponding quantum numbers like isospin and stangeness which satisfy the appropriate Lie algebra. As the symmetry is approximate only, the corresponding charges are not strictly conserved and thus are time-dependent to some extent.

Surprisingly the chiral symmetry group (5.1) is not just $SU(N_F)$ which is what we observe in the hadron spectrum. The chiral group is doubled by the axial part, in fact in the massless limit left–handed and right–handed fields satisfy an $SU(N_F)$ independently and thus we obtain $SU(N_F)_L \otimes SU(N_F)_R$ which is equivalent to vector times axial-vector $SU(N_F)_V \otimes SU(N_F)_A$ (see Sec. 12 for more details). The only explanation for the absence of the parity doublers in the spectrum is that in fact the $SU(N_F)_A$ is spontaneously broken, i.e., the symmetry is manifest in the dynamics, represented by the massless QCD–Lagrangian, but absent in the ground state (vacuum) and hence in the space of states built up about the non–symmetric vacuum (see Sec. 9). In turn this implies the existence of a set of Goldstone bosons, which must be massless: for $SU(2)$ we must have 3 Goldstone bosons the three pions $\pi^\pm, \pi^0$, for $SU(3)$ we must have 8 Goldstone bosons the pions plus $K^\pm$, K^0, $\bar{K}^0$ and η. Since in reality quark masses are non-vanishing the "would be Goldstone bosons" aquire a mass and one calls them *preudo Goldstone bosons*. This mechanism explains why the pseudoscalar mesons are the lightest hadrons and why they have masses substantially lower than the other hadrons.

We finally mention that the $U(1)_V$ factor in (5.1) in fact corresponds to the baryon number conservation, while $U(1)_A$ is not a symmetry at all. The latter is broken by quantum corrections, the famous Adler–Bell–Jackiw anomaly.

For a more detailed discussion we refer to Sec. 12 for the chiral group and to Sec. 9 for the spontaneous symmetry breaking and the Goldstone phenomenon.

□

We have seen that we are often dealing with imperfect symmetries in nature. The various possibilities we have encountered may be classified as follows:

[23]Note the in each doublet the quarks with the larger charge magnitude like c and t have also larger mass than s and b, respectively. The lightest two quarks are an exception, the u quark is lighter than the d quark which is of existential importance as it makes the proton to be lighter than the neutron and the neutron to decay into protons and not vice versa. Thus the inversion is crucial for the stability of the proton and hence for all structure in the universe.

- Symmetries broken by weaker interactions: What are the enhanced symmetries we get if we switch of gravity, the weak forces and the electromagnetic interactions?
- Symmetries broken by mass terms or other terms in the Lagrangian of dimension less than four. In general such breakings disappear if we go to higher energies only if they are generated by
- spontaneous symmetry breaking.
- What is barely considered in the literature: symmetries may show up at low energies because one does not see the short distance details, which means that at high enough energies when probing short distances symmetries observed at low energy may be violated. Typically, interaction terms of dimension larger than four (non-renormalizable terms which naturally arise in low energy effective theories but which are suppressed if the high-energy cut-off is large enough) could violate symmetries we see at low energies. Low energy here means present accelerator energies as we expect the Planck scale $M_{\mathrm{Planck}} \sim 10^{19}$ GeV to be the fundamental reference scale. Such non-renormalizable terms are absent in the SM and possibly only show up when we go much closer to the Planck scale M_{Planck}. Effects at the 1 ppm level we may expect only at energies $E/M_{\mathrm{Planck}} \sim 10^{-3}$, i.e., $E \sim M_{\mathrm{GUT}} \sim 10^{16}$GeV.

Appendix to Section 5:

Some useful formulae for matrix transformations

1.

$$\boxed{\begin{aligned} e^{iA}Be^{-iA} &= B + i[A,B] + \tfrac{i^2}{2!}[A,[A,B]] + \dots \\ &\quad + \tfrac{i^n}{n!}[A,[A,\dots,[A,[A,B\underbrace{]]\dots]}_{n} + \dots \end{aligned}}$$

holds for any two operators or matrices A and B .

Proof:

Replace A by λA and perform a Taylor expansion in λ

$$F(\lambda) = e^{i\lambda A}Be^{-i\lambda A} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \left(\frac{\partial^n F}{\partial \lambda^n} \right) |_{\lambda=0}$$

and evaluate the Taylor coefficients. Using that A commutes with $e^{i\lambda A}$ we get

$$\frac{\partial F}{\partial \lambda} = e^{iA}\; i\,[A,B]e^{-i\lambda A}$$

and by repeated differentiation

$$\frac{\partial^n F}{\partial \lambda^n} = e^{i\lambda A}\; i^n\; [A,A,\dots[A,B\underbrace{]]\dots]}_{n}\, e^{-i\lambda A}.$$

For $\lambda = 1$ the result follows.

2.

$$e^{i\sum_l T_l\omega_l}\,T_i\,e^{-i\sum_l T_l\omega_l} = T_k\left(e^{i\sum_l t_l\omega_l}\right)_{ki}$$

i.e. T_i transforms as a vector under the adjoint representation $(t_l)_{ik} = -i\,c_{lik}$ or $(t_l)_{ki} = i\,c_{lik}$.

Proof:

$$\begin{aligned}
& e^{i\sum_l T_l\omega_l}\,T_i\,e^{-i\sum_l T_l\omega_l} \\
= \;& T_i + i\,[T_l\omega_l, T_i] + \frac{i^2}{2!}\,[T_{l_2}\omega_{l_2}\,[T_{l_1}\omega_{l_1}, T_i]] + \ldots \\
& + \frac{i^n}{n!}[T_{l_n}\omega_{l_n}, [T_{l_{n-1}}\omega_{l_{n-1}}, \ldots, [T_{l_1}\omega_{l_1}, T_i\underbrace{]\ldots]]}_{n} + \ldots \\
= \;& T_k\delta_{ki} + T_k\, i\;(t_l\omega_l)_{ki} + T_k\frac{i^2}{2!}\,(t_l\omega_l)^2_{ki} + \ldots + T_k\frac{i^n}{n!}\,(t_l\omega_l)^n_{ki} + \ldots \\
= \;& T_k\left(e^{i\sum_l t_l\omega_l}\right)_{ki}
\end{aligned}$$

We have used 1. and $[T_l, T_i] = T_k\, i\, c_{lik} = T_k(t_l)_{ki}$ such that

$$\begin{aligned}
[T_l\omega_l, T_i] &= T_k(t_l\omega_l)_{ki} \\
[T_{l_2}\omega_{l_2}, [T_{l_1}\omega_{l_1}, T_i]] &= [T_{l_2}\omega_{l_2}, T_{k_1}]\,(t_{l_1}\omega_{l_1})_{k_1 i} \\
&= T_k\,(t_{l_2}\omega_{l_2})_{kk_1}\,(t_l, \omega_{l_1})_{k_1 i} \\
&= T_k(t\omega)^2_{ki} \quad etc.
\end{aligned}$$

where repeated indices have to be summed over.

3.

$$e^{-i\sum_l T_l\omega_l}\partial_\mu\left(e^{i\sum_l T_l\omega_l}\right) = T_k\left(\frac{1-e^{-i\sum_l t_l\omega_l}}{\sum_l t_l\omega_l}\right)_{ki}\partial_\mu\omega_i$$

is a linear combination of the generators and thus an element of the Lie-algebra

Proof:

$$\begin{aligned}
& e^{-i\sum_l T_l\omega_l}\partial_\mu\left(e^{i\sum_l T_l\omega_l}\right) \\
= \;& -i\,[T_l\omega_l, \partial_\mu] + \frac{i^2}{2!}\,[T_{l_2}\omega_{l_2}, [T_{l_1}\omega_{l_1}, \partial_\mu]] + \ldots \\
& + \frac{(-i)^n}{n!}\,[T\omega, [T\omega, \ldots, [T\omega, \partial_\mu]\ldots]] + \ldots \\
= \;& -i\,T_k\,\delta_{ki}\,(-\partial_\mu\omega_i) + \frac{i^2}{2!}\,[T_l\omega_l, T_i]\,(-\partial_\mu\omega_i) + \ldots \\
& + \frac{(-i)^n}{n!}\,[T\omega, [T\omega, \ldots, [T\omega, T_i\underbrace{]\ldots]]}_{n-1}(-\partial_\mu\omega_i) + \ldots \\
= \;& -T_k\sum_{n=1}^{\infty}\frac{(-i)^n}{n!}\,(t\omega)^{n-1}_{ki}\partial_\mu\omega_i \\
= \;& T_k\left(\frac{1-e^{-i\sum_l t_l\omega_l}}{\sum_l t_l\omega_l}\right)_{ki}\partial_\mu\omega_i = T_k\Lambda_{ki}(\omega)\,\partial_\mu\omega_i
\end{aligned}$$

We have used 1. and $[T_l\omega_l, \partial_\mu] = T_l\omega_l\partial_\mu - \partial_\mu T_l\omega_l = -T_l(\partial_\mu\omega_l)$ and then proceed as in the proof of 2.

Exercises: Section 5

① Show that $8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus 10^* \oplus 27$.

② Discuss the isospin properties of the triplet of pions (π^+, π^0, π^-) .

The isospin symmetry of the scattering operator S not only leads to relations between matrix elements but also to selection rules: Suppose

(a) T is a generator of a symmetry transformation such that $[T, S] = 0$,

(b) $\mid \alpha >$ and $\mid \beta >$ are eigenstates of T i.e. $T \mid \alpha >= t_\alpha \mid \alpha >, T \mid \beta >= t_\beta \mid \beta >$

What does this imply for the S-matrix elements

$$S_{\beta\alpha} =< \beta \mid S \mid \alpha > \ ?$$

Find a few examples.

③ Use the Young tableaux to construct the meson states in

$$3 \otimes 3^*$$

and the baryon states in

$$3 \otimes 3 \otimes 3 \ .$$

The states in the pseudoscalar meson octet of flavor $SU(3)$ are characterized by the 3^{rd} component of isospin and by hypercharge $Y = B + S$ (B baryon number $B = 0$ for mesons, S strangeness $S = 0$ for pions). Display the weight diagram ($I_3 - Y$ plot) of the meson states. How are they composed of u, d and s quarks in the $SU(3)_{\text{flavor}}$ quark model ?

④ The structure constants c_{ikl} of a Lie-algebra $[T_i, T_k] = ic_{ikl}T_l$ satisfy the Jacobi identity.

$$c_{ikn}c_{nlm} + \text{ terms cyclic in } (ikl) = 0$$

Use this to show that $(\tilde{T}_i)_{kl} = -ic_{ikl}$ also satisfies the Lie-algebra (adjoint representation).

⑤ Lepton number L_e is another additive quantum number which is strictly conserved. $L_e(e^-) =$

1 by convention. Determine L_e for the other particles from the observed reactions:

1. $L_e(e^+) = -1,\ L_e(\gamma) = 0$:

$$\begin{aligned} p + e &\rightarrow p + e + \gamma \\ \gamma^* &\rightarrow e^+ + e^- \end{aligned}$$

2. $L_e(\pi^0) = L_e(\pi^\pm) = 0$:

$$\begin{aligned} \pi^0 &\rightarrow 2\gamma,\ \gamma + e^+ + e^- \\ p + \pi^- &\rightarrow n + \pi^0 \\ p + \pi^0 &\rightarrow n + \pi^+ \end{aligned}$$

3. $L_e(\bar{\nu}_e) = -1,\ L_e(\nu_e) = 1$:

$$\begin{aligned} \pi^- &\rightarrow e^- + \bar{\nu}_e \\ \pi^+ &\rightarrow e^+ + \nu_e \end{aligned}$$

From the last two reactions we learn the important result $\nu_e \neq \bar{\nu}_e$!

⑥ Baryon number conservation is responsible for the stability of the proton. By convention $B(p) = 1,\ B(e^-) = 0$. Determine the baryon numbers of particles from the observation of the following reactions:

a.) Baryons and mesons:

1. $B(\pi^0) = 0$:

$$p + p \rightarrow p + p + \pi^0$$

2. $B(n) = B(p),\ B(\pi^\pm) = B(\pi^0) = 0$:

$$\begin{aligned} p + p &\rightarrow p + n + \pi^+ \\ \pi^- + p &\rightarrow n + \pi^0 \end{aligned}$$

3. $B(K^\pm) = B(K^0) = 0$:

$$\begin{aligned} K^\pm &\rightarrow \pi^\pm + \pi^0 \\ K^0 &\rightarrow \pi^+ + \pi^-,\ \pi^+ + \pi^- + \pi^0 \end{aligned}$$

4. $B(\Lambda),\ B(\Sigma) = 1$:

$$\begin{aligned} \pi^- + p &\rightarrow \Lambda^0 + K^0,\ \Sigma^- + K^+ \\ \pi^+ + p &\rightarrow \Sigma^+ + K^+,\ \Sigma^0 + \Lambda^0 \end{aligned}$$

5. $B(\Xi),\ B(\Omega^-) = 1$:

$$K^- + p \rightarrow \Xi^- + K^+,\ \Xi^0 + K^0,\ \Omega^- + K^+ + K^0$$

b.) Antibaryons:

6. $B(\bar{p}) = -1$:

$$p + p \quad \rightarrow \quad p + p + p + \bar{p}$$

7. $B(\bar{B}) = -1$:

$$p + \bar{p} \quad \rightarrow \quad \bar{n} + n,\ \bar{\Lambda}^0 + \Lambda^0,\ \bar{\Sigma}^0 + \Sigma^0,\ \bar{\Sigma}^{\pm} + \Sigma^{\mp},\ \bar{\Xi}^{+} + \Xi^{-}$$

c.) Photon:

8. $B(\gamma) = 0$:

$$p \quad \rightarrow \quad p + \gamma$$

d.) Leptons: All leptons are produced in pairs, $B(e^-) = 0$ by convention.

9. $B(e) = B(\mu) = 0$:

$$\gamma^* \quad \rightarrow \quad e^+ + e^-,\ \mu^+ + \mu^-$$

10. $B(\nu_e) = B(\nu_\mu) = 0$:

$$\begin{aligned} n &\rightarrow p + e^- + \bar{\nu}_e \\ \mu^- &\rightarrow e^- + \bar{\nu}_e + \nu_\mu \\ \mu^+ &\rightarrow e^+ + \nu_e + \nu_\mu^- \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu \end{aligned}$$

6 Local gauge invariance, Yang-Mills theories

In QED we learn that a photon field (Abelian gauge field) is required, which minimally couples to electrons (charged particles), if we "enlarge" the global $U(1)$ gauge symmetry of the free electron system to a local gauge symmetry (Weyl 1929). **Yang-Mills fields** (**non-Abelian gauge fields**) are needed if we require a global $SU(n)$ symmetry of free matter fields to be promoted to a local $SU(n)$ gauge symmetry (Yang and Mills 1954).

As we shall see, the presence of global symmetries on the one hand and of local symmetries on the other hand has very different physical consequences. Astonishingly, according to our present knowledge, all fundamental interactions of quarks and leptons derive from a gauge principle with respect to a particular gauge group.

6.1 Global symmetries and Noether currents

We consider a multiplet of n free spin 1/2 matter fields

$$\Psi(x) := \quad \Psi_{\alpha a}(x) = \begin{pmatrix} \psi_{\alpha 1} \\ \vdots \\ \psi_{\alpha n} \end{pmatrix} .$$

Each $\psi_{\alpha a}(a = 1, \ldots, n)$ is a Dirac field with spinor index $\alpha = 1, \ldots, 4$. The free Lagrangian of the field $\Psi(x)$ is given by

$$\mathcal{L}_0^{\Psi} = \bar{\Psi}(x)\,(i\gamma^\mu\partial_\mu - m)\,\Psi(x) \equiv \sum_a \bar{\Psi}_{\alpha a}(x)\,(i\gamma^\mu\partial_\mu - m)_{\alpha\beta}\,\Psi_{\beta a}(x)$$

and the Euler-Lagrange equations of motion

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \bar{\Psi}} - \frac{\partial \mathcal{L}}{\partial \bar{\Psi}} = 0 \quad \text{and} \quad \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \Psi} - \frac{\partial \mathcal{L}}{\partial \Psi} = 0$$

are equivalent to n uncoupled Dirac equations

$$(i\gamma^\mu\partial_\mu - m)\,\Psi_a = 0\ ; \quad (a = 1, \ldots, n)\ .$$

This Lagrangian has global $U(n)$ symmetry

$$\Psi \to \Psi' = \tilde{U}\Psi\ ; \quad \tilde{U} \in U(n).$$

Since $\tilde{U} = Ue^{i\phi}$ with $U \in SU(n)$ we may restrict ourselves to consider $SU(n)$ transformations only, because the phase is common to all fields in the multiplet. Accordingly, in the following, we assume the field Ψ to transform as a **vector** in the fundamental representation of $SU(n)$. It will be convenient to write the $SU(n)$ matrices in the form

$$U = U(\omega) = e^{ig\sum_{i=1}^{n^2-1} T_i\omega_i}$$

with a common real positive scale factor g (it will be identified with the gauge coupling later on) split off from the group parameters ω_i. For **global** transformations the parameters ω_i are independent of the space-time point x.

By Noether's theorem the global $SU(n)$ symmetry of $\mathcal{L}_0^\Psi$ implies the existence of $r = n^2 - 1$ conserved currents $j_i^\mu(x)$ and the "charges" $Q_i = \int d^3x j_i^0(t,\vec{x})$ are time independent and represent the generators of the $SU(n)$ Lie-algebra $[Q_i, Q_k] = ic_{ikl}Q_l$. Noether's theorem and the specific form of the currents can be obtained by applying an infinitesimal transformation

$$\begin{aligned}\Psi \to \Psi' = \Psi + \delta\Psi \quad &; \quad \delta\Psi = \quad ig\sum_i T_i\Psi\delta\omega_i \\ \bar{\Psi} \to \bar{\Psi}\,' = \bar{\Psi} + \delta\bar{\Psi} \quad &; \quad \delta\bar{\Psi} = -ig\sum_i \bar{\Psi}T_i\delta\omega_i\end{aligned}$$

to the field and by using the invariance of $\mathcal{L}_0^\Psi$.

The variations of the fields imply a variation of $\mathcal{L}_0^\Psi$ by

$$\delta\mathcal{L}_0^\Psi = \delta\bar{\Psi}\,\frac{\partial\mathcal{L}}{\partial\bar{\Psi}} + \delta\left(\partial_\mu\bar{\Psi}\right)\frac{\partial\mathcal{L}}{\partial\partial_\mu\bar{\Psi}} + \frac{\partial\mathcal{L}}{\partial\Psi}\,\delta\Psi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\Psi}\,\delta\left(\partial_\mu\Psi\right) \quad .$$

Since the transformation is global (x-independent) $\delta\partial_\mu\Psi = \partial_\mu\delta\Psi$ and by the Euler-Lagrange equation $\frac{\partial\mathcal{L}}{\partial\Psi} = \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\Psi}$ we find

$$\begin{aligned}\delta\mathcal{L}_0^\Psi &= \partial_\mu\left\{\delta\bar{\Psi}\frac{\partial\mathcal{L}}{\partial\partial_\mu\bar{\Psi}} + \frac{\partial\mathcal{L}}{\partial\partial_\mu\Psi}\delta\Psi\right\} \\ &= -g\sum_i\left(\partial_\mu j_i^{\mu\Psi}\right)\delta\omega_i\end{aligned}$$

where

$$j_i^{\mu\Psi} = \bar{\Psi}(x)\gamma^\mu T_i\Psi(x) \equiv \bar{\Psi}_{\alpha a}(x)\left(\gamma^\mu\right)_{\alpha\beta}\left(T_i\right)_{ab}\Psi_{\beta b}(x)$$

are the fermionic $SU(n)$ Noether currents.

Since $\delta\mathcal{L}_0^\Psi = 0$ for arbitrary $\delta\omega_i$, we indeed must have the currents being conserved

$$\partial_\mu j_i^{\mu\Psi}(x) = 0\;;\quad i = 1,\ldots,n^2-1\;.$$

For a multiplet of complex scalar fields

$$\Phi_a = \begin{pmatrix}\phi_1\\ \vdots \\ \phi_n\end{pmatrix}$$

and a free Lagrangian

$$\mathcal{L}_0^\Phi = \left(\partial_\mu\Phi\right)^+\left(\partial^\mu\Phi\right) - \mu^2\Phi^+\Phi$$

we find the Noether currents

$$j_i^{\mu\Phi} \quad = \quad i\,\Phi^+T_i\overset{\leftrightarrow}{\partial^\mu}\,\Phi$$

where

$$f\overset{\leftrightarrow}{\partial_\mu}g = f(\partial_\mu g) - (\partial_\mu f)g\;.$$

It is important to notice that the form of the conserved currents does not change if symmetric interactions without derivatives of fields are present.

6.2 Local symmetries and gauge fields

While global symmetries give raise to a classification of states according to irreducible representations of the symmetry group and to conserved quantum numbers, local symmetries have a rather different physical implication. The requirement of local symmetries is a **dynamical principle** which implies that matter fields **must** be in interaction with massless spin 1 gauge fields in a specific way. The property of local gauge invariance corresponds to the validity of an **equivalence principle:** the n degrees of freedom of the local field $\Psi_a(x)$ are locally indistinguishable in the sense that $\Psi(x)$ and $U(x)\Psi(x)$ describe the same physics and

$$\Psi(x) \to \Psi'(x) = U(x)\Psi(x)\ , \quad U(x) \in SU(n)$$

is a symmetry of the dynamics. This means that an observer at space-time point x may choose a coordinate frame for the internal degrees of freedom in the multiplet $\Psi_a(x)$ independently from an observer at a different space-time point x'. At first sight this requirement might look quite natural and harmless. It has dramatic consequences, however. Essentially, it dictates the form of the dynamics once the local transformation laws of the matter fields are known. The equivalence principle for internal symmetries is very similar to the classical equivalence principle of gravity, which implies that gravity emerges as the geometry of space-time.

The freedom to have associated with each space-time point an independent frame for the internal symmetry space only makes sense if we are able to "synchronize" local frames at different space-time points. In order to be able to actually perform such a synchronization we need carriers of physical signals traveling at the universal speed of light, namely, $n^2 - 1$ massless spin 1 bosons described by a matrix $\boldsymbol{V}_\mu(x) = \sum_i T_i V_{\mu i}(x)$ which is an element of the Lie-algebra. A change of local frames between space-time points x and $x + dx$ must be correlated with a local gauge transformation of $\boldsymbol{V}_\mu(x)$. What we need is a statement saying when "a field $\Psi(x)$ does not change between x and $x+dx$". With other words, we need a definition of **parallel displacement**. When there are no internal symmetries and $\Psi(x)$ is a real field we would say that the field does not change between x and $x + dx$ if

$$\Psi(x + dx) = \Psi(x) + \partial_\mu \Psi(x)\, dx^\mu = \Psi(x)$$

or $\partial_\mu \Psi(x) = 0$. In QED the complex electron field $\psi(x)$ is coupled to the photon field $A_\mu(x)$, a $U(1)$ gauge field, and in the coupled electron-photon system the phase $\psi(x) = e^{ie\alpha(x)} \mid \psi(x) \mid$ has no physical significance. However, the statement that $\mid \psi(x) \mid$ is constant over dx:

$$\partial_\mu \mid \psi(x) \mid = \partial_\mu \left(e^{-ie\alpha(x)} \psi(x) \right) = 0$$

or $(\partial_\mu - ie\partial_\mu \alpha(x))\psi(x) = 0$ is not gauge invariant and must be replaced by the condition that the **covariant derivative**

$$D_\mu \psi(x) = (\partial_\mu - ieA_\mu(x))\, \psi(x) = 0$$

vanishes. For the non-Abelian $SU(n)$ gauge symmetry this generalizes to

$$\left(D_\mu \Psi(x)\right)_a = \left(\partial_\mu - ig\boldsymbol{V}_\mu(x)\right)_{ab} \Psi_b(x) = 0$$

In Fig. 6.1 we have illustrated the geometrical meaning of the covariant derivative. For infinitesimal dx we compare the fields $\Psi(x)$ at point x and $\Psi(x + dx)$ at point $x + dx$. Consider the covariant expansion $\Psi(x + dx) = \Psi(x) + D_\mu \Psi(x)\, dx^\mu$. If $\Psi(x)$ satisfies $D_\mu \Psi(x) = 0$ the field is "parallel". We denote by $\Psi^{||}_{(dx)}(x)$ the field $\Psi(x+dx)$ which has been shifted parallel from $x+dx$ to x along the path dx. Then $D_\mu\Psi(x)\, dx^\mu = \Psi^{||}_{(dx)}(x) - \Psi(x)$.

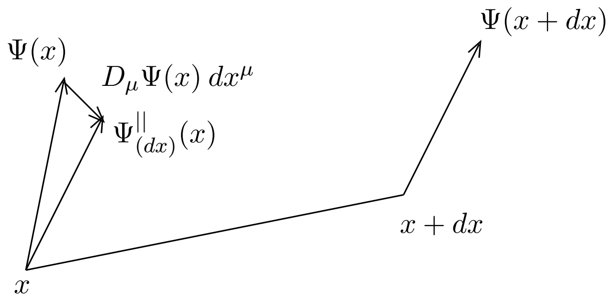


Fig. 6.1: Geometrical interpretation of the parallel displacement.

We now discuss in more detail how locally gauge invariant field theories can be constructed.

6.2.1 Minimal couplings of the matter fields

Global G $(= SU(n))$ invariance of $\mathcal{L}_0^\Psi$ follows from the fact that with $\Psi(x)$ also $\partial_\mu\Psi(x)$ transforms as a vector:

$$\begin{aligned}\Psi(x) &\rightarrow \Psi'(x) = U\Psi(x)\\ \partial_\mu\Psi(x) &\rightarrow \partial_\mu\Psi'(x) = U\partial_\mu\Psi(x)\end{aligned}$$

when $\partial_\mu U = 0$. Under local transformations $\partial_\mu U \neq 0$

$$\begin{aligned}\partial_\mu\Psi(x) \rightarrow \partial_\mu\Psi'(x) &= U(x)\partial_\mu\Psi(x) + (\partial_\mu U(x))\,\Psi(x)\\ &= U(x)\left(\partial_\mu + U^{-1}(x)\,(\partial_\mu U(x))\right)\Psi(x)\\ &\neq U(x)\partial_\mu\Psi(x)\end{aligned}$$

does no longer transform vector-like because of the extra term

$$U^{-1}(x)\,(\partial_\mu U(x)) = ig\sum_i T_i\tilde{V}_{i\mu}\ \in\ G'.$$

As indicated, this term is an element of the Lie-algebra G' of the symmetry group G and a four-vector under Lorentz transformation. Neglecting higher order terms, for infinitesimal transformations $U = 1 + ig\sum_i T_i\omega_i$ we easily calculate

$$U^{-1}(x)\,(\partial_\mu U(x)) = ig\sum_i T_i\partial_\mu\omega_i(x)$$

such that $\tilde{V}_{i\mu}(x) = \partial_\mu\omega_i(x)$. For finite transformations we may write $\tilde{V}_{i\mu}(x) = \sum_l \Lambda_{il}(\omega)\partial_\mu\omega_l(x)$ where the matrix $\Lambda_{il}(\omega)$ is given in the Appendix.

When applied to $\mathcal{L}_0^\Psi$ local gauge transformations induce a non-invariant term:

$$\begin{aligned}\mathcal{L}_0^\Psi &\rightarrow \mathcal{L}_0^\Psi + \bar{\Psi}i\gamma^\mu U^{-1}\,(\partial_\mu U)\,\Psi\\ &= \mathcal{L}_0^\Psi - g\sum_i \bar{\Psi}\gamma^\mu T_i\Psi\tilde{V}_{i\mu}(x)\\ &= \mathcal{L}_0^\Psi - g\sum_i j_i^\mu(x)\tilde{V}_{i\mu}(x)\ .\end{aligned}$$

This term describes the coupling of a set of $r = n^2 - 1$ real vector fields $\tilde{V}_{i\mu}(x)$ to the Noether currents $j_i^\mu(x)$, which are conserved under global transformations.

In order to obtain a locally gauge invariant **extension** of the free system we must introduce a set of r real vector fields $V_{i\mu}(x)$ as dynamical variables (physical degrees of freedom) which couple to the Noether currents:

$$
\begin{aligned}
\mathcal{L}_0^\Psi \to \mathcal{L}^\Psi &= \mathcal{L}_0^\Psi + g \sum_i j_i^\mu(x) V_{i\mu}(x) \\
&= \mathcal{L}_0^\Psi + g \sum_i \bar{\Psi} \gamma^\mu T_i V_{i\mu}(x) \Psi \\
&= \bar{\Psi} \left(i\gamma^\mu D_\mu - m \right) \Psi \ .
\end{aligned}
$$

The fields $V_{i\mu}(x)$ are called **Yang-Mills fields** or **non-Abelian gauge fields**.

Formally $\mathcal{L}^\Psi$ follows from $\mathcal{L}_0^\Psi$ by **minimal substitution**

$$
\partial_\mu \to D_\mu = \partial_\mu - ig \, \boldsymbol{V}_\mu
$$

where

$$
\boldsymbol{V}_\mu = \sum_i T_i V_{i\mu}(x) \ \in G'
$$

is an element of the Lie-algebra.

D_μ defines the **covariant derivative**. $\mathcal{L}^\Psi$ is locally gauge invariant provided $D_\mu \Psi$ transforms as a vector

$$
\begin{aligned}
\Psi(x) &\to \Psi'(x) = U(x)\Psi(x) \\
D_\mu \Psi(x) &\to (D_\mu \Psi)' = U(x) D_\mu \Psi(x)
\end{aligned}
$$

and hence

$$
D'_\mu U(x) = U(x) D_\mu \ .
$$

This condition fixes the transformation law of the fields $V_{i\mu}(x)$:

$$
\begin{aligned}
D'_\mu &= \partial_\mu - ig\boldsymbol{V}_\mu{}' = U(x) D_\mu U^{-1}(x) \\
&= U(x) \left(\partial_\mu - ig\boldsymbol{V}_\mu \right) U^{-1}(x) \\
&= \partial_\mu - igU(x)\boldsymbol{V}_\mu U^{-1}(x) + U(x)\partial_\mu U^{-1}(x) \\
&= \partial_\mu - igU(x) \left(\boldsymbol{V}_\mu - \frac{i}{g} U^{-1}(\partial_\mu U) \right) U^{-1}(x) \ .
\end{aligned}
$$

Here, we have used $\partial_\mu \left(UU^{-1} \right) = \left(\partial_\mu U \right) U^{-1} + U \left(\partial_\mu U^{-1} \right) = 0$ or $U \left(\partial_\mu U^{-1} \right) = - \left(\partial_\mu U \right) U^{-1}$.

Consequently, we find

$$
\boldsymbol{V}_\mu \to \boldsymbol{V}_\mu{}' = U(x) \left(\boldsymbol{V}_\mu - \frac{i}{g} U^{-1}(\partial_\mu U) \right) U^{-1}(x) \ .
$$

Like $\partial_\mu \Psi(x)$, $\boldsymbol{V}_\mu(x)$ does not transform as a vector since the local transformation law is different from the global one. In fact $\boldsymbol{V}_\mu$ has been **required** to produce a compensating term for the non-covariant term obtained for $\partial_\mu \Psi$, in order that $(\partial_\mu - ig\boldsymbol{V}_\mu)\Psi$ is a vector.

For infinitesimal transformations $U = 1 + ig \sum_i T_i \omega_i(x)$ we obtain to linear order in $\omega_i(x)$:

$$\begin{aligned} \boldsymbol{V}_\mu{}' = \sum_i T_i V'_{i\mu}(x) &= \sum_i T_i \left(V_{i\mu}(x) + \partial_\mu \omega_i(x)\right) + ig \sum_{k,\,l} [T_k, T_l]\, \omega_k V_{l\mu} \\ &= \sum_i T_i \left(V_{i\mu}(x) - g c_{ikl} V_{l\mu} \omega_k + \partial_\mu \omega_i\right) \end{aligned}$$

or

$$V'_{i\mu} = V_{i\mu} - g c_{ikl} V_{l\mu}\, \omega_k + \partial_\mu \omega_i \; .$$

Since $(\tilde{T}_k)_{il} = i c_{ikl}$ represent the generators in the adjoint representation we have

$$\begin{aligned} \delta V_{i\mu} &= -g c_{ikl} V_{l\mu}\, \omega_k + \partial_\mu \omega_i \\ &= ig(\tilde{T}_k)_{il} V_{l\mu}\, \omega_k + \partial_\mu \omega_i \end{aligned}$$

which compares to

$$\delta \Psi_a = ig(T_k)_{ab} \Psi_b \omega_k$$

for the matter fields. We notice that $V_{i\mu}(x)$ transforms under the adjoint representation up to a divergence term[24]. Accordingly the fields $V_{i\mu}(x)$ carry $SU(n)$ charge, which is obvious also from their coupling to the charged Noether currents.

As a result, local gauge invariance requires the matter fields to interact with $n^2 - 1$ massless gauge fields via minimal coupling

$$\mathcal{L}^\Psi = \bar{\Psi}\left(i\gamma^\mu D_\mu - m\right)\Psi = \mathcal{L}_0^\Psi + g \sum_i \bar{\Psi}\gamma^\mu T_i \Psi V_{i\mu} \; .$$

The gauge coupling constant g is a free parameter. The interaction vertex is depicted in Fig. 6.2.

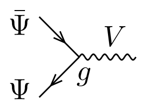


Fig. 6.2: Matter field couplings of a gauge theory.

6.2.2 Non-Abelian field strength tensor

Because the gauge fields $V_{i\mu}(x)$ must be dynamical variables (i.e. describe physical degrees of freedom) the complete Lagrangian density must include a kinetic term for these fields. By their Lorentz structure the $V_{i\mu}(x)$ are gauge potentials describing massless spin 1 fields. Therefore, the Lagrangian must include a term

$$\mathcal{L}_0^V = -\frac{1}{4} \overset{\circ}{G}_{i\mu\nu} \overset{\circ}{G}{}_i^{\mu\nu}$$

[24]As we know from QED, massless four–component gauge fields necessarily exhibit non–physical degrees of freedom because there are only two physical states the transversal one's. Hence, the gauge potentials describe among the physics also redundant stuff. Thats why the transformation laws of the gauge fields under local gauge transformations are anomalous (by the disturbing divergence term). Attempts to describe gauge interactions directly in terms of the more physical field strength tensor (see below) seem not to be possible. By the construction presented before somehow the gauge potentials are quantities which show up in a natural way. As in QED, at the end one has to show that the physical transition matrix elements are gauge invariant and do not depend on the redundancies of the formalism

with $\overset{\circ}{G}_{i\mu\nu} = \partial_\mu V_{i\nu}(x) - \partial_\nu V_{i\mu}(x)$. Like for free matter fields the free Lagrangian $\mathcal{L}_0^V$ is not locally gauge invariant, however. A gauge invariant generalization of $\mathcal{L}_0^V$ is obtained if we substitute $\overset{\circ}{G}_{i\mu\nu}$ by a non-Abelian **field strength tensor** $G_{i\mu\nu}$ which is a 2^{nd} rank antisymmetric Lorentz tensor, transforming as a vector under the adjoint representation of $SU(n)$. This means that

$$\boldsymbol{G}_{\mu\nu} = \sum_i T_i \, G_{i\mu\nu}(x) \ \in \ G'$$

has to transform like

$$\boldsymbol{G}_{\mu\nu} \to \boldsymbol{G}_{\mu\nu}{}' = U(x)\boldsymbol{G}_{\mu\nu}U^{-1}(x)$$

without a divergence term! Since D_μ is a Lorentz vector and satisfies $D'_\mu = U(x)D_\mu U^{-1}(x)$ the commutator $[D_\mu, D_\nu]$ satisfies all the properties required for $\boldsymbol{G}_{\mu\nu}$.

It is now easy to calculate $\boldsymbol{G}_{\mu\nu}$. Using

$$(D_\mu)_{ik} = \partial_\mu \delta_{ik} - ig \sum_j (T_j)_{ik} V_{j\mu}(x)$$

we have

$$\begin{aligned}(D_\mu)_{ik}(D_\nu)_{kl} &= \partial_\mu\partial_\nu\delta_{il} - ig\sum_j (T_j)_{il}\,(V_{j\mu}\partial_\nu + V_{j\nu}\partial_\mu) \\ &- ig\sum_j (T_j)_{il}\partial_\mu V_{j\nu} - g^2 \sum_{j,\,j'} (T_j)_{ik}(T_{j'})_{kl} V_{j\mu}V_{j'\nu}\end{aligned}$$

and hence (the symmetric terms drop)

$$\begin{aligned}[D_\mu, D_\nu]_{il} &= -ig\sum_j (T_j)_{il}\,(\partial_\mu V_{j\nu} - \partial_\nu V_{j\mu}) \\ &\quad -g^2 \sum_{j',\,j''} \left[T_{j'}, T_{j''}\right]_{il} V_{j'\mu}V_{j''\nu} \\ &= -ig\sum_j (T_j)_{il} G_{j\mu\nu} = -ig \ (\boldsymbol{G}_{\mu\nu})_{il}\end{aligned}$$

where, using $[T_{j'}, T_{j''}] = ic_{j'j''j}T_j$,

$$G_{i\mu\nu} = \partial_\mu V_{i\nu} - \partial_\nu V_{i\mu} + g c_{ijk} V_{j\mu} V_{k\nu} \ .$$

This is indeed the gauge covariant generalization of $\overset{\circ}{G}_{i\mu\nu}$. In absence of matter fields the Lagrangian density

$$\mathcal{L}_{YM} = -\frac{1}{4}\sum_i G_{i\mu\nu}G_i^{\mu\nu}$$

defines the so called **pure Yang-Mills theory** .

If $\boldsymbol{V}_\mu = \frac{i}{g}U^{-1}(x)\,(\partial_\mu U(x))$ the field is called a "**pure gauge**" field. For a pure gauge field $\boldsymbol{G}_{\mu\nu} \equiv 0$.

In contrast to the Abelian case (QED), the fields $V_{i\mu}$ and $G_{i\mu\nu}$ transform nontrivially under gauge transformations because they carry nontrivial $SU(n)$-charge. As charged fields they must couple to themselves in order to allow for a locally gauge invariant action. The self-interactions have the following prescribed form:

$$\begin{aligned} \mathcal{L}_{YM} &= -\frac{1}{4}\overset{\circ}{G}_{i\mu\nu}\overset{\circ}{G}{}_i^{\mu\nu} \\ &\quad -\frac{1}{2}g c_{ikl}\left(\partial^\mu V_i^\nu - \partial^\nu V_i^\mu\right) V_{k\mu} V_{l\nu} \\ &\quad -\frac{1}{4}g^2 c_{ikl} c_{ik'l'} V_k^\mu V_l^\nu V_{k'\mu} V_{l'\nu} \end{aligned}$$

The corresponding interaction vertices are shown in Fig. 6.3.

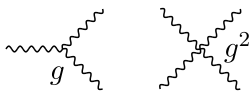


Fig. 6.3: Yang-Mills couplings

If we include the matter fields we have the complete locally gauge invariant Lagrangian density

$$\mathcal{L}_{\text{inv}} = -\frac{1}{4}\sum_i G_{i\mu\nu} G_i^{\mu\nu} + \bar{\Psi}\left(i\gamma^\mu D_\mu - m\right)\Psi$$

with **one** coupling constant g as a free parameter. The strengths of the three different interaction vertices are fixed by the same gauge coupling constant.

Let me add a remark about the geometrical interpretation of the field strength tensor which derives from the one of the covariant derivative (see Fig. 6.1). To this end we consider an infinitesimal parallelogram of points x, $x+dx$, $x+dy$ and $x+dx+dy$ and a field $\Psi(x_i)$ at the different points. In order to shift $\Psi(x+dx+dy)$ parallel to the point x we have two possible paths to follow along the sides of the parallelogram. We denote these paths by (dx,dy) and (dy,dx). The parallel displaced fields are $\Psi^{||}_{(dx,dy)}(x)$ and $\Psi^{||}_{(dy,dx)}(x)$. We now calculate the difference of these two fields. To this end we perform a covariant expansion along the two paths:

$$\begin{aligned} \Psi(x+dx+dy) &= \Psi(x+dy) + D_\mu\Psi(x+dy)\,dx^\mu \\ &= \Psi(x) + D_\nu\Psi(x)\,dy^\nu + D_\mu\Psi(x)\,dx^\mu + D_\mu D_\nu\Psi(x)\,dx^\mu\,dy^\nu \\ \Psi(x+dx+dy) &= \Psi(x+dx) + D_\nu\Psi(x+dx)\,dy^\nu \\ &= \Psi(x) + D_\mu\Psi(x)\,dx^\mu + D_\nu\Psi(x)\,dy^\nu + D_\nu D_\mu\Psi(x)\,dx^\mu\,dy^\nu \end{aligned}$$

For the difference we obtain

$$\Psi^{||}_{(dy,dx)}(x) - \Psi^{||}_{(dx,dy)}(x) = [D_\mu, D_\nu]\Psi(x)\,dx^\mu\,dy^\nu = -ig\boldsymbol{G}_{\mu\nu}\Psi(x)\,dx^\mu\,dy^\nu$$

exhibiting the field strength tensor as a **curvature tensor** . If the field strength is non-vanishing the parallel-displacements of a vector along different paths yield a different result. For infinitesimal shifts the difference vector is proportional to the original vector to the field strength and to the area of the parallelogram. The curvature is illustrated in Fig. 6.4.

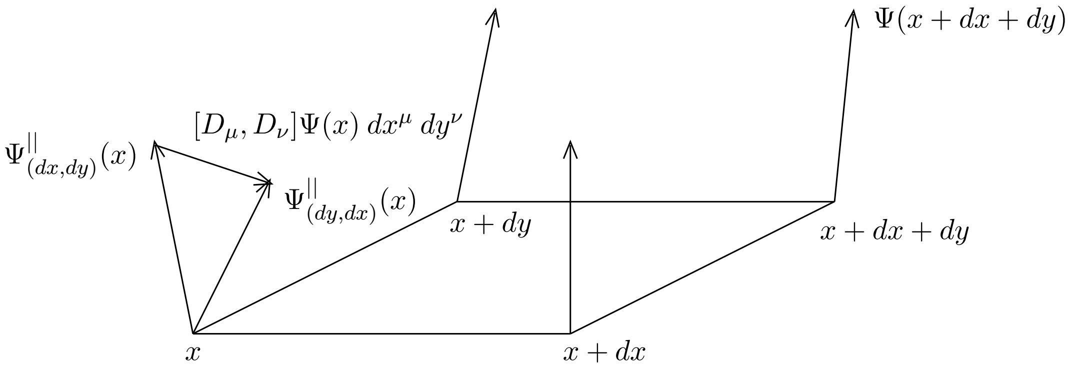


Fig. 6.4: Geometrical interpretation of the field strength tensor.

6.2.3 Equations of motion and currents

Given the invariant Lagrangian

$$\mathcal{L}_{\text{inv}} = -\frac{1}{4}\sum_i G_{i\mu\nu}G_i^{\mu\nu} + \bar{\Psi}\left(i\gamma^\mu\left(\partial_\mu - ig\sum_i T_i V_{i\mu}\right) - m\right)\Psi$$

the Euler-Lagrange equations

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \bar{\Psi}} = \frac{\partial \mathcal{L}}{\partial \bar{\Psi}} \quad \text{and} \quad \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu V_{i\nu}} = \frac{\partial \mathcal{L}}{\partial V_{i\nu}}$$

read

$$(i\gamma^\mu \partial_\mu - m)\,\Psi = -g\gamma^\mu T_i V_{i\mu}\Psi$$

and

$$\partial_\mu G_i^{\mu\nu} = -g\left(j_i^{\nu\Psi} + j_i^{\nu V}\right) = -g j_i^\nu$$

with

$$\begin{aligned} j_i^{\nu\Psi} &= \bar{\Psi}\gamma^\nu T_i \Psi \quad \text{Noether current of the matter field} \\ j_i^{\nu V} &= c_{ikl} G_k^{\nu\rho} V_{l\rho} \quad \text{current of the gauge field} \end{aligned}$$

The total current $j_i^\nu = j_i^{\nu\Psi} + j_i^{\nu V}$ is obviously conserved

$$\partial_\nu j_i^\nu = -\frac{1}{g}\partial_\nu \partial_\mu G_i^{\mu\nu} = 0$$

because $\partial_\nu\partial_\mu$ is symmetric whereas $G_i^{\mu\nu}$ is antisymmetric in $(\mu\nu)$.

The equations of motion may be written in a manifestly gauge invariant form

$$\begin{aligned} (i\gamma^\mu D_\mu - m)\,\Psi(x) &= 0 \\ D_\mu G_i^{\mu\nu}(x) &= -g J_i^\nu(x) \end{aligned}$$

where

$$D_\mu G_i^{\mu\nu} = \partial_\mu G_i^{\mu\nu} - g c_{ikl} V_{k\mu} G_l^{\mu\nu} .$$

The gauge invariant form is simply obtained by adding the appropriate terms on both sides of the equations of motion given before. Notice that $D_\mu G_i^{\mu\nu}$ and hence J_i^ν, given by

$$J_i^\nu = j_i^\nu - c_{ikl} V_{k\mu} G_l^{\mu\nu} = j_i^{\nu\Psi} = \bar{\Psi}\gamma^\nu T_i \Psi,$$

are vectors under local gauge transformations.

The covariant current coincides with the matter field current which is not conserved: $\partial_\nu J_i^\nu \neq 0$. On the other hand the conserved current

$$j_i^\nu = \bar{\Psi}\gamma^\nu T_i \Psi + c_{ikl} V_{k\mu} G_l^{\mu\nu} = -\frac{1}{g}\partial_\mu G_i^{\mu\nu}$$

is obviously not covariant because it is the ordinary derivative of a vector. We then arrive at the conclusion:

> In a locally gauge invariant theory a covariant conserved current with respect to the gauge symmetry does not exist.

This tells us that local symmetries are not symmetries in the usual global sense, like Poincaré invariance, Isospin invariance etc. . Local symmetries are **dynamical symmetries** and a consequence of the validity of an equivalence principle. Global symmetries describe algebraic properties of a system, only.

Summary (of subsection (6.2))

1. If we require n matter fields $(\psi_1, \ldots \psi_n)$ to be locally indistinguishable, such that

$$\Psi(x) = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \to \Psi'(x) = U(x)\Psi(x)\ ;\ \ U(x) = \exp ig \sum_i T_i \omega_i(x)\ \in\ SU(n)$$

   is a local symmetry of the system, the matter fields must couple minimally to a set of $r = n^2 - 1$ massless spin 1 gauge fields $V_{i\mu}(x)$:

$$\mathcal{L}_{\mathrm{inv}}^\Psi = \bar{\Psi}(x)\left(i\gamma^\mu D_\mu - m\right)\Psi(x)$$

   with covariant derivative

$$D_\mu = \partial_\mu - ig\sum_i T_i V_{i\mu}(x)\ .$$

   All matter fields $\psi_1 \ldots \psi_n$ must have identical mass and spin.

2. The locally gauge invariant Lagrangian of the gauge fields must be of the form

$$\mathcal{L}_{YM} = -\frac{1}{4}\sum_i G_{i\mu\nu} G_i^{\mu\nu}$$

with covariant field strength tensor

$$G_{i\mu\nu} = \partial_\mu V_{i\nu}(x) - \partial_\nu V_{i\mu}(x) + g c_{ikl} V_{k\mu}(x) V_{l\nu}(x)$$

The non-Abelian gauge fields must be self-interacting in the specific way as prescribed by $\mathcal{L}_{YM}$. A mass term

$$\frac{M^2}{2} \sum_i V_{i\mu}(x) V_i^\mu(x)$$

for the gauge bosons is not admitted.

3. Under infinitesimal gauge transformations the fields transform as

$$\delta\Psi_a = \ ig\, (T_i)_{ab} \Psi_b \delta\omega_i \ , \ \ \delta\bar{\Psi}_a = -ig\ \bar{\Psi}_b (T_i)_{ba} \delta\omega_i$$

with $(T_i)_{ab}$ the generators of $SU(n)$ in the fundamental representation.

$$\delta V_{k\mu} = ig \left(\tilde{T}_i\right)_{kl} V_{l\mu} \delta\omega_i + \partial_\mu \delta\omega_i \ , \ \ \delta G_{k\mu\nu} = ig \left(\tilde{T}_i\right)_{kl} G_{l\mu\nu} \delta\omega_i$$

with $\left(\tilde{T}_i\right)_{kl} = -ic_{ikl}$ the generators of $SU(n)$ in the adjoint representation.

A final remark concerns the generalization of the Yang-Mills construction to other gauge groups G. The necessary and sufficient conditions for the Yang-Mills construction to be possible are that the structure constants c_{ikl}

i) satisfy the Jacobi-identity

ii) are fully antisymmetric.

Whereas i) is automatic for finite matrices (finite dimensional representations) but may not hold for infinite matrices, the condition ii) is more restrictive. It holds for compact semi-simple Lie groups. Compact mean that the parameter space has a finite volume $\int \prod_i d\omega_i = V_\omega < \infty$. Semi-simple means G is a product of simple Lie-groups e.g. $G = SU(3) \otimes SU(2) \otimes U(1)$. Simple Lie-groups are those which cannot be decomposed into invariant subgroups. For each simple subgroup of a non-simple group G there would be a set of non-Abelian or Abelian ($U(1)$) gauge fields and an independent coupling constant.

Notice that the conditions i) and ii) are really non-trivial. As an example the Poincaré group is non-compact and has no finite dimensional unitary representation.

Outlook: For a long time, after Yang and Mills had proposed to extend local gauge invariance from Abelian to non–Abelian symmetry groups, non–Abelian gauge theories were considered to be unphysical, because they required the existence of multiplets of massless spin 1 bosons, which were known not to exist in Nature. At that time it was not known that there are two ways out of the dilemma. One is the Higgs mechanism where the gauge bosons acquire a mass by "spontaneous symmetry breaking of the local gauge symmetry". Today we know that the electroweak gauge group $SU(2)_L \otimes U(1)_Y$, of weak isospin and weak hypercharge Y, is broken down to the Abelian electromagnetic gauge group $U(1)_{em}$. Three out of the 3+1 gauge bosons acquire a mass and the remaining massless state is the photon. The weak gauge bosons $W^\pm$ and Z in fact turned out to be very heavy (about 80 GeV and 91 GeV, respectively) and were discovered in 1983 at the CERN $p\bar{p}$ collider. The LEP e^+e^- storage ring at CERN, in operation since 1989, is a Z factory and produces millions of Z's.

The other "solution" is **confinement**. Unbroken non-Abelian gauge theories are asymptotically free (Politzer, Gross and Wilzcek 1973), which means that they have small effective coupling at high energies (short distances) but strong effective coupling at low energies (large distances). We know that the strong interactions of hadrons are described by an unbroken color $SU(3)_c$ local gauge theory, called quantum chromodynamics (QCD) (Gell-Mann, Fritzsch and Leutwyler 1973). The matter fields are the colored quarks triplets which interact through the octet of massless gauge fields, called gluons. Quarks and gluons are permanently confined inside of the hadrons. This is another mechanism which hides massless gauge bosons from the physical spectrum.

Obviously the earlier conclusion that Yang-Mills theories are not of relevance for physics was premature.

Exercises: Section 6

① If $\boldsymbol{V}_\mu = \frac{i}{g} U^{-1}(x)\,(\partial_\mu U(x))$ the field is called a " pure gauge" field. Show that in this case $\boldsymbol{G}_{\mu\nu} \equiv 0$.

② Show that a mass term $\frac{M^2}{2}\sum_i V_{i\mu}V_i^\mu$ cannot be locally gauge invariant.

③ Show that $\mathcal{L}_{YM}$ can be written in the form of a trace

$$\mathcal{L}_{YM} = -\frac{1}{2} Tr\,(\boldsymbol{G}_{\mu\nu}\boldsymbol{G}^{\mu\nu})$$

④ Prove the validity of the "homogeneous Maxwell equation" or "Bianchi-identity":

$$D_\rho \boldsymbol{G}_{\mu\nu} + \text{ terms cyclic in } (\rho\mu\nu) \equiv 0$$

7 Path integral quantization

The prescription how to quantize properly a non-Abelian gauge theory and how to obtain the correct Feynman rules is described in the next section. Thus the reader interested more in applications may skip this somewhat technical section.

The path integral quantization was originally invented by Feynman and it was the main tool to quantize fields. Later path integral quantization was displaced by canonical quantization, which is more closely related to ordinary quantum mechanics with states and operators as basic objects. While being conceptually simple and having a direct physical interpretation, the disadvantage of the canonical quantization is its reliance on the "interaction picture", i.e., on the splitting of the Lagrangian into a free part $\mathcal{L}_0$ and an interaction part $\mathcal{L}_{\rm int}$. In particular for the correct quantization of non-Abelian gauge theories this necessary splitting is not gauge invariant and the control of gauge invariance remained obscure for a long time. The break–through came with the rejuvenation of the path integral formalism by Faddeev, Popov, Fradkin, Tyupkin, 't Hooft and others. The advantages of the path integral formulation will be discussed in more detail below. The most important fact is that the path integral representation provides a **non–perturbative definition of quantum field theory** and is the starting point for investigating phenomena like confinement and bound states in QCD, or non–perturbative effects like instantons, sphalerons etc. The path integral formulation of field theory provides a much more general framework, than the scattering theory inspired canonical quantization approach.

7.1 Functional integral for bosons

Our aim here is to reformulate a canonically quantized theory in terms of a "path integral". Let $\varphi(x)$ be a quantized free real scalar field of mass m. The corresponding Lagrangian reads:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi(x)\,\partial^\mu\varphi(x) \; - \frac{m^2}{2}\varphi^2(x) \;\; .$$

Since the dynamics of a system is governed by the principle of least action the physically relevant quantity is actually the action $\int d^4x\mathcal{L}(x)$. We are thus free to perform a partial integration and write the free Lagrangian as a quadratic form in the field with the Klein-Gordon operator as a kernel:

$$\mathcal{L} = -\frac{1}{2}\varphi(x)\,(\,\Box_x + m^2)\,\varphi(x) \;\; . \tag{7.1}$$

The time ordered two point function, the Feynman propagator, is given by

$$< 0 \mid T\left\{\varphi(x), \varphi(y)\right\} \mid 0 > \;\; = \;\; i\,\Delta_F(x-y; m^2) \tag{7.2}$$

with

$$\Delta_F(z; m^2) = \frac{1}{(2\pi)^d}\int d^dq\, e^{-iqz}\,\frac{1}{q^2-m^2+i\epsilon} \;\; , \tag{7.3}$$

or, in Fourier space,

$$\tilde{\Delta}_F(q) = \frac{1}{q^2-m^2+i\epsilon} \;\; . \tag{7.4}$$

The space–time dimension is taken to be d. Note that due to translation invariance $\Delta_F(x-y; m^2)$ is a function of $x-y$ only. Since the Feynman propagator is a solution of the inhomogeneous Klein-Gordon equation with a point source term

$$(\,\Box_x + m^2)\,\Delta_F(x-y; m^2) = -\delta^{(d)}(x-y) \tag{7.5}$$

we note that up to a sign (our convention) the Feynman propagator is the inverse of the kernel of the bilinear Lagrangian written in the form Eq. (7.1).

We define a **scalar product** by

$$(f_1, \Delta_F f_2) \doteq \int d^dx\, d^dy\, f_1(x)\, \Delta_F(x-y)\, f_2(y)$$

where $f_i(x)$ are suitable test functions, e.g., from the space $\mathcal{S}(\mathbf{R}^d)$ of functions which are smooth (infinitely differentiable) and fall off at infinity faster than any power. In physics terms the test functions are smooth localized wave packets.

7.1.1 Generating functional for bosons

A **generating functional** is then defined by

$$\begin{aligned} < 0 \mid T\, e^{i \int d^dx\, \varphi(x)\, J(x)} \mid 0 > \;\; &\doteq \;\; \sum_{n=0}^{\infty} \frac{i^n}{n!} \int d^dx_1 \ldots \int d^dx_n\, J(x_1) \ldots J(x_n) \;\times \\ & \quad\; < 0 \mid T\,\{\varphi(x_1) \ldots \varphi(x_n)\} \mid 0 > \quad . \end{aligned} \tag{7.6}$$

The ordinary function $J(x)$ is a classical external field, which allows us to represent the Green functions as "moments" with respect to $J(x)$ (r.h.s. of Eq. (7.6)). The generating functional allows us to extract all time–ordered Green functions and these allow us to reconstruct the whole quantum field theory. For a free theory, i.e., $\varphi(x)$ a free field, we may easily calculate all Green functions: By Wick's theorem (see Sec. 3.4.2)

$$< 0 \mid T\,\{\varphi(x_1) \ldots \varphi(x_n)\} \mid \; 0 >= \sum \text{ all possible complete contractions } \; .$$

A contraction means a pairing of two fields which represents a Feynman propagator. Non vanishing contributions require n to be even.

From the n fields we chose successively a pair and replace it by a Feynman propagator until all fields are used up. Given the ordered set of points $(x_1, x_2, \ldots, x_n)$ the number of choices in performing the pairings are given by

$$\begin{array}{llcl} \text{1st} & \text{pair} & \binom{n}{2} & \text{possibilities} \\ \text{2nd} & \text{pair} & \binom{n-2}{2} & \text{possibilities} \\ & & \vdots & \\ \text{last} & \text{pair} & \binom{2}{2} & \text{possibilities} \quad , \end{array}$$

which yields,

$$\left(\begin{array}{c} n \\ 2 \end{array}\right) \left(\begin{array}{c} n-2 \\ 2 \end{array}\right) \cdots \left(\begin{array}{c} 2 \\ 2 \end{array}\right) = \frac{n!}{(n-2)!2!}\, \frac{(n-2)!}{(n-4)!2!} \cdots \frac{2!}{0!2!} = \frac{n!}{(2)^{n/2}} \quad .$$

The number of propagators is $n/2$. So far we have over counted the $(n/2)!$ permutations of the $n/2$ propagators as different and we have to correct for that by dividing the above result by this factor. The $n!$ permutations of $(x_1, x_2, \ldots, x_n)$ are equivalent because of the permutation symmetry (as integration variables the x_i's are indistinguishable). We obtain

$$< 0 \mid T\,\{\varphi(x_1) \ldots \varphi(x_n)\} \mid 0 >= (i)^{n/2} \frac{n!}{(2)^{n/2}(n/2)!}\, \Delta_F(x_1 - x_2)\, \Delta_F(x_3 - x_4) \; \cdots \; \Delta_F(x_{n-1} - x_n)$$

for n even, and zero for n odd. We thus find

$$\left\| \begin{array}{rcl} Z_0\{J\} &=& <0\,|\,T\, e^{i\int d^dx\,\varphi(x)\,J(x)}\ \,|\,0> \\ &=& \sum_{n=0}^{\infty} \frac{(-i)^{n/2}}{(n/2)!}\,[\frac{1}{2}(J,\Delta_F\,J)]^{n/2} = \exp -\frac{i}{2}(J,\Delta_F\,J)\ \ . \end{array}\right. \tag{7.7}$$

This is a **(free) Gaussian functional** :

$$\begin{array}{rcl} Z_0\{J\} &=& e^{-\frac{i}{2}\int d^dx\,d^dy\,J(x)\,\Delta_F(x-y)\,J(y)} \\ Z_0\{0\} &=& 1 \qquad \text{normalization} \quad . \end{array}$$

The time ordered Green functions may be obtained from the generating functional by taking **functional derivatives** with respect to the source function $J(x)$:

$$\begin{array}{rcl} -i\,\frac{\delta}{\delta J(x_1)}\quad Z_0\{J\}\ |_{J=0} &=& <0\,|\,T\,\{\varphi(x_1)\}\,|\,0>=0 \\ && \text{“mean”} = 0 \\ (-i)^2\,\frac{\delta^2}{\delta J(x_1)\,\delta J(x_2)}\quad Z_0\{J\}\ |_{J=0} &=& <0\,|\,T\,\{\varphi(x_1)\varphi(x_2)\}\,|\,0> \\ && \text{“covariance”} = i\Delta_F(x_1-x_2) \\ \vdots && . \end{array} \tag{7.8}$$

These relations are immediately derived from the basic property

$$\frac{\delta}{\delta J(x)}J(y) = \delta^{(d)}(x-y) \text{ or equivalently } \frac{\delta}{\delta J(x)}\int d^dy\,F(y)J(y) = F(x) \tag{7.9}$$

of the functional derivative.

7.1.2 Wick rotation, imaginary time, Euclidean functional

The momentum space representation of the Feynman propagator Eq. (7.4) is an analytic function in q^0 with poles at $q^0 = \pm(\,\omega_p - i\epsilon)$ where $\omega_p = \sqrt{\vec{q}^{\,2}+m^2}$ (see Sec. 3.4.3). This allows us to rotate by $\frac{\pi}{2}$ the integration path in q^0, going from $-\infty$ to $+\infty$, without crossing any singularity. In doing so, we rotate from Minkowski space to Euclidean space

$$q^0 \to -iq^d \Rightarrow q = (q^0,q^1,\dots,q^{d-2},q^{d-1}) \to \underline{q} = (q^1,q^2,\dots,q^{d-1},q^d) \text{ where } q^2 \to -\underline{q}^2\ \ . \tag{7.10}$$

This rotation to the Euclidean region is called **Wick rotation** .

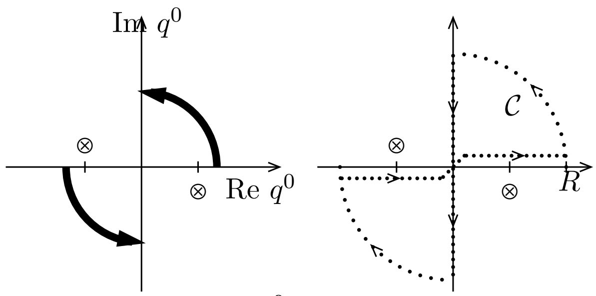


Fig. 7.1: Wick rotation in the complex q^0–plane. The poles of the Feynman propagator are indicated by $\otimes$'s.

More precisely: analyticity of a function $\tilde{f}(q^0,\vec{q}\,)$ in q^0 implies that the **contour integral**

$$\oint_{\mathcal{C}(R)} dq^0\,\tilde{f}(q^0,\vec{q}\,)\ = 0$$

for the closed path $C(R)$ in Fig. 7.1.2 vanishes. If the function $\tilde{f}(q^0, \vec{q}\,)$ falls off sufficiently fast at infinity, then the contribution from the two "arcs" goes to zero when the radius of the contour $R \to \infty$. In this case we obtain

$$\int_{-\infty}^{\infty} dq^0\, \tilde{f}(q^0, \vec{q}\,) + \int_{+i\infty}^{-i\infty} dq^0\, \tilde{f}(q^0, \vec{q}\,) = 0$$

or

$$\int_{-\infty}^{\infty} dq^0\, \tilde{f}(q^0, \vec{q}\,) = \int_{-i\infty}^{+i\infty} dq^0\, \tilde{f}(q^0, \vec{q}\,) = -i \int_{-\infty}^{+\infty} dq^d\, \tilde{f}(-iq^d, \vec{q}\,) \ ,$$

which is the Wick rotation. At least in perturbation theory, one can prove that the conditions required to allow us to perform a Wick rotation are fulfilled.

We notice that the Euclidean Feynman propagator obtained by the Wick rotation

$$\frac{1}{q^2 - m^2 + i\epsilon} \to -\frac{1}{\underline{q}^2 + m^2} \tag{7.11}$$

has no singularities (poles) and an $i\epsilon$–prescription is not needed any longer.

In configuration space a Wick rotation implies going to **imaginary time** $x^0 \to ix^0 = x^d$ such that $qx \to -\underline{q}\underline{x}$ and hence

$$x^0 \to -ix^d \Rightarrow x^2 \to -\underline{x}^2 \ , \quad \Box_x \to -\Delta_{\underline{x}} \ , \quad i\int d^dx \cdots \to \int d^d\underline{x} \cdots \quad . \tag{7.12}$$

While in Minkowski space $x^2 = 0$ defines the light–cone $x^0 = \pm|\vec{x}|$, in the Euclidean region $\underline{x}^2 = 0$ implies $\underline{x} = 0$. Note that possible singularities on the light–cone like $1/x^2$, $\delta(x^2)$ etc. turn into singularities at the point $\underline{x} = 0$. This simplification of the singularity structure is the merit of the positive definite metric in Euclidean space.

All quantities considered above have their Euclidean versions. The Lagrangian (7.1) takes the form

$$\mathcal{L} = -\frac{1}{2}\varphi(\underline{x})\,(\,-\Delta_{\underline{x}} + m^2)\,\varphi(\underline{x}) \ . \tag{7.13}$$

For the Euclidean two–point function we obtain

$$\begin{aligned} i\Delta_F(x-y;m^2) &= \frac{i}{(2\pi)^d}\int d^dq\, e^{-iq\,(x-y)}\,\frac{1}{q^2-m^2+i\epsilon} \\ &= \frac{1}{(2\pi)^d}\int d^d\underline{q}\, e^{i\underline{q}\,(\underline{x}-\underline{y})}\,\frac{1}{\underline{q}^2+m^2} \\ &= G_0(\underline{x}-\underline{y}) \ . \end{aligned} \tag{7.14}$$

Note that G_0 is the inverse of $(\,-\Delta + m^2)$, thus

$$(\,-\Delta_{\underline{x}} + m^2)\,G_0(\underline{x}-\underline{y}) = \delta^{(d)}(\underline{x}-\underline{y}) \ .$$

The Euclidean Green functions are also called **Schwinger functions**.

With

$$-i\,(J, \Delta_F\, J) = \int d^d\underline{x}\, d^d\underline{y}\, J(\underline{x})\, G_0(\underline{x}-\underline{y})\, J(\underline{y}) = (J, G_0\, J)$$

we may write

$$Z_0^E\{J\} = \exp\frac{1}{2}(J, G_0\, J) \tag{7.15}$$

for the generating functional of the Euclidean Green functions.

In the following, unless stated otherwise, we work in Euclidean space and denote d–dimensional Euclidean vectors by simply $x, y, z, p, q, \ldots$ In Fourier space we obtain the following representation: The scalar product

$$(J, G_0\, J) = \int d^dx\, d^dy\, J(x)\, G_0(x-y)\, J(y)$$

upon Fourier decomposition of the source function

$$J(x) = \int \frac{d^dq}{(2\pi)^d}\, e^{-iqx}\, \tilde{J}(q)$$

reads

$$\begin{array}{rcccl}
(J, G_0\, J) & = & \int \frac{d^dq}{(2\pi)^d} \frac{d^dp}{(2\pi)^d}\, \tilde{J}(q)\, \tilde{J}(p) & \times & \int d^dx\, d^dy\, e^{-iqx}\, e^{-ipy}\, G_0(x-y) \\
& = & \cdots & \times & \int d^dz\, d^dy\, e^{-iqz}\, e^{i(p+q)y}\, G_0(z) \\
& = & \cdots & \times & \tilde{G}_0(q)\, (2\pi)^d\, \delta^{(d)}(p+q)
\end{array}$$

and hence

$$\begin{aligned}
(J, G_0\, J) &= \int \frac{d^dq}{(2\pi)^d} \frac{d^dp}{(2\pi)^d}\, (2\pi)^d\, \delta^{(d)}(p+q) \tilde{J}(q)\, \tilde{J}(p)\, \tilde{G}_0(q) \\
&= \int \frac{d^dq}{(2\pi)^d}\, \tilde{J}(q)\, \tilde{J}(-q)\, \tilde{G}_0(q) \\
&= \int \frac{d^dq}{(2\pi)^d}\, |\tilde{J}(q)|^2\, \tilde{G}_0(q) \quad .
\end{aligned}$$

Here we have used the reality property of the source

$$J(x)^* = J(x)$$

which implies

$$\begin{aligned}
\tilde{J}(q) &= \int d^dx\, e^{\,iqx}\, J(x) \\
\tilde{J}(q)^* &= \int d^dx\, e^{-iqx}\, J(x) = \tilde{J}(-q) \quad .
\end{aligned}$$

We notice that in momentum space $(J, G_0\, J)$ is diagonal and manifestly positive, and we have the result that $Z_0^E\{J\}$ is a Gauss functional with positive kernel G_0.

7.1.3 Gauss integrals and Gauss functionals

The basic Gauss integral reads

$$\int_{-\infty}^{+\infty} dx\, e^{-(ax^2+2bx+c)} = \sqrt{\frac{\pi}{a}}\, e^{\frac{b^2-ac}{a}} \quad ; \quad a > 0 \ ,$$

and with $c = 0$ and $2b = J$ we obtain

$$F\{J\} = \int_{-\infty}^{+\infty} dx\, e^{-(ax^2+Jx)} = \sqrt{\frac{\pi}{a}}\; e^{\frac{J^2}{4a}} \quad ; \quad a > 0 \ ,$$

which is the generating function of the moments

$$\int_{-\infty}^{+\infty} dx\, x^n\, e^{-ax^2} = (-1)^n \frac{d^n}{dJ^n}\; F\{J\}|_{J=0} \ .$$

The generalization to the multidimensional case is obvious: Let K_{ij} be a positive $N \times N$ matrix and let its inverse K_{ij}^{-1} be the kernel of the quadratic form $\sum_{i,j} \varphi_i K_{ij}^{-1} \varphi_j$. We may then consider the multi–integral

$$\int_{-\infty}^{+\infty} d\varphi_1 \ldots \varphi_N\, e^{-\frac{1}{2}\sum_{i,j} \varphi_i K_{ij}^{-1} \varphi_j + \sum_i \varphi_i J_i} \tag{7.16}$$

which can be calculated in a simple way, by diagonalizing the quadratic form. Thus, let S be the orthogonal transformation $S^{-1} = S^T$ which is diagonalizing the kernel K^{-1}:

$$(SK^{-1}S^{-1})_{ij} = \lambda_i \delta_{ij} \ .$$

Since $SS^T = 1$ we have $\det(SS^T) = \det S \det S^T = (\det S)^2 = 1$. We may chose $\det S = 1$. A change of integration variables

$$S_{ij}\varphi_j = \varphi'_i \quad ; \quad \varphi = S^{-1}\varphi' \ , \quad J = SJ'$$

$$\prod d\varphi_i = \left| \frac{\partial \varphi_i}{\partial \varphi'_j} \right| \prod d\varphi'_j = \det S^{-1} \prod d\varphi'_i = \prod d\varphi'_i$$

leads to

$$\begin{aligned}
& \int_{-\infty}^{+\infty} \prod_i d\varphi'_i e^{-\frac{1}{2}\sum_j \varphi'^2_j \lambda_j + \sum_j \varphi'_j J'_j} \\
= \ & \int_{-\infty}^{+\infty} \prod_i d\varphi'_i e^{-\frac{1}{2}\lambda_i \varphi'^2_i + \varphi'_i J'_i} \ = \prod_i \sqrt{\frac{2\pi}{\lambda_i}}\ e^{\sum_j \frac{J'^2_j}{2\lambda_j}} \\
= \ & \frac{(2\pi)^{N/2}}{\sqrt{\det(K'^{-1})}}\ e^{\frac{1}{2}\sum_{i,j} J'_i K'_{ij} J'_j} \ = \ \frac{(2\pi)^{N/2}}{\sqrt{\det(K^{-1})}}\ e^{\frac{1}{2}(J,KJ)}
\end{aligned} \tag{7.17}$$

where we used

$$K'_{ij} = \frac{1}{\lambda_i}\, \delta_{ij} \text{ and } \det K^{-1} = \prod_i \lambda_i \ .$$

The last step, transforming back to the original variables, makes use of the invariance of the scalar product (J, KJ) and of the determinant $\det K$ under orthogonal transformations.

7.1.4 Minkowski space and Fresnel integrals

In Minkowski space, where we usually work in relativistic physics, the Gaussian integrals encountered for the Euclidean boson fields turn into oscillatory **Fresnel integrals**. We list a few results for illustration:

For real variables we have:

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \frac{d\varphi}{\sqrt{2\pi i}}\, e^{\frac{i}{2}\varphi A\varphi} = A^{-1/2} \quad ; \quad A > 0 \\
&\int_{-\infty}^{+\infty} \prod_x \frac{d\varphi_x}{\sqrt{2\pi i}}\, e^{\frac{i}{2}\sum_x \varphi_x A_x \varphi_x} = \left(\prod_x A_x\right)^{-1/2} \quad ; \quad A_x > 0\,, \quad \text{diagonal} \\
&\int_{-\infty}^{+\infty} \prod_x \frac{d\varphi_x}{\sqrt{2\pi i}}\, e^{\frac{i}{2}\sum_{x,y} \varphi_x A_{xy} \varphi_y} = (\det A)^{-1/2} \quad ; \quad A_{xy} \text{ symmetric, positive definite.}
\end{aligned} \tag{7.18}$$

The last result follows from the previous one by performing an orthogonal transformation of A to diagonal form.

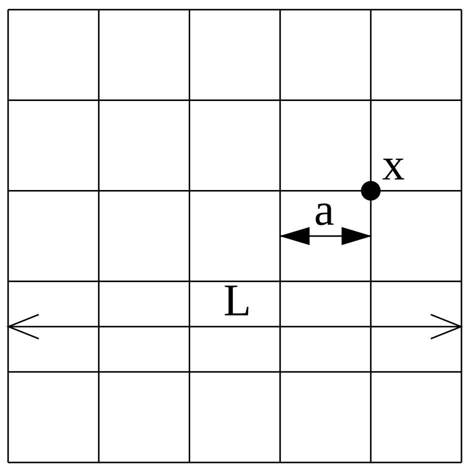
For complex variables the corresponding result reads:

$$\left\| \int_{-\infty}^{+\infty} \prod_x \frac{d\varphi_x}{\sqrt{2\pi i}} \frac{d\varphi_x^*}{\sqrt{2\pi i}}\, e^{i\sum_{x,y} \varphi_x^* A_{xy} \varphi_y} = (\det A)^{-1} \right. \quad , \tag{7.19}$$

where A_{xy} is hermitian and positive definite. Again, this last result can be obtained by diagonalizing A by an appropriate unitary transformation.

7.1.5 Lattice field theory.

We now consider the N variables of the previous example as N variables associated with the points x of a d–dimensional Euclidean "space–time" lattice $V_{L,a}$ (see the Figure)



By a we denote the lattice spacing and ℓ is the number of points in one direction. Thus, $\ell\, a = L$ is the size and $L^d = V$ is the volume. The number of points is $N = \ell^d$, which corresponds to the number of variables. We then consider the generating function

$$\begin{aligned}
Z_0^E\{J\} &= \text{const} \int \prod_{x \in V_{L,a}} d\varphi_x\, e^{-\frac{a^{2d}}{2}\sum_{x,y}\varphi_x G_{a\,x,y}^{-1}\varphi_y + a^d \sum_x \varphi_x J_x} \\
&= \exp \frac{1}{2}(J, G_a J) \ .
\end{aligned} \tag{7.20}$$

The only difference relative to (7.16) are some factors, powers of the lattice spacing a, chosen such that the continuum limit exists trivially. For example, we have

$$(J, G_a J) = a^{2d} \sum_{x,y} J_x G_{a\,x,y} J_y \tag{7.21}$$

and, with the correspondence $a \leftrightarrow dx^i$ $(i = 1, \ldots, d)$, the Riemann sums converge to the integral

$$(J, G_0 J) \quad = \quad \int dx^1 \ldots dx^d \int dy^1 \ldots dy^d J(x) G_0(x-y) J(y)$$

when $a \to 0$. The "unity operator" in Euclidean configuration space on the lattice is

$$a^{-d}\,\delta^{(d)}_{x,y} \overset{a\to 0}{\to} \delta^{(d)}(x-y) \tag{7.22}$$

with $\delta^{(d)}_{x,y} = \prod_{i=1}^{d} \delta_{x^i,y^i}$ the Kronecker identity. Thus

$$\sum_{x'} G^{-1}_{a\,x,x'} G_{a\,x',y} = a^{-d}\,\delta^{(d)}_{x,y} \quad .$$

Finally, we chose the normalization constant "const" in (7.20) such that

$$\text{const} \;\hat{=}\; a^{dN}\,\frac{\sqrt{G_a^{-1}}}{(2\pi)^{N/2}} \;\Leftrightarrow\; Z_0^E\,\{0\} = 1 \quad .$$

Divergence and Laplace operator on the lattice:

In relating a continuum field theory to a field theory on a lattice the differential operators, like the Laplace operator appearing in the Euclidean Lagrangian (7.13), must be replaced by finite difference operators.

The partial derivative on a discrete lattice is defined by the nearest neighbor difference operator

$$\nabla^{\pm}_{a,i} \doteq \pm\frac{1}{a}\left\{f_{x\pm ae_i} - f_x\right\}$$

where e_i is a unit vector in the direction i. Obviously, in the continuum limit

$$\lim_{a\to 0} \nabla^{\pm}_{a,i} f_x = \frac{\partial f}{\partial x^i} \quad .$$

The discrete Laplace operator is defined by

$$\Delta_a \doteq \sum_{i=1}^{d} \nabla^{+}_{a,i}\nabla^{-}_{a,i}$$

such that

$$\Delta_a f_x = -a^{-2}\left\{2d f_x - \sum_{|x-y|=a} f_y\right\} \quad , \tag{7.23}$$

with the limit

$$\lim_{a\to 0} \Delta_a f_x = \Delta f(x) = \sum_{i=1}^{d} \frac{\partial^2 f}{\partial x^{i2}} \quad .$$

We are ready, now, to consider the lattice version of the Euclidean Lagrangian Eq. (7.13). We choose G_a^{-1} in (7.20) such that

$$a^d \sum_y G^{-1}_{a\,x,y}\varphi_y = (-\Delta_a + m^2\,)\varphi_x \tag{7.24}$$

and thus obtain

$$-\frac{a^{2d}}{2}\sum_{x,y}\varphi_x G^{-1}_{a\,x,y}\varphi_y = -\frac{a^d}{2}\sum_x \varphi_x(-\Delta_a + m^2\,)\varphi_x \quad . \tag{7.25}$$

For $a \to 0$ this goes to

$$\int d^dx\,\mathcal{L}_0^E(\varphi) \quad \text{with} \quad \mathcal{L}_0^E(\varphi) = -\frac{1}{2}\varphi(x)(-\Delta + m^2\,)\varphi(x) \quad . \tag{7.26}$$

We thus have constructed a finite–dimensional lattice approximation in Euclidean space for the free real scalar field theory defined by the Lagrangian (7.1). The generating functional for the correlation functions is given by Eq. (7.20) with the identification Eq. (7.24). Noting the identity

$$e^{-\frac{a^d}{2}\sum_x \varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x} = \prod_x e^{-\frac{a^d}{2}\varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x}$$

we obtain

$$\begin{aligned} &\int \prod_{x\in V_{L,a}} d\varphi_x\; e^{-\frac{a^{2d}}{2}\sum_{x,y}\varphi_x G^{-1}_{a\,x,y}\varphi_y} \\ =&\int \prod_{x\in V_{L,a}} d\varphi_x\; e^{-\frac{a^d}{2}\varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x} = \frac{(2\pi)^{N/2}}{a^{dN}\sqrt{\det G_a^{-1}}} \quad . \end{aligned} \tag{7.27}$$

This gives raise to the definition of the **normalized measure**

$$\prod_x d\mu_{N,a}(\varphi_x) \doteq \frac{\prod_x\; e^{-\frac{a^d}{2}\varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x} d\varphi_x}{\int\prod_x\; e^{-\frac{a^d}{2}\varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x}\;d\varphi_x} \tag{7.28}$$

with **mean zero** and **covariance** $(-\Delta_a + m^2\,)^{-1}$**.** With the help of this Gaussian measure we finally obtain the representation

$$\begin{aligned} Z_0^E\,\{J\} &= \int \prod_{x\in V_{L,a}} d\mu_{N,a}(\varphi_x)\; e^{a^d\sum_x \varphi_x J_x} \\ &= \exp\frac{1}{2}(J, G_a J) \quad . \end{aligned} \tag{7.29}$$

for the generating functional. Here, G_a denotes the Euclidean Feynman propagator on the lattice, satisfying

$$(-\Delta_{a,x} + m^2\,)\,G_{a\,x,y} = a^{-d}\,\delta^{(d)}_{x,y} \quad . \tag{7.30}$$

Explicit form of the lattice propagator

Before we are going to construct the continuum limit let us explicitly calculate the Euclidean Feynman propagator on the lattice: We consider a system in a finite box and fields satisfying **periodic boundary conditions**

$$\varphi_{x+\ell e_i} = \varphi_x \quad , \quad i = 1,\ldots d \quad , \quad \ell \;\text{ integer} \quad .$$

We perform a Fourier transformation

$$a^d\sum_{x\in V_{L,a}}\varphi_x e^{-iqx} \doteq \tilde{\varphi}_q \quad . \tag{7.31}$$

By the periodicity of φ_x the momenta q take values on the **reciprocal lattice**, which is obtained as follows: Periodicity requires $x = \vec{n} \cdot a$ with n_i integer. If we denote by $\vec{e}_i$ the basis vectors in x–space, furthermore periodicity implies $\vec{q} \cdot \vec{e}_i \cdot a = 2\pi m_i$ with m_i integer. We may introduce basis vectors in q–space by

$$\vec{f}_i \cdot \vec{e}_i = \frac{2\pi}{a}\delta_{ij}$$

such that

$$\vec{q} = \vec{f}_i m_i \ \text{ with } \ |q_i| \in \left[-\frac{\pi}{a}, +\frac{\pi}{a}\right] \ .$$

This region in q–space is called **Brillouin zone** and we denote it by $\Lambda_{L,a}$. The inverse Fourier transformation the may be written as

$$\frac{1}{Na^d} \sum_{q \in \Lambda_{L,a}} \tilde{\varphi}_q e^{iqx} = \varphi_x \tag{7.32}$$

which we may easily check to be true. In Fourier space Eq. (7.25) takes the form

$$\frac{1}{2}\frac{1}{Na^d} \sum_{q \in \Lambda_{L,a}} |\tilde{\varphi}_q|^2 \, \tilde{G}_{aq}^{-1}$$

where

$$\tilde{G}_{aq}^{-1} = m^2 + 4a^{-2} \sum_{i=1}^{d} \sin^2 \frac{aq_i}{2} \ \ \stackrel{a \to 0}{\to} \ \ m^2 + q^2 + O(a^2 q^4) \tag{7.33}$$

is the **free lattice propagator for Euclidean bosons**. As it should be, this propagator approaches the free Euclidean propagator in the continuum limit. This result follows from

$$\begin{aligned}
a^d \sum_{x \in V_{L,a}} (-\Delta_{a,x} + m^2\,)\, e^{-iqx} &= a^{(d-2)} \sum_{x \in V_{L,a}} \left\{ [(ma)^2 + 2d] e^{-iqx} - \sum_{|x-y|=a} e^{-iqy} \right\} \\
&= a^{(d-2)} \sum_{x \in V_{L,a}} \left\{ [(ma)^2 + 2d] e^{-iqx} - \sum_{i=1}^{d} 2\cos(aq_i)\, e^{-iqx} \right\} \\
&= a^d \sum_{x \in V_{L,a}} \left\{ m^2 + \sum_{i=1}^{d} \frac{4}{a^2} \sin^2 \frac{aq_i}{2} \right\} e^{-iqx}
\end{aligned}$$

where we have used (7.23) and in the last step $\cos\theta_i = 1 - 2\sin^2 \frac{\theta_i}{2}$.

7.1.6 Addendum: Euclidean field theory and statistical mechanics

A simple typical model in statistical mechanics is the **Ising–model** of a ferromagnet. The atoms of a solid are located on a regular lattice, which we assume to be d–dimensional cubic, and carry spins σ_x which may point into $\pm$ z–direction: $\sigma_x = \pm 1 \in Z_2$[25]. Only nearest neighbor spins interact with each other, parallel spins are attractive while anti-parallel spins are repulsive. Thus the Ising–Hamiltonian reads

$$H_N(\sigma) = -k \sum_{\substack{x,y \in V_{L,a} \\ |x-y|=a}} \sigma_x \sigma_y \ .$$

[25] Z_n denotes the discrete multiplicative group of the n-th unit roots $e^{i\frac{2\pi}{n}}$ in the complex plane

The partition sum from which all statistical and thermodynamic properties of the model may be derived, reads

$$Z_N = \sum_{\{\sigma\}} e^{-H_N(\sigma)} = \int \prod_{x \in V_{L,a}} d\sigma_x \, \delta(\sigma_x^2 - 1) \, e^{-H_N(\sigma)} \quad .$$

Here we have absorbed the Boltzmann factor $\beta = 1/(k_B T)$ into the definition of the nearest neighbor coupling k. We may write the discrete distribution, described by the δ–function, by a limit of starting from smooth continuous spin distributions

$$\delta(\sigma_x^2 - 1) = \lim_{u_0 \to \infty} \sqrt{\frac{u_0}{\pi}} e^{-u_0 (\sigma_x^2 - 1)^2} \quad .$$

For finite but sufficiently large u_0, this provides model with smearing of the sharp values $\sigma = \pm$ of the spins. We may write the Hamiltonian in terms of the Laplace operator on the lattice

$$H_N(\sigma) = \frac{1}{2} \sigma K \sigma = a^{-2} k \sum_x \left\{ 2d\sigma_x^2 + \sigma_x \Delta_a \sigma_x \right\}$$

and we obtain for finite u_0:

$$Z_N = \left(\frac{u_0}{\pi} \right)^{N/2} e^{u_0 N} \int \prod_{x \in V_{L,a}} d\sigma_x \, e^{-\bar{H}_N(\sigma)}$$

where

$$H_N(\sigma) = \frac{Z_a}{2} a^d \sum_x \sigma_x \left(-\Delta_a + m^2 \right) \sigma_x + u Z_a^2 a^d \sum_x \sigma_x^4$$

$$\begin{aligned} Z_a &= 2ka^{2-d} > 0 \\ m^2 &= -2a^{-2} \left(\frac{u_0}{k} + d \right) < 0 \\ u &= \frac{u_0}{4k^2} a^{d-4} > 0 \end{aligned}$$

We now perform a field renormalization

$$\varphi_x = \sqrt{Z_a} \sigma_x$$

and obtain

$$Z\{J\} = \int \prod_{x \in V_{L,a}} d\varphi_x \, e^{-\bar{H}_N(\varphi) + a^d \sum_x \varphi_x J_x} / \int \prod_{x \in V_{L,a}} d\varphi_x \, e^{-\bar{H}_N(\varphi)}$$

as the generating function for the correlation functions of the Ising–model. We have

$$\begin{aligned} \bar{H}_N(\varphi) &= \bar{H}_0(\varphi) + \bar{H}_{\text{int}}(\varphi) \\ \bar{H}_0(\varphi) &= \frac{1}{2} a^d \sum_x \varphi_x \left(-\Delta_a + m^2 \right) \varphi_x \\ \bar{H}_{\text{int}}(\varphi) &= u a^d \sum_x \varphi_x^4 \quad . \end{aligned}$$

In the continuum limit

$$\begin{aligned} \bar{H}_0(\varphi) &= \stackrel{a \to 0}{\to} \int d^d x \, \mathcal{L}_0(\varphi) \\ \bar{H}_{\text{int}}(\varphi) &= \stackrel{a \to 0}{\to} \int d^d x \, \mathcal{L}_{\text{int}}(\varphi) \end{aligned}$$

with

$$\begin{aligned}\mathcal{L}_0(\varphi) &= \frac{1}{2}(-\partial\varphi\partial\varphi + m^2\varphi^2)\,(x)\\ \mathcal{L}_{\text{int}}(\varphi) &= u\varphi^4(x)\end{aligned}$$

we obtain the familiar φ^4 self–interacting Euclidean field theory, which by a Wick rotation is identical to a local relativistic quantum field theory! We note that in QFT language and in statistical mechanics language the terms interaction part and free part of the Hamiltonian are essentially exchanged

statistical mechanics		quantum field theory
$-\frac{1}{2}\sigma K \sigma$	$\Leftrightarrow$	$\mathcal{L}_0(\varphi)$
$\delta(\sigma_x^2-1)$	$\Leftrightarrow$	$\mathcal{L}_{\text{int}}(\varphi)$

On can show that the Ising model has a critical point, which corresponds to vanishing renormalized mass, where the long range behavior of the Ising model $|x| >> a$ is indeed precisely described by an Euclidean φ^4 field theory.

7.1.7 Continuum limit, infinite volume limit

Starting from the Gaussian (free) measure $\prod_x d\mu_{N,a}(\varphi_x)$ we may construct in the standard fashion the function space of square integrable functions:

$$L^2\left(\mathbf{R}^N(V_{L,a}),\;\prod_x d\mu_{N,a}(\varphi_x)\right)$$

and φ_x is called **Gaussian random variable** labeled by $x \in V_{L,a}$. How do these well–defined finite dimensional approximants relate to a continuum field theory?

Originally the problem with taking limits is the following: Formal starting point is the Lebesgue measure

$$\prod_x d\varphi_x$$

which does not make sense in both, **the thermodynamic limit** or **infinite volume limit** $L \to \infty$ with a fixed, and/or, **the continuum limit** $a \to 0$ with L fixed.

The way out is to include into the integration measure the Gaussian factors which damp the Lebesgue measure at $\pm\infty$ exponentially. This is how we obtained the measure Eq. (7.28). For this measure on a suitable function space the limits exist. The limiting space is Feynman's **"path space"** "$\mathbf{R}^\infty(\mathbf{R}^d)$" . More precisely, we shall see that the space of tempered distributions $\mathcal{S}'(\mathbf{R}^d)$ is a suitable space for defining the continuum limit. Which means that the limits

$$d\omega(\varphi) = \lim_{\substack{a\to 0\\ L\to\infty}} \prod_x d\mu_{N,a}(\varphi_x) \tag{7.34}$$

exist as Gaussian measures over $\mathcal{S}'(\mathbf{R}^d)$, i.e., for all $\varphi_x \in \mathcal{S}'(\mathbf{R}^d)$.

In order to show this we note that

$$L^2\left(\mathbf{R}^N(V_{L,a}),\;\prod_x d\mu_{N,a}(\varphi_x)\right) \subset L^2\left(\mathcal{S}'(\mathbf{R}^d),\;d\omega\right) = \left\{F(\varphi)\;;\;\int |F(\varphi)|^2\,d\omega(\varphi) < \infty\right\} \tag{7.35}$$

such that all finite dimensional "approximations" are contained in a fixed space.

This can be seen as follows: consider the **functional Fourier transformation** of the Gauss measure

$$
\begin{aligned}
Z_0\left\{J=ig\right\} &= < e^{i\varphi(g)} >= \int \prod_x d\mu_{N,a}(\varphi_x) e^{i\varphi(g)} \\
&= e^{-\frac{1}{2}||g||^2} \; . \qquad (7.36)
\end{aligned}
$$

We observe that the inverse Fourier transform exists because it is exponentially damped, too. The norm suitable for our purpose is given by

$$
||g||^2 = a^{2d} \sum_{x,y} g(x) G_{a\;x,y}\, g(y) \;\; ; \quad g(x) \in \mathcal{S}(\mathbf{R}^d) \qquad (7.37)
$$

and we used the scalar product

$$
\varphi(g) = a^d \sum_x \varphi_x g(x) \;\; ; \quad g(x) \in \mathcal{S}(\mathbf{R}^d)
$$

in performing the Fourier transform. $\mathcal{S}(\mathbf{R}^d)$ is the test function space of $\mathcal{C}^\infty$–functions (differentiable of infinite order) falling off stronger than any power at $\pm\infty$. Note the difference between the scalar product in Eq. (7.21), where J_x is an external field on the lattice, and the norm in Eq. (7.37), where $g(x) \in \mathcal{S}(\mathbf{R}^d)$ is a test function in the continuum, already before taking the continuum limit.

We first note that $||g||^2$ and $\varphi(g)$ have well defined limits $a \to 0/L \to \infty$

$$
||g||^2 = \int d^dx\, d^dy\; g(x) G_0(x-y) g(y) \quad \text{and} \quad \varphi(g) = \int d^dx\; \varphi(x) g(x)
$$

$\forall g(x) \in \mathcal{S}(\mathbf{R}^d)$. $G_0(x-y)$ is the Euclidean free field propagator in the continuum, and we may easily check directly that $||g||^2$ exists in this limit. The existence of the limit for $\varphi(g)$ is ascertained provided $\varphi(x) \in \mathcal{S}'(\mathbf{R}^d)$. This is just by the definition of the distribution space $\mathcal{S}'(\mathbf{R}^d)$ which is the dual space to $\mathcal{S}(\mathbf{R}^d)$.

Note that the representation formally defined by Eq. (7.34) looses its meaning in the limit. However, $d\omega(\varphi)$ still makes sense as a functional Fourier transform of $e^{-\frac{1}{2}||g||^2}$. The Green functions exist as tempered distributions

$$
< \varphi(x_1) \; \ldots \; \varphi(x_n) >= \int d\omega(\varphi) \varphi(x_1) \; \ldots \; \varphi(x_n) \in \mathcal{S}'(\mathbf{R}^d)^n \qquad (7.38)
$$

and the polynomials are dense in the function space $L^2\left(\mathcal{S}'(\mathbf{R}^d),\; d\omega\right)$. This implies that for the purpose of perturbation theory everything is well under control.

With this in mind, we may perform the following formal steps leading to the continuum form of the path integral: Starting from the Gauss measure Eq. (7.28) we may write

$$
\begin{aligned}
\prod_x e^{-\frac{a^d}{2}\varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x}\, d\varphi_x &= \prod_x d\varphi_x e^{-\frac{a^d}{2}\sum_x \varphi_x\,(-\Delta_a+m^2\,)\,\varphi_x} \\
&\sim \mathcal{D}\varphi(x)\, e^{-\int d^dx\, \frac{1}{2}\,\varphi(x)\,(-\Delta+m^2\,)\,\varphi(x)} \quad ,
\end{aligned}
$$

and rewrite the exponent as follows:

$$
-\int d^dx\, \frac{1}{2}\, \varphi(x)\,(-\Delta+m^2\,)\,\varphi(x) \;\; = \;\; \int d^dx\, \mathcal{L}_0^E(\varphi) = i \int d^dx\, \mathcal{L}_0(\varphi) \;\; .
$$

In the last step we changed back from Euclidean to Minkowski space by setting $x^d = ix^0$. As a result we obtain the well–known suggestive formal expression

$$\boxed{Z_0\{J\} = \int \mathcal{D}\varphi(x)\; e^{i\int d^dx\,(\mathcal{L}(\varphi)+\varphi(x)\,J(x))} / \int \mathcal{D}\varphi(x)\; e^{i\int d^dx\,\mathcal{L}(\varphi)}} \tag{7.39}$$

which is the **path integral** representation of the generating functional, we were looking for. The denominator serves to normalize to $Z_0\{0\} = 1$. Often we will also use the equivalent notation

$$\mathcal{D}\varphi(x) \equiv \prod_x d\varphi(x) \quad , \tag{7.40}$$

because, we think that the second kind of notation is more suggestive; as it is indicating that, the precise meaning is always that of a limit, starting from finite dimensional lattice approximants.

The existence of the Gaussian measure $d\omega(\varphi)$ in the abstract sence, as a Fourier transform, is not very convenient from the point of view of practical calculations. For the latter purpose only the finite dimensional approximations are suitable. In fact, all numerical studies of non-perturbative phenomena in quantum field theory which are based on the path integral representation are starting from lattice approximations.

One might wonder about states and operators in the path integral formulation of field theory. We are not going to develop in more detail how the canonical quantization is related to path integral quantization, however, we briefly mention how contact can be made with the Hilbert space structure.

As we have mentioned already, the (smeared) polynomials in the field are dense in the function space $L^2\left(\mathcal{S}'(\mathbf{R}^d),\; d\omega\right)$. We may define Wick monomials, corresponding to normal ordered products of fields (see Sec. 3.4.2), with the help of the generating functional

$$e^{t\varphi(g)} = e^{-\frac{1}{2}t^2 <\varphi(g)^2>} \;:\; e^{t\varphi(g)} \;: \quad .$$

One obtains the orthogonality relations

$$<:\; \varphi^n(f)\; :\,:\; \varphi^m(g)\; :>= n!\,\delta_{nm}\; <\; \varphi(f)\,\varphi(g)\; >^n$$

which means that the subspace $\mathcal{H}_n \subset L^2$ spanned by

$$\text{“}:\; \varphi^n(f)\; :>\text{”} = \mathcal{H}_n$$

when f runs over all of $\mathcal{S}$, are orthogonal

$$\mathcal{H}_n \perp \mathcal{H}_m\;;\;\; n \neq m \;\; .$$

Thus: L^2 has Fock space structure

$$L^2\left(\mathcal{S}'(\mathbf{R}^d),\; d\omega\right) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n \;\; ,$$

and this provides the bridge back to canonical quantization!

7.1.8 The principle of least action

Generally, a functional $F\{J\}$ is a quantity F which depends on a space-time function $J(x)$. The functional derivative $\frac{\delta F\{J\}}{\delta J(x)}$ is defined by the infinitesimal variation of $F\{J\}$ as a result of an infinitesimal variation of $J(x)$:

$$\delta F = \int d^4x \, \frac{\delta F\{J\}}{\delta J(x)} \, \delta J(x)$$

which is equivalent to

$$\frac{\delta F\{J\}}{\delta J(x)} = \lim_{\epsilon \to 0} \frac{F\{J'\} - F\{J\}}{\epsilon}$$

with $J'(y) = J(y) + \epsilon \, \delta^{(4)}(y - x)$ or, equivalently,

$$\int d^4x \, f(x) \frac{\delta F\{J\}}{\delta J(x)} = \lim_{\epsilon \to 0} \frac{F\{J + \epsilon \, f\} - F\{J\}}{\epsilon}$$

for smooth test functions $f(x)$. Note that $F\{J + \epsilon f\}$ is an ordinary function of ϵ and provided it is sufficiently differentiable may be expanded in ϵ

$$\begin{aligned} F\{J + \epsilon f\} &= F\{J\} + \epsilon \left(\frac{\partial F\{J + \epsilon f\}}{\partial \epsilon} \right)_{\epsilon = 0} + \cdots \\ &= F\{J\} + \epsilon \int d^4x \, f(x) \frac{\delta F\{J\}}{\delta J(x)} + \cdots \quad . \end{aligned}$$

Let $\mathcal{L}\{\varphi\}(x)$ be the Lagrangian density depending on some dynamical classical variables $\varphi(x)$, then

$$S(\varphi) \doteq \int d^dx \, \mathcal{L}\{\varphi\}(x) \tag{7.41}$$

is the **classical action**. The dynamical behavior of a system is determined by **the principle of least action**, which requires the physically acceptable values of a dynamical variable to have **stationary action**:

$$\frac{\delta S}{\delta \varphi(x)} = 0 \quad . \tag{7.42}$$

Each solution of this "equation of motion" corresponds to a classically allowed "history", i.e., a **time evolution**.

Remark: The explicit form of $\mathcal{L}\{\varphi\}(x)$ is highly non–unique. One may replace $\varphi(x)$ by an arbitrary functional of $\varphi(x)$ without changing the dynamics (stationary points). This fact is usually referred to as the **equivalence theorem**.

We are now ready to present the functional derivation of the equation of motion. Consider the generating functional

$$Z_0\{J\} = \int_{\otimes} \prod d\varphi(x) \; e^{i \int d^dx \, \mathcal{L}(\varphi) + i \int d^dx \, \varphi(x) \, J(x)}$$

where $\otimes$ indicates normalization. We now perform an infinitesimal transformation of the integration variables

$$\varphi(x) \to \varphi'(x) = \varphi(x) + \delta\varphi(x)$$

with

$$\det\left(\frac{\partial\varphi(x)}{\partial\varphi'(x)}\right)=1\ \Rightarrow\ \prod d\varphi'(x)=\prod d\varphi(x)\quad .$$

The variation of the action is given by

$$S(\varphi')\ =\ \int d^dx\,\mathcal{L}(\varphi')\ =\ S(\varphi)+\delta S(\varphi)$$

with

$$\delta S(\varphi)\ =\ \int d^dx\ \left(\frac{\partial\mathcal{L}}{\partial\varphi}\,\delta\varphi+\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\,\delta\partial_\mu\varphi\right)\ =\ \int d^dx\ \left(\frac{\partial\mathcal{L}}{\partial\varphi}\,\delta\varphi-\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\,\delta\varphi\right)$$

where we have performed a partial integration, and assumed that $\delta\partial_\mu\varphi=\partial_\mu\delta\varphi$. For the source term we have

$$\varphi'(x)J(x)=\varphi(x)J(x)+J(x)\delta\varphi(x)\quad .$$

Since the value of the functional integral does not depend on the choice of the integration variable its value remains constant and hence:

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\ -\ \frac{\partial\mathcal{L}}{\partial\varphi}\ =\ J(x) \tag{7.43}$$

up to a total divergence. Since the Lagrangian itself is determined only up to a divergence one can chose the latter appropriately. For the self–interacting $\lambda\varphi^4$–theory considered in Sec. 3 (see Eq. (3.5)) the Lagrangian reads

$$\mathcal{L}(x)=\frac{1}{2}\left(\partial\varphi\partial\varphi+m^2\varphi^2\right)(x)+\frac{\lambda}{4!}\varphi^4(x)$$

and we obtain the known field equation Eq. (3.6)

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\ -\ \frac{\partial\mathcal{L}}{\partial\varphi}\ =\ J(x)\ =\ (\ \Box+m^2\ )\,\varphi(x)+\frac{\lambda}{3!}\varphi^3(x)$$

As a result we note the equivalence statement:

The dynamical variable $\varphi(x)$ satisfies the classical equation of motion (with source) $\Leftrightarrow$ the principle of least action holds.

Functional form of the equation of motion:

The generating functional $Z\{J\}$ may be understood as the **vacuum functional** in the presence of an external source. We therefore write

$$Z\{J\}\quad=\quad<1>_J$$

and similarly

$$\frac{\delta Z\{J\}}{\delta J(x)}\quad=\quad i<\ \varphi(x)\ >_J$$

etc.

The equation of motion then may be written in the form

$$(\,\Box + m^2\,)\,\frac{\delta Z\,\{J\}}{\delta J(x)} \;=\; i\,(\,\Box + m^2\,)\;<\,\varphi\,>_J = \;i\,J(x)\;<1>_J - i < \;\frac{\partial \mathcal{L}_{\mathrm{int}}}{\partial \varphi}(x)\;>_J \tag{7.44}$$

The free functional satisfies

$$\begin{aligned}(\,\Box + m^2\,)\,\frac{\delta Z_0\,\{J\}}{\delta J(x)} &= -\frac{i}{2} Z_0\,\{J\}\;(\,\Box + m^2\,)\,\frac{\delta}{\delta J(x)}\,(J, \Delta_F J) \\ &= i\,J(x)\,Z_0\,\{J\}\;\;. \end{aligned}\tag{7.45}$$

This may be checked easily. Using

$$\begin{aligned}\frac{\delta}{\delta J(x)}\int d^d y_1\, d^d y_2\, J(y_1)\,\Delta_F(y_1 - y_2)\,J(y_2)& \\ = 2\int d^d y\,\Delta_F(x-y)\,J(y)&\end{aligned}$$

and Eq. (7.5) we have

$$\begin{aligned}(\,\Box + m^2\,)\,\frac{\delta}{\delta J(x)}\,(J, \Delta_F J) &= 2\int d^d y\,(\,\Box + m^2\,)\,\Delta_F(x-y) \\ &= -2\,J(x)\;\;,\end{aligned}$$

which verifies Eq. (7.45).

The field equation (7.44) may be rewritten in the form of an integral equation

$$\left\|\;\frac{\delta Z\{J\}}{\delta J(x)} \;=\; -i\int d^d y\,\Delta_F(x-y)\,\left\{J(x)\,Z\,\{J\} + <\;\frac{\partial \mathcal{L}_{\mathrm{int}}}{\partial\varphi}(y)\;>_J\right\}\;\;. \right. \tag{7.46}$$

The original form (7.44) is obtained by applying the Klein–Gordon operator to the integral equation and by using (7.5). Expanding (7.46) into a power series in J and matching powers of J, yields the equations of motion for all Green functions.

7.1.9 Path integral quantization of an interacting theory

In canonically quantized quantum field theory we are mostly dealing with S–matrix elements

$$<\;k_1\alpha_1,\;\ldots\;,k_n\alpha_n\,|\,\mathbf{S}\,|\,k'_1\alpha'_1,\;\ldots\;,k'_{n'}\alpha'_{n'}\;>$$

which are related to the time–ordered Green functions

$$<0|\;T\{\varphi(x_1)\;\cdots\;\varphi(x_m)\}|0>$$

as inferred by the LSZ reduction formula. The Gell-Mann–Low formula,

$$<0|\;T\{\varphi(x_1)\;\cdots\;\varphi(x_m)\}|0> = \;{}_{in}<0|T\;\{\varphi_{in}(x_1)\;\cdots\;\varphi_{in}(x_m)\,\mathbf{S}\}|0>_{in\,\otimes}$$

with

$$\mathbf{S} = Te^{i\int d^d x\,\mathcal{L}^{(in)}_{\mathrm{int}}(\varphi_{in})}$$

the scattering matrix represented in terms of the free *in*–fields, then allows us to calculate Green functions and hence S–matrix elements perturbatively, by expanding the exponential form of **S** into a power series in $\mathcal{L}_{\text{int}}$. Here we note that one treated the free part $\mathcal{L}_0$ and the interaction part $\mathcal{L}_{\text{int}}$ of the Lagrangian on unequal footing.

We now consider the generating functional for the time–ordered Green functions of the interacting theory: By definition and applying the Gell-Mann–Low formula we obtain in a first step

$$\begin{aligned} Z\{J\} &= \, < 0 \,|\, T\, e^{i\int d^dx\, \varphi(x)\, J(x)} \,|\, 0 > \\ &= {}_{in}\!< 0 \,|\, T\, e^{i\int d^dx\, \varphi_{in}(x)\, J(x)}\, \mathbf{S} \,|\, 0 >_{in} \\ &= {}_{in}\!< 0 \,|\, T\, e^{i\int d^dx\, \mathcal{L}^{(in)}_{\text{int}}(\varphi_{in})}\, e^{i\int d^dx\, \varphi_{in}(x)\, J(x)} \,|\, 0 >_{in\,\otimes} \quad . \end{aligned} \tag{7.47}$$

We then may use the fact that, any functional of the free field φ_{in}, except for the free functional itself, may be evaluated by replacing the free field by the functional derivative with respect to the source $-i\frac{\delta}{\delta J(x)}$. For non-polynomial functionals we may easily show this to be true by expanding the functional into a series of polynomials. We thus may write

$$Z\{J\} = e^{i\int d^dx\, \mathcal{L}_{\text{int}}(-i\frac{\delta}{\delta J(x)})}\; {}_{in}\!< 0 \,|\, T\, e^{i\int d^dx\, \varphi_{in}(x)\, J(x)} \,|\, 0 >_{in\,\otimes} \tag{7.48}$$

and then use the representation of the free functional as a path integral

$${}_{in}\!< 0 \,|\, T\, e^{i\int d^dx\, \varphi_{in}(x)\, J(x)} \,|\, 0 >_{in\,\otimes} = Z_0\{J\} = e^{-\frac{i}{2}\,(J,\Delta_F J)} = \int_{\otimes} \mathcal{D}\varphi(x)\, e^{i\int d^dx\, (\mathcal{L}_0(\varphi)+\varphi J)} \quad .$$

In doing so, we find

$$\begin{aligned} Z\{J\} &= e^{i\int d^dx\, \mathcal{L}^{(in)}_{\text{int}}(-i\frac{\delta}{\delta J(x)})} \int_{\otimes} \mathcal{D}\varphi(x)\, e^{i\int d^dx\, (\mathcal{L}_0(\varphi)+\varphi J)} \\ &= \int_{\otimes} \mathcal{D}\varphi(x)\, e^{i\int d^dx\, (\mathcal{L}_0(\varphi)+\mathcal{L}_{\text{int}}(\varphi)+\varphi J)} \\ &= \int_{\otimes} \mathcal{D}\varphi(x)\, e^{i\int d^dx\, (\mathcal{L}(\varphi)+\varphi J)} \end{aligned} \tag{7.49}$$

the promised path integral for the interacting theory. Most remarkably, only the total Lagrangian appears in the path integral. The splitting into a free and an interacting part of the Lagrangian is not necessary any more. The derivation presented above showed at the same time how the path integral is used in perturbation theory, where the complete equivalence of both schemes are manifest. The prescription $\otimes$ again indicates the normalization $Z\{0\} = 1$ to be imposed. The time–ordered Green functions are now given by

$$\begin{aligned} (-i)^n \frac{\delta^n}{\delta J(x_1)\, \ldots\, J(x_n)}\; Z\{J\}\, |_{J=0} &= \, < \varphi(x_1)\, \ldots\, \varphi(x_n) > \\ &= \int_{\otimes} \mathcal{D}\varphi(x)\, \varphi(x_1)\, \ldots\, \varphi(x_n)\, e^{i\int d^dx\, (\mathcal{L}(\varphi))} \\ &\equiv \, < 0 \,|\, T\{\varphi(x_1) \cdots \varphi(x_m)\} | \, 0 > \end{aligned} \tag{7.50}$$

which is our final result for the path integral quantization of scalar fields. Of course in the interacting case we will encounter the same renormalization problems, as in the canonical quantization. The finite dimensional lattice approximation

$$Z\{J\}^{\text{reg}} = \int_{\otimes} \prod_{x \in V_{L,a}} d\varphi_x\, e^{ia^d \sum_x (\mathcal{L}(\varphi)_x + \varphi_x J_x)} \tag{7.51}$$

with

$$ia^d \sum_x \mathcal{L}(\varphi)_x = -a^d \sum_x \left(\frac{1}{2}\varphi_x\, (-\Delta_a + m^2\,)\, \varphi_x + \mathcal{L}_{\text{int}}(\varphi)_x \right) \tag{7.52}$$

(rotated to Euclidean space) is a regularization for both ultraviolet singularities showing up in continuum limit and infinite system divergences which are related to taking the thermodynamical limit. This regularization is called **lattice regularization**. The "lattice theory" usually is the starting point for investigations beyond perturbation theory.

Advantages of the Feynman path integral representation:

❶ Only classical fields appear, everything commutes, there is no time ordering etc. The role of non–trivial classical solutions of the field equations, like instantons, for example, within the path integral formulation is much more transparent, than in the canonical quantization approach. They are just particular contributions to the path integral in accord with the action principle.

❷ Only the full Lagrangian $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}}$ enters the path integral representation for the generating functional of the Green functions. This representation is therefore intrinsically non-perturbative! In particular, it does not depend on the splitting into a free and an interaction part. For non–abelian gauge theories this will be crucial for controlling gauge invariance, which is a property of the full invariant Lagrangian, only (see Sec. 6).

✘ A related observation is that the path integral does not refer to LSZ interpolating fields, which bridge to the free scattering states of the physical spectrum. Such a "particle interpretation" of fields used to parametrize the Lagrangian need not be a priori known. The path integral formulation thus is more flexible as a starting point. QCD is the prototype example of an intrinsically non–perturbative theory. We know that quarks and gluons, in terms of which the QCD Lagrangian must be written, do not relate to scattering states in any simple way. Rather we know that quarks and gluons are permanently confined inside the hadrons and only the latter appear as scattering states in Nature. In a restricted sense perturbation theory may still apply. Due to asymptotic freedom quarks and gluons "show up" as jets of hadrons in highly energetic processes.

Problems with the path integral:

❶ In Minkowski space the integration over the time component x^0 is not a priori well–defined (oscillatory integrals) and one has to apply an $i\epsilon$–prescription, at least. In fact one is dealing with distributions and not with functions. This problem may be circumvented by using analyticity in x^0 and to Wick rotate from Minkowski to Euclidean space, where one works with positive definite metric. At the end results have to be analytically continued back to Minkowski space. For numerical methods this may be an unsurmountable obstacle since errors usually run out of control when we try to do the analytical continuation numerically.

❷ For interacting theories the existence of the continuum limit is a serious problem. Either the continuum limit is defined in a purely formal manner, like for example in perturbation theory, where one defines the objects order by order in perturbation theory. Or, one starts from the finite dimensional approximants (finite elements, lattice) where one cannot avoid to violate basic principles like Poincaré covariance, chiral symmetry etc. One must prove that the violated properties are **restored** in the continuum limit. In Wilson's formulation of lattice gauge theories gauge invariance is preserved if there are no chiral fermions.

❸ The "path integral" formulation of fermions is a notorious problem as we shall see below.

Perturbation theory based on the path integral:

In the canonical quantization formulation the perturbation expansion is based on the Gell-Mann–Low formula, which reads, in the expanded form,

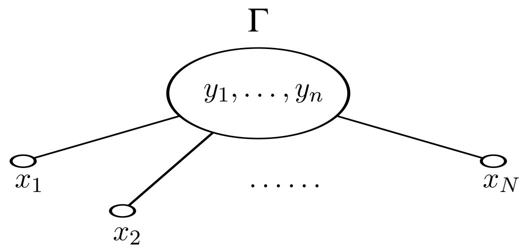
$$< 0\,|\,T\{\varphi(x_1)\,\cdots\,\varphi(x_m)\}|\,0> = \sum_{n=0}^{\infty} \tfrac{i^n}{n!} \int d^4y_1\,\ldots\,d^4y_n\,\times$$
$$_{in}< 0\,|\,T\left\{\varphi_{in}(x_1)\,\cdots\,\varphi_{in}(x_m)\,\mathcal{L}^{(in)}_{\rm int}(y_1)\,\ldots\,\mathcal{L}^{(in)}_{\rm int}(y_n)\right\}|\,0>_{in\,\otimes}$$

and expresses the Green functions in terms of free quantized fields $\varphi_{in}(x)$. The time-ordered Green functions of free fields can be evaluated by using Wick's theorem: which is based upon representing the fields and the states in terms of creation and annihilation operators $a_{in}(p,r)$ and $a^+_{in}(p,r)$ and using the canonical commutation relations together with the vacuum annihilation property $a_{in}(p,r)\,|\,0>= 0$. This is the way we derived the Feynman rules previously in see Sec. 3.

In the functional formulation the analogous expansion takes the form:

$$\begin{aligned} &< 0\,|\,T\{\varphi(x_1)\,\cdots\,\varphi(x_N)\}|\,0> = <\varphi(x_1)\,\ldots\,\varphi(x_n)\,> \\ = \;& (-i)^N\,\frac{\delta^N}{\delta J(x_1)\,\ldots\,J(x_N)}\;Z\left\{J\right\}\,|_{J=0} = (-i)^N \sum_{n=0}^{\infty}\frac{i^n}{n!}\int d^4y_1\,\ldots\,d^4y_n\,\times \\ & \frac{\delta^N}{\delta J(x_1)\,\ldots\,J(x_N)}\;\mathcal{L}_{\rm int}(-i\frac{\delta}{\delta J(y_1)})\,\ldots\,\mathcal{L}_{\rm int}(-i\frac{\delta}{\delta J(y_n)})\;e^{-\frac{i}{2}(J,\Delta_F J)}\Big|_{J=0,\otimes} \end{aligned} \tag{7.53}$$

$$= \sum_{\Gamma \atop \otimes}$$



where the sum extends over all non-vacuum ($\otimes$) Feynman diagrams Γ with external vertices ∘—— , internal vertices $\times$ all lines contracted pairwise by propagators ∘——∘ (see Sec. 3.4.4).

The evaluation is purely a matter of differentiating the Gauss functional and setting $J = 0$. Evaluating the non–vanishing contributions in the perturbation expansion leads to Feynman rules which are identical to the one's obtained in the canonical quantization approach.

Exercises: Subsection 7.1

① Show that the free functional

$$Z_0\left\{J\right\} = \exp -\frac{i}{2}\,(J,\Delta_F J)$$

can be defined as a generalized Gauss integral in Minkowski space provided we replace the mass m^2 by $m^2 - i\epsilon$ with $\epsilon > 0$, infinitesimal, and then perform the limit $\epsilon \to 0$ after integration.

② Let $\varphi(x)$ be a complex classical field and

$$Z\left\{J^*,J\right\} = \int\limits_{-\infty}^{+\infty}\!\!\int\limits_{\otimes} \prod_x d\varphi^*_x\,d\varphi_x\;e^{i\int d^dx\,(\mathcal{L}_0(\varphi,\varphi^*)+J^*\varphi+J\varphi^*)}$$

the corresponding free functional. Transform this to a functional integral for the two real fields Reφ and Imφ. Write down the appropriate lattice regularization.

③ Describe your own understanding of how the functional integral

$$\int \mathcal{D}\varphi(x) e^{i\int d^d x \, (\mathcal{L}_0(\varphi)+J\varphi)}$$

can be given an unambiguous meaning.

④ Investigate, in the context of the action principle, the "Lagrangian densities" $\mathcal{L}(x) = \varphi(x)$ and $\mathcal{L}(x) = \frac{1}{2}\left(\partial\varphi\partial\varphi + m^2\varphi^2\right)(x) + \frac{\lambda}{3!}\varphi^3(x)$. Compare the latter with the $\lambda\varphi^4$–theory.

⑤ Derive the Feynman rules for the $\lambda\varphi^4$–theory with the help of the functional method. Use

$$\exp -\frac{i}{2}(J, \Delta_F J) = \sum_{m=0}^{\infty} \frac{(-\frac{i}{2})^m}{m!}(J, \Delta_F J)^m$$

and evaluate the terms which do not vanish for $J = 0$. For N external fields and n interaction vertices we have to take $N + 4n$ derivatives. Hence, only the term with $m = (N + 4n)/2$ in the above sum gives a non-vanishing term.

⑥ Using Eq. (7.53), evaluate explicitly the contributions for $N = 0, 1, 2, 3, 4$ and $n = 0, 1, 2$ in $\lambda\varphi^4$–theory. Draw all corresponding Feynman diagrams.

7.2 Functional integral for fermions

7.2.1 Generating functional for fermions

The analogue of the generating functional for time ordered Green functions for Bose fields, in the case of fermions takes the form:

$$\begin{aligned} Z_0\{\bar{\zeta}\ ,\zeta\} &= \ <0|Te^{i\int d^d x\,\left(\bar{\zeta}(x)\psi(x)+\bar{\psi}(x)\zeta(x)\right)}|0> \\ &= \textstyle\sum_{n=0}^{\infty}\frac{i^{2n}}{(n!)^2}\int dz_1\dots dz_n dy_1\dots dy_n \ \times \qquad\qquad (7.54) \\ \times\bar{\zeta}(z_n)\dots\bar{\zeta}(z_1) & \ <0|T\left\{\psi(z_1)\cdots\psi(z_n)\bar{\psi}(y_n)\cdots\bar{\psi}(y_1)\right\}|0> \ \zeta(y_n)\dots\zeta(y_1)\ . \end{aligned}$$

Here $\psi(x)$ is a quantized free Dirac field and satisfies Fermi statistics and corresponding anti-commuting rules. The time ordered Green functions $<0|T\left\{\psi(x)\cdots\bar{\psi}(y)\cdots\right\}|0>$ are totally antisymmetric , since under the time ordering prescription the fields are strictly anticommuting. Therefore, the classical external **sources** $\bar{\zeta}(x)$ and $\zeta(x)$ must be **anticommuting classical variables**, which are called **Grassmann variables**. A simple mnemonic rule for dealing with anticommuting fields and variables is the following: Consider ζ and $\bar{\zeta}$ to be free Dirac fields in the classical limit $\hbar \to 0$, where

$$\{\zeta(x)\ ,\ \zeta(y)\} = 0\ ; \ \ \{\bar{\zeta}(x)\ ,\ \bar{\zeta}(y)\} = 0 \ \text{ and } \ \ \{\zeta(x)\ ,\ \bar{\zeta}(y)\} = O(\hbar) \to 0\ .$$

It is a useful conventions to take Fermi fields describing different particles (electrons, muons, protons,....) anticommuting relative to each other. Consequently, it is natural and convenient to also consider the source functions ζ (as Dirac fields in the classical limit) to be anticommuting with the field operators ψ. Note that the source terms

$$\bar{\zeta}\psi \ \text{ and } \ \bar{\psi}\zeta$$

appearing in the generating functional are written as Lorentz scalars.

Grassmann algebras:

We first consider the algebra of the ζ's and $\bar{\zeta}$'s for a finite number of variables: The $2N$ algebraic quantities ζ_i and $\bar{\zeta}_i$, $i = 1, \ldots, N$ are required to satisfy the anti commutation relations

$$\{\zeta_i\ ,\ \zeta_k\} = 0\ ;\ \ \{\bar{\zeta}_i\ ,\ \bar{\zeta}_k\} = 0\ ;\ \ \{\zeta_i\ ,\ \bar{\zeta}_k\} = 0\ \ .$$

Therefore, they satisfy:

$$\zeta_i^2 = 0 = \bar{\zeta}_i^2\ \ .$$

The elements ζ_i, $\bar{\zeta}_i$ are the **generators** of the **Grassmann algebra**, the elements of which are polynomials in the ζ_i's and $\bar{\zeta}_i$'s with complex antisymmetric tensor coefficients $T_{m,n}(i_1 \ldots i_m | i'_1 \ldots i'_n)$:

$$F(\zeta, \bar{\zeta}) = \sum_{m,n} \bar{\zeta}_{i_m} \cdots \bar{\zeta}_{i_1} T_{m,n}(i_1 \ldots i_m | i'_1 \ldots i'_n) \zeta_{i'_1} \cdots \zeta_{i'_n}\ \ .$$

As usual in tensor calculus we will adopt the summation convention: doubly curing indices are summed over if not stated otherwise.

Elements of the Grassmann algebra can be added and multiplied with scalars ($\in C$) and hence form a **vector space** of dimension:

$$2^{2N} = \left(\sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} \right) \cdot \left(\sum_{n=0}^{N} \begin{pmatrix} N \\ n \end{pmatrix} \right)$$

An element is called **even** if it **commutes** and **odd** if it **anticommutes**, respectively, with all generators. By the anticommutation relation also products

$$F_1(\zeta, \bar{\zeta}) \cdot F_2(\zeta, \bar{\zeta})$$

are defined. A list of simple monomials in the ζ_i's is:

$$1,\ \zeta_1,\ \ldots,\ \zeta_N,\ \zeta_1\zeta_2,\ \ldots,\ \zeta_{N-1}\zeta_N,\ \ldots,\ \zeta_1\zeta_2 \ldots \zeta_N\ \ .$$

The **differentiation** of Grassmann variables is defined algebraically as follows: For generating elements we have the **left–derivatives** :

$$\begin{array}{ll} \frac{\partial}{\partial \zeta_i} \zeta_k = \delta_{ik}\ ; & \frac{\partial}{\partial \bar{\zeta}_i} \bar{\zeta}_k = \delta_{ik} \\ \frac{\partial}{\partial \zeta_i} \bar{\zeta}_k = 0\ ; & \frac{\partial}{\partial \bar{\zeta}_i} \zeta_k = 0 \\ \frac{\partial}{\partial \zeta_i} 1 = 0\ s; & \frac{\partial}{\partial \bar{\zeta}_i} 1 = 0 \end{array}$$

and as a consequence

$$\left(\frac{\partial}{\partial \zeta_i} \right)^2 = 0 = \left(\frac{\partial}{\partial \bar{\zeta}_i} \right)^2$$

on the entire Grassmann algebra. For products of monomials $F_1 \cdot F_2$ the product rule reads:

$$\frac{\partial}{\partial \zeta_i} (F_1 \cdot F_2) = \frac{\partial F_1}{\partial \zeta_i} \cdot F_2 \pm F_1 \cdot \frac{\partial F_2}{\partial \zeta_i}\ \ ,\quad \text{for} \begin{cases} F_1 & even \\ F_1 & odd\ . \end{cases}$$

The set of elements ζ_i, $\bar{\zeta}_i$, $\frac{\partial}{\partial \zeta_i}$, $\frac{\partial}{\partial \bar{\zeta}_i}$ defines a Grassmann algebra of $2 \times 2N$ elements, the **extended algebra**. A right–derivative may be defined accordingly.

In the **continuum limit** $N \to \infty$ we have the correspondence

$$\sum_{\cdots i_k \cdots} \cdots \bar{\zeta}_{i_k} \cdots T_{m,n}(\cdots i_k \cdots | \cdots) \cdots \Rightarrow \int d^d z_k \; \cdots \bar{\zeta}(z_k) \cdots T_{m,n}(\cdots z_k \cdots | \cdots) \cdots$$

$$\frac{\partial}{\partial \zeta_i} \Rightarrow \frac{\delta}{\delta \zeta(y)} \text{ with } \frac{\delta^2}{\delta \zeta(x_1) \delta \zeta(x_2)} = -\frac{\delta^2}{\delta \zeta(x_2) \delta \zeta(x_1)}$$

$$\frac{\delta}{\delta \zeta(x)} \zeta(x_1) \zeta(x_2) = \delta(x - x_1) \zeta(x_2) - \delta(x - x_2) \zeta(x_1)$$

After this short digression to Grassmann algebras we turn back to our original problem the

Calculation of the free generating functional $Z_0 \{\bar{\zeta}\ , \zeta\}$**:**

For free Dirac fields we can calculate

$$< 0 | T \left\{ \psi(z_1) \cdots \psi(z_n) \bar{\psi}(y_n) \cdots \bar{\psi}(y_1) \right\} | 0 >$$

easily. All ψ's have to be contracted with $\bar{\psi}$'s in pairs in all possible ways. To each contraction there corresponds a Feynman propagator $i\, S_F(z_i - y_k)$ up to a sign. The sign is obtained by noting that $\bar{\zeta}\psi$ and $\bar{\psi}\zeta$ commute with fermionic variables. For convenience one therefore writes the sources always together with the fields and performs the contractions:

$$\begin{array}{lll} \text{1st} & \text{pair} & n \text{ possible } z\text{'s}, \qquad\qquad n \text{ possible } y\text{'s} \\ & & \Rightarrow n^2 \times i\bar{\zeta}(z_n) S_F(z_n - y_n) \zeta(y_n) \\ \text{2nd} & \text{pair} & n-1 \text{ possible } z\text{'s}, \qquad n-1 \text{ possible } y\text{'s} \\ & & \Rightarrow (n-1)^2 \times i\bar{\zeta}(z_{n-1}) S_F(z_{n-1} - y_{n-1}) \zeta(y_{n-1}) \\ \vdots & & \end{array}$$

with integration over all z's and y's. This yields

$$(n!)^2 \left(i \left(\bar{\zeta}, S_F\, \zeta \right) \right)^n$$

where

$$(\bar{\zeta}, S_F\, \zeta) = \int d^d x\, d^d y\, \bar{\zeta}(x) S_F(x - y) \zeta(y) \quad .$$

So far we have counted the $n!$ equivalent permutations of the n factors as different such that we still have to divide by $n!$. We thus obtain for the generating functional the following

Result:

$$\begin{aligned} Z_0 \{\bar{\zeta}\ , \zeta\} &= < 0 | T e^{i \int d^d x \left(\bar{\zeta}(x) \psi(x) + \bar{\psi}(x) \zeta(x) \right)} | 0 > \\ &= \sum_{n=0}^{\infty} \frac{(i)^{2n}}{(n!)^2} \frac{(i)^n (n!)^2}{n!} (\bar{\zeta}, S_F\, \zeta)^n = \exp -i (\bar{\zeta}, S_F\, \zeta) \quad . \end{aligned} \tag{7.55}$$

This is again an exponential of a bilinear form, however, with the crucial difference that the form is **antisymmetric** and not symmetric as in the Bose case. Nevertheless, it seems suggestive to ask the question whether we can write the generating functional as an integral

$$\text{``}\quad Z_0 \{\bar{\zeta}\ , \zeta\} = \int_{\otimes} \prod_x d\bar{\psi}_x d\psi_x e^{i \int d^d x \left(\mathcal{L}_0(\bar{\psi}, \psi) + \bar{\zeta}\psi + \bar{\psi}\zeta \right)} \quad \text{''}$$

and if yes, it is obvious that $\bar{\psi}_x$ and ψ_x must be Grassmann variables, classical anticommuting c–number fields, which satisfy the same algebra as the $\bar{\zeta}_x$ and ζ_x and must anticommute with them.

7.2.2 Matrices, integral representations of determinants

The main problem in dealing with fermions is the calculation of functional determinants. We therefore have to study first some general properties of $n \times n$ matrices.

Properties of square matrices:

1.

Let A be an arbitrary $n \times n$ matrix, then there exists a unitary transformation S: $SS^+ = S^+S = 1$ such that

$$S^{-1}AS = \Delta_A$$

has trigonal form

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad ; \quad \Delta_A = \begin{pmatrix} a'_{11} & \cdots & a'_{1n} \\ 0 & \ddots & \\ 0 & \cdots & a'_{nn} \end{pmatrix} \quad .$$

2.

We note that

$$\det A = \det \left(S^{-1}AS\right) = \det\left(\Delta_A\right) = \prod_i a'_{ii}$$

is a simple consequence of $\det(AB) = \det(A)\det(B)$ and $\det S^{-1} = (\det S)^{-1}$ for $\det S \neq 0$.

3.

Similarly,

$$\mathrm{Tr}\,(A) = \mathrm{Tr}\,(S^{-1}AS) = \mathrm{Tr}\,(\Delta_A) = \sum_i a'_{ii}$$

follows from the property $\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$ of a trace.

4.

For triangular matrices we have

$$(\Delta_A\,\Delta_B)_{ii} = (\Delta_A)_{ii}\,(\Delta_B)_{ii}$$

as may be easily checked.

5.

If Δ is trigonal then also $I \pm \Delta$ is a triangular matrix, with I the $n \times n$ unit matrix.

6.

The exponential mapping

$$B = e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

exists and is one–to–one and invertible for $||A||$ sufficiently small. As a norm we may take $||A|| \doteq n \times \max_{i,k} |a_{ik}|$ and it is sufficient to have $||B|| < ||I|| = n$ in order to obtain

$$A = \ln B = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} (B - I)^m \quad .$$

Furthermore we easily verify the properties:

$$\left\| \begin{array}{c} B^T = e^{A^T} \quad , \quad B^* = e^{A^*} \quad , \quad B^+ = e^{A^+} \quad , \quad B^{-1} = e^{-A} \\ S^{-1}BS = e^{S^{-1}AS} \end{array} \right.$$

Theorem 1: For <u>any</u> square matrix A the identity

$$\boxed{\det e^A = e^{\mathrm{Tr}\,A}} \tag{7.56}$$

holds and $\det e^A$ is never singular.

Proof: There exists a unitary transformation S of A to trigonal form: $S^{-1}AS = \Delta_A$. By the properties listed above we have: $\det(e^A) = \det(S^{-1}e^A S) = \det(e^{S^{-1}AS}) = \det(e^{\Delta_A}) = \prod_i e^{a'_{ii}} = e^{\sum_i a'_{ii}} = e^{\mathrm{Tr}\,\Delta_A} = e^{\mathrm{Tr}\,A}$. Note that

$$\Delta_B = e^{\Delta_A} = \sum_{n=0}^{\infty} \frac{\Delta_A^n}{n!}$$

is a trigonal matrix with $\Delta_{B\,ii} = e^{\Delta_{A\,ii}}$ q.e.d.

Theorem 2: For <u>any</u> square matrix L the identity

$$\boxed{\det(I+L) = e^{\mathrm{Tr}\,\ln(I+L)}} \tag{7.57}$$

holds, for $||L||$ sufficiently small. Equivalently,

$$\boxed{\det M = e^{\mathrm{Tr}\,\ln M}}$$

for $||M|| < ||I|| = n$. By the latter condition we have

$$\ln M = \ln(I + (M-I)) = -\sum_{m=1}^{\infty} \frac{(I-M)^m}{m} \quad .$$

Proof: Let $M = I + L$ such that $\ln M = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} L^m$. There exists a unitary transformation such that $S^{-1}MS = \Delta_M$ is triangular. Now we have

$$\begin{aligned} & \mathrm{Tr}\,\ln \Delta_M = \mathrm{Tr}\,\ln(S^{-1}MS) = \mathrm{Tr}\,\ln(I + S^{-1}LS) \\ = \;& \mathrm{Tr} \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} S^{-1}L^m S = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \mathrm{Tr}\, S^{-1}L^m S \\ = \;& \mathrm{Tr}\,\ln(I+L) = \mathrm{Tr}\,\ln M \quad . \end{aligned}$$

Also, we note that

$$\mathrm{Tr}\,\ln \Delta_M = \mathrm{Tr}\,\ln(I + \Delta_L) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \mathrm{Tr}\,\Delta_L \quad .$$

Since Δ_M is triangular, also $\Delta_L = \Delta_M - I$ is triangular. In addition: $(\Delta_L^m)_{ii} = (\Delta_{L\,ii})^m$ such that

$$\begin{aligned} \mathrm{Tr}\,\ln \Delta_M &= \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \left[(\Delta_{L\,11})^m + (\Delta_{L\,22})^m + \cdots + (\Delta_{L\,nn})^m \right] \\ &= \ln(I + \Delta_{L\,11}) + \cdots + \ln(I + \Delta_{L\,nn}) \\ &= \ln \Delta_{M\,11} + \cdots + \ln \Delta_{M\,nn} \\ &= \ln \prod_i \Delta_{M\,ii} = \ln\det \Delta_M \\ &= \ln\det M \end{aligned}$$

q.e.d.

This closes our digression to general properties of square matrices. We now consider how a determinant may be represented by Feynman diagrams.

Determinants versus diagrams: bosons

We consider the vacuum to vacuum transition amplitude in the path integral representation for a complex scalar field in interaction with an external potential $V(x)$ (see Eq. (7.19)):

$$< 1 >_V = \int_{-\infty}^{+\infty} \prod_x \frac{d\varphi_x}{\sqrt{2\pi i}} \frac{d\varphi_x^*}{\sqrt{2\pi i}} \, e^{i\int d^dx\, \varphi^*(x)\left(-(\Box+m^2)+V(x)\right)\varphi(x)} = (\det M(x,y))^{-1}$$

with

$$M(x,y) = \Big(-(\Box + m^2) + V(x)\Big)\ \delta^{(d)}(x-y)$$

in terms of which the action reads

$$\int d^dx\, d^dy\, \varphi^*(x)\, M(x,y)\, \varphi(y) \ \ .$$

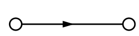
The perturbation expansion with respect to $V(x)$ yields:

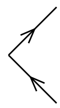
$$< 1 >_V = \sum_{m=0}^{\infty} \frac{i^m}{m!} \int dx_1 \cdots dx_n\, V(x_1) \cdots V(x_m)\ < \varphi^*\varphi(x_1) \cdots \varphi^*\varphi(x_m) >_0$$

The correlation functions $< \cdots >_0$ are expectation values with respect to the free field, with kernel

$$M_0(x,y) = -(\Box + m^2)\, \delta^{(d)}(x-y) \ \ .$$

The elements of the Feynman rules of our potential model are:

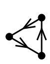
propagators :  $i\Delta_F(x-y)$

vertices : V  $\int d^dx\, V(x)$.

The connected graphs are generated by the logarithm of the normalized vacuum functional:

$$G(V) = \ln < 1 >_V - \ln < 1 >_0$$

where

$G(V) =$ v  + v  v + v  V V $+ \cdots$

The direct perturbative evaluation proceeds as follows: Start at x_1. We have $(m-1)$ possibilities and a representative term is $i\Delta_F(x_1 - x_2)$. Continuing at x_2 we have $(m-2)$ possibilities and a representative is $i\Delta_F(x_2 - x_3)$ and so on. This yields

$$(m-1)!\, i^m\, \Delta_F(x_1 - x_2) \cdots \Delta_F(x_m - x_1)$$

and hence

$$G(V) \ = \ \sum_{m=1}^{\infty} \frac{(-1)^m}{m} \int dx_1 \cdots dx_m\, V(x_1) \cdots V(x_m)\, \Delta_F(x_1 - x_2) \cdots \Delta_F(x_m - x_1)$$

The same result we may obtain from the non–perturbative result for $(\det M(x,y))^{-1}$ by expanding $\det M(x,y)$ in powers of $V(x)$. To this end we write

$$\begin{aligned} M(x,y) &= \int dx' M_0(x,y) \, \left(\delta^{(d)}(x'-y) + \Delta_F(x'-y)\, V(y)\right) \\ &= M_0 * (I+L) \end{aligned}$$

as a convolution integral. This is an identity, since

$$\int dx' \left(-(\Box_x + m^2)\right) \, \delta^{(d)}(x-x')\, \delta^{(d)}(x'-y) = -(\Box_x + m^2)\, \delta^{(d)}(x-y)$$

and

$$\begin{aligned} \int dx' \left(-(\Box_x + m^2)\right) \, \delta^{(d)}(x-x')\, \Delta_F(x'-y)\, V(y) &= -(\Box_x + m^2)\, \Delta_F(x-y)\, V(y) \\ &= V(x)\, \delta^{(d)}(x-y) \;\; . \end{aligned}$$

We thus have

$$(\det M(x,y))^{-1} = (\det M_0(x,y))^{-1} \cdot \left(\det(\delta^{(d)}(x-y) + \Delta_F(x-y)\, V(y))\right)^{-1}$$

where for the second factor we can apply the formula

$$\det(I+L) = e^{\mathrm{Tr}\ \ln((I+L)}$$

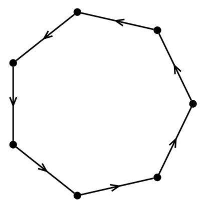
with

$$\mathrm{Tr}\ \ln((I+L) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \mathrm{Tr}\, L^m \;\; .$$

For our specific form of L we obtain

$$\mathrm{Tr}\, L^m = \int dx_1 \cdots dx_m\, V(x_1) \cdots V(x_m)\, \Delta_F(x_1 - x_2) \cdots \Delta_F(x_m - x_1)$$

which represents a loop with m interaction points:

$\mathrm{Tr}\, L^m =$ 

In matrix notation this reads:

$$\mathrm{Tr}\, L^m = L_{i_1 i_2} L_{i_2 i_3} \cdots L_{i_m i_1}$$

and we observe that $\mathrm{Tr}\ \ln((I+L)$ is precisely given by the sum of all connected one–loop graphs. Thus

$$\begin{aligned} G(V) &= \ln <1>_V - \ln <1>_0 \\ &= -\ln(\det M(x,y)) + \ln(\det M_0(x,y)) \\ &= -\mathrm{Tr}\ \ln((I+L) = \sum_{m=1}^{\infty} \frac{(-1)^m}{m} \mathrm{Tr}\, L^m \\ &= \sum_{m=1}^{\infty} \frac{(-1)^m}{m} \int dx_1 \cdots dx_m\, V(x_1) m \cdots V(x_m)\, \Delta_F(x_1 - x_2) \cdots \Delta_F(x_m - x_1) \end{aligned}$$

which is identical with the result obtained above.

We now are prepared to consider the fermions:

Determinants versus diagrams: fermions

We now consider the vacuum to vacuum transition amplitude in the path integral representation for a free Dirac field in interaction with an external potential $V(x)$:

$$< 1 >_{V,\psi} = < 0|T e^{i\int d^dx\, \bar{\psi}(x)\, V(x)\, \psi(x)}\, |0 >$$

which can be easily calculated as a perturbation expansion in V. We obtain analogue expressions as in the case of bosons with the replacement $\Delta_f(x) \to S_F(x)$ and each loop has a factor (-1) due to Fermi statistics. Hence we have

$$\begin{aligned} G(V)_\psi &\doteq \ln < 1 >_V - \ln < 1 >_0 \\ &= -\sum_{m=1}^{\infty} \frac{(-1)^m}{m} \int dx_1 \cdots dx_m\, V(x_1) \cdots V(x_m)\, S_F(x_1 - x_2) \cdots S_F(x_m - x_1) \end{aligned}$$

and since all connected graphs are one-loop diagrams the Fermi statistics just yields an overall sign.

Corresponding matrices read

$$\begin{aligned} M_\psi(x, y) &= (-(i\gamma^\mu \partial_m u - m) + V(x))\ \delta^{(d)}(x - y) \\ M_{0,\psi}(x, y) &= -(i\gamma^\mu \partial_m u - m)\, \delta^{(d)}(x - y)\ . \end{aligned}$$

and

$$I + L_\psi = \delta^{(d)}(x - y) + S_F(x - y)\, V(x)$$

With the change in sign we have

$$\begin{aligned} G(V)_\psi &= \ln(\det M_\psi(x, y)) - \ln(\det M_{0,\psi}(x, y)) \\ &= \mathrm{Tr}\ \ln((I + L_\psi) = -\sum_{m=1}^{\infty} \frac{(-1)^m}{m} \mathrm{Tr}\, L_\psi^m \end{aligned}$$

We thus can conclude that a functional integral representation for a fermionic system, if it exists, must read

$$\begin{aligned} < 1 >_{V,\psi} &= \int_{-\infty}^{+\infty} \prod_x \frac{d\psi_x}{\sqrt{2\pi i}} \frac{d\bar{\psi}_x}{\sqrt{2\pi i}}\ e^{i\int d^dx\, \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m + V(x))\psi(x)} \\ &= (\det M_\psi(x, y)) \\ &= (\det M_{0,\psi}(x, y)) \cdot \exp \mathrm{Tr}\,(I + S_F V) \end{aligned}$$

7.2.3 Path integral for fermions

For Fermions, being described by spinors, the Wick rotation to Euclidean space is not completely trivial. We therefore prefer to stay in Minkowski space. Hence our consideration will directly compare to the integral representations in terms of Fresnel integrals Eqs. (7.18,7.19).

The above consideration lead us to the conclusion that integration over the anticommuting fields (Grassmann variables), in case of the Dirac field, must have the property

$$\int_{-\infty}^{+\infty} \frac{d\psi}{\sqrt{2\pi i}} \frac{d\bar{\psi}}{\sqrt{2\pi i}}\ e^{i\bar{\psi} A \psi} = A$$

which compares to the Fresnel integral

$$\int_{-\infty}^{+\infty} \frac{d\varphi}{\sqrt{2\pi i}} \frac{d\varphi^*}{\sqrt{2\pi i}} \, e^{i\varphi^* A\varphi} = A^{-1}$$

obtained for a complex scalar field.

Since ψ and $\bar{\psi}$ are anticommuting we have $(\bar{\psi}A\psi)^n = 0 \;\; ; \;\; n > 1$! and hence

$$e^{i\bar{\psi}A\psi} = 1 + i\bar{\psi}A\psi \;\; .$$

In order to obtain the required result we must define the “integration” as follows:

Integration over Grassmann variables

$$\begin{gathered} \int \tfrac{d\psi}{\sqrt{2\pi i}} \tfrac{d\bar{\psi}}{\sqrt{2\pi i}} \, 1 = 0 \;\; , \;\; \int \tfrac{d\psi}{\sqrt{2\pi i}} \tfrac{d\bar{\psi}}{\sqrt{2\pi i}} \, i\bar{\psi}A\psi = A \\ \int \tfrac{d\psi}{\sqrt{2\pi i}} \tfrac{d\bar{\psi}}{\sqrt{2\pi i}} \, (\bar{\psi}A\psi)^n = 0 \;\; ; \;\; n > 1 \;\; . \end{gathered}$$

Since normalization factors are uninteresting, they are normalized away in any case, we may define “integration over Grassmann variables” (Berezin 1966), which we will call “Berezin–integration” in the sequel, as follows:

Let ψ_i and $\bar{\psi}_i$ be the generating elements of a Grassmann algebra. Then integration over this algebra is defined algebraically by:

$$\begin{array}{rr} \int d\psi_i \psi_k = \delta_{ik} \; ; & \int d\bar{\psi}_i \bar{\psi}_k = \delta_{ik} \\ \int d\psi_i \bar{\psi}_k = 0 \; ; & \int d\bar{\psi}_i \psi_k = 0 \\ \int d\psi_i \; 1 = 0 \; ; & \int d\bar{\psi}_i \; 1 = 0 \end{array}$$

and we notice that these rules are algebraically isomorphic to differentiation of Grassmann variables. Thus for Grassmann variables we have

$$\boxed{\text{integration} \equiv \text{differentiation}}$$

The elements ψ_i, $\bar{\psi}_i$, $\frac{\partial}{\partial \psi_i}$, $\frac{\partial}{\partial \bar{\psi}_i}$, $d\psi_i$ and $d\bar{\psi}_i$ are all anticommuting and form an extended algebra.

Which properties does the Grassmann variable integration have in common with normal integration? Not too much! It is in no way related to Riemann integrals or measure theory. Finite integration boundaries are not defined, one has always to understand improper integrals $\int_{-\infty}^{+\infty} \cdots$. However it is

1. linear

2. the integral of a derivative vanishes

$$\int d\psi_i \frac{\partial}{\partial \psi_i} \, F(\psi) \equiv 0$$

3. it allows us to perform partial integration provided integration goes over the whole algebra

$$\int \prod_k d\psi_k \; F_1(\psi) \, \left(\frac{\partial}{\partial \psi_i} \, F_2(\psi) \right) = \mp \int \prod_k d\psi_k \, \left(\frac{\partial}{\partial \psi_i} \, F_1(\psi) \right) \, F_2(\psi)$$

Otherwise the integration rules are identical to differentiation rules. For example: for one variable $F(\psi) = F_0 + F_1\psi$ is the general element with F_0 and F_1 complex numbers. We have

$$\frac{d}{d\psi}F(\psi) = F_1 = \int d\psi F(\psi) \ .$$

Similarly, under a transformation of variables: $\psi' = a\psi$ we obtain

$$\frac{d}{d\psi}\cdots = a\frac{d}{d(a\psi)}\cdots \quad \text{and} \quad \int d\psi \cdots = a\int d(a\psi)\cdots \quad ,$$

which means that the integral over Grassmann variables transforms with the **inverse** of the Jacobian!

The multidimensional case reads:

$$\begin{aligned}\int \prod_x \frac{d\psi_x}{\sqrt{2\pi i}}\,\frac{d\bar\psi_x}{\sqrt{2\pi i}}\; e^{i\sum_x \bar\psi_x A_x \psi_x} &= \prod_x \int \frac{d\psi_x}{\sqrt{2\pi i}}\,\frac{d\bar\psi_x}{\sqrt{2\pi i}}\; e^{i\bar\psi_x A_x \psi_x} \\ = \prod_x \int \frac{d\psi_x}{\sqrt{2\pi i}}\,\frac{d\bar\psi_x}{\sqrt{2\pi i}}\; 1 + (i\bar\psi_x A_x\psi_x) &= \prod_x A_x\end{aligned}$$

in the diagonal case and

$$\Big\| \int \prod_x \tfrac{d\psi_x}{\sqrt{2\pi i}}\,\tfrac{d\bar\psi_x}{\sqrt{2\pi i}}\; e^{i\sum_{x,y}\bar\psi_x A_{xy}\psi_y} = \det A \tag{7.58}$$

where A is hermitian and positive definite. With this property A may be transformed to diagonal form by a unitary transformation. This representation is the direct analogue of Eq. (7.19) for charged bosons, with the crucial difference that the determinant is replaced by its inverse value.

So far we have not considered the external sources. For a "conjugate" pair of variables we have

$$\int \frac{d\psi}{\sqrt{2\pi i}}\,\frac{d\bar\psi}{\sqrt{2\pi i}}\; e^{i\left(\bar\psi A\psi + \bar\zeta\psi + \bar\psi\zeta\right)} = A\, e^{-i\bar\zeta A^{-1}\zeta} \quad .$$

A shift of variables (quadratic completion)

$$\psi \to \psi + A^{-1}\zeta \ , \quad \bar\psi \to \bar\psi + \bar\zeta A^{-1}$$

yields

$$\bar\psi A\psi + \bar\zeta\psi + \bar\psi\zeta = (\bar\psi + \bar\zeta A^{-1})A(\psi + A^{-1}\zeta) - \bar\zeta A^{-1}\zeta$$

and with

$$\int d\psi d\bar\psi \cdots = \int d\psi' d\bar\psi' \cdots$$

the result follows.

In the multidimensional case we obtain

$$\Big\| \int \prod_x \tfrac{d\psi_x}{\sqrt{2\pi i}}\,\tfrac{d\bar\psi_x}{\sqrt{2\pi i}}\; e^{i\left(\sum_{x,y}\bar\psi_x A_{xy}\psi_y + \sum_x \bar\zeta_x\psi_x + \bar\psi_x\zeta_x\right)} = \det A\; e^{-i\sum_{x,y}\bar\zeta_x A^{-1}_{xy}\zeta_y} \quad . \tag{7.59}$$

Note that since

$$\int \frac{d\psi}{\sqrt{2\pi i}}\,\frac{d\bar\psi}{\sqrt{2\pi i}}\;\cdots \quad \text{transforms as} \quad \frac{\partial^2}{\partial\psi_x\partial\bar\psi_x}.$$

a non-unitary transformation $(\bar{\psi}, \psi) = A(\bar{\psi}', \psi')$ transforms with the inverse Jacobian!

$$F(\bar{\psi}, \psi) = F(A\bar{\psi}', A\psi') = F'(\bar{\psi}', \psi')$$

$$\frac{\partial^2}{\partial\psi\partial\bar{\psi}} = \det\left(\frac{\partial(\bar{\psi}', \psi')}{\partial(\bar{\psi}, \psi)}\right) \frac{\partial^2}{\partial\psi'\partial\bar{\psi}'} = \left(\det A^{-1}\right) \frac{\partial^2}{\partial\psi'\partial\bar{\psi}'}$$

$$\int d\psi d\bar{\psi} F(\bar{\psi}, \psi) = \left(\det A^{-1}\right) \int d\psi' d\bar{\psi}' F'(\bar{\psi}', \psi')$$

Result:

The path integral representation for free fermions reads

$$\begin{aligned} Z_0\left\{\bar{\zeta}\ ,\zeta\right\} &= \int\limits_{\otimes} \prod_x d\bar{\psi}_x d\psi_x e^{i\int d^dx\,\left(\mathcal{L}_0(\bar{\psi},\psi)+\bar{\zeta}\psi+\bar{\psi}\zeta\right)} \\ &= \exp -i(\bar{\zeta}, S_F\,\zeta) \end{aligned} \tag{7.60}$$

where integration is understood as Berezin–integration. For an interacting theory by virtue of the properties of functional differentiation we obtain

$$\begin{aligned} Z\left\{\bar{\zeta},\zeta\right\} &= <0|Te^{i\int d^dx\,\left(\bar{\zeta}\psi+\bar{\psi}\zeta\right)}|0> \\ &= {}_{in}<0|Te^{i\int d^dx\,\mathcal{L}^{(in)}_{\rm int}(\bar{\psi},\psi)}\;e^{i\int d^dx\,\left(\bar{\zeta}\psi_{in}+\bar{\psi}_{in}\zeta\right)}|0>_{in\otimes} \\ &= e^{i\int d^dx\,\mathcal{L}_{\rm int}(-i\frac{\delta}{\delta\zeta},-i\frac{\delta}{\delta\bar{\zeta}})}\;Z_0\left\{\bar{\zeta}\ ,\zeta\right\} \\ &= \int\limits_{\otimes} \prod_x d\bar{\psi}_x d\psi_x e^{i\int d^dx\,\left(\mathcal{L}_{tot}(\bar{\psi},\psi)+\bar{\zeta}\psi+\bar{\psi}\zeta\right)} \end{aligned} \tag{7.61}$$

as our final result.

Comment on the Berezin–integration:

❶ The terminology and the formal way of writing the "Berezin–integration" only serves a **formally analogue** representation of fermions and bosons. Properties, which we usually associate with the notion of integration or in particular with a "definite integral" do not carry over to Berezin–integration. This becomes immediately clear as integration over Grassmann variables in defined algebraically for discrete algebras.

❷ To consider Berezin–integration actually to mean differentiation comes closer to standard associations.

✘ The key property needed is not "integration", but the representation of a determinant as an exponential of a trace:

$$\det(I + L) = \exp(\mathrm{Tr}\ \ln(I + L))$$

which opens the way to a perturbative calculation of $\det(I + L)$ provided $||L|| < ||I||$.

✙ Nevertheless, the representation of the generating functional for time ordered Green functions for Fermi fields as a path integral does not lack *formal elegance*. The main reason are the equal properties for fermions and bosons of the path integral under unitary or orthogonal transformations and under shifts (translations) of variables, where in any case the functional determinant is unity!.

Exercises: Subsection 7.2

① Calculate the integral

$$\int \prod_x \frac{d\psi_x}{\sqrt{2\pi i}}\, e^{i \sum_{x,y} \psi_x A_{xy} \psi_y}$$

where ψ_x are the generators of a real Grassmann algebra. Hint: A_{xy} may be assumed to be antisymmetric. If A_{xy} is hermitian antisymmetric, then $A = iA_R$ with A_R real antisymmetric. Furthermore, by an orthogonal transformation S we may transform A_R to a standard matrix $\mathbf{C}$, where $\mathbf{C}$ has 2×2 matrices

$$C = i\sigma_2 = \left(\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right)$$

along the diagonal and zero elsewhere. Thus we may assume $A = iS^T\mathbf{C}S$ where $S^TS = 1$ and det$\mathbf{C}$= -1. Remark: $\mathbf{C}$ is the standard form of the metric of a simplectic space (see Sec. B).

② a). Show that

$$\begin{aligned} & e^{i(\bar{\zeta}\psi+\bar{\psi}\zeta)} = e^{i\bar{\zeta}\psi}\, e^{i\bar{\psi}\zeta} \\ = \;& 1 + i\bar{\zeta}\psi + i\bar{\psi}\zeta - i^2\bar{\psi}\psi\bar{\zeta}\zeta \end{aligned}$$

b). Check the validity of the result obtained by performing a shift variable, with the one obtained by a direct calculation, i.e., by expanding the exponential and integration.

c). Verify the rule integration≡differentiation

$$\int d\bar{\psi} d\psi\; e^{i(\bar{\psi}A\psi+\bar{\zeta}\psi+\bar{\psi}\zeta)} \propto \frac{\partial^2}{\partial\psi\partial\bar{\psi}}\; e^{i(\bar{\psi}A\psi+\bar{\zeta}\psi+\bar{\psi}\zeta)}$$

③ Prove that the value of the functional integral does not depend on the splitting of the Lagrangian into a free and an interacting part. Hint: compare one splitting $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\rm int}$ with another one $\mathcal{L} = \mathcal{L}'_0 + \mathcal{L}'_{\rm int}$ with $\mathcal{L}'_{\rm int} = \mathcal{L}_{\rm int} + \mathcal{L}_0 - \mathcal{L}'_0$.

④ Derive the equation of motion for ψ by utilizing the formal translational invariance of the Lebesgue measure.

7.3 Path integral for non-Abelian gauge fields

The proper quantization of non-Abelian gauge theories can be understood best in the path integral approach, as discussed earlier.

We consider a non–Abelian gauge theory with matter fields $\psi(x)$ transforming under a representation of the group $SU(n)$ and the Yang–Mills fields $V_{\mu i}$, the invariant Lagrangian reads

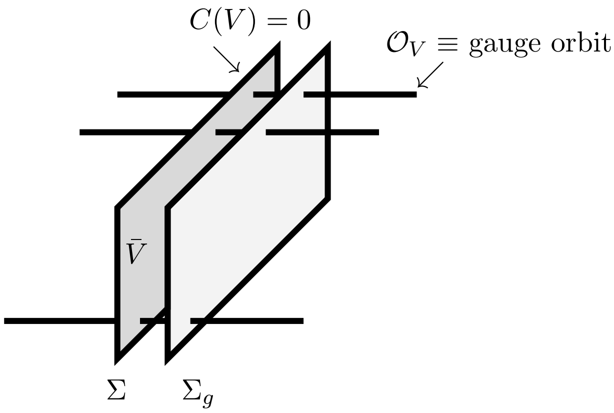
$$\mathcal{L}_{\rm inv} = -\frac{1}{4}\sum_i G_{i\mu\nu}G_i^{\mu\nu} + \bar{\psi}\;(i\gamma_\mu D_\mu - m)\;\psi \tag{7.62}$$

In this section we need consider the pure Yang–Mills part only. As in the Abelian case discussed in Sec. 4, the equations of motion which derive from $\mathcal{L}_{\text{inv}}$, due to gauge degeneracy, do not determine the non–Abelian gauge potentials $V_{\mu i}$ $(i = 1, \ldots, r)$ in a unique manner. This is obvious, since a gauge transformation changes $V_{\mu i}$, while $\mathcal{L}_{\text{inv}}$ remains the same. The problem is, as we already know, that the $V_{\mu i}$'s include unphysical degrees of freedom. Massless spin 1 bosons have two physical degrees of freedom, while $V_{\mu i}$ has four independent components. As a consequence the formal "path integral"

$$\text{``} \int \mathcal{D}V_{\mu i} e^{i \int d^d x \, \mathcal{L}_{\text{inv}}} = \infty \text{ ''}$$

is infinite, because we try to integrate form $-\infty$ to $+\infty$ over variables on which the integrand does <u>not</u> depend. Apparently we attempt to integrate along directions in the space of the gauge potentials, which are physically equivalent. We thus have to distinguish, among the variables $V_{\mu i}$, between a subset of variables $\bar{V}_{\mu i}$ which cannot be transformed into each other by a gauge transformation and a complementary set of variables which correspond to the pure gauge degrees of freedom. For this second set of variables we may choose the parameters ω_i which parametrize the gauge group.

The physical variables $\bar{V}_{\mu i}$ may be chosen to lie in a surface Σ the complementary space is spanned by the **gauge orbits** of physically equivalent fields. This is illustrated in the following Figure:



A gauge orbit is formally defined by

$$\mathcal{O}_V = \left\{ V'_g \mid V'_g \sim V \; ; \;\; g \in G \right\}$$

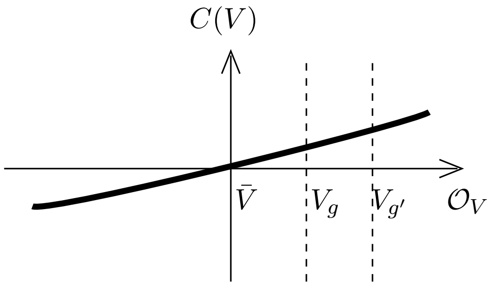
and we have

$$\mathcal{L}_{\text{YM}} = \text{ constant on } \mathcal{O}_V$$

and Σ is the hyper-surface $\perp$ to the gauge orbits. Σ is determined by the gauge fixing condition

$$\Sigma \; : \;\; C_i(V) = 0 \;\; , \quad i = 1, \ldots, r$$

In order that the gauge function $C_i(V)$ fixes the gauge uniquely we must require it to be a strictly monotonic function along the gauge orbits as shown in the Figure:



This suggests to define the path integral by integration over the hyper-surface Σ only:

$$\int \prod_{x,i} dV_{\mu i}\; e^{i\int d^dx\, \mathcal{L}_{\rm inv}} \prod_{x,i} \delta\left(C_i(V)(x)\right)$$

or equivalently by integration with a Gaussian damping factor along the gauge orbit directions

$$I(C) = \int \prod_{x,i} dV_{\mu i}\; e^{i\int d^dx\, \mathcal{L}_{\rm inv}}\; e^{-\frac{i}{2\xi}\int d^dx\, \sum_i C_i(x)^2}$$

The two forms are equivalent in the sense that

$$\begin{aligned} e^{-\frac{i}{2\xi}\int d^dx\, \sum_i C_i(x)^2} &= e^{-\frac{1}{2\xi}\int d^d\underline{x}\, \sum_i C_i(\underline{x})^2} = \prod_{\underline{x},i} e^{-\frac{1}{2\xi} C_i(\underline{x})^2} \\ \delta\left(C_i(V)(\underline{x})\right) &= \lim_{\xi\to 0} \frac{1}{\sqrt{2\pi\xi}} e^{-\frac{1}{2\xi} C_i(\underline{x})^2} \end{aligned}$$

such that for gauge invariant and hence ξ–independent quantities the equivalence is obvious. Since only the normalized measure is of interest the extra factors in front of the exponential is taken care of by properly normalizing the path integral. Note that

$$\int_{-\infty}^{+\infty} dC_i(V)(\underline{x})\; e^{-\frac{1}{2\xi} C_i(\underline{x})^2} = \sqrt{2\pi\xi}\ .$$

The path integral proposed cures the problem of the infinities from integration along the gauge orbits, however, in the non–Abelian case the result depends on the gauge fixing function in an essential way such that gauge invariance of observables is definitely lost. This is not acceptable and the above result cannot be the answer we where looking for.

In order to get an idea about how the "path integral" can be modified such that gauge invariance gets restored in an appropriate way we write symbolically

$$I(C) = \int_\Sigma \mathcal{D}\bar{V}_{\mu i}\; e^{i\int d^dx\, \mathcal{L}_{\rm inv}} \int_{\mathcal{O}_V} \mathcal{D}\omega_i\; e^{-\frac{i}{2\xi}\int d^dx\, \sum_i C_i(x)^2}\ .$$

This makes clear that we must arrange the second integral to be independent of the choice of C_i. This can be achieved easily if we introduce the Jacobian determinant as follows:

$$\int_{\mathcal{O}_V} \mathcal{D}\omega_i\; \det\left(\frac{\partial C}{\partial \omega}\right)\; e^{-\frac{i}{2\xi}\int d^dx\, \sum_i C_i(x)^2}$$

Then, since the determinant is the Jacobian of the transformation $\omega_i \to C_i$, we indeed obtain the integral

$$\int \mathcal{D}C_i\; e^{-\frac{i}{2\xi}\int d^dx\, \sum_i C_i(x)^2} = \left(\sqrt{2\pi\xi}\,\right)^N$$

over all possible C_i obtained from a particular one, the original C_i, by a gauge transformation, parametrized by ω_i. Here we have to resort to a finite dimensional approximation of N degrees of freedom. This integral is trivially independent of a particular choice of C_i. The correct path integral for a non-Abelian gauge theory therefore reads:

$$\int \mathcal{D}\bar{V}_{\mu i} \int \mathcal{D}\omega_i \, e^{i\int d^dx \, \left(\mathcal{L}_{\mathrm{inv}} - \frac{i}{2\xi}\sum_i C_i^2\right)} \det\left(\frac{\partial C}{\partial \omega}\right) = \int \mathcal{D}V_{\mu i} \, \det\left(\frac{\partial C}{\partial \omega}\right) \, e^{i\int d^dx \, (\mathcal{L}_{\mathrm{inv}} + \mathcal{L}_{\mathrm{GF}})}$$

where $\det\left(\frac{\partial C}{\partial \omega}\right)$ is the Faddeev–Popov determinant. The crucial trick is to write this determinant as an exponential, see Eq. (7.57), which allows us to perform a perturbative treatment. Equivalently, according to Eq. (7.58), the Faddeev–Popov determinant may be written as a path integral over Grassmann fields, the so called **Faddeev–Popov ghost fields**:

$$\det\left(\frac{\partial C}{\partial \omega}\right) = \int d\bar{\eta}_i d\eta_i \, e^{i\int d^dx \, \mathcal{L}_{\mathrm{FP}}}$$

where

$$\mathcal{L}_{\mathrm{FP}} = \bar{\eta}_i M_{ik} \eta_k \quad , \quad M_{ik} = \frac{\partial C_i}{\partial \omega_k} \quad .$$

As already mentioned, this kind of representation is absolutely crucial because it allows us to treat the Faddeev–Popov determinant in perturbation theory. The bilinear part defines a FP–ghost propagator, while the exponential containing the interaction part can be expanded into a perturbation series (see Sec. 8).

Result:

One can prove that the following ansatz for the path integral of a non-Abelian gauge theory provides the proper quantization of such a theory and yields gauge invariant observables:

- $\mathcal{L}_{\mathrm{eff}} = \mathcal{L}_{\mathrm{inv}} + \mathcal{L}_{\mathrm{GF}} + \mathcal{L}_{\mathrm{FP}}$ is the quasi invariant effective Lagrangian
- $\mathcal{L}_{\mathrm{GF}} = -\frac{1}{2\xi}\sum_i \left(\partial_\mu V_i^\mu(x)\right)^2$ is the simplest (bilinear) manifestly Lorentz invariant gauge fixing term. This so called Lorentz gauge, in spite of the fact that it does not fix the gauge uniquely (Gribov ambiguity), serves all practical purposes in perturbation theory.
- the Faddeev-Popov term which corresponds to this choice of the gauge fixing term, is constructed as follows: consider a gauge transformation of $V_{\mu i}$:

  $$V_{\mu i} \to V_{\mu i} + \partial_\mu \omega_i - g c_{ikl} V_{\mu l} \omega_k$$

  the gauge function

  $$C_i(V) = -\partial_\mu V_i^\mu$$

  changes to

  $$C_i \to C_i - \Box \omega_i + g c_{ikl} \partial^\mu \left(V_{\mu l} \omega_k\right)$$

  such that

  $$M_{ik} = \frac{\partial C_i}{\partial \omega_k} = -\Box \delta_{ik} + g c_{ikl} \partial^\mu \left(V_{\mu l}\right.$$

  which uniquely fixes the Faddeev–Popov Lagrangian $\mathcal{L}_{\mathrm{FP}} = \bar{\eta}_i M_{ik} \eta_k$ for the Lorentz gauge, generally utilized in perturbative calculations.

- The generating functional now reads (Faddeev–Popov, 't Hooft 1971)

$$\boxed{\begin{array}{l} Z\left\{J,\bar{\zeta},\zeta,\bar{\beta},\beta\right\} = \int\limits_{\otimes} \mathcal{D}V_{\mu i}\mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}\bar{\eta}_i\mathcal{D}\eta_i \times \\ e^{i\int d^dx\,\left\{\mathcal{L}_{\text{eff}}(V,\bar{\psi},\psi,\bar{\eta},\eta)+J_i^\mu V_{\mu i}+\bar{\zeta}\psi+\bar{\psi}\zeta+\bar{\beta}\eta+\bar{\eta}\beta\right\}} \end{array}} \tag{7.63}$$

where we have included the fermions according to Eq. (7.61), and added source terms for the Faddeev–Popov ghosts.

The perturbation expansion and the corresponding Feynman rules follow as usual by separation of the bilinear part $\mathcal{L}_0 = \mathcal{L}_{\text{eff}}^{\text{bilinear}}$ and the interaction part $\mathcal{L}_{\text{int}} = \mathcal{L}_{\text{eff}} - \mathcal{L}_0$. Accordingly, we have a free functional

$$\begin{aligned} Z_0\left\{J,\bar{\zeta},\zeta,\bar{\beta},\beta\right\} &= \int\limits_{\otimes} \mathcal{D}V_{\mu i}\mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}\bar{\eta}_i\mathcal{D}\eta_i\; e^{i\int d^dx\,\left\{\mathcal{L}_0(V,\bar{\psi},\psi,\bar{\eta},\eta)+J_i^\mu V_{\mu i}+\bar{\zeta}\psi+\bar{\psi}\zeta+\bar{\beta}\eta+\bar{\eta}\beta\right\}} \\ &= \exp\frac{i}{2}(J,\Delta_V\,J)\;\cdot\exp -i(\bar{\zeta},S_F\,\zeta)\;\cdot\exp -i(\bar{\beta},\Delta_{\text{FP}}\,\beta) \end{aligned} \tag{7.64}$$

and the perturbation expansion follows from

$$\begin{aligned} Z\left\{J,\bar{\zeta},\zeta,\bar{\beta},\beta\right\} = &= \;<0|Te^{i\int d^dx\,\left(JV+\bar{\zeta}\psi+\bar{\psi}\zeta+\bar{\beta}\eta+\bar{\eta}\beta\right)}\,|0> \\ &= \;_{in}<0|Te^{i\int d^dx\,\mathcal{L}_{\text{int}}^{(in)}(V,\bar{\psi},\psi,\bar{\eta},\eta)}\;e^{i\int d^dx\,\left(JV_{in}+\bar{\zeta}\psi_{in}+\bar{\psi}_{in}\zeta+\bar{\beta}\eta_{in}+\bar{\eta}_{in}\beta\right)}\,|0>_{in\otimes} \\ &= \;e^{i\int d^dx\,\mathcal{L}_{\text{int}}(-i\frac{\delta}{\delta V},-i\frac{\delta}{\delta\zeta},-i\frac{\delta}{\delta\zeta}-i\frac{\delta}{\delta\beta},-i\frac{\delta}{\delta\beta})}\;Z_0\left\{J,\bar{\zeta},\zeta,\bar{\beta},\beta\right\}\quad . \end{aligned} \tag{7.65}$$

For further details we refer to Sec. 8.

8 Quantization, perturbation expansion, Feynman rules

The starting point for the quantization of a non–Abelian gauge theory is the invariant Lagrangian density

$$\mathcal{L}_{\rm inv} = -\frac{1}{4}\sum_i G_{i\mu\nu}\, G_i^{\mu\nu} + \bar{\psi}\,(i\gamma^\mu D_\mu - m)\,\psi \ \ .$$

We shall restrict ourselves to discuss the perturbative approach. Thus phenomena related to instantons (Belavin, Polyakov, Shvarts and Tyupkin 1975) and the Gribov ambiguity (Gribov 1978) are not considered here. We split $\mathcal{L}_{\rm inv}$ into a free part $\mathcal{L}_0$ and an interaction part $\mathcal{L}_{\rm int}$ which is taken into account as a formal power series expansion in the gauge coupling g. The perturbation expansion is an expansion in terms of the free fields described by $\mathcal{L}_0$. Notice that this splitting of $\mathcal{L}_{\rm inv}$ is **not** gauge invariant, a fact which causes technical problems in controlling gauge invariance and renormalizability. At this stage we are confronted with another serious problem: The problem of quantizing massless spin 1 fields, which is familiar from QED. Since $\mathcal{L}_{\rm YM}$ is gauge invariant, the gauge potentials $V_{i\mu}$ cannot be uniquely determined from the gauge invariant field equations. Again one has to break the gauge invariance, now, by a sum of $r = n^2-1$ gauge fixing conditions

$$C_i(V) = 0\ ,\ \ i = 1, \cdots, r\ \ .$$

It is known from QED that the only relativistically invariant condition linear in the gauge potential which we can write is the Lorentz condition. Correspondingly we require

$$C_i(V) = -\partial_\mu\, V_i^\mu\,(x)\ = 0\ ,\ \ i = 1, \cdots, r\ \ .$$

It should be stressed that a covariant formulation is mandatory for calculations beyond the tree level. We are thus lead to break the gauge invariance of the Lagrangian by adding the **gauge fixing** term

$$\mathcal{L}_{\rm GF} = -\frac{1}{2\xi}\sum_i\ (\partial_\mu\, V_i^\mu\,(x))^2$$

with ξ a free gauge parameter. Together with the term $\mathcal{L}_0^V$ from $\mathcal{L}_{\rm inv}$ we obtain for the bilinear gauge field part

$$\mathcal{L}_{0,i}^{V,\xi} = -\frac{1}{4}\,(\partial_\mu\, V_{i\,\nu} - \partial_\nu\, V_{i\,\mu})^2 - \frac{1}{2\xi}\,(\partial_\mu\, V_i^\mu\,(x))^2$$

which now uniquely determines a free gauge field propagator. Unlike in QED, however, $\mathcal{L}_{\rm GF}$ breaks local gauge invariance explicitly. Scattering matrix elements are now gauge dependent (depend explicitly on ξ) and renormalizability and unitarity are lost. This can be shown be explicit calculations. In order to cure these unacceptable diseases one has to add another term to the Lagrangian which restores gauge invariance somehow. That this is possible at all was a surprise. The term which does this job has been found by Faddeev and Popov and is called **Faddeev-Popov term** (Faddeev and Popov 1967). We first remember that in a relativistically covariant quantization the field $V_{i\mu}(x)$ describes besides the two physical transversal modes also longitudinal and scalar **ghost** "particles". In non-Abelian gauge theories these ghosts carry $SU(n)$ charge and thus transform nontrivially under gauge transformations. The Faddeev-Popov trick consists in adding further charged ghost fields $\bar{\eta}_i(x)$ and $\eta_i(x)$, the so called **Faddeev-Popov ghosts**, which conspire with the other ghosts in such a way that physical matrix elements remain gauge invariant. Unitarity and renormalizability are then restored. The FP–ghosts must be

massless spin 0 fermions. For the unphysical ghosts this wrong spin–statistics assignment is no obstacle. The Faddeev-Popov term must be of the form

$$\mathcal{L}_{\rm FP} = \bar{\eta}_i(x) M_{ik} \eta_k(x)$$

where

$$\begin{aligned} M_{ik} &= \frac{\partial C_i(V)}{\partial V_{j\mu}(x)} \left(D_\mu\right)_{jk} \\ &= -\partial^\mu \left(\partial_\mu \delta_{ik} - g c_{ikj} V_{j\mu}(x)\right) \\ &= -\Box \delta_{ik} + g c_{ikj} V_{j\mu}(x)\, \partial^\mu + g c_{ikj} \left(\partial^\mu V_{j\mu}(x)\right) \quad . \end{aligned}$$

By partial integration of $S_{FP} = \int d^4x\, \mathcal{L}_{\rm FP}(x)$ we may write

$$\mathcal{L}_{\rm FP} = \partial_\mu \bar{\eta}_i \partial^\mu \eta_i - g c_{ikj} \left(\partial^\mu \bar{\eta}_i\right) V_{j\mu} \eta_k$$

which describes massless scalar fermions in interaction with the gauge fields.

The complete Lagrangian for a quantized Yang–Mills theory is

$$\mathcal{L}_{\rm eff} = \mathcal{L}_{\rm inv} + \mathcal{L}_{\rm GF} + \mathcal{L}_{\rm FP} \quad .$$

The free (bilinear) part

$$\mathcal{L}_0 = \mathcal{L}_0(V) + \mathcal{L}_0(\psi) + \mathcal{L}_0(\eta)$$

with

$$\begin{aligned} \mathcal{L}_0(V) &= \frac{1}{2} V_{i\mu} \left[\left(\Box g^{\mu\nu} - \left(1 - \frac{1}{\xi}\right) \partial^\mu \partial^\nu \right) \delta_{ik} \right] V_{k\nu} \\ \mathcal{L}_0(\psi) &= \bar{\psi}_{\alpha a} \left[\left(\left(i\gamma^\mu\right)_{\alpha\beta} \partial_\mu - m \delta_{\alpha\beta} \right) \delta_{ab} \right] \psi_{\beta b} \\ \mathcal{L}_0(\eta) &= \bar{\eta}_i \left[\left(-\Box\right) \delta_{ik} \right] \eta_k \end{aligned}$$

determines the free propagators, the differential operators in the square brackets being the inverses of the propagators. By Fourier transformation the free propagators are obtained in algebraic form (i.e. the differential operators are represented by c-numbers) in momentum space. Using the convention

$$\phi(x) = \frac{1}{(2\pi)^4} \int d^4p\, e^{-ipx}\, \tilde{\phi}(p) \; , \quad \phi = V, \psi, \eta$$

the partial derivative ∂_μ goes into $-ip_\mu$ and hence

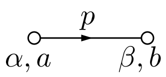
$$\begin{aligned} \left(\Box g^{\mu\nu} - \left(1 - \frac{1}{\xi}\right) \partial^\mu \partial^\nu \right) \delta_{ik} &\rightarrow -\left(g^{\mu\nu} - \left(1 - \frac{1}{\xi}\right) \frac{p^\mu p^\nu}{p^2} \right) p^2\, \delta_{ik} \\ \left(i\gamma^\mu \partial_\mu - m\right)_{\alpha\beta} \delta_{ab} &\rightarrow \left(\not{p} - m\right)_{\alpha\beta} \delta_{ab} \\ \left(-\Box\right) \delta_{ik} &\rightarrow p^2\, \delta_{ik} \quad . \end{aligned}$$

Inverting these c-number matrices we obtain the results depicted in Fig. 8.1.

The interaction part of the Lagrangian is given by

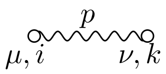
$$\begin{aligned} \mathcal{L}_{\rm int} &= g \bar{\psi} \gamma^\mu T_i \psi V_{i\mu} - \frac{1}{2} g c_{ikl} \left(\partial^\mu V_i^\nu - \partial^\nu V_i^\mu\right) V_{k\mu} V_{l\nu} \\ &- \frac{1}{4} g^2 c_{ikl} c_{ik'l'} V_k^\mu V_l^\nu V_{k'\mu} V_{l'\nu} - g c_{ikj} \left(\partial^\mu \bar{\eta}_i\right) V_{j\mu} \eta_k \end{aligned}$$

with a single coupling constant g for the four different vertices. In momentum space the vertices are represented by the following expressions: (For convenience we have symmetrized by permutations of the summation indices of the V–vertices. All momenta have been chosen incoming) *a*). Massive fermion propagator



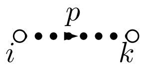
$$: \qquad \tilde{\Delta}_F^{\psi}(p)_{\alpha\beta,\,ab} = \left(\frac{1}{\not{p}-m+i\varepsilon}\right)_{\alpha\beta}\delta_{ab}$$

b). Massless gauge boson propagator



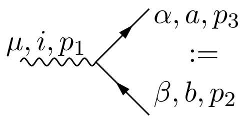
$$: \qquad \tilde{\Delta}_F^{V}(p,\xi)_{ik}^{\mu\nu} = -\left(g^{\mu\nu} - (1-\xi)\frac{p^\mu p^\nu}{p^2}\right)\frac{1}{p^2+i\varepsilon}\,\delta_{ik}$$

c). Massless FP–ghost propagator



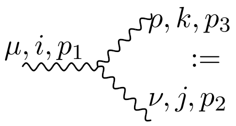
$$: \qquad \tilde{\Delta}_F^{\eta}(p)_{ik} = \frac{1}{p^2+i\varepsilon}\,\delta_{ik}$$

d). Fermion gauge field coupling



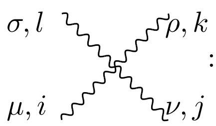
$$:= \qquad g\,(\gamma^\mu)_{\alpha\beta}\,(T_i)_{ab}$$

e). Triple gauge field coupling



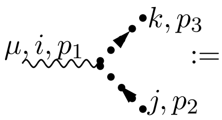
$$:= \quad -igc_{ijk}\,\{g^{\mu\nu}\,(p_2-p_1)^\rho \;+\; g^{\mu\rho}\,(p_1-p_3)^\nu \;+\; g^{\nu\rho}\,(p_3-p_2)^\mu\}$$

f). Quartic gauge field coupling



$$:= \qquad -g^2\left\{\begin{array}{l} \;\;c_{nij}c_{nkl}\,(g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\ +c_{nik}c_{njl}\,(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\ +c_{nil}c_{njk}\,(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma}) \end{array}\right.$$

g). FP–ghost gauge field coupling



$$:= \qquad -igc_{ijk}\,(p_3)^\mu$$

Fig. 8.1: Feynman rules for $\mathcal{L}_{\text{eff}}$. Momenta at vertices are chosen ingoing.

Here and in the following we do not explicitly write the $i\varepsilon$-prescription for the Feynman propagators and include it in the mass. Thus m always stands for $m - i\varepsilon$.

Given the set of well-defined propagators together with $\mathcal{L}_{\text{int}}$ we are able to calculate S-matrix elements and Green functions in perturbation theory. The S-matrix elements are obtained using

$$\begin{aligned} & < \text{out } p_1', j_1', \lambda_1', \ldots \mid p_1, j_1, \lambda_1, \ldots \text{in} > \\ = \; & < \text{in } p_1', j_1', \lambda_1', \ldots \mid S \mid p_1, j_1, \lambda_1, \ldots \text{in} > \end{aligned}$$

where the S–operator is represented in terms of the free in–fields

$$S = T e^{i \int d^4x \, \mathcal{L}_{\text{int}}\left(V^{(in)}, \psi^{(in)}, \cdots\right)(x)}$$

by expanding the exponential into a power series. Since the states and the the fields appearing in $\mathcal{L}_{\text{int}}$ may be represented in terms of the free in-state creation and annihilation operators formally we in principle know how to calculate the S-matrix elements to any order of the perturbation expansion. It is convenient however to work at an intermediate step with the closely related time ordered Green functions

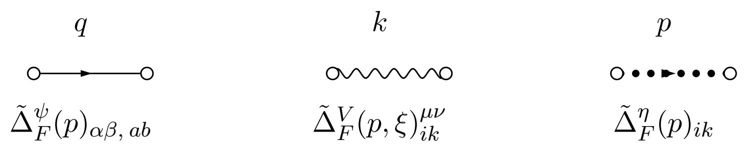
$$\begin{aligned} & < 0|T\left\{\phi_{i_1}(x_1)\, \phi_{i_2}(x_2) \cdots\right\}|0> \\ & =_{in}< 0|T\left\{\phi_{i_1}^{(in)}(x_1)\, \phi_{i_2}^{(in)}(x_2) \cdots S\right\}|0>_{in \otimes} \\ & = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int d^4y_1 \cdots d^4y_n \;{}_{in}< 0|T\left\{\phi_{i_1}^{(in)}(x_1)\, \phi_{i_2}^{(in)}(x_2) \cdots \mathcal{L}_{\text{int}}^{(in)}(y_1) \cdots \mathcal{L}_{\text{int}}^{(in)}(y_n)\right\}|0>_{in \otimes} \end{aligned}$$

and using the LSZ reduction formulae discussed in Sec. 3.3. The vacuum expectation values appearing in the last equation can be evaluated using Wick's theorems (see Sec. 3.4): Express all (free) fields in terms of the creation and annihilation operators and commute (anticommute) all annihilation operators to the right until they act onto the vacuum $|0>_{in}$ and yield zero. The only non-vanishing contributions are those where all pairs of operators have been contracted i.e. replaced by a corresponding c-number commutator (anticommutator).

The bookkeeping for the non-vanishing contributions may be organized diagrammatically using the following **Feynman rules**:

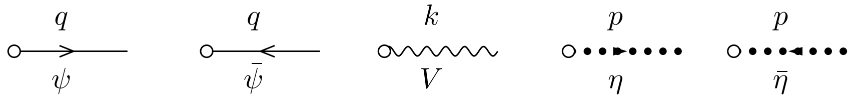
Each contribution at n-th order perturbation theory is characterized by a Feynman diagram with N external and n internal vertices (drawn as points in a plane) which are completely contracted i.e. connected by propagators:

❶ Lines:

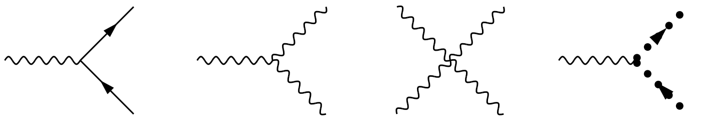


❷ Vertices:

a) external:



b) internal:



$g\,(\gamma^\mu)_{\alpha\beta}\,(T_i)_{ab}$ $\quad igV^{\alpha\beta\gamma}_{abc}$ $\quad -g^2T^{\alpha\beta\gamma\delta}_{abcd}$ $\quad -igc_{abc}\,(p_3)^\alpha$

❸ Integrals: All external momenta by convention are chosen incoming. At each vertex we have four-momentum conservation which allows to eliminate dependent momenta, the remaining **internal momenta** are **loop momenta** which must be integrated over

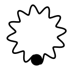
$$\frac{1}{(2\pi)^d}\int d^dl_i\cdots$$

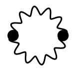
d is the space–time dimension.

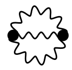
❹ Factors: Multiply the integrals over the products of propagators and coupling matrices obtained so far by the following factors:

total four-momentum conservation	:	$(2\pi)^d\delta^{(d)}\left(\sum p_{i\,ext}\right)$
each interaction vertex	:	i
each propagator	:	i
each fermion loop	:	-1
each FP–ghost loop	:	-1
combinatorial factors	:	see below

For the combinatorial factors the rules are as follows: Take out a permutation symmetry factor $\frac{1}{n!}$ from the vertex for any n identical fields in that vertex. These symmetry factors are omitted in the Feynman rules. Then multiple lines have the weight factors

 : $\frac{1}{2!}$

 : $\frac{1}{2!}$

 : $\frac{1}{3!}$

The proof is the same as to one given for the φ^4 theory at the end of Sec. 3.4.

Note that following the conventional assignment of the fermion wave functions Table 3.1, the product of spinors and γ–matrices which corresponds to an open string or to a closed loop of fermion lines are ordered from left to right, in opposite direction of the fermion line arrows, which show the flow of fermion number.

The correct Feynman rules for a non–Abelian gauge theory were derived for the first time using the Faddeev–Popov trick by G.'t Hooft in 1971.

9 Spontaneous symmetry breaking and Goldstone bosons

In many cases symmetries are not ideally realized in nature. As an example, isospin invariance of strong interactions is not exact. It is broken by small mass-splittings among the members of isospin multiplets. For the nucleon doublet one has $(m_n - m_p)/m_p \simeq 1.29/938.27$, for the pion triplet $(m_{\pi\pm} - m_{\pi 0})/m_{\pi 0} \simeq 4.59/134.97$ and so on. In addition, electromagnetic and weak interactions violate the isospin symmetry of the strong interactions. Isospin is not conserved, for example, in decays like $\pi^0 \to \gamma\gamma$ (electromagnetic) or $n \to p + e^- + \bar{\nu}_e$ (weak).

Besides such **explicit** breaking one distinguishes **spontaneous** breaking of symmetries. A symmetry is said to be spontaneously broken if the ground state exhibits less symmetry than the Lagrangian (or Hamiltonian) of the system. In fact it may happen that symmetric equations of motion have non-symmetric solutions representing states of lower energy than the symmetric ones. Though the ground state exhibits non-trivial structure, in the spontaneously broken case, we still call it vacuum.

The phenomenon of spontaneously broken symmetries is well known from condensed matter physics. Let us consider the **Heisenberg model** of a ferromagnet, as an example. The model assumes nearest neighbor spin-spin interactions of spins $\vec{S}_{\vec{r}} = (S^x_{\vec{r}}, S^y_{\vec{r}}, S^z_{\vec{r}})$ attached to the sites $\vec{r}$ of a cubic lattice. The corresponding interaction Hamiltonian

$$H = -J \sum_{<\vec{r},\vec{r}\,'>} \vec{S}_{\vec{r}} \cdot \vec{S}_{\vec{r}\,'}$$

is rotationally symmetric. Since parallel spins are favored energetically, below the critical temperature, the system exhibits "spontaneous" magnetization. We may choose it to point in the direction of the z-axis. The magnetization is the ground state expectation value of the spin variable $\vec{S}$

$$< S^z_{\vec{r}} >=< S^z_{\vec{0}} >= M \neq 0 \quad \text{or} \quad < \vec{S}_{\vec{0}} >= (0,0,M) \neq 0$$

Generally, a local variable which has a non-vanishing ground state expectation value is called **order parameter**. In a state with non-zero magnetization the symmetry of the system is reduced to rotations around the magnetization axis. The original symmetry is "spontaneously broken". Of course, the magnetization may point in any direction, which means that there are infinitely many physically equivalent ground states. The different possible ground states are related to each other by rotations. Once we have chosen a particular ground state to describe the system the symmetry is spontaneously broken.

9.1 The Goldstone theorem

In abstract terms, we may characterize the situation as follows: Let the Lagrangian of a system be symmetric with respect to a group G of transformations. If a unique vacuum $\mid 0 >$ exists, then it must be invariant under G, meaning that $\mid 0 >$ is a singlet, and the symmetry is exact. In general, however, the ground state may be degenerate such that there exists more than one state of lowest energy. The set of ground states then must transform as a multiplet of G. For a continuous symmetry group a continuous "orbit" of vacua is obtained. Each ad hoc choice of one of the ground states as the "physical vacuum" of the system breaks the symmetry spontaneously. Typically, there exists a local field which transforms non-trivially under G and which has a non-vanishing vacuum expectation value (order parameter):

$$< 0 \mid \varphi_a(x) \mid 0 >= F_a \neq 0.$$

Since the symmetry is spontaneously broken, there is no symmetry principle which forces a field (or any operator) which has the quantum numbers of the vacuum to have vanishing vacuum expectation value. Of course we require the vacuum to be symmetric under translations and Lorentz transformations. Therefore, the field $\varphi_a(x)$ must be scalar and F_a must be a constant vector in the internal symmetry space. Let us suppose that φ_a transforms according to the fundamental representation of $G = SU(n)$. By a **global** transformation

$$\varphi_a \to \varphi'_a = e^{i\sum_i Q_i\omega_i}\varphi_a e^{-i\sum_i Q_i\omega_i} = \left(e^{i\sum_i T_i\omega_i}\right)_{ab}\varphi_b$$

we may arrange things such that only the real part of the nth component has non-vanishing vacuum expectation value:

$$< 0 \mid \varphi_a(x) \mid 0 >= 0\ ,\ a = 1, \ldots, n-1\ ;\ \ < 0 \mid \varphi_n(x) \mid 0 >\neq 0\ .$$

Since we assume the Lagrangian to have **global** $SU(n)$ symmetry, we have $n^2 - 1$ conserved Hermitian currents $j_i^\mu(x)$:

$$\partial_\mu j^\mu(x) = 0\ ;\ \ j_\mu^+ = j_\mu$$

and the generators of the group are represented by the charge operators

$$Q_i = \int d^3x\, j^0(\vec{x}, t)\ ,\quad \frac{dQ_i}{dt} = 0.$$

For infinitesimal transformations the transformation law for the field, given above, takes the form of a set of commutation relations

$$[Q_i, \varphi_a] = (T_i)_{ab}\,\varphi_b.$$

The T_i's are the generators in the $n \times n$ fundamental matrix representation given in Sec. 5. If we take the vacuum expectation value we obtain

$$< 0 \mid [Q_i, \varphi_a] \mid 0 >= (T_i)_{ab} < 0 \mid \varphi_b(x) \mid 0 >= (T_i)_{an}\, v.$$

If the symmetry is **not** broken spontaneously $v = 0$ the vacuum must be invariant $Q_i \mid 0 >= 0$ for $i = 1, \ldots, n^2 - 1$. Otherwise $< 0 \mid [Q_i, \varphi_a] \mid 0 >\neq 0$ for some i. We denote the subset of Q_i's for which $< 0 \mid [Q_i, \varphi_a] \mid 0 >\neq 0$ by $\tilde{Q}_i$. Since $< 0 \mid \tilde{Q}_i\varphi_a \mid 0 > - < 0 \mid \varphi_a\tilde{Q}_i \mid 0 >\neq 0$ we must have $\tilde{Q}_i \mid 0 >=\mid 0' >_i\neq 0$. On the other hand, by the symmetry of the Lagrangian, the generators commute with the Hamiltonian

$$[Q_i\ ,\ H] = 0\ ,\ \ i = 1, \ldots, n^2 - 1.$$

This implies

$$\left[\tilde{Q}_i, H\right] \mid 0 >= \tilde{Q}_i H \mid 0 > - H\tilde{Q}_i \mid 0 >= 0$$

i.e. if $\mid 0 >$ is an eigenstate of H then $\tilde{Q}_i \mid 0 >=\mid 0' >_i$ must be an eigenstate of H with the same eigenvalue if $\tilde{Q}_i \mid 0 >\neq 0$. The vacuum must be degenerate in this case. Another important consequence follows if we consider

$$\begin{aligned} < 0 \mid \left[\tilde{Q}_i, H\right] \mid p > &= < 0 \mid \tilde{Q}_i H \mid p > - < 0 \mid H\tilde{Q}_i \mid p > \\ &= p^0 < 0 \mid \tilde{Q}_i \mid p >= p^0\ {}_i< 0' \mid p >= 0. \end{aligned}$$

for a complete set of eigenstates $\mid p>$ of P^μ.

This implies that the new vacuum $\mid 0' >_i = \tilde{Q}_i \mid 0 >$ is orthogonal to all the states $\mid p >$ belonging to the Hilbert space with vacuum $\mid 0 >$. For each vacuum we get an **inequivalent representation** of the physical states i.e. the states created from different vacua cannot be mapped by unitary transformations. Each choice of a fixed vacuum $\mid 0 >$ yields a physically equivalent description of the system, however.

Next we consider the conditions

$$< 0 \mid \left[\tilde{Q}_i, \varphi_a\right] \mid 0 >= \int d^3x \; < 0 \mid \left[\tilde{j}_i^0(x), \varphi_a\right] \mid 0 >= \tilde{c}_{ia} v \neq 0$$

which imply that

$$< 0 \mid \tilde{j}_i^\mu(x) \varphi_a(y) \mid 0 >\neq 0.$$

$\varphi_a(y)$ contains a creation operator which creates from the vacuum $\mid 0 >$ a state $\mid p >_a$. The above condition thus is equivalent to

$$< 0 \mid \tilde{j}_i^\mu(x) \mid p >_a \neq 0.$$

Using translation invariance and $P^\mu \mid p >_a = p^\mu \mid p >_a$ we obtain

$$\begin{aligned} < 0 \mid \tilde{j}_i^\mu(x) \mid p >_a &= < 0 \mid e^{-iPx} \tilde{j}_i^\mu(0) e^{iPx} \mid p >_a \\ &= < 0 \mid \tilde{j}_i^\mu(0) \mid p >_a e^{ipx} \end{aligned}$$

where $< 0 \mid \tilde{j}_i^\mu(0) \mid p >_a = f_{ia} p^\mu$ since it is a function of p only and must be a Lorentz vector. Here it is important that $\mid p >$ is a scalar state. Hence

$$< 0 \mid \tilde{j}_i^\mu(x) \mid p >_a = f_{ia} p^\mu e^{ipx} \quad \text{with} \quad f_{ia} \neq 0.$$

Since $\tilde{j}_i^\mu(x)$ is a conserved current we must have

$$\begin{aligned} \partial_\mu < 0 \mid \tilde{j}_i^\mu(x) \mid p >_a &= < 0 \mid \partial_\mu \tilde{j}^\mu(x) \mid p >_a \\ &= i f_{ia} p^2 e^{ipx} = 0 \end{aligned}$$

and hence $p^2 = 0$. This is a very interesting result saying that the state $\mid p >_a$ must be a mass zero state and $\varphi_a(x)$ must be a massless scalar field. We thus have proven the **Goldstone theorem:** Spontaneous breaking of a continuous symmetry implies the existence of zero mass bosons, so called **Goldstone bosons**. How many Goldstone bosons are there? This question may be answered easily if we inspect the conditions

$$< 0 \mid [Q_i, \varphi_a] \mid 0 >= (T_i)_{an} v; i = 1, \ldots, n^2 - 1, a = 1, \ldots, n$$

more closely. Only generators T_i with a non-zero element in the nth column yield a symmetry breaking condition. Using the representation of the T_i's given in Sec. 5 we have **one** of the $n-1$ diagonal T_i' with a non-zero element in the nth column. We denote the corresponding Q_i by $\tilde{Q}_0$ and obtain

$$< 0 \mid \left[\tilde{Q}_0, \varphi_n\right] \mid 0 >= -\frac{n-1}{\sqrt{2n(n-1)}} v \neq 0.$$

In addition, there are $2(n-1)$ off-diagonal matrices which have either a $-i$ or a 1 at one position in the last column of rows $a = 1$ to $n-1$. The corresponding Q_i's we denote by $\tilde{Q}_a^1$ and $\tilde{Q}_a^2$, $a = 1, \ldots, n-1$. Thus, we have

$$< 0 \mid \left[\tilde{Q}_a^1, \varphi_a\right] \mid 0 >= -iv \; ; \;\; < 0 \mid \left[\tilde{Q}_a^2, \varphi_a\right] \mid 0 >= v \; .$$

The remaining $n^2 - 1 - 2(n-1) - 1 = (n-1)^2 - 1$ generators have $< 0 \mid [Q_a, \varphi_a] \mid 0 >= 0$ and hence leave the vacuum invariant. These are precisely the generators of the subgroup $SU(n-1)$ which leaves the vector

$$F_a = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ v \end{pmatrix}$$

invariant:

$$UF = F \quad \Rightarrow \quad U = \left(\begin{array}{c|c} \hat{U} & 0 \\ \hline 0 & 1 \end{array}\right) \quad , \quad \hat{U} \in SU(n-1).$$

This group is the stability group (or little group) of the vacuum expectation value $< 0 \mid \varphi_a(x) \mid 0 >= F_a$ of the scalar multiplet $\varphi_a(x)$.

The $2(n-1)+1$ broken generators require $2n-1$ of the $2n$ real fields contained in the n complex fields φ_a to be massless i.e. there must be $2n-1$ Goldstone bosons.

If we introduce real fields by $\varphi_a = \varphi_a^1 + i\varphi_a^2$ ($\varphi_a^1 = Re\varphi_a, \varphi_a^2 = Im\varphi_a$) we obtain

$$< 0 \mid [Q_i, \varphi_a] \mid 0 >=< 0 \mid \left[Q_i, \varphi_a^1\right] \mid 0 > +i < 0 \mid \left[Q_i, \varphi_a^2\right] \mid 0 >$$

Now, for a real field (and Hermitian generators $Q_i^+ = Q_i$) we have

$$\begin{aligned} < 0 \mid \left[Q_i, \varphi_a^k\right] \mid 0 > &= < 0 \mid Q_i \varphi_a^k \mid 0 > - < 0 \mid \varphi_a^k Q_i \mid 0 > \\ &= < 0 \mid Q_i \varphi_a^k \mid 0 > - < 0 \mid Q_i \varphi_a^k \mid 0 >^* \\ &= i \, 2 \, Im < 0 \mid Q_i \varphi_a^k \mid 0 > \end{aligned}$$

and hence

$$< 0 \mid [Q_i \varphi_a] \mid 0 >= i \, 2 \, Im < 0 \mid Q_i \varphi_a^1 \mid 0 > -2 \, Im < 0 \mid Q_i \varphi_a^2 \mid 0 >$$

The non-vanishing expectation values are then

$$\begin{aligned} 2 \, Im < 0 \mid \tilde{Q}_0, \varphi_n^2 \mid 0 > &= \frac{n-1}{\sqrt{2n(n-1)}} v \\ 2 \, Im < 0 \mid \tilde{Q}_a^1, \varphi_a^1 \mid 0 > &= -v \\ 2 \, Im < 0 \mid \tilde{Q}_a^2, \varphi_a^2 \mid 0 > &= -v \end{aligned}$$

and the corresponding fields must be massless by the argument given before.

As a result we have: The complex multiplet $\varphi_a(x)$ in the fundamental representation of $SU(n)$ exhibits $2n$ real fields. In the spontaneously broken phase we can choose the internal space frame such that exactly **one** field, $\varphi_n^1 = Re\varphi_n$ has a non-vanishing vacuum expectation value $< 0 \mid \varphi_n^1 \mid 0 >= v > 0$. The remaining $2n-1$ must be massless Goldstone boson fields. The field φ_n' may have any mass and is not a Goldstone boson even if it would be massless by accident.

A final remark concerning the currents: If a global symmetry is exact the Noether currents

$$j_i^\mu(x) =: j_i^\mu(x) :$$

represented in terms of creation and annihilation operator exhibit terms a^+b only (possibly $b=a$) and no a^+b^+ or $a\,b$ terms. Therefore

$$j_i^\mu(x) \mid 0 >= 0.$$

In the spontaneously broken phase conditions

$$< 0 \mid \tilde{j}_i^\mu(x)\varphi_a(y) \mid 0 > \neq 0$$

must hold, such that $\tilde{j}_i^\mu(x)$ cannot annihilate the vacuum! How can we understand this? In our case $\varphi_n = \varphi_n' + v$ with $< 0 \mid \varphi_n' \mid 0 >= 0$ i.e. φ_n' is the field having "normal" properties if expanded in terms of creation and annihilation operators. The contribution of φ_a to the Noether current is

$$j_i^\mu = i : \varphi_a^* \, (T_i)_{ab} \overleftrightarrow{\partial}_\mu \, \varphi_b :$$

If we shift the field $\varphi_n \to \varphi_n + v$ we obviously get terms proportional to $v\partial_\mu\varphi_a$ which are **linear** in the fields. This obviously explains why

$$j_i^\mu(x) \mid 0 > \neq 0$$

now.

There are $2n-1$ currents which get a linear term

$$\begin{array}{lll} \tilde{Q}_0 : & \tilde{j}_0^\mu & = i : \varphi_a^* \left(\tilde{T}_0\right)_{ab} \overleftrightarrow{\partial}_\mu \, \varphi_b : \quad +2\frac{n-1}{\sqrt{2n(n-1)}} v\partial_\mu\varphi_n^2 \\ \tilde{Q}_a^1 : & \tilde{j}_a^{\mu 1} & = i : \varphi_a^* \left(\tilde{T}_a^1\right)_{ab} \overleftrightarrow{\partial}_\mu \, \varphi_b : \quad -2v\partial_\mu\varphi_a^1 \\ \tilde{Q}_a^2 : & \tilde{j}_a^{\mu 2} & = i : \varphi_a^* \left(\tilde{T}_a^2\right)_{ab} \overleftrightarrow{\partial}_\mu \, \varphi_b : \quad -2v\partial_\mu\varphi_a^2 \,. \end{array}$$

and create $2n-1$ different Goldstone bosons from the vacuum.

9.2 Models of spontaneous symmetry breaking

The first field theory models which exhibited spontaneous symmetry breaking have been invented around the year 1960 as models for describing the pion-nucleon ($\pi - N$) system. Nambu (1960) and Nambu-Jona-Lasinio (1961) proposed the so called **Nambu-Jona-Lasinio model**. At about the same time Gell-Mann-Lévy (1960) proposed the linear **σ-model** as a description of the same system. In both models pions appear as Goldstone bosons. An approximately realized Goldstone mechanism still is the only way to understand why pions are so much lighter than nucleons [26]. Later Goldstone (1961) noticed that the appearance of massless states is a general phenomenon related to spontaneous symmetry breaking. One distinguishes two different schemes of spontaneous symmetry breaking. In the **linear realization** (prototype linear σ-model) a scalar field, which develops a non-vanishing vacuum expectation value, is introduced as a fundamental field in the Lagrangian. Such models with elementary scalars are often considered to be doubtful as viable physical models because elementary scalars have not yet been found in nature. The other possibility is the **dynamical symmetry breaking** scheme (prototype Nambu-Jona-Lasinio model) which has no elementary scalars. The order parameter is a composite field like $\bar{\psi}\psi$ and the Goldstone bosons are composite fermion – antifermion states similar to the real pions (if they would be massless). The problem with such models is that our unfortunate inability to treat the relativistic bound state problem makes it hard to make unambiguous predictions.

In the following we consider models with elementary scalars for illustration of spontaneous symmetry breaking and the Goldstone mechanism.

9.2.1 A model with spontaneous breaking of a discrete symmetry

Consider a real scalar field with self-interaction described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \varphi\right)^2 - V(\varphi) = \frac{1}{2}\left(\partial_\mu \varphi\right)^2 + \frac{\mu^2}{2}\varphi^2 - \frac{\lambda}{4!}\varphi^4 \;.$$

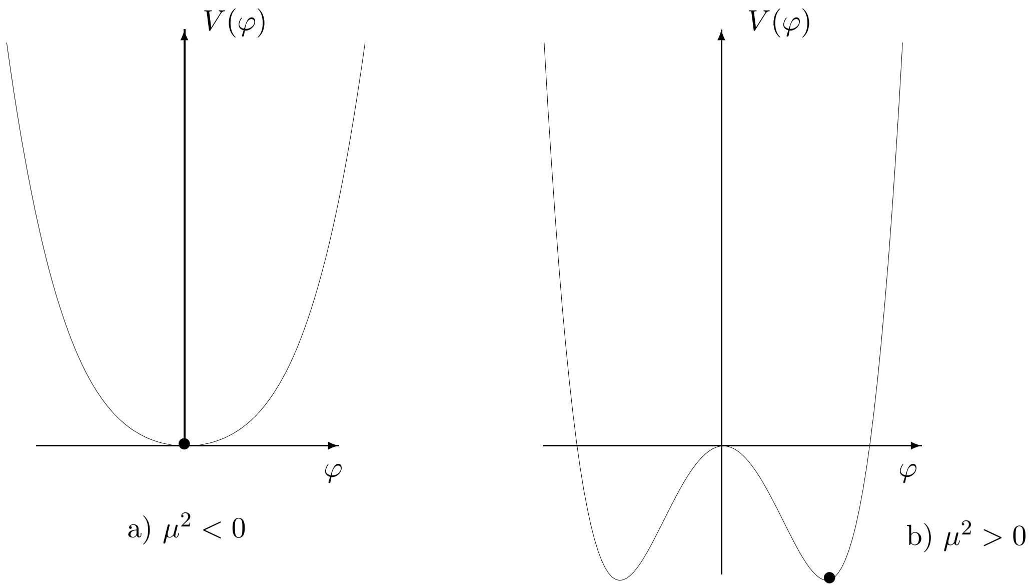


Fig. 9.1: Scalar potential a) in the symmetric and b) in the spontaneously broken phase.

Stability of the system requires $\lambda > 0$. The equation of motion for the field φ is

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi} = \frac{\partial \mathcal{L}}{\partial \varphi} \quad \text{or} \quad \Box \varphi = -\frac{\partial V}{\partial \varphi} = \mu^2 \varphi - \frac{\lambda}{3!}\varphi^3 \;.$$

The Lagrangian ha a discrete symmetry $Z_2 = \{\pm 1\} : \varphi \to -\varphi$.

[26]In modern QCD language the Goldstone boson picture of the pions is realized in the **chiral limit** of vanishing light quark masses. In this limit the pions are bona fide massless Goldstone particles

Let us treat φ as a classical field for the moment. The equilibrium solution of the system is the one for which $\varphi(x)$ is a constant determined by $\frac{\partial V}{\partial \varphi}=0$. If $\mu^2<0$, we have a unique ground state solution $\varphi(x)=\varphi_0=0$. In the quantized version we then have a unique vacuum $\mid 0>$ and $<0 \mid \varphi(x) \mid 0>=0$. $m^2=-\mu^2>0$ is the physical mass of the field φ. Now choose $\mu^2>0$. Then

$$\frac{\partial V}{\partial \varphi}=\frac{\lambda}{3!}\varphi^3-\mu^2\varphi=0$$

has besides the trivial solution $\varphi_0=0$ two non-trivial solutions (see Fig. 9.1)

$$\varphi_\pm=v_\pm=\pm\sqrt{\frac{6\mu^2}{\lambda}}\ .$$

Only these two solutions correspond to a minimum of the potential and hence to a ground state solutions. Obviously now we have two degenerate ground states.

If $\varphi(x)$ is treated as a quantum field we have to choose **one** of the ground states as the vacuum $\mid 0>$. This choice is fixed if we require, in first approximation the vacuum expectation value to be equal to one of the classical results

$$<0\mid \varphi(x)\mid 0>=v_+=v>0$$

for example.

Notice that a perturbation expansion based on the splitting $\mathcal{L}=\mathcal{L}_0+\mathcal{L}_{\rm int}$ with $\mathcal{L}_0=\frac{1}{2}\left(\partial_\mu\varphi\right)^2+\frac{\mu^2}{2}\varphi^2$ and $\mathcal{L}_{\rm int}=-\frac{\lambda}{4!}\varphi^4$ does not make sense because this would correspond to an expansion in terms of negative $m^2=-\mu^2<0$ solutions. Such Tachyons, neither satisfy the spectrum condition nor local causality. As we shall see there is a simple way to circumvent this problem. If we would perform a perturbation expansion about the fake free field solution, in every finite order of perturbation expansion, we would have a tachyon as an artifact of a nonsensical expansion. Only be infinite resummation techniques we would be able to recover the right physical answer. A field which allows for a normal particle interpretation must satisfy

$$<0\mid \varphi'(x)\mid 0>=0.$$

Such a field we simply obtain by a shift

$$\varphi=\varphi'+v$$

from the original field φ (B. W. Lee 1969). When we rewrite the Lagrangian in terms of φ' we obtain

$$\begin{aligned}\mathcal{L}(\varphi)&=\mathcal{L}(\varphi'+v)=\mathcal{L}'(\varphi')\\&=\frac{1}{2}\left(\partial_\mu\varphi'\right)^2+\frac{\mu^2}{2}\left(\varphi'+v\right)^2-\frac{\lambda}{4!}\left(\varphi'+v\right)^4\\&=\frac{1}{2}\left(\partial_\mu\varphi'\right)^2+\frac{1}{2}\left(\mu^2-\frac{\lambda v^2}{2}\right)\varphi'^2\\&\quad-\frac{\lambda}{4!}\varphi'^4-\frac{\lambda v}{3!}\varphi'^3+\left(\mu^2 v-\frac{\lambda v^3}{6}\right)\varphi'+\frac{\mu^2v^2}{2}-\frac{\lambda v^4}{4!}\ .\end{aligned}$$

The mass term of φ' now is given by

$$m^2_{\rm eff}=\frac{\lambda v^2}{2}-\mu^2=\frac{\lambda v^2}{2}-\frac{\lambda v^2}{6}=\frac{\lambda v^2}{3}>0\ \ !$$

where we have used $v = \sqrt{\frac{6\mu^2}{\lambda}}$ such that $\lambda v^2 = 6\mu^2$.

Now we can set up a perturbation expansion with

$$\begin{aligned}
\mathcal{L}_0(\varphi') &= \frac{1}{2}\left(\partial_\mu \varphi'\right)^2 - \frac{m_{\text{eff}}^2}{2}\varphi'^2 \\
\mathcal{L}_{\text{int}}\left(\varphi'\right) &= -\frac{\lambda}{4!}\varphi'^4 - \frac{\lambda v}{3!}\varphi'^3 - c_1\varphi' + c_0
\end{aligned}$$

and obtain the Feynman rules depicted in Fig. 9.2.

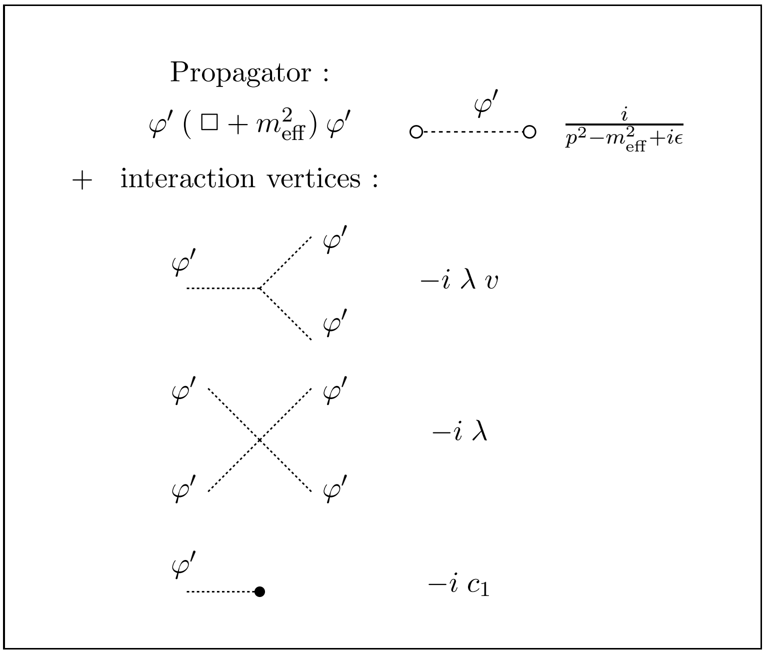


Fig. 9.2: Feynman rules for $\mathcal{L}\left(\varphi'\right)$

The vacuum energy must be adjusted to zero: $c_0 = 0$. The free parameters are λ and m_{eff}^2 as in the symmetric case. The vacuum expectation value is determined by

$$v = +\sqrt{\frac{3m_{\text{eff}}^2}{\lambda}}$$

and $c_1 = v\left(\frac{\lambda v^2}{6} - \mu^2\right) = 0$! by the condition that $< 0 \mid \varphi'(x) \mid 0 >= 0$.

The symmetry $\varphi' \to -\varphi'$ is broken now and we have learned that spontaneous symmetry breaking can **reveal a physical mass** to a particle.

9.2.2 The Goldstone model

The Goldstone model illustrates spontaneous breakdown of a continuous symmetry. We may obtain it from the previous model by replacing the real field by a complex field $\varphi \neq \varphi^*$. The Lagrangian has the form

$$\mathcal{L} = \partial_\mu \varphi^* \partial^\mu \varphi - V(\varphi\varphi^*) = \partial_\mu \varphi^* \partial^\mu \varphi + \mu^2 \varphi^* \varphi - \frac{\lambda}{3!}(\varphi^* \varphi)^2$$

and is symmetric under **global** $U(1)$ transformations. We use the equivalent representation in terms of a doublet of two real fields

$$\hat{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \; ; \; \varphi_i = \varphi_i^* \;\; (i = 1, 2)$$

which transforms under $O(2) \simeq U(1)$ rotations. The complex (charged) field is then given by

$$\varphi = \frac{1}{\sqrt{2}} (\varphi_1 + i \, \varphi_2) \; .$$

With $\hat{\varphi}^2 = \varphi_1^2 + \varphi_2^2 = 2\varphi^* \varphi$ the Lagrangian takes the form

$$\mathcal{L}(\hat{\varphi}) = \frac{1}{2} (\partial_\mu \hat{\varphi})^2 + \frac{\mu^2}{2} \hat{\varphi}^2 - \frac{\lambda}{4!} \hat{\varphi}^4 \; ,$$

which obviously is symmetric under 2-dimensional rotations

$$\hat{\varphi} \to \hat{\varphi}' = \begin{pmatrix} \varphi_1' \\ \varphi_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} .$$

The equation of motion for $\hat{\varphi}$ is

$$\Box \hat{\varphi} = \mu^2 \hat{\varphi} - \frac{\lambda}{3!} \hat{\varphi}^2 \hat{\varphi} \; .$$

Again for $\mu^2 < 0$ we find a unique ground state solution $\varphi_1 = \varphi_2 = 0$ and $m^2 = -\mu^2$ is the common mass of φ_1 and φ_2 . If $\mu^2 > 0$, the solution $\hat{\varphi}_0 = \binom{0}{0}$ is unstable. The potential has a minimum at (see Fig. 9.3)

$$\hat{\varphi}_0^2 = \frac{6\mu^2}{\lambda}$$

and we have a continuous orbit of ground state solutions related by rotations. We now choose, ad hoc, **one** particular ground state solution as the vacuum $\mid 0 >$. $O(2)$-invariance is spontaneously broken now. The particular choice of $\mid 0 >$ is fixed by specifying the vacuum expectation value of $\hat{\varphi}$:

$$< 0 \mid \hat{\varphi}(x) \mid 0 > = \begin{pmatrix} 0 \\ v \end{pmatrix} \; ; \; v > 0 \; ; \; v = \sqrt{\frac{6\mu^2}{\lambda}}$$

The physical effect of this breaking again becomes transparent if we write the Lagrangian in terms of the shifted field

$$\hat{\varphi}'(x) = \hat{\varphi}(x) - \begin{pmatrix} 0 \\ v \end{pmatrix} \;\; \text{i.e.} \;\; \begin{array}{l} \varphi_1' = \varphi_1 \\ \varphi_2' = \varphi_2 - v \; . \end{array}$$

We obtain

$$\begin{aligned} \mathcal{L}(\hat{\varphi}) &= \mathcal{L}(\hat{\varphi}' + (\begin{smallmatrix} 0 \\ v \end{smallmatrix})) = \mathcal{L}'(\hat{\varphi}') \\ &= \frac{1}{2} (\partial_\mu \hat{\varphi}')^2 - \frac{1}{2} \left(\frac{\lambda v^2}{6} - \mu^2 \right) \varphi_1'^2 - \frac{1}{2} \left(\frac{\lambda v^2}{2} - \mu^2 \right) \varphi_2'^2 \\ & \quad - \frac{\lambda}{4!} (\hat{\varphi}')^4 - \frac{\lambda v}{2 \cdot 3!} \hat{\varphi}'^2 \varphi_2' - c_1 \varphi_2' + c_0 \end{aligned}$$

with the mass terms

$$
\begin{aligned}
m^2_{1\mathrm{eff}} &= \frac{\lambda v^2}{6} - \mu^2 = \frac{\lambda v^2}{6} - \frac{\lambda v^2}{6} = 0 \\
m^2_{2\mathrm{eff}} &= \frac{\lambda v^2}{2} - \mu^2 = \frac{\lambda v^2}{2} - \frac{\lambda v^2}{6} = \frac{\lambda v^2}{3} > 0 \ !
\end{aligned}
$$

This is a remarkable result: By spontaneous symmetry breaking the particles have acquired different masses and one of the particles is massless. As we know, the appearance of a Goldstone boson is a necessary consequence of spontaneous breaking of a continuous symmetry and is not peculiar to the particular model.

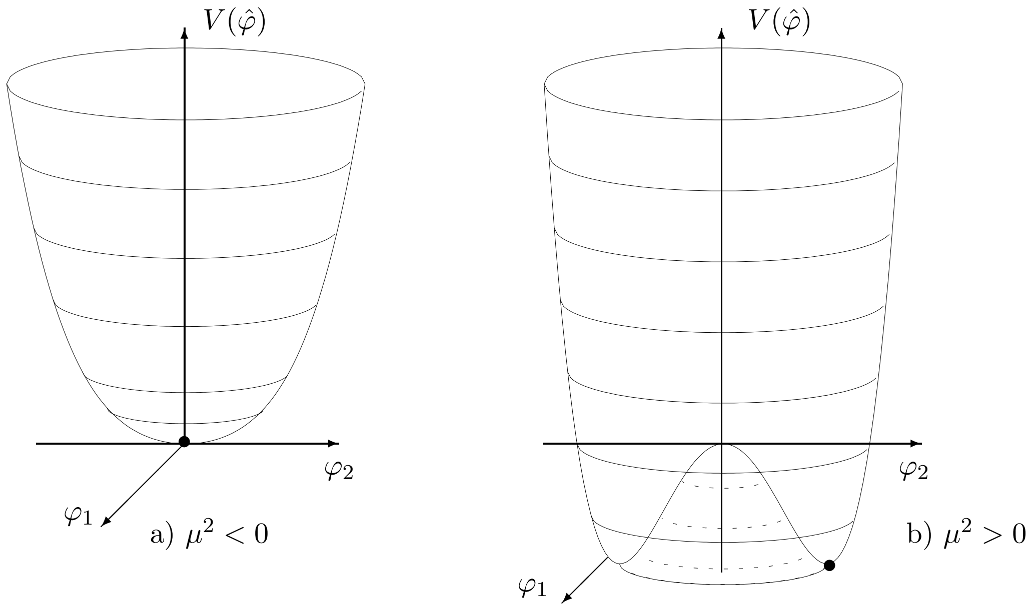


Fig. 9.3: Potential of the Goldstone model a) in the symmetric and b) in the spontaneously broken phase.

The original Lagrangian which seems to describe the two fields φ_1 and φ_2 in a completely symmetric way, has been shown to describe two neutral scalar particles of unequal mass. This is a clear manifestation of symmetry breaking.

This symmetry breaking is very different, however, from an explicit breaking of the symmetry which would result if we would add ad hoc two independent mass terms for the fields φ_1 and φ_2. This would yield a model with three independent parameters m_1, m_2 and λ. In the spontaneously broken case one of the fields **must** be massless and the model has the original number of parameter λ and m_2 in spite of the fact that new interaction vertices have been generated by the shift of the field. v is determined by $v = \sqrt{\frac{3m_2^2}{\lambda}}$ and $c_1 = v\left(\frac{\lambda v^2}{6} - \mu^2\right) = 0$.

There is an intuitive way of understanding the existence of a Goldstone boson from the fact that the vacuum is not unique. Since the vacuum is a state of zero energy and momentum different vacuum states can differ only by the presence of a Bose condensate i.e. the presence of an unspecified number of quanta of zero energy and momentum. Such quanta are possible only if there exist massless particles with the quantum numbers of the vacuum. In our example we can understand intuitively that there will be **one** such type of Goldstone bosons. The different possible values of $< 0 \mid \hat{\varphi}(x) \mid 0 >$, lie on a circle of radius v. Each point on the circle correspond to a particular choice of the vacuum state. Thus there is one degree of freedom (motion along the circle) by which different vacuum states may be connected, and hence **one** type of massless bosons.

The curvature of the potential

$$\left(\frac{\partial^2 V}{\partial\varphi_i \partial\varphi_k}\right)\bigg|_{\hat{\varphi}=\hat{\varphi}_0}$$

at the classical minimum actually represents the mass matrix of the system. If we diagonalize the matrix we find an eigenvalue zero corresponding to zero curvature of the potential in direction tangential to the ground state orbit.

A final remark may be in order here: we have shown in Subsec. 9.1 that the different degenerate vacua carry physical Hilbert spaces which are totally orthogonal (inequivalent representations). Thus the existence of the valley in the potential which seems to explain the come about of a Goldstone boson as the mode which moves along the bottom of the valley is misleading. There are no Goldstone bosons moving from one vacuum to another one. Note that all Goldstone bosons move at the speed of light and carry some momentum. The fact that this momentum may be arbitrarily small does not change the fact that there cannot be physical transitions between different degenerate vacua. Similarly, if we consider a Heisenberg ferromagnet which has a net spontaneous magnetization in a given direction (always in the thermodynamic limit, i.e., at infinite volume) the spin waves which correspond to the Goldstone excitations of course do not affect (i.e., rotate) the given ground state of the system.

Exercises: Section 9

① Discuss the symmetry breaking of $SU(2)_{\text{flavor}}$ (Isospin) and $SU(3)_{\text{flavor}}$ (Isospin and Strangeness) in the spin 1/2 baryon octet. Comment the decays $\Sigma^0 \to \Lambda\gamma$ and $\Sigma^+ \to p\gamma$ and compare them to the strong decays $\Sigma^- \to \Lambda\pi^-(?)$ and $\Sigma^- \to n\pi^-$. Use the quark model schema for the discussion.

② In the Goldstone model of spontaneous symmetry breaking two real fields $\hat{\varphi} = \binom{\varphi_1}{\varphi_2}$ interact by the Lagrangian $\mathcal{L} = \frac{1}{2}\left(\partial_\mu \hat{\varphi}\right)^2 + \frac{\mu^2}{2}\hat{\varphi}^2 - \frac{\lambda}{4!}\hat{\varphi}^4$. For $\mu^2 > 0$ let the ground state be given by $\hat{\varphi}_0 = \binom{0}{v}, v > 0$. Show that the curvature of the potential $\frac{\partial^2 V}{\partial\varphi_i\partial\varphi_k}$ at the ground state $\hat{\varphi}_0$ represents the mass matrix of the system.

③ Write down the Feynman rules for the Goldstone model (see Fig. 9.2).

10 The Higgs mechanism

or

spontaneously broken local gauge symmetries

In Sec. 6 we learned that global symmetries on the one hand and local symmetries on the other hand lead to very different physical consequences. Not surprisingly, therefore, the physics of spontaneously broken symmetries also is very different for local and global symmetries. While for global symmetries the Goldstone mechanism implies the existence of Goldstone bosons (massless scalar particles) the analogue so called Higgs mechanism of spontaneously breaking a gauge symmetry reveals that the corresponding "would be Goldstone bosons" are unphysical and can be eliminated from the theory by a local gauge transformation. At the same time the spin 1 gauge bosons acquire a mass in a way which does not conflict with the renormalizability of the theory. As we shall see, the number of gauge bosons which become massive is equal to the number of scalars which become unphysical (Higgs ghosts). Since each massive spin 1 particle has one additional degree of freedom (one longitudinal besides the two transversal ones) in comparison with the massless case the number of physical degrees of freedom remains conserved when a system undergoes a Higgs mechanism. The interesting property of the Higgs mechanism is that it provides a tool which allows to generate masses of particles in Yang-Mills theories without spoiling the renormalizability and without implying the existence of massless scalars. Such scalars have never been observed in nature and hence are not acceptable in a realistic theory of elementary particles.

Since the basic lesson about the Higgs mechanism physicists have learned from super conductivity in condensed matter physics we will start our discussion with a brief account of the Ginzburg-Landau model of a super-conductor before we consider the effect in a field theoretic context.

10.1 Superconductivity and the Meissner effect

Superconductivity (*Kammerlingh Onnes, 1911*) is always accompanied by the Meissner effect (*Meissner Ochsenfeld, 1933*) which is that electromagnetic fields must vanish inside a superconductor. This means that a mass is generated for the photon in a superconducting state. In a normal conductor, like in the vacuum, local gauge invariance requires the photon to be massless and hence the electromagnetic field to be long ranged. Below a critical temperature T_c (Curie temperature) the electron-phonon (lattice vibrations of the solid) interactions lead to an effective attractive interaction between electrons by exchange of virtual phonons (*Fröhlich, 1950*). In the presence of an attractive interaction the Fermi sea which describes the ground state of a normal metal (electron gas) is unstable against formation of bound states between electron pairs of opposite spin and equal total momentum for all pairs, no matter how weak the interaction (*Cooper, 1950*). According to Ginzburg-Landau (1950) a superconductor may be described by an order parameter $\Psi_s(\vec{r})$ which describes the macroscopic wave function of the quantum mechanical ground state. $\mid \Psi_s(\vec{r}\,) \mid^2 = \frac{n_s(\vec{r}\,)}{2}$ is the density of Cooper-pairs of electrons, i.e. $n_s(\vec{r}\,)$ is the density of superconducting electrons. At constant temperature the system is described by the free energy $F[\Psi_s]$ as a function of the order parameter. In the normal state $n_s \equiv 0$ for all $T > T_c$. Since n_s vanishes as we approach T_c from below we may expand $F[\Psi_s]$ in terms of Ψ_s and grad Ψ_s if we are not to far below T_c . Because Ψ_s is a charged field (carrying charge $e^* = 2e$ and with mass $m^* = 2m_e$) $F[\Psi_s]$ only can depend on $\Psi_s\Psi_s^*$. In addition, grad Ψ_s in the presence of a magnetic field $\vec{B}$ must enter in the form of a canonical momentum

$$\frac{\hbar}{i}\vec{\nabla} \to \frac{\hbar}{i}\vec{\nabla} - \frac{e^*}{c}\vec{A}$$

where $\vec{A}$ is the vector potential. This guarantees a gauge invariant coupling between Ψ_s and $\vec{A}$. Accordingly, a stationary superconducting state is described by the Ginzburg-Landau Ansatz

$$\begin{aligned} F[\Psi_s] = F_n(T,0) \quad &+ \int d^3r \, \frac{\vec{B}^2(\vec{r}\,)}{8\pi} \\ &+ \int d^3r \left(a \mid \Psi_s \mid^2 + \frac{b}{2} \mid \Psi_s \mid^4 + \ldots \right) \\ &+ \int d^3r \left(\frac{1}{2m^*} \left| \left(\frac{\hbar}{i}\vec{\nabla} - \frac{e^*}{c}\vec{A}(\vec{r}\,) \right) \Psi_s(\vec{r}\,) \right|^2 + \ldots \right) \end{aligned}$$

with $\vec{B} = \text{rot}\vec{A}$. The first term is the normal metal contribution at vanishing $\vec{B}$ and the second term is the magnetic field free energy. The ground state wave functions Ψ_s must be determined such that $F[\Psi_s]$ is at the minimum. This implies that Ψ_s must be a solution of the Ginzburg-Landau equations. These are

i) the Maxwell equation

$$\vec{j}_s = \frac{c}{4\pi} \text{ rot } \vec{B} \quad \text{(which implies div } \vec{j}_s \equiv 0)$$

for the superconducting current

$$\vec{j}_s = \frac{e^*\hbar}{2m^* i} \Psi_s^* \overleftrightarrow{\nabla} \Psi_s - \frac{e^{*2}}{m^* c} \mid \Psi_s \mid^2 \vec{A}(\vec{r}\,)$$

ii) the Schrödinger equation

$$\left(\frac{\vec{p}^{\,2}}{2m^*} + b \mid \Psi_s \mid^2 \right) \Psi_s = -a\Psi_s$$

with

$$\vec{p} = \frac{\hbar}{i}\vec{\nabla} - \frac{e^*}{c}\vec{A}.$$

Thus $-a$ is the energy eigenvalue and $b \mid \Psi_s \mid^2$ $(b > 0)$ is a repulsive self–interaction potential which forces Ψ_s to spread over the whole system. Ψ_s is a condensate wave function describing the condensate of Cooper-pairs. By gauge invariance

$$\begin{aligned} \vec{A} \to \vec{A}\,' &= \vec{A} + \vec{\nabla}\chi(\vec{r}\,) \\ \Psi_s \to \Psi_s' &= e^{i\frac{e^*}{\hbar c}\chi(\vec{r}\,)}\Psi_s \end{aligned}$$

of the free energy the phase φ of Ψ_s

$$\Psi_s = \mid \Psi_s \mid e^{i\varphi}$$

has no physical meaning. It may be arranged to vanish by a local gauge transformation.

We are now ready to discuss the solutions of the Ginzburg-Landau equations. For the interior of a superconductor we have $\vec{B} \equiv 0$ and $n_s =$ constant. Thus for the free energy density $f = F/V$, V the volume of the system, we obtain

$$\Delta f = f_s - f_n = a \mid \Psi_s \mid^2 + \frac{b}{2} \mid \Psi_s \mid^4$$

with

$$\left(a + b \mid \Psi_s \mid^2\right) \Psi_s = 0.$$

If $a > 0$ $\Psi_s \equiv 0$ and hence $n_s \equiv 0$ i.e. the system is in a normal state. This occurs if $T > T_c$. For $T < T_c$ we must have $a < 0$ and the ground state is non-trivial:

$$\mid \Psi_{s0} \mid^2 = \frac{n_s}{2} = -a/b > 0$$

such that[27]

$$\Delta f_{\min} = -\frac{a^2}{2b} < 0$$

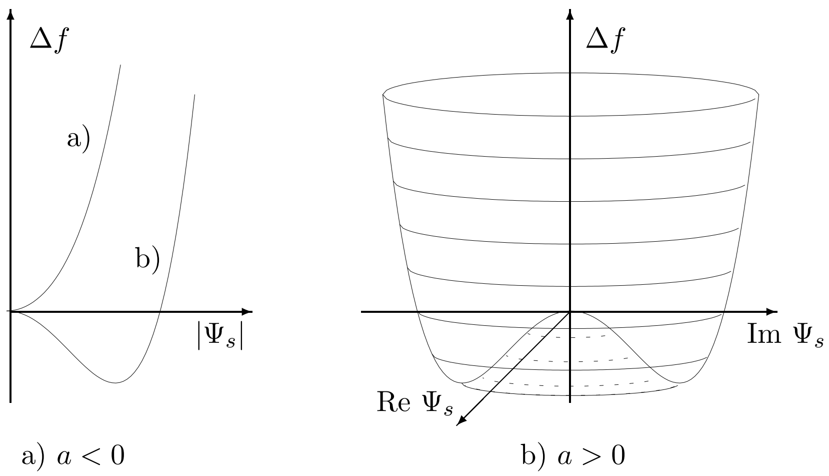


Fig. 10.1: Potential of the Ginzburg-Landau model a) in the normal and b) in the superconducting phase.

The superconducting current then reads

$$\vec{j}_s = -\frac{e^{*2}}{m^* c} \mid \Psi_{s0} \mid^2 \vec{A} = -\frac{e^2 n_s}{mc} \vec{A}$$

and hence we have

$$\text{rot } \vec{j}_s = -\frac{e^2 n_s}{mc} \vec{B}.$$

[27] If we apply an external magnetic field H then f gets increased by $\frac{H^2}{8\pi}$. At the critical field $H_c = \frac{H_c^2}{8\pi} = \frac{a^2}{2b}$ a transition to the normal state takes place.

Using the Maxwell equation rot $\vec{B} = \frac{4\pi}{c}\vec{j}_s$ we find

$$\text{rot rot } \vec{B} = \text{grad div}\vec{B} - \Delta\vec{B} = \frac{4\pi}{c} \text{ rot } \vec{j}_s = -\frac{4\pi n_s e^2}{mc^2}\vec{B}$$

and since div$\vec{B} =$ div rot $\vec{A} \equiv 0$ this reads

$$\Delta\vec{B} = \frac{1}{\lambda_L^2}\vec{B}$$

with

$$\lambda_L = \sqrt{\frac{mc^2}{4\pi n_s e^2}}$$

the London penetration depth (empirically $\lambda_L \simeq 100 - 500 \overset{\circ}{A}$). This last equation exhibits the Meissner effect of superconductivity i.e. the magnetic field gets expelled from the superconducting state. At the boundary the field $\vec{B}$ is exponentially screened (see Exercise).

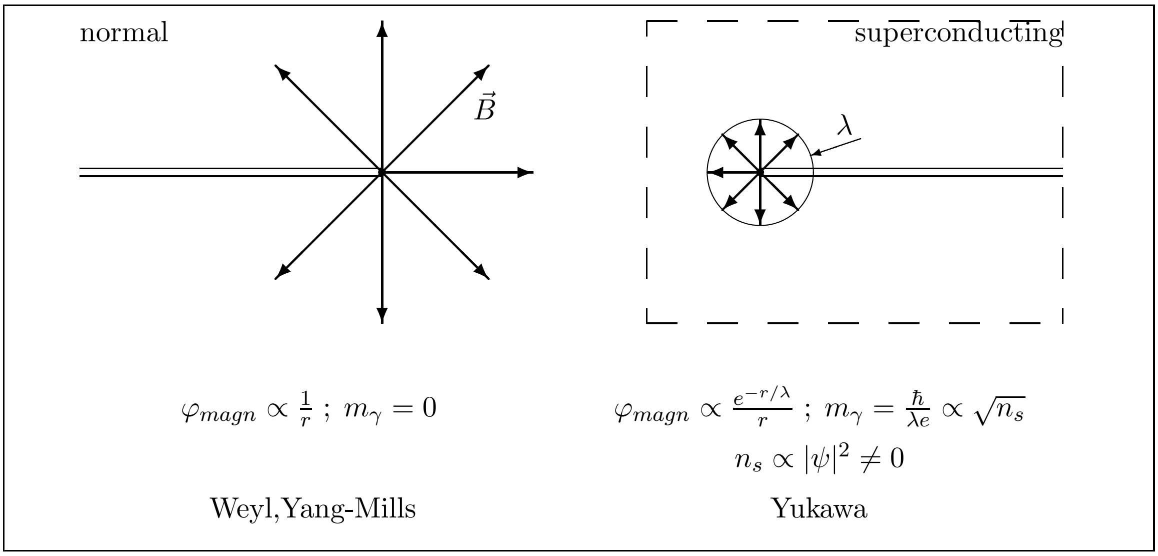


Fig. 10.2: The Meissner effect for a magnetic pole.

There is a contribution from $\vec{B}$ to the free energy in the superconducting state from the boundary layer:

$$\Delta F_s^B = \frac{1}{8\pi}\int d^3r\, \vec{B}^2 = \frac{1}{8\pi}\int d^3r\, \vec{B} \text{ rot } \vec{A}$$

Using div$(\vec{B} \times \vec{A}) = \vec{A}$ rot $\vec{B} - \vec{B}$ rot $\vec{A}$ and the fact that the integral over a divergence vanishes (by partial integration) if no long ranged fields are present we obtain

$$\Delta F_s^B = \frac{1}{8\pi}\int d^3r\, \vec{A} \text{ rot } \vec{B} = \frac{1}{2c}\int d^3r\, \vec{j}_s\vec{A} = -\frac{1}{8\pi\lambda_L^2}\int d^3r\, \vec{A}^{\,2} \quad .$$

Here we have utilized the Maxwell equation for rot $\vec{B}$ and the explicit form for $\vec{j}_s$. This result is an alternative form of the Meissner effect. The "photon field" $\vec{A}$ in the superconducting phase acquired a mass

$$m_\gamma = \frac{1}{\sqrt{4\pi}\lambda_L} = \sqrt{\frac{n_s e^2}{mc^2}}$$

So far we have not shown that the Ginzburg-Landau superconductor exhibits infinite conductivity. For completeness we briefly discuss also this basic property. Heuristically, ideal conductivity follows from free (frictionless) motion

$$m \dot{\vec{v}} = e\vec{E}$$

of the superconducting electrons. By $\vec{v}_s$ we denoted the average drift velocity and the superconducting current is

$$\vec{j}_s = e n_s \vec{v}_s \ .$$

This implies the London equation

$$\dot{\vec{j}} = \frac{n_s e^2}{m}\vec{E}$$

and replaces Ohms law

$$\vec{j}_n = \sigma \vec{E}$$

with finite conductivity σ valid for a normal metal. For a normal conductor the electric field energy is compensated by energy dissipation caused by friction (production of heat) which, using $\vec{j}_n = e n_e \vec{v}_n$, implies

$$\vec{v}_n = \frac{\sigma}{e n_e}\vec{E} = \text{ constant for } \vec{E} = \text{ constant}$$

which describes a stationary current flow.

In the Ginzburg-Landau model ideal conductivity derives from the results given before and Maxwell's equations. We assume that external charges and currents are absent and also neglect possible displacement currents proportional to $\dot{\vec{E}}$. Then, the homogeneous Maxwell equations

$$\text{div } \vec{E} = 0 \ , \ \ \text{div } \vec{B} = 0$$

infer that $\vec{E}$ and $\vec{B}$ derive from a vector potential $\vec{A}$

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} \ , \ \ \vec{B} = \text{rot}\vec{A}$$

and we assume the London gauge

$$\text{div } \vec{A} = 0$$

in order to determine $\vec{A}$ uniquely. Using our previous result

$$\vec{j}_s = -\frac{e^2 n_s}{mc}\vec{A}$$

and the inhomogeneous Maxwell equations

$$\text{rot } \vec{E} = -\frac{1}{c} \dot{\vec{B}} \ , \ \ \text{rot}\vec{B} = \frac{4\pi}{c} \vec{j}_s$$

we obtain London's equations (*London, 1935*)

$$\dot{\vec{j}}_s = \frac{n_s e^2}{m} \vec{E} \ , \ \ \text{rot}\vec{j}_s = -\frac{n_s e^2}{m_c} \vec{B} \ .$$

These indeed exhibit infinite conductivity and the Meissner effect. Notice that both $\vec{E}$ and $\vec{B}$ get exponentially screened at the phase boundary (see Exercise).

10.2 The Abelian Higgs model (gauged Goldstone model)

The Abelian Higgs model is a relativistic quantum field theory version of the Ginzburg-Landau model of superconductivity and at the same time a local gauge symmetry version of the Goldstone model discussed in Sec. 9.2.2 (*Higgs, 1964, Englert, Brout, 1964*). We consider a self–interacting charged scalar field $\varphi \neq \varphi^*$ described by the Lagrangian

$$\mathcal{L} = \partial_\mu \varphi^* \partial^\mu \varphi - V(\varphi^* \varphi) = \partial_\mu \varphi^* \partial^\mu \varphi + \mu^2 \varphi^* \varphi - \frac{\lambda}{3!} (\varphi^* \varphi)^2$$

which is symmetric under global $U(1)$-transformations

$$\varphi(x) \to \varphi'(x) = e^{ie\omega} \varphi(x) \ .$$

For the moment we assume $\mu^2 = -m^2 < 0$ such that m is the mass of the field φ . Let us now couple φ to a massless gauge field A_μ in a locally gauge invariant manner. As usual this is achieved by performing a minimal substitution

$$\partial_\mu \varphi(x) \to D_\mu \varphi(x) = (\partial_\mu - ieA_\mu(x))\, \varphi(x)$$

in the original Lagrangian and by adding a kinetic term for the gauge field. In this way we obtain the Abelian Higgs model described by the Lagrangian

$$\mathcal{L}_{\text{inv}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \varphi)^* (D^\mu \varphi) - V(\varphi^* \varphi)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the Abelian field strength tensor. By construction $\mathcal{L}_{\text{inv}}$ is invariant under local $U(1)$-transformations

$$\begin{aligned} A_\mu(x) \ &\to \ A'_\mu(x) = A_\mu(x) - \partial_\mu \omega(x) \\ \varphi(x) \ &\to \ \varphi'(x) = e^{ie\omega(x)} \varphi(x) \ \ . \end{aligned}$$

Let us now assume that $\mu^2 > 0$. Then the ground state is characterized by a non-zero value of φ . We may chose a gauge and a corresponding vacuum state such that

$$< 0 \mid \varphi(x) \mid 0 >= \frac{v}{\sqrt{2}}$$

with v real and positive. The question now is what physical particles the model describes.

On a classical level we may express the complex field in polar coordinates

$$\varphi(x) = \frac{\rho(x)}{\sqrt{2}} e^{-i\frac{\theta(x)}{v}} \; ; \; \rho(x) = \rho'(x) + v$$

where $\rho(x)$ and $\theta(x)$ are real scalar fields which take values in the ranges

$$\rho(x) \geq 0 \; , \quad 0 \leq \theta(x) < 2\pi \; .$$

The normalization by v in the exponential is needed in order to have the usual dimension for the scalar field $\theta(x)$.

In the new coordinates we have

$$\varphi^* \varphi = \frac{\rho^2}{2}$$

$$\left(D_\mu \varphi\right)^* \left(D^\mu \varphi\right) = \frac{1}{2} \left(\partial \rho\right)^2 + \frac{e^2}{2} \left(A_\mu + \frac{1}{ev} \partial_\mu \theta\right)^2 \rho^2$$

since the covariant derivative takes the form

$$D_\mu \varphi = \left(\partial_\mu - ieA_\mu\right) \varphi = \frac{\left(\partial_\mu \rho\right)}{\sqrt{2}} e^{-i\frac{\theta}{v}} - ie \left(A_\mu + \frac{1}{ev} \left(\partial_\mu \theta\right)\right) \frac{\rho}{\sqrt{2}} e^{-i\frac{\theta}{v}} \; .$$

The form of the Lagrangian now is

$$\mathcal{L}_{\rm inv} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} \left(\partial \rho'\right)^2 + \frac{e^2}{2} \left(A_\mu + \frac{1}{ev} \partial_\mu \theta\right)^2 \left(\rho' + v\right)^2 - V\left(\frac{(\rho' + v)^2}{2}\right)$$

and, remarkably, the field θ only enters like gauge transforming A_μ . Indeed, since $\mathcal{L}_{\rm inv}$ is gauge invariant, we may perform a gauge transformation

$$\begin{aligned} A_\mu \;&\rightarrow\; A_\mu - \frac{1}{ev} \partial_\mu \theta \\ \varphi \;&\rightarrow\; e^{i\frac{\theta}{v}} \varphi = \frac{\rho}{\sqrt{2}} \end{aligned}$$

which eliminates θ completely from the Lagrangian. Notice that $F_{\mu\nu}$ and ρ remain unaffected by a gauge transformation. Consequently there exists a gauge where $\mathcal{L}_{\rm inv}$ can be represented in terms of the two fields A_μ and ρ only.

$$\mathcal{L}_{\rm inv} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} \left(\partial \rho'\right)^2 + \frac{e^2 v^2}{2} A_\mu A^\mu + e^2 v A_\mu A^\mu \rho' + \frac{e^2}{2} A_\mu A^\mu \rho'^2 - V\left(\frac{(\rho' + v)^2}{2}\right) .$$

This Lagrangian describes a **massive** spin 1 boson A_μ and a massive spin 0 boson, called **Higgs boson**. The masses are

$$m_A = ev \; , \quad m_\rho = \sqrt{\frac{\lambda}{3}}\, v$$

where the latter comes from $V(\varphi^* \varphi)$. The θ-field, the "would be" Goldstone boson, has disappeared from the physical spectrum. Somehow the field θ got absorbed by the gauge field A_μ, which acquired a mass and hence the extra (third) degree of freedom required for a massive spin 1 particle. Therefore the number of physical degrees of freedom has been conserved in the process.

The non-observability of the θ mode is reminiscent of the fact that in quantum mechanics the phase of the wave function of a charged particle is not observable. In an arbitrary gauge θ is an unphysical ghost particle, called a **Higgs ghost**. One can prove that there are no non-vanishing transition matrix elements between θ-particle states and physical states.

The non-appearance of Goldstone bosons and the generation of a mass for the gauge bosons, instead, is known as the **Higgs mechanism**. The Goldstone theorem no longer holds for gauged symmetries.

So far our argumentation has been rather formal and mainly on the classical field theory level. In quantum field theory the use of the polar decomposition of φ is problematic for two reasons. Firstly, using non-polynomial fields like $e^{-i\frac{\theta}{v}}$ is in conflict with perturbative renormalizability which only allows monomials of fields with dimension 4 at most in 4 space-time dimensions. In our Abelian model this seems not to be a problem since $\mathcal{L}(A,\rho,\theta)$ looks like to be renormalizable. However, and this is the second problem, we do not know how to maintain the conditions $\rho \geq 0$ and $0 \leq \theta < 2\pi$ which have to be satisfied by the classical fields. If we just treat ρ and θ as normal real scalar fields ($-\infty < \rho, \theta < +\infty$) we are not sure that we are still talking about the original model. We therefore will discuss the model in a more precise way in the following.

We proceed by following closely our discussion of the Goldstone model. We represent φ by a **linear** transformation

$$\varphi = \frac{1}{\sqrt{2}}\,(\varphi_1 + i\varphi_2)$$

in terms of two real fields φ_1 and φ_2 . For the fields $\varphi_i \quad (i = 1,2)$ the $U(1)$ transformations correspond to $O(2)$ rotations in the (φ_1,φ_2)-plain. Accordingly an infinitesimal local $O(2)$ transformation reads

$$\begin{array}{lclcl} A_\mu & \to & A'_\mu(x) & = & A_\mu(x) - \partial_\mu\omega(x) \\ \varphi_1(x) & \to & \varphi'_1(x) & = & \varphi_1(x) + e\omega(x)\varphi_2(x) \\ \varphi_2(x) & \to & \varphi'_2(x) & = & \varphi_2(x) - e\omega(x)\varphi_1(x)\ . \end{array}$$

For $\mu^2 > 0$ we may chose the ground state (vacuum) such that [2]

$$< 0 \mid \varphi_1(x) \mid 0 >= 0\ , \quad < 0 \mid \varphi_2(x) \mid 0 >= v > 0$$

and represent the Lagrangian in terms of

$$\varphi'_1 = \varphi_1\ , \quad \varphi'_2 = \varphi_2 - v\ \ .$$

To this end, we first notice that we may write

$$(D_\mu\varphi)^*(D^\mu\varphi) = \partial_\mu\varphi^*\partial^\mu\varphi - \hat{j}^\mu A_\mu = \frac{1}{2}(\partial\varphi_1)^2 + \frac{1}{2}(\partial\varphi_2)^2 - \hat{j}^\mu A_\mu$$

[2]In order to conform with our discussion of the Goldstone model we use a convention which differs by a phase from the one used so far in this section.

with[3]

$$\hat{j}^\mu = -ie\left\{\varphi^* \overleftrightarrow{\partial^\mu} \varphi - ie\varphi^*\varphi A^\mu\right\} = e\varphi_1 \overleftrightarrow{\partial^\mu} \varphi_2 - \frac{e^2}{2}A^\mu\left(\varphi_1^2+\varphi_2^2\right)$$

Using this we easily evaluate $\mathcal{L}_{\rm inv}$ in terms of φ_1' and φ_2' .

The resulting $\mathcal{L}$ reads

$$\begin{aligned}\mathcal{L}_{inv} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{m_A^2}{2}A_\mu A^\mu + \frac{1}{2}(\partial\varphi_1')^2 + \frac{1}{2}(\partial\varphi_2')^2 - \frac{m_2^2}{2}\varphi_2'^2 \\ &\quad + evA_\mu\partial^\mu\varphi_1' + \frac{e^2}{2}A_\mu A^\mu\left(\varphi_1'^2+\varphi_2'^2\right) - eA_\mu\varphi_1' \overleftrightarrow{\partial^\mu} \varphi_2' \\ &\quad + evA_\mu A^\mu\varphi_2' - \frac{\lambda}{4!}\left(\varphi_1'^2+\varphi_2'^2\right)^2 - \frac{\lambda v}{2\ 3!}\left(\varphi_1'^2+\varphi_2'^2\right)\varphi_2' \\ &\quad - c_1\varphi_2' + c_0\end{aligned}$$

where m_2, c_1 and c_0 are the same as in Sec. 9.2.2. As before A_μ now has a mass $m_A = ev$ and φ_2' is the massive physical Higgs scalar. φ_1' is the "would be" Goldstone field (Higgs ghost). The crucial difference between the Abelian Higgs model and the Goldstone model is the presence of the field A_μ and the appearance of a mixing term $A_\mu\partial^\mu\varphi_1'$ in the bilinear part of the Lagrangian. Since the latter is **not** diagonal in the fields A_μ, φ_2' and φ_1' we cannot yet write down the Feynman rules. In any case we first have to fix a gauge for the vector potential A_μ . If we would chose simply the linear covariant gauge $\partial_\mu A^\mu = 0$ and correspondingly add a gauge fixing term $\mathcal{L}_{\rm GF} = -\frac{1}{2\xi}(\partial_\mu A^\mu)^2$ to the invariant Lagrangian, as appropriate for QED, we still would be left with a non-diagonal term in the bilinear part $\mathcal{L}_0$ of the Lagrangian which obscures the particle interpretation. We have to modify the gauge condition in such a way that we achieve at the same time a diagonalization of $\mathcal{L}_0$. There is a unique linear covariant gauge condition, the t'Hooft gauge,

$$\partial_\mu A^\mu(x) + \xi m_A \varphi_1' = 0$$

which does the job. Indeed by adding the gauge fixing term

$$\begin{aligned}\mathcal{L}_{\rm GF} &= -\frac{1}{2\xi}\left(\partial_\mu A^\mu + \xi ev\varphi_1'\right)^2 \\ &= -\frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2 - ev\partial_\mu A^\mu\varphi_1' - \frac{\xi m_A^2}{2}\varphi_1'^2\end{aligned}$$

for ξ an arbitrary gauge parameter we achieve:

---

[3]Notice that $\hat{j}^\mu$ is **not** the conserved Noether current. In any local gauge theory the latter is determined by the divergence of the field strength tensor (Maxwell equation).

$$\partial_\nu F^{\nu\mu} = j^\mu \leftrightarrow -\partial_\nu\frac{\partial\mathcal{L}}{\partial\partial_\nu A_\mu} = -\frac{\partial\mathcal{L}}{\partial A_\mu}\ .$$

This current is trivially conserved

$$\partial_\mu j^\mu = \partial_\mu\partial_\nu F^{\mu\nu} = 0$$

because $\partial_\mu\partial_\nu$ is symmetric while $F^{\mu\nu}$ is antisymmetric. In our model we find

$$j^\mu = -ie\left\{\varphi^* \overleftrightarrow{\partial^\mu} \varphi - 2ieA^\mu\varphi^*\varphi\right\} = e\varphi_1 \overleftrightarrow{\partial^\mu} \varphi_2 - e^2A^\mu\left(\varphi_1^2+\varphi_2^2\right)$$

which differs from $\hat{j}^\mu$ by a factor 2 in the second term. This difference is due to the fact that the interaction is not linear in A_μ as it is in QED for example.

i) the standard term for fixing the gauge of the gauge field propagator

ii) by a partial integration in the action $\int d^4x\, \mathcal{L}(x)$:

$$-ev\partial_\mu A^\mu \varphi_1' = -evA_\mu \partial^\mu \varphi_1'$$

we precisely get the term needed to cancel the non-diagonal term present in $\mathcal{L}_{inv}$, and,

iii) as a result of the diagonalization, the field φ_1' acquires a mass $m_1 = \sqrt{\xi} m_A$ which however is gauge dependent. This reveals that φ_1' cannot be a physical field because physics must be gauge invariant and hence cannot depend on ξ .

Because our model is Abelian we do not need a Faddeev-Popov term. According to the Faddeev-Popov construction, described in Sec. 8, we would get simply

$$\mathcal{L}_{\rm FP} = \partial_\mu \bar{\eta} \partial^\mu \eta \;;\;\; (\text{since}\;\; c_{ijk} \equiv 0)\;.$$

i.e. the FP-ghost do not couple to any other field and hence are completely free ghosts and can be ignored.

As a result the correct Lagrangian for the Abelian Higgs model is

$$\mathcal{L}_{eff} = \mathcal{L}_{\rm inv} + \mathcal{L}_{\rm GF},$$

from which we can setup the Feynman rules in the standard manner. The diagonal bilinear terms are given by

$$\begin{aligned}
\mathcal{L}_0(A) &= \frac{1}{2} A_\mu \left[\left(\Box + m_A^2 \right) g^{\mu\nu} - \left(1 - \frac{1}{\xi} \right) \partial^\mu \partial^\nu \right] A_\nu \\
\mathcal{L}_0(\varphi_2') &= \frac{1}{2} \varphi_2' \left[- \left(\Box + m_2^2 \right) \right] \varphi_2' \\
\mathcal{L}_0(\varphi_1') &= \frac{1}{2} \varphi_1' \left[- \left(\Box + \xi m_A^2 \right) \right] \varphi_1'
\end{aligned}$$

where we have performed the appropriate partial integrations.

The propagators, given by the inverses of the operators (kernels) in square brackets, are easily determined in momentum space (see Sec. 8):

$$\tilde{\Delta}_F^A(p,\xi)^{\mu\nu} = - \left(g^{\mu\nu} - (1-\xi) \frac{p^\mu p^\nu}{p^2 - \xi m_A^2} \right) \frac{1}{p^2 - m_A^2 + i\epsilon}$$

massive gauge boson propagator

$$\tilde{\Delta}_F^{\varphi_2'}(p) = \frac{1}{p^2 - m_2^2 + i\epsilon}$$

physical Higgs propagator

$$\tilde{\Delta}_F^{\varphi_1'}(p) = \frac{1}{p^2 - \xi m_A^1 + i\epsilon}$$

unphysical Higgs ghost propagator

Physical quantities must be gauge invariant and hence cannot depend on the gauge parameter ξ . In particular we may consider the limit $\xi \to \infty$. In this limit the A-propagator takes the form

$$-\left(g^{\mu\nu} - \frac{p^\mu p^\nu}{m_A^2}\right)\frac{1}{p^2 - m_A^2 + i\epsilon}$$

which describes a physical massive spin 1 boson. At the same time the ghost field φ_1' becomes infinitely heavy and thus freezes out from the dynamics. Therefore the gauge $\xi \to \infty$ is called the **physical or unitary gauge**. In this gauge ghosts are absent and massive spin 1 particles are described by the standard physical propagators (as listed in the Appendix to Sec. 2).

Our discussion of the quantization of the model confirms the physical interpretation we found earlier by heuristic arguments on the classical level.

One problem remains to be mentioned. In the unitary gauge we loose renormalizability because in this gauge the A-propagators, for $p^2 \to \infty$, behaves like $O(1)$ and not like $O(p^{-2})$ as it is necessary for a field theory to be renormalizable. We notice that for any finite ξ the A-propagator behaves as $O(p^{-2})$ and one indeed can prove that the model is renormalizable. The gauges with finite ξ therefore are called **renormalizable gauges**. In order to be able to control the high energy (ultraviolet) behavior calculations of higher order effects (loops) must be performed in a renormalizable gauge. Since physical quantities like S-matrix elements are gauge independent (this one can prove) we know that they are independent of the gauge parameter ξ . Thus, for physical quantities a calculation performed in a renormalizable gauge yields the same answer we would obtain in the unitary gauge. This means that we are able to control at the same time renormalizability and unitarity of the model.

10.3 Higgs mechanism for Yang-Mills theories (Higgs-Kibble mechanism)

In this chapter we briefly illustrate the Higgs mechanism for non-Abelian theories (*Kibble, 1967*). As an example we consider a model with the symmetry group $SU(2)$ which will play a central role in the standard model of electroweak interactions (mass generation for the weak gauge bosons).

Let us consider a $SU(2)$-doublet of complex scalar fields (Higgs field) with hypercharge $Y = +1$:

$$\Phi_b = \begin{pmatrix} \varphi^+ \\ \varphi_0 \end{pmatrix}$$

where the charge assignment follows from the Gell-Mann Nishijima relation $Q = T_3 + \frac{Y}{2}$. We also define a Y-charge conjugate field with hypercharge $Y = -1$:

$$\Phi_t = i\tau_2 \Phi_b^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \varphi^- \\ \varphi_0^* \end{pmatrix} = \begin{pmatrix} \varphi_0^* \\ -\varphi^- \end{pmatrix}$$

where $\varphi^- = (\varphi^+)^*$. In terms of the matrix

$$\tilde{\phi} = (\Phi_t, \Phi_b) = \begin{pmatrix} \varphi_0^* & \varphi^+ \\ -\varphi^- & \varphi_0 \end{pmatrix}$$

we may represent the doublet fields as

$$\Phi_b = \tilde{\phi}\chi_b \ , \ \ \Phi_t = \tilde{\phi}\chi_t$$

with isospinor basis vectors

$$\chi_b = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ , \ \chi_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ .$$

We also may write $\tilde{\phi}$ in terms of real fields H and φ_i $(i = 1, 2, 3)$

$$\tilde{\phi} = \frac{1}{\sqrt{2}} \left(\mathbf{1}\, H + i \tau_i \varphi_i \right) = \begin{pmatrix} \frac{H + i\varphi_3}{\sqrt{2}} & i \frac{\varphi_1 - i\varphi_2}{\sqrt{2}} \\ i \frac{\varphi_1 + i\varphi_2}{\sqrt{2}} & \frac{H - i\varphi_3}{\sqrt{2}} \end{pmatrix} \ .$$

By $\mathbf{1}$ we denoted the 2×2 unit matrix, τ_i are the Pauli matrices. The real and the complex representation are related by

$$\varphi^+ = i\, \frac{\varphi_1 - i\varphi_2}{\sqrt{2}} \ , \ \ \varphi_0 = \frac{H - i\varphi}{\sqrt{2}} \ , \ \ \varphi = \varphi_3 \ .$$

The matrix field $\tilde{\phi}$ has the following important properties:

$$\begin{aligned} \tilde{\phi}\tilde{\phi}^+ \quad &= \quad \tilde{\phi}^+\tilde{\phi} \quad &&= \quad S \cdot \mathbf{1} \\ \mathrm{Det}\ \tilde{\phi} \quad &= \quad \mathrm{Det}\ \tilde{\phi}^+ \quad &&= \quad S \end{aligned}$$

with $S = \varphi^+\varphi^- + \varphi_0^*\varphi_0$.
We notice that the singlet field S is given by

$$S = \Phi_b^+ \Phi_b = \varphi^+\varphi^- + \varphi_0^*\varphi_0 = \varphi^+\varphi^- + \frac{1}{2}\varphi^2 + \frac{1}{2}H^2 .$$

Now let us suppose that Φ_b has a non-vanishing vacuum expectation value

$$< 0 \mid \Phi_b \mid 0 >= \Phi_0 = \frac{v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

which is equivalent to

$$< 0 \mid H \mid 0 >= v \ ; \ \ < \varphi_i >= 0 \ .$$

In this case (Higgs phase) $\Phi_b^+ \Phi_b$ has a classical positive background term

$$S = \frac{v^2}{2}(1 + \mathcal{X})$$

where

$$\mathcal{X} = 2\, \frac{H'}{v} + \frac{H'^2 + \varphi^2}{v^2} + 2\, \frac{\varphi^+\varphi^-}{v^2} \ ; \ \ H' = H - v \ , \ \ < 0 \mid H' \mid 0 >= 0 \ .$$

The fields in S appear as a perturbation (quantum fluctuations) about the classical background field and in $R_e S > 0$ analytic functions $F(S)$, like S^{-1}, $\sqrt{S}$ or $1/\sqrt{S}$ are well defined as perturbation series in v^{-1} (in the standard model of electroweak interactions $v^{-2} = \sqrt{2} G_\mu$ where G_μ is the weak interaction Fermi constant).

In this sense the matrix field

$$\tilde{U} = \frac{\tilde{\phi}}{\sqrt{S}}$$

is well defined and has the property

$$\tilde{U}\tilde{U}^{+} = \tilde{U}^{+}\tilde{U} = \mathbf{1}\ , \ \ \mathrm{Det}\,\tilde{U} = \mathrm{Det}\,\tilde{U}^{+} = 1$$

which infers that $\tilde{U}$ is an $SU(2)$ matrix which we may write as

$$\tilde{U} = e^{i\frac{\tau_i}{2}\frac{\theta_i}{v}}$$

with three real fields θ_i . As a result we find the representation

$$\Phi_b = \tilde{U}\frac{\rho' + v}{\sqrt{2}}\chi_b\ ; \ \ \Phi_b^{+}\Phi_b = \frac{(\rho' + v)^2}{2}$$

for the Higgs doublet field.[28]

After this formal digression we are going to discuss the physics of the global and local $SU(2)$ Higgs model. We first study the Goldstone model with Lagrangian

$$\mathcal{L} = (\partial_\mu\Phi_b)^{+}(\partial^\mu\Phi_b) - V\left(\Phi_b^{+}\Phi_b\right) = (\partial_\mu\Phi_b)^{+}(\partial^\mu\Phi_b) + \mu^2\Phi_b^{+}\Phi_b - \lambda\left(\Phi_b^{+}\Phi_b\right)^2$$

exhibiting global $SU(2)$ invariance

$$\Phi_b(x) \to \Phi_b(x) = e^{i\frac{\tau_i}{2}\omega_i}\Phi(x)\ .$$

If $\mu^2 > 0$ the symmetry is spontaneously broken and we chose the ground state such that Φ_b has the vacuum expectation value Φ_0 given above. Our main interest again is the mass spectrum of the model. The square mass matrix is given by

$$M_{ik}^2 = \left.\left(\frac{\partial^2 V}{\partial\varphi_i\partial\varphi_k}\right)\right|_{\Phi_b=\Phi_0}\ .$$

Since V is already diagonal in the fields (depending on $\Phi_b^{+}\Phi_b$ only) it can be evaluated easily. For the ground state $V|_{\Phi_b=\Phi_0}$ is at the minimum and therefore

$$\frac{\partial V}{\partial\varphi_i} = \frac{\partial V}{\partial\left(\Phi_b^{+}\Phi_b\right)}\frac{\partial\left(\Phi_b^{+}\Phi_b\right)}{\partial\varphi_i} = \left(2\lambda\Phi_b^{+}\Phi_b - \mu^2\right)\tilde{\varphi}_i$$

at $\Phi_b = \Phi_0$ must vanish. This implies that the first factor must vanish for $\Phi_b = \Phi_0$, since for $\tilde{\varphi}_i = H$ the second factor is non-zero. Thus μ^2 takes the value $\mu^2 = \lambda v^2$. Then

$$\left.\frac{\partial^2 V}{\partial\varphi_k\partial\varphi_i}\right|_{\Phi_b=\Phi_0} = \left.2\lambda\tilde{\varphi}_k\tilde{\varphi}_i\right|_{\varphi_i=0\,,\,H=v} = \begin{cases} 2\lambda v^2 = m_H^2 & ;\ \varphi_i = \varphi_k = H \\ 0 & ;\ \text{otherwise}\end{cases}$$

[28] By expanding the exponential and utilizing the properties of the Pauli matrices one finds

$$H' + v = (\rho' + v)\cos\frac{\theta}{2v}\ ; \ \ \varphi_i = (\rho' + v)\frac{\theta_i}{v}\sin\frac{\theta}{2v}$$

with $\theta = \sqrt{\theta_i\theta_i}$.

Thus, as ascertained by the Goldstone theorem, the model exhibits 3 Goldstone bosons $\varphi^{\pm}$ and φ and a massive neutral scalar H with mass $m_H^2 = 2\lambda v^2$. The situation dramatically changes if we "gauge" the symmetry. If we couple Φ_b in a locally gauge invariant way to $SU(2)$ Yang-Mills fields we have

$$\mathcal{L}_{inv} = -\frac{1}{4} W_{a\mu\nu} W_a^{\mu\nu} + (D_\mu \Phi_b)^+ (D_\mu \Phi_b) - V\left(\Phi_b^+ \Phi_b\right)$$

as a resulting Lagrangian. $W_{a\mu\nu}$ is the non-Abelian field-strength tensor

$$W_{a\mu\nu} = \partial_\mu W_{a\nu} - \partial_\nu W_{a\mu} + g\epsilon_{abc} W_{b\mu} W_{c\nu}$$

and

$$D_\mu \Phi_b = \left(\partial_\mu - ig\frac{\tau_a}{2} W_{a\mu}\right) \Phi_b$$

the covariant derivative of the Higgs field.

Using the result we have derived at the beginning of this chapter, we may write the Higgs field in the form

$$\Phi_b = e^{i\frac{\tau_i}{2}\frac{\theta_i}{v}} \frac{\rho' + v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

This representation tells us that Φ_b is a local gauge transform of

$$\Phi_b^u = \frac{\rho' + v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with $SU(2)$ transformation matrix

$$\tilde{U}(\theta) = e^{i\frac{\tau_i}{2}\frac{\theta_i}{v}} .$$

With other words by a local gauge transformation

$$\begin{aligned} \Phi_b \rightarrow \Phi_b^u &= \tilde{U}^+(\theta) \Phi_b \\ W_{a\mu}\tau_a \rightarrow W_{a\mu}^u \tau_a &= \tilde{U}^+(\theta) W_{a\mu} \tau_a \tilde{U}(\theta) + \frac{2i}{g} \tilde{U}^+(\theta) \partial_\mu \tilde{U}(\theta) , \end{aligned}$$

which leaves the Lagrangian invariant (see Sec. 6.2), the fields θ_i can be eliminated completely from the invariant Lagrangian. The Yang-Mills term $W_{a\mu\nu} W_a^{\mu\nu}$ reads the same in terms of the new (unitary gauge) fields $W_{a\mu}^u$. The Higgs potential now depends on one single field, the ρ-field, $V\left(\Phi_b^+ \Phi_b\right) = V\left(\frac{(\rho'+v)^2}{2}\right)$ and provides the mass and the self-interactions. The most interesting change shows up in the Higgs "kinetic" term:

$$\begin{aligned} (D_\mu \Phi_b^u)^+ (D^\mu \Phi_b^u) &= \frac{1}{2}(0,1) \left(\partial_\mu \rho' + ig\frac{\tau_a}{2} W_{a\mu}(\rho' + v)\right) \left(\partial^\mu \rho' - ig\frac{\tau_b}{2} W_b^\mu (\rho' + v)\right) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \frac{1}{2} (\partial \rho')^2 + \frac{g^2}{8} (\rho' + v)^2 W_{a\mu} W_a^\mu \end{aligned}$$

exhibiting a mass term for the triplet of Yang-Mills fields [29]. Again the mass term

$$\frac{M_W^2}{2} W_{a\mu} W_a^\mu = \frac{g^2 v^2}{8} W_{a\mu} W_a^\mu$$

is generated owing to the presence of a vacuum condensate v and the mass is given by

$$M_W = \frac{gv}{2} \ .$$

The triplet of massive gauge bosons exhibits three additional degrees of freedom, while the three scalars θ_i (the would-be Goldstone bosons) have disappeared from the physical spectrum. We notice that the component of the Higgs doublet which picks the non-vanishing vacuum expectation value always describes a physical massive spin 0 boson, the physical Higgs boson. Such a particle so far has not been found in nature. This means that, provided it exists at all, it must be heavy enough such that its production was not possible up to presently accessible energies.

As a result we have learned that in the Higgs phase ($v > 0$) there exists a gauge, the physical or unitary gauge, for which the invariant Lagrangian takes a particularly simple and physically transparent (ghost free) form.

$$\begin{aligned} \mathcal{L}_{\rm inv} &= -\frac{1}{4} W_{a\mu\nu} W_a^{\mu\nu} + \frac{M_W^2}{2} W_{a\mu} W_a^\mu \\ &\quad + \frac{M_W^2}{v} H W_{a\mu} W_a^\mu + \frac{M_W^2}{v^2} H^2 W_{a\mu} W_a^\mu \\ &\quad + \frac{1}{2} (\partial H)^2 - \frac{m_H^2}{2} H^2 - \frac{\lambda}{4} H^4 - \lambda v H^3 \end{aligned}$$

where we dropped the constant term $\frac{\lambda v^4}{4}$. H denotes the physical Higgs (denoted by ρ' so far). Another novel feature of models in the Higgs phase is that masses and couplings are no longer independent parameters. We now have typical **mass-couplings** relationships

$$M_W = \frac{gv}{2} \ , \quad m_H = \sqrt{2\lambda} v \ .$$

In the unbroken phase $\mu^2 < 0$, $v = 0$ we have the independent parameters

$$g, \lambda, m^2 = -\mu^2$$

and the model describes 3 massless gauge bosons and 4 physical scalars of equal mass m . In contrast, in the Higgs phase, where 3 massive gauge bosons and one massive physical scalar are present, the independent parameters are

$$g, \lambda, v = \sqrt{\frac{\mu^2}{\lambda}} \quad \text{or} \quad M_W, m_H = \sqrt{2\mu^2} \ , \ v \ .$$

The effect of the Higgs mechanism may be summarized as follows:

[29] We have used

$$\tau_a \tau_b W_{a\mu} W_b^\mu = \frac{1}{2} \{\tau_a, \tau_b\} W_{a\mu} W_b^\mu = \delta_{ab} W_{a\mu} W_b^\mu .$$

Notice that the mixed term $(\partial_\mu \rho') W_a^\mu$ drops out.

i) it generates the mass of the gauge bosons

ii) it implies the existence of a Higgs boson the mass of which is a free parameter (equivalent to the scalar self coupling of the Higgs potential)

iii) the Higgs boson acts as a "physical cut-off" for the massive vector boson gauge theory which without the Higgs would not be renormalizable.

iv) it implies mass-couplings relationships which can be tested experimentally since the masses and couplings can be determined independently.

For the standard model of electroweak interactions besides the $SU(2)$ also the weak hypercharge $U(1)_Y$ will be a local symmetry. This will complicate matters slightly. In addition also the fermion masses will be generated by the Higgs mechanism.

Exercises: Section 10

② Solve $\Delta\vec{B} = \frac{1}{\lambda_L^2}\vec{B}$ for a boundary between a normal state in the half-space $x < 0$ and a super conducting state in the complementary half-space $x \geq 0$. In the normal state an external magnetic field $\vec{H}$ parallel to the positive z-axis is applied. How does the superconducting current look like? Also discuss the electric field $\vec{E} = \frac{4\pi}{c^2}\lambda_L^2\dot{\vec{j}}_s$ present if $\vec{H}$ varies with time.

1. (see Sec. 10.3) The $SU(2)$ triplet of gauge fields $W_{a\mu}$ do not carry hypercharge. Use $Q = T_3 + \frac{Y}{2}$ to show that

$$W_\mu^+ = \frac{W_{1\mu} - iW_{2\mu}}{\sqrt{2}} \ , \quad W_\mu^0 = W_{3\mu}$$

carry electrical charge +1 and 0, respectively. Hint: Use the commutation relation

$$[T_3, T_\pm] = \pm T_\pm \ , \quad T_\pm = T_1 \pm iT_2$$

of the generators $T_i = \frac{\tau_i}{2}$ of $SU(2)$.

2. (see Sec. 10.3) Evaluate $(D_\mu\Phi_b)^+(D^\mu\Phi_b)$ in terms of $\varphi^\pm, \varphi$ and $H' = H - v$ (for non-unitary gauges) and discuss the result. Determine the gauge fixing function

$$C_a = \partial_\mu W_a^\mu + \dots$$

and add

$$\mathcal{L}_{\mathrm{GF}} = -\frac{1}{2\xi}C_a^2$$

to $\mathcal{L}_{\mathrm{inv}}$ in such a way that the bilinear part

$$\mathcal{L}_{\mathrm{inv}} + \mathcal{L}_{\mathrm{GF}} = \mathcal{L}_0 + \mathcal{L}_{\mathrm{int}}$$

is diagonal in the fields. Discuss the particle spectrum of the model.

11 Weak interactions at low energies

11.1 Introduction

The history of the theory of weak interaction processes started around 1930 at a time when the photon (γ) (Einstein 1905) the electron (e^-) (Wiechert, Thomas) and the proton (p) (H-atom) were the only known elementary particles. These together with their antiparticles and the neutrinos are all the stable particles we know. It was not known at that time that nuclei contain besides protons also neutrons. Weak processes had to do with decays, the only known weak processes were β–decays, like

$$^6He \to ^6 Li + e^- + \text{ missing energy } ,$$

were very puzzling for the following reason. The observed α– and γ–spectra of nuclei are discrete, which tells us that the nuclear states are discrete. Therefore one expected the electrons from β–decays to be mono-energetic with energy

$$E_e \simeq \Delta = (M_i - M_f)\, c^2$$

determined by the mass difference between the initial (i) and final (f) nuclei. Note that, for a two body decay, in the CM frame (where $\vec{p}_f = -\vec{p}_e$)

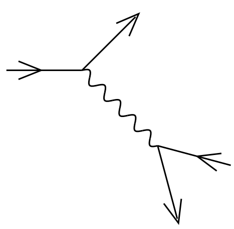
$$E_f^{\rm kin} = \frac{\vec{p}_e^{\,2}}{2M_f} \ll E_e + \sqrt{\vec{p}_e^{\,2} + m_e^2}$$

the kinetic energy of the final nucleus is completely negligible. In contrast to these theoretical expectations, experimentally a continuous electron energy spectrum with $E_e^{max} \simeq \Delta$ was observed (Chadwick 1914). In addition to this *crisis of energy conservation* also angular momentum conservation was in conflict since in $^6He \to ^6 Li + e^-$ the nuclear spin of ^6He is $J = 0$ while ^6Li has $J = 1$. This lead Pauli to propose the existence of a neutral very light fermion, later called the neutrino, and to explain the β–decays as three body decays, for our example,

$$^6He \to ^6 Li + e^- + \bar{\nu}_e \ ,$$

where $\bar{\nu}_e$ escapes observation. After the discovery of the neutron (n) by Chadwick in 1932 it became clear that nuclear β–decay was related to neutron β–decay

$$n \to p + e^- + \bar{\nu}_e$$



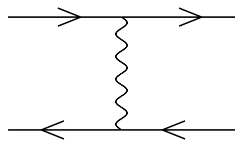
which is a **four fermion process**. For free nucleons the crossed reaction

$$p \to n + e^+ + \nu_e$$

is forbidden kinematically ($M_n > M_p$), it can take place inside a nucleus (K-capture), as for example, ${}^{10}C \to^{10} B^* + e^+ + \nu_e$.

In 1934 Fermi proposed a four fermion theory of β–decay in close analogy to a four fermion effective interaction in QED. In QED the transition matrix element for a four fermion process like Coulomb scattering of a proton and an electron

$$p + e^- \to p + e^-$$



is given, in the Feynman gauge, by

$$T = e^2 \left(\bar{u}_p \gamma_\mu u_p\right) \left(-\frac{g_{\mu\nu}}{q^2}\right) \left(\bar{u}_e \gamma_\nu u_e\right)$$

which, in the one-photon exchange approximation, exhibits a current–current form of effective interaction described by an effective Lagrangian

$$\mathcal{L}^{\text{eff}}_{QED} \simeq -\frac{e^2}{q^2} j^{(p)}_{\mu\,\text{em}}\, j^{(p)\mu}_{\text{em}} \quad .$$

Fermi postulated that weak interaction responsible for β–decay is of an analogous form, however, short ranged i.e. it has no propagator. The matrix element for the reaction then Eq. 11.1 reads

$$T = \frac{G_F}{\sqrt{2}} \left(\bar{u}_p \gamma_\mu u_p\right) \left(\bar{u}_e \gamma^\mu u_e\right)$$

which derives from an effective interaction Lagrangian

$$\mathcal{L}^{\text{eff}}_{\text{weak}} = \frac{G_F}{\sqrt{2}} \left(\bar{\psi}_p \gamma_\mu \psi_n\right) \left(\bar{\psi}_e \gamma^\mu \psi_{\nu_e}\right)$$

obtained by replacing the spinors by the corresponding fields. In contrast to the electromagnetic current $\bar{\psi}_e \gamma_\mu \psi_e$, which is electrically neutral in that it creates or destroys an e^+e^-–pair or destroys and recreates an electron etc., the weak current

$$J^+_{\mu\,\text{hadron}} = \bar{\psi}_p \gamma_\mu \psi_n$$

in β–decay is a **charged current (CC)** carries charge 1, in units of the positron charge, as it destroys a neutron and creates a proton. The Hermitean conjugate current

$$J_\mu \equiv J^-_\mu = \left(J^+_\mu\right)^+$$

is

$$J^-_{\mu\,\mathrm{hadron}} = \bar{\psi}_n \gamma_\mu \psi_p$$

and destroys one unit of charge. For the leptons, similarly, we have

$$J^+_{\mu\,\mathrm{lepton}} = \bar{\psi}_{\nu_e} \gamma_\mu \psi_e$$

such that we may write the Fermi Lagrangian in the form

$$\mathcal{L}^{\mathrm{eff}}_{\mathrm{weak}} = \frac{G_F}{\sqrt{2}} \left(J^+_{\mu\,\mathrm{hadron}}(x)\; J^\mu_{\mathrm{lepton}}(x) \;+\; \mathrm{h.c.} \right)$$

where the Hermitean conjugate (h.c.) term describes inverse β–decay . While the individual currents do not conserve the charge the Lagrangian does. The interaction must have overall charge conservation and it must be a Lorentz scalar.

In 1936 Anderson and Neddermeyer discovered the muon....

From the investigations of nuclear β–decay it soon became clear the the Fermi theory was not able to explain all the β–decays . Subsequently the Fermi Ansatz was generalized (Gamow, Teller and others) to the most general form of a four fermion interaction

$$\begin{array}{rclcl} \mathcal{L}^{\mathrm{eff}}_{\mathrm{weak}} & = & \frac{G_F}{\sqrt{2}} \sum_i \{ & & c_i \; (\bar{\psi}_1 \Gamma^i \psi_2)\, (\bar{\psi}_3 \Gamma_i \psi_4) \\ & & & + & c'_i \; (\bar{\psi}_1 \Gamma^i \psi_2)\, (\bar{\psi}_3 \Gamma_i \gamma_5 \psi_4) \\ & & & + & \mathrm{h.c.} \qquad \} \end{array}$$

with

$$\Gamma^i = \{\mathbf{1}, \gamma^\mu, \sigma^{\mu\nu}, \gamma^\mu \gamma_5, i\gamma_5\}$$

a basis of 4×4 matrices. By $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ as usual we denote the antisymmetric tensor. The matrices Γ^i have the Hermitecity property $\Gamma^i = \gamma^0 \Gamma^{i+} \gamma^0$ such that $\bar{\psi}\Gamma^i\psi$ is Hermitean and the transforms under L–transformations as **scalar (S)**, **vector (V)**, **tensor (T)**, **axial vector (A)** and **pseudoscalar (P)**, respectively. The couplings S, V and T are parity even the couplings A and P are parity odd.

Before 1957 the physics community was taking parity and time reversal invariance for granted and the requirement of these symmetries leads to

$$\begin{array}{ll} \text{P invariance}: & c_i,\; c'_i \text{ real} \\ \text{T invariance}: & c'_i \;=\; 0 \end{array}$$

The existence of parity violating weak interactions was proposed by Lee and Yang in 1956 to solve the so called θ–τ puzzle. At that time there apparently existed two strange particles, which were called θ and τ, and decayed as

$$\theta^\pm \to \pi^\pm + \pi^0 \quad , \quad \tau^\pm \to \pi^\pm + \pi^+ \pi^-$$

which had identical mass and lifetime, close to the values known for the K^0, and identical quantum numbers except from their opposite intrinsic parities. This follows from the observation that both decays lead to a zero angular momentum state and by the fact that the pions are spin zero particles with negative intrinsic parity. Lee and Yang pointed out the possibility to identify

$$\theta^{\pm} \equiv \tau^{\pm} \equiv K^{\pm}$$

if weak interactions do not preserve parity. The crucial experiment performed by Wu in 1957 was the study of the decay of polarized

$$^{60}Co \to ^{60} N_i^* + e^- + \bar{\nu}_e$$

revealed that parity indeed was violated maximally, with equal strength of parity conserving and parity violating interactions. The result could be interpreted in such a way that only left handed neutrinos exist:

$$\psi_\nu \to \psi_{\nu L} = \frac{1-\gamma_5}{2}\psi_\nu$$

This two component theory of the neutrino or Weyl fermion theory (Weyl 1929) was proposed in 1957 independently by Salam, Landau and Lee and Yang. In 1958 Marshak and Sudarshan, Feynman and Gell–Mann, Sakurai and Theis independently showed that a vector minus axial–vector form of the charge changing weak current, the so called **V–A interaction**, was able to explain all experimental facts.

As a matter of fact 24 years of intense activities in weak interaction physics revealed that a “simple” modification, V $\to$ V–A of the 1934 Fermi theory was the right answer. Thus the effective charge changing weak interaction Lagrangian reads

$$\mathcal{L}^{\mathrm{eff}}_{\mathrm{weak}} = \frac{G_F}{\sqrt{2}}\, J^+_\mu(x)\, J^\mu(x)$$

where the current has a leptonic and a hadronic piece of the V–A form

$$J_\mu = J_{\mu\,\mathrm{lepton}} + J_{\mu\,\mathrm{hadron}}$$

with

$$\begin{aligned} J_{\mu\,\mathrm{lepton}} &= \bar{\psi}_e\gamma_\mu\,(1-\gamma_5)\,\psi_{\nu_e}\bar{\psi}_\mu\gamma_\mu\,(1-\gamma_5)\,\psi_{\nu_\mu} + \cdots \\ J_{\mu\,\mathrm{hadron}} &= \end{aligned}$$

This Lagrangian describes at the same time processes like μ–decay

$$\mu \to e^- + \bar{\nu}_e + \nu_\mu$$

and β–decay . It exhibits $e-\mu$ universality and $\mu-\beta$ universality ($G_\beta \simeq G_\mu$). The existence of the neutrino was experimentally established in the reaction

$$\bar{\nu}_e + p \to n + e^+$$

by Reines and Cowens in 1953. Experiments with ν_μ's from π–decays

$$\begin{aligned} \pi^- \quad &\rightarrow \quad \mu^- + \nu_\mu \\ &\qquad ?? \rightarrow e^- \bar{\nu}_e + \nu_\mu \end{aligned}$$

showed that

$$\nu_\mu + Z \rightarrow (Z+1) + \mu^-$$

takes place, while

$$\nu_\mu + Z \not\rightarrow (Z+1) + e^-$$

is not found. Here Z is a nucleus with Z protons and N neutrons and $(Z+1)$ is a nucleus with Z+1 protons and N-1 neutrons. This proves that

$$\nu_\mu \neq \nu_e \ \ !$$

In addition the reaction

$$\mu \rightarrow e + \gamma$$

was never found, which leads to the conclusion that , within experimental limits, there exist conserved **lepton numbers** L_e and L_μ.

The validity of the V–A hypothesis ($\mu - \beta$ universality) for the hadronic current which includes no neutrino and therefore cannot be "explained" just by the absence of the right handed ν_R's in Nature was clarified by the investigation of the purely hadronic parity violating reaction

$$\Lambda^0 \rightarrow p + \pi^- \quad .$$

11.2 μ–decay

A detailed investigation of the μ–decay

$$\mu^-(p_0, n_0) \ \rightarrow e^-(p_1, n_1) \ + \bar{\nu}_E(p_2) \ + \nu_\mu(p_3)$$

is particularly tricky and challenging due to the fact that the two neutrinos escape any direct observation such that all information we can have must be derived from the electron spectrum. The p_i denote the particle momenta and the n_i possible polarization vectors. Thus all experimental information must derive from

$$d\Gamma = \frac{1}{2m_\mu} \frac{(2\pi)^4}{(2\pi)^6} d\mu(p_1) \int \frac{d^3p_2}{2E_2} \frac{d^3p_3}{2E_3} \delta^{(4)}(Q - p_2 - p_3) \sum_{r_2, r_3} |T|^2$$

with both neutrinos integrated out and summed over the spins. We will see that for massless neutrinos the square of the matrix element T is proportional to $p_{2\rho} p_{3\sigma}$

$$\sum_{r_2, r_3} |T|^2 = p_{2\rho} p_{3\sigma} F^{\rho\sigma}(p_0, p_1, n_0, n_1)$$

such that the integral over the two neutrinos may be performed independent of the form of the interaction. One obtains

$$\begin{aligned}
(2\pi)^2\, I_{\rho\sigma} &= \int \frac{d^3p_2}{2E_2}\frac{d^3p_3}{2E_3}\delta^{(4)}(Q-p_2-p_3)\, p_{2\rho}p_{3\sigma} \\
&= \int \frac{d^3p_2}{2|\vec{p}_2|}\frac{1}{2|\vec{Q}-\vec{p}_2|}\delta(Q^0-|\vec{p}_2|-|\vec{Q}-\vec{p}_2|)\, p_{2\rho}(Q-p_2)_\sigma \\
&= \frac{\pi}{24}\left(Q^2 g_{\rho\sigma}+2Q_\rho Q_\sigma\right)
\end{aligned}$$

such that I is a symmetric tensor:

$$I_{\rho\sigma} = I_{\sigma\rho} \quad .$$

The matrix element can be easily calculated. The relevant part of the effective Lagrangian reads

$$\mathcal{L}_{eff,int} = -\frac{G}{\sqrt{2}}\left(\bar{e}\gamma^\alpha\left(1-\gamma_5\right)\nu_e\right)\left(\bar{\nu}_\mu\gamma_\alpha\left(1-\gamma_5\right)\mu\right)$$

and thus

$$\begin{aligned}
T &= {}_{out}\!< e^-, \bar{\nu}_e\nu_\mu|\mu^- >_{in} \\
&= \frac{G}{\sqrt{2}}\left(\bar{u}_e\gamma^\alpha\left(1-\gamma_5\right)v_{\nu_e}\right)\left(\bar{u}_{\nu_\mu}\gamma_\alpha\left(1-\gamma_5\right)u_\mu\right)
\end{aligned}$$

and

$$T^* = \frac{G}{\sqrt{2}}\left(\bar{v}_{\nu_e}\gamma^\beta\left(1-\gamma_5\right)u_e\right)\left(\bar{u}_\mu\gamma_\beta\left(1-\gamma_5\right)u_{\nu_\mu}\right)$$

The transition probability is thus proportional to

$$|T|^2 = \frac{G^2}{2}\bar{u}_e\gamma^\alpha\left(1-\gamma_5\right)v_{\nu_e}\bar{v}_{\nu_e}\gamma^\beta\left(1-\gamma_5\right)u_e\bar{u}_{\nu_\mu}\gamma_\alpha\left(1-\gamma_5\right)u_\mu\bar{u}_\mu\gamma_\beta\left(1-\gamma_5\right)u_{\nu_\mu}$$

11.3 Neutrino scattering, and the weak mixing angle.

Exercises: Section 11

① XX

② In the Fermi type model of low energy effective weak interactions, the weak interaction Lagrangian for charged current (cc) processes is

$$\mathcal{L}_{\text{eff,int}} = -\frac{G_F}{\sqrt{2}}J^+_\mu J^{\mu-}$$

where

$$\begin{aligned}
J^2_\mu &= \Sigma\bar{\nu}_l(x)\gamma_\mu\left(1-\gamma_5\right)l(x) + \text{ hadronic terms} \\
& \quad\; l = e,\mu,\tau
\end{aligned}$$

is the charged $SU(2)_L$ current. $l(x)$ and $\nu_l(x)$ are the Dirac fields describing a lepton l and its neutrino $\nu : k$, respectively. Calculate, within this model, the μ-decay rate for unpolarized states:

$$\mu^-(p^0) \to e^-(p_1) + \bar{\nu}_l(p_2) + \nu_\mu(p_3).$$

Use

$$d\Gamma = \frac{1}{2m_\mu}\frac{1}{(2\pi)^2} d\mu(p_1) \int \frac{d^3p_2}{2E_2} \int \frac{d^3p_3}{2E_3} \delta^{(4)}\left(Q - p_2 - p_3\right) \times \sum_{\text{spins}} \mid T \mid^2$$

where $Q = p_0 - p_1$.

12 Chiral transformations, chiral symmetry and the axial-vector anomaly

12.1 Chiral fields and the $U(1)$-axial current

In the zero mass limit a free Dirac particle decouples into two chiral states described by the Weyl fields ψ_L and ψ_R. This becomes evident if we write the free Dirac Lagrangian in terms of ψ_L and $\psi_R : \psi = \psi_L + \psi_R$[30]

$$\begin{aligned}\mathcal{L} &= \bar{\psi} i\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \\ &= \bar{\psi}_L i\gamma^\mu \partial_\mu \psi_L + \bar{\psi}_R i\gamma^\mu \partial_\mu \psi_R - m\left(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L\right)\end{aligned}$$

If $m = 0$ we observe that $\mathcal{L}$ decomposes into two independent Lagrangians for the fields ψ_L and ψ_R. In this case $\mathcal{L}$ is not only invariant under global phase transformations

$$\psi \to e^{i\alpha}\psi \; ; \quad e^{i\alpha} \in U(1)_V \quad \text{vector group}$$

(infinitesimal: $\delta\psi = i\alpha\psi,\ \delta\bar{\psi} = -i\alpha\bar{\psi}$) but also under global **chiral transformations**

$$\psi \to e^{i\beta\gamma_5}\psi \; ; \quad e^{i\beta\gamma_5} \in U(1)_A \quad \text{axial group}$$

(infinitesimal: $\delta\psi = i\beta\gamma_5\psi, \delta\bar{\psi} = i\beta\bar{\psi}\gamma_5$; note the change of sign for $\delta\bar{\psi}$!). Since

$$\gamma_5\ \psi_L = -\psi_L \ , \quad \gamma_5\ \psi_R = \psi_R$$

the chiral transformation for the fields ψ_L and ψ_R reads:

$$\begin{aligned}\psi_L \to e^{i\beta\gamma_5}\psi_L &= e^{-i\beta}\psi_L \\ \psi_R \to e^{i\beta\gamma_5}\psi_R &= e^{\,i\beta}\psi_R\end{aligned}$$

and hence the "chiral" fields transform with opposite chirality (handedness). The conserved Noether currents

$$j^\mu(x) = -\left\{\frac{\delta\bar{\psi}}{\delta\omega}\frac{\partial\mathcal{L}}{\partial\partial_\mu\bar{\psi}} + \frac{\partial\mathcal{L}}{\partial\partial_\mu\psi}\frac{\delta\psi}{\delta\omega}\right\}$$

for the $U(1)_V \otimes U(1)_A$ symmetry group are the vector current $\left(\frac{\delta\psi}{\delta\omega} = i\psi\right)$[31]

$$j^\mu = \bar{\psi}\gamma^\mu\psi \; ; \quad \partial_\mu j^\mu = 0$$

---

[30] Remember: γ_5 is Hermitian and anticommutes with all γ-matrices. Therefore $\Pi_\pm = (1 \pm \gamma_5)/2$ are Hermitian projection operators: $\Pi_+ + \Pi_- = 1$, $\Pi_\pm^2 = \Pi_\pm$, $\Pi_+\Pi_- = \Pi_-\Pi_+ = 0$. Since by definition $\psi_L = \Pi_-\psi$ and $\psi_R = \Pi_+\psi$ we have $\bar{\psi}_L = \psi_L^+\gamma^0 = \psi^+\Pi_-\gamma^0 = \psi^+\gamma^0\Pi_+ = \bar{\psi}\Pi_+$ and similarly $\bar{\psi}_R = \bar{\psi}\Pi_-$. Therefore $\bar{\psi}_L\gamma^\mu\psi_R = \bar{\psi}_R\gamma^\mu\psi_L = 0$ and $\bar{\psi}_L\psi_L = \bar{\psi}_R\psi_R = 0$.

[31] For massive fields and charged (non-diagonal) currents $j^\mu = \bar{\psi}_1\gamma^\mu\psi_2$ and $j_5^\mu = \bar{\psi}_1\gamma^\mu\gamma_5\psi_2$ the divergences of the currents can be easily calculated for free fields: By the Dirac equation $\gamma^\mu\partial_\mu\psi_i = -im_i\psi_i$, $\partial_\mu\bar{\psi}_i\gamma^\mu = im_i\bar{\psi}_i$ and hence

$$\partial_\mu j^\mu = \left(\partial_\mu\bar{\psi}_1\right)\gamma^\mu\psi_2 + \bar{\psi}_1\gamma^\mu\left(\partial_\mu\psi_2\right) = i(m_1 - m_2)\bar{\psi}_1\psi_2$$

and

$$\partial_\mu j_5^\mu = \left(\partial_\mu\bar{\psi}_1\right)\gamma^\mu\gamma_5\psi_2 + \bar{\psi}_1\gamma^\mu\gamma_5\left(\partial_\mu\psi_2\right) = i(m_1 + m_2)\bar{\psi}_1\gamma_5\psi_2 \; .$$

Thus j^μ is conserved only if $m_1 = m_2$, for j_5^μ to be conserved we must require $m_1 = m_2 = 0$.

and the axial-vector current $\left(\frac{\delta\psi}{\delta\omega} = i\gamma_5\psi\right)$

$$j_5^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi \; ; \;\; \partial_\mu j_5^\mu = 0 \; .$$

Both vectors $V^\mu = \bar{\psi}\gamma^\mu\psi$ and axial vectors $A^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi$ are invariant under chiral transformations whereas scalars $S = \bar{\psi}\psi$, pseudoscalars $P = i\bar{\psi}\gamma_5\psi$ and tensors $T^{\mu\nu} = \bar{\psi}\sigma^{\mu\nu}\psi$ are not.

Now, suppose that ψ describes a massless charged Dirac particle which couples to photons. The massless QED Lagrangian

$$\mathcal{L} = \bar{\psi}\gamma^\mu\left(\partial_\mu - ieA_\mu\right)\psi$$

formally still looks chirally invariant. In fact, however, global chiral symmetry is broken now by the **axial-vector anomaly** *(Adler, Bell and Jackiw, 1969)*

$$\partial_\mu j_5^\mu = \frac{e^2}{8\pi^2}\tilde{F}_{\mu\nu}F^{\mu\nu} \neq 0 \tag{12.1}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is the electromagnetic field strength tensor and $\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$ its dual pseudotensor (parity odd). The pseudoscalar density is a divergence of a gauge dependent pseudovector

$$\tilde{F}_{\mu\nu}F^{\mu\nu} = \partial^\mu K_\mu \; ; \;\; K_\mu = 2\epsilon_{\mu\rho\nu\sigma}A^\rho\partial^\nu A^\sigma . \tag{12.2}$$

This anomaly is a quantum effect which cannot be removed. In particular it cannot be compensated by adding a counterterm to the Lagrangian which would restore the symmetry. For QED the non-conservation of the j_5^μ current poses no problems because photons do not couple to an axial-vector current. For gauge theories for which gauge fields couple to axial-vector currents γ_5-anomalies are disastrous. If they are present they destroy renormalizability and unitarity. The problems are evident if we consider, for example, the Abelian subgroup $U(1)_Y$ of hypercharge Y of the electroweak standard model. The latter is not parity conserving and thus does not treat left-handed and right-handed fields in a democratic way. As we shall see in Sec. 13, the leptonic hypercharge current has the form

$$j_Y^\mu = \frac{1}{2}\left(\bar{\ell}_L\gamma^\mu\ell_L + \bar{\nu}_{\ell L}\gamma^\mu\nu_{\ell L}\right) + \bar{\ell}_R\gamma^\mu\ell_R$$

where ℓ denotes a lepton field and ν_ℓ its associated neutrino field. This current contributes to the action a term

$$\mathcal{A}_{int}^{(Y)} = g'\int d^4x \; j_Y^\mu(x)\; B_\mu(x) \; .$$

B_μ is the $U(1)_Y$ gauge field and g' the corresponding gauge coupling. Under a local gauge transformation

$$B_\mu(x) \to B'_\mu(x) = B_\mu(x) - \partial_\mu\omega(x)$$

the action changes by

$$\delta\mathcal{A}_{int}^{(Y)} = -g'\int d^4x \; j_Y^\mu(x)\,\partial_\mu\omega(x)$$

and after a partial integration

$$\delta\mathcal{A}_{int}^{(Y)} = -g'\int d^4x \; \left(\partial_\mu j_Y^\mu(x)\right)\,\omega(x)$$

we see that the action cannot be gauge invariant unless the current is conserved. In fact

$$(\partial_\mu j_Y^\mu(x)) \neq 0$$

has an anomaly as we shall see below. In any case, for a gauge theory in order to be renormalizable we must make sure that axial anomalies are absent. For QED and QCD, which are parity conserving, anomalies are absent because the gauge fields couple to vector currents only. The $SU(2)_L \otimes U(1)_Y$ electroweak theory is **not** anomaly save. It can be rendered renormalizable only by **anomaly cancellation** between leptons and quarks. This leads to **lepton-quark duality**: the electroweak standard model is renormalizable only if fermions appear in lepton-quark families.

12.2 The chiral group $U(n)_V \otimes U(n)_A$

In general we have to consider currents associated with non-Abelian symmetries. What we said about Abelian models carries over to the non-Abelian ones. Of particular interest is the so called **chiral group**

$$\mathcal{G}_F = U(N_F)_V \otimes U(N_F)_A \simeq SU(N_F)_V \otimes SU(N_F)_A \otimes U(1)_V \otimes U(1)_A$$

with N_F the number of quark flavors. As we know the pions have odd inner parity which means that an effective pion field has the transformation properties of a $SU(2)_I$-triplet

$$\vec{\pi}(x) = \bar{\Psi} i\gamma^5 \frac{\vec{\tau}}{2}\,\Psi(x)\ ;\ \ \Psi \simeq \left(\begin{array}{c} p \\ n \end{array}\right)$$

where Ψ is a $SU(2)_I$-doublet of fermions with the quantum numbers of the nucleons, as indicated. This shows that although QCD is parity conserving non-trivial chiral properties are crucial in the theory of strong interaction. QCD distinguishes between leptons (which do not participate in strong interactions) and colored triplets of quarks and antiquarks it does not distinguish between different flavors however. Strong interactions therefore exhibit a flavor symmetry. For free quarks known are the f=u, d, c, s, t and b (6 flavors) quarks which show up in three colors each c=red (r), green (g) and blue(b), we have a Lagrangian

$$\mathcal{L}^q = \sum_{f=1}^{6}\sum_{c=1}^{3} (\bar{\psi}_{cf} i\,\gamma^\mu \partial_\mu \psi_{cf} - m_f \bar{\psi}_{cf}\psi_{cf})$$

To the extend that we can neglect the quark masses this Lagrangian has a global $U(18)_V \otimes U(18)_A$ symmetry. If strong interaction is switched on by the minimal substitution

$$\partial_\mu \to D_\mu = \partial_\mu \delta_{cc'} - i g_s \left(\frac{\lambda_i}{2}\right)_{cc'} G_{i\mu}\ ,$$

where the gauge fields $G_{i\mu}$ $(i = 1, \ldots, 8)$ are called gluons, g_s is the QCD coupling constant and the λ_i are the Gell-Mann matrices (see Sec. 5), the subgroup $SU(3)_c$ of color is promoted to a local symmetry. QCD requires the quark masses to be degenerate in color space. Again, provided we discard quark masses, there is a symmetry flavor symmetry

$$[\mathcal{G}_F, SU(3)_c] = 0$$

where $N_F = 6$. As we increase the number of flavors from $N_F = 2$ to 6, the above symmetry is broken more and more by increasingly heavy quark masses

quark flavor	u	d	s	c	b	t
mass (MeV)	∼5	∼9	190	1650	4750	174200

In practice at a given energy scale only those flavors are effective in a physical process which have masses lower than the given energy. Degrees of freedom with masses heavier than the given energy scale do not influence the which means that the flavor symmetry is preserved the better the lower the number of flavors is which participate in a given physical process.

In the following we will discuss the general condition for anomaly freedom of a theory.

In perturbation theory the axial anomaly shows up in closed fermion loops with an odd number of axial-vector couplings if a non-vanishing γ_5-odd trace of γ-matrices like [32]

$$Tr\ (\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_5) = 4i\epsilon^{\mu\nu\rho\sigma}$$

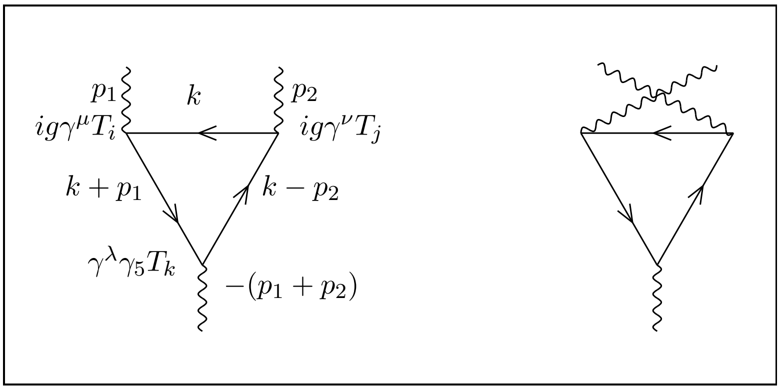


Figure 5. 5: Triangle diagram exhibiting the axial anomaly

is involved and if the corresponding Feynman integral is not ultraviolet convergent such that it requires regularization. The simplest diagram exhibiting the axial anomaly is the triangle diagram (see Fig. 5) which leads to the amplitude (1st diagram)

$$\tilde{T}^{\mu\nu\lambda}_{ijk}(p_1,p_2) = (-1)\, i^5\ Tr\ (T_jT_iT_k)\frac{g^2}{(2\pi)^4}\int d^4k\ Tr\ \left(\frac{1}{\not k - \not p_2 + i\epsilon}\gamma^\nu\frac{1}{\not k + i\epsilon}\gamma^\mu\frac{1}{\not k + \not p_1 + i\epsilon}\gamma^\lambda\gamma_5\right) \tag{12.3}$$

If we include the bose symmetric contribution (2^{nd} diagram)

$$T^{\mu\nu\lambda}_{ijk}(p_1,p_2) = \tilde{T}^{\mu\nu\lambda}_{ijk}(p_1,p_2) + \tilde{T}^{\nu\mu\lambda}_{jik}(p_2,p_1) \tag{12.4}$$

and **impose** vector current conservation

$$p_{1\mu}T^{\mu\nu\lambda}_{ijk}(p_1,p_2) = p_{2\nu}T^{\mu\nu\lambda}_{ijk}(p_1,p_2) = 0 \tag{12.5}$$

we obtain the unambiguous regularization independent result

$$-(p_1+p_2)_\lambda\, T^{\mu\nu\lambda}_{ijk}(p_1,p_2) = i\,\frac{g^2}{16\pi^2}D_{ijk}\ 4\,\epsilon^{\mu\nu\rho\sigma}p_{1\rho}p_{2\sigma} \neq 0 \tag{12.6}$$

with $D_{ijk} = Tr\ (\{T_i,T_j\}\,T_k)$.

[32]Notice that $Tr\ \left(\prod_{i=1}^n \gamma^{\mu_i}\gamma_5\right) = 0$ for $n < 4$ and for all $n =$ odd.

This result is independent on the masses of the fermion lines and is not changed by higher order corrections. Therefore the result is **exact** beyond perturbation theory! *(Adler and Bardeen, 1969).*

We may represent the result as an operator identity in configuration space: $p_{1\rho}(p_{2\sigma})$ corresponds to a derivative $i\partial_\rho(i\partial_\sigma)$ acting on a gauge field (external leg) $V_{i\mu}(x)\,(V_{j\nu}(x))$ while $-(p_1+p_2)_\lambda$ corresponds to the vertex $i\partial_\lambda j_5^\lambda(x)$. Because of the permutation symmetry in the two gauge fields we have to divide by a factor 2. We then obtain:

$$\partial_\lambda j_{5k}^\lambda(x) = -\frac{g^2}{32\pi^2} D_{ijk}\, 4\,\epsilon^{\mu\nu\rho\sigma}\,(\partial_\rho V_{i\mu}(x))\,(\partial_\sigma V_{j\nu}(x))\,.$$

Using the antisymmetry of the ϵ-tensor and renaming summation indices we may rewrite this result

$$\begin{aligned}\partial_\lambda j_{5k}^\lambda(x) &= -\frac{g^2}{32\pi^2} D_{ijk}\,\epsilon^{\mu\nu\rho\sigma}\,(\partial_\rho V_{i\mu}-\partial_\mu V_{i\rho})\,(\partial_\sigma V_{j\nu}-\partial_\nu V_{j\sigma}) \\ &= \frac{g^2}{16\pi^2} D_{ijk}\;\overset{\circ}{\tilde{G}}{}_i^{\mu\nu}\,\overset{\circ}{G}_{j\mu\nu}\end{aligned}$$

with $\overset{\circ}{G}_{j\mu\nu}= \partial_\mu V_{j\nu}-\partial_\nu V_{j\mu}$ and $\overset{\circ}{\tilde{G}}{}_i^{\mu\nu}= \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\,\overset{\circ}{G}_{i\rho\sigma}$. The expression for $-(p_1+p_2)_\lambda\,T_{ijk}^{\mu\nu\lambda}(p_1,p_2)$ is a matrix element of $i\partial_\lambda j_5^\lambda(0)$. If terms from other diagrams contributing to other possible matrix-elements of the axial current are included one finds for the final form of the anomaly

$$\partial_\lambda j_{5k}^\lambda(x) = \frac{g^2}{16\pi^2} D_{ijk}\,\tilde{G}_i^{\mu\nu}(x) G_{j\mu\nu}(x) \tag{12.7}$$

where $G_{i\mu\nu}(x)$ is the non-Abelian field strength tensor and $\tilde{G}_i^{\mu\nu}$ its dual pseudotensor. As a result the condition for the absence of an anomaly reads

$$D_{ijk} = Tr\ (\{T_i,T_j\}\,T_k) = 0\ \ \forall\ \ (ijk)\ . \tag{12.8}$$

The matrices T_i are the generators of a gauge group in a representation R under which the fermions transform. What are the general conditions for the absence of anomalies? To answer this question we need the following basic properties of traces:

i) trace of the transpose A^T of a matrix A:

$$Tr\ (A) = \sum_i A_{ii} = Tr\ (A^T)$$

ii) trace of the equivalent $A' = SAS^{-1}$ of a matrix A

$$Tr\ (SAS^{-1}) = Tr\ (S^{-1}SA) = Tr\ (A)$$

since $Tr\ (AB) = Tr\ (BA)$. S must be nonsingular.

We have to distinguish two different types of representations, real and complex ones. A representation R is called **real** if its **complex conjugate** representation R^* is equivalent to $R : R^* \simeq R$ i.e. $R^* = SRS^{-1}$ for some fixed non-singular matrix S. Non-real representation $R \not\simeq R^*$ are called **complex**. The following general statements hold:

A) Real representations are anomaly free: $D_{ijk} \equiv 0$. Since $R^* \simeq R$ we have $D_{ijk}(R^*) = D_{ijk}(R)$. On the other hand in the canonical basis $T_i^+ = T_i$, if T_i are the generator of R then $-T_i^* = -T_i^T$ are the generators of R^*. Therefore

$$D_{ijk}(R^*) = -D_{ijk}(R).$$

Consequently $D_{ijk}(R) \equiv 0$ for real representation. The groups $SO(2\ell+1), \ell > 1, Sp(2\ell), G_2, F_4, E_7$ and E_8 have only real representations and hence are anomaly free. The groups $SO(2\ell), \ell > 1$, except $SO(6) \simeq SU(4)$ are also anomaly free.

B) Groups for which the fundamental representation is real are anomaly free.

One can show that for any representation R

$$D_{ijk}(R) = D_{ijk}(R_0) \cdot K(R)$$

where R_0 is the fundamental representation. Only the invariant quantity $K(R)$ depends on the representation, $K(R_0) = 1$.

For the groups $SU(2) \simeq SO(3)$ and E_6 one has $D_{ijk}(R_0) = 0$; and thus all representations of these groups are anomaly free.

C) The groups $SU(n), n \geq 3$ have complex representations and $D_{ijk}(R_0) \neq 0$. Fermions transforming under representations of these groups in general lead to anomalies. In order to avoid anomalies in this case one has to find those representations R for which $K(R) = 0$ i.e. for the $SU(n)$, $n \geq 3$, groups one can avoid anomalies only by putting the fermions into particular representations. One can easily find these representations. We always can write the fermion currents j_i^μ which couple to the gauge fields in terms of left-handed and right-handed fields

$$\begin{aligned} j_i^\mu &= \bar{\psi}_L \gamma^\mu T_{Li} \psi_L + \bar{\psi}_R \gamma^\mu T_{Ri} \psi_R \\ &= \bar{\psi} \gamma^\mu \frac{1-\gamma_5}{2} T_{Li} \psi + \bar{\psi} \gamma^\mu \frac{1+\gamma_5}{2} T_{Ri} \psi. \end{aligned}$$

Since $\bar{\psi}_L \gamma_\mu \psi_R = \bar{\psi}_R \gamma_\mu \psi_L = \bar{\psi}_L \gamma_\mu \gamma_5 \psi_R = \bar{\psi}_R \gamma_\mu \gamma_5 \psi_L = 0$ the contributions from the left-handed and right-handed fields in closed fermion loops decouple. The contribution to the anomaly thus is given by the sum of the contributions from left-handed fields and from right-handed fields. If a particular left-handed loop gives an anomaly proportional to $D_{ijk}(R_L)$ then the right-handed loop gives an anomaly contribution proportional to $-D_{ijk}(R_R)$ because the two contributions differ by a sign at the $\gamma^\mu \gamma_5$ vertex as $\gamma_5 \frac{1\pm\gamma_5}{2} = \pm \frac{1\pm\gamma_5}{2}$. Thus the condition that no anomalies arise from gauge interactions is that for all i, j, k :

$$Tr\ (\{T_{Li}, T_{Lj}\} T_{Lk}) - Tr\ (\{T_{Ri}, T_{Rj}\} T_{Rk}) = 0\ . \tag{12.9}$$

For $SU(2)$ the terms are individually zero for any representation. For $SU(3)$ there is no anomaly if the ψ_L and ψ_R transform under the same representation. This is the case for QCD where due to $T_{Ri} = T_{Li}$ no axial currents couple to the gluons. Since the left-handed antiquarks are in the complex conjugate representation 3^* of the fundamental representation 3 the corresponding right-handed particle fields also transform under 3: This follows, because under antiparticle conjugation

$$\psi_L \xrightarrow{C} \psi_L^c = i\gamma^2 \psi_R^*\ .$$

D) The Abelian group $U(1) \simeq SO(2)$ are not anomaly save. The previous argument for $SU(n)$ groups for the $U(1)$ group leads to

$$D = Tr\, T_{L0}^3 - Tr\, T_{R0}^3 = 0$$

as a condition for anomaly cancellation. Here T_0 denotes the Abelian generator of hyper charge.

For non-simple groups the generators of the factor groups commute with each other. Therefore, the anomaly is given by the sum of the anomalies of the individual subgroups. This means that anomalies must be absent for each factor group.

For the standard model only the $U(1)_Y$ yields a non trivial condition which must be satisfied in order to have an anomaly free theory! Let us check this condition now.

The matter field are in

$$\text{doublets} \quad \left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array} \right)_L \quad \text{and singlets} \quad \psi_{R1}, \psi_{R2} \; .$$

The hypercharge and charge assignments satisfy:

$$Y_i = 2(Q_i - T_{3i}) \; ; \;\; Q_{Ri} = Q_{Li} = Q_i \; ; \;\; Q_1 - Q_2 = 1$$

such that we have

$$\begin{aligned} Y_{L1} &= 2(Q_1 - \frac{1}{2}) = 2Q_1 - 1 \; , \;\; Y_{R1} = 2Q_1 \\ Y_{L2} &= 2(Q_2 + \frac{1}{2}) = 2Q_1 - 1 \; , \;\; Y_{R2} = 2Q_2 = 2Q_1 - 2 \;\; . \end{aligned}$$

Thus the anomaly contribution per doublet is given by

$$D = Y_{L1}^3 + Y_{L2}^3 - Y_{R1}^3 - Y_{R2}^3 = 2\,(2Q_1 - 1)^3 - (2Q_1)^3 - (2Q_1 - 2)^3 = -12\, Q_1 + 6 \;\; .$$

Consequently, the $SU(2)_L \otimes U(1)_Y$ electroweak model with leptons only is **not** renormalizable (not anomaly free). The anomaly must be canceled by a contribution of opposite size coming from the quarks!

For the anomaly factor D we get for a lepton doublet and $N_c = 3$ colored quark doublets and the associated singlets

$$D_{\text{leptons}} + D_{\text{quarks}} = 6 - 6N_c\,(2Q_1^q - 1) = 0$$

and since

$$Q_1^q = \frac{1}{2}\left(1 + \frac{1}{N_c}\right) = 2/3 \quad \text{for} \quad N_c = 3 \;\; .$$

Notice that one could satisfy the anomaly condition also with the nucleon doublet: $N_c = 1 \; , \;\; Q_1 = 1$ i.e.

$$\left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array} \right) = \left(\begin{array}{c} p \\ n \end{array} \right) .$$

This is an example of 't Hooft's anomaly condition: A composite particle must reproduce the axial anomaly of its fermionic constituents. This may be understood as a consequence of the Adler-Bardeen theorem which states that the anomaly is not renormalized by higher order effects and hence that the axial anomaly is a known non-perturbation effect.

As a result we find: The electroweak standard model is renormalizable only if fermions are grouped into lepton-quark families.

Exercises: Section 12

① In the SM the leptonic hypercharge current has the form

$$j_Y^\mu = \frac{1}{2}\left(\bar{\ell}_L \gamma^\mu \ell_L + \bar{\nu}_{\ell L} \gamma^\mu \nu_{\ell L}\right) + \bar{\ell}_R \gamma^\mu \ell_R \ .$$

Write it in terms of vector plus axialvector contributions.

② Calculate the trace and the Feynman integral (12.3) and verify (12.6).

The closer you look the more there is to see.

13 The Standard Model of fundamental interactions

According to present day knowledge, the known "fundamental" interactions of "elementary" particles follow from a local gauge principle (*Weyl, Yang-Mills*) with the gauge group

$$G^{\mathrm{SM}}_{\mathrm{local}} = SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \tag{13.1}$$

which is broken through a Higgs mechanism to $SU(3)_c \otimes U(1)_{\mathrm{em}}$. $SU(3)_c$ is the color gauge group which determines the interaction of the color triplets of quarks via an octet of colored gluons. This unbroken color gauge theory is called quantum chromodynamics (QCD) (Fritzsch, Gell-Mann, Leutwyler 1973) and determines strong interaction physics, and in particular, the hadron spectrum and the residual strong interaction between hadrons. Quarks originally emerged as the building blocks of hadrons in the attempt to classify the hadronic states according to their flavor (Gell-Mann, Ne'eman, Zweig 1964). The quark model hypothesis required the hadrons to be composite states of quarks with the quarks being permanently confined in the hadrons. The confinement hypothesis declares **hadrons** to be **composite elementary particles**. More specifically, baryons are three quark states, mesons are quark-antiquark states

Baryons:, $p, n, \Lambda, \Sigma, \Delta, \ldots$

$$\begin{array}{lllllll} B = (q_1 q_2 q_3) : \underline{s = 1/2} : & p: & uud & \underline{s = 3/2} : & \Delta^{++} : & uuu \\ & n: & udd & & \vdots & \vdots \\ & & & & \Delta^- & : & ddd \end{array}$$

Mesons: $\pi, K, \eta, \rho, \omega, \ldots$

$$\begin{array}{llllll} M = (q_1 \bar{q}_2) : \underline{s = 0} : & \pi^+ : & u\bar{d} & \underline{s = 1} : & \rho^+ : & u\bar{d} \\ & \pi^0 : & \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}) & & \rho^0 : & \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}) \\ & \pi^- : & d\bar{u} & & \rho^- : & d\bar{u} \\ & \eta : & \frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d}) & & \omega : & \frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d}) \end{array}$$

Quarks must then be fractionally charged: $Q_u = 2/3, Q_d = -1/3$. A crucial problem showed up for the states which are totally symmetric under permutations of spin and flavor (e.g. Δ^{++} : $u(\uparrow)u(\uparrow)u(\uparrow)$). The spin-statistic theorem requires both the quarks ($s = 1/2$) and the baryons ($s = 1/2, 3/2, \ldots$) to satisfy the Pauli principle. This requires a totally antisymmetric quark wave function for the quarks in the baryon. This **spin-statistics crisis** could be solved only by assigning a new quantum number, called **color**, to quarks. Each quark must exist in three copies, the red(r), green(g), and blue(b) quarks. Since color never has been observed it was natural to require $SU(3)_c$ color symmetry to be a local symmetry, which implies that colors are indistinguishable. The confinement hypothesis now requires that **physical states** must be **color singlets**. The singlet condition leads in a very natural way to the baryons

$$(q_1 q_2 q_3)_{\text{color singlet}} = \frac{1}{\sqrt{3!}} \varepsilon_{c_1 c_2 c_3} q_{1c_1} q_{2c_2} q_{3c_3}$$

and mesons

$$(q_1\bar{q}_2)_{\text{color singlet}} = \frac{1}{\sqrt{3}}\delta_{c_1c_2}q_{1c_1}\bar{q}_{2c_2}$$

The wave functions of hadronic states are given now by a product of spatial, color, flavor and spin wave functions. The ground state hadrons have no orbital angular momentum, such that the spatial wave function is symmetric under permutations of the constituents. Simultaneous eigenstates of color (singlet), flavor and spin are then easily constructed by considering factors. The calculation of the hadrons as bound states of quarks (i.e. the actual calculation of the spatial wave function) is an unsolved problem. After this digression to hadronic physics we come back to the other part of of $G^{\text{SM}}_{\text{local}}$.

The other unbroken subgroup $U(1)_{\text{em}}$ of $SU(2)_L \otimes U(1)_Y$ defines quantum electrodynamics (QED), describing the interaction of all charged particles with the photon. $SU(2)_L$ is the **weak isospin** gauge group which determines the interaction of the left–handed $(V-A)$ fermion currents, known from weak interaction processes, with a triplet of weak gauge bosons also called intermediate vector bosons. Finally, the $U(1)_Y$ subgroup is needed in order to recover $U(1)_{\text{em}}$ with $Q = T_3 + \frac{Y}{2}$ in the broken phase as we shall see below.

The standard model (SM) is determined essentially by specifying the matter fields and their transformation properties under local gauge transformations. A way to "understand" the emergence of the SM is described in the following.

13.1 The matter fields

The "real world" is build from **massless** spin 1/2 particles the quarks and leptons. Spin 1/2 particles in a sense are more fundamental than other particles because they allow to compose particles of any spin (e.g. $1/2 \otimes 1/2 = 0 \oplus 1$ etc.). Massless Fermi fields have fixed handedness called chirality or helicity. If ψ is a massless Dirac field the left–handed field $\psi_L = \frac{1-\gamma_5}{2}\psi$ and the right–handed field $\psi_R = \frac{1+\gamma_5}{2}\psi$ do not mix under Lorentz-transformations, rotations and space-time translation.

$$\underset{\psi_L}{\overset{\vec{s}}{\Longleftarrow}\!\circ\!\underset{\vec{p}}{\longrightarrow}} \qquad \overset{P}{\longleftrightarrow} \qquad \underset{\psi_R}{\circ\!\overset{\vec{s}}{\underset{\vec{p}}{\Rightarrow}}}$$

The fields ψ_L and ψ_R are interchanged under parity transformations. Since ψ is a local field satisfying Einstein causality the left–handed Dirac field of a massless electron denoted by e^-_L describes a left–handed electron and, simultaneously, a right–handed positron. Similarly, if e^-_R is the local field describing a right–handed electron, this field also describes a left–handed positron. Thus $e^-_R \equiv e^+_L$. Therefore, we may consider all massless fields to be left–handed.

According to todays knowledge, matter is made out of colored quarks and leptons which are grouped into three families. The not too long ago observed neutrino oscillations require the neutrinos to have a tiny mass which must be different for the different flavors. This requires the existence of right–handed neutrinos $\nu_{\ell R} \equiv \bar{\nu}_{\ell L}$ in spite of the fact that they do not couple directly to gauge fields (i.e., they are singlets with respect to the SM gauge group).

As we have argued earlier quarks of different flavors must show up in three replica of red(r), green(g), and blue(b) color. The first family of fermions are

$$\nu_{eL}, \bar{\nu}_{eL}, e^-_L, e^+_L, u_{Lr}, u_{Lg}, u_{Lb}, u^c_{Lr}, u^c_{Lg}, u^c_{Lb}, d_{Lr}, d_{Lg}, d_{Lb}, d^c_{Lr}, d^c_{Lg}, d^c_{Lb}$$

where u^c denotes the antiparticle of the u-quarks and so on. All **stable matter** is built from these first family quarks and leptons. The two additional families (who has ordered them?) we obtain by replacing (ν_e, e) with (ν_μ, μ) and (u, d) with (c, s) and (ν_e, e) with (ν_τ, τ) and (u, d) with (t, b) . They supplement Nature by forms of unstable matter and allow for the phenomena of flavor mixing and *CP*-violation. Altogether we know 3×16 degrees of freedom if we include the top–quark t for which we only have indirect evidence. If there would be no interactions the (free) Lagrangian of the world would be

$$\mathcal{L}_{\text{matter}} = \sum_a \bar{\psi}_{La} i\gamma^\mu \partial_\mu \psi_{La}$$

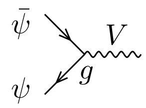
which exhibits a huge symmetry of global $U(48)$. In nature a subgroup of this large global symmetry group turns out to be a local symmetry

$$\psi_{La} \to U(x)_{ab}\psi_{Lb} \; ; \;\; U(x) \in G_{\text{local}}$$

where $G_{\text{local}} = SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. This requires the fields to couple to massless spin 1 gauge fields via minimal coupling

$$\partial_\mu \psi_L \to D_\mu \psi_L = \left(\partial_\mu - i \sum_\alpha g_\alpha T_{\alpha i} V_{\mu\alpha i} \right) \psi_L$$

with interaction vertices:



Here, α labels the factors $SU(3)_c, SU(2)_L$ and $U(1)_Y$ to which the Yang-Mills construction applies individually. The emergence of G_{local} as a direct product of two simple Lie groups and the Abelian group $U(1)_Y$ is closely related to the phenomenological appearance of strong, weak and electromagnetic interactions as separate phenomena differing in strength and symmetry. The $SU(3)_c$ which determines strong interactions distinguishes between triplets of quarks, "antitriplets" of antiquarks and singlets of leptons and neutrinos. Since the latter do not carry color they do not participate in strong interactions i.e. they do not talk to quarks and gluons, the gauge quanta of $SU(3)_c$. The $SU(2)_L$ distinguishes between doublets of left–handed particles $\binom{\nu_e}{e^-}_L, \ldots$ and singlets of left–handed antiparticles which are usually identified by the right–handed particles $e^+_L \equiv e^-_R, \ldots$. This undemocratic treatment of particles and antiparticles is what we know as maximal parity violation of weak interactions. The Abelian $U(1)_Y$ only affects the phases of the fields according to the **weak hypercharge** assignment $Y = 2(Q - T_3)$.

Notice that a fermion mass term

$$\bar{\psi}\psi = \bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L$$

cannot be $SU(2)_L \otimes U(1)_Y$ invariant. It is the parity violating nature of weak interactions which forbid fermion masses. We summarize the local multiplet structure in the following tables:

"Weak quantum numbers": $Q = T_3 + \frac{1}{2}Y$

Doublets:

	$(\nu_\ell)_L$	$(\ell^-)_L$	$(u,c,t)_L$	$(\tilde{d},\tilde{s},\tilde{b})_L$
Q	0	- 1	$\frac{2}{3}$	- $\frac{1}{3}$
T_3	$\frac{1}{2}$	- $\frac{1}{2}$	$\frac{1}{2}$	- $\frac{1}{2}$
Y	- 1	- 1	$\frac{1}{3}$	$\frac{1}{3}$

Singlets:

	$(\nu_\ell)_R$	$(\ell^-)_R$	$(u,c,t)_R$	$(d,s,b)_R$
Q	0	- 1	$\frac{2}{3}$	- $\frac{1}{3}$
T_3	0	0	0	0
Y	0	- 2	$\frac{4}{3}$	- $\frac{2}{3}$

`group`	`representation`		
$SU(3)_c$	3	$\begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix}$	quark color triplets
	3^*	$\begin{pmatrix} q^c_{\bar{r}} \\ q^c_{\bar{g}} \\ q^c_{\bar{b}} \end{pmatrix}$	antiquark color triplets
	1		leptons, neutrinos
$SU(2)_L$	$2 = 2^*$	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} u \\ \tilde{d} \end{pmatrix}_L$ $\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} c \\ \tilde{s} \end{pmatrix}_L$ $\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} c \\ \tilde{b} \end{pmatrix}_L$	left–handed weak isospin doublets of leptons and quarks (flavor doublets)
	1	$\nu_{eR}, e^-_R, u_R, d_R$ $\nu_{\mu R}, \mu^-_R, c_R, s_R$ $\nu_{\tau R}, \tau^-_R, t_R, b_R$	right–handed weak isospin singlets of leptons and quarks
$U(1)_Y$	phase transformations	weak hypercharge	$Y = 2(Q - T_3)$

Notice:

- $\nu_{eR}, \nu_{\mu R}$ and $\nu_{\tau R}$ all have zero quantum numbers with respect to G_{local} (i.e. no couplings to gauge fields).
- the bottom components of the quark doublets are Cabibbo-Kobayashi-Maskawa rotated fields $\tilde{d}, \tilde{s}, \tilde{b}$ (see below).

In the following we denote the weak doublets by L_ℓ ($\ell = e,\ \mu,\ \tau$) for the leptons and by L_q for the quarks.

13.2 The gauge fields

For each factor group of G_{local}, local gauge invariance requires the existence of a set of massless gauge fields in the adjoint representation, which couple minimally to the matter fields. We denote the gauge fields $V_{\mu\alpha i}(x)$ and the gauge couplings as follows:

group		**gauge fields**	**name**	**coupling**
$SU(3)_c$	:	$G_{\mu i}\ ;\ \ i = 1, \ldots, 8$	gluons	g_s
$SU(2)_L$	:	$W_{\mu a}\ ;\ \ a = 1, 2, 3$	weak gauge bosons	g
$U(1)_Y$	:	B_μ		g'

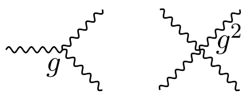
The Yang-Mills Lagrangian is given by

$$\mathcal{L}_{\text{YM}} = -\frac{1}{4} G_{\mu\nu i} G_i^{\mu\nu} - \frac{1}{4} W_{\mu\nu a} W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

with field strength tensors

$$\begin{aligned}
SU(3)_c \ &: \ G_{\mu\nu i} = \partial_\mu G_{\nu i} - \partial_\nu G_{\mu i} + g_s f_{ikl} G_{\mu k} G_{\nu l} \\
SU(2)_L \ &: \ W_{\mu\nu a} = \partial_\mu W_{\nu a} - \partial_\nu W_{\mu a} + g \varepsilon_{abc} W_{\mu b} W_{\nu c} \\
U(1)_Y \ &: \ B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \ .
\end{aligned}$$

The non-abelian fields must exhibit self–interactions described by the vertices



and with coupling strength equal the matter field coupling

$$\mathcal{L}_{\text{matter, int}} = g \bar{\psi}_L \gamma^\mu T_a \psi_L W_{\mu a}$$

where $J_a^\mu = \bar{\psi}_L \gamma^\mu T_a \psi_L$ are the fermion currents associated with the gauge group.

Since the QCD part, the unbroken $SU(3)_c$, represents the final answer and the relevant Feynman rules have been given in Sec. 8, we will omit the strong interactions and concentrate to the discussion of the electroweak SM $SU(2)_L \otimes U(1)_Y$ in the following.

The non–abelian gauge field interaction terms for the electroweak gauge group $SU(2)_L \otimes U(1)_Y$ have been experimentally confirmed by LEP experiments [?].

13.3 The electroweak gauge bosons, $\gamma - Z$–mixing

In the real world $SU(2)_L \otimes U(1)_Y$ is broken to $U(1)_{\rm em}$ by the masses of the physical particles. In particular, the weak gauge bosons which mediate weak interactions must be massive as the weak interactions are short ranged. Also the weak equivalence principle which claims the left–handed neutrino to be indistinguishable from the left–handed electron is obviously not true in the real world. In 1961 Glashow proposed a massive vector boson gauge theory by simply adding the symmetry breaking mass terms:

$$\mathcal{L}_{\rm broken} = \mathcal{L}_{\rm inv} + \mathcal{L}_{\rm mass} \;.$$

Before we can add mass terms we have to know which **physical fields** we have. They have to be mass eigenstates and must carry electric charge $Q = T_3 + \frac{Y}{2}$. Q is the only good quantum number left after the breaking. The fields $W_{\mu a}$ have $Y = 0$ and hence the physical fields must be eigenstates of $T_3 = \frac{\tau_3}{2}$. Using the $SU(2)$ algebra $[T_i, T_k] = i\varepsilon_{ikl} T_l$, $T_i = T_i^+$ one finds that the ladder operators

$$T_\pm = T_1 \mp i T_2 \;,\;\; T_+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \;,\;\; T_- = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

satisfy

$$[T_3, T_\pm] = \pm T_\pm$$

therefore the fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_{\mu 1} \mp i W_{\mu 2} \right)$$

carry charge ± 1 . These are the charged W bosons. The "charged" part of the matter field Lagrangian reads for the leptons

$$\begin{aligned}
\mathcal{L}^{\ell,\text{charged}}_{\text{matter, int}} &= \frac{g}{2} \left(\bar{\nu}_\ell, \bar{\ell} \right)_L \gamma^\mu \left(\tau_1 W_{\mu 1} + \tau_2 W_{\mu 2} \right) \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \\
&= \frac{g}{2} \left(\bar{\nu}_\ell, \bar{\ell} \right)_L \gamma^\mu \left(T_+ + T_- \right) W_{\mu 1} + i \left(T_+ - T_- \right) W_{\mu 2} \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \\
&= \frac{g}{\sqrt{2}} \left(\bar{\nu}_\ell, \bar{\ell} \right)_L \gamma^\mu \left(T_+ \frac{W_{\mu 1} + i W_{\mu 2}}{\sqrt{2}} + T_- \frac{W_{\mu 1} - i W_{\mu 2}}{\sqrt{2}} \right) \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \\
&= \frac{g}{\sqrt{2}} \left(J_\ell^{\mu +} W_\mu^- + J_\ell^{\mu -} W_\mu^+ \right)
\end{aligned}$$

where ($J^{\mu\pm} = J_1^\mu \mp i\, J_2^\mu$)

$$J_\ell^{\mu +} = \bar{\ell}_L \gamma^\mu \nu_{\ell L} \quad \text{and} \quad J_\ell^{\mu -} = \left(J_\ell^{\mu +} \right)^+ = \bar{\nu}_{\ell L} \gamma^\mu \ell_L$$

is the leptonic contribution to the **weak charge changing current**.

The fields $W_{\mu 3}$ and B_μ both have $Y = 0$ and $T_3 = 0$ and hence $Q = 0$. Since they have the same physical (unbroken) quantum numbers they can mix. To find out the physical linear combinations, we have to inspect their interactions with the matter fields. For the leptons we have

$$\begin{aligned}
\mathcal{L}^{\ell,\text{neutral}}_{\text{matter}} &= \bar{L}_\ell i\gamma^\mu \left(\partial_\mu - igT_3W_{\mu 3} + i\frac{g'}{2}B_\mu\right) L_\ell \\
&\quad + \bar{\ell}_R i\gamma^\mu \left(\partial_\mu + ig'B_\mu\right)\ell_R + \bar{\nu}_{\ell R} i\gamma^\mu\partial_\mu\nu_{\ell R} \\
&= \bar{\ell} i\gamma^\mu\partial_\mu\ell + \bar{\nu}_\ell i\gamma^\mu\partial_\mu\nu_\ell \\
&- \frac{g}{2}\left(\bar{\ell}_L\gamma^\mu\ell_L - \bar{\nu}_{\ell L}\gamma^\mu\nu_{\ell L}\right)W_{3\mu} \\
&- \frac{g'}{2}\left(\bar{\ell}_L\gamma^\mu\ell_L + \bar{\nu}_{\ell L}\gamma^\mu\nu_{\ell L}\right)B_\mu \\
&- g'\left(\bar{\ell}_R\gamma^\mu\ell_R\right)B_\mu
\end{aligned}$$

The electromagnetic lepton current must be

$$j^{\mu\ell}_{\text{em}} = -\left(\bar{\ell}\gamma^\mu\ell\right) = -\left(\bar{\ell}_L\gamma^\mu\ell_L + \bar{\ell}_R\gamma^\mu\ell_R\right)$$

without a neutrino contribution. The field Z_μ which couples to the neutrino is

$$Z^\mu \propto gW_{\mu 3} - g'B_\mu$$

and the photon must be orthogonal to Z^μ

$$A^\mu \propto gB_\mu + g'W_{\mu 3}.$$

The transformation

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \frac{1}{\sqrt{g^2 + g'^2}}\begin{pmatrix} g & -g' \\ g' & g \end{pmatrix}\begin{pmatrix} W_{\mu 3} \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} W_{\mu 3} \\ B_\mu \end{pmatrix}$$

with

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}\ , \quad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}\ ; \quad \cos^2\theta_W + \sin^2\theta_W = 1$$

is (must be) orthogonal such that the real kinetic terms (bilinear parts) of the gauge fields satisfy

$$\overset{\circ}{G}_{\mu\nu 3}\overset{\circ}{G}{}^{\mu\nu}{}_3 + B_{\mu\nu}B^{\mu\nu} = \overset{\circ}{Z}_{\mu\nu}\overset{\circ}{Z}{}^{\mu\nu} + F_{\mu\nu}F^{\mu\nu}$$

with

$$\begin{aligned}
\overset{\circ}{Z}_{\mu\nu} &= \partial_\mu Z_\nu - \partial_\nu Z_\mu \\
F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu\ .
\end{aligned}$$

If we replace now

$$\begin{aligned}
W_{3\mu} &= \cos\theta_W Z_\mu + \sin\theta_W A_\mu \\
B_\mu &= -\sin\theta_W Z_\mu + \cos\theta_W A_\mu
\end{aligned}$$

in the Lagrangian we obtain

$$\mathcal{L}^{\ell,\text{neutral}}_{\text{matter, int}} = \frac{g}{\cos\theta_W} J^\mu_Z Z_\mu + e j^\mu_{\text{em}} A_\mu$$

with weak neutral current

$$J^\mu_Z = J^\mu_3 - \sin^2\theta_W j^\mu_{\text{em}} \quad .$$

The relation $e = g\sin\theta_W$ must hold. This is the **unification condition**. Above we have used relations like

$$\begin{aligned} e = g\sin\theta_W = g'\cos\theta_W \\ g\cos\theta_W + g'\sin\theta_W = \sqrt{g^2+g'^2} = \tfrac{g}{\cos\theta_W} \\ g'\sin\theta_W = \tfrac{g'^2}{\sqrt{g^2+g'^2}} = \sin^2\theta_W\sqrt{g^2+g'^2} = \sin^2\theta_W \tfrac{g}{\cos\theta_W} \quad . \end{aligned}$$

The mixing angle θ_W is called weak mixing angle and $\sin^2\theta_W$ **weak mixing parameter** (*Glashow, 1961*). Notice that in the limit $g' \to 0$ i.e. no $U(1)_Y$ interaction $e \to 0$ and we would not get the electromagnetic interactions. This demonstrates the need for the Abelian subgroup $U(1)_Y$. If we include the quarks

$$\begin{aligned} \mathcal{L}_q &= \bar{q}_{uR} i\gamma^\mu \left(\partial_\mu - i\frac{4}{3}\frac{g'}{2}B_\mu\right) q_{uR} + \bar{q}_{dR} i\gamma^\mu \left(\partial_\mu + i\frac{2}{3}\frac{g'}{2}B_\mu\right) q_{dR} \\ &\quad + \bar{L}_q i\gamma^\mu \left(\partial_\mu - i\frac{1}{3}\frac{g'}{2}B_\mu - ig\frac{\tau_a}{2}W_{\mu a}\right) L_q \end{aligned}$$

we obtain the $SU(2)_L \otimes U(1)_Y$ fermion currents:

$$\textbf{i)} \qquad J^+_\mu = \bar{\nu}_\ell \gamma_\mu \frac{1-\gamma_5}{2}\ell^- + \bar{q}_u \gamma_\mu \frac{(1-\gamma_5)}{2} U_{\text{CKM}} q_d$$

is the charged current (CC) which is of pure $V-A$ type and exhibits Cabibbo-Kobayashi-Maskawa **flavor mixing** :

$$q_u = \begin{pmatrix} u \\ c \\ t \end{pmatrix} ; \;\; q_d = \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

are "horizontal" family vectors and

$$U_{\text{CKM}} = \begin{pmatrix} V_{ud}V_{us}V_{ub} \\ V_{cd}V_{cs}V_{cb} \\ V_{td}V_{ts}T_{tb} \end{pmatrix}$$

is a unitary 3×3 mixing matrix.

Strict lepton number conservation (still true within exerimental limits) is only possible if the neutrinos are strictly massless. Non-vanishing neutrino masses lead to **neutrino-oscillations**. Neutrino mixing searches (ν-oscillations $\nu_\ell \leftrightarrow \nu_{\ell'}$) have confirmed the effect recently, which implies the existence of non–vanishing neutrino masses. Present direct upper limits on the neutrino masses are:

$$\begin{aligned} m_{\nu_e} &< \quad 3.0 \text{ eV} \quad (\text{from } {}^3H \to {}^3He\, e^- \, \bar{\nu}_e) \\ m_{\nu_\mu} &< \quad 190 \text{ keV} \quad (\text{from } \pi \to \mu \, \nu_\mu) \\ m_{\nu_\tau} &< \quad 18.2 \text{ MeV} \quad (\text{from } \tau^- \to 3\pi \, \nu_\tau) \end{aligned}$$

Lower bounds are not yet so easy to establish at present but observed neutrino mixing phenomena indicate values of about two to three orders of magnitude lower than the above direct upper limits. In any case this implies corresponding lepton numbers L_ℓ ($\ell = e, \mu, \tau$)–violations.

ii) $$J_{\mu Z} = \sum_f \bar{\psi}_f \gamma_\mu \left(v_f - a_f \gamma_5\right) \psi_f \equiv J_{3\mu} - \sin^2\theta_W j_{\mu\,\mathrm{em}}$$

is the weak neutral current (NC). It is strictly **flavor conserving** (GIM mechanism). The vector and axial vector couplings are:

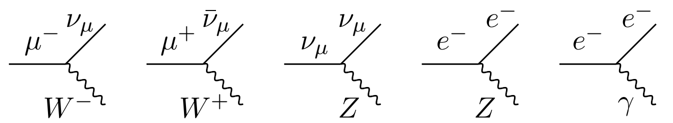
$$v_f = -Q_f \sin^2\theta_W + \frac{T_{3f}}{2} \; ; \; a_f = \frac{T_{3f}}{2} \; .$$

iii) $$j_{\mu\,\mathrm{em}} = \sum_f Q_f \bar{\psi}_f \gamma_\mu \psi_f$$

is the electromagnetic current.

In terms of the fermion currents we get

$$\begin{aligned} \mathcal{L}_{\mathrm{matter}} &= \sum_f \bar{\psi}_f i \gamma^\mu \partial_\mu \psi_f \\ &\quad + \frac{g}{\sqrt{2}} \left(J_\mu^+ W^{\mu -} + J_\mu^- W^{\mu +} \right) + \frac{g}{\cos\theta_W} J_{\mu Z} Z^\mu + e j_{\mu\,\mathrm{em}} A^\mu \end{aligned}$$



as a final form for the electroweak interactions of matter with the physical gauge fields. Some typical vertices in the leptonic sector are shown.

Now we also have to rewrite the Yang-Mills term

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4} \left(\partial_\mu B_\nu - \partial_\nu B_\mu\right)^2 - \frac{1}{4} \left(\partial_\mu W_{\nu a} - \partial_\nu W_{\mu a} + g \varepsilon_{abc} W_{\mu b} W_{\nu c}\right)^2$$

in terms of the physical fields $W_\mu^\pm, Z_\mu$ and A_μ .

For the charged fields we use the relations

$$\begin{aligned} W_{\mu 1} = \frac{1}{\sqrt{2}} \left(W_\mu^+ + W_\mu^-\right) \quad &, \quad W_{\mu 2} = \frac{i}{\sqrt{2}} \left(W_\mu^+ - W_\mu^-\right) \\ W_{\mu 1} W_{\nu 1} + W_{\mu 2} W_{\nu 2} &= W_\mu^+ W_\nu^- + W_\nu^+ W_\mu^- \\ W_{\mu 1} W_{\nu 2} - W_{\nu 1} W_{\mu 2} &= -i \left(W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-\right) \\ W_{\mu 1} W_1^\mu + W_{\mu 2} W_2^\mu &= 2 W_\mu^+ W^{\mu -} \end{aligned}$$

for the neutral fields

$$\begin{aligned}
B_\mu &= -\sin\theta_W Z_\mu + \cos\theta_W A_\mu \\
W_{\mu 3} &= \quad \cos\theta_W Z_\mu + \sin\theta_W A_\mu \\
W_{\mu 3} W_{\nu 3} &= \cos^2\theta_W Z_\mu Z_\nu + \cos\theta_W \sin\theta_W \left(Z_\mu A_\nu + Z_\nu A_\mu\right) + \sin^2\theta_W A_\mu A_\nu \\
W_{\mu 3} W_3^\mu + B_\mu B^\mu &= Z_\mu Z^\mu + A_\mu A^\mu \quad .
\end{aligned}$$

For the free (bilinear) part we obtain

$$\begin{aligned}
\mathcal{L}^{(0)}_{\text{YM}} &= -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} \overset{\circ}{W}_{\mu\nu a} \overset{\circ}{W}{}^{\mu\nu}{}_a \\
&= -\frac{1}{2}\left(\partial_\mu W^+_\nu - \partial_\nu W^+_\mu\right)\left(\partial^\mu W^{\nu -} - \partial^\nu W^{\mu -}\right) \\
&\quad -\frac{1}{4}\left(\partial_\mu Z_\nu - \partial_\nu Z_\mu\right)^2 - \frac{1}{4}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)^2 \quad .
\end{aligned}$$

As it should be the charged fields only appear in the neutral bilinear combination $W^+_\mu W^-_\nu$.

For the interaction part

$$\mathcal{L}_{\text{YM, int}} = -\frac{g}{2}\varepsilon_{abc} \overset{\circ}{W}_{\mu\nu a} W^\mu_b W^\nu_c - \frac{g^2}{4}\varepsilon_{abc}\varepsilon_{ab'c'} W_{\mu b} W_{\nu c} W^\mu_{b'} W^\nu_{c'}$$

the transformation is slightly more complicated. We first consider the trilinear term, which in terms of the charged fields reads

$$\begin{aligned}
\mathcal{L}^{\text{triple}}_{\text{YM}} = i\tfrac{g}{2} \Big[& \overset{\circ}{W}_{\mu\nu 3} \left(W^{\mu +} W^{\nu -} - W^{\mu -} W^{\nu +}\right) \\
& + \overset{\circ}{W}{}^+{}_{\mu\nu} \left(W^{\mu -} W^\nu_3 - W^\mu_3 W^{\nu -}\right) \\
& - \overset{\circ}{W}{}^-{}_{\mu\nu} \left(W^{\mu +} W^\nu_3 - W^\mu_3 W^{\nu +}\right)\Big]
\end{aligned}$$

Since the antisymmetric rotations $\overset{\circ}{W}{}^{\cdot}{}_{\mu\nu} = \partial_\mu W^{\cdot}_\nu - \partial_\nu W^{\cdot}_\mu$ are contracted with an antisymmetric expression we may replace it by $-2\partial_\nu W^{\cdot}_\mu$. Furthermore we may write $W^{\mu +} = g^{\mu\rho} W^+_\rho$, $W^{\mu -} = g^{\mu\sigma} W^-_\sigma$ and $W^\mu_3 = g^{\mu\lambda} W_{\lambda 3}$, $W^+_\mu = \delta^\rho_\mu W^+_\rho$, $W^-_\mu = \delta^\sigma_\mu W^-_\sigma$ and $W_{\mu 3} = \delta^\lambda_\mu W_{\lambda 3}$ etc. such that the fields appear as a common factor $W^+_\rho W^-_\sigma W_{\lambda 3}$.

We thus obtain

$$\begin{aligned}
\mathcal{L}^{\text{triple}}_{\text{YM}} = ig \Big[& \left(g^{\rho\lambda} g^{\sigma\nu} - g^{\rho\sigma} g^{\lambda\nu}\right)\left(\partial_\nu W^+_\rho\right) W^-_\sigma W_{\lambda 3} \\
& + \left(g^{\rho\sigma} g^{\lambda\nu} - g^{\rho\nu} g^{\sigma\lambda}\right) W^+_\rho \left(\partial_\nu W^-_\sigma\right) W_{\lambda 3} \\
& + \left(g^{\rho\nu} g^{\sigma\lambda} - g^{\rho\lambda} g^{\sigma\nu}\right) W^+_\rho W^-_\sigma \left(\partial_\nu W_{\lambda 3}\right)\Big] \quad .
\end{aligned}$$

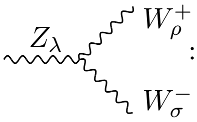
In momentum space we assign the incoming momenta p_1, p_2 and p_3 to the fields W^+, W^- and W_3 respectively, such that the derivative is given by a factor $\partial_\nu \to -ip_{i\nu}$:

$$\begin{aligned}
\tilde{\mathcal{L}}^{\text{triple}}_{\text{YM}} &= g\Big[\left(g^{\rho\lambda} g^{\sigma\nu} - g^{\rho\sigma} g^{\lambda\nu}\right) p_{1\nu} \\
&\quad + \left(g^{\rho\sigma} g^{\lambda\nu} - g^{\rho\nu} g^{\sigma\lambda}\right) p_{2\nu} \\
&\quad + \left(g^{\rho\nu} g^{\sigma\lambda} - g^{\rho\lambda} g^{\sigma\nu}\right) p_{3\nu}\Big] \tilde{W}^+_\rho(p_1)\, \tilde{W}^-_\sigma(p_2)\, \tilde{W}_{\lambda 3}(p_3) \\
&= g\, V^{\rho\sigma,\lambda}(p_1, p_2, p_3)\, \tilde{W}^+_\rho(p_1)\, \tilde{W}^-_\sigma(p_2)\, \tilde{W}_{\lambda 3}(p_3)
\end{aligned}$$

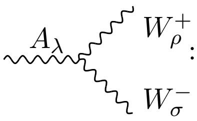
with

$$V^{\rho\sigma,\lambda}(p_1,p_2,p_3) = g^{\rho\sigma}(p_2-p_1)^\lambda + g^{\rho\lambda}(p_1-p_3)^\sigma + g^{\sigma\lambda}(p_3-p_2)^\rho \quad .$$

The replacement of $W_{\lambda 3}$ by Z_μ and A_μ is trivial and we find for the triple gauge vertices:



$$: \quad g\cos\theta_W \; V^{\rho\sigma,\lambda}(p_1,p_2,p_3)$$



$$: \quad e \; V^{\rho\sigma,\lambda}(p_1,p_2,p_3) \quad .$$

Notice that by $g\sin\theta_W = e$ the $W^\pm$ fields indeed couple with charge unit 1, as it should be. For the quadrilinear term, finally, we use the identity

$$\varepsilon_{abc}\varepsilon_{ab'c'} = \delta_{bb'}\delta_{cc'} - \delta_{bc'}\delta_{cb'}$$

such that

$$\mathcal{L}^{\text{quartic}}_{\text{YM}} = -\frac{g^2}{4}\left[\left(\vec{W}_\mu \cdot \vec{W}^\mu\right)^2 - \left(\vec{W}_\mu \cdot \vec{W}_\nu\right)^2\right].$$

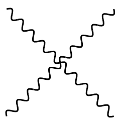
Inserting the charged fields we get

$$\mathcal{L}^{\text{quartic}}_{\text{YM}} = -\frac{g^2}{2}\left(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma}\right)\left(W^+_\mu W^-_\nu + 2W_{\mu 3}W_{\nu 3}\right)W^+_\rho W^-_\sigma$$

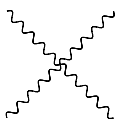
The term quadrilinear in W_3 has dropped out in the difference. This must be so, since otherwise we would get photon self–interactions. This cannot be the case for the Abelian photon field. If we symmetrize in the identical fields (i.e. in the $W_{\mu 3}W_{\nu 3}$ and the $W^-_\nu W^-_\sigma$ terms) we may express the vertices in terms of the tensor

$$T^{\rho\sigma,\mu\nu} \doteq 2g^{\rho\sigma}g^{\mu\nu} - g^{\rho\mu}g^{\sigma\nu} - g^{\rho\nu}g^{\sigma\mu}.$$

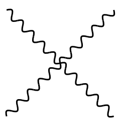
After substituting the $W_{\mu 3}W_{\nu 3}$ term we have for the quartic gauge interaction vertices:



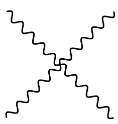
$$: \qquad g^2 \qquad T^{\rho\mu,\sigma\nu}\left(\tfrac{1}{2!}\,\tfrac{1}{2!}\right)$$



$$: \quad -g^2\cos^2\theta_W \quad T^{\rho\sigma,\mu\nu}\left(\tfrac{1}{2!}\right)$$



$$: \quad -eg\cos\theta_W \quad T^{\rho\sigma,\mu\nu}$$



$$: \qquad -e^2 \qquad T^{\rho\sigma,\mu\nu}\left(\tfrac{1}{2!}\right)$$

$$Z_\mu \qquad W^+_\rho$$
$$A_\nu \qquad W^-_\sigma$$
$$A_\mu \qquad W^+_\rho$$
$$A_\nu \qquad W^-_\sigma$$

The permutation symmetry factors $\frac{1}{2!}$ for each pair of identical fields are omitted in the Feynman rules since diagrams obtained by permutations of identical lines are not counted separately.

If we add the mass terms

$$\mathcal{L}_{\rm mass} = -\sum_f m_f \bar{\psi}_f \psi_f + \frac{1}{2} M_Z^2 Z_\mu Z^\mu + M_W^2 W^+_\mu W^{\mu -}$$

we have a model which essentially describes well all experimental data. **But** it is not an acceptable theory, because the mass term breaks $SU(2)_L \otimes U(1)_Y$ in a catastrophic way which makes the theory non-renormalizable. This means that beyond the Born approximation perturbative calculations are not unambiguously defined. A renormalizable massive vector boson gauge theory can only be obtained if masses are generated by the Higgs mechanism[33].

13.4 The Higgs field and mass generation

A **minimal renormalizable extension** of the massive vector boson gauge theory can be obtained only if physical particle masses are generated by a Higgs mechanism. Starting from the massless $SU(2)_L \otimes U(1)_Y$ invariant gauge theory one couples all fields which should acquire a mass in an $SU(2)_L \otimes U(1)_Y$ invariant manner to a Higgs field, a scalar field which develops a non-vanishing vacuum expectation value. Since the Higgs mechanism has to break $SU(2)_L \otimes U(1)_Y$ to $U(1)_{\rm em}$, the Higgs field must transform in a non-trivial way under both factors of the gauge group. Again, the simplest possibility is to assume the Higgs field to be a $SU(2)_L$ doublet Φ_b transforming according to the fundamental representation. What is the transformation property of Φ_b under $U(1)_Y$? The component of Φ_b which develops a non-vanishing vacuum expectation value must be electrically neutral, since the charge Q must be a good quantum number in the broken phase. By <u>convention</u> we choose the lower component to be neutral. Since $Q = T_3 + \frac{Y}{2}$, the field Φ_b must have $Y = 1$ and the upper component must have charge $Q = 1$. Therefore we write

$$\Phi_b = \left(\begin{array}{c} \varphi^+ \\ \varphi_0 \end{array} \right) .$$

We may represent this complex doublet in terms of **four** real fields H_s and φ_i ; $i = 1, 2, 3$.

We first introduce the 2×2 matrix field

$$\frac{1}{\sqrt{2}}(H_s + i\tau_i\varphi_i) = \left(\begin{array}{cc} \frac{H_s + i\varphi_3}{\sqrt{2}} & i\frac{\varphi_1 - i\varphi_2}{\sqrt{2}} \\ i\frac{\varphi_1 + i\varphi_2}{\sqrt{2}} & \frac{H_s - i\varphi_3}{\sqrt{2}} \end{array} \right) = (\Phi_t, \Phi_b) \ .$$

[33] As we will see later, in the electroweak SM except for the Higgs mass all other masses are generated by the Higgs field acquiring a vacuum expectation value. Besides the weak gauge bosons $W^\pm$ and Z, also all fermions aquire their mass via the minimal Higgs mechanisms. Note that this does not mean that the mass of ordinary matter in our universe is solely due to the Higgs mechanism. The major part, about 99 %, of the mass of matter fom which stars planets and interstellar matter is made, namely, protons and neutrons (baryonic matter), get their mass through strong interaction binding energy. In fact, the nucleon masses (~ 1 GeV) are completely dominated by the binding energy, while the light u and d quarks nucleons are made of only contribute a tiny amount (as $m_u \sim 3$ MeV, $m_d \sim 8$ MeV). The Higgs mechanism is responsible only for the quark masses not for the nucleon masses.

and iso-doublet vectors $\chi_t = \binom{1}{0}$ and $\chi_b = \binom{0}{1}$.

Then, we may write

$$\Phi_b = \begin{pmatrix} \varphi^+ \\ \varphi_0 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(H_s + i\tau_i \varphi_i \right) \chi_b$$

with $\varphi^\pm = \pm \frac{i}{\sqrt{2}} \left(\varphi_1 \mp i\varphi_2 \right)$ and $\varphi_0 = \frac{1}{\sqrt{2}} \left(H_s - i\varphi_3 \right)$.

The field

$$\Phi_t = i\tau_2 \Phi_b^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Phi_b^* = \begin{pmatrix} \varphi_0^* \\ -\varphi^- \end{pmatrix} = \frac{1}{\sqrt{2}} \left(H_s + i\tau_i \varphi_i \right) \chi_t$$

is the Y-charge conjugate doublet field with $Y = -1$. For the physical interpretation of the Higgs field it is important to notice that $H_s + i\tau_i\varphi_i$ can be written in "polar" form: [34]

$$H_s + i\tau_i \varphi_i = \rho_s e^{i\tau_i \theta_i} = \rho_s \left(\cos\theta + i\, \frac{\sin\theta}{\theta} \tau_i \theta_i \right)$$

with $\theta = \sqrt{\theta_i \theta_i}$ i.e. $H_s = \rho_s \cos\theta, \varphi_i = \rho_s \frac{\sin\theta}{\theta} \theta_i$ or for infinitesimal $\theta :$ $H_s \simeq \rho_s, \varphi_i \simeq \rho_s \theta_i$. We may write therefore

$$\Phi_b = \frac{\rho_s}{\sqrt{2}} U(\theta) \chi_b \quad \text{with} \quad U(\theta) \in SU(2) \ .$$

The transformation properties of Φ_b and the requirement of local $SU(2)_L \otimes U(1)_Y$ invariance completely determine the couplings of the gauge fields to the Higgs field: The invariant Higgs Lagrangian reads

$$\mathcal{L}_{\text{Higgs}} = \left(D_\mu \Phi_b \right)^+ \left(D^\mu \Phi_b \right) - \lambda \left(\Phi_b^+ \Phi_b \right)^2 + \mu^2 \Phi_b^+ \Phi_b$$

with covariant derivative

$$D_\mu \Phi_b = \left(\partial_\mu - i\frac{g'}{2} B_\mu - ig\frac{\tau_a}{2} W_{\mu a} \right) \Phi_b \ .$$

Invariant couplings of the Higgs field with the fermions may easily be constructed. Noting that $L_\psi = \binom{\psi_t}{\psi_b}_L, \Phi_b$ and Φ_t transform identically under $SU(2)_L : L_\psi \to U(x) L_\psi, \Phi_i \to U(x) \Phi_i$, while $\psi_{iR} \to \psi_{iR}$, we obtain scalars $\bar{L}_\psi \Phi_i \psi_{iR}$ $(i = t, b)$ which have hypercharge zero if ψ_{iR} is chosen as indicated. Explicitly the two kind of terms read

$$\begin{array}{ccc} \bar{\psi}_{tL} \varphi^+ \psi_{bR} & + & \bar{\psi}_{bL} \varphi_0 \psi_{bR} \\ \bar{\psi}_{tL} \varphi_0^* \psi_{tR} & - & \bar{\psi}_{bL} \varphi^- \psi_{tR} \ . \end{array}$$

[34] Use $(\tau_i\theta_i)^2 = \tau_i\tau_k\theta_i\theta_k = \frac{1}{2}\left\{\tau_i,\tau_k\right\}\theta_i\theta_k = \delta_{ik}\theta_i\theta_k = \theta_i\theta_i = \theta^2$ with $\theta = \sqrt{\theta_i\theta_i}$. Separating even and odd powers in the exponential series defining $e^{i\tau_i\theta_i}$ we have

$$\begin{aligned} e^{i\tau_i\theta_i} &= \sum_{k=0}^{\infty} \frac{i^{2k}}{2k!} (\tau_i\theta_i)^{2k} + \sum_{k=0}^{\infty} \frac{i^{2k+1}}{2k+1!} (\tau_i\theta_i)^{2k} \tau_i\theta_i \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k!} \theta^{2k} + i\tau_i \frac{\theta_i}{\theta} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1!} \theta^{2k+1} = \cos\theta + i\tau_i \frac{\theta_i}{\theta} \sin\theta \ . \end{aligned}$$

In order to obtain Hermitean operators we have to add the Hermitean conjugate (h.c.) to each of the terms. Thus the Yukawa type Higgs-fermion Lagrangian for one lepton–quark family reads

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -G_{b\ell}\,(\bar{L}_\ell \Phi_b \ell_R \; + \text{h.c.}) - G_{t\ell}\,(\bar{L}_\ell \Phi_t \nu_{\ell R} + \text{h.c.}) \\ &\quad -G_{bq}\,(\bar{L}_q \Phi_b b_R + \text{h.c.}) - G_{tq}\,(\bar{L}_q \Phi_t t_R \; + \text{h.c.})\end{aligned}$$

where $G_{i\ell}$ and G_{iq} are the Yukawa couplings for the leptons and quarks, respectively. A more general form for $\mathcal{L}_{\text{Yukawa}}$ will be discussed below. For massless neutrinos we must set $G_{t\ell} \equiv 0$.

It is quite non-trivial that the masses of the weak gauge bosons and the masses of the fermions may be generated with the same single Higgs field. This only works because besides the fermion doublets there exist fermion singlets as it is obvious from the Yukawa terms given above. With this minimal choice of elementary scalars the SM often is called more specifically the **minimal standard model**.

Above we have noted that Φ_b may be written as a $SU(2)$ transform of

$$\Phi_b^u = \frac{\rho_s}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

i.e.

$$\Phi_b = \frac{\rho_s}{\sqrt{2}} U(\theta) \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

If we perform a **particular** gauge transformation of the fields

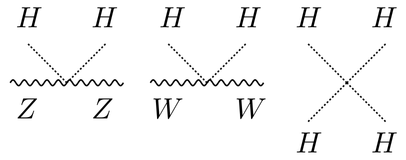
$$\begin{aligned}\Phi_b^u &= U^+(\theta)\,\Phi_b = \frac{\rho_s}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ W_{\mu a}^u \tau_a &= U^+(\theta)\, W_{\mu a}\tau_a U(\theta) + \frac{2i}{g} U^+(\theta)\partial_\mu U(\theta) \\ L_\psi^u &= U^+(\theta)\, L_\psi\end{aligned}$$

the three real fields θ_i and hence φ_i are eliminated from the invariant Lagrangian. The fields φ_i therefore cannot be physical, they are called **Higgs ghosts**. Since $\mathcal{L}_{\text{matter}}$ and $\mathcal{L}_{\text{YM}}$ are invariant their form is unaffected. In contrast, $\mathcal{L}_{\text{Higgs}}$ and $\mathcal{L}_{\text{Yukawa}}$ take a particularly simple form since only **one** real scalar field is left in Φ_b^u, the **physical Higgs field** $\rho_s = H_s$. Notice that the number of fields which can be eliminated by a $SU(2)$ transformation is given by the order of the group, which is 3 in our case. The particular gauge we have chosen to eliminate the Higgs ghosts is called **unitary gauge** (U-gauge). The general case in which the Higgs ghosts are present is called **renormalizable gauge** (R-gauge) for reasons which will become clear later on.

Since physics is more transparent in the U-gauge let us consider the Higgs mechanism first in this gauge. If we replace Φ_b by $\frac{H_s}{\sqrt{2}}\chi_b$ in $\mathcal{L}_{\text{Higgs}}$ and replace the gauge fields by the physical fields $W^\pm, Z$ and A we obtain the simple form

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} &= \frac{1}{2}\,(\partial_\mu H_s \partial^\mu H_s) + \frac{1}{8} H_s^2 \left\{ \left(\frac{g}{\cos\theta_W}\right)^2 Z_\mu Z^\mu + 2g^2\, W_\mu^+ W^{\mu-} \right\} \\ &\quad -\frac{\lambda}{4} H^4 + \frac{\mu^2}{2} H^2\end{aligned}$$

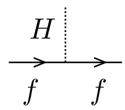
with vertices



and similarly

$$\mathcal{L}_{\text{Yukawa}} = -\sum_f \frac{G_f}{\sqrt{2}} H_s \bar{\psi}_f \psi_f$$

with vertex



If $\mu^2 > 0$ by the Higgs mechanism $< 0 \mid H_s \mid 0 >= v > 0$ in the physical ground state $\mid 0 >$. The existence

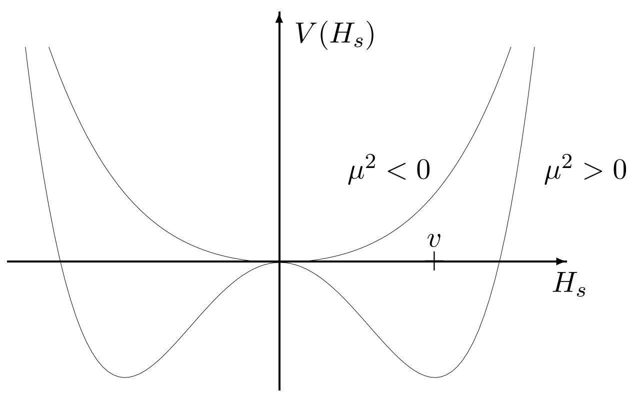


Fig. 13.1: Higgs potential of the SM a) in the symmetric ($\mu^2 < 0$) and b) in the broken phase ($\mu^2 > 0$).

of a **Bose condensate** density v breaks the global Z_2 symmetry $H_s \leftrightarrow -H_s$ of the Higgs Lagrangian and as we shall see only a $U(1)_{\text{em}}$ local symmetry remains from the original $SU(2)_L \otimes U(1)_Y$ symmetry.

The correct mass spectrum is obtained by the shift

$$H = H_s - v$$

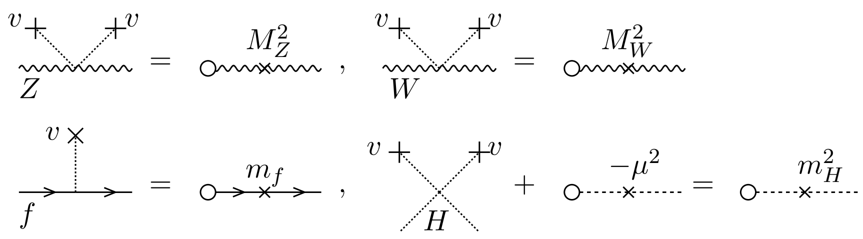
to the physical Higgs field H which satisfies $< 0|H|0 >= 0$. The Lagrangian then takes the form

$$\begin{aligned}
\mathcal{L}_{\text{Higgs}} &= \frac{1}{2}\left(\partial_\mu H \partial^\mu H\right) + \frac{1}{2}\left(1 + \frac{H}{v}\right)^2 \left\{M_Z^2 Z_\mu Z^\mu + 2M_W^2 W_\mu^+ W^{\mu -}\right\} \\
&\quad -\frac{\lambda}{4} H^4 - \lambda v H^3 - \frac{1}{2} m_H^2 H^2 \\
\mathcal{L}_{\text{Yukawa}} &= -\sum_f m_f \left(1 + \frac{H}{v}\right) \bar{\psi}_f \psi_f .
\end{aligned}$$

with the characteristic **mass-coupling relations**

$$M_Z = \frac{gv}{2\cos\theta_W} \;,\;\; M_W = \frac{gv}{2} \;,\;\; m_f = \frac{G_f}{\sqrt{2}}\, v \quad \text{and} \quad m_H = \sqrt{2\lambda}\, v$$

The masses are generated through interaction with the condensate: Diagrammatically, with $v =< 0|H_s|0 >=$ $\times\, v$



As a result we have found

$$\mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} = \mathcal{L}_{\text{mass}} + \mathcal{L}_{\text{Higgs, free}} + \mathcal{L}_{\text{Higgs, int}}$$

where the existence of the Higgs and its interaction is experimentally unverified. However, as we shall see below, the vacuum expectation value of the Higgs field is very precisely determined

$$v = \left(\sqrt{2}G_\mu\right)^{-1/2} = 246.221(3)\ \text{GeV}$$

by the Fermi constant G_μ . v is called the **Fermi scale**.

If we take serious the SM and hence the Higgs mechanism, we have to conclude that **the existence of the Higgs Bose condensate has been established experimentally**.

Like in the Ginzburg-Landau theory of superconductivity, where the Bose condensate field represents an effective description only, of the Cooper-pairs of the underlying microscopic BCS theory, the Higgs field could turn out to be a low energy effective description of a composite fermion-antifermion condensate. Generally, it is believed that the SM should provide the correct low energy effective theory for physics below the 1 TeV scale.

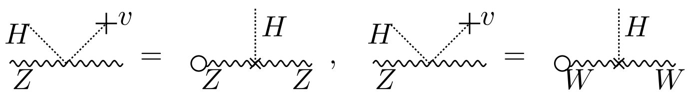
What the Higgs mechanism has done for us may be summarized as follows: The Higgs mechanism

- breaks $SU(2)_L \otimes U(1)_Y$ to $U(1)_{\text{em}}$ as required by phenomenology
- generates the masses of $W^\pm, Z$ and the fermions (without conflicting renormalizability)
- provides a "physical cut-off" to the massive vector boson gauge theory i.e. it "restores" renormalizability and the associated smooth high energy behavior of physical transition amplitudes.
- implies the existence of a **new** massive scalar particle, the Higgs particle, the mass of which is essentially unconstrained. Some general constraints follow, however, since neither the limit $m_H \to 0$ (causing an infrared problem) nor the limit $m_H \to \infty$ (causing ultraviolet problems) exists.

Notice that **the Higgs couples universally to the masses of particles**. The couplings are proportional to the mass and the mass-square for fermions and bosons, respectively, and in proportion to v:

$$
\begin{array}{ll}
m_f/v & \text{for fermions} \\
M_V^2/v & \text{for bosons and triple vertices} \\
M_V^2/v^2 & \text{for bosons and quartic vertices}
\end{array}
$$

where $V = Z, W$ or H. Thus the coupling of the Higgs to ordinary matter consisting of u-quarks, d-quarks and electrons is minuscule. In fact all couplings to the light fermions including the b–quark are too small for getting reasonable production cross-sections. The Higgs is therefore very hard to find. It can only be produced at reasonable rates in association with heavy particles like W- and Z-bosons or the top–quark, which remains to be discovered first. In this context we note an important consequence of the "spontaneous symmetry breaking", namely, the generation of the triple vertices HZZ and HWW



which are the most important vertices for **Higgs production** since, firstly, the coupling is large and, secondly, they allow the Higgs to be **singly produced** which is particularly important for energetic (phase space) reason when the Higgs is heavy. Detecting the Higgs by these couplings is very promising and would be an unambiguous signal of the Higgs mechanism at work. Processes of this kind are, H-production in Z-decay $Z \to H\mu^+\mu^-$ which is allowed energetically provided the Higgs is lighter than the Z, or in associated production with a Z $e^+e^- \to Z^* \to ZH$ at sufficiently high energy which strongly depends on the unknown mass of the Higgs.

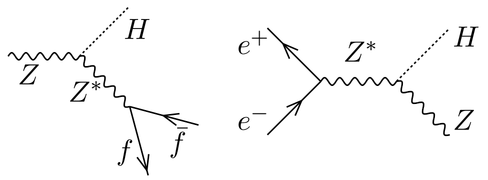


Fig. 13.2: Higgs production processes

At present the limit for m_H from LEP experiments is

$$m_H > 60 \text{ GeV } \ (95\% \text{ CL}) .$$

Possible windows for a light Higgs have been excluded all the way down to $m_H = 0$. At LEP2 the Higgs search can be extended to about $m_H \simeq M_Z$. If the Higgs should be heavier, and this is likely to be the case, a discovery is possible only at future colliders like SSC or LHC.

Notice that higher order predictions depend on the unknown mass of the Higgs boson, the remnant from spontaneous symmetry breaking, and the mass of the unknown top quark, the missing member of the 3rd fermion family and other possible unknown physics. While higher order predictions of physical quantities depend substantially on the unknown top mass the dependence on the unknown Higgs mass is much weaker as we shall discuss below. At present, the Higgs sector is completely unverified and its confirmation is a big challenge for experimental particle physics.

13.5 The Higgs sector in the R-gauge

This chapter should be skipped by the reader who is not interested in problems of renormalizability or higher order calculations.

In the previous subsection we used the U-gauge to discuss the physical implications of the Higgs mechanism. The problem with the U-gauge is, that it is not manifestly renormalizable. This will become clear later on. In order to stay within a manifestly renormalizable framework we have to keep the Higgs as a complex doublet and stay within a manifestly gauge invariant scheme. The Higgs mechanism can be implemented exactly as before, but the Higgs and Yukawa Lagrangians are more complicated due to the extra Higgs ghosts $\varphi^\pm$ and $\varphi = \sqrt{2}\,\mathrm{Im}\,\varphi_0$. We assume $\mu^2 > 0$ in the Higgs Lagrangian and we may choose the vacuum $\mid 0 >$ such that

$$< 0 \mid \Phi_b \mid 0 >= \frac{v}{\sqrt{2}}\chi_b \quad \text{with} \quad v = v^* > 0 \ .$$

Again we have to rewrite the Lagrangian in terms of a shifted field

$$\Phi'_b = \Phi_b - \frac{v}{\sqrt{2}}\chi_b$$

for which

$$< 0 \mid \Phi'_b \mid 0 >= 0$$

before we can set up a perturbative approach. We first calculate the vector boson (VB) mass term

$$\begin{aligned}
\mathcal{L}_{\text{mass, VB}} &= \mathcal{L}_{\text{Higgs}}\Big|_{\Phi_b=\frac{v}{\sqrt{2}}\chi_b} + V\left(\frac{v^2}{2}\right) = \frac{v^2}{2}\left(D_\mu\chi_b\right)^+\left(D^\mu\chi_b\right) \\
&= \frac{v^2}{8}\chi_b^+\left(g'B_\mu + g\tau_{a'}W_{\mu a'}\right)\left(g'B^\mu + g\tau_a W_a^\mu\right)\chi_b \\
&= \frac{v^2}{8}\left(g'^2 B_\mu B^\mu - 2g'gB_\mu W_3^\mu + g^2 W_{\mu 3}W_3^\mu + g^2\left(W_{\mu 1}W_1^\mu + W_{\mu 2}W_2^\mu\right)\right) \\
&= \frac{v^2(g^2+g'^2)}{8}\,Z_\mu Z^\mu + \frac{v^2 g^2}{4}\,W_\mu^+ W^{\mu -} \\
&= \frac{1}{2}M_Z^2\,Z_\mu Z^\mu + M_W^2\,W_\mu^+ W^{\mu -} + \text{“}\, 0\cdot A_\mu A^\mu \,\text{”} \ .
\end{aligned}$$

We have used $\tau_{a'}\tau_a = \delta_{a'a} + i\varepsilon_{a'ac}\tau_c$ yielding $\tau_{a'}\tau_a W_{\mu a'}W_a^\mu = W_{\mu a}W_a^\mu$ since the ε-term cannot contribute due to the symmetry of $W_{\mu a'}W_a^\mu$. From the mixed terms only τ_3 contributes since $\chi_b^+\tau_i\chi_b = -\delta_{i3}$. The fields $W^\pm, Z$ and A follow from diagonalizing the mass-matrix i.e. by setting $gW_{\mu 3} - g'B_\mu = \sqrt{g^2+g'^2}Z_\mu$ and $W_{\mu 1}W_1^\mu + W_{\mu 2}W_2^\mu = 2W_\mu^+W^{\mu -}$. A_μ must be the field orthogonal to Z_μ, which turns out to be massless, and this fact ascertains the existence of a residual Abelian $U(1)_{\text{em}}$ symmetry. Of course the physical fields found here are identical to the ones we introduced earlier by different arguments.

The remaining terms of $\mathcal{L}_{\text{Higgs}}$ may be worked out by using $\Phi_b = \Phi + \frac{v}{\sqrt{2}}\chi_b$ with $\Phi = \binom{\varphi^+}{\varphi_0}$ where

$$\varphi_0 = \frac{H - i\varphi}{\sqrt{2}} \ , \quad H, \varphi \ \text{real} \ .$$

Then, we obtain

$$\begin{aligned}
\Phi_b^+\Phi_b - \frac{v^2}{2} &= \Phi^+\Phi + \frac{v}{\sqrt{2}}\left(\chi_b^+\Phi + \Phi^+\chi_b\right) \\
&= \varphi^+\varphi^- + \frac{1}{2}\varphi^2 + \frac{1}{2}H^2 + vH \ .
\end{aligned}$$

By the ground state condition

$$\left.\frac{\partial V}{\partial \Phi_b}\right|_{\Phi_b = \frac{v}{\sqrt{2}}\chi_b} = 0$$

we have $\mu^2 = \lambda v^2 = \frac{m_H^2}{2}$. We then get for the part bilinear in the scalars:

$$\begin{aligned}\mathcal{L}_{\text{Higgs},0} &= \left(\partial^\mu \Phi^+\right)\left(\partial_\mu \Phi\right) - \lambda v^2 H^2 \\ &= \partial_\mu \varphi^+ \partial^\mu \varphi^- + \frac{1}{2}\partial_\mu \varphi \partial^\mu \varphi + \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{m_H^2}{2} H^2 \quad .\end{aligned}$$

The ghost fields $\varphi^\pm$ and φ (the "would be Goldstone bosons") are massless due to $\mu^2 - \lambda v^2 = 0$. By the same condition linear terms in the scalars are absent.

For the scalar interactions we find

$$\begin{aligned}\mathcal{L}_{\text{Higgs, int}} &= -\lambda\left(\Phi^+\Phi\right) - 2\lambda v\left(\Phi^+\Phi\right) H \\ &= -\lambda\left(\varphi^+\varphi^-\right)^2 - \lambda\varphi^+\varphi^-\left(H^2+\varphi^2\right) - \frac{\lambda}{4}\left(H^2+\varphi^2\right)^2 \\ &\quad -2\lambda v\varphi^+\varphi^- H - \lambda v\varphi^2 H - \lambda v H^3 \quad .\end{aligned}$$

There is a bilinear term mixing scalar and vector fields (p.i. means partial integration)

$$\begin{aligned}\mathcal{L}^{(0)}_{\text{Higgs, VB}} &= \frac{v}{\sqrt{2}}\chi_b^+\left(i\frac{g'}{2}B_\mu + ig\frac{\tau_a}{2}W_{\mu a}\right)\partial_\mu\Phi + \text{ h.c.} \\ &= -M_Z Z_\mu \partial^\mu\varphi + iM_W W_\mu^- \partial^\mu\varphi^+ - iM_W W_\mu^+\partial^\mu\varphi^- \\ &\overset{\text{p.i.}}{=} M_Z\left(\partial^\mu Z_\mu\right)\varphi - iM_W\left(\partial^\mu W_\mu^-\right)\varphi^+ + iM_W\left(\partial^\mu W_\mu^+\right)\varphi^-\end{aligned}$$

which can be compensated by a similar term of opposite sign from the gauge fixing Lagrangian $\mathcal{L}_{\text{GF}}$ (see below). Finally, there remains the quite ugly term

$$\begin{aligned}\mathcal{L}^{(\text{int})}_{\text{Higgs, VB}} &= \frac{i}{2}\Phi^+\left(g'B_\mu + g\tau_a W_{\mu a}\right)\partial^\mu\Phi + \text{ h.c.} \\ &\quad +\frac{1}{4}\Phi^+\left(g'B_\mu + g\tau_{a'}W_{\mu a'}\right)\left(g'B^\mu + g\tau_a W_a^\mu\right)\Phi \\ &\quad +\frac{1}{4}\frac{v}{\sqrt{2}}\chi_b^+\left(g'B_\mu + g\tau_{a'}W_{\mu a'}\right)\left(g'B^\mu + g\tau_a W_a^\mu\right)\Phi + \text{ h.c.}\end{aligned}$$

which has to be worked out in terms of $W_\mu^\pm, Z_\mu, A_\mu$ and $H, \varphi^\pm, \varphi$. The result is part of the Feynman rules, which will be given later.

The Yukawa term is much simpler. We consider one doublet only. Thus

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -G_{b\psi}\left(\bar\psi_{tL}\varphi^+\psi_{bR} + \bar\psi_{bL}\varphi_0\psi_{bR}\right) + \text{ h.c.} \\ &\quad -G_{t\psi}\left(\bar\psi_{tL}\varphi_0^*\psi_{tR} - \bar\psi_{bL}\varphi^-\psi_{tR}\right) + \text{ h.c.}\end{aligned}$$

with $\varphi_0 = \frac{v}{\sqrt{2}} + \frac{H - i\varphi}{\sqrt{2}}$ we get

$$
\begin{aligned}
\mathcal{L}_{\text{Yukawa}} &= -\frac{v}{\sqrt{2}} G_{b\psi}\left(\bar{\psi}_{bL}\psi_{bR}+\bar{\psi}_{bR}\psi_{bL}\right) \\
&\quad -\frac{v}{\sqrt{2}} G_{t\psi}\left(\bar{\psi}_{tL}\psi_{tR}+\bar{\psi}_{tR}\psi_{tL}\right) \\
&\quad -G_{b\psi}\bar{\psi}_{tL}\psi_{bR}\varphi^{+}+G_{t\psi}\bar{\psi}_{tR}\psi_{bL}\varphi^{+} \\
&\quad +G_{t\psi}\bar{\psi}_{bL}\psi_{tR}\varphi^{-}-G_{b\psi}\bar{\psi}_{bR}\psi_{tL}\varphi^{-} \\
&\quad +i\frac{G_{b\psi}}{\sqrt{2}}\left(\bar{\psi}_{bL}\psi_{bR}-\bar{\psi}_{bR}\psi_{bL}\right)\varphi \\
&\quad -\frac{G_{b\psi}}{\sqrt{2}}\left(\bar{\psi}_{bL}\psi_{bR}+\bar{\psi}_{bR}\psi_{bL}\right)H \\
&\quad -i\frac{G_{t\psi}}{\sqrt{2}}\left(\bar{\psi}_{tL}\psi_{tR}-\bar{\psi}_{tR}\psi_{tL}\right)\varphi \\
&\quad -\frac{G_{t\psi}}{\sqrt{2}}\left(\bar{\psi}_{tL}\psi_{tR}+\bar{\psi}_{tR}\psi_{tL}\right)H
\end{aligned}
$$

where we used $\left(\bar{\psi}_{iL}\psi_{kR}\right)^{*}=\bar{\psi}_{kR}\psi_{iL}$. Working out the chiral projectors $\psi_L=\frac{1-\gamma_5}{2}\psi$ etc. and replacing $G_{t\psi}=\sqrt{2}\frac{m_t}{v}, G_{b\psi}=\sqrt{2}\frac{m_b}{v}$ we arrive at

$$
\begin{aligned}
\mathcal{L}_{\text{Yukawa}} &= -m_b\bar{\psi}_b\psi_b-m_t\bar{\psi}_t\psi_t-\frac{m_b}{v}\bar{\psi}_b\psi_bH-\frac{m_t}{v}\bar{\psi}_t\psi_tH \\
&\quad +i\frac{m_b}{v}\bar{\psi}_b\gamma_5\psi_b\varphi-i\frac{m_t}{v}\bar{\psi}_t\gamma_5\psi_t\varphi \\
&\quad -\frac{1}{\sqrt{2}v}\left(m_b\bar{\psi}_t\left(1+\gamma_5\right)\psi_b-m_t\bar{\psi}_t\left(1-\gamma_5\right)\psi_b\right)\varphi^{+} \\
&\quad -\frac{1}{\sqrt{2}v}\left(m_b\bar{\psi}_b\left(1-\gamma_5\right)\psi_t-m_t\bar{\psi}_b\left(1+\gamma_5\right)\psi_t\right)\varphi^{-} \quad .
\end{aligned}
$$

Due to the many unphysical vertices obtained in a renormalizable gauge higher order calculations get extremely involved in general. Tree level calculations may be done in the U-gauge without problems.

13.6 The Yukawa sector and flavor mixing

So far we have considered the Higgs–fermion couplings for each quark–lepton family separately. The fact that there exist three families of fermions which are made up of fields with identical $SU(2)_L \otimes U(1)_Y$ transformation properties allows us to form invariant Yukawa couplings for arbitrary combinations of fields from the different families. Such **flavor mixing** is known to occur in the quark sector while for leptons all searches for family mixing have been negative so far.

For the quarks we have four **horizontal vectors** in "family space" with identical quantum numbers with respect to the local gauge group. These are the left–handed and right–handed versions of the up and down family vectors $u_i=(u,c,t)$ and $d_i=(d,s,b)$ and we denote them by

u_{iL}, d_{iL}, u_{iR}, d_{iR} . The general form of the Yukawa term reads

$$\mathcal{L}^q_{\text{Yukawa}} = -\sum_{i,j=1}^{3} \left[G^{ij}_{tq} \bar{L}_{qi} \Phi_t u_{jR} + G^{ij}_{bq} \bar{L}_{qi} \Phi_b d_{jR} + \text{h.c.} \right]$$

with G^{ij}_{tq} and G^{ij}_{bq} arbitrary complex 3 x 3 matrices. The mass–matrix we obtain by inserting $\Phi_b = \frac{v}{\sqrt{2}}\binom{0}{1}$ and $\Phi_t = \frac{v}{\sqrt{2}}\binom{1}{0}$. Thus

$$\begin{aligned} \mathcal{L}^q_{\text{mass}} &= -\frac{v}{\sqrt{2}} \sum_{i,j=1}^{3} \left[G^{ij}_{tq} \bar{u}_{iL} u_{jR} + G^{ij}_{bq} \bar{d}_{iL} d_{jR} + \text{h.c.} \right] \\ &= -\sum_{i,j=1}^{3} \left[m^{ij}_{tq} \bar{u}_{iL} u_{jR} + m^{ij}_{bq} \bar{d}_{iL} d_{jR} + \text{h.c.} \right] \end{aligned}$$

with

$$m^{ij}_{\cdot} = \frac{v}{\sqrt{2}} G^{ij}_{\cdot} \quad .$$

The quark fields considered up to now are in the **weak interaction basis** in which the matter field Lagrangian is diagonal in the families. We now have to find the **physical fields** in the **mass eigenstate basis** in which the mass–matrix is diagonal such that each field has a fixed mass[35]. To this end we have to perform global unitary transformation on the horizontal vectors. Unitary because they have to leave the kinetic terms of the quarks invariant and global because we have to diagonalize the constant coupling matrices $G_{\cdot\, q}$.

We first perform the transformations

$$\begin{aligned} u_{jR} &\to (V_{uR})_{jk}\, u_{kR} \\ d_{jR} &\to (V_{dR})_{jk}\, d_{kR} \\ \begin{pmatrix} u \\ d \end{pmatrix}_{jL} &\to (V_L)_{jk} \begin{pmatrix} u \\ d \end{pmatrix}_{kL} \end{aligned}$$

of the singlets and the doublet. Since the doublet fields are transformed as a doublet these unitary transformations do not change the matter field Lagrangian which exhibits terms of the form $\bar{L}_{qi} \cdots L_{qi}$, $\bar{u}_{iR} \cdots u_{iR}$ and $\bar{d}_{iR} \cdots d_{iR}$! For the mass matrices we obtain

$$\begin{aligned} m_{tq} &\to V_L^+\, m_{tq}\, V_{uR} \quad \text{real diagonal} \\ m_{bq} &\to V_L^+\, m_{bq}\, V_{dR} \end{aligned}$$

[35]Notice that the mass-matrices m_{bq} and m_{tq} are not Hermitean in general. We make use of the fact that any square matrix can be diagonalized with the help of two unitary transformations. In our case we thus need four matrices such that

$$\begin{aligned} V^+_{uL} m_{tq} V_{uR} = D_{tq} = diag\,(m_u, m_c, m_t) \\ V^+_{dL} m_{bq} V_{dR} = D_{bq} = diag\,(m_d, m_s, m_b) \end{aligned}$$

The matrices may be determined if we multiply the equations with its Hermitean conjugate, for example,

$$V^+_{uL} m_{tq} V_{uR} V^+_{uR} m^+_{tq} V_{uL} = V^+_{uL} m_{tq} m^+_{tq} V_{uL} = D^2_{tq} = diag\,(m^2_u, m^2_c, m^2_t)$$

and we see that V_{uL} is the matrix which diagonalizes the Hermitean matrix $m_{tq} m^+_{tq}$. This reduces our problem to a standard eigenvalue problem for Hermitean operators.

and the transformed matrices are required to be **real diagonal** each with 3 free parameters (the 3 independent masses) left out of the 18 we started with. A unitary 3 x 3 matrix has 9 parameters and with the two matrices V_{uR} and V_L we have enough parameters to satisfy the 15 conditions to make G_{tq} real diagonal. However, the remaining 3 parameters not used up from V_L plus the 9 of V_{dR} are not sufficient to diagonalize G_{bq}. We thus have to transform the (by convention) lower components of the doublets by an independent unitary transformation:

$$d_{jL} \quad \rightarrow \quad \tilde{d}_{jL} = (U_{\mathrm{CKM}})_{jk}\, d_{kL}$$

This however changes the form of $\mathcal{L}_{\mathrm{matter}}$ and generates a family–mixing in the charged current. This leads us to the form of the charged current

$$J_\mu^{CC} = (\bar{u}, \bar{c}, \bar{t})\, \gamma_\mu\, (1-\gamma_5)\, U_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

given earlier with the unitary 3×3 matrix

$$U_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} .$$

Part of the parameters are used up to diagonalize

$$m_{bq} \quad \rightarrow \quad U^{+}_{\mathrm{CKM}}\, V_L^{+}\, m_{bq}\, V_{dR} \quad \text{real diagonal} \;\; .$$

After diagonalization we always assume the fields to be relabeled such that the fields with identical quantum numbers are ordered with increasing mass:

$$m_u \leq m_c \leq m_t \;\; \text{and} \;\; m_d \leq m_s \leq m_b \;\; .$$

How many observable parameters are there left ? The phase of a charged field is not observable. In the charged current only left–handed fields are present. If we change the phases of the left–handed fields

$$q_L \rightarrow e^{i\phi_q} q_L \; , \;\; \phi_q \quad \text{an arbitrary real number}$$

the CKM-matrix changes according to

$$V \rightarrow \begin{pmatrix} e^{-i\phi_u} & 0 & 0 \\ 0 & e^{-i\phi_c} & 0 \\ 0 & 0 & e^{-i\phi_t} \end{pmatrix} \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} e^{i\phi_d} & 0 & 0 \\ 0 & e^{i\phi_s} & 0 \\ 0 & 0 & e^{i\phi_b} \end{pmatrix}$$

and thus

$$V_{u_i d_j} \rightarrow \exp\, -i\, (\phi_{u_i} - \phi_{d_j})\, V_{u_i d_j} \quad .$$

This generalizes to any number N of families. If all of the 2N fields are transformed with the same phase $V_{u_i d_j}$ is not affected and hence there are 2N-1 phases which may be transformed away by rephasing the left–handed fields. This means that 2N-1 phases of $U_{\rm CKM}$ are not measurable. This is only true if the remaining parts of the Lagrangian are not affected. The neutral current is not affected by this rephasing because it is diagonal in flavor and handedness. The Yukawa Lagrangian obviously is not invariant under left–handed rephasing because it necessarily connects left–handed and right–handed fields. However, we have the phases of the right–handed fields at our disposal. We may choose these phases to be the same as the one's of the left–handed fields (for each individual flavor) such that $\mathcal{L}_{\rm Yukawa}$ remains invariant.

We may count now the number of free parameters which affect the physics. Let us consider N families. A unitary N $\times$ N matrix has N^2 parameters. We may compare it with an orthogonal N $\times$ N matrix, describing a rotation in N dimensional Euclidean space and exhibiting N(N-1)/2 parameters, which may be taken to be the Euler angles. We may parametrize the unitary matrix by the N(N-1)/2 Euler angles, describing the rotations, plus N^2-N(N-1)/2 phases. As we have argued before 2N-1 of these phases are unobservable. Thus we end up with (N-1)(N-2)/2 physical phases:

number of families N	number of angles N(N-1)/2	number of phases (N-1)(N-2)/2
2	1	0
3	3	1
4	6	3

For N=3 we have 4 relevant parameters. Besides 3 (real) 3-dimensional rotations there remains one phase undetermined. Thus $U_{\rm CKM}$ may be parametrized in terms of 3 rotation angles and one phase[36]:

$$U_{\rm CKM} = \left(\begin{array}{ccc} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23}-c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23}-s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}c_{23}-c_{12}s_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23}-s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{array} \right) .$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ with i and j being the family labels. Without loss of generality one may assume all c_{ij} and s_{ij} to be positive and the phase δ_{13} to lie in the range $0 \leq \delta_{13} < 2\pi$.

A non-degenerate phase

$$\delta_{13} \neq 0\;,\;\;\pi$$

leads to complex effective couplings in the charge changing current which violate *CP*-invariance. It is important to notice that this specific kind of "standard model *CP*-violation" is only possible for more than two families. In fact, the mixing matrix has to satisfy a number of conditions in order that *CP*-violation occurs. The basic observation is that a unitary 3 by 3 matrix which has a zero matrix element somewhere is necessarily real. Therefore mixing must be non-degenerate:

$$\theta_{ij} \neq 0\;,\;\;\frac{\pi}{2}\quad ij=12,23,13\;\;.$$

[36]There are many equivalent parametrizations, here we present the one advocated by the Particle Data Group.

For the same reason, CP-violation can only occur this way if all the states are distinguishable. This means that **all states with the same charge must have different masses**:

$$m_u < m_c < m_t \quad \text{and} \quad m_d < m_s < m_b \ .$$

which happens to be so in Nature. Otherwise, suppose the s and the b quarks would have the same mass, for example, then the Lagrangian would be invariant under $U(2)$ in (s,b)-flavor space. If we perform the $U(2)$ rotation

$$\begin{pmatrix} s' \\ b' \end{pmatrix} = \frac{1}{X} \begin{pmatrix} V_{us} & V_{ub} \\ |V_{ub}|e^{i\delta'} & |V_{us}|e^{i\delta} \end{pmatrix} \begin{pmatrix} s \\ b \end{pmatrix}$$

with $X = \sqrt{|V_{us}|^2 + |V_{ub}|^2}$ and the phases constraint by unitarity, $\delta' = \delta + \delta_{us} - \delta_{ub} + \pi$ we obtain

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & X & 0 \\ V_{cd} & V'_{cs} & V'_{cb} \\ V_{td} & V'_{ts} & V'_{tb} \end{pmatrix} \begin{pmatrix} d \\ s' \\ b' \end{pmatrix}$$

the u-quark only couples to d and s' but no longer to the b' such that the CKM-matrix has a zero element and hence would be real.

Due to unitarity, there is no mixing effect in the neutral current. In the weak interaction basis the neutral current is diagonal in helicities, flavors and families and the unitary CKM-transformation

$$\sum_{i=1}^{3} \bar{d}_{iL} \cdots d_{iL} \ \to \ \sum_{i=1}^{3} \bar{\tilde{d}}_{iL} \cdots \tilde{d}_{iL} \equiv \sum_{i=1}^{3} \bar{d}_{iL} \cdots d_{iL}$$

leaves its form invariant. This is called the GIM-mechanism explaining the absence of flavor-changing neutral currents (FCNC). In fact, in order to explain the absence of FCNC's, Glashow, Iliopoulos and Maiani had to propose, in 1970, the existence of a fourth quark, the charm quark c as a doublet partner of the s quark. At that time only three quarks where known.

The discovery of the charmonium ($c\bar{c}$) state J/ψ in 1974 revealed the completeness of the 2nd family with the charm quark c. The first 3rd family member showed up in 1975 with the discovery of the τ . With the observation of the bottonium ($b\bar{b}$) state Υ the existence of the b quark could be established.

First indications of an unexpectedly heavy top quark came from the first observation of $B^0 \leftrightarrow \overline{B^0}$ oscillations by the ARGUS Collaboration in 1987 [**?**]. LEP precision measurements together with SM fits in 1995 had constrained the top mass to

$$m_t = 170 \pm 10^{+17}_{-19} \text{ GeV}$$

assuming a Higgs mass in the range 60 GeV $< m_H <$ 1000 GeV. Shortly after in 1995 the top was discovered at the Tevatron (CDF and D0 at Fermilab) by $t\bar{t}$–production [**?**]. More recent top mass measurement at the Tevatron [**?**] determined the rather precise result

$$m_t = 172.6 \pm 1.4 \text{ GeV}\,, \tag{13.2}$$

in excellent agreement with the final indirect determination $m_t = 172.3^{+10.2}_{-7.6}$ GeV from the LEP precision measurements of Z resonance parameters [**?**].

Note about neutrino masses and mixings: we notice that, according to Tab. **??**, the right-handed neutrinos (now established to exist in reality) are *sterile*. They do not carry any gauge charge and hence, in contrast to all other particles, do not interact via the spin 1 gauge bosons with the rest of the world. Hence, the right-handed neutrinos could be absent altogether. This would imply the leptonic CC to exhibit some very special properties: if $\nu_{\ell R}$ would not exist, then $m_{\nu_\ell} = 0$ and lepton numbers L_ℓ would be strictly conserved individually for each flavor $\ell = e, \mu, \tau$. For a long time this seemed to be supported by experiments. Today we know that this is true approximately only, although lepton-number violating processes such as $\mu \to e\gamma$ still are expected to have extremely small probability (see below). The observed neutrino mixing implies that neutrino masses must be non-vanishing and non-degenerate. Indeed, for small neutrino masses, the lepton-number violation is expected to be seen first in neutrino oscillations, which have been subject of extensive experimental searches (Davis since 1968, ... , Super-Kamiokande 1998-2001, SNO 2002, KamLAND 2003) [**?**, **?**, **?**, **?**, **?**].

We summarize the following **important consequences:**

- i) all masses of quarks, leptons and neutrinos are independent
- ii) the coupling of the Higgs boson to the fermions is universally proportional to each fermion mass, for bosons proportional to the square of each boson mass
- iii) there is quark flavor violation in charge exchange weak interactions; similarly, the different neutrino flavors mix in charged current interactions
- iv) the phase in $U_{\rm CKM}$ is CP-violating and thus potentially capable of explaining the observed CP-violation in K-decays (Cronin and Fitch 1964). At least 3 families are needed to "explain" CP-violation in this way. In fact it predicts $O(1)$ (in units of Cabibbo's $\lambda = \sin\theta_c$) CP-violation in the B-system, which has been fully confirmed experimentally[37].
- v) flavor is conserved in neutral currents at tree level (GIM mechanism). This is strikingly supported by experiment, at least for the light flavors.
- vi) neutral currents are all CP-conserving at tree level.

These "predictions" of the SM must be considered a great success. Many extensions of the SM encounter difficulties with flavor changing and CP-violation in neutral currents.

The leptonic CC has some very special properties, which derive from the absence of right-handed neutrinos in them. Among the unsolved neutrino-puzzles we mention: Why are neutrino masses so small (see saw mechanism?)? Do neutrinos have unusual magnetic moments? Are there neutrinos which are their own antiparticles (Majorana neutrinos)?

The properties of the weak currents have been established in a long history which started with Fermi in 1934. Here, we only mention some more recent of the fundamental experimental tests:

- V-A structure of the CC:
  μ-decay provides the most sensitive clean direct tests for right-handed currents (e.g. $SU(2)_R \otimes SU(2)_L \otimes U(1)_{B-L}$ extension of the SM). The best limit for the transition amplitude is

$$\frac{A_{V+A}}{A_{V-A}} < 0.029 \quad (90\% CL)$$

[37] The particle-antiparticle mixing of the neutral kaons $K^0 \leftrightarrow \bar{K}^0$ (Gell-Mann und Pais 1955) played a key role in revealing CP violation as an observable effect. In the B-meson system $B^0 \leftrightarrow \bar{B}^0$ mixing plays an analogous role. CP violation in the B-system was found to be precisely as predicted by CKM mixing (BABAR at SLAC, Belle at KEK 2001): $\sin(2\phi_1) = 0.78 \pm 0.08$ where $\phi_1 = \arg(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*})$ [**?**]. For its prediction Kobayashi and Maskawa have been awarded the Noble Pize in 2008 [**?**].

- absence of flavor-changing NC at tree level:

$$\begin{array}{rcc}\Gamma(K_L \to \mu^+\mu^-)/\Gamma(K_L \to all) & = & (9.5^{+2.4}_{-1.5}) \times 10^{-9} \\ \Gamma(D^0 \to \mu^+\mu^-)/\Gamma(D^0 \to all) & < & 1.1 \times 10^{-5} \\ \Gamma(B^0 \to e^+e^-)/\Gamma(B^0 \to all) & < & 3 \times 10^{-5}\end{array}$$

  Flavor-changing NC processes are allowed in higher orders (rare processes).

- special properties of the lepton current:
  The lepton numbers L_ℓ ($\ell = e, \mu, \tau$) are other additive quantum numbers which seem to be strictly conserved at first sight. By convention $L_\ell(\ell^-) = 1$. That L_μ is separately conserved follows from the non-observations of the decays

$$\begin{array}{cc}\mu^+ \to e^+ + \gamma & \Gamma(\mu \to e\gamma)/\Gamma(\mu \to \mathrm{all}) < 1.2 \times 10^{-11} \\ \mu^+ \to e^+ + e^- + e^+ & \Gamma(\mu \to 3e)/\Gamma(\mu \to \mathrm{all}) < 1.0 \times 10^{-12} \\ K_L \to e + \mu & \Gamma(K_L \to e\mu)/\Gamma(K_L \to \mathrm{all}) < 4.7 \times 10^{-12} \\ K^+ \to \pi^+ + e + \mu & \Gamma(K^+ \to \pi^+ e\mu)/\Gamma(K^+ \to \mathrm{all}) < 2.1 \times 10^{-10} \\ \mu^- + (Z,A) \to e^- + (Z,A) & \Gamma(\mu^-\mathrm{Ti} \to e^-\mathrm{Ti})/\Gamma(\mu^-\mathrm{Ti} \to \mathrm{all}) < 4.0 \times 10^{-12} \\ \mu^- + (Z,A) \to e^+ + (Z-2,A) & \Gamma(\mu^-\mathrm{Ti} \to e^+\mathrm{Ca})/\Gamma(\mu^-\mathrm{Ti} \to \mathrm{all}) < 3.6 \times 10^{-11}\end{array} .$$

  Tests of the separate conservation of L_τ are much less stringent: The best limits are:

$$\Gamma(\tau \to e\gamma)/\Gamma(\tau \to \mathrm{all}) < 2.7 \times 10^{-6} \quad \text{and} \quad \Gamma(\tau \to \mu\gamma)/\Gamma(\tau \to \mathrm{all}) < 1.1 \times 10^{-6} .$$

  Within the experimentally well established electroweak standard model strict lepton number conservation is only possible if the neutrinos are strictly massless. Non-vanishing neutrino masses lead to **neutrino-oscillations**. Neutrino mixing searches (ν-oscillations $\nu_\ell \leftrightarrow \nu_{\ell'}$) have confirmed the effect recently which implies the existence of non–vanishing neutrino masses. Present direct upper limits on the neutrino masses are:

$$\begin{array}{lcl}m_{\nu_e} < & 3.0 \text{ eV} & (\text{from } {}^3H \to {}^3He\, e^-\, \bar\nu_e) \\ m_{\nu_\mu} < & 190 \text{ keV} & (\text{from } \pi \to \mu\, \nu_\mu) \\ m_{\nu_\tau} < & 18.2 \text{ MeV} & (\text{from } \tau^- \to 3\pi\, \nu_\tau)\end{array}$$

  Lower bounds are not yet so easy to establish at present but observed neutrino mixing phenomena indicate values of about two to three orders of magnitude lower than the above direct upper limits. In any case this implies corresponding lepton numbers L_ℓ ($\ell = e, \mu, \tau$)–violations.

  Another important limit is the absence of $\Delta L_e = 2$ transitions. The limit from neutrinoless double beta decay $(Z,A) \to (Z+2,A) + e^+ + e^-$ is $t_{1/2} > 1.6 \times 10^{25}$ years for ^{76}Ge . The observation of such reactions would imply that the electron-neutrino is a massive **Majorana neutrino**, a self-conjugate fermion which is its own antiparticle.

Neutrino mixing (ν-oscillations $\nu_\ell \leftrightarrow \nu_{\ell'}$) turned out to be very hard to establish. Ater decades of intense research finally neutrino oscillations were confirmed [**?**, **?**, **?**, **?**, **?**], which means that neutrinos have tiny non–degenerate masses and mix [**?**] similar to the quarks, although the mixing pattern looks very different from quark-mixing[38]. This also requires the right–handed singlet neutrinos to exist, which, however, are sterile with respect to any gauge interaction. In 2002 R. Davis Jr. and M. Koshiba awarded the Nobel Prize for pioneering contributions to astrophysics, in particular for the detection of cosmic neutrinos.

At about the same time, when the neutrino puzzle was resolved, the measurements of direct CP-violation (ε') in the K-meson system [CP-LEAR[NA48]/CERN (2002), kTEV/FNL (2008)] and CP-violation in the B-meson system [B factories BaBar/SLAC and Belle/KEK (2001)] was experimentally proven to follow excatly the pattern of the Kobayashi-Maskawa 3–family quark mixing and the resulting CP violation mechanism [**?**]. M. Kobayashi and T. Maskawa were awarded the Nobel Prize in 2008, for their prediction (made in 1973) of $O(1)$ (in units of Cabibbo's $\lambda = \sin\theta_c$) CP-violaton in the B–meson system.

13.7 Flavor mixing pattern

Historically flavor mixing was "observed" first by a comparison of the decays $K \to \mu\nu$ and $\pi \to \mu\nu$. Only u, d and s flavors (isospin and strangeness) were known at that time and the approximate $SU(3)_{flavor}$ symmetry was established. This symmetry is substantially broken by mass splittings within $SU(3)$ multiplets like for the pseudoscalar mesons with $m_K \simeq 494$ MeV and $m_\pi \simeq 140$ MeV. Hadronic transition matrix-elements however satisfy $SU(3)$ relations quite well. Denoting the matrix element between the pseudoscalar meson P and the weak hadronic current $h^\mu(x)$ by

$$< 0|h^\mu(0)|P(p) >= ip^\mu f_P$$

one obtains for the ratio of the decay widths

$$\frac{\Gamma(K \to \mu\nu)}{\Gamma(\pi \to \mu\nu)} = \frac{m_K}{m_\pi}\left(\frac{1-m_\mu^2/m_K^2}{1-m_\mu^2/m_\pi^2}\right)^2\left(\frac{f_K}{f_\pi}\right)^2 \simeq 1.3$$

and thus

$$\left(\frac{f_K}{f_\pi}\right)^2 \simeq 0.075$$

and not $O(1)$ as suggested by approximate $SU(3)$ symmetry! Cabibbo solved this puzzle by noting that the strangeness conserving $\Delta S = 0$ part and the strangeness changing $\Delta S = 1$ part of the hadronic current mix in a specific way, described by a rotation:

$$h^\mu = h^\mu_{\Delta S=0}\cos\theta_c + h^\mu_{\Delta S=1}\sin\theta_c$$

such that the effective couplings for the two processes are

$$\begin{aligned} K \to \mu\nu \quad &: \quad G_F \sin\theta_c \\ \pi \to \mu\nu \quad &: \quad G_F \cos\theta_c \end{aligned}$$

[38]Present results may be summarized as follows: **a)** Solar neutrinos: $\Delta m^2_{12} \approx (7)\times 10^{-5}$ eV^2, $\tan^2\Theta_{12} \approx 0.4$ (large $\nu_e \leftrightarrow \nu_\mu$ mixing), $\sin^2 2\Theta_{13} < 0.067$. **b)** Atmospheric neutrinos: $\Delta m^2_{23} \approx (1.3-3.0)\times 10^{-3}$ eV^2, $\sin^2 2\Theta_{23} > 0.9$ (large angle $\nu_\mu \leftrightarrow \nu_\tau$ mixing). Main features are:
- smallness of ν masses: $m_\nu < 1-2$ eV, at least for one mass $m_\nu > \sqrt{\Delta m^2_{23}} > 0.04$ eV,
- hierarchy of Δm^2's : $|\Delta m^2_{12}/\Delta m^2_{23}| = 0.01-0.15$,
- no strong hierarchy of masses: $|m_2/m_3| > |\Delta m_{12}/\Delta m_{23}| = 0.18^{+0.22}_{-0.08}$,
- bi–large or maximal mixing between neighboring families (1-2) and (2-3),
- small mixing between remote families (1-3),

in any case $m_\nu \ll m_\ell, m_q$.

with $\sin\theta_c \simeq 0.22$ and thus

$$\left(\frac{f_K}{f_\pi}\right)^2 \to \tan^2\theta_c \left(\frac{f_K}{f_\pi}\right)^2 \simeq 0.0795 \left(\frac{f_K}{f_\pi}\right)^2$$

such that the $SU(3)$ relation $f_K = f_\pi$ is satisfied quite well. Of course if one uses the above ratio to fix the Cabibbo angle one has to consider other processes in order to see whether the above hypothesis makes sense or not.

Similarly for the Baryons. $\underline{\beta\text{–decay}}$: $G_n = G_F \cos\theta_c$, $\underline{\Lambda\text{–decay}}$: $G_\Lambda = G_F \sin\theta_c$.

In addition, as a result of CVC, for $\underline{\mu\text{–decay}}$: $G_\mu = G_F$!

As a result, Cabibbo universality works for $\Delta S = 0$ and $\Delta S = 1$ transitions, however, not for $\Delta S = 2$ transitions, which require the c–quark!

As a next step Glashow, Iliopoulos and Maiani introduces the c quark in order to explain the absence of FCNC's and thus for the first time considered a 2 family world:

$$\mathcal{L}^{CC} = \frac{g}{2\sqrt{2}}(\bar{u},\bar{c})\,\gamma_\mu\,(1-\gamma_5)\,U_{\mathrm{CKM}} \begin{pmatrix} d \\ s \end{pmatrix} W^\mu$$

with the unitary 2×2 matrix

$$U = \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} .$$

For N=2 the quark mixing matrix is automatically real and given by a simple rotation, the Cabibbo rotation matrix

$$\begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix}$$

In the 2 family world the hadronic currents are:

$$\begin{aligned}
\text{CC:}\quad J_\mu^+ &= \bar{u}\gamma_\mu\,(1-\gamma_5)\,\underbrace{(\;d\cos\theta_c + s\sin\theta_c)}_{\tilde{d}} && \text{Cabibbo} \\
&+\;\; \bar{c}\gamma_\mu\,(1-\gamma_5)\,\underbrace{(-d\sin\theta_c + s\cos\theta_c)}_{\tilde{s}} && \text{GIM piece} \\
\text{NC:}\quad J_\mu^Z &= \bar{u}\gamma_\mu\,(v_u - a_u\gamma_5)\,u + \underbrace{\bar{\tilde{d}}\gamma_\mu\,(v_d - a_d\gamma_5)\,\tilde{d}}_{\text{FCNC}} \\
&+\;\; \bar{c}\gamma_\mu\,(v_u - a_u\gamma_5)\,c + \bar{\tilde{s}}\gamma_\mu\,(v_d - a_d\gamma_5)\,\tilde{s} && \text{GIM piece} \\
&= \bar{u}\gamma_\mu\,(v_u - a_u\gamma_5)\,u + \bar{d}\gamma_\mu\,(v_d - a_d\gamma_5)\,d \\
&+\;\; \bar{c}\gamma_\mu\,(v_u - a_u\gamma_5)\,c + \bar{s}\gamma_\mu\,(v_d - a_d\gamma_5)\,s\;\; .
\end{aligned}$$

Without the c quark $\tilde{s}$ would be absent in the CC and if one assumes that in the NC only the fields already present in the CC enter one ends up with a flavor changing NC. Although NC's had not been observed at all (before 1973) such FCNC's would have had observable consequences. The N=2 mixing scheme sometimes is called *Cabibbo universality.* Due to the existence of a third family Cabibbo universality is violated, because the 2 by 2 sub-matrix of the CKM-matrix is not unitary. A comparison of the N=2 and the N=3 mixing schemes in the 2 family world yields:

U	N=2	N=3
V_{ud}	$\cos\theta_c$	c_{12}
V_{us}	$\sin\theta_c$	s_{12}
V_{cd}	$-\sin\theta_c$	$-s_{12}c_{23}$
V_{cs}	$\cos\theta_c$	$c_{12}c_{23}$

where we used the excellent approximation $c_{13} = 1$, as c_{13} is known to deviate from unity only in the fifth decimal place.

The N=2 mixing scheme was extended to N=3 by Kobayashi and Maskawa in 1973 in order to incorporate CP-violation in a natural way.

Empirically the CKM matrix elements have the approximate moduli

$$|V| \simeq \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}$$

with $\lambda \simeq \sin\theta_c \simeq 0.22$ given by the sine of the Cabibbo angle. This suggests the Wolfenstein parametrization (by unitarization up to higher order terms)

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3\,(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3\,(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4)$$

where $A \sim 1$ and $\rho^2 + \eta^2 < 1$. The corresponding quark decay pattern is illustrated in the following diagram:

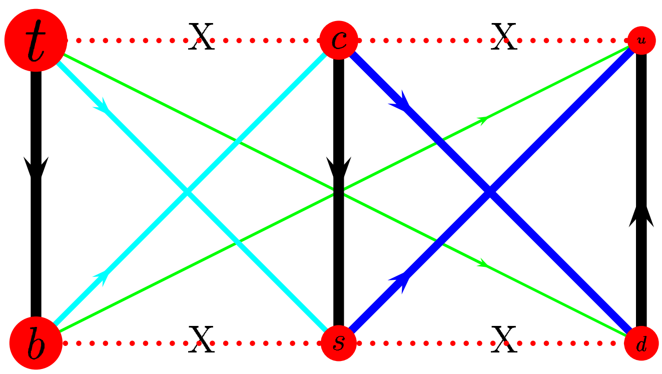


Figure 13.11: The CKM mixing hierarchy (**??**). FCNCs at tree level are forbidden [X].

Note: the u quark is stable, the s and b quarks are metastable. Flavor changing neutral current transitions are allowed only as second (or higher) order transitions: e.g. $b \to s$ is in fact $b \to (t^*, c^*, u^*) \to s$, where the asterix indicates "virtual transition".

Exercises: Section 13

① The left (L)- and right (R)-handed fields are given by

$$\psi_L = \frac{1-\gamma_5}{2}\psi\ , \quad \psi_R = \frac{1+\gamma_5}{2}\psi\ .$$

Show that a mass term is given by

$$\bar{\psi}\psi = \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L\ .$$

Discuss helicity mixing for the terms

$$S = \bar{\psi}\psi\ , \quad P = \bar{\psi}\gamma_5\psi\ , \quad V^\mu = \bar{\psi}\gamma^\mu\psi\ , \quad A^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi\ , \quad T^{\mu\nu} = \bar{\psi}\sigma^{\mu\nu}\psi$$

② Consider two free fields ψ_1 and ψ_2 with different masses m_1 and m_2 . Calculate the divergences of the vector and axial vector currents

$$V^\mu = \bar{\psi}_1\gamma^\mu\psi_2 \quad \text{and} \quad A^\mu = \bar{\psi}_1\gamma^\mu\gamma_5\psi_2$$

$$\partial_\mu V^\mu = ? \qquad\qquad \partial_\mu A^\mu = ?$$

What is the consequence of this result for the weak currents ? (Hint: Use the Dirac equations)

③ Find the form of the quark currents which couple to W and Z and the photon. As a starting point use the $SU(2)_L \otimes U(1)_Y$ covariant derivative form

$$\begin{aligned}\mathcal{L}_q &= \bar{b}_R i\gamma^\mu\left(\partial_\mu + i\frac{1}{3}g'B_\mu\right)b_R + \bar{t}_R i\gamma^\mu\left(\partial_\mu - i\frac{2}{3}g'B_\mu\right)t_R \\ &\quad + \bar{L}_q i\gamma^\mu\left(\partial_\mu - i\frac{1}{3}\frac{g'}{2}B_\mu - ig\frac{\tau_a}{2}W_{\mu a}\right)L_q\end{aligned}$$

Check the correctness of the covariant derivatives with the quantum numbers assigned to the quarks. We have denoted by $L_q = \binom{t}{b}_L$ the left–handed doublet. t_R and b_R are the corresponding right–handed singlets. Quark mixing should be ignored.

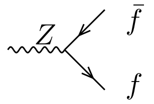
④ Work out the Yang-Mills couplings for $SU(2)_L$ in terms of the physical fields $W^\pm, Z$ and A .

⑤ Give numerical values for the Higgs couplings to the various particles. Calculate the decay width of the Higgs into a fermion pair and estimate the branching ratios for the different flavors. Discuss the dependence of the Higgs mass for the range $0 < m_H < 2M_W$. Compare the Higgs width with the width of known particles.

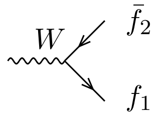
14 Physics at the Z resonance

14.1 Production and Decay of the Weak Vector Bosons

At lowest order, production and decay of massive vector bosons are described by the Born diagrams



$$:= \quad \frac{g}{2\cos\Theta_W}\bar{f}_{\bar{c}\alpha}\left(\gamma^\mu(v_f - a_f\gamma_5)\right)_{\alpha\beta} f_{c\beta}\,\delta_{\bar{c}c}$$



$$:= \quad \frac{g}{2\sqrt{2}}\,V_{12}\,\bar{f}_{2\bar{c}\alpha}\left(\gamma^\mu(1-\gamma_5)\right)_{\alpha\beta} f_{1_{c\beta}}\,\delta_{\bar{c}c}\ , \tag{14.1}$$

where V_{12} denotes the Cabibbo-Kobayashi-Maskawa mixing matrix. We have explicitly indicated the color and spinor indices. Z and $W^\pm$ production and decay may be described by *one* general vertex $\hat{g}\bar{f}_2\gamma^\mu(v - a\gamma_5)f_1$. The production cross section for unpolarized beams in the zero-width approximation is

$$f_1(p_1) + \bar{f}_2(p_2) \to V(p)$$

$$\sigma(f_1 + \bar{f}_2 \to V) = \frac{1}{N_{c_1}}\frac{1}{N_{c_2}}\frac{1}{(2s_1+1)}\frac{1}{(2s_2+1)}\,\pi\,\delta\left((p_1+p_2)^2 - M_V^2\right)\cdot\frac{\sum|T_{12}|^2}{M_V^2} \tag{14.2}$$

where T_{12} is the transition matrix element and the sum extends over color (c_i) and spin (s_i) of initial and final state particles. With respect to the initial states, the cross section is determined by the color and spin averages of $|T_{12}|^2$.

Equation (14.2) follows from Eq. (3.35)

$$d\sigma = \frac{(2\pi)^4\delta^{(4)}\left(P_f - P_i\right)}{2\sqrt{\lambda\left(s, m_1^2, m_2^2\right)}}|\,T_{12}\,|^2 d\mu(p)$$

by taking the spin (color) average over the initial spins (colors) and sum over the final spins (colors) $|\,T_{12}\,|^2 \to \frac{1}{4}\sum_{s_1 s_2 \lambda}|\,T_{12}\,|^2$ and integration over the phase space. If we neglect the electron mass ($m_e^2 << M_Z^2, s = (p_1+p_2)^2$), the total unpolarized cross–section for Z–production in e^+e^-–annihilation is

$$\begin{aligned}\sigma(e^+e^- \to Z)\,|_{s\simeq M_Z^2} &= \frac{1}{4}\int\frac{d^3p}{(2\pi)^3\,2E_Z}\frac{(2\pi)^4\delta^{(4)}(p_Z - p_1 - p_2)}{2s}\sum|\,T_{12}\,|^2 \\ &= \frac{2\pi\,\delta(E_Z - E)}{16sE}\sum|\,T_{12}\,|^2 = \frac{1}{4}\frac{\sum|\,T_{12}\,|^2}{M_Z^2}\,\pi\,\delta(s - M_Z^2)\ .\end{aligned}$$

In the c.m. frame $E_Z = M_Z$ and $E = E_1 + E_2 = \sqrt{s}$. Furthermore, we used the following properties of the δ–function: $\delta(x) = \delta(-x)$ and $\delta(ax) = 1/a\,\delta(x)$ with $x = E - M_Z$ and $a = 2E = 2M_Z = E + M_Z$ for $E = M_Z$. Hence $(E - M_Z)(E + M_Z) = s - M_Z^2$.

The decay width is determined by the same $\sum|T_{12}|^2$, since the processes are related by crossing, and given by

$$V(p) \to f_1(p_1) + \bar{f}_2(p_2)$$

$$\Gamma(V \to f_1 + \bar{f}_2) = \frac{1}{(2s_V+1)} \frac{M_V}{16\pi} \frac{2|\vec{p}\,|}{M_V} \frac{\sum |T_{12}|^2}{M_V^2} \tag{14.3}$$

where $\vec{p}$ is the decay momentum of a fermion in the center of mass frame.

Equation (14.3) follows from Eq. (3.38)

$$d\Gamma = \frac{(2\pi)^4 \delta^{(4)}(P_f - P_i)}{2M} \mid T_{fi} \mid^2 d\mu(p_1) d\mu(p_2)$$

by taking the spin average over the initial spin and sum over the final spins (colors) $|T_{12}|^2 \to \frac{1}{3}\sum_{s_1 s_2 \lambda} |T_{12}|^2$ and by integration of the two–body phase space in the rest frame of the decaying particle:

$$\Gamma \;=\; \frac{(2\pi)^4}{2M} \frac{1}{(2\pi)^6} \frac{1}{3} \int \frac{d^3p_1}{2E_1} \frac{d^3p_2}{2E_2} \delta^{(4)}(M - p_1 - p_2) \sum \mid T_{12} \mid^2 \;.$$

The integration over d^3p_2 is trivial by three momentum conservation and yields $\vec{p}_2 = -\vec{p}_1$ and we denote $|\vec{p}_2\,| = |\vec{p}_1\,| = p$. We then write d^3p_1 in polar coordinates and note that the remaining integrand does not depend on the angles. The angular integration then yields 4π and hence

$$\Gamma \;=\; \frac{1}{2M} \frac{4\pi}{(2\pi)^2} \frac{1}{12} \int\limits_0^\infty \frac{dp\, p^2}{\sqrt{p^2+m_1^2}\sqrt{p^2+m_2^2}} \,\delta\left(M - \sqrt{p^2+m_1^2} - \sqrt{p^2+m_2^2}\right) \sum \mid T_{12} \mid^2 \;.$$

The last integration may be performed by choosing the argument of the δ–function as an integration variable[39]. Thus let

$$f(p) \doteq \sqrt{p^2+m_1^2} + \sqrt{p^2+m_2^2} - M$$

and p_0 be the solution of $f(p_0) = 0$. The derivative with respect to p reads

$$f'(p) = \frac{p}{\sqrt{p^2+m_1^2}} + \frac{p}{\sqrt{p^2+m_2^2}} = \frac{p(\sqrt{p^2+m_1^2} + \sqrt{p^2+m_2^2}\,)}{\sqrt{p^2+m_1^2}\sqrt{p^2+m_2^2}} = \frac{pM}{\sqrt{p^2+m_1^2}\sqrt{p^2+m_2^2}} \;,$$

where in the last step we have used the fact that under the integral we have $p = p_0$ and hence $(\sqrt{p^2+m_1^2} + \sqrt{p^2+m_2^2}) = M$. We note that

$$\frac{dp\, p^2}{\sqrt{p^2+m_1^2}\sqrt{p^2+m_2^2}} \,\delta\left(M - \sqrt{p^2+m_1^2} - \sqrt{p^2+m_2^2}\right) = f'(p)\, dp\, \frac{p}{M}\, \delta(f) = df\, \frac{p_0}{M}\, \delta(f)$$

and thus obtain

$$\Gamma = \frac{1}{24\pi M^2} \sum \mid T_{12} \mid^2 p_0$$

with $p_0 = |\vec{p}\,|$ the magnitude of the three momentum of the decay products in the c.m. frame.

Finally, let us calculate p_0: p_0 is the solution of $\sqrt{p^2+m_1^2} + \sqrt{p^2+m_2^2} = M$. If we square the defining equation we obtain

$$p_0^2 + m_1^2 + p_0^2 + m_2^2 + 2\sqrt{p_0^2+m_1^2}\sqrt{p_0^2+m_2^2} = M^2$$

and squaring

$$2\sqrt{p_0^2+m_1^2}\sqrt{p_0^2+m_2^2} = M^2 - (p_0^2+m_1^2) - (p_0^2+m_2^2)$$

---

[39] As a rule

$$\delta(f(x)) = \sum_{x_0} |f'(x_0)^{-1}|\, \delta(x - x_0)$$

where x_0 are the solutions of $f(x_0) = 0$.

once more, we get

$$4(p_0^2+m_1^2)(p_0^2+m_2^2) = (M^2-(p_0^2+m_1^2)-(p_0^2+m_2^2))^2$$

which may be written as

$$M^4+(m_1^2-m_2^2)^2-2M^2(m_1^2+m_2^2)-4M^2p_0^2 = 0 \ .$$

such that

$$p_0^2 = \left[\left(M^2-(m_1+m_2)^2\right)\ \left(M^2-(m_1-m_2)^2\right)\right]/(4M^2)$$

from which p_0 is trivially obtained. Only the positive solution is physical. Our final result for the width of a two body decay is then given by

$$\boxed{\Gamma = \tfrac{M}{16\pi}\sqrt{1-\tfrac{(m_1-m_2)^2}{M^2}}\sqrt{1-\tfrac{(m_1+m_2)^2}{M^2}}\tfrac{1}{2s_V+1}\tfrac{\sum\mid T_{12}\mid^2}{M^2}} \tag{14.4}$$

This formula exhibits explicitly the kinematic limit for the two body phase space: $M > m_1+m_2$.

The relevant T-matrix element is given by

$$T_{12} = \hat{g}\bar{v}_{2_{\bar{c}\alpha}}(p_2,s_2)(\gamma^\mu(v-a\gamma_5))_{\alpha\beta}u_{1_{c\beta}}(p_1,s_1)\varepsilon^*_\mu(p,\lambda)\delta_{\bar{c}c} \ .$$

The color and spin sums $\sum|T_{12}|^2$ may be calculated using the completeness relations

$$\begin{array}{rcl}
\sum_{s_1} u_{1c_1\alpha_1}(p_1,s_1)\bar{u}_{1\bar{c}_1\beta_1}(p_1,s_1) & = & (\not{p}_1+m_1)_{\alpha_1\beta_1}\ \delta_{\bar{c}_1c_1} \\
\sum_{s_2} v_{2c_2\alpha_2}(p_2,s_2)\bar{v}_{2\bar{c}_2\beta_2}(p_2,s_2) & = & (\not{p}_2-m_2)_{\alpha_2\beta_2}\ \delta_{\bar{c}_2c_2} \\
\sum_\lambda \varepsilon^*_\mu(p,\lambda)\ \varepsilon_\nu(p,\lambda) & = & -g_{\mu\nu}+\frac{p_\mu p_\nu}{M_V^2}
\end{array}$$

and the trace of a product of Kronecker $\delta_{cc'}$ in color space

$$\sum_{\text{color}} = N_{cf} \ .$$

As usual $\not{p} \doteq p_\alpha\gamma^\alpha$. We obtain

$$\begin{array}{rcl}
\sum\mid T_{12}\mid^2 & = & \hat{g}^2\ N_{cf}\cdot\left(-g_{\mu\nu}+\dfrac{p_\mu p_\nu}{M_V^2}\right) \\
 & \cdot & \mathrm{Tr}\,(\not{p}_2-m_2)\ \gamma^\mu\ (v-a\gamma_5)\ (\not{p}_1+m_1)\ \gamma^\nu\ (v-a\gamma_5) \ .
\end{array}$$

For the calculation of the trace we first anticommute γ_5 to the right using $\gamma_5\gamma^\alpha = -\gamma^\alpha\gamma_5$:

$$\begin{array}{rcl}
X & = & (\not{p}_2-m_2)\ \gamma^\mu\ (v-a\gamma_5)\ (\not{p}_1+m_1)\ \gamma^\nu\ (v-a\gamma_5) \\
 & = & \not{p}_2\ \gamma^\mu\ (v-a\gamma_5)\ \not{p}_1\ \gamma^\nu\ (v-a\gamma_5)+\ m_1\ \not{p}_2\ \gamma^\mu\ (v-a\gamma_5)\ \gamma^\nu\ (v-a\gamma_5) \\
 & - & m_2\ \gamma^\mu\ (v-a\gamma_5)\ \not{p}_1\ \gamma^\nu\ (v-a\gamma_5)-m_1m_2\ \gamma^\mu\ (v-a\gamma_5)\ \gamma^\nu\ (v-a\gamma_5) \\
 & = & \not{p}_2\ \gamma^\mu\ \not{p}_1\ \gamma^\nu\ (v-a\gamma_5)^2+\ m_1\ \not{p}_2\ \gamma^\mu\ \gamma^\nu\ (v+a\gamma_5)\ (v-a\gamma_5) \\
 & - & m_2\ \gamma^\mu\ \not{p}_1\ \gamma^\nu\ (v-a\gamma_5)^2-m_1m_2\ \gamma^\mu\ \gamma^\nu\ (v+a\gamma_5)\ (v-a\gamma_5)
\end{array}$$

and use $(v-a\gamma_5)^2\ = v^2+a^2-2va\gamma_5$ and $(v+a\gamma_5)\ (v-a\gamma_5)\ = v^2-a^2$. Taking traces,

$$\begin{array}{rcl}
\mathrm{Tr}\ (\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) & = & 4\,(g^{\mu\nu}g^{\rho\sigma}+g^{\mu\sigma}g^{\nu\rho}-g^{\mu\rho}g^{\nu\sigma}) \\
\mathrm{Tr}\ (\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_5) & = & 4\,i\,\varepsilon^{\mu\nu\rho\sigma} \\
\mathrm{Tr}\ (\gamma^\mu\gamma^\nu) & = & 4\,g^{\mu\nu} \\
\mathrm{Tr}\ (\gamma^\mu\gamma^\nu\gamma^\rho) & = & \mathrm{Tr}\ (\gamma^\mu\gamma^\nu\gamma^\rho\gamma_5)\ =\ 0
\end{array}$$

we find

$$
\begin{aligned}
\mathrm{Tr}\, X &= 4\,(v^2+a^2)\,[p_2^\mu p_1^\nu + p_1^\mu p_2^\nu - g^{\mu\nu}\,(p_1 p_2)] \\
&- \; 8va\, i\, \varepsilon^{\alpha\mu\beta\nu} p_{2\alpha} p_{1\beta} - 4\,(v^2-a^2)\, m_1 m_2\, g^{\mu\nu} \quad .
\end{aligned}
$$

Since this trace is contracted with the symmetric tensor $-g_{\mu\nu} + \frac{p_\mu p_\nu}{M_V^2}$ the ε–tensor term cannot contribute and we obtain, using $g_{\mu\nu} g^{\mu\nu} = g_\nu{}^\nu = 4$,

$$
\begin{aligned}
\sum |\, T_{12}\,|^2 &= \hat{g}^2\, N_{cf} \left\{ 4\,(v^2+a^2)\, \left[\left(2 - \frac{p^2}{M_V^2}\right)\,(p_1 p_2) \,+ 2\,\frac{(pp_1)\,(pp_2)}{M_V^2} \right] \right. \\
&+ \left. 4\,(v^2-a^2)\, m_1\, m_2\, \left(4 - \frac{p^2}{M_V^2}\right) \right\} \quad .
\end{aligned}
$$

We now use the kinematic relations: $p = p_1 + p_2$, $p^2 = p_1^2 + 2p_1p_2 + p_2^2 = s$. We allow the vector boson to be off-shell, such that $p^2 = s \neq M_V^2$. The on-shell conditions for the fermions read: $p_1^2 = m_1^2$ and $p_2^2 = m_2^2$ and we get

$$
p_1 p_2 = \frac{1}{2}\,(s - m_1^2 - m_2^2)\ , \quad pp_1 = \frac{1}{2}\,(s + m_1^2 - m_2^2)\ , \quad pp_2 = \frac{1}{2}\,(s - m_1^2 + m_2^2)
$$

for the scalar products. Thus

$$
\begin{aligned}
\frac{\sum |T_{12}|^2}{M_V^2} &= 4\hat{g}^2 N_{cf} \left\{ (v^2+a^2) \left(\tfrac{s}{M_V^2} - (2 - \tfrac{s}{M_V^2}) \tfrac{m_1^2+m_2^2}{2M_V^2} - \tfrac{(m_1^2-m_2^2)^2}{2M_V^4} \right) \right. \\
&\left. + (v^2-a^2)(4 - \tfrac{s}{M_V^2}) \tfrac{m_1 m_2}{M_V^2} \right\} \ .
\end{aligned}
$$

The transition matrix element $\sum |\, T_{12}\,|^2$ for on-shell vector bosons $s = (p_1+p_2)^2 = M_V^2$ is given by

$$
\begin{aligned}
\frac{\sum |T_{12}|^2}{M_V^2} &= 4\hat{g}^2 N_{cf} \left\{ (v^2+a^2) \left(1 - \tfrac{m_1^2+m_2^2}{2M_V^2} - \tfrac{(m_1^2-m_2^2)^2}{2M_V^4} \right) \right. \\
&\left. + (v^2-a^2)\, 3\, \tfrac{m_1 m_2}{M_V^2} \right\} \ .
\end{aligned}
$$

The coupling $\hat{g} = (1/\cos\Theta_W, 1)g/2$ may be written in the form

$$
\hat{g}^2 = \frac{1}{v^2} \hat{g}^2 v^2 = \sqrt{2}\, G_\mu\, M_V^2
$$

since $\sqrt{2} G_\mu = 1/v^2$, $g^2 v^2 = 4M_W^2$ and $g^2 v^2/\cos^2\Theta_W = 4M_Z^2$.

For $V = Z$ with $f_1 = f_2 = f$ we have

$$
\hat{g}^2 = \frac{1}{v^2} \frac{g^2 v^2}{4\cos^2\Theta_W} = \sqrt{2}\, G_\mu\, M_Z^2
$$

$$
\begin{aligned}
v_f &= -2Q_f\, \sin^2\Theta_W + T_{3f} \;=\; -2Q_f \sin^2\Theta_W \pm \frac{1}{2} \\
a_f &= T_{3f} \;=\; \pm\frac{1}{2}
\end{aligned}
$$

and hence, using $2Q_f\, T_{3f} = |Q_f|$,

$$
\begin{aligned}
(v_f^2 + a_f^2) &= \frac{1}{2} - 2\,|Q_f| \sin^2\Theta_W + 4\, Q_f^2 \sin^4\Theta_W \\
(v_f^2 - a_f^2) &= \qquad - 2\,|Q_f| \sin^2\Theta_W + 4\, Q_f^2 \sin^4\Theta_W \quad .
\end{aligned}
$$

For $V = W^{\pm}$ and $f_1 \neq f_2$ we have simply

$$\hat{g}^2 = \frac{1}{v^2}\frac{g^2 v^2}{4} = \sqrt{2}\, G_\mu\, M_W^2$$

$$v = a = \frac{1}{\sqrt{2}}$$

and therefore

$$(v^2 + a^2) = 1 \ , \quad (v^2 - a^2) = 0 \ .$$

As a result the formula for the partial widths are given by

$$\begin{aligned} \Gamma_{V \to f_1 \bar{f}_2} &= \tfrac{\sqrt{2} G_\mu M_V^3}{12\pi} \tfrac{2|\vec{p}|}{M_V} N_{cf} \left\{ (v^2 + a^2) \left(1 - \tfrac{1}{2}\tfrac{(m_1^2 + m_2^2)}{M_V^2} - \tfrac{1}{2}\tfrac{(m_1^2 - m_2^2)^2}{M_V^4} \right) \right. \\ & \left. + (v^2 - a^2) 3 \tfrac{m_1 m_2}{M_V^2} \right\} \end{aligned} \tag{14.5}$$

or, for light fermions $m_i \ll M_V$, $|\vec{p}| \simeq \frac{M_V}{2}$,

$$\Gamma_{W \to f_1 \bar{f}_2} = \tfrac{\sqrt{2} G_\mu M_W^3}{12\pi} N_{cf} |V_{12}|^2 \ ; \ \Gamma_{Z \to f \bar{f}} = \tfrac{\sqrt{2} G_\mu M_Z^3}{12\pi} N_{cf} (v_f^2 + a_f^2) \tag{14.6}$$

where we have reinserted the Cabibbo-Kobayashi-Maskawa matrix element V_{12} for the charged current. Notice that for $\sin^2 \Theta_W = 0$ (i.e. $g' = 0$) $M_Z = M_W$ and

$$\Gamma_{W \to f_1 \bar{f}_2} = \Gamma_{Z \to f_1 \bar{f}_1} + \Gamma_{Z \to f_2 \bar{f}_2} \ .$$

Table 14.1 Lowest order predictions for Γ_W and Γ_Z for $\sin^2 \Theta_W = 0.23$, $M_W = 80.19(32)$ GeV and $M_Z = 91.174(21)$ GeV.

$W \to f_1 \bar{f}_2$			$Z \to f \bar{f}$			
$f_1 \bar{f}_2$	$\Gamma_0(MeV)$	$B_r(\%)$	$f\bar{f}$	$f(\sin^2 \Theta_W)^+$	$\Gamma_0(MeV)$	$B_r(\%)$
$\ell \bar{\nu}_\ell$	225.6	11.1	$\nu_\ell \bar{\nu}_\ell$	1	165.8	6.8
			inv		497.5	20.5
			$\ell \bar{\ell}$	0.5032	83.4	3.4
$u \bar{d}$	676.7	33.3	$u \bar{u}$	0.5748	286.0	11.8
			$d \bar{d}$	0.7404	368.3	15.2
had	1353.4	66.6	had		1676.6	69.2
tot	2030.1	100.0	tot		2424.3	100.0

Since $m_t > M_W$, M_Z, the decays $Z \to t\bar{t}$ and $W \to tb$ are energetically forbidden. LEP has excluded the existence of a fourth family neutrino of mass $m_\nu < M_Z/2$. Since the b quark is the heaviest of the produced fermions, with a mass $m_b \simeq 5\ GeV$, we can safely neglect all mass effects in calculating the widths.

By $\Gamma_{\text{inv}} = 3\Gamma_{Z \to \nu\bar{\nu}}$ we denote the **invisible** width for the decays into ν_e, ν_μ and ν_τ. Γ_{had} is the total hadronic width for the decays into u, d, s, c and b quarks or the corresponding hadronic states. The total Z-width (and similarly the W-width) is given with high accuracy by the sum over the two body decays

$$\Gamma_{\rm tot} = \Gamma_Z \simeq \sum_f \Gamma_f \ ; \ \ \Gamma_f = \Gamma_{Z\to f\bar{f}} = \Gamma_\nu \ f_f(\sin^2\Theta_W)$$

with $f_f(\sin^2\Theta_W) \doteq 1-4|Q_f|\sin^2\Theta_W + 8Q_f^2\sin^4\Theta_W$ normalized to the ν channel. Since the Higgs is known not to be light ($m_H > 60$ GeV), the contribution $\Gamma_{Z\to Hf\bar{f}}$ is insignificant. It would only be non-negligible for a very light Higgs.

We now consider

$Z-$production in e^+e^-collisions.

In the light fermion approximation (see Eq. (14.3))

$$\frac{\sum|T_{12}|^2}{M_V^2} = 3\cdot 16\pi\ \frac{\Gamma_{Z\to f\bar{f}}}{M_Z}$$

and hence, in the narrow width approximation,

$$\sigma(e^+e^-\to Z)|_{s\simeq M_Z^2} = 12\pi\frac{\Gamma_{Z\to e^+e^-}}{M_Z}\pi\delta(s-M_Z^2)$$

with $s=(p_1+p_2)^2 = 4E_b^2$ and E_b the beam energy.

Using the relation (the Breit-Wigner form will be "derived" below)

$$\pi\delta(s-M_Z^2) = \lim_{\Gamma_Z\to 0}\frac{M_Z\Gamma_Z}{(s-M_Z^2)^2+\Gamma_Z^2M_Z^2}$$

we easily obtain the cross section for (finite width) resonance production, described by a Breit-Wigner line-shape,

$$\sigma(e^+e^-\to Z) = \frac{12\pi}{M_Z^2}\frac{\Gamma_{Z\to ee}\Gamma_Z M_Z^2}{(s-M_Z^2)^2+\Gamma_Z^2M_Z^2}\quad .$$

Near resonance, the cross section for $e^+e^-\to Z\to f\bar{f}$ is

$$\sigma(e^+e^-\to Z\to f\bar{f}) = \sigma(e^+e^-\to Z)\cdot\frac{\Gamma_{Z\to ff}}{\Gamma_Z}$$

where the last factor is the branching fraction, and hence

$$\sigma_{0Z}^f = \sigma(e^+e^-\to Z\to f\bar{f}) = \sigma_{\rm peak}^f\frac{\Gamma_Z^2M_Z^2}{(s-M_Z^2)^2+\Gamma_Z^2M_Z^2}\quad . \tag{14.7}$$

where $\sigma_{\rm peak}^f$ is the peak cross section, evaluated at $s=M_Z^2$,

$$\sigma_{\rm peak}^f = \frac{12\pi}{M_Z^2}\frac{\Gamma_e}{\Gamma_Z}\frac{\Gamma_f}{\Gamma_Z}\quad . \tag{14.8}$$

Table 14.2 Lowest order peak cross section $\sigma_{\rm peak}^f$. M_Z and Γ_f as in Tab. 14.1. ($1GeV^{-2} = 0.38938\times 10^6\ nb$)

f	ν	μ	u	d	inv	had	tot
$\sigma_{\rm peak}^f(nb)$	4.16	2.09	7.17	9.23	12.47	42.03	60.77

3.2 The process $e^+e^- \to f\bar{f}, (f\bar{f}\gamma)(f \neq e)$

We now consider in detail the process

$$e^+(p_+) + e^-(p_-) \to \bar{f}(q_+) + f(q_-) + \text{“}\gamma(k)\text{”}$$

in the Born approximation given by the diagrams in Fig. 14.1 Real γ emission will be considered later.

In the center of mass frame in terms of the beam energy $s = (p_+ + p_-)^2 = 4E_b^2$ and $t = (p_+ - q_+)^2 = 2E_b^2(1 - \cos\theta)$ with θ the angle between p_+ and q_+. By the arguments given in the previous chapter we can safely neglect the fermion masses if we assume $s \gg m_b^2$ (the bottom quark is the heaviest of the final state fermions at LEP energies). $t\bar{t}$ production is not considered.

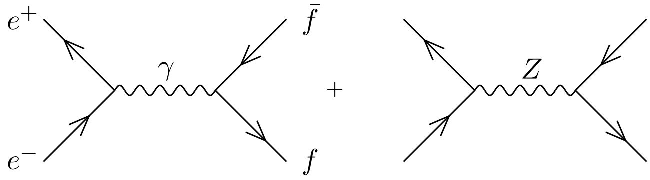
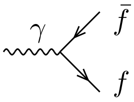
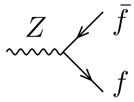


Fig. 14.1: Born diagrams for the process $e^+e^- \to f\bar{f}$

Since we are considering beam energies E_b far above the Υ threshold the fermions are essentially massless and helicity is a good quantum number. It is therefore convenient to use left- and right-handed fields $f_L = \frac{1-\gamma_5}{2} f$ and $f_R = \frac{1+\gamma_5}{2} f$ which describe polarized fermion states: The couplings are



$$:= \; eQ_f \bar{f}\gamma^\mu f \; ;$$



$$:= \; \frac{M_Z}{v} \bar{f}\gamma^\mu(v_f - a_f\gamma_5) f \tag{14.9}$$

where $\bar{f}\gamma^\mu f = \bar{f}_L\gamma^\mu f_L + \bar{f}_R\gamma^\mu f_R$ and $\bar{f}\gamma^\mu(v_f - a_f\gamma_5)f = \varepsilon_{Lf}\bar{f}_L\gamma^\mu f_L + \varepsilon_{Rf}\bar{f}_R\gamma^\mu f_R$ with $\varepsilon_{Lf} = v_f + a_f$ and $\varepsilon_{Rf} = v_f - a_f$.

We notice that for the vector-like couplings (i.e. vector or axial-vector) no $\bar{f}_R \ldots f_L$ or $\bar{f}_L \ldots f_R$ terms are present. This is a general feature of *any* gauge interaction (coupling via spin 1 vector fields). Since e_L^- describes a left-handed electron and a right-handed positron etc. there are no transitions from equal helicity e^+e^- into equal helicity $\bar{f}f$ states! Therefore there are only four possible transition amplitudes $T_{h_i h_f}$ for polarized states: $T_{LL}, T_{LR}, T_{RL}, T_{RR}$, where $h_i = -1\ (L),\ +1\ (R)$ is the electron helicity and $h_f = -1\ (L),\ +1\ (R)$ the final state helicity of the fermion f. The helicity of the antiparticle in each case is fixed to have the the opposite value of that of the particle. This is true for any gauge theory!

The differential cross section is given by

$$\frac{d\sigma}{d\cos\theta} = \frac{1}{32\pi s} \frac{N_{cf}}{4} \sum |T_{h_i h_f}|^2 \tag{14.10}$$

with

$$|T_{h_e h_f}|^2 = (1 \pm \cos\theta)^2 \left| \varepsilon_{h_e e}\varepsilon_{h_f f} \, \frac{\sqrt{2}G_\mu M_Z^2 s}{s - M_Z^2 + iM_Z\Gamma_Z} + Q_e Q_f 4\pi\alpha \right|^2 \tag{14.11}$$

The sign $+(-)$ is for T_{LL}, T_{RR} (T_{LR}, T_{RL}).

We give a brief derivation of the above result Eq. (14.11). The matrix elements for the two graphs of Fig. 14.1 are given by

$$\begin{aligned}
T^Z_{h_e h_f} &= \frac{ig^2}{4\cos^2\Theta_W}\varepsilon_{h_e}\varepsilon_{h_f} \quad \frac{\{\bar v_e(-h_e)\gamma^\mu u_e(h_e)\}\{\bar u_f(h_f)\gamma_\mu v_f(-h_f)\}}{s-M_Z^2+iM_Z\Gamma_Z} \\
T^\gamma_{h_e h_f} &= ie^2 Q_f Q_e \quad \frac{\{\bar v_e(-h_e)\gamma^\mu u_e(h_e)\}\{\bar u_f(h_f)\gamma_\mu v_f(-h_f)\}}{s}
\end{aligned}$$

and the transition amplitudes $|T_{h_e h_f}|^2$ are proportional to

$$\begin{aligned}
A_{h_e h_f} &\doteq \mathrm{Tr}\,\{v_{e^+}\bar v_{e^+}(-h_e)\gamma^\mu u_{e^-}\bar u_{e^-}(h_e)\gamma^\nu\}\cdot \mathrm{Tr}\,\{u_f\bar u_f(h_f)\gamma_\mu v_{\bar f}\bar v_{\bar f}(-h_f)\gamma_\nu\} \\
&= \mathrm{Tr}\,\{\not p_+\gamma^\mu \not p_-\left(\frac{1+h_e\gamma_5}{2}\right)^2\gamma^\nu\}\cdot \mathrm{Tr}\,\{\not q_-\gamma_\mu\not q_+\left(\frac{1+h_f\gamma_5}{2}\right)^2\gamma_\nu\}\;.
\end{aligned}$$

In the last step we have used the following projection operators for fixed helicity:

$$\begin{aligned}
u_h(p,s)\bar u_h(p,s) &= (\not p + m)\,\frac{1+h\gamma_5\not s}{2} \\
v_h(p,s)\bar v_h(p,s) &= (\not p - m)\,\frac{1-h\gamma_5\not s}{2}\;.
\end{aligned}$$

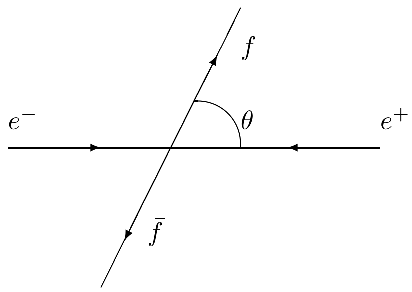
Then, using

$$\mathrm{Tr}\,\{\not p_+\gamma^\mu\not p_-\left(\frac{1+h\gamma_5}{2}\right)^2\gamma^\nu\} = 2\{p_-^\mu p_+^\nu + p_+^\mu p_-^\nu - g^{\mu\nu}(p_+p_-) - ih_e\varepsilon^{\alpha\mu\beta\nu}(p_+)_\alpha(p_-)_\beta\}$$

we obtain

$$A_{h_e h_f} = 8\,\{(p_-q_+)(p_+q_-)+(p_-q_-)(p_+q_+)+h_e h_f[(p_-q_+)(p_+q_-)-(p_-q_-)(p_+q_+)]\}$$

It is natural to evaluate this amplitude in the c.m. frame



where we have

$$\begin{aligned}
(p_+q_+) &= (p_-q_-) = p^2(1-\cos\theta) = \frac{s}{4}(1-\cos\theta) \\
(p_+q_-) &= (p_-q_+) = p^2(1+\cos\theta) = \frac{s}{4}(1+\cos\theta)
\end{aligned}$$

and hence

$$\begin{aligned}
A_{h_e h_f} &= \frac{s^2}{2}\left[(1-\cos\theta)^2+(1+\cos\theta)^2+h_e h_f\{(1+\cos\theta)^2-(1-\cos\theta)^2\}\right] \\
&= s^2(1+h_e h_f\cos\theta)^2
\end{aligned}$$

The total transition amplitude, which is the square of the sum of the two terms the $Z-$ and the γ–exchange diagrams is then given by

$$\begin{aligned}
|T_{h_e h_f}|^2 &= s^2(1+h_e h_f\cos\theta)^2\left|\varepsilon_{h_e}\varepsilon_{h_f}\frac{\sqrt{2}G_\mu M_Z^2}{s-M_Z^2+iM_Z\Gamma_Z}+Q_eQ_f\frac{4\pi\alpha}{s}\right|^2 \\
&= (1+h_e h_f\cos\theta)^2\left\{\varepsilon_{h_e}^2\varepsilon_{h_f}^2 2G_\mu^2 M_Z^4|\chi(s)|^2+16\pi^2\alpha^2Q_f^2-8\pi\alpha\sqrt{2}G_\mu M_Z^2 Q_f\varepsilon_{h_e}\varepsilon_{h_f}\mathrm{Re}\chi(s)\right\}
\end{aligned}$$

where $\chi(s) := \frac{s}{s-M_Z^2+iM_Z\Gamma_Z}$ is the resonance factor.

The differential cross section in the c.m. frame then is given by

$$\frac{d\sigma}{d\cos\theta} = \frac{|T|^2}{32\pi s}$$

For unpolarized beams we still have to perform the appropriate averaging and spin summation. The tree terms of $|T|^2$

$$\begin{aligned}
\textstyle\sum_{h_e=\pm1,h_f=\pm1}(1+h_eh_f\cos\theta)^2 &= (1+\cos^2\theta)\\
\textstyle\sum_{h_e=\pm1,h_f=\pm1}\epsilon_{h_e}\epsilon_{h_f}(1+h_eh_f\cos\theta)^2 &= v_ev_f(1+\cos^2\theta)+2a_ea_f\cos\theta\\
\textstyle\sum_{h_e=\pm1,h_f=\pm1}\epsilon^2_{h_e}\epsilon^2_{h_f}(1+h_eh_f\cos\theta)^2 &= (v_e^2+a_e^2)(v_f^2+a_f^2)(1+\cos^2\theta)+8a_fv_fa_ev_e\cos\theta
\end{aligned}$$

In the cross section we distinguish three pieces, the pure QED cross section, the $\gamma-Z$ interference term and the pure Z-exchange cross section:

$$\frac{d\sigma}{d\cos\theta}(e^+e^-\to\gamma^*,Z^*\to\bar{f}f) = \frac{d\sigma^\gamma}{d\cos\theta}+\frac{d\sigma^{\gamma Z}}{d\cos\theta}+\frac{d\sigma^Z}{d\cos\theta}\ . \tag{14.12}$$

Of course, in general, there is no way to measure these terms individually. However, individual terms may dominate as for example the QED piece at low s or the Z-exchange term near the Z-resonance.

For unpolarized beams and final states we obtain

$$\begin{aligned}
\frac{d\sigma^\gamma}{d\cos\theta} &= \frac{\pi\alpha^2Q_f^2N_{cf}}{2s}(1+\cos^2\theta)\\
\frac{d\sigma^{\gamma Z}}{d\cos\theta} &= -\frac{\alpha Q_f\sqrt{2}G_\mu M_Z^2}{4s}N_{cf}Re\chi(s)\cdot\{v_ev_f(1+\cos^2\theta)\\
&\quad +2a_ea_f\cos\theta\}\\
\frac{d\sigma^Z}{d\cos\theta} &= \frac{G_\mu^2M_Z^4}{16\pi s}N_{cf}|\chi(s)|^2\cdot\left\{(v_e^2+a_e^2)(v_f^2+a_f^2)(1+\cos^2\theta)\right.\\
&\quad \left.+8a_ea_fv_ev_f\cos\theta\right\}
\end{aligned} \tag{14.13}$$

with the resonance factor

$$\chi(s) = \frac{s}{s-M_Z^2+iM_Z\Gamma_Z}\ .$$

Near the Z-resonance, the process $e^+e^-\to f\bar{f}$ is predominantly a parity violating weak interaction transition. The axial couplings a_f lead to *asymmetries* at the tree level.

i) asymmetry in the angular distribution due to terms linear in $\cos\theta$ called forward-backward asymmetry or charge asymmetry A_{FB}, e.g. in $e^+e^-\to\mu^+\mu^-$ the μ^+ is produced with different probability in opposite directions relative to the incoming e^+. [40]

ii) asymmetry between cross sections for (polarized) L and R states, the so called left-right asymmetries A_{LR}.

Before we discuss the asymmetries in more detail, we briefly consider the total cross section.

[40] This type of asymmetry (though much smaller) is also present in pure QED (parity conserving) coming from higher order effects (box diagrams).

Total cross section

The total cross section, with respect to the final state couplings, may be split into a pure vector and a pure axial-vector piece

$$\sigma_0^f = \int_{-1}^{+1} d\cos\theta \frac{d\sigma}{d\cos\theta} = \sigma_{0f}^{VV} + \sigma_{0f}^{AA} \tag{14.14}$$

with

$$\begin{aligned}
\sigma_{0f}^{VV} &= \tfrac{4\pi\alpha^2 Q_f^2 N_{cf}}{3s} - \tfrac{2\alpha Q_f \sqrt{2} G_\mu M_Z^2 N_{cf}}{3s} Re\chi(s) v_e v_f + \tfrac{G_\mu^2 M_Z^4 N_{cf}}{6\pi s} |\chi|^2 (v_e^2 + a_e^2) v_f^2 \\
\sigma_{0f}^{AA} &= \tfrac{G_\mu^2 M_Z^4 N_{cf}}{6\pi s} |\chi|^2 (v_e^2 + a_e^2) a_f^2 \ .
\end{aligned}$$

Near the Z-resonance the pure Z-exchange term dominates and we may rewrite the cross section in the form:

$$\sigma_{0f} = \sigma_{0f}^\gamma + \sigma_{0f}^{\gamma Z} + \sigma_{0f}^Z = \sigma_{0f}^Z \cdot \left(1 + \frac{\sigma_{0f}^{\gamma Z}}{\sigma_{0f}^Z}\right) + \sigma_{0f}^\gamma = \sigma_{0f}^Z \left(1 + \mathcal{R}_f \frac{s - M_Z^2}{s}\right) + \sigma_{0f}^\gamma \ .$$

For $\mathcal{R}_f$ we find a $\gamma - Z$ interference correction

$$\mathcal{R}_f = \frac{8\pi\alpha Q_e Q_f}{\sqrt{2} G_\mu M_Z^2} \frac{v_e v_f}{(v_e^2 + a_e^2)(v_f^2 + a_f^2)} \ .$$

At resonance ($s = M_Z^2$) this correction does not contribute. σ_{0f}^γ is the QED background term

$$\sigma_{0f}^\gamma = \frac{4\pi\alpha^2 Q_f^2 N_{cf}}{3s}$$

which leads to a correction below 1% at resonance. Finally, using formula (39) for the width we find Eqs. (40) and (41) in agreement with our simplified derivation of the previous chapter. Closer inspection shows that the cross section formula

$$\sigma_{\mathrm{eff}}^f(s) = \frac{12\pi\Gamma_e\Gamma_f}{(s-M_Z^2)^2 + s^2\frac{\Gamma_Z^2}{M_Z^2}} \left\{ \frac{s}{M_Z^2} + \mathcal{R}_f \frac{s-M_Z^2}{M_Z^2} + \mathcal{I}_f \frac{\Gamma_Z}{M_Z} + \dots \right\} + \sigma_{0,\mathrm{QED}}^f \tag{14.15}$$

yields a model independent fit of the Z-line-shape provided Γ_e, Γ_f and Γ_Z are the physical (partial) widths, i.e. they include higher order corrections. We have included possible corrections proportional to Γ_Z/M_Z and the ellipses represent higher order terms in the expansion in $\frac{s-M_Z^2}{M_Z^2}$ and $\frac{\Gamma_Z}{M_Z}$. One important point (see below) is that $\Gamma_Z(s)$ defined in terms of the Z self-energy $\Pi_Z(s)$ by $M_Z\Gamma_Z(s) = Im\Pi_Z(s)$ is to high accuracy proportional to s:

$$\Gamma_Z(s) \simeq \frac{s}{M_Z^2}\Gamma_Z.$$

3.3 Asymmetries

A. Forward-backward asymmetry:

The differential cross section has a $\cos\theta$ even and a $\cos\theta$ odd term:

$$\frac{d\sigma}{d\cos\theta} = \sigma_{0f} \cdot \frac{3}{8}(1 + \cos^2\theta) + \Delta_{0f}\cos\theta \tag{14.16}$$

where σ_{0f} is the total cross section and

$$\begin{array}{rcl}\Delta_{0f} & = & \Delta_{0f}^{\gamma Z}+\Delta_{0f}^{Z} \\ \Delta_{0f}^{\gamma Z} & = & -\frac{\alpha Q_f N_{cf}}{2s}\sqrt{2}G_\mu M_Z^2 Re\chi a_e a_f \\ \Delta_{0f}^{Z} & = & \frac{G_\mu^2 M_Z^4}{2\pi s}N_{cf}|\chi|^2 a_e a_f v_e v_f \end{array} .$$

The forward-backward asymmetry is

$$A_{FB}(s,\cos\theta)=\frac{d\sigma(\theta)-d\sigma(\pi-\theta)}{d\sigma(\theta)+d\sigma(\pi-\theta)}=\frac{8}{3}\frac{\Delta_{0f}}{\sigma_{0f}}\frac{\cos\theta}{1+\cos^2\theta}$$

or in integrated form

$$A_{FB}(s)=\frac{\left(\int_0^1-\int_{-1}^0\right)d\cos\theta\frac{d\sigma}{d\cos\theta}}{\int_{-1}^{+1}d\cos\theta\frac{d\sigma}{d\cos\theta}}=\frac{\Delta_{0f}}{\sigma_{0f}} . \tag{14.17}$$

Particular regimes of interest are the following:

i) For small $s \ll M_Z^2$ we have

$$R_{0f}=\frac{\sigma_{0f}}{\sigma_{\mu\mu}}\simeq\left\{Q_f^2N_{cf}+\frac{\sqrt{2}G_\mu}{2\pi\alpha}N_{cf}\frac{v_e(Q_fv_f)}{1-s/M_Z^2}\right\} \tag{14.18}$$

where

$$\sigma_{\mu\mu}=\sigma_0(e^+e^-\to\gamma^*\to\mu^+\mu^-)=\frac{4\pi\alpha^2}{3s}$$

is the QED "point" cross section used to normalize the hadronic cross section

$$\begin{array}{rcl}\sigma_{\rm had} & = & \sigma(e^+e^-\to\text{ hadrons })=\sum\limits_{\text{quark }q}\sigma_{0q} \\ R(s) & \doteq & \dfrac{\sigma_{\rm had}}{\sigma_{\mu\mu}}=\sum\limits_q R_q\simeq 3\sum\limits_{\substack{q\\ m_q\leq\sqrt{s}}}Q_q^2\end{array} \tag{14.19}$$

Notice that in this quantity the color factor 3 can be directly measured! For the asymmetry we get

$$A_{FB}^{f\bar f}(s)\simeq\frac{3}{8}a_e\left(\frac{a_f}{Q_f}\right)\frac{\sqrt{2}G_\mu}{\pi\alpha}\frac{s}{1-s/M_Z^2} \tag{14.20}$$

an expression which vanishes for $s\to 0$.

ii) For $s\simeq M_Z$ we find

$$A_{FB}(M_Z^2)=\frac{\Delta_{0f}^Z(M_Z^2)}{\sigma_{0f}^Z(M_Z^2)}=\frac{3}{4}\cdot\frac{2v_ea_e}{v_e^2+a_e^2}\cdot\frac{2v_fa_f}{v_f^2+a_f^2} . \tag{14.21}$$

Asymmetries at the $Z-$resonance can all be expressed in terms of the coupling ratios

$$A_f\doteq\frac{2v_fa_f}{v_f^2+a_f^2}=\frac{\varepsilon_{Lf}^2-\varepsilon_{Rf}^2}{\varepsilon_{Lf}^2+\varepsilon_{Rf}^2}. \tag{14.22}$$

From the representation in terms of left- and right-handed couplings introduced earlier we see that A_f measures the normalized difference between left-handed and right-handed transition amplitudes. For A_{FB} we have

$$A_{FB}^{f\bar{f}}(M_Z^2) = \frac{3}{4} A_e A_f. \tag{14.23}$$

It is important to notice that

$$A_e = \frac{2\xi}{1+\xi^2} \text{ with } \xi = \frac{v_e}{a_e} = 1 - 4\sin^2\Theta_W$$

is a quantity which would vanish for $\sin^2\Theta_W = 0.25$. Since the experimental value for $\sin^2\Theta_W$ is about 0.23 A_e is unfortunately rather small. A small difference of large numbers is difficult to determine precisely. By universality A_e is the same for $\ell = e, \mu$ and τ and hence

$$A_{FB}^{\mu\bar{\mu}}(M_Z^2) = \frac{3}{4} A_e^2 \simeq 3\xi^2$$

Table 14.3 $e^+e^- \to f\bar{f}$ forward-backward asymmetry at the Z-peak for various values of $\sin^2\Theta_W$

$\sin^2\Theta_W$	0.22	0.23	0.24	0.25
ξ	0.12	0.08	0.04	0
A_e	0.237	0.159	0.0799	0
$A_{FB}^{\mu\bar{\mu}}$	0.0420	0.0190	0.0048	0
$A_{FB}^{c\bar{c}}$	0.1253	0.0802	0.0382	0
$A_{FB}^{b\bar{b}}$	0.1673	0.1117	0.0557	0

B. Final state polarization asymmetry

We only consider the integrated asymmetry

$$A_{\text{pol}}^f \doteq \frac{\sigma(e^+e^- \to f_L\bar{f}) - \sigma(e^+e^- \to f_R\bar{f})}{\sigma(e^+e^- \to f_L\bar{f}) + \sigma(e^+e^- \to f_R\bar{f})} \tag{14.24}$$

at the Z-resonance. Using the helicity amplitudes $|T_{h_i h_e}|^2$ we obtain:

$$A_{\text{pol}}^f(M_Z^2) = \frac{(\varepsilon_{Lf}^2 - \varepsilon_{Rf}^2)}{(\varepsilon_{Lf}^2 + \varepsilon_{Rf}^2)} = A_f \tag{14.25}$$

which is independent of the initial state couplings. This asymmetry cannot be measured for quarks which hadronize into hadron showers. The only case which can be investigated is the τ-polarization where it is possible to reconstruct the τ-polarization from the decays $\tau \to \pi\nu$, $\tau \to \rho\nu$ and $\tau \to a_1\nu$ with the subsequent decays $\rho \to \pi\pi$ and $a_1 \to \pi\pi\pi$. For $e^+e^- \to \tau^+\tau^-$ we have

$$A_{\text{pol}}^\tau(M_Z^2) = A_e = \frac{2\xi}{1+\xi^2} \tag{14.26}$$

which is linear in the vector coupling ξ. Some numerical values have been given in Tab. 14.3.

C. Polarized beams

With polarized beams one can measure a number of additional asymmetries:

a) Initial state transversal polarization asymmetry
(azimuthal asymmetry for natural polarization)

In the magnetic field of a ring collider the e^+-spins tend to line up with the magnetic field ($-y$-direction) such that a natural transverse polarization is set up. If we assume the electron to move in the z-direction we may write the e^+-polarization vector in the form

$$\vec{P}^{\pm} = \left(P_{\perp}^{\pm} \cos\varphi^{\pm}, P_{\perp}^{\pm}\sin\varphi^{\pm}, P_{L}^{\pm} \right)$$

where $P_{\perp}^{\pm}$ measures the transverse and $P_L^{\pm}$ the longitudinal degree of polarization. $\varphi^{\pm} = -\pi/2$ and $P_L^{\pm} = 0$ means natural polarization. If beams are transversely polarized one has an azimuthal asymmetry and one defines

$$A_{\perp} = \frac{4}{P_{\perp}^{+} P_{\perp}^{-}} \frac{\int d\Omega \; \cos 2\varphi \frac{d\sigma}{d\Omega} \left(e^+(P_{\perp}^{+}) + e^-(P_{\perp}^{-}) \to f\bar{f} \right)}{\int d\Omega \frac{d\sigma}{d\Omega} \left(e^+(P_{\perp}^{+}) + e^-(P_{\perp}^{-}) \to f\bar{f} \right)} \,. \tag{14.27}$$

We just give, without derivation, the result one obtains for $s = M_Z^2$:

$$A_{\perp}(M_Z^2) = \frac{v_e^2 - a_e^2}{v_e^2 + a_e^2} = -\frac{1-\xi^2}{1+\xi^2} \tag{14.28}$$

a quantity which is independent of the final state.

b) Initial state longitudinal polarization asymmetry

In this case, assuming longitudinally polarized beams, one measures the total cross section with left-handed and right-handed electrons separately and defines

$$A_{LR} = \frac{\sigma(e_L^- e^+ \to f\bar{f}) - \sigma(e_R^- e^+ \to f\bar{f})}{\sigma(e_L^- e^+ \to f\bar{f}) + \sigma(e_R^- e^+ \to f\bar{f})} \,. \tag{14.29}$$

At $s = M_Z^2$ one finds

$$A_{LR}(M_Z^2) = \frac{\varepsilon_{Le}^2 - \varepsilon_{Re}^2}{\varepsilon_{Le}^2 + \varepsilon_{Re}^2} = A_e \simeq 2\xi \tag{14.30}$$

for the integrated asymmetry.

The left-right asymmetry is a very important observable due to the following properties: It is

- a ratio of <u>total</u> cross sections
- independent of the final state [41]
- linear in ξ

The first property is very important since the notoriously large QED and QCD corrections are almost identical for left- and right-handed states and therefore drop out in the ratio almost completely. The second property implies that one can sum over all flavors gaining enormously in statistics. The third property tells us that A_{LR} is enhanced by a factor $2/(3\xi)$ relative to $A_{FB}^{\mu\bar{\mu}}$.

[41] This is true only for the integrated asymmetry. The angular distributions $A_{LR}^f(\cos\theta)$ depend on the flavor f.

In addition the polarized forward-backward asymmetries can be measured:

$$A^f_{FB,\mathrm{pol}} = \frac{2}{P_L^+ + P_L^-} \frac{(\int_0^1 - \int_{-1}^0) d\cos\theta \frac{d\sigma}{d\cos\theta}(P_L^+ \neq 0) - (\int_0^1 - \int_{-1}^0) d\cos\theta \frac{d\sigma}{d\cos\theta}(P_L^- \neq 0)}{\int_{-1}^{+1} d\cos\theta \frac{d\sigma}{d\cos\theta}(P_L^+ \neq 0) + \int_{-1}^{+1} d\cos\theta \frac{d\sigma}{d\cos\theta}(P_L^- \neq 0)} \quad (14.31)$$

which on the Z-resonance yields:

$$A^f_{FB,\mathrm{pol}}(M_Z^2) = \frac{3}{4} A_f \quad (14.32)$$

3.4 Conclusions:

The measurement of asymmetries opens up the possibility to determine many independent observables. This is crucial for precision tests of the SM. Two points make asymmetries at the Z-peak very interesting. First, at the Z-peak very high rates of events are available which makes high precision tests possible. Second, at the Z-peak one is dealing almost purely with a weak NC process! No detailed clean tests of the NC were possible before LEP. The clean $\nu_\mu e$-scattering processes suffer from low rates and the deep inelastic $\nu_\mu N$-scattering data from hadronic uncertainties.

Longitudinally polarized beams are highly desired, since only in this case one has good observables that can test

$$A_f = \frac{2 v_f a_f}{v_f^2 + a_f^2}$$

for individual flavors to a good accuracy. This is an important supplement to the measurement of the partial widths which yields tests of

$$v_f^2 + a_f^2 .$$

For precision tests it will be crucial to carefully analyze the following types of corrections:

i) QED corrections, bremsstrahlung;

ii) electroweak "non-QED" corrections;

iii) QCD corrections for hadronic final states;

iv) mass effects.

These corrections will be discussed in the following. Particularly interesting are the "non-QED" higher order corrections since they are the key in finding deviations from the SM at its quantum level.

A Properties of free fields

Notation for Green functions:

a) Causal commutator:

$$\Delta(z;m^2) = -i\,(2\pi)^{-3} \int d^4p\, \epsilon(p^0)\, \delta(p^2 - m^2)\, e^{-ipz}$$

$$\left(\Box + m^2\right)\, \Delta(z) = 0\ ,\quad \Delta(0,\vec{z}\,) = 0\ ,\quad \partial^0\Delta(z^0,\vec{z}\,)|_{z^0=0} = -\delta^{(3)}(\vec{z}\,)$$

b) Feynman propagator:

$$\Delta_F(z;m^2) = (2\pi)^{-4} \int d^4p\, \frac{1}{p^2 - m^2 + i\epsilon}\, e^{-ipz}$$

$$\left(\Box + m^2\right)\, \Delta_F(z) = -\delta^{(4)}(z)$$

Important remark: The Green functions are distributions (generalized or singular functions) and the δ–function is the identity in the space of distributions, which we may choose to be the space of tempered distributions $\mathcal{S}'(\mathbf{R}^4)$. The last equation above tells us that up to a sign the Feynman propagator is the inverse of the Klein-Gordon operator $\Box + m^2$.

Notice that in the following, except in the Feynman propagator, p^2 is always on-shell and hence $p^0 = \sqrt{\vec{p}^{\,2} + m^2}$. Therefore, $e^{\pm ipx}\,|_{p^0=\sqrt{\vec{p}^{\,2}+m^2}}$ is always a solution of the Klein-Gordon equation: $(\Box + m^2)\ e^{\pm ipx} = 0$.

A.1 Real scalar field: representation: (0,0)

$$\left(\Box + m^2\right) \varphi(x) = 0\ ;\quad \varphi^* = \varphi$$

Describes a particle of mass m, spin 0, particle $\equiv$ antiparticle.
Examples: π^0, η, Higgs particle H .

$$\underline{\varphi(x) = \int d\mu(p) \left\{ a(\vec{p}\,)\, e^{-ipx} + a^+(\vec{p}\,)\, e^{ipx} \right\}}$$

Canonical commutation relations:

$$\begin{array}{rcl} [\ a(\vec{p}\,), a^+(\vec{p}\,')] &=& (2\pi)^3\, 2\omega_p\, \delta^{(3)}(\vec{p} - \vec{p}\,') \\ [\ a(\vec{p}\,),\ \ a(\vec{p}\,')] &=& 0 \\ [a^+(\vec{p}\,), a^+(\vec{p}\,')] &=& 0 \end{array}$$

Green functions: solutions of the homogeneous (no source) or inhomogeneous (point source) Klein-Gordon equations

$$\begin{array}{rcl} [\varphi(x), \varphi(y)] &=& i\,\Delta(x - y; m^2) \\ <0\,|\,T\,\{\varphi(x), \varphi(y)\}\,|\,0> &=& i\,\Delta_F(x - y; m^2) \end{array}$$

$$(\ \Box_x + m^2\ )\,\Delta(x - y; m^2) = 0\ \ ,\quad (\ \Box_x + m^2\ )\,\Delta_F(x - y; m^2) = -\delta^{(4)}(x - y)$$

A.2 Complex scalar field: representation (0,0)

$$\left(\Box + m^2\right)\varphi(x) = 0\ ;\quad \varphi^* \neq \varphi$$

Describes a particle of mass m, spin 0, particle $\neq$ antiparticle.
Examples: (π^+, π^-), $(K^0, \bar{K}^0)$.

$$\underline{\varphi(x) = \int d\mu(p) \left\{ a(\vec{p}\,)\, e^{-ipx} + b^+(\vec{p}\,)\, e^{ipx} \right\}}$$

A complex field is equivalent to a doublet (φ_1, φ_2) of real fields: $\varphi_i = \varphi_i^*$, $i = 1, 2$ related by $\varphi^\pm = \frac{1}{\sqrt{2}}(\varphi_1 \mp i\,\varphi_2)$ with $\varphi^- = \varphi$, $\varphi^+ = \varphi^*$.

Canonical commutation relations:

$$\left[a(\vec{p}\,), a^+(\vec{p}\,')\right] \;=\; \left[b(\vec{p}\,), b^+(\vec{p}\,')\right] = (2\pi)^3\, 2\omega_p\, \delta^{(3)}(\vec{p} - \vec{p}\,')$$

and all other commutators vanishing.

Green functions:

$$\begin{aligned} [\varphi(x), \varphi^*(y)] &= i\,\Delta(x - y; m^2) \\ < 0 \mid T\,\{\varphi(x), \varphi^*(y)\} \mid 0 > &= i\,\Delta_F(x - y; m^2) \end{aligned}$$

A.3 Dirac field: representation $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ (reducible)

$$(\Box + m^2)\psi_\alpha(x) = 0\ ;\quad (i\gamma^\mu \partial_\mu - m)_{\alpha\beta}\psi_\beta(x) = 0 \quad \text{(Dirac equation)}$$

Describes a particle of mass m, spin 1/2, particle $\neq$ antiparticle.
Examples: $(e^-, e^+), (p, \bar{p}), (n, \bar{n})$, "quarks".

$$\underline{\psi_\alpha(x) = \sum_{r=\pm 1/2} \int d\mu(p) \left\{ u_\alpha(\vec{p}, r)\, a(\vec{p}, r)\, e^{-ipx} + v_\alpha(\vec{p}, r)\, b^+(\vec{p}, r)\, e^{ipx} \right\}}$$

Canonical anticommutation relations:

$$\left\{a(\vec{p}, r), a^+(\vec{p}\,', r')\right\} = \left\{b(\vec{p}, r), b^+(\vec{p}\,', r')\right\} = (2\pi)^3\, 2\omega_p\, \delta^{(3)}(\vec{p} - \vec{p}\,')\, \delta_{rr'}$$

and all other anticommutators vanishing.

Dirac-algebra: Algebra of 4 x 4 $\underline{\gamma\text{-matrices}}$

$$\begin{aligned} \{\gamma^\mu, \gamma^\nu\} = &\quad \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = \quad 2g^{\mu\nu} \\ \{\gamma_5, \gamma^\mu\} = &\quad \gamma_5\gamma^\mu + \gamma^\mu\gamma_5 = \quad 0 \end{aligned}$$

$$\begin{aligned} \gamma_5 &= i\gamma^0\gamma^1\gamma^2\gamma^3\ ;\ \gamma_5^2 = 1 \\ \sigma^{\mu\nu} &= \frac{i}{2}\,(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) = \frac{i}{2}\,[\gamma^\mu, \gamma^\nu] \end{aligned}$$

$$\gamma_\mu = g_{\mu\nu}\gamma^\nu$$

$$\gamma_0 = \gamma^0 \ , \ \ \gamma_i = -\gamma^i \ , \ \ \gamma^5 = \gamma_5$$

$$\gamma^\mu = \left(\gamma^0, \vec{\gamma}\right) \ , \ \ \gamma_\mu = \left(\gamma^0, -\vec{\gamma}\right)$$

Equivalence relations:

$$\begin{array}{rclcrcr} \gamma_\mu^+ & = & A\gamma_\mu A^{-1} & \qquad & \gamma_5^+ & = & -A\gamma_5 A^{-1} \\ \gamma_\mu^T & = & B\gamma_\mu B^{-1} & & \gamma_5^T & = & B\gamma_5 B^{-1} \\ \gamma_\mu^* & = & C\gamma_\mu C^{-1} & & \gamma_5^* & = & -C\gamma_5 C^{-1} \end{array}$$

Pauli matrices:

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

Representations of the Dirac algebra:

a. Standard representation

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$A = \gamma^0, \quad B = \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad C = \begin{pmatrix} \sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}$$

b. Helicity representation

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$A = \gamma^0, \quad B = \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}$$

$$\gamma^\mu_{\text{standard}} = S \ \gamma^\mu_{\text{helicity}} \ S^{-1}$$

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = S^{-1}$$

Spinors: Classical solutions of the Dirac equation

$$\begin{array}{lc}
\text{Notation:} & \not p \doteq p_\mu \gamma^\mu \\
\text{Adjoints:} & \bar u = u^+\gamma^0 \ , \ \ \bar v = v^+\gamma^0 \\
\text{Dirac equation:} & (\not p - m)\, u = 0 \ , \ \ (\not p + m)\, v = 0
\end{array}$$

$$\begin{array}{lrcl}
\text{Normalization:} & \bar u(p,r)\gamma^\mu u(p,r') & = & 2\,p^\mu \delta_{rr'} \\
& \bar u(p,r) v(p,r') & = & 0 \\
& \bar v(p,r) u(p,r') & = & 0 \\
& \bar v(p,r)\gamma^\mu v(p,r') & = & 2\,p^\mu \delta_{rr'} \\
& \bar u(p,r) u(p,r) & = & 2m\,\delta_{rr'} \\
& \bar v(p,r) v(p,r) & = & -2m\,\delta_{rr'} \\
\text{Completeness:} & \sum_r u(p,r)\bar u(p,r) & = & \not p + m \\
& \sum_r v(p,r)\bar v(p,r) & = & \not p - m
\end{array}$$

Relation to spinors in rest frame:

$$\begin{array}{rcl}
u(p,r) & = & \dfrac{1}{\sqrt{p^0+m}}(\not p + m)\, \tilde u(0,r) \\
v(p,r) & = & \dfrac{1}{\sqrt{p^0+m}}(\not p - m)\, \tilde v(0,r) \\
\tilde u(0,r) & = & \begin{pmatrix} U_n(r) \\ 0 \end{pmatrix} \ ; \ \ \tilde v(0,r) = \begin{pmatrix} 0 \\ V_n(r) \end{pmatrix}
\end{array}$$

Properties of two-spinors:

$$\begin{array}{rcl}
V_n(r) & = & -i\sigma_2 U_n(r) \\
U^+(r)\, U(r') & = & \delta_{rr'} \\
V^+(r)\, V(r') & = & \delta_{rr'} \\
\vec\sigma\cdot\vec n\, U_n(\pm) & = & \pm\, U_n(\pm) \\
\vec\sigma\cdot\vec n\, V_n(\pm) & = & -\pm\, V_n(\pm)
\end{array}$$

$$\vec n \text{ quantization axis} \begin{cases} \text{standard representation } \vec n = (0,0,1) \\ \text{helicity representation } \vec n = \vec p/|\vec p\,| \end{cases}$$

Representations for spinors:

Notation for two-spinors:

$$\begin{array}{ll}
U(\tfrac12) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & U(-\tfrac12) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
V(\tfrac12) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & V(-\tfrac12) = -\begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{array}$$

a. Standard representation:

$$u(p,r) = \sqrt{p^0+m} \quad \begin{pmatrix} U(r) \\ \frac{\vec{\sigma}\cdot\vec{p}}{p^0+m} U(r) \end{pmatrix}$$

$$v(p,r) = \sqrt{p^0+m} \quad \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{p^0+m} V(r) \\ V(r) \end{pmatrix}$$

b. Helicity representation: $\sigma_\mu \doteq (\mathbf{1}, \vec{\sigma}\,), \hat{\sigma}_\mu = (\mathbf{1}, -\vec{\sigma}\,)$

$$u_h(p,r) = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{p^0+m}} \begin{pmatrix} (\sigma_\mu p^\mu + m) & U(r) \\ (\hat{\sigma}_\mu p^\mu + m) & U(r) \end{pmatrix}$$

$$v_h(p,r) = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{p^0+m}} \begin{pmatrix} (\sigma_\mu p^\mu + m) & V(r) \\ -(\hat{\sigma}_\mu p^\mu + m) & V(r) \end{pmatrix}$$

Projection operators:

1. <u>Covariant spin projection operators</u> $\Pi_\pm$:
   For a fermion of momentum p

$$\Pi_\pm = \frac{1}{2}\,(1 \pm \gamma_5 \not{n})$$

define covariant spin projection operators. n is a space like unit vector orthogonal to p

$$n^2 = -1\;;\;\; n \cdot p = 0\;.$$

Its general form is

$$n = L_{\vec{p}}\,(0, \vec{\xi}\,) = \left(\frac{\vec{p}\cdot\vec{\xi}}{m}\;,\;\; \vec{\xi} + \frac{\vec{p}\cdot\vec{\xi}}{m(p^0+m)}\vec{p} \right)$$

where $\vec{\xi}^{\,2} = 1$, $\vec{\xi}$ the direction of polarization in the rest frame.

$$\text{Projection property:} \quad \begin{aligned} &\Pi_+ + \Pi_- = 1 \\ &\Pi_\pm^2 = \Pi_\pm \\ &\Pi_+\Pi_- = \Pi_-\Pi_+ = 0 \end{aligned}$$

The projection operators are self–adjoint

$$\Pi_\pm = \gamma^0 \;(\Pi_\pm)^+ \;\gamma^0$$

with respect to the scalar product $\bar{v}u = v^+\gamma^0 u$ in four–spinor space[42].

Physical interpretation:

$$\begin{aligned} \Pi_\pm u(p,s) &= u(p,s)\,\delta_{s,\pm\frac{1}{2}} \\ \Pi_\pm v(p,s) &= v(p,s)\,\delta_{s,\pm\frac{1}{2}} \end{aligned}$$

[42]The Lorentz invariant scalar product for the space of four–spinors is given by $(v,u) \to \bar{v}u \equiv v^+\gamma^0 u$. The adjoint $\mathcal{O}^\dagger$ of an operator $\mathcal{O}$ is defined by $\mathcal{O}^\dagger = \gamma^0\mathcal{O}^+\gamma^0$. On four–spinor space the usual Hermitecity is replaced by the self–adjointness $\mathcal{O}^\dagger = \mathcal{O}$. As usual, then (u, Hu) is real if $H = \gamma^0 H^+ \gamma^0$.

2. <u>Particle and antiparticle projection operators</u> $\Lambda_\pm$
   The matrices

$$\Lambda_\pm = \frac{1}{2m}\left(\pm\not{p} + m\right)$$

   project to particle solutions (positive frequencies) and antiparticle solutions (negative frequencies), respectively.

$$\text{Projection property:}\quad \begin{aligned} &\Lambda_+ + \Lambda_- = 1 \\ &\Lambda_\pm^2 = \Lambda_\pm \\ &\Lambda_+\Lambda_- = \Lambda_-\Lambda_+ = 0 \end{aligned}$$

Physical interpretation:

particle	antiparticle
$\Lambda_+ u(p,s) = u(p,s)$	$\Lambda_+ v(p,s) = 0$
$\Lambda_- u(p,s) = 0$	$\Lambda_- v(p,s) = v(p,s)$

Representation in terms of spinors:

particle with polarization r:	$u_\alpha(p,r)\bar{u}_\beta(p,r) = \frac{1}{2}\left\{(\not{p}+m)\left(1+\gamma_5\not{n}_r\right)\right\}_{\alpha\beta}$
antiparticle with polarization r:	$v_\alpha(p,r)\bar{v}_\beta(p,r) = \frac{1}{2}\left\{(\not{p}-m)\left(1+\gamma_5\not{n}_r\right)\right\}_{\alpha\beta}$
unpolarized particle:	$\sum_r u_\alpha(p,r)\bar{u}_\beta(p,r) = (\not{p}+m)_{\alpha\beta}$
unpolarized antiparticle:	$\sum_r v_\alpha(p,r)\bar{v}_\beta(p,r) = (\not{p}-m)_{\alpha\beta}$

Linearly independent basis of 4x4 matrices:

$$\Gamma_{(r)} = \{1, i\gamma_5, \gamma_\mu, \gamma_\mu\gamma_5, \sigma_{\mu\nu}\}\ ;\ \ r = S, P, V, A, T$$

The labels r denote scalar (S), pseudoscalar (P), vector (V), axial vector (A) and tensor (T) matrices.

Adjoint spinor field:

$$\bar{\psi} \doteq \psi^+\gamma_0$$

defined such that

$$\bar{\psi}\Gamma_{(r)}\psi \text{ is hermitean}$$

and, as indicated by r, transform as scalar (S), pseudoscalar (P), vector (V), axial vector (A) and tensor (T) under Lorentz transformations.

Green functions: solutions of the homogeneous (no source) or inhomogeneous (point source) Dirac equations

$$\begin{aligned} \{\psi_\alpha(x), \psi_\beta(y)\} &= i\,S_{\alpha\beta}(x-y;m) \\ <0\,|\,T\,\{\psi_\alpha(x), \bar{\psi}_\beta(y)\}\,|\,0> &= i\,S_{F\alpha\beta}(x-y;m) \\ S_{\alpha\beta}(z;m) &= \left(i\gamma^\mu\partial_\mu + m\right)_{\alpha\beta}\Delta(z;m^2) \\ S_{F\alpha\beta}(z;m) &= \left(i\gamma^\mu\partial_\mu + m\right)_{\alpha\beta}\Delta_F(z;m^2) \end{aligned}$$

$$\begin{aligned}(\, i\gamma^\mu \partial_{x\mu} - m\,)_{\alpha\alpha'} S_{\alpha'\beta}(x-y;m) &= 0 \\ (\, i\gamma^\mu \partial_{x\mu} - m\,)_{\alpha\alpha'} S_{F\alpha'\beta}(x-y;m) &= \delta_{\alpha\beta}\delta^{(4)}(x-y)\end{aligned}$$

Charge conjugation is mapping particle into antiparticle creation and annihilation operators and vice versa:

$$a(\vec{p},r) \leftrightarrow b(\vec{p},r) \quad , \qquad a^+(\vec{p},r) \leftrightarrow b^+(\vec{p},r) \quad ,$$

up to a phase. For the Dirac field charge conjugation reads (see 2.38)

$$\psi_\alpha(x) \to \mathcal{C}_{\alpha\beta}\bar{\psi}^T_\beta(x)$$

with

$$\mathcal{C} = i\left(\gamma^2\gamma^0\right) = -i\begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix} \quad .$$

Properties of $\mathcal{C}$ are:

$$\mathcal{C}^T = -\mathcal{C} \quad , \quad \mathcal{C}\gamma^\mu\mathcal{C}^{-1} = -\gamma^\mu \quad ,$$

and for the spinors charge conjugation takes the form

$$(\mathcal{C}u)^T = \bar{v} \;\; \text{and} \quad (\mathcal{C}v)^T = \bar{u} \quad ,$$

which may be verified by direct calculation.

A.4 Neutrino field: representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$

$$i\gamma^\mu\partial_\mu\psi(x) = 0$$

Dirac field in limit $m \to 0$ (exists if normalized appropriately). Describes a particle of mass 0, spin $\frac{1}{2}$, particle $\neq$ antiparticle, spin orientation $r = \pm\frac{1}{2}$ relative to $\vec{p}$ (helicity states). Subspaces with r fixed are invariant under $\mathcal{P}^\uparrow_+$ i. e. a massless Dirac field decomposes into two fields of fixed chirality ($\equiv$ handedness). Applying the chiral projectors $\Pi_\pm = \frac{1}{2}(1 \pm \gamma_5)$ (which satisfy $\Pi_+ + \Pi_- = 1$, $\Pi^2_\pm = \Pi_\pm$ and $\Pi_+\Pi_- = \Pi_-\Pi_+ = 0$) we obtain the two fields:

$\psi_R = \Pi_+\psi$ right-handed field ($\Pi_-\psi_R = 0$)

describes $\begin{cases} \text{particle with spin} \uparrow\uparrow \text{momentum} \\ \text{antiparticle with spin} \uparrow\downarrow \text{momentum} \end{cases}$

$\psi_L = \Pi_-\psi$ left-handed field ($\Pi_+\psi_L = 0$)

describes $\begin{cases} \text{particle with spin} \uparrow\downarrow \text{momentum} \\ \text{antiparticle with spin} \uparrow\uparrow \text{momentum} \end{cases}$

in Nature: $\psi^{\text{phys}}_\nu = \psi_{\nu L}$ “two component theory of the neutrino.”

A.5 Real vector field: representation $(\frac{1}{2},\, \frac{1}{2})$

$$(\Box + m^2)V_\mu(x) = 0\ ;\ \ \partial^\mu V_\mu(x) = 0\ ;\ \ V_\mu^* = V_\mu$$

Describes a particle of mass m, spin 1, particle = antiparticle.
Examples: Z, ρ^0 .

$$\underline{V_\mu(x) = \sum_{r=\pm 1,0} \int d\mu(p) \left\{ \epsilon_\mu(\vec{p}, r)\, a(\vec{p}, r)\, e^{-ipx} + \epsilon_\mu^*(\vec{p}, r)\, a^+(\vec{p}, r)\, e^{ipx} \right\}}$$

$$\begin{aligned} r &= \pm 1 \quad \text{transversal degrees of freedom} \\ r &= 0 \quad \text{longitudinal degree of freedom} \end{aligned}$$

Canonical commutation relations:

$$[a(\vec{p}, r), a^+(\vec{p}\,', r')] = [b(\vec{p}, r), b^+(\vec{p}\,', r')] = (2\pi)^3\, 2\omega_p\, \delta^{(3)}(\vec{p} - \vec{p}\,')\, \delta_{rr'}$$

and all other commutators vanishing.

Polarization vectors: Classical solutions of the field equations

Normalization:	$\epsilon_\mu(p, r)\, \epsilon^{\mu *}(p, r') = -\delta_{rr'}$
Completeness:	$\sum_r \epsilon_\mu(p, r)\, \epsilon_\nu^*(p, r) = -g_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}$
Absence of scalar mode:	$p_\mu \epsilon^\mu(\vec{p}, r) = 0$

The explicit form of the polarization vectors $\epsilon_\mu(p, r)$ depends on the choice of the quantization axis $\vec{n}$ for the spin:

a) Canonical (Cartesian) basis: J_3 diagonal in the rest frame

$$\epsilon^0(p, r) = \frac{p_r}{m}\ ;\ \ \epsilon^k(p, r) = \delta_{kr} + p_k p_r \frac{1}{m(p^0 + m)}$$

b) Helicity basis: the relationship to the canonical basis is analogous to the relation between Cartesian coordinates and the spherical harmonics Y_{1m} :

$$\begin{aligned} (x, y, z) \to r\, (Y_{11}, Y_{1-1}, Y_{10}) &= \sqrt{\frac{3}{4\pi}} \left(\frac{-x - iy}{\sqrt{2}}, \frac{x - iy}{\sqrt{2}}, z \right) \\ \epsilon_h^\mu(p, +) &= \frac{1}{\sqrt{2}} (-\epsilon^\mu(p, 1) - i\epsilon^\mu(p, 2)) \\ \epsilon_h^\mu(p, -) &= \frac{1}{\sqrt{2}} (\ \ \epsilon^\mu(p, 1) - i\epsilon^\mu(p, 2)) \\ \epsilon_h^\mu(p, 0) &= \epsilon^\mu(p, 0) \end{aligned}$$

Green functions: solutions of the homogeneous (no source) or inhomogeneous (point source) Proca equations

$$\begin{aligned} [V^\mu(x), V^\nu(y)] &= iD^{\mu\nu}(x - y; m^2) \\ <0 \mid T\, \{V^\mu(x), V^\nu(y)\} \mid 0> &= iD_F^{\mu\nu}(x - y; m^2) \\ D^{\mu\nu}(z; m^2) &= \left\{ -g^{\mu\nu} - \frac{\partial^\mu \partial^\nu}{m^2} \right\} \Delta(z; m^2) \\ D_F^{\mu\nu}(z; m^2) &= \left\{ -g^{\mu\nu} - \frac{\partial^\mu \partial^\nu}{m^2} \right\} \Delta_F(z; m^2) \end{aligned}$$

$$\begin{aligned}
\{(\,\Box_x + m^2\,)\,g_{\mu\mu'} - \partial_{x\mu}\partial_{x\mu'}\,\}\,D^{\mu'\nu}(x-y;m^2) &= 0 \\
\{(\,\Box_x + m^2\,)\,g_{\mu\mu'} - \partial_{x\mu}\partial_{x\mu'}\,\}\,D_F^{\mu'\nu}(x-y;m^2) &= \delta_\mu^{\ \nu}\delta^{(4)}(x-y)
\end{aligned}$$

A.6 Complex vector field: representation $(\frac{1}{2},\,\frac{1}{2})$

Like a complex scalar field, a complex vector field may be defined in terms of two real fields. Let $V_1^\mu(x)$ and $V_2^\mu(x)$ be two real vector fields of equal mass m. Then

$$W_\mu^\pm(x) = \frac{1}{\sqrt{2}}\,(V_{1\mu}(x) \mp iV_{2\mu}(x))$$

are charged vector fields. Describe particles of mass m, spin 1, particle $\neq$ antiparticle. Examples: $W^\pm$, $\rho^\pm$.

$$\begin{aligned}
W^{\mu-}(x) &= \sum_{r=\pm1,0}\int d\mu(p)\left\{\epsilon_-^\mu(\vec{p},r)\,a(\vec{p},r)\,e^{-ipx} + \epsilon_+^\mu(\vec{p},r)\,b^+(\vec{p},r)\,e^{ipx}\right\} \\
W^{\mu+}(x) &= \left(W^{\mu-}(x)\right)^*
\end{aligned}$$

The polarization vectors are given by

$$\epsilon_\pm^\mu = \frac{1}{\sqrt{2}}\,(\epsilon_1^\mu \mp i\epsilon_2^\mu)\ ;\ \ \epsilon_+^\mu = \epsilon_-^{\mu *}$$

and satisfy

Normalization:	$\epsilon_{\mu-}(p,r)\,\epsilon_+^\mu(p,r') = -\delta_{rr'}$
Completeness:	$\sum_r \epsilon_{\mu-}(p,r)\,\epsilon_{\nu+}(p,r) = -g_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}$
Absence of scalar mode:	$p_\mu\epsilon_\pm^\mu(\vec{p},r) = 0$

A.7 Photon:

In principle the photon may be described in a gauge invariant manner by the antisymmetric electromagnetic field strength tensor$F^{\mu\nu}$ and the field equations

$$\Box F_{\mu\nu} = 0\ ;\ \ \partial_\mu F^{\mu\nu} = 0\ ;\ \ \partial_\mu\tilde{F}^{\mu\nu} = 0$$

where $\tilde{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$ is the dual pseudotensor.

Describes a particle of mass 0, spin 1 with two transversal degrees of freedom.
Examples: γ, "gluons".

Because local gauge invariance requires the matter fields to couple to the gauge potentials (interaction $j_{em}^\mu A_\mu$), a massless spin 1 gauge field must be described by a vector potential A_μ (see Sec. 4) [43]. Since $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the homogeneous Maxwell-equation $\partial_\mu\tilde{F}^{\mu\nu} = 0$ is satisfied

[43] Another point is that the inhomogeneous Maxwell equation

$$\partial_\mu F^{\mu\nu} = j^\nu$$

as a vector equation and the homogeneous Maxwell equation

$$\partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} + \partial_\mu F_{\nu\lambda} = 0$$

as a $3rd$ rank tensor equation cannot be obtained as the Euler-Lagrange equations from an invariant Lagrangian if we have available a as a dynamical variable the $2nd$ rank tensor $F_{\mu\nu}$ only as an identity. A_μ is not unique (gauge dependent) and in general transforms as a Lorentz vector up to a divergence only!

$$\underline{A_\mu(x) = \sum_{r=\pm} \int d\mu(p) \left\{ \varepsilon_\mu(p,r)\, a(\vec{p},r)\, e^{-ipx} + h.c. \right\}}$$

Polarization vectors: Solution of the classical field equations.

Normalization:	$\varepsilon_\mu(p,r)\varepsilon^{\mu *}(p,r') = -\delta_{rr'}$
Completeness:	$\sum_{r=\pm} \varepsilon_\mu(p,r)\varepsilon^*_\nu(p,r) = -g_{\mu\nu} + p_\mu f_\nu + p_\nu f_\mu$
Absence of scalar mode:	$p_\mu \varepsilon^\mu(p,r) = 0$

The arbitrary "vector" f_μ depends on the choice of the gauge. We must require physical amplitudes to be gauge invariant and hence to be independent of f_μ. Gauge invariance, i.e. invariance under $A_\mu \to A_\mu - \partial_\mu \alpha(x)$, amounts to the invariance under the substitutions

$$\varepsilon_\mu \to \varepsilon_\mu + \lambda\, p_\mu \;;\;\; \lambda \text{ an arbitrary constant}$$

of the polarization vectors. One can prove that the polarization "vectors" for massless spin 1 fields can not be covariant. The non-covariant terms are always proportional to p_μ, however.
Green functions:

$$\begin{aligned} [A_\mu(x), A_\nu(y)] &= i\, D_{\mu\nu}\,(x-y) \\ < 0 \mid T\,\{A_\mu(x), A_\nu(y)\} \mid 0 > &= i\, D_{F\mu\nu}\,(x-y) \end{aligned}$$

where

$$\begin{aligned} D_{\mu\nu}\,(z) &= -g_{\mu\nu}\,\Delta(z;0) - \partial_\mu F_\nu(z) - \partial_\nu F_\mu(z) \\ D_{F\mu\nu}\,(x-y) &= -g_{\mu\nu}\,\Delta_F(z;0) - \partial_\mu F_{F\nu}(z) - \partial_\nu F_{F\mu}(z) \end{aligned}$$

with $F_\mu(z)$ and $F_{F\mu}(z)$ gauge dependent functions (not covariant in general).

B The group $SL(2,C)$ and spinor representations

When discussing relativistic invariance we have to distinguish between the classical and the quantum mechanical level:

- classical: Lorentz group $\Lambda \in L_+^{\uparrow}$ and covariant quantities like scalars, vectors, tensors
- quantum mechanical: unimodular group $U \in SL(2,C)$ with "covariant" quantities called spinors.

There will be two types of spinors:

- spinors of even rank which may be represented by tensors, and which transform one-to-one under $L_+^{\uparrow}$,
- spinors of odd rank, which cannot be represented by tensors, and which transform two-to-one under $L_+^{\uparrow}$.

Starting point are the non-trivial representations of lowest dimension: namely, the spin 1/2 representations by 2×2 matrices. The corresponding angular momentum algebra is

$$\begin{aligned} \left[J_i^{(\frac{1}{2})}, J_k^{(\frac{1}{2})} \right] &= i\varepsilon_{ikl} J_l^{(\frac{1}{2})} \\ \left(\vec{J}^{(\frac{1}{2})} \right)^2 &= j\,(j+1)\,\mathbf{1} = \frac{3}{4} \cdot \mathbf{1} \ . \end{aligned}$$

The explicit representation is provided by the Pauli matrices: $\vec{J}^{(\frac{1}{2})} = \vec{\sigma}/2$.

B.1 The group $SL(2,C)$

There is a direct relationship between the proper orthochronous Lorentz group $L_+^{\uparrow}$ and the group $SL(2,C)$ of complex 2×2 matrices with determinant 1. The two groups have the same Lie algebra (infinitesimal transformations) but have a different parameter space (global transformations). The relationship between $L_+^{\uparrow}$ and $SL(2,C)$ may be constructed in the following way: let us denote by

$$\sigma_\mu = (\mathbf{1}, \vec{\sigma}\,)$$

the four-vector of Hermitean 2×2 matrices composed from the unit matrix and the Pauli matrices[44]:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$

[44] They have the properties:

$$\begin{aligned} \sigma_i^+ = \sigma_i \ , \quad \sigma_i^2 = \mathbf{1} \ , \quad \vec{\sigma}^2 = 3 \cdot \mathbf{1} \\ \sigma_i \sigma_k = \frac{1}{2}\{\sigma_i, \sigma_k\} + \frac{1}{2}[\sigma_i, \sigma_k] = \delta_{ik} + i\,\varepsilon_{ikl}\,\sigma_l \\ \sum_{k=1}^{3} (\sigma_k)_{ij}\,(\sigma_k)_{mn} = 2\,\delta_{in}\delta_{jm} - \delta_{ij}\delta_{mn} \end{aligned}$$

To any real four-vector x^μ we can associate a Hermitean 2×2 matrix given by

$$x_\mu \to X = x^\mu \sigma_\mu = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \quad .$$

Conversely, every Hermitean 2×2 matrix X determines a real four-vector

$$X \to x^\mu = \frac{1}{2} \mathrm{Tr}\ (X\sigma^\mu) \quad .$$

Thus we have a linear one-to-one correspondence between real four-vectors and Hermitean 2×2 matrices. Furthermore we note that

$$\det X = x^2 = x^\mu x_\mu \quad .$$

This establishes the following

Theorem 1: The mapping

$$x^\mu = \frac{1}{2} \mathrm{Tr}\ (X\sigma^\mu) \ \in \mathbf{M_4} \leftrightarrow X = x^\mu \sigma_\mu \ \in \mathbf{H(2)}$$

establishes a linear isomorphism (one-to-one mapping) between the Minkowski space $\mathbf{M_4}$ and the space of two-dimensional Hermitean matrices $\mathbf{H(2)}$.

Definition 1: The complex 2×2 matrices U with determinant $\det U = 1$ with matrix multiplication as a composition law form a group: the **special linear group** or unimodular linear group in two complex dimensions $SL(2, C)$.

In the following we will discuss the relation between the groups $SL(2, C)$ and $L_+^\uparrow$. An element $U \in SL(2, C)$ provides a mapping

$$X \to X' = UXU^+ \quad \text{i.e.} \quad x'^\mu \sigma_\mu = x^\nu U \sigma_\nu U^+$$

between Hermitean matrices, which preserves the determinant

$$\det X' = \det U \det X \det U^+ = \det X \quad .$$

The transformation

$$X \to X' = UXU^+ \quad \text{on } \mathbf{H(2)}$$

corresponds to the real linear transformation

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu \, x^\nu \quad \text{on } \mathbf{M_4}$$

which satisfies $x'^\mu x'_\mu = x^\mu x_\mu$ and therefore is a Lorentz transformation. The correspondence

$$x'^\mu = \Lambda^\mu{}_\nu \, x^\nu \to x'^\mu \sigma_\mu = \Lambda^\mu{}_\nu \, x^\nu \sigma_\mu = x^\nu U_\Lambda \sigma_\nu U_\Lambda^+$$

implies

$$\boxed{U_\Lambda \sigma_\nu U_\Lambda^+ = \Lambda^\mu{}_\nu \, \sigma_\mu}$$

and proves the following

Theorem 2a: The Hermitean 2×2 matrices $\sigma_\mu = (\mathbf{1}, \vec{\sigma}\,)$ form a covariant four-vector with respect to the representation $U_\Lambda = D^{(\frac{1}{2},0)} \equiv D$.

The transformation matrix U_Λ may be written in the canonical form[45]

$$U_\Lambda = U(\vec{\chi}, \vec{\omega}\,) = e^{\vec{\chi}\frac{\vec{\sigma}}{2}}\, e^{i\vec{\omega}\frac{\vec{\sigma}}{2}} \tag{B.1}$$

as a unitary rotation times a Hermitean boost.

From the above representation we see that the two *different* $SL(2,C)$ transformations

$$U(\vec{\chi}, \vec{n}\theta) \;\text{ and }\; U(\vec{\chi}, \vec{n}\,(\theta+2\pi)) = -U(\vec{\chi}, \vec{n}\theta)$$

represent the same Lorentz transformation. Therefore, U_Λ is a double-valued representation of $L_+^\uparrow$, i.e., the mapping

$$SL(2,C) \to L_+^\uparrow$$

is two-to-one. $SL(2,C)$ itself is simply connected. The two elements $Z_2 = \{\mathbf{1}, -\mathbf{1}\}$ form an invariant subgroup, $UZ_2U^{-1} = Z_2\;\; \forall U \in SL(2,C)$, and we find the isomorphism

$$SL(2,C)/Z_2 \simeq L_+^\uparrow \;\; .$$

The main results obtained so far may we summarized in the following

Theorem 3: The group $SL(2,C)$ is simply connected. The homomorphism (many-to-one mapping) $U_\Lambda \in SL(2,C) \to \Lambda \in L_+^\uparrow$ is two-to-one.

A group G is called **universal covering group** of another group $\tilde{G}$ if G is simply connected and has the same Lie-algebra as $\tilde{G}$. Thus

- $SL(2,C)$ is the universal covering group of $L_+^\uparrow$.

---

[45] The following theorem holds:
Theorem: Each matrix $U \in SL(2,C)$ may be written as a product $U = HV$ where $H = H^+$ is Hermitean and $V = V^{+-1}$ is unitary.
Corollary: Let $U = HV$ then $\bar{U} = \left(U^+\right)^{-1} = \left((HV)^+\right)^{-1} = \left(V^+H\right)^{-1} = H^{-1}V$
Proof: We note the following properties:

1) $UU^+ = e^{\vec{\chi}\vec{\sigma}}$ is Hermitean, $\vec{\chi}$ real,
2) $\left(UU^+\right)^{-1/2} = e^{-\vec{\chi}\frac{\vec{\sigma}}{2}}$ exists,
3) $\left[\left(UU^+\right)^{-1/2} U\right] \left[\left(UU^+\right)^{-1/2} U\right]^+ = 1$,
   which implies that $\left(UU^+\right)^{-1/2} U$ is unitary and has determinant 1; consequently:
4) $\left(UU^+\right)^{-1/2} U = e^{-\vec{\chi}\frac{\vec{\sigma}}{2}} U = e^{i\vec{\omega}\frac{\vec{\sigma}}{2}}$, $\vec{\omega}$ real .

The result is

$$U_\Lambda = U(\vec{\chi}, \vec{\omega}\,) = e^{\vec{\chi}\frac{\vec{\sigma}}{2}}\, e^{i\vec{\omega}\frac{\vec{\sigma}}{2}}$$

with $\vec{\chi}$ and $\vec{\omega}$ real.
Special cases are:

U Hermitean: $U = U^+ = H = e^{\vec{\chi}\frac{\vec{\sigma}}{2}}$ pure Lorentz boost,

U unitary: $U = U^{+-1} = V = e^{i\vec{\omega}\frac{\vec{\sigma}}{2}}$ pure rotation.

The unitary elements U form the subgroup $SU(2)$ of $SL(2,C)$.

The main point is that

- unitary representations of $L_+^\uparrow$ may always be considered as faithful (one-to-one) representations of $SL(2,C)$.

The actual meaning of this fact will become clear below, when considering spinors.

Remark: The relationship between $L_+^\uparrow$ and $SL(2,C)$ is analogue to the one between the rotation group O_3 and the unimodular unitary group $SU(2)$, which is the universal covering group of O_3.

Definition 2: With D also $\bar{D} \equiv (D^+)^{-1}$ is a representation of a group, which we call **conjugate representation** . Let D be the irreducible representation $D = D^{(\frac{1}{2},0)}$ of $SL(2,C)$ then $D^{(0,\frac{1}{2})} \equiv \bar{D}$ is an inequivalent (a new) representation of $SL(2,C)$ [46].

If U_Λ is given by Eq. (B.1) we obtain for the conjugate representation

$$\bar{U}_\Lambda = U^+_{\Lambda^{-1}} = e^{-\vec{\chi}\frac{\vec{\sigma}}{2}}\, e^{i\vec{\omega}\frac{\vec{\sigma}}{2}} \quad . \tag{B.2}$$

While σ_μ is a covariant vector with respect to the representation $D = U_\Lambda$, the vector

$$\hat{\sigma}_\mu = (\mathbf{1}, -\vec{\sigma}\,)$$

is a covariant vector with respect to the representation $\bar{D} = \bar{U}_\Lambda$:

$$\boxed{\bar{U}_\Lambda \hat{\sigma}_\mu \bar{U}^+_\Lambda = \Lambda^\nu{}_\mu\, \hat{\sigma}_\nu}$$

and thus we have

Theorem 2b: The Hermitean 2×2 matrices $\hat{\sigma}_\mu = (\mathbf{1}, -\vec{\sigma}\,)$ form a covariant four-vector with respect to the representation $\bar{U}_\Lambda = D^{(0,\frac{1}{2})} \equiv \bar{D}$.

The representations D and $\bar{D}$ are the two inequivalent non-trivial representations of lowest dimension, called **fundamental spinor representations** .

Interrelation between D and $\bar{D}$: We first notice the the inverse of $U = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ with $\det U = \alpha\delta - \beta\gamma = 1$ is $U^{-1} = \left(\begin{smallmatrix}\delta & -\beta\\ -\gamma & \alpha\end{smallmatrix}\right)$ and we observe that the matrix

$$C = e^{i\pi J^{(\frac{1}{2})}} = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -C^+ = -C^{-1} \tag{B.3}$$

provides the transformations

$$U^{-1} = C^{-1} U^T C \quad , \quad \bar{U} = U^{+-1} = C^{-1} U^* C \tag{B.4}$$

and

$$\hat{\sigma}_\mu = C^{-1} \sigma_\mu^T C \quad . \tag{B.5}$$

A scalar product on the space $\mathbf{H(2)}$ is defined by

$$X = x^\mu \sigma_\mu \quad , \quad \hat{Y} = y^\mu \hat{\sigma}_\mu \Rightarrow \frac{1}{2} \mathrm{Tr}\,(X\hat{Y}) = x\cdot y \quad .$$

[46] Two representations of a group G are called **equivalent representations** if one follows from the other by a change of basis in the representation space: Thus there exists a non-singular transformation S such that $\bar{D}(g) = S^{-1} D(g) S\ \forall g \in G$.

It is thus identical to the known scalar product in Minkowski space.

One easily checks the following useful identities:

$$\begin{aligned}
\mathrm{Tr}\,(\sigma_\mu \hat{\sigma}_\nu) &= 2g_{\mu\nu} && \text{(B.6)}\\
\sigma_\mu \hat{\sigma}_\nu + \hat{\sigma}_\nu \sigma_\mu &= 2g_{\mu\nu} \cdot \mathbf{1} && \text{(B.7)}\\
\frac{1}{2}\mathrm{Tr}\left(\sigma^\mu U_\Lambda \hat{\sigma}_\nu U_\Lambda^+\right) &= \Lambda^\mu{}_\nu \quad . && \text{(B.8)}
\end{aligned}$$

The last identity tells us, once more, that there are always two different $SL(2,C)$ elements $\pm U_\Lambda$ which map into one element Λ in $L_+^\uparrow$.

We finally mention that the group $SL(2,C)$ extended by the translations is denoted by $iSL(2,C)$, the inhomogeneous linear unimodular group, which is related to $\mathcal{P}_+^\uparrow$ in the same way as $SL(2,C)$ is related to $L_+^\uparrow$.

B.2 Spinors

The advantage to work directly with $SL(2,C)$ becomes obvious once we deal with spinors: A transformation $U \in SL(2,C)$ may be understood as a mapping

$$U : \ V \to V$$

of a two-dimensional complex vector space V onto itself. The "complex two-vectors"

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in V$$

are called **spinors**. The spinors

$$u(\uparrow) = \begin{pmatrix} u_1 \\ 0 \end{pmatrix} \quad \text{and} \quad u(\downarrow) = \begin{pmatrix} 0 \\ u_2 \end{pmatrix}$$

are eigenvectors to the eigenvalues $\pm 1/2$ of the 3rd component of the spin matrix $\sigma_3/2$:

$$\frac{\sigma_3}{2}u(\uparrow) = +\frac{1}{2}u(\uparrow) \ \text{ and } \ \frac{\sigma_3}{2}u(\downarrow) = -\frac{1}{2}u(\downarrow) \quad .$$

Since there are two types of inequivalent representations we have to distinguish two types of spinors, called undotted and dotted spinors.

Definition 3a: A covariant **undotted spinor** is a spinor which transforms according to the representation $(\frac{1}{2}, 0)$, i.e., it is characterized by the transformation law

$$u_a \in V \to u'_a = U_{\Lambda a}{}^b \, u_b \in V \ \ ; \ \ a, b = 1, 2 \quad . \tag{B.9}$$

Definition 3b: A contravariant **dotted spinor** is a spinor which transforms according to the representation $(0, \frac{1}{2})$, i.e., it is characterized by the transformation law

$$u^{\dot{a}} \in V \to u'^{\dot{a}} = \bar{U}_\Lambda{}^{\dot{a}}{}_{\dot{b}} \, u^{\dot{b}} \in V \ \ ; \ \ \dot{a}, \dot{b} = 1, 2 \quad . \tag{B.10}$$

Next we consider **scalar products** and the **metric** in spinor space:

Definition 4a: Let v_a be a covariant undotted spinor. A *contravariant undotted spinor* is defined to be a undotted spinor u^a with the property

$$u^a v_a = \text{ invariant } \ (v \in V) \ .$$

From

$$u'^a v'_a = u'^a \, U_{\Lambda a}{}^b \, v_b = u^b v_b$$

we find the transformation law

$$u'^a \, U_{\Lambda a}{}^b \ = u^b \Rightarrow u'^a = U_\Lambda^{-1T a}{}_b \, u^b = \bar{U}_\Lambda^{*}{}^a{}_b \, u^b$$

for the contravariant undotted spinor.

Definition 4b: Let $v^{\dot{a}}$ be a contravariant dotted spinor. A *covariant dotted spinor* is defined to be a dotted spinor $u_{\dot{a}}$ with the property

$$v^{\dot{a}} u_{\dot{a}} = \text{ invariant } \ (v \in V) \ .$$

Here, from

$$v'^{\dot{a}} u'_{\dot{a}} = \bar{U}_\Lambda{}^{\dot{a}}{}_{\dot{b}} \, v^{\dot{b}} u'_{\dot{a}} = v^{\dot{b}} u_{\dot{b}}$$

we obtain

$$\bar{U}_\Lambda{}^{\dot{a}}{}_{\dot{b}} \, u'_{\dot{a}} = u_{\dot{b}} \Rightarrow u'_{\dot{a}} = \bar{U}_\Lambda^{-1T}{}_{\dot{a}}{}^{\dot{b}} \, u_{\dot{b}} = U^*_{\Lambda \dot{a}}{}^{\dot{b}} \, u_{\dot{b}}$$

for the transformation law of the covariant dotted spinor.

Equipped with these scalar products, we may define metric spinors g^{ab} and $g^{\dot{a}\dot{b}}$, which allow us to obtain the contravariant spinor's from the covariant one's by

$$u^a = g^{ab} u_b \ , \quad u^{\dot{a}} = g^{\dot{a}\dot{b}} u_{\dot{b}} \ .$$

Since the transpose of the inverse of a transformation matrix is related to the original transformation matrix according to Eq. (B.4), we infer that the metric we are looking for must be given by the matrix C Eq. (B.3). Thus the metric is given by the two-dimensional antisymmetric tensor

$$g^{ab} = (i\sigma_2)^{ab} = \varepsilon^{ab} = -\varepsilon_{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

which is numerically invariant. It is the only numerically invariant tensor for the group $SL(2, C)$, besides the Kronecker symbol δ_a^b. The scalar product thus may be written as

$$u^a v_a = \varepsilon^{ab} u_a v_b = u^1 v_2 + u^2 v_1 = u_1 v_2 - u_2 v_1 = \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \ .$$

Similarly, with

$$g^{\dot{a}\dot{b}} = (i\sigma_2)^{\dot{a}\dot{b}} = \varepsilon^{\dot{a}\dot{b}} = -\varepsilon_{\dot{a}\dot{b}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

we may write

$$u^{\dot{a}} v_{\dot{a}} = \varepsilon^{\dot{a}\dot{b}} u_{\dot{a}} v_{\dot{b}} = u^{\dot{1}} v_{\dot{2}} + u^{\dot{2}} v_{\dot{1}} = u_{\dot{1}} v_{\dot{2}} - u_{\dot{2}} v_{\dot{1}} = \det \begin{pmatrix} u_{\dot{1}} & v_{\dot{1}} \\ u_{\dot{2}} & v_{\dot{2}} \end{pmatrix} \quad .$$

Notice the antisymmetry of the scalar product

$$\boxed{u^a v_a = -u_a v^a}$$

valid for undotted or dotted indices. Thus $u^a u_a = 0$ for any $u \in V$.

We finally list the different types of fundamental spinors and their transformation properties:

spinor	representation
u_a	$D = D^{(\frac{1}{2},0)}$
$u_{\dot{a}}$	D^*
u^a	$\bar{D}^* = (D^{-1})^T$
$u^{\dot{a}}$	$\bar{D} = D^{(0,\frac{1}{2})} = (D^+)^{-1}$

Notice that by Eq. (B.4) D^* is equivalent to $\bar{D}$ and $\bar{D}^*$ is equivalent to D.

Spinors of higher rank

Since we have to distinguish two different inequivalent representations as $SL(2,C)$ transformation laws, a higher rank spinor is characterized by two integers n and m. A "spinor–tensor" of rank (n, m) is an object which transforms like a product of n undotted two–spinors and m dotted two–spinors:

$$\psi_{a_1 \ldots a_n \dot{b}_1 \ldots \dot{b}_m} \sim u_{a_1} \ldots u_{a_n} v_{\dot{b}_1} \ldots v_{\dot{b}_m} \quad .$$

Properties which hold for each type of indices separately are the following:

- Indices may be raised and lowered by means of the metric tensor
- Lower indices may be contracted with upper indices of the same kind in an invariant way

Thus contraction of mixed indices is not an invariant procedure and hence is not meaningful. Also, note that from one spinor one cannot get an non-trivial Hermitean from, because $u_a u^a = 0$ due to the antisymmetry.

All higher spins may be constructed from the fundamental representations. As an example we consider the construction of a spin 1 field.

Spin 1 representation

The usual representation for a spin 1 field is $(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$. It may be constructed as follows:

From an undotted two-spinor u_a a vector field may be defined by

$$V^\mu \doteq u^+ \hat{\sigma}^\mu u \quad .$$

Under a Lorentz transformation $u \to u' = U_\Lambda u$ we obtain

$$V'^\mu = u'^+ \hat{\sigma}^\mu u' = u^+ U_\Lambda^+ \hat{\sigma}^\mu U_\Lambda u = \Lambda^\mu{}_\nu \, u^+ \hat{\sigma}^\nu u = \Lambda^\mu{}_\nu \, V^\nu$$

and hence, V^μ indeed is a contravariant vector. The ordinary Lorentz transformation is thus equivalent to a transformation $u, u^* = (u^+)^T \to U_\Lambda u, U_\Lambda^* u^*$, which is equivalent to the representation $(\frac{1}{2}, \frac{1}{2})$.

An equivalent representation may be obtained using a dotted spinor. Let $v^{\dot{a}}$ be a dotted two-spinor, which transforms like $v \to v' = \bar{U}_\Lambda v$, then the vector field defined by

$$\hat{V}^\mu \doteq v^+ \sigma^\mu v$$

again is a contravariant vector (i.e., in the representation $(\frac{1}{2}, \frac{1}{2})$).

Parity and reflexion of spinors

The parity P, as we know, acts on four–vectors like $Px = (x^0, -\vec{x}\,)$ and satisfies $P^2 = 1$. With respect to the rotation group O_3 P is just a rotation by the angle 2π and thus in the context of the rotation group P has no special meaning. This is different for the Lorentz group. While

$$U_P \vec{J} = \vec{J} U_P$$

commutes

$$U_P \vec{K} = -\vec{K} U_P$$

does not commute. As a consequence, considering the representations Eqs. (B.1) and (B.2), we learn that

$$U_P U(\vec{\chi}, \vec{n}\,\theta) = U(-\vec{\chi}, \vec{n}\,\theta) U_P$$

and hence

$$U_P U_\Lambda = \bar{U}_\Lambda U_P \quad . \tag{B.11}$$

We thus have the

Theorem 4: Under parity a undotted spinor is transformed into a dotted spinor and vice versa.

The position of the indices of U_P may be read off from the position of the indices of the transformation matrices U_Λ and $\bar{U}_\Lambda$. Since U_P commutes with $\vec{J}$, and in particular with the z–component of the spin J_3, each component transforms individually, which means that U_P acts diagonal. Furthermore, $P^2 = 1$ implies $U_P^2 = \pm 1$ or

$$\boxed{U_P = u_P = \pm 1 \quad \text{or} \quad \pm\, i} \quad .$$

This follows from the correspondences $x^\mu \to X$, $Px^\mu \to U_P X U_P$ and $P^2 x^\mu = x^\mu \to U_P^2 X U_P^2 = X$ and U_P diagonal, which implies that we may write it as a c–number phase u_P, simply.

B.3 Bispinors

Parity can only be a symmetry if undotted and dotted spinors enter a theory in a symmetric fashion. This means that P–invariance requires undotted and dotted spinors to be grouped in pairs into a single object which is called bispinor.

Definition 5: A **bispinor** is a four–spinor

$$u_\alpha \doteq \begin{pmatrix} u_a \\ v^{\dot{b}} \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ v^{\dot{1}} \\ v^{\dot{2}} \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

which transforms according to the reducible representation $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$

$$u_\alpha \to u'_\alpha = \begin{pmatrix} U_\Lambda & 0 \\ 0 & \bar{U}_\Lambda \end{pmatrix}_{\alpha\beta} u_\beta = (S_\Lambda u)_\alpha$$

and under parity

$$u_\alpha = \begin{pmatrix} u_a \\ v^{\dot{b}} \end{pmatrix} \to u_P \begin{pmatrix} v^{\dot{a}} \\ u_b \end{pmatrix} = u_P \begin{pmatrix} v^{\dot{1}} \\ v^{\dot{2}} \\ u_1 \\ u_2 \end{pmatrix} = u_P \begin{pmatrix} u_3 \\ u_4 \\ u_1 \\ u_2 \end{pmatrix}$$

We may rewrite the parity operation acting on a bispinor as follows:

Definition 6: γ^0 denotes the 4×4 matrix

$$\gamma^0 \doteq \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \quad . \tag{B.12}$$

It acts on bispinors as

$$\gamma^0 \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} u_3 \\ u_4 \\ u_1 \\ u_2 \end{pmatrix} .$$

With the help of γ^0 a parity transformation may be written in the compact form

$$\boxed{u_\alpha \to u_P\, {\gamma^0}_{\alpha\beta} u_\beta} \quad . \tag{B.13}$$

Properties of γ^0 are

$$(\gamma^0)^2 = \mathbf{1} \ , \quad \gamma^0 = \gamma^{0\ +} = \gamma^{0\ -1} \quad . \tag{B.14}$$

Definition 7: γ_5 denotes the 4×4 matrix

$$\gamma_5 \doteq \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \quad . \tag{B.15}$$

It acts on bispinors as

$$\gamma_5 \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ -u_3 \\ -u_4 \end{pmatrix} .$$

Thus the undotted part of a bispinor is an eigenvector to the eigenvalue +1, the dotted part is eigenvector to the eigenvalue −1:

$$\gamma_5 \begin{pmatrix} u_a \\ 0 \end{pmatrix} = \begin{pmatrix} u_a \\ 0 \end{pmatrix} \quad ; \quad \gamma_5 \begin{pmatrix} 0 \\ v^{\dot b} \end{pmatrix} = - \begin{pmatrix} 0 \\ v^{\dot b} \end{pmatrix} \ .$$

Other properties of γ_5 are

$$(\gamma_5)^2 = \mathbf{1} \ , \quad \gamma_5 = \gamma_5^+ = \gamma_5^{-1} \ . \tag{B.16}$$

Definition 8: The 4×4 matrices

$$\Pi_\pm \doteq \frac{1}{2}\,(1 \pm \gamma_5) \tag{B.17}$$

are projection operators

$$\Pi_+ = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & 0 \end{pmatrix} \quad ; \quad \Pi_- = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{1} \end{pmatrix}$$

for the upper and lower components of a bispinor, respectively; they are called **chiral projectors**. Thus

$$\Pi_+ \begin{pmatrix} u_a \\ v^{\dot b} \end{pmatrix} = \begin{pmatrix} u_a \\ 0 \end{pmatrix} \quad ; \quad \Pi_- \begin{pmatrix} u_a \\ v^{\dot b} \end{pmatrix} = \begin{pmatrix} 0 \\ v^{\dot b} \end{pmatrix}$$

are projections to the irreducible blocks of the bispinors.

Right–handed and left–handed bispinors

With the help of the chiral projectors acting on bispinors we may reformulate the notion of undotted and dotted spinors as right–handed and left–handed bispinors, respectively.

Definition 9: Given a bispinor u_α, we define

$$u_{R\alpha} \doteq \Pi_+ u_\alpha \quad \text{and} \quad u_{L\alpha} \doteq \Pi_- u_\alpha \ . \tag{B.18}$$

We call u_R **right-handed** (bi)spinor and u_L **left–handed** (bi)spinor.

Later, when discussing spinor fields, we will see that the "handedness" is directly related to the helicity of a particle. Right–handed refers to positive helicity and left–handed refers to negative helicity.

We note that there is a perfect equivalence between right–handed (bi)spinors and undotted spinors and left–handed (bi)spinors and dotted spinors. In the physics literature commonly the bispinor language together with the handedness is used instead of the undotted and dotted two–spinor terminology.

Properties of the representation S_Λ

The reducible representation

$$S_\Lambda \equiv \begin{pmatrix} U_\Lambda & 0 \\ 0 & \bar{U}_\Lambda \end{pmatrix} = e^{i(\vec{\chi}\,\vec{\mathcal{K}} + \vec{\omega}\,\vec{\mathcal{J}}\,)} \tag{B.19}$$

has the generators (see (B.1) and (B.2))

$$\begin{array}{llcl} 1) & \text{Rotations} & : \vec{\mathcal{J}} & \equiv \begin{pmatrix} \vec{J} & 0 \\ 0 & \vec{J} \end{pmatrix} \\ 2) & \text{Boosts} & : \vec{\mathcal{K}} & \equiv i \begin{pmatrix} \vec{J} & 0 \\ 0 & -\vec{J} \end{pmatrix} \end{array}$$

with $\vec{J} = \vec{\sigma}/2$. With the help of γ_5 we may write

$$\vec{\mathcal{K}} = i\,\gamma_5\,\vec{\mathcal{J}} \ .$$

A crucial property of the representation S_Λ is the following:

Theorem 5: S_Λ^{+-1} is equivalent to S_Λ .

Indeed,

$$S_\Lambda^{+-1} = S_{\Lambda^{-1}}^{+} = \gamma^0 S_\Lambda \gamma^0 \ , \tag{B.20}$$

which is easily verified. Thus, there exist only one such representation . One consequence is that we may obtain an invariant Hermitean form from one single bispinor. It is convenient to first introduce an adjoint bispinor:

Definition 10: to a bispinor u_α we associate the **adjoint bispinor**

$$\bar{u}_\alpha \doteq \left(u^+\gamma^0\right)_\alpha = (u_3^*, u_4^*, u_1^*, u_2^*) \ .$$

We then have the

Theorem 6: the bilinear form

$$\bar{u}_\alpha u_\alpha = u_\alpha^+ {\gamma^0}_{\alpha\beta} u_\beta \ ,$$

where repeated spinor indices are summed over, is *Hermitean and Lorentz invariant.*

The proof is simple and follows from the properties of γ^0 and S_Λ which we have stated above.

B.4 Boosts and rotations

Let us consider now the lowest dimensional non-trivial representation of a Lorentz boost

$$D^{(\frac{1}{2})}\left(L(\vec{p}\,)\right) = e^{\vec{\chi}\frac{\vec{\sigma}}{2}} = \mathbf{1}\,\cosh\frac{\chi}{2} + \vec{\sigma}\cdot\vec{n}\,\sinh\frac{\chi}{2}$$

where $\chi = |\vec{\chi}\,|$ and $\vec{n} = \vec{\chi}/|\vec{\chi}\,|$ the unit vector in direction of $\vec{\chi}$. To prove the last equality we note that the boost matrix is Hermitean and that any Hermitean 2×2 matrix may be written in the form: $h = \alpha\mathbf{1} + \vec{\beta}\,\vec{\sigma}$ with real coefficients α, β_i. Expanding the exponential we have

$$e^{\vec{\chi}\frac{\vec{\sigma}}{2}} = \sum_{n=0}^{\infty} \frac{(\chi/2)^n}{n!} n^{i_1}\sigma_{i_1}\cdots n^{i_n}\sigma_{i_n}$$

Using $n^i n^k \sigma_i \sigma_k = \frac{1}{2} n^i n^k \{\sigma_i, \sigma_k\} = n^i n^k \delta_{ik}\cdot\mathbf{1} = \vec{n}^2\ \cdot\mathbf{1} = \mathbf{1}$ we obtain

$$e^{\vec{\chi}\frac{\vec{\sigma}}{2}} \ = \ \mathbf{1}\sum_{k=0}^{\infty}\frac{1}{(2k)!}\left(\frac{\chi}{2}\right)^{2k} + \vec{\sigma}\cdot\vec{n}\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\left(\frac{\chi}{2}\right)^{2k+1}$$

from which the result follows.

For a special Lorentz transformation $L(\vec{p}\,)$ the parameter χ is the hyperbolic angle determined by

$$\cosh\chi = \frac{p^0}{m} \quad , \quad \sinh\chi = \frac{|\vec{p}\,|}{m} \quad , \quad \vec{n} = \hat{\vec{p}} = \frac{\vec{p}}{|\vec{p}\,|} \quad .$$

The hyperbolic functions satisfy $\cosh^2\chi - \sinh^2\chi = 1$ and for half-angles we have

$$\cosh\frac{\chi}{2} = \left(\frac{\cosh\chi + 1}{2}\right)^{1/2} \quad , \quad \sinh\frac{\chi}{2} = \left(\frac{\cosh\chi - 1}{2}\right)^{1/2} \quad .$$

Using these relations we find

$$\cosh\frac{\chi}{2} = \sqrt{\frac{p^0+m}{2m}} \quad , \quad \sinh\frac{\chi}{2} = \sqrt{\frac{p^0+m}{2m}}\frac{|\vec{p}\,|}{p^0+m}$$

in terms of the boost momentum. We thus obtain the important representation

$$\begin{aligned} D^{(\frac{1}{2})}\left(L(\vec{p}\,)\right) = e^{\vec{\chi}\frac{\vec{\sigma}}{2}} &= \frac{1}{\sqrt{2m\,(p^0+m)}}\left((p^0+m)\mathbf{1} + \vec{p}\cdot\vec{\sigma}\,\right) \\ &= \frac{1}{\sqrt{2m\,(p^0+m)}}\left(p^\mu\sigma_\mu + m\right) \end{aligned} \qquad \text{(B.21)}$$

for the Lorentz boost, and furthermore we may write

$$D\left(L(\vec{p}\,)\right) D^+\left(L(\vec{p}\,)\right) = \left(D\left(L(\vec{p}\,)\right)\right)^2 = e^{\vec{\chi}\cdot\vec{\sigma}} = \mathbf{1}\,\cosh\chi + \hat{\vec{p}}\cdot\vec{\sigma}\,\sinh\chi = \frac{1}{m}\,p^\mu\sigma_\mu \quad , \qquad \text{(B.22)}$$

which is *Lorentz invariant.*

The second inequivalent lowest dimensional non-trivial representation of a Lorentz boost is

$$\begin{aligned} \bar{D}^{(\frac{1}{2})}\left(L(\vec{p}\,)\right) = e^{-\vec{\chi}\frac{\vec{\sigma}}{2}} &= \frac{1}{\sqrt{2m\,(p^0+m)}}\left((p^0+m)\mathbf{1} - \vec{p}\cdot\vec{\sigma}\,\right) \\ &= \frac{1}{\sqrt{2m\,(p^0+m)}}\left(p^\mu\hat{\sigma}_\mu + m\right) \end{aligned} \qquad \text{(B.23)}$$

and

$$\bar{D}\left(L(\vec{p}\,)\right)\bar{D}^+\left(L(\vec{p}\,)\right) = \left(D\left(L(\vec{p}\,)\right)\right)^2 = e^{-\vec{\chi}\cdot\vec{\sigma}} = \frac{1}{m}\left(p^0\,\mathbf{1} - \vec{p}\cdot\vec{\sigma}\,\right) = \frac{1}{m}\,p^\mu\hat{\sigma}_\mu \qquad \text{(B.24)}$$

again is *Lorentz invariant.*

Finally, the lowest dimensional non-trivial representation of a rotation reads

$$D^{(\frac{1}{2})}\left(R(\vec{\omega}\,)\right) = e^{i\vec{\omega}\frac{\vec{\sigma}}{2}} = \mathbf{1}\,\cos\frac{\omega}{2} + i\,\vec{\sigma}\cdot\hat{\vec{\omega}}\,\sin\frac{\omega}{2}$$

where $\omega = |\vec{\omega}\,|$ and $\hat{\vec{\omega}} = \vec{\omega}/|\vec{\omega}\,|$ a unit vector in direction of $\vec{\omega}$. Notice that this $SU(2) \subset SL(2,C)$ rotation is a rotation by half of the angle, only, of the corresponding classical O_3 rotation.

B.5 Transformation laws of the annihilation and creation operators

Let us consider massive states in the canonical basis $|\vec{p}, j_3>$ which transform according to Eq. (2.2), which implies

$$U(\Lambda, a)\ a^{+}(\vec{p}, j_3)\, U^{-1}(\Lambda, a) = e^{ia\Lambda p} D^{(j)}\left(R_{\Lambda,p}\right)_{j_3' j_3} a^{+}(\Lambda\vec{p}, j_3')\ , \tag{B.25}$$

where we used $U(\Lambda, a)|0>=|0>$ and dropped the vacuum state, from which the one–particle state was created by the creation operator. For the annihilation operator we thus have

$$U(\Lambda, a)\ a(\vec{p}, j_3)\, U^{-1}(\Lambda, a) = e^{-ia\Lambda p} D^{(j)}\left(R^{-1}_{\Lambda,p}\right)_{j_3 j_3'} a(\Lambda\vec{p}, j_3')\ . \tag{B.26}$$

A first obvious problem is that the annihilation and the creation operator transform in a different way and we have to bring the transformation laws into the same form in order to be able to get a field with simple transformation properties. To this end we use Eq. (B.4). Since for a pure rotation R the representation $D(R)$ is unitary, we have

$$D^{(j)}(R)^{*} = C D^{(j)}(R) C^{-1}\ . \tag{B.27}$$

The matrix C for arbitrary spin is a $(2j+1)\times(2j+1)$ matrix with the properties

$$C^{*}C = (-1)^{2j}\ ;\quad C^{+}C = 1\ ,$$

and, explicitly, we have

$$C_{j_3 j_3'} = \left(e^{i\pi J_2^{(j)}}\right)_{j_3 j_3'} = (-1)^{j+j_3}\delta_{j_3,-j_3'}\ . \tag{B.28}$$

Utilizing these properties of C together with the unitarity of $D^{(j)}(R)$ we observe that we may write

$$D^{(j)}(R)_{j_3' j_3} = \left\{C D^{(j)}(R^{-1}) C^{-1}\right\}_{j_3 j_3'}\ . \tag{B.29}$$

Thus we may rewrite Eq.(B.25) in the form

$$U(\Lambda, a)\ a^{+}(\vec{p}, j_3)\, U^{-1}(\Lambda, a) = e^{ia\Lambda p}\left\{C D^{(j)}\left(R^{-1}_{\Lambda,p}\right) C^{-1}\right\}_{j_3' j_3} a^{+}(\Lambda\vec{p}, j_3')\ , \tag{B.30}$$

which directly compares to the transformation law of the creation operator Eq. (B.26).

We notice that the transformation matrix has a complicated p–dependence, such that the Fourier transforms of the creation and the annihilation operators have themselves no simple transformation law in configuration space. Therefore, one has to construct suitable p–independent linear combinations before taking the Fourier transform. The new basis of creation and annihilation operators is called **spinor basis**. We will denote operators in the spinor basis by a tilde: $\tilde{a}^{+}(\vec{p}, j_3)$, $\tilde{a}(\vec{p}, j_3)$ and require them to transform under $SL(2, C)$ like

$$\begin{aligned}
U(\Lambda)\ \tilde{a}\ (\vec{p}, j_3)\, U^{-1}(\Lambda) &= D^{(j)}\left(\Lambda^{-1}\right)_{j_3 j_3'} \tilde{a}\ (\Lambda\vec{p}, j_3') \\
U(\Lambda)\ \tilde{a}^{+}(\vec{p}, j_3)\, U^{-1}(\Lambda) &= D^{(j)}\left(\Lambda^{-1}\right)_{j_3 j_3'} \tilde{a}^{+}(\Lambda\vec{p}, j_3')\ .
\end{aligned}$$

We note that $D^{(j)}(\Lambda)$ cannot be unitary unless $j=0$. How is this possible? In order to understand this we have just to look at the relationship between the canonical and the spinor basis. We have

$$\begin{aligned}
D^{(j)}\left(R^{-1}_{\Lambda,p}\right) &= D^{(j)}\left(L^{-1}(\vec{p})\Lambda^{-1}L(\Lambda\vec{p})\right) \\
&= D^{(j)}\left(L^{-1}(\vec{p})\right) D^{(j)}\left(\Lambda^{-1}\right) D^{(j)}\left(L(\Lambda\vec{p})\right)
\end{aligned}$$

where the left hand side is a $(2j+1)$ dimensional representation of a rotation, which is unitary, while on the right hand side we have a product of $(2j+1)$ dimensional non–unitary representations of the $SL(2,C)$. The point is that only the product must be unitary not the individual factors. If we multiply the last relation from the left with $D^{(j)}\left(L(\vec{p}\,)\right)$ we indeed achieve that

$$
\begin{aligned}
U\left(\Lambda\right) D^{(j)}\left(L(\vec{p}\,)\right)_{j_3 j_3'}\, a\left(\vec{p}, j_3'\right) U^{-1}\left(\Lambda\right) &= D^{(j)}\left(\Lambda^{-1}\right)_{j_3 j_3'} D^{(j)}\left(L(\Lambda\vec{p}\,)\right)_{j_3' j_3''}\, a\left(\Lambda\vec{p}, j_3''\right) \\
U\left(\Lambda\right) \left\{D^{(j)}\left(L(\vec{p}\,)\right) C^{-1}\right\}_{j_3 j_3'} a^{+}(\vec{p}, j_3')\, U^{-1}\left(\Lambda\right) &= D^{(j)}\left(\Lambda^{-1}\right)_{j_3 j_3'} \left\{D^{(j)}\left(L(\Lambda\vec{p}\,)\right) C^{-1}\right\}_{j_3' j_3''} a^{+}(\Lambda\vec{p}, j_3'')
\end{aligned}
$$

such that we can identify the appropriate operators in the spinor basis

$$
\begin{aligned}
\tilde{a}\left(\vec{p}, j_3\right) &= \left(D^{(j)}\left(L(\vec{p}\,)\right)\right)_{j_3 j_3'} a\left(\vec{p}, j_3'\right) \\
\tilde{a}^{+}(\vec{p}, j_3) &= \left(D^{(j)}\left(L(\vec{p}\,)\right) C^{-1}\right)_{j_3 j_3'} a^{+}(\vec{p}, j_3')\ .
\end{aligned}
\tag{B.31}
$$

This is the crucial result. We now have creation and annihilation operators[47] which have the same p–independent transformation law and the Fourier transform of a linear combination of the general form

$$
\xi\, \tilde{a}\ \left(\vec{p}, j_3; m, j\right) e^{-ipx} + \eta\, \tilde{b}^{+}\left(\vec{p}, j_3; m, j\right) e^{ipx}
$$

and thus

$$
\int d\mu(p) \left\{\xi\, \tilde{a}(\vec{p}, j_3) e^{-ipx} + \eta\, \tilde{b}^{+}(\vec{p}, j_3) e^{ipx}\right\}
$$

satisfy the required simple transformation law. As indicated, the creation operator b^+ only need have the same mass and spin as the annihilation operator a, not however the same charge–like quantum numbers.

B.6 Fields

Now basic fields may be written down for all the irreducible representations of $SL(2,C)$. The simplest one's are the following:

Fields to the representation $(j,0)$

$$
\varphi_{j_3}(x) = \sum_{j_3'} \int d\mu(p) \left\{\xi D^{(j)}\left(L(\vec{p}\,)\right)_{j_3 j_3'} a(\vec{p}, j_3')\, e^{-ipx} + \eta\, \hat{D}^{(j)}\left(L(\vec{p}\,)\right)_{j_3 j_3'} b^{+}(\vec{p}, j_3')\, e^{ipx}\right\}
\tag{B.32}
$$

where $D^{(j)}\left(L(\vec{p}\,)\right)$ and $\hat{D}^{(j)}\left(L(\vec{p}\,)\right) \equiv D^{(j)}\left(L(\vec{p}\,)\right) C^{-1}$ represent one particle wave functions. We will write them down in a more familiar form below.

Fields to the representation $(0,j)$

$$
\chi_{j_3}(x) = \sum_{j_3'} \int d\mu(p) \left\{\xi' D^{(j)}\left(L(-\vec{p}\,)\right)_{j_3 j_3'} a(\vec{p}, j_3')\, e^{-ipx} + \eta' \hat{D}^{(j)}\left(L(-\vec{p}\,)\right)_{j_3 j_3'} b^{+}(\vec{p}, j_3')\, e^{ipx}\right\}
\tag{B.33}
$$

This field is obtained from the $(j,0)$ field by the substitutions $D^{(j)}\left(L(\vec{p}\,)\right) \to D^{(j)}\left(L(-\vec{p}\,)\right)$ and by choosing independent linear combination coefficients ξ' and η'. Both fields satisfy the Klein–Gordon equation (2.18) and transform according to the $(2j+1)$–dimensional non–unitary irreducible representations specified (see Eq. (2.19)). These irreducible fields, also called Joos–Weinberg fields, do not obey any field equation other than the Klein–Gordon equation.

[47] Note that $\tilde{a}^{+} \neq (\tilde{a})^{+}$ in the spinor basis.

With the help of the commutation relations between annihilation and creation operators the field commutator (anti–commutator) becomes

$$\left[\varphi_{j_3}(x), \varphi^+_{j'_3}(x')\right]_\mp = \int d\mu(p) \left(D^{(j)}\left(L(\vec{p}\,)\right) D^{(j)}\left(L(\vec{p}\,)\right)^+\right)_{j_3 j'_3} \left\{|\xi|^2 e^{-ip(x-x')} \mp |\eta|^2 e^{ip(x-x')}\right\} \quad . \quad \text{(B.34)}$$

This result may be obtained as follows:

1) Each field exhibits a summation over j_3 and an integration over d^3p

$$\sum_{j''_3} \sum_{j'''_3} \int d\mu(p) \int d\mu(p') \cdots$$

2) The non–vanishing commutators (anti–commutators) between annihilation and creation operators under the integrals are

$$\left[a(\vec{p}, j''_3), a^+(\vec{p}\,', j'''_3)\right]_\mp \;=\; \delta_{j''_3 j'''_3} \, (2\pi)^3 \, 2\omega_p \, \delta^{(3)}\left(\vec{p} - \vec{p}\,'\right)$$

and

$$\left[b^+(\vec{p}, j''_3), b(\vec{p}\,', j'''_3)\right]_\mp \;=\; \mp\delta_{j''_3 j'''_3} \, (2\pi)^3 \, 2\omega_p \, \delta^{(3)}\left(\vec{p} - \vec{p}\,'\right) \quad .$$

3) The coefficients of the two terms are

$$D^{(j)}\left(L(\vec{p}\,)\right)_{j_3 j''_3} D^{+(j)}\left(L(\vec{p}\,')\right)_{j'''_3 j'_3} \xi\xi^* e^{-ipx} e^{ip'x'}$$

and

$$\left(D^{(j)}\left(L(\vec{p}\,)\right) C^{-1}\right)_{j_3 j''_3} \left(D^{(j)}\left(L(\vec{p}\,')\right) C^{-1}\right)^+_{j'''_3 j'_3} \eta\eta^* e^{ipx} e^{-ip'x'} \quad ,$$

respectively.

4) One of the integrations is trivial:

$$\int d\mu(p')(2\pi)^3 \, 2\omega_p \, \delta^{(3)}\left(\vec{p} - \vec{p}\,'\right) F(\vec{p}, \vec{p}\,') = F(\vec{p}, \vec{p}\,) \quad ,$$

and because of $m = m'$, we note that $\vec{p} = \vec{p}\,'$ implies $\omega_{p'} = \omega_p$ and $p' = p$. Furthermore, also the j'''_3 summation is trivial, as $\sum_{j'''_3} \cdots \delta_{j''_3 j'''_3} = \cdots\Big|_{j'''_3 = j''_3}$.

5) What remains is the integral $\int d\mu(p) \cdots$ over the two terms as given in Eq. (B.34).

Using the known form of the boost matrices we obtain

$$\left(D^{(j)}\left(L(\vec{p}\,)\right) D^{(j)}\left(L(\vec{p}\,)\right)^+\right) = e^{2\vec{\chi}\vec{J}^{\,(j)}} = e^{2\chi\hat{\vec{p}}\vec{J}^{\,(j)}}$$

with $\cosh\chi = \omega_p/m$ and $\hat{\vec{p}} = \vec{p}/|\vec{p}\,|$. This matrix may be easily computed for arbitrary j, by expanding the exponential and using the angular momentum algebra. For $j = 0$ we have 1 for $j = 1/2$ we found before $\frac{1}{m}\, p^\mu \sigma_\mu$ and for arbitrary spin we find a result of the form

$$\left(e^{2\chi\hat{\vec{p}}\vec{J}^{\,(j)}}\right)_{j_3 j'_3} = \frac{1}{m^{2j}}\, t^{\mu_1 \ldots \mu_{2j}}_{j_3 j'_3} p_{\mu_1} \ldots p_{\mu_{2j}}$$

where t is a constant, symmetric and traceless tensor. Equation (B.34) thus may be written in the form

$$\left[\varphi_{j_3}(x), \varphi^+_{j'_3}(x')\right]_\mp = \frac{i^{2j}}{m^{2j}}\, t^{\mu_1 \ldots \mu_{2j}}_{j_3 j'_3} \partial_{\mu_1} \ldots \partial_{\mu_{2j}} \int d\mu(p) \left\{|\xi|^2 e^{-ip(x-x')} \mp (-1)^{2j} |\eta|^2 e^{ip(x-x')}\right\} \quad . \quad \text{(B.35)}$$

Note the appearance of the crucial factor $(-1)^{2j}$ in the second term now, which comes from the different sign of the argument in the exponential, when we replace the p_μ factors by derivatives with respect to x: $(p_\mu \to \pm i\partial_\mu)\ \exp \mp ipx$.

We thus have derived the form Eq. (2.25) of the commutator (anti–commutator) and the requirement of locality implies the spin statistics (2.26) and the crossing (2.27) theorems, which require $\pm(-1)^{2j} = 1$ and $|\xi| = |\eta|$, respectively. For further discussion we refer to Sec. 2.4.

For the commutator (anti–commutator) of the $(0, j)$ fields we just have to replace D by $\bar{D}$ in (B.34) and we obtain

$$\left[\chi_{j_3}(x), \chi^+_{j'_3}(x')\right]_\mp = \frac{i^{2j}}{m^{2j}}\, \hat{t}^{\mu_1 \ldots \mu_{2j}}_{j_3 j'_3}\, \partial_{\mu_1} \ldots \partial_{\mu_{2j}} \int d\mu(p) \left\{|\xi'|^2 e^{-ip(x-x')} \mp (-1)^{2j} |\eta'|^2 e^{ip(x-x')}\right\} \quad . \tag{B.36}$$

where $\hat{t}$ is obtained from t by replacing $\vec{p}$ by $-\vec{p}$ in the definition of t. Thus we may write $\hat{t}^{\mu_1 \ldots \mu_{2j}}_{j_3 j'_3} \equiv t_{j_3 j'_3\, \mu_1 \ldots \mu_{2j}}$. The locality requirement for the $(0, j)$ fields yield restrictions on the coefficients ξ' and η', which have the same forms as the one's obtained for ξ and η in case of the $(j, 0)$ fields. So far the constraints on ξ' and η' are independent from the one's on ξ and η.

Interesting is the relative locality of the $(j, 0)$ and the $(0, j)$ fields. For this we have to look at

$$\left[\varphi_{j_3}(x), \chi^+_{j'_3}(x')\right]_\mp = \delta_{j_3 j'_3} \int d\mu(p) \left\{\xi \xi'^* e^{-ip(x-x')} \mp \eta \eta'^* e^{ip(x-x')}\right\} \quad , \tag{B.37}$$

which is obtained in the same way as (B.34). The only change is the replacement $D \to \bar{D}$ in the second factor of the D's in DD^+ and the substitutions $\xi \to \xi'$ and $\eta \to \eta'$ for one of the fields. Since $D\bar{D}^+ = DD^{-1} = 1$ the result follows.

Locality requires $\eta\eta'^* = \xi\xi'^*$ for bosons and $\eta\eta'^* = -\xi\xi'^*$ for fermions, thus

$$\eta\eta'^* = (-1)^{2j} \xi\xi'^* \quad .$$

This implies that the antiparticle operator of the $(0, j)$ field must carry a phase $(-1)^{2j}$ with respect to the antiparticle operator in the $(j, 0)$ field ! An admissible normalization thus is

$$\xi' = \xi = \eta = 1 \quad \text{and} \quad \eta' = (-1)^{2j} \quad . \tag{B.38}$$

We thus may set $\xi = \eta = 1$ in the $(j, 0)$ field (B.32) and obtain the $(0, j)$ field (B.33) from the $(j, 0)$ one by the substitutions $D^{(j)}(L(\vec{p})) \to D^{(j)}(L(-\vec{p}))$ and $b^+ \to (-1)^{2j} b^+$.

With this choice of the phases the fields are local and local relative to each other, i.e., the commutators (anti–commutators)

$$[\ \varphi_{j_3}(x),\ \varphi^{(+)}_{j'_3}(y)] \ , \quad [\ \chi_{j_3}(x),\ \chi^{(+)}_{j'_3}(y)] \quad \text{and} \quad [\ \varphi_{j_3}(x),\ \chi^{(+)}_{j'_3}(y)]$$

vanish for $(x - y)^2 < 0$. This requirement is crucial for the Dirac field, which combines the two fields $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ into one reducible local field.

B.7 One particle wave functions for spin 1/2

In the following we use the labels $r = j_3$ for the spinor components. Alternatively, we may use the labeling $r = 1, 2$ or $\uparrow, \downarrow$ or $+\frac{1}{2}, -\frac{1}{2}$ or $+, -$.

In the fields spinors of the form

$$D^{(j)}(L(\vec{p}))_{rr'}\ a(\vec{p}, r')$$

show up, and if we write

$$\begin{pmatrix} a(\vec{p},\uparrow) \\ a(\vec{p},\downarrow) \end{pmatrix} = U(\uparrow)\, a(\vec{p},\uparrow) + U(\downarrow)\, a(\vec{p},\downarrow)$$

we obtain the more familiar from

$$u(p,r'\,)\, a(\vec{p},r'\,) = D^{(j)}\left(L(\vec{p}\,)\right) U(r'\,)\, a(\vec{p},r'\,)$$

where summation over r' is understood. This identifies

$$u(p,r) = D^{(j)}\left(L(\vec{p}\,)\right) U(r)$$

as the relativistic wave function for the particles described by $a(\vec{p},r)$.

The conventional argumentation for the construction of the wave functions goes as follows: the quantization for massive particles is performed by convention in the rest frame and the states are eigenstates of the 3rd component of angular momentum J_3. As mentioned above, this yields the basic normalized two-component spinors:

$$U(\uparrow) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad U(\downarrow) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad ; \quad U^+(r)\, U(r') = \delta_{rr'} \quad .$$

The wave function for the r component of $a(\vec{p},r)$, in the $(\frac{1}{2},0)$ field, is the <u>undotted spinor</u> obtained from the spinor $U(r)$ at rest by the appropriate boost. Thus, using Eq. (B.21), we obtain

$$\begin{aligned} u(p,r) &= D^{(\frac{1}{2})}\left(L(\vec{p}\,)\right)\, U(r) \\ &= \frac{1}{\sqrt{2m\,(p^0+m)}}\,(p^\mu\sigma_\mu + m)\, u(0,r) \end{aligned} \tag{B.39}$$

with

$$u(0,r) = U(r)$$

in the rest frame.

Also the wave function for the r component of $b^+(\vec{p},r)$ is an <u>undotted spinor</u>. Again it is obtained from the spinor $U(r)$ at rest by a corresponding boost Eq. (B.31):

$$\begin{aligned} v(p,r) &= D^{(\frac{1}{2})}\left(L(\vec{p}\,)\right)\, (-i\sigma_2)\, U(r) \\ &= D^{(\frac{1}{2})}\left(L(\vec{p}\,)\right)\, V(r) \\ &= \frac{1}{\sqrt{2m\,(p^0+m)}}\,(p^\mu\sigma_\mu + m)\, v(0,r) \end{aligned} \tag{B.40}$$

with

$$v(0,r) = V(r) \equiv -i\sigma_2 U(r)$$

the rest frame spinors. Explicitly,

$$V(\uparrow) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad V(\downarrow) = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad ; \quad V^+(r)\, V(r') = \delta_{rr'} \quad .$$

We may summarize our considerations by writing down the following

Result: Spin $\frac{1}{2}$ fields

The representation $(\frac{1}{2},0)$ is described by the undotted spinor field

$$\varphi_a(x)=\sum_r\int d\mu(p)\left\{u_a(p,r)\,a(\vec{p},r)\,e^{-ipx}+v_a(p,r)\,b^+(\vec{p},r)\,e^{ipx}\right\}\quad . \tag{B.41}$$

The conjugate representation $(0,\frac{1}{2})$, obtained by the replacements $D^{(\frac{1}{2})}\left(L(\vec{p}\,)\right)\to D^{(\frac{1}{2})}\left(L(-\vec{p}\,)\right)$ and $b^+\to(-1)^{2j}b^+$ in the $(\frac{1}{2},0)$ field, is described by the dotted spinor field

$$\chi^{\dot{a}}(x)=\sum_r\int d\mu(p)\left\{\hat{u}^{\dot{a}}(p,r)\,a(\vec{p},r)\,e^{-ipx}+\hat{v}^{\dot{a}}(p,r)\,b^+(\vec{p},r)\,e^{ipx}\right\}\quad . \tag{B.42}$$

where

$$\begin{aligned}\hat{u}^{\dot{a}}(p,r)&=\frac{1}{\sqrt{2m\,(p^0+m)}}\,(p^\mu\hat{\sigma}_\mu+m)\,u(0,r)\\ \hat{v}^{\dot{a}}(p,r)&=\frac{-1}{\sqrt{2m\,(p^0+m)}}\,(p^\mu\hat{\sigma}_\mu+m)\,v(0,r)\end{aligned} \tag{B.43}$$

are the dotted spinor one–particle wave functions. Note that $\vec{p}\to-\vec{p}$ is accounted for by going from σ_μ to $\hat{\sigma}_\mu$, while $b^+\to-b^+$ is taken into account by a change of sign in the corresponding wave function $\hat{v}$.

The two–spinor fields satisfy the following linear coupled system of first order differential equations:

$$\begin{aligned}i(\hat{\sigma}^\mu)^{\dot{a}a}\partial_\mu\varphi_a(x)&=m\chi^{\dot{a}}(x)\\ i(\sigma^\mu)_{a\dot{a}}\partial_\mu\chi^{\dot{a}}(x)&=m\varphi_a(x)\ .\end{aligned} \tag{B.44}$$

As we shall see below this is just an unusual form of the Dirac equation. It expresses the fact that both fields are associated with the same states, respectively, with the same one–particle operators $a(p,r)$ and $b^+(p,r)$. Consequently, the two fields φ and χ are not independent, in spite of the fact that they transform in a different way under $SL(2,C)$.

We can easily derive the Eqs. (B.44). We may write $\hat{\sigma}_\mu p^\mu=\omega_p\,\mathbf{1}-\vec{\sigma}\,\vec{p}=2|\vec{p}\,|(\frac{\omega_p}{2|\vec{p}\,|}\,\mathbf{1}-h)$ where $h\equiv\frac{\vec{\sigma}}{2}\,\frac{\vec{p}}{|\vec{p}\,|}$ is the helicity operator, and for massless states, where $\omega_p=|\vec{p}\,|$, we have $\hat{\sigma}_\mu p^\mu=2|\vec{p}|\,(\frac{1}{2}-h)$ as a projection operator on states with helicity $-\frac{1}{2}$, while $\sigma_\mu p^\mu=2|\vec{p}|\,(\frac{1}{2}+h)$ is a projection operator on states with helicity $+\frac{1}{2}$.

For massive states we observe that $p^\mu p^\nu\hat{\sigma}_\mu\sigma_\nu=p^\mu p^\nu\sigma_\mu\hat{\sigma}_\nu=p^2\cdot\mathbf{1}=m^2\cdot\mathbf{1}$ (see Eq. (B.7)) and therefore

$$\begin{aligned}\varphi(x):\quad \hat{\sigma}_\mu p^\mu u(p,r)&=\tfrac{m}{\sqrt{2m(p^0+m)}}(\hat{\sigma}_\mu p^\mu+m)\,U(r)=m\hat{u}(p,r)\\ -\hat{\sigma}_\mu p^\mu v(p,r)&=\tfrac{-m}{\sqrt{2m(p^0+m)}}(\hat{\sigma}_\mu p^\mu+m)\,V(r)=m\hat{v}(p,r)\\ \chi(x):\quad \sigma_\mu p^\mu\hat{u}(p,r)&=\tfrac{m}{\sqrt{2m(p^0+m)}}(\sigma_\mu p^\mu+m)\,U(r)=mu(p,r)\\ -\sigma_\mu p^\mu\hat{v}(p,r)&=\tfrac{-m}{\sqrt{2m(p^0+m)}}(\sigma_\mu p^\mu+m)\,V(r)=mv(p,r)\qquad .\end{aligned}$$

Finally, using

$$(i\partial_\mu=\pm p_\mu)\ e^{\mp ipx}$$

implies Eqs. (B.44) for the local fields.

In the massless limit $m \to 0 \quad : p^0 = \omega_p = |\vec{p}\,|$ we obtain two decoupled equations

$$\begin{aligned} i(\hat{\sigma}^\mu)^{\dot{a}a}\partial_\mu \varphi_a(x) &= 0 \\ i(\sigma^\mu)_{a\dot{a}}\partial_\mu \chi^{\dot{a}}(x) &= 0 \; . \end{aligned} \tag{B.45}$$

These are the **Weyl equations** and the massless spinor fields are the **Weyl fields** or **Weyl spinors**. They describe the neutrino, for example, and we have

$$\begin{aligned} \hat{\sigma}_\mu p^\mu u(p,r) = 0 \;\; &\Rightarrow \;\; (p^0 - \vec{\sigma}\,\vec{p}\,)u(p,r) = 0 \\ \sigma_\mu p^\mu \hat{u}(p,r) = 0 \;\; &\Rightarrow \;\; (p^0 + \vec{\sigma}\,\vec{p}\,)\hat{u}(p,r) = 0 \end{aligned}$$

with $p^0 = |\vec{p}\,|$ and $\vec{\sigma}\,\hat{\vec{p}} = 2h$ and h the helicity operator, we may write $hu = u$ and $h\hat{u} = -\hat{u}$ from which we learn that φ_a is a right–handed massless spin 1/2 field, while $\chi^{\dot{a}}$ is its left–handed parity partner.

B.8 The Dirac field: parity doubled spin $\frac{1}{2}$ field

The Dirac field is the bispinor field obtained by combining the irreducible fields $\varphi_a(x)$ and $\chi^{\dot{a}}(x)$ into one reducible field $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$. It is the natural field to be used to describe fermions participating parity conserving interactions like QED and QCD. Explicitly, the Dirac field is given by

$$\psi_\alpha(x) = \left(\begin{array}{c} \varphi_a \\ \chi^{\dot{a}} \end{array} \right)(x) = \sum_r \int d\mu(p) \left\{ u_\alpha(p,r)\, a(\vec{p},r)\, e^{-ipx} + v_\alpha(p,r)\, b^+(\vec{p},r)\, e^{ipx} \right\} \; . \tag{B.46}$$

where

$$u_\alpha = \left(\begin{array}{c} u_a \\ \hat{u}^{\dot{a}} \end{array} \right) \;\; ; \;\; v_\alpha = \left(\begin{array}{c} v_a \\ \hat{v}^{\dot{a}} \end{array} \right) \; .$$

Utilizing the explicit representation Eqs. (B.39), (B.40) and (B.43) of the two–spinors we obtain

$$u_\alpha(p,r) = \frac{1}{\sqrt{2m\,(p^0+m)}} \left(\begin{array}{c} (p^\mu \sigma_\mu + m)\, U(r) \\ (p^\mu \hat{\sigma}_\mu + m)\, U(r) \end{array} \right) \tag{B.47}$$

$$v_\alpha(p,r) = \frac{1}{\sqrt{2m\,(p^0+m)}} \left(\begin{array}{c} (p^\mu \sigma_\mu + m)\, V(r) \\ -(p^\mu \hat{\sigma}_\mu + m)\, V(r) \end{array} \right) \tag{B.48}$$

for the four-spinor one–particle wave functions.

The Dirac field has the following properties:

① it transforms according to the $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ representation of $SL(2,C)$

$$U(\Lambda,a)\psi_\alpha(x)U^{-1}(\Lambda,a) = S(\Lambda^{-1})_{\alpha\beta}\psi_\beta(\Lambda x + a) \; , \tag{B.49}$$

where $S(\Lambda) \equiv S_\Lambda$ is given by Eq. (B.19).

② $\psi_\alpha(x)$ satisfies the Klein-Gordon equation:

$$(\Box + m^2)\psi_\alpha(x) = 0$$

③ $\psi_\alpha(x)$ satisfies the **Dirac equation**:

$$(i\gamma^\mu \partial_\mu - m)_{\alpha\beta}\, \psi_\beta(x) = 0$$

where

$$\gamma^\mu \doteq \begin{pmatrix} 0 & \sigma^\mu \\ \hat{\sigma}^\mu & 0 \end{pmatrix}$$

are the Dirac matrices in the helicity representation. The Dirac equation is nothing but Eqs. (B.44) written in terms of the bispinor ψ. The crucial algebraic (representation independent) properties of the Dirac matrices are the anti–commutator relations

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \cdot \mathrm{Id}$$

and

$$\{\gamma^\mu, \gamma_5\} = 0 \ \ .$$

They follow directly from (B.7) and the definition of γ_5 in Sec. B.3. By "Id" we denoted the 4×4 unit matrix. Usually the unit matrix is not explicitly written if it is just a factor of some other expression, like $g^{\mu\nu}$ on the r.h.s. of the Dirac algebra or as a factor of the mass m in the Dirac equation.

④ **Definition:** The **adjoint field** $\bar{\psi}_\alpha(x)$ is defined such that

$$\bar{\psi}_\alpha(x)\psi_\alpha(x) \quad \text{is a } \textbf{scalar}$$

under Lorentz transformations. By Theorem 6 in Sec. B.3, the adjoint spinor is given by

$$\bar{\psi}_\alpha(x) \equiv \psi^+_\beta(x)(\gamma^0)_{\beta\alpha} \ \ .$$

⑤ the adjoint field satisfies:

$$\begin{array}{ll} 1) & U(\Lambda,a)\bar{\psi}_\alpha(x)U^{-1}(\Lambda,a) = \bar{\psi}_\beta(\Lambda x + a)\, S(\Lambda)_{\beta\alpha} \\ 2) & \bar{\psi}_\alpha(x)(\overleftarrow{\Box} + m^2) = 0 \\ 3) & \bar{\psi}_\alpha(x)\left(i\gamma^\mu \overleftarrow{\partial}_\mu + m\right)_{\alpha\beta} = 0 \end{array}$$

Dirac equation:

A spin 1/2 particle has two degrees of freedom and hence should be described by a two component field. For the given states described by the one–particle operators $a(\vec{p},r)$ and $b^+(\vec{p},r)$ we may associate two different local fields a undotted or a dotted spinor. The reducible field $\psi_\alpha(x)$ has twice the number of components of the fields $\varphi_a(x)$ and $\chi^{\dot{a}}(x)$, which means that two components are linearly dependent now. This linear dependence among the components requires the validity of a linear relationship, which is known as the Dirac equation. Using $(i\partial_\mu = \pm p_\mu)\, e^{\mp ipx}$ the Dirac equation yields, in the integrand of the Fourier decomposition of the Dirac field,

$$\begin{aligned} (\gamma^\mu p_\mu - m)_{\alpha\beta}\, u_\beta(p,r) &= 0 \\ (\gamma^\mu p_\mu + m)_{\alpha\beta}\, v_\beta(p,r) &= 0 \ \ . \end{aligned}$$

The four–spinors thus are classical solutions of the Dirac equation.

Remarks:

1) The normalizations which we obtained in a natural way for the massive particles are not convenient because they do not allow us to take smooth limits $m \to 0$. This is obvious if we look at the commutator (anti–commutator) formulae Eqs. (B.35) and (B.36), or, at the representations of the spinors Eqs. (B.39),(B.40) and (B.43). Therefore, in the following, we are going to chose a different normalization by multiplying the Weyl or Dirac spinors considered so far by $\sqrt{m}$. This amounts to choose the normalization

$$\xi' = \xi = \eta = m^j \quad \text{and} \quad \eta' = (-1)^{2j}\, m^j \ . \tag{B.50}$$

in place of (B.38). For the spin 1/2 case, formally, we have to set $u \to \sqrt{m}u$, $v \to \sqrt{m}v$ etc., for all one–particle wave functions. This is the normalization adapted henceforth (see Appendix A.3).

2) The representation obtained for the Dirac field is the so called **helicity representation** in four–spinor space (spinors, Dirac algebra). This should not be confused with the choice of the basis for the physical states in Hilbert space, where we distinguish between helicity states and canonical states. In the helicity representation the spinors and spinor–fields transform according to the reducible representation in block diagonal form (B.19). There is no a priori reason not to use some other representation. In fact for massive fermions one commonly uses the **standard representation**, which is chosen such that a particle at rest in momentum space is described by spinors u and v which have, respectively, upper and lower components only. This is achieved by the choice

$$\psi_\alpha^{\text{standard}} = \begin{pmatrix} \xi \\ \eta \end{pmatrix}_\alpha$$

where

$$\xi = \frac{1}{\sqrt{2}}\,(\varphi + \chi) \quad , \quad \eta = \frac{1}{\sqrt{2}}\,(\varphi - \chi) \quad .$$

In the rest frame we then have (using the new normalization)

$$u(0,r)^{\text{standard}} = \sqrt{m}\begin{pmatrix} U(r) \\ 0 \end{pmatrix} \quad , \quad v(0,r)^{\text{standard}} = \sqrt{m}\begin{pmatrix} 0 \\ V(r) \end{pmatrix} \quad .$$

Thus the relationship between the helicity and the standard representation is determined by the transformation of basis

$$\begin{aligned} \psi_\alpha^{\text{standard}} &= S\, \psi_\alpha^{\text{helicity}} \\ \gamma^\mu_{\text{standard}} &= S\, \gamma^\mu_{\text{helicity}}\, S^{-1} \end{aligned}$$

where

$$S = S^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

In Appendix A.3 of Sec. 2 we have summarized properties of Dirac fields and the Dirac algebra.

B.9 Majorana field:

This is a spin 1/2 field for which the antiparticle is identical to the particle ($b = a$). In order that the two Weyl equations remain consistent we must have a further relationship, a **reality condition**, between $\chi^{\dot{a}}$ and φ_a. A solution is given by

$$\chi^{\dot{a}} = (i\sigma_2)^{\dot{a}a}\varphi_a{}^* \quad . \tag{B.51}$$

This may be checked by using the complex conjugation matrix $C = i\sigma_2$ discussed earlier (see Eq. (B.3) and (B.5)). Since we are more familiar with the four–spinor notation, we write the **Majorana field** like a Dirac field

$$\psi_\alpha(x) = \begin{pmatrix} \varphi_a \\ \chi^{\dot{a}} \end{pmatrix}(x) = \begin{pmatrix} \varphi_a \\ (i\sigma_2\varphi^*)^{\dot{a}} \end{pmatrix}(x) \quad . \tag{B.52}$$

Explicitly we have

$$\psi_\alpha(x) = \begin{pmatrix} \varphi \\ i\sigma_2\varphi^* \end{pmatrix}_\alpha (x) = \sum_r \int d\mu(p) \left\{ u_\alpha(p,r)\, a(\vec{p},r)\, e^{-ipx} + v_\alpha(p,r)\, a^+(\vec{p},r)\, e^{ipx} \right\} \quad . \tag{B.53}$$

Note: the Majorana field is formally identical to a Dirac field in the helicity representation and setting $b = a$. It satisfies the Dirac equation is self–conjugate, however. It necessarily describes a neutral spin 1/2 particle, usually called **Majorana–neutrino**.

B.10 One particle wave functions for spin 1: Polarization vectors

Although the irreducible fields $(j,0)$ and $(0,j)$ are the simplest fields describing a spin j particle, in physical applications, it is advantageous to use slightly different fields. The reason is that strong and electromagnetic interactions are parity conserving and it is much more convenient to work with fields which exhibits a symmetric pairing of parity partners. We already know that spin 1/2 particles are described by Dirac fields rather than by undotted or dotted two–spinors. Even the parity violating weak interactions are written in terms of left– and right–handed Dirac fields and not in terms of Weyl fields. Similarly, spin 1 particles are not described by the irreducible $2j+1$–component fields $(1,0)$ or $(0,1)$, usually. One rather uses a four–vector field $V^\mu(x)$, which corresponds to a $(\frac{1}{2},\frac{1}{2})$ representation , and which incorporates both parities in one field.

The representation $(\frac{1}{2},\frac{1}{2})$ is not irreducible with respect to the subgroup of rotations. The angular momentum decomposition (see (2.30))

$$D(\frac{1}{2},\frac{1}{2}) \supset D_R^{(\frac{1}{2},\frac{1}{2})} = D_R^{(\frac{1}{2}+\frac{1}{2})} \oplus D_R^{(\frac{1}{2}-\frac{1}{2})} = D_R^{(1)} \oplus D_R^{(0)}$$

exhibits a spin 1 and a spin 0 part. The spin 0 part is described by a scalar field $\varphi(x) \doteq \partial_\mu V^\mu(x)$. We require $V^\mu(x)$ not to create or destroy scalar particles, which is the case only if the field satisfies $\varphi(x) \equiv 0$. This may be achieved for a massive spin 1 field by requiring it to be a solution of the Proca equation $(\Box + m^2)V_\mu(x) - \partial_\mu(\partial_\nu V^\nu) = 0$. A real (neutral) massive spin 1 field is then given by

$$V^\mu(x) = \sum_{r=\pm 1,0} \int d\mu(p) \left\{ \epsilon^\mu(p,r)\, a(\vec{p},r)\, e^{-ipx} + \epsilon^{*\mu}(p,r)\, a^+(\vec{p},r)\, e^{ipx} \right\} \quad ,$$

where $r = \pm 1$ label the transversal degrees of freedom and $r = 0$ the longitudinal degree of freedom.

Construction of the polarization vectors:

Quantization may be performed in the rest frame and, conventionally, one of the following coordinate systems is chosen:

a) Cartesian basis (real), J_3 diagonal:

Generators:

$$J_1^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad , \quad J_2^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & i \\ i & -i & 0 \end{pmatrix} \quad , \quad J_3^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Eigenvectors:

$$\epsilon(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \epsilon(2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad \epsilon(3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

with eigenvalues $\lambda = +1, -1, 0$.

b) Helicity basis (complex), $\vec{J}\,' = U \vec{J} U^+$, $\epsilon'(r) = U \epsilon(r)$:

The transition from Cartesian to spherical coordinates reads $(x, y, z) \Rightarrow (\frac{-x-iy}{\sqrt{2}}, \frac{x-iy}{\sqrt{2}}, z)$. The corresponding transformation matrix U is then given by

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 & 0 \\ -i & -i & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} .$$

In the new basis the generators take the form:

$$J_1^{\prime(1)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad , \quad J_2^{\prime(1)} = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad , \quad J_3^{\prime(1)} = \begin{pmatrix} 0 & i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Eigenvectors:

$$\epsilon'(1) = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ -i \\ 0 \end{pmatrix} \quad , \quad \epsilon'(2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} \quad , \quad \epsilon'(3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

with eigenvalues $\lambda = +1, -1, 0$. One usually denotes $\epsilon'(1) = \epsilon_+$, $\epsilon'(2) = \epsilon_-$ and $\epsilon'(3) = \epsilon_0$ such that $J_3 \epsilon_\lambda = \lambda \epsilon_\lambda$.

We are looking for three covariant polarization vectors $\epsilon^\mu(p, r)$ which depend on one vector p, only. The polarization vectors have to satisfy the conditions:

1) Absence of scalar mode: $p_\mu \epsilon^\mu(p, r) = 0$

2) Normalization: $\epsilon_\mu(p, r)\, \epsilon^{*\mu}(p, r') = -\delta_{rr'}$

3) Completeness: $\sum_r \epsilon_\mu(p, r)\, \epsilon^*_\nu(p, r) = -g_{\mu\nu} + \frac{p_\mu p_\nu}{M^2}$

Note that the r.h.s. of the completeness relation had to be chosen compatible with the transversality condition 1). Indeed, we find

$$p^\mu \left(-g_{\mu\nu} + \frac{p_\mu p_\nu}{M^2}\right) = -p_\nu + p_\nu \frac{p^2}{M^2} = 0 \text{ as } p^2 = M^2 \ .$$

We are going now to construct the polarization vectors in the artesian basis:

• In the rest frame:

$$p_\mu \epsilon^\mu(p_0, r) = p^0 \epsilon^0 = M\epsilon^0 = 0$$

which implies

$$\epsilon^0(p_0, r) = 0 \quad \forall \ r \ .$$

The three–vectors $\vec{\epsilon}\,(p_0, r)$ are required to be real eigenvectors of the 3rd component of the angular momentum

$$J_3\, \vec{\epsilon}\,(p_0, r) = \lambda_r\, \vec{\epsilon}\,(p_0, r)$$

which yields

$$\epsilon^k(p_0, r) = \delta_{kr}$$

for the normalized eigenvectors with the eigenvalues $\lambda_r = 1, -1, 0$ for $r = 1, 2, 3$.

• In an arbitrary frame:

The crucial point is that for a given mass M and a given three–momentum $\vec{p}$ there exist exactly one boost matrix, such that

$$p^\mu = L^\mu{}_\nu(p)\, p_0^\nu \ , \quad p_0^\nu = (M, \vec{0}\,)$$

and the polarization vectors at momentum p are obtained be the Lorentz boost

$$\epsilon^\mu(p, r) \doteq L^\mu{}_\nu(p)\, \epsilon^\nu(p_0, r) \ .$$

Explicitly, we have

$$\begin{aligned} L^0{}_0(p) &= \cosh\chi \\ L^i{}_0(p) &= L^0{}_i(p) = \hat{p}_i \sinh\chi \\ L^i{}_j(p) &= \delta_{ij} + \hat{p}_i\hat{p}_j(\cosh\chi - 1) \end{aligned}$$

with $\cosh\chi = \omega_p/M$, $\sinh\chi = |\vec{p}\,|/M$ and $\hat{\vec{p}} = \vec{p}/|\vec{p}\,|$. As a result we obtain

$$\begin{pmatrix} \epsilon^0(p, r) \\ \epsilon^i(p, r) \end{pmatrix} = \begin{pmatrix} \frac{\omega_p}{M} & \frac{p_j}{M} \\ \frac{p_i}{M} & \delta_{ij} + \hat{p}_i\hat{p}_j(\frac{\omega_p}{M} - 1) \end{pmatrix} \begin{pmatrix} \epsilon^0(p_0, r) \\ \epsilon^j(p_0, r) \end{pmatrix} \ ,$$

or,

$$\epsilon^0(p, r) = \frac{p_r}{M} \ , \quad \epsilon^k(p, r) = \delta_{kr} + p_k p_r \frac{1}{M(\omega_p + M)}$$

where we used $(\omega_p - M)/(M\vec{p}^{\,2}) = 1/(M(\omega_p + M))$.

We have to show that the constructed vectors satisfy the conditions 1) to 3) stated above.

Proofs:

0) We first prove the covariance of the polarization vectors. Under a Lorentz transformation

$$p^\mu \to \Lambda^\mu{}_\nu \, p^\nu = \Lambda^\mu{}_\nu \, L^\nu{}_\rho(p) \, p_0^\rho \equiv L^\mu{}_\rho(\Lambda p) \, p_0^\rho$$

and therefore

$$\epsilon^\mu(p,r) \to \epsilon^\mu(\Lambda p, r) = L^\mu{}_\nu(\Lambda p) \, \epsilon^\nu(p_0, r) = \Lambda^\mu{}_\nu \, L^\nu{}_\rho(p) \, \epsilon^\rho(p_0, r) = \Lambda^\mu{}_\nu \, \epsilon^\nu(p,r) \ \ ,$$

which demonstrates the proper transformation law of a four–vector.

1) $p_\mu \epsilon^\mu(p,r) = 0$:

Since

$$p_{0\nu} \doteq p_\mu L^\mu{}_\nu = \omega_p L^0{}_\nu - p^i L^i{}_\nu = (M, \vec{0}\,)$$

we have

$$p_\mu \epsilon^\mu(p,r) = p_{0\nu} \epsilon^\nu(p_0, r) = M \epsilon^0(p_0, r) = 0 \ \ ,$$

which proves the assertion.

2) Consider the tensor $X_{\mu\nu} = \sum_r \epsilon_\mu(p,r)\epsilon^*_\nu(p,r)$ which must satisfy

$$p^\mu X_{\mu\nu} = p^\nu X_{\mu\nu} = 0$$

and for the r.h.s of the completeness relation this holds, as we have checked before. Hence it remains to be shown that the completeness relation itself is true. Using Lorentz covariance it remains to be shown that the completeness relation is true in the rest frame: with $p_0 = (M, \vec{0}\,)$ we have

$$\begin{aligned} X_{0\nu}(p_0) &= 0 \ \text{ obviously since } \ \epsilon^0(p_0, r) = 0 \\ X_{00}(p_0) &= 0 = -g_{00} + \frac{M^2}{M^2} = 0 \\ X_{ik}(p_0) &= \sum_r \epsilon_i(p_0, r) \, \epsilon^*_k(p_0, r) = -g_{ik} = \delta_{ik} \end{aligned}$$

and by the covariance the completeness holds in any frame.

3) Using the property

$$L^\mu{}_\rho = (L^{-1})_\rho{}^\mu \Rightarrow (L^{-1})_\rho{}^\mu L_\mu{}^\nu = \delta_\rho{}^\nu$$

of a Lorentz transformation, we obtain

$$\begin{aligned} \epsilon_\mu(p,r) \, \epsilon^{*\mu}(p,r') &= L_\mu{}^\nu L^\mu{}_\rho \, \epsilon_\nu(p_0, r) \, \epsilon^{*\rho}(p_0, r') \\ &= \epsilon_\nu(p_0, r) \, \epsilon^{*\nu}(p_0, r') = \epsilon^0 \epsilon^{*0} - \vec{\epsilon}\,\vec{\epsilon}^{\,*} = -\delta_{rr'} \ \ . \end{aligned}$$

This completes the necessary proofs of the properties of the polarization vectors.

q.e.d.

The transition to helicity basis and the application to charged fields is straight forward and given in Appendix A.5 and A.6 of Sec. 2

C Massless particles

Peculiarities of massless particles and fields occur for particles with spin ≥ 1. The reason is simple: for spin < 1 massless and massive states have the same number of degrees of freedom for a given spin.

Spin 0: Taking the limit $\lim_{m\to 0}\varphi(x)$ of a massive scalar field poses no problems. The massless scalar field

$$\varphi(x) = \int d\mu(p) \left\{ a(\vec{p}\,)\, e^{-ipx} + a^{+}(\vec{p}\,)\, e^{ipx} \right\} \quad ,$$

is a solution of the **d'Alembert equation** $\Box\varphi(x) = 0$. In momentum space p now is a light–like vector, i.e., on the mass–shell $p^2 = 0$ and hence $p^0 = \omega_p = |\vec{p}\,|$. The invariant integration volume is

$$d\mu(p) = (2\pi)^{-3} \frac{d^3p}{2|\vec{p}\,|} \quad , \quad d^3p = |\vec{p}\,|^2\, d|\vec{p}\,|\, d\Omega_3$$

where $d\Omega_3 = \sin\theta d\theta d\phi$ with $0 \leq \theta \leq \pi$, $0 \leq \phi < 2\pi$ the surface element of the compact 3–sphere.

Spin 1/2: Also the massive spinor fields $\varphi_a, \chi^{\dot{a}}$ or ψ_α have proper massless limits, if they are properly normalized.

Spin 1: (and higher) the zero mass limits lead to problems. While the massive spin j particle has $2j+1$ degrees of freedom, the massless one has two, only. Starting point are the helicity states. Since there is no rest frame for a massless particle, quantization is performed relative to a light-like standard vector in z–direction $q^\mu = (q, 0, 0, q)$, usually.

C.1 Massless states

We first consider the subgroup of $L_+^\uparrow$ which leaves the standard vector q invariant. This subgroup is called **little group** of q:

$$\Lambda q = q \quad \leftrightarrow \Lambda \in LG_q \quad .$$

We consider the Pauli–Lubansky generators L_μ, defined in Eq. (2.4), which satisfies $L_\mu P^\mu = 0$ such that

$$L_\mu P^\mu |q, \alpha >= q^\mu L_\mu |q, \alpha >= q\,(L^0 - L^3)\, |q, \alpha >$$

and, furthermore,

$$L^3 |q, \alpha >= \frac{1}{2} q\, \varepsilon^{30\rho\sigma} M_{\rho\sigma} |q, \alpha >= q J_3 |q, \alpha >$$

such that $L^0 = L^3$ and $L^3 = qJ_3$ on standard states. We are left with the following commutation relations for L^1, L^2, L^3

$$[L_1, L_2] = 0 \quad , \quad [L_3, L_1] = iqL_2 \quad , \quad [L_3, L_1] = -iqL_1$$

which compares to

$$[P_1, P_2] = 0 \quad , \quad [J_3, P_1] = iP_2 \quad , \quad [J_3, P_1] = -iP_1$$

which is the Lie algebra of the group $E(2)$, the group of rotations and translations in the xy–plane. A crucial property of the group $E(2)$ is its **non–semi–simple** nature, i.e., it has an invariant

Abelian subgroup. The consequence of the cummutativity of L_1 and L_2 is that $L_1^2 + L_2^2$ and hence $L_\mu L^\mu$ may take any real value, which means that we may have "continuous spin". We are interested only in representations with discrete spin, i.e., for finite dimensional representations . Defining the ladder operators $L_\pm = L_1 \pm iL_2$ we obtain the algebra

$$[L_+, L_-] = 0 \ , \quad [L_3, L_\pm] = qL_\pm \ .$$

The procedure then is the same as in the case of angular momentum. A finite dimensional representation must have a lowest state $|\lambda\rangle \Rightarrow L_-|\lambda\rangle = 0$. There also must exist a highest state $(L_+)^p|\lambda\rangle \Rightarrow (L_+)^{p+1}|\lambda\rangle = 0$ for some integer $p \geq 0$. However, L_- and L_+ commute such that $L_-(L_+)^r|\lambda\rangle = (L_+)^r L_-|\lambda\rangle = 0$ for any integer $r \geq 0$ and hence, $L_- = 0$ and $L_+ = (L_-)^+ = 0$ on the entire representation space. $|\lambda\rangle$ is eigenstate of $L_3 = qJ_3$: $J_3|\lambda\rangle = \lambda|\lambda\rangle$. As a result we have the

Theorem 7: All finite dimensional representations of $L_+^\uparrow$ for $m = 0$ must be 1-dimensional.

States with arbitrary momentum

Let p be an arbitrary vector with $p^2 = 0$ and $p^0 = |\vec{p}\,| > 0$. The vector p can be obtained from the standard vector q by a boost $L_z(p, q)$ along the z–axis, which transforms $(q, 0, 0, q)$ into $\tilde{p} = (p^0, 0, 0, p^0)$, and a rotation $R_{\varphi,\theta}(p)$ which rotates $\tilde{p}$ into p about an axis in the xy–plane. Thus

$$p = R_{\varphi,\theta}(p)\, L_z(p, q)\, q$$

and we obtain arbitrary massless states by the

Definition:

$$\begin{aligned} |p, \lambda\rangle &\doteq U(R_{\varphi,\theta}(p)\, L_z(p, q))\, |q, \lambda\rangle \\ &= U(R_{\varphi,\theta}(p))\, |L_z q, \lambda\rangle \end{aligned} \tag{C.54}$$

which is a rotated standard state of the requested momentum.

Theorem 8: $|p, \lambda\rangle$ is a helicity state:

$$h|p, \lambda\rangle = \lambda|p, \lambda\rangle$$

Proof: as in massive case.

Theorem 9: The massless states $|p, \lambda\rangle$ transform diagonal in λ, according to

$$U(\Lambda, a)\, |p, \lambda\rangle = e^{ia\Lambda p}\, e^{-i\phi\lambda}\, |\Lambda p, \lambda\rangle \tag{C.55}$$

Proof:

1. $U(\Lambda, a)\, |p, \lambda\rangle = U(1, a)\, U(\Lambda R_{\varphi,\theta}(p)\, L_z(p, q))\, |q, \lambda\rangle$
2. $U(\Lambda R_{\varphi,\theta}(p)\, L_z(p, q))\, |q, \lambda\rangle = ?$

   We first note that the momentum Λp is obtained from q by

$$\begin{aligned} \Lambda p &= \Lambda R_{\varphi,\theta}(p) L_z(p, q) q \\ &= R_{\varphi,\theta}(\Lambda p) L_z(\Lambda p, q) q \ . \end{aligned}$$

Therefore

$$\Lambda R_{\varphi,\theta}(p) L_z(p,q) = R_{\varphi,\theta}(\Lambda p) L_z(\Lambda p, q)\, S_q(\Lambda, p)$$

with $S_q \in LG_q$, i.e., $S_q(\Lambda, p)q = q$. S_q may be represented as a product

$$S = S_1 S_2 R_z \tag{C.56}$$

where S_1, S_2 and R_z are generated by L_1, L_2 and $L_3/q = J_3$, respectively.

3. On standard states we then have

$$\begin{aligned} U(R_z)\,|q,\lambda\rangle &= e^{-i\phi J_3}|q,\lambda\rangle = e^{-i\phi\lambda}|q,\lambda\rangle \\ U(S_1)\,|q,\lambda\rangle = U(S_2)\,|q,\lambda\rangle &= |q,\lambda\rangle \end{aligned}$$

since $L_1 = L_2 = 0$ on all standard states. Hence

$$U(S_q(\Lambda,p))\,|q,\lambda\rangle = e^{-i\phi\lambda}|q,\lambda\rangle \ .$$

4.

$$\begin{aligned} U(\Lambda, a)\,|p,\lambda\rangle &= U(1,a)\,U(R_{\varphi,\theta}(\Lambda p))\,U(L_z(\Lambda p,q))\,U(S(\Lambda,p))\,|q,\lambda\rangle \\ &= U(1,a)\,U(R_{\varphi,\theta}(\Lambda p))\,U(L_z(\Lambda p,q))\,e^{-i\phi\lambda}\,|q,\lambda\rangle \\ &= U(1,a)\,e^{-i\phi\lambda}\,|\Lambda p,\lambda\rangle \\ &= e^{ia\Lambda p}\,e^{-i\phi\lambda}\,|\Lambda p,\lambda\rangle \end{aligned}$$

q.e.d.

Admitted λ–values for the discrete representations

Consider a rotation about the z–axis: $R_z q = q$

$$U(R_z)|q,\lambda\rangle = e^{-i\phi\lambda}|q,\lambda\rangle \ .$$

A rotation by 2π yields

$$e^{-i2\pi\lambda} = \pm 1$$

for a true representation of $SL(2,C)$. Thus the allowed values are

$$\lambda = 0, \pm\frac{1}{2}, \pm 1, \ldots$$

and hence for each value of λ there exists *exactly one* irreducible representation . $|\lambda|$ is the spin of the massless particle.

As a result we have the following

Theorem 10: The finite dimensional irreducible and unitary representations of $SL(2,C)$ to mass 0 and spin j are 1–dimensional and characterized by the helicity $\lambda = \pm j$. To a given spin $j > 0$ there exist exactly two helicity states. Each of the two possible states is invariant by itself under $\mathcal{P}_+^\uparrow$. The states get interchanged under parity transformations:

$$U(P)\,h\,U(P)^{-1} = -h \ , \quad P \in \mathcal{P}_-^\uparrow \ .$$

C.2 Massless fields

The construction of local massless fields is based on the transformation laws of the massless states Eq. (C.55). Under homogeneous Lorentz transformations the states transform according to an Abelian group, and the phase function is satisfying

$$\begin{aligned} \phi(S_1 S_2) &= \phi(S_1) + \phi(S_2) \\ \phi(S^{-1}) &= -\phi(S) \end{aligned} \tag{C.57}$$

where

$$S = S(\Lambda, p) = \tilde{L}^{-1}(\Lambda p)\, \Lambda \tilde{L}(p) \tag{C.58}$$

and

$$\tilde{L}(p) = R_{\varphi,\theta}(p))\, L_z(p, q) \quad . \tag{C.59}$$

The states transform under $L_+^{\uparrow}$ with a p–dependent phase ! Under complex conjugation we have a change of sign and, consequently, we note

$$\boxed{a^+(\vec{p}, \lambda) \text{ and } a(\vec{p}, -\lambda) \text{ transform identically}}$$

namely, by

$$U(\Lambda)\, a(\vec{p}, \lambda)\, U^{-1}(\Lambda) = e^{i\lambda\phi(S(\Lambda,p))} a(\Lambda\vec{p}, \lambda)$$

and hence, there are two linear combinations with a simple transformation law:

$$\begin{aligned} & U(\Lambda) \left\{ \xi a(\vec{p}, \lambda)\, e^{-ipx} + \eta b^+(\vec{p}, -\lambda)\, e^{ipx} \right\} U^{-1}(\Lambda) \\ &= e^{\lambda i\phi(S(\Lambda,p))} \left\{ \xi a(\Lambda\vec{p}, \lambda)\, e^{-ipx} + \eta b^+(\Lambda\vec{p}, -\lambda)\, e^{ipx} \right\} \end{aligned}$$

which are suitable for the construction of local fields. The real local field with normal inner parity $\eta_P = (-1)^j = -1$ obtained is

$$A^\mu(x) = \sum_{\pm} \int d\mu(p) \varepsilon^\mu_\pm(p) \left\{ a(\vec{p}, \pm)\, e^{-ipx} - a^+(\vec{p}, \mp) e^{ipx} \right\} \tag{C.60}$$

where the polarization vectors $\varepsilon^\mu_\pm(p)$ will be constructed below. Here, we only mention the transformation property we read off if we consider $U(\Lambda)\, A^\mu(x)\, U^{-1}(\Lambda)$ and change integration variable $\Lambda p \to p$ we obtain

$$\varepsilon^\mu_\pm(p) \to e^{\pm i\phi(S(\Lambda,p))} \varepsilon^\mu_\pm(\Lambda^{-1} p) \tag{C.61}$$

which we will have to compare with the transformation law of a vector

$$\varepsilon^\mu_\pm(p) \to (\Lambda^{-1})^\mu{}_\nu\, \varepsilon^\nu_\pm(p) \tag{C.62}$$

expected if $A^\mu(x)$ would transform as a four–vector field $U(\Lambda)\, A^\mu(x)\, U^{-1}(\Lambda) = \left(\Lambda^{-1}\right)^\mu{}_\nu\, A^\nu(\Lambda x)$.

Polarization vectors for massless spin 1 particles:

The procedure to obtain the polarization vectors for massless spin 1 particles is similar, in spirit, to the one applied in case of the spinors in Sec. B.7 and of the polarization vectors for massive spin 1 particles in Sec. B.10: One starts from a standard frame and then performs a boost to a state with arbitrary momentum. The construction is as follows:

1) Standard vector in z–direction. Helicity basis:

$$\varepsilon^1_\pm = \mp\frac{1}{\sqrt{2}} \ , \quad \varepsilon^2_\pm = -\frac{i}{\sqrt{2}} \ , \quad \varepsilon^3_\pm = \varepsilon^0_\pm = 0$$

describe the two helicity states in the standard frame:

$$\varepsilon^\mu_+ = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ -i \\ 0 \end{pmatrix} \ , \quad \varepsilon^\mu_- = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \ ,$$

and with the helicity operator

$$\hat{\vec{p}}\,\vec{J}^{(1)} = J^{(1)}_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ ,$$

we find indeed

$$J_3\varepsilon_\pm = \pm\varepsilon_\pm$$

2) Arbitrary direction of the momentum:

Let $R(p)$ denote the rotation of the z-axis into the direction of $\vec{p}$, i.e., $p^\mu = R^\mu{}_\nu(p)\,p^\nu_0$ where $p^\nu_0 = p_0\,(1,0,0,1)$. Then

$$\varepsilon^\mu_\pm(p) = R^\mu{}_\nu(p)\,\varepsilon^\nu_\pm$$

and one easily verifies

$$\begin{aligned} \varepsilon^\mu_\pm(p)^* &= -\varepsilon^\mu_\mp(p) \\ \varepsilon_{\pm\mu}(p)\,\varepsilon^\mu_\pm(p) &= \varepsilon_{\pm\mu}\varepsilon^\mu_\pm = 0 \\ \varepsilon_{\pm\mu}(p)\,\varepsilon^\mu_\mp(p) &= \varepsilon_{\pm\mu}\varepsilon^\mu_\mp = 1 \ . \end{aligned}$$

The orthogonality of the general polarization vectors follows from the one's in the standard frame by rotational invariance:

$$\begin{aligned} \varepsilon_{\lambda\mu}(p)\,\varepsilon^\mu_{\lambda'}(p) &= g_{\mu\nu}\varepsilon^\nu_\lambda(p)\,\varepsilon^\mu_{\lambda'}(p) = \\ g_{\mu\nu}R^\mu{}_\rho R^\nu{}_\sigma \varepsilon^\sigma_\lambda \varepsilon^\rho_{\lambda'} &= g_{\rho\sigma}\varepsilon^\sigma_\lambda\varepsilon^\rho_{\lambda'} \ . \end{aligned}$$

3) Time component in standard representation:

$$\varepsilon^0_\pm(p) = 0 \ : \ R^0{}_\nu(p)\,\varepsilon^\nu_\pm = \delta^0{}_\nu\,\varepsilon^\nu_\pm = \varepsilon^0_\pm = 0$$

4) Transversality:

$$p_\mu \varepsilon^\mu_\pm(p) \;=\; p_\mu R^\mu{}_\nu\,(p)\,\varepsilon^\nu_\pm = p_{0\mu}\varepsilon^\mu_\pm = 0 \;\;,$$

where we used the property that for a general Lorentz transformation $\Lambda^\mu{}_\nu = \left(\Lambda^{-1}\right)_\nu{}^\mu$ and hence $p_\mu R^\mu{}_\nu\,(p) = \left(R^{-1}\right)_\nu{}^\mu\; p_\mu = p_{0\nu}$, since, $p^\mu = R^\mu{}_\nu\,(p)\,p_0^\nu$ with $p_0^\mu = p^0\,(1,0,0,1)$.

5) Transformation properties of $\varepsilon^\mu_\pm(p)$:

So far we have considered the construction of the polarization vectors relative to a light–like standard four–vector $q^\mu = q^0\;(1,0,0,1)$ with $q^0 = p^0$ and obtained $\varepsilon^\mu_\pm(p) = R^\mu{}_\nu\,(p)\,\varepsilon^\nu_\pm$ by performing a rotation. If $p^0 \neq q^0$ we first have to boost the momentum appropriately in the z–direction. Thus in general we have

$$\varepsilon^\mu_\pm(p) = \tilde{L}^\mu{}_\nu\,(p)\,\varepsilon^\nu_\pm$$

with $\tilde{L}(p)$ defined in Eq. (C.59). We first consider Eq. (C.62)

$$\begin{aligned}(\Lambda^{-1})^\mu{}_\nu\,\varepsilon^\nu_\pm(p) \;&=\; \left(\Lambda^{-1}\tilde{L}(p)\right)^\mu{}_\nu\;\varepsilon^\nu_\pm \\ &=\; \left(\tilde{L}(\Lambda^{-1}p)\,S(\Lambda^{-1},p)\right)^\mu{}_\nu\;\varepsilon^\nu_\pm \end{aligned} \tag{C.63}$$

where

$$S(\Lambda^{-1},p) = \tilde{L}^{-1}(\Lambda^{-1}p)\Lambda^{-1}\tilde{L}(p) \in LG_q \;\;! \tag{C.64}$$

Since $S(\Lambda^{-1},p)$ leaves invariant the standard vector q, which is our reference vector for the quantization, we must be able to choose $S(\Lambda^{-1},p) \in LG_q$ arbitrary. With other words, the physics cannot depend on the choice of a particular element from LG_q. We observe that the polarization vectors of massless spin 1 particles cannot be unique, they must be considered to be **equivalence classes**.

Generally, the transformation law for the polarization vectors may be written as

$$(S(\Lambda^{-1},p))^\mu{}_\nu\;\varepsilon^\nu_\pm = e^{\pm i\phi(S(\Lambda,p))}\varepsilon^\mu_\pm + X_\pm(\Lambda,p)\,q^\mu \;\;. \tag{C.65}$$

Because the phase factors $e^{\pm i\phi}$ show up in a canonical way via the transformation of the physical states, we are interested here in the extra terms $X_\pm q^\mu$, only, which occur, when $S(\Lambda,p)$ has the property that

$$\phi(S(\Lambda,p)) = 0 \;\;. \tag{C.66}$$

This condition defines a subgroup of LG_q of elements $\overset{\circ}{\Lambda}$ of the form

$$\overset{\circ}{\Lambda} = \begin{pmatrix} 1+|\alpha|^2 & u_1 & -u_2 & -|\alpha|^2 \\ u_1 & 1 & 0 & -u_1 \\ -u_2 & 0 & 1 & u_2 \\ |\alpha|^2 & u_1 & -u_2 & 1-|\alpha|^2 \end{pmatrix} \tag{C.67}$$

where u_1 and u_2 are two real parameters and $\alpha \doteq \frac{u_1+iu_2}{\sqrt{2}}$. One easily checks, that

1^0. $\quad \overset{\circ}{\Lambda}\; q^\nu = q^\mu$

2^0. $\quad \overset{\circ}{\Lambda}\ \varepsilon^\nu_\pm = \varepsilon^\mu_\pm + X_\pm(\Lambda, q)\, q^\mu$

with

$$X_+ = -\frac{\alpha^*}{q} \quad , \quad X_- = \frac{\alpha}{q} \quad .$$

Herewith, we have determined the equivalence classes of polarization vectors. They are given by

$$\boxed{\tilde{\varepsilon}^\mu_\pm = \tfrac{1}{\sqrt{2}}\,(\zeta_\pm, \mp 1, -i, \zeta_\pm)} \tag{C.68}$$

where $\zeta_\pm$ are arbitrary parameters.

The general form of $\varepsilon^\mu_\pm(p)$ then is given by

$$\begin{aligned} \tilde{\varepsilon}^\mu_\pm(p) &= \tilde{L}^\mu{}_\nu\,(p)\,\tilde{\varepsilon}^\nu_\pm \\ &= \varepsilon^\mu_\pm(p) + \lambda\, p^\mu \end{aligned} \tag{C.69}$$

since $\tilde{L}(p)q = p$ and we have set $X_\pm \to \lambda$ which is an arbitrary parameter.

For the Lorentz transformed vectors $(\Lambda^{-1})^\mu{}_\nu\ \varepsilon^\nu_\pm(p)$ we obtain

$$\begin{aligned} (\Lambda^{-1})^\mu{}_\nu\, \varepsilon^\nu_\pm(p) &= \Big(\tilde{L}(\Lambda^{-1}p)\, S(\Lambda^{-1}, p)\Big)^\mu{}_\nu\ \varepsilon^\nu_\pm \\ &= e^{\pm i\phi(S(\Lambda,p))} \varepsilon^\mu_\pm(\Lambda^{-1}p) + X_\pm(\Lambda, p)(\Lambda^{-1}p)^\mu \end{aligned}$$

or

$$e^{\pm i\phi(S(\Lambda,p))}\ \varepsilon^\mu_\pm(\Lambda^{-1}p) = (\Lambda^{-1})^\mu{}_\nu\ \left(\varepsilon^\nu_\pm(p) - X_\pm(\Lambda, p)\, p^\nu\right)$$

We know that $\varepsilon^0_\pm(\Lambda^{-1}p)$ vanishes identically by construction. We therefore are able to compute $X_\pm(\Lambda, p)$:

$$0 = (\Lambda^{-1})^0{}_\nu\ \left(\varepsilon^\nu_\pm(p) - X_\pm(\Lambda, p)\, p^\nu\right)$$

or

$$X_\pm(\Lambda, p) = (\Lambda^{-1})^0{}_\nu\, \varepsilon^\nu_\pm(p)/(\Lambda^{-1}p)^0$$

and we have the final result

$$e^{\pm i\phi(S(\Lambda,p))}\ \varepsilon^\mu_\pm(\Lambda^{-1}p) = \left[(\Lambda^{-1})^\mu{}_\nu - (\Lambda^{-1})^0{}_\nu\, \frac{(\Lambda^{-1}p)^\mu}{(\Lambda^{-1}p)^0}\right] \varepsilon^\nu_\pm(p) \quad . \tag{C.70}$$

This demonstrates the non-covariant transformation properties of the polarization vectors ! This of course does not mean that the "physics" is not Lorentz invariant. But it is an unavoidable technical complication, which is controlled be the requirement of **gauge invariance**.

6) Gauge invariance:

① Polarization vectors:

$\varepsilon^\mu_\pm(p)$ is a 4–component quantity which we try to use to describe **two** physical degrees of freedom. One of the 4 components can be eliminated in a manifestly covariant way by the Lorentz condition (transversality) $p_\mu \varepsilon^\mu_\pm(p) = 0$.

② Gauge invariance:

In any case we must require the gauge invariance of physical predictions, i.e., the substitution

$$\boxed{\varepsilon^{\mu}_{\pm}(p) \to \varepsilon^{\mu}_{\pm}(p) \, + \, \lambda \, p^{\mu} \quad , \quad \lambda \text{ arbitrary}} \tag{C.71}$$

does not affect physical quantities. Which is the same as the well known invariance under **local gauge transformations**

$$A_{\mu}(x) \to A_{\mu}(x) \, - \, \partial_{\mu}\alpha(x) \tag{C.72}$$

of the vector potential.

Note: An immediate consequence of the non-covariant transformation properties of the polarization vectors is the following: In the representation Eq. (C.60) $A^{\mu}(x)$ is <u>not a Lorentz vector</u>! Instead, we have L–invariance up to a total divergence

$$U(\Lambda) \, A^{\mu}(x) \, U^{-1}(\Lambda) = (\Lambda^{-1})^{\mu}{}_{\nu} \, A^{\nu}(\Lambda x) \, + \, \partial^{\mu}\Phi(x) \ \ , \tag{C.73}$$

of an arbitrary scalar function $\Phi(x)$. For the construction of a covariant photon field we refer to Sec. 4.1.

C.3 Admitted representations for massless particles

The following theorem holds:

Theorem: Massless particles of helicity λ only can transform according to the representations

$$(A, B) = (A, A + \lambda) \ \ ,$$

where $2A$ and $2(A + \lambda)$ are non–negative integer numbers.

Note: This means that for massless particles there is an essential restriction of the general angular momentum rule valid for massive particles presented in Sec. 2.6.

Thus the simplest representations for massless fields are:

$$\begin{array}{rcll} \lambda = -\frac{1}{2} & : & (\frac{1}{2}, 0), (1, \frac{1}{2}), \ldots & \text{left} - \text{handed} \\ +\frac{1}{2} & : & (0, \frac{1}{2}), (\frac{1}{2}, 1), \ldots & \text{right} - \text{handed} \\ -1 & : & (1, 0), (\frac{3}{2}, \frac{1}{2}), (2, 1), \ldots & \text{left} - \text{handed} \\ +1 & : & (0, 1), (\frac{1}{2}, \frac{3}{2}), (1, 2), \ldots & \text{right} - \text{handed} \end{array}$$

Not admitted is the representation $(\frac{1}{2}, \frac{1}{2})$, which is precisely the one needed for the description of massless spin 1 fields, which show up in physics as **gauge fields** associated with local gauge symmetries as we shall see in Secs. 4 and 6.

Other representations are possible only if $iSL(2, C)$ is no longer unitarily represented on the Hilbert space of physical states. This means that, creation and annihilation operators, and the states they describe, do not transform according to unitary representations. The representation space then necessarily has **indefinite metric** and hence differs from the physical Hilbert space. For a detailed discussion we refer to Sec. 4.1.

Proof of the theorem:

The restriction obtained for massless fields is a consequence of the properties of the little group $E(2)$ of a light–like four–vector. $E(2)$ is not semi–simple, i.e., it has an Abelian subgroup. This is very different from the massive case, where the standard vector $(m, \vec{0}\,)$ has $O(3)$ as a stability group, which is semi–simple.

We consider a representation $U(S)$ of the little group of the light–like vector q: $Sq = q$. The infinitesimal form reads

$$U(S) = 1 - i\hat{\chi}_1 L_1 - i\hat{\chi}_2 L_2 - i\phi J_3$$

where we used the form (C.56). From our discussion in Sec. C.1 we know that for standard states $L^0 = L^3 = qJ_3$ and furthermore

$$\begin{aligned} L^1 &= \frac{q}{2}[\varepsilon^{10\rho\sigma} M_{\rho\sigma} + \varepsilon^{13\rho\sigma} M_{\rho\sigma}] = q(J_1 + K_2) \\ L^2 &= \frac{q}{2}[\varepsilon^{20\rho\sigma} M_{\rho\sigma} + \varepsilon^{23\rho\sigma} M_{\rho\sigma}] = q(J_2 - K_1) \ . \end{aligned}$$

Setting $\chi_i = \hat{\chi}_i\, q$ we find for the representations (A, B) of $SL(2, C)$ the following:

$$\begin{aligned} U(S) &= 1 - i\chi_1(J_1 + K_2) - i\chi_2(J_2 - K_1) - i\phi J_3 \\ &= 1 - i\phi(A_3 + B_3) - i\chi_+ A_- i\chi_- B_+ \end{aligned}$$

with $A_- = A_1 - iA_2$ and $B_+ = B_1 + iB_2$ are the lowering and raising ladder operators for the "angular momenta" A and B, respectively, and $\chi_\pm = \chi_1 \pm i\chi_2$. Remember that $\vec{J} = \vec{A} + \vec{B}$, $\vec{K} = -i(\vec{A} - \vec{B}$ and A_i and B_i satisfy two independent angular momentum algebras.

For a representation this translates into

$$D^{(A,B)}(S) = 1 - i\varphi(\mathcal{A}_3 + \mathcal{B}_3) - i\chi_+ \mathcal{A}_- i\chi_- \mathcal{B}_+$$

Now, the crucial point is that physics cannot depend on the specific choice of $S \in LG_q$. More specifically, $D^{(A,B)}(S)$ acting on a one–particle wave function $u(\lambda)$ describing a massless particle of helicity λ must yield:

$$D^{(A,B)}(S)u(\lambda) = e^{-i\lambda\phi(S)}u(\lambda)$$

and for infinitesimal transformations we have

$$e^{-i\lambda\phi(S)} = e^{-i\phi(S)\, J_3} \simeq e^{-i\phi\, J_3}$$

and hence

$$\phi(S) \simeq \varphi$$

$$\begin{aligned} (\mathcal{A}_3 + \mathcal{B}_3)\, u(\lambda) &= \lambda u(\lambda) \\ (\mathcal{A}_-)\, u(\lambda) &= 0 \\ (\mathcal{B}_+)\, u(\lambda) &= 0 \end{aligned}$$

Since $[C_3, C_\pm] = \pm C_\pm$ where $C_\pm = C_1 \pm iC_2$ independently for $C_i = A_i, B_i$ and the same algebra holds for the representations $\mathcal{A}_i, \mathcal{B}_i$ we infer the following: $(\mathcal{A}_-)\, u(\lambda) = 0$ tells us that $u(\lambda)$ is an eigenvector of $\mathcal{A}_3$ to the lowest eigenvalue, which is $(\mathcal{A}_3)u(\lambda) = -Au(\lambda)$. Similarly, $(\mathcal{B}_+)u(\lambda) = 0$ implies that $u(\lambda)$ is an eigenvector of $\mathcal{B}_3$ to the highest eigenvalue, which is $(\mathcal{B}_3)\, u(\lambda) = Bu(\lambda)$. Thus $(\mathcal{A}_3 + \mathcal{B}_3)\, u(\lambda) = (B - A)\, u(\lambda) = \lambda u(\lambda)$ such that $B = A + \lambda$.

q.e.d.

D.4 Manipulations of perturbation series

1. Vacuum graphs and the normalization of perturbative amplitudes

In quantum field theory normalizations, such as $< 0|S|0 >= 1$ of the S–matrix or $Z\{0\} = 1$ of a generating functional $Z\{J\}$, are usually nontrivial, i.e., they are not automatic and must be performed explicitly. In perturbation theory the rule how to get the proper normalization, for the examples mentioned at least, is very simple: proper normalization is achieved by omitting all vacuum diagrams.

In order to show this we introduce the following notation:

$$<< T\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\,\} >>={}_{in}\!< 0|T\{\phi^{(in)}(x_1)\,\phi^{(in)}(x_2)\cdots\phi^{(in)}(x_n)\;e^{i\int d^dx\,\mathcal{L}^{(in)}_{\rm int}(x)}\}|0 >_{in}$$
$$< 0|T\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\,\}\,|0 >={}_{in}\!< 0|T\{\phi^{(in)}(x_1)\,\phi^{(in)}(x_2)\cdots\phi^{(in)}(x_n)\;e^{i\int d^dx\,\mathcal{L}^{(in)}_{\rm int}(x)}\}|0 >_{in\,\otimes}$$

and

$$Z ={}_{in}\!< 0|T\{e^{i\int d^dx\,\mathcal{L}^{(in)}_{\rm int}(x)}\}|0 >_{in}$$

Proposition: Omission of all vacuum diagrams (the $\otimes$ prescription) is identical to the division by Z, i.e.,

$$Z\cdot < 0|T\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\,\}\,|0 >\equiv<< T\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\,\} >>$$

Proof: in the perturbation expansion a Feynman graph contributing to $<< \cdots >>$ consists of two parts factorizing parts

$$\begin{array}{c}\int d^dy_1\cdots d^dy_m << T\{\phi(x_1)\;\cdots\phi(x_n)\mathcal{L}_{\rm int}(y_1)\cdots\mathcal{L}_{\rm int}(y_m)\} >> \\ =\sum_{k=0}^m \left(\begin{array}{c} m \\ k\end{array}\right)\int d^dy_1\cdots d^dy_m \\ < 0|T\{\phi(x_1)\;\cdots\phi(x_n)\mathcal{L}_{\rm int}(y_1)\cdots\mathcal{L}_{\rm int}(y_{m-k})\}\,|0 >_\otimes\cdot \\ < 0|T\{\mathcal{L}_{\rm int}(y_{m-k+1})\cdots\mathcal{L}_{\rm int}(y_m)\}\,|0 >\end{array}$$

The first factor belongs to a group of k internal vertices where each connected component connects to at least one external vertex x_i. The remaining $m-k$ internal vertices do not connect to any external vertex and hence belong to **vacuum graphs**. Because all internal vertices are integrated out, there are $\binom{m}{k}$ equal possibilities to choose k vertices out of m.

Since the double summation over m and k may be written as a double sum over $l = m-k$ and k as follows:

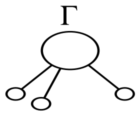
$$\begin{array}{rl} & \sum\limits_{m=0}^{\infty}\frac{i^m}{m!}\sum\limits_{k=0}^{m}\left(\begin{array}{c} m \\ k\end{array}\right)\cdots \\ = & \sum\limits_{l=0}^{\infty}\frac{i^l}{l!}\cdots\sum\limits_{k=0}^{\infty}\frac{i^k}{k!}\cdots\end{array}$$

the proposition is proven.

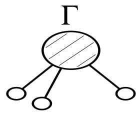
2. Connected Green functions

A full Green function is determined by a sum of graphs which may have several connected components:

$$< 0|T\left\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\right\}|0> \quad = \sum_{\substack{\Gamma \\ \otimes}}$$



$$< 0|T\left\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\right\}|0>^{\text{conn}} \quad \equiv \sum_{\substack{\Gamma \\ \otimes}\ \text{conn}}$$



and they are related by the

Theorem 1:

$$< 0|T\left\{\phi(x_1)\,\phi(x_2)\cdots\phi(x_n)\right\}|0>$$

$$= \sum_{\substack{\text{partitions}\\ i}} < 0|T\{\phi(x_{i_{11}})\,\phi(x_{i_{12}})\cdots\phi(x_{i_{1n_1}})\}|0>^{\text{conn}}\cdots < 0|T\{\phi(x_{i_{k1}})\,\phi(x_{i_{k2}})\cdots\phi(x_{i_{kn_k}})\}|0>^{\text{conn}}$$

where the sum extends over all <u>partitions</u> of $(1,\cdots,n)$ into k classes $(i_{11},\cdots,i_{1n_1})\cdots(i_{k1},\cdots,i_{kn_k})$ with $n_1+n_2+\cdots+n_k=n$.

Proposition: If $Z\{J\}$ is the generating functional of the full time ordered Green functions , then $G\{J\}=\ln Z\{J\}$ is the generating functional of the connected Green functions:

$$G\{J\} \doteq \sum_{n=1}^{\infty}\frac{1}{n!}\int d^dx_1\,\ldots\int d^dx_n\,J(x_1)\,\ldots J(x_n)\,G^{(n)}(x_1,\ldots,x_n)$$

where

$$G^{(n)}(x_1,\ldots,x_n)\equiv< 0\mid T\left\{\varphi(x_1)\ldots\varphi(x_n)\right\}\mid 0>^{\text{conn}}$$

Proof:

$$\begin{aligned}
\left.\frac{\delta^n Z\{J\}}{\delta J(x_1)\cdots\delta J(x_n)}\right|_{J=0} &= \left.\frac{\delta^n}{\delta J(x_1)\cdots\delta J(x_n)}\exp G\{J\}\right|_{J=0}\\
< 0|T\left\{\phi(x_1)\cdots\phi(x_n)\right\}|0> &= \sum_{k=0}^{n}\frac{1}{k!}\left.\frac{\delta^n}{\delta J(x_1)\cdots\delta J(x_n)}\left(G\{J\}\right)^k\right|_{J=0}\\
= \sum_{k=0}^{n}\sum_{\substack{\text{partitions}\\ i}} G^{(n_1)}(x_{i_{11}},\cdots,x_{i_{1n_1}})&\cdots G^{(n_k)}(x_{i_{k1}},\cdots,x_{i_{kn_k}})
\end{aligned}$$

with the sum over all partitions of n into k classes. The inverse relationship is obtained by calculating

$$G^{(n)}(x_1,\ldots,x_n)=\left.\frac{\delta^n}{\delta J(x_1)\cdots\delta J(x_n)}\ln Z\{J\}\right|_{J=0}$$

The proof that $G^{(n)}(x_1,\ldots,x_n)$ is given by a sum of connected diagrams only proceeds by assuming that $G^{(n)}(x_1,\ldots,x_n)$ contains two disconnected pieces and thus factorizes into two factors. This can be shown to lead to a contradiction to the above recurrence relation.

3. Vertex functions

The generating functional of the vertex functions is the functional **Legendre transform** of the generating functional of the connected Green functions:

Definition: The functional $\Gamma\{K\}$ is defined be the functional Legendre transform

$$\Gamma\{K\} = G\{J\} - i \int d^d y \, \left(K(y) + G^{(1)} \right) \, J(y)$$

with

$$K(y) = -i \frac{\delta G\{J\}}{\delta J(y)} - G^{(1)} \; ; \;\; J(y) = i \frac{\delta \Gamma\{K\}}{\delta K(y)} \quad .$$

where

$$G^{(1)} = < 0|\phi(y)|0 > = \text{constant}$$

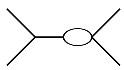
and

$$\Gamma\{0\} = G\{0\} \; ; \quad K(y)|_{J=0} = 0$$

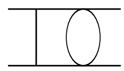
Proposition:

$$\Gamma\{K\} \doteq \sum_{n=2}^{\infty} \frac{1}{n!} \int d^d y_1 \, \ldots \int d^d y_n \, K(y_1) \, \ldots K(y_n) \, \Gamma^{(n)}(y_1, \ldots, y_n)$$

generates the amputated one–particle irreducible ($1pi$) Green functions, called **vertex functions**. A Feynman diagram is $1pi$ if it cannot be divided into two disconnected parts by cutting one line. Examples:



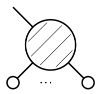
one particle reducible



one particle irreducible

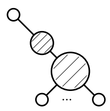
Amputation: is defined by convolution with the inverse propagator

$$\int dy_1 \, G^{(2)-1}(x_1 - y_1) \, G^{(n)}(y_1, x_2, \cdots, x_n) \equiv G^{(n)}(\underline{x_1}, x_2, \cdots, x_n) =$$

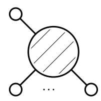


such that conversely

$$\int dy_1 \, G^{(2)}(x_1 - y_1) \, G^{(n)}(\underline{y_1}, x_2, \cdots, x_n) \qquad \equiv \qquad G^{(n)}(x_1, x_2, \cdots, x_n)$$



$=$



The

Inverse propagator: satisfies

$$\int dy \, G^{(2)-1}(x - y) \, G^{(2)}(y - z) = \delta^{(d)}(x - z)$$

Proof of the proposition:

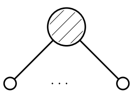
$$G^{(n)}(x_1, \cdots, x_n) = \frac{\delta^n}{\delta J(x_1) \cdots \delta J(x_n)} \left\{ \Gamma\{K\} + i \int dy \, \left(K(y) + G^{(1)} \right) \, J(y) \right\} \Big|_{J=0}$$

$$= \sum_{k=2}^{n} \sum_{\ell=1}^{k} \frac{(-1)^{k-\ell}}{(k-\ell)!} \sum_{\substack{\text{partitions} \\ i}} \int dy_1 \cdots dy_k \, \Gamma^{(k)}(y_1, \cdots, y_k)$$

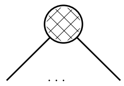
$$\times \quad G^{(n_1+1)}(x_{i_{11}}, \cdots, x_{i_{1n_1}}; y_1) \cdots G^{(n_\ell+1)}(x_{i_{\ell 1}}, \cdots, x_{i_{\ell n_\ell}}; y_\ell)$$

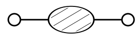
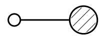
$$\times \quad G^{(1)}(y_{\ell+1}) \cdots G^{(1)}(y_k)$$

Graphically this reads:

$$G^{(n)} = \sum_{\substack{\Gamma \\ \otimes \text{ conn}}}$$  $\underline{n > 2}$

is a connected **tree** (no loops) in the $\Gamma^{(i)}$'s:

$$\Gamma^{(i)} = \sum_{\substack{\Gamma \text{ conn} \\ \otimes \ 1pi}}$$  $(n \geq i > 2)$

connected by <u>propagators</u> $G^{(2)} =$  and <u>tadpoles</u> $G^{(1)} =$  . All loops are contained in the vertices $\Gamma^{(i)}$, which do not exhibit any "one particle lines".

<u>Reverse relationship:</u>

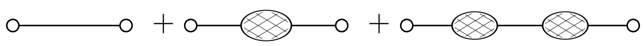
$$\begin{aligned}
&\Gamma^{(n)}(x_1, \cdots, x_n) = \frac{\delta^n}{\delta K(x_1) \cdots \delta K(x_n)} \left\{ G\left\{J\right\} - i \int dx \, \left(K(x) + G^{(1)} \right) \, J(x) \right\} \Bigg|_{K=0} \\
= \; &\sum_{k=2}^{n} (-1)^k \sum_{\substack{\text{partitions} \\ i}} \int dy_1 \cdots dy_k \, G^{(k)}(y_1, \cdots, y_k) \times \\
&\Gamma^{(\ell_1+1)}(x_{i_{11}}, \cdots, x_{i_{1\ell_1}}; y_1) \cdots \Gamma^{(\ell_k+1)}(x_{i_{k1}}, \cdots, x_{i_{k\ell_k}}; y_k)
\end{aligned}$$

where $0 < \ell_i < n - 1$.

The proof proceeds as follows: $\Gamma^{(i)}$ is given by $1pi$ diagrams, if not, then there exists a line such that omission of that line implies that $\Gamma^{(i)}$ consists of disconnected parts and thus factorizes. But this leads to a contradiction to the recurrence relation.

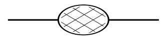
By definition:

$$\Gamma^{(2)} = -G^{(2)-1} \tag{D.1}$$

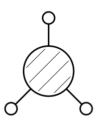
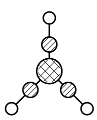
$$G^{(2)} =$$  $+ \cdots$

$$\begin{aligned}
&= G_0^{(2)} \left(1 + \tilde{\Gamma}^{(2)} G_0^{(2)} + \left(\tilde{\Gamma}^{(2)} G_0^{(2)} \right)^2 + \cdots \right) \\
&= G_0^{(2)} \cdot \frac{1}{1 - \tilde{\Gamma}^{(2)} G_0^{(2)}} = \frac{1}{G_0^{(2)-1} - \tilde{\Gamma}^{(2)}} = \frac{1}{p^2 - m^2 - \tilde{\Gamma}^{(2)}}
\end{aligned}$$

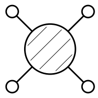
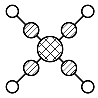
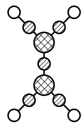
$$\Gamma^{(2)} = -G^{(2)-1} + \tilde{\Gamma}^{(2)} = -G^{(2)-1}$$

$= -$ $^{-1} + \sum_{\substack{\Gamma \text{ conn} \\ \otimes \ 1pi}}$ 

Examples: we assume $G^{(1)} = 0$ for simplicity

$G^{(3)} =$  $=$ 

$$= \int dy_1\, dy_2\, dy_3\, G^{(2)}(x_1, y_1)\, G^{(2)}(x_2, y_2)\, G^{(2)}(x_3, y_3)\, \Gamma^{(3)}(y_1.y_2, y_3)$$

$G^{(4)}$ =  =  +  + crossed diagrams

D.5 Unitarity and Locality

The aim of this section is to discuss the tools by which unitarity and locality can be proven to all orders in perturbation theory. In order to keep notation simple we will often restrict ourselves to prove the relevant relationships for the scalar self-interacting ϕ^4-model. The relevant properties which allow to generalize the proofs to an arbitrary renormalizable field theory will be discussed as well.

D.5.1 Generalized (off-shell) unitarity

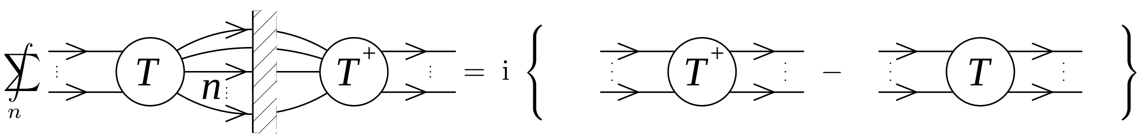
1. On-shell unitarity

The physical unitarity is a property of S –matrix elements between scattering states with the particles on their mass-shells ($p^2 = m^2$) and thus having energy $p^0 = \omega_p = \sqrt{\vec{p}^2 + m^2}$. The unitarity is usually formulated for the related T–matrix elements and reads (see Eq. 3.19):

$$\begin{aligned} &\textstyle\sum_n \sum_{\alpha_1 \cdots \alpha_n} \int \prod_{i=1}^n d\mu(p_i)\, (2\pi)^4\, \delta^{(4)}(\sum_{i=1}^{n_i} p_i^{'} - \sum_{i=1}^n p_i) \\ &T^*(p_1\alpha_1, \ldots, p_n\alpha_n | p_1^{''}\alpha_1^{''}, \ldots) \,\cdot\, T(p_1\alpha_1, \ldots, p_n\alpha_n | p_1^{'}\alpha_1^{'}, \ldots) \\ &\quad = \mathrm{i}\left\{ T^*(p_1^{'}\alpha_1^{'}, \ldots | p_1^{''}\alpha_1^{''}, \ldots) \,-\, T(p_1^{''}\alpha_1^{''}, \ldots | p_1^{'}\alpha_1^{'}, \ldots) \right\} \end{aligned} \tag{D.1}$$

where $\{p_i\alpha_i\}$ denotes a complete set of quantum numbers characterizing particle number i in the scattering state. By $|p_1^{'}\alpha_1^{'}, \ldots, p_{n_i}^{'}\alpha_{n_i}^{'} >$ we denoted the *in*–state, $|p_1^{''}\alpha_1^{''}, \ldots, p_{n_i}^{''}\alpha_{n_i}^{''} >$ denotes the *out*–state, while $|p_1\alpha_1, \ldots, p_n\alpha_n >$ is a n–particle intermediate state.

The on–shell unitarity relation, known as the optical theorem, graphically may be represended as:



2. Off-shell generalized unitarity

The aim is to turn the above relationship between S–matrix elements into a relationship between the corresponding time ordered Green functions from which the S–matrix elements follow via the LSZ reduction formulae (see Sec. 3.3). Thus, the external on–shell particles are replaced by virtual off–shell particles, with p^0 as an independent variable. We first consider the time ordered Green functions in configuration space and later transform the result to momentum space.

Let us denote by $T(x_1, \ldots, x_n) \doteq T\{\phi(x_1) \cdots \phi(x_n)\}$ the time–ordered , and by $\bar{T}(x_1, \ldots, x_n) \doteq \bar{T}\{\phi(x_1) \cdots \phi(x_n)\}$ the anti time–ordered products of interacting fields. The following important **identity** holds:

$$\Big\| \sum_{(i_1,\ldots,i_\nu)} (-1)^\nu\, T(x_{i_1}, \ldots, x_{i_\nu})\, \bar{T}(x_{i_{\nu+1}}, \ldots, x_{i_n}) = 0 \tag{D.2}$$

where $(i_1, \ldots, i_\nu, i_{\nu+1}, \ldots, i_n)$ is a permutation of $(1, \ldots, n)$, the indices labeling the external lines (fields) .

Proof (by induction):
For $n = 2$ the identity reads

$$T(x_1, x_2) \; - T(x_1)\, \bar{T}(x_2) \; - T(x_2)\, \bar{T}(x_1) \; + \bar{T}(x_1, x_2) = 0$$

and is true by definition of the time ordering operation:

$$\begin{array}{lcl} x_1^0 > x_2^0 & : & \phi(x_1)\phi(x_2) \; - \phi(x_1)\phi(x_2) \; - \phi(x_2)\phi(x_1) \; + \phi(x_2)\phi(x_1) \; = 0 \\ x_2^0 > x_1^0 & : & \phi(x_2)\phi(x_1) \; - \phi(x_1)\phi(x_2) \; - \phi(x_2)\phi(x_1) \; + \phi(x_1)\phi(x_2) \; = 0 \; . \end{array}$$

The identity then follows by **induction**: Let $x_n^0 > x_j^0$ for all $j \neq n$. Then we may write

$$\begin{array}{cl} & \sum_{(i_1,\ldots,i_\nu)} (-1)^\nu \, T(x_{i_1}, \ldots, x_{i_\nu})\, \bar{T}(x_{i_{\nu+1}}, \ldots, x_{i_n}) \\ = & \sum_{n\in\,(i_1,\ldots,i_\nu)} (-1)^\nu \, \phi(x_n)\, T(x_{i_1}, \ldots, \check{x}_n, \ldots, x_{i_\nu})\, \bar{T}(x_{i_{\nu+1}}, \ldots, x_{i_n}) \\ + & \sum_{n\notin\,(i_1,\ldots,i_\nu)} (-1)^\nu \, T(x_{i_1}, \ldots, x_{i_\nu})\, \bar{T}(x_{i_{\nu+1}}, \ldots, \check{x}_n, \ldots, x_{i_n})\, \phi(x_n) \\ = & -\phi(x_n) \, \sum_{(i_1',\ldots,i_{\nu'}')} (-1)^{\nu'} \, T(x_{i_1'}, \ldots, x_{i_{\nu'}'})\, \bar{T}(x_{i_{\nu'+1}'}, \ldots, x_{i_{n-1}'}) \\ + & \sum_{(i_1',\ldots,i_\nu')} (-1)^\nu \, T(x_{i_1'}, \ldots, x_{i_\nu'})\, \bar{T}(x_{i_{\nu+1}'}, \ldots, x_{i_{n-1}'})\, \phi(x_n) \; = 0 \end{array}$$

where $(i_1', \ldots, i_{n-1}')$ is a permutation of $(1, \ldots, n-1)$ and $\nu' = \nu - 1$. By $\check{x}_i$ we indicate that the argument x_i has to be omitted at that position. The result is zero since the sums $\sum_{(i_1,\ldots,i_{\nu''})} (-1)^{\nu''} \, T(x_{i_1}, \ldots, x_{i_{\nu''}})\, \bar{T}(x_{i_{\nu''+1}}, \ldots, x_{i_{n-1}}) \;=\; 0$ by the induction hypothesis. This proves the identity as the induction hypothesis has shown to be true for $n = 2$.

We know may apply the reduction formula (see Eq. 3.20):

$$\begin{array}{l} < 0|T\{\phi(x_1) \cdots \phi(x_m)\}|p_1, \cdots, p_n >_{in} = i^n \displaystyle\int dy_1 \cdots dy_n e^{-i(y_1 p_1 + \cdots + y_n p_n)} \times \\ (\Box_{y_1} + m^2) \; \cdots \; (\Box_{y_n} + m^2) \; < 0|T\{\phi(x_1) \cdots \phi(x_m) \, \phi(y_1) \cdots \phi(y_n)\}|0 > \quad . \end{array}$$

We will denote $< 0|T\{\phi(x_1) \cdots \phi(x_m) \, \phi(y_1) \cdots \phi(y_n)\}|0 >= \tau(x_1, \ldots, x_m, y_1, \ldots, y_n)$ and define amputated arguments by

$$(\Box_{y_i} + m^2)\, \tau(x_1, \ldots, y_i, \ldots, y_n) \doteq \tau(x_1, \ldots, \underline{y_i}, \ldots, y_n) \; .$$

Inserting the **completeness relation** (we suppress explicit summation over the other quantum numbers α which were exhibited in Eq. D.1)

$$\sum_n \int d\mu(p_1) \cdots d\mu(p_n) \, |p_1, \ldots, p_n >_{in} \; {}_{in}\!< p_1, \ldots, p_n| = 1$$

into our identity D.2 and taking the vacuum expectation value we have

$$\begin{array}{rcl} 0 & = & \sum_{(i_1,\ldots,i_\nu)} (-1)^\nu \sum_m \int d\mu(p_1) \cdots d\mu(p_m) \\ & & < 0|T(x_{i_1}, \ldots, x_{i_\nu})\, |p_1, \ldots, p_m >_{in} \; {}_{in}\!< p_1, \ldots, p_m|\bar{T}(x_{i_{\nu+1}}, \ldots, x_{i_n})|0 > \\ & = & \int dy_1 \cdots dy_m dy_1' \cdots dy_m' \; \sum_m \sum_{(i_1,\ldots,i_\nu)} (-1)^\nu \int d\mu(p_1) \cdots d\mu(p_m) \, e^{-i\sum_1^m p_i(y_i - y_i')} \\ & & \tau(x_{i_1}, \ldots, x_{i_\nu}, \underline{y_1}, \ldots, \underline{y_m}) \; \cdot \bar{\tau}(x_{i_{\nu+1}}, \ldots, x_{i_n}, \underline{y_1'}, \ldots, \underline{y_m'}) \\ & = & \sum_m \sum_{(i_1,\ldots,i_\nu)} (-1)^\nu \int dy_1 \cdots dy_m' \\ & & \tau(x_{i_1}, \ldots, x_{i_\nu}, \underline{y_1}, \ldots, \underline{y_m}) \; \prod_{i=1}^m (-i\Delta^+(y_i - y_i'))\, \bar{\tau}(x_{i_{\nu+1}}, \ldots, x_{i_n}, \underline{y_1'}, \ldots, \underline{y_m'}) \end{array} \tag{D.3}$$

this is the **generalized unitarity** relation in configuration space we were looking for. Note that

$$\int d\mu(p)\, e^{-ipy} = -i\Delta^{+}(y)$$

is the positive frequency part of the free field commutator. Performing the Fourier transformation

$$\tau(x_1,\dots,x_n,\underline{y_1},\dots,\underline{y_m}) = \int \frac{d^4q_1}{(2\pi)^4}\cdots\frac{d^4p_1}{(2\pi)^4}\cdots e^{-i\sum_{i=1}^{n} q_i x_i} e^{+i\sum_{i=1}^{m} p_i y_i}\ \tau(q_1,\dots,q_n,\underline{p_1},\dots,\underline{p_m})$$

$$\bar{\tau}(x_1,\dots,x_n,\underline{y_1^{'}},\dots,\underline{y_m^{'}}) = \int \frac{d^4q_1^{'}}{(2\pi)^4}\cdots\frac{d^4p_1^{'}}{(2\pi)^4}\cdots e^{-i\sum_{i=1}^{n} q_i^{'} x_i} e^{-i\sum_{i=1}^{m} p_i^{'} y_i^{'}}\ \bar{\tau}(-q_1^{'},\dots,-q_n^{'},\underline{p_1^{'}},\dots,\underline{p_m^{'}})$$

inserting this into D.3 and taking the Fourier transform

$$\int \prod_{i=1}^{n} \frac{d^4 q_i}{(2\pi)^4} e^{-i\sum_{i=1}^{n} q_i x_i} \dots$$

then yields

$$\begin{aligned}
0 &= \textstyle\sum_m \sum_{(i_1,\dots,i_\nu)} (-1)^\nu \int d\mu(p_1)\cdots d\mu(p_m) \\
&\quad \tau(q_{i_1},\dots,q_{i_\nu},\underline{p_1},\dots,\underline{p_m})\, \bar{\tau}(-q_{i_{\nu+1}},\dots,-q_{i_n},\underline{p_1},\dots,\underline{p_m}) \\
&= \textstyle\sum_m \sum_{(i_1,\dots,i_\nu)} (-1)^\nu \int \prod_{i=1}^{m} \frac{d^4p_i}{(2\pi)^4} \\
&\quad \tau(q_{i_1},\dots,q_{i_\nu},\underline{p_1},\dots,\underline{p_m})\ \left(\prod_{i=1}^{m} 2\pi\Theta(p_i^0)\,\delta(p_i^2-m_i^2)\,\right) \bar{\tau}(-q_{i_{\nu+1}},\dots,-q_{i_n},\underline{p_1},\dots,\underline{p_m})
\end{aligned} \tag{D.4}$$

which is the **generalized unitarity in momentum space**.
Note: For real fields (neutral particles) we have

$$\begin{aligned}
&{}_{in}< p_1,\dots,p_m|\bar{T}\{\phi(x_1)\cdots\phi(x_n)\}|0> = < 0|T\{\phi(x_1)\cdots\phi(x_n)\}|p_1,\dots,p_m>_{in}^{*} \\
&= (-i)^n \int dy_1\cdots dy_m e^{i\sum_{i=1}^{m} y_i p_i} (\Box_{y_1}+m^2)\cdots(\Box_{y_m}+m^2)\ < 0|\bar{T}\{\phi(x_1)\cdots\phi(y_m)\}|0> \\
&= (-i)^n \int dy_1\cdots dy_m e^{i\sum_{i=1}^{m} y_i p_i} \bar{\tau}(x_1,\dots,x_n,\underline{y_1},\dots,\underline{y_m}) \\
&= (-i)^n \int dy_1\cdots dy_m e^{i\sum_{i=1}^{m} y_i p_i} \tau(x_1,\dots,x_n,\underline{y_1},\dots,\underline{y_m})^{*}
\end{aligned}$$

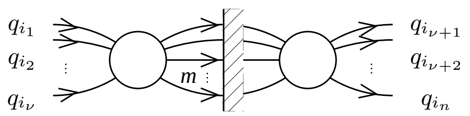
and thus the following reality properties

$$\begin{aligned}
\bar{\tau}(x_1,\dots,x_n,\underline{y_1},\dots,\underline{y_m}) &= \tau(x_1,\dots,x_n,\underline{y_1},\dots,\underline{y_m})^{*} \\
\bar{\tau}(q_1,\dots,q_n,\underline{p_1},\dots,\underline{p_m}) &= \tau(q_1,\dots,q_n,\underline{p_1},\dots,\underline{p_m})^{*}
\end{aligned}$$

hold for the τ–functions of neutral particles.

Finally, we may write Eq. D.4 in diagrammatic form as follows:

$$\sum_{m} \sum_{(i_1,\dots,i_\nu)} (-1)^\nu$$



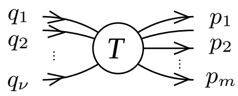
$= 0$ (D.5)

For a given diagram with n fixed external fields the sum extends over all possible **cuts** of lines which separate the diagram into two parts. For each cut diagram ("unitarity graph") we distinguish three types of lines:

1. To the left of the cut, excluding the cut lines, the normal Feynman rules for time-ordered Green functions apply with lines representing standard Feynman propagators.

Graphically:

$$\tau(q_1,\ldots,q_\nu,\underline{p_1},\ldots,\underline{p_m}) :=$$

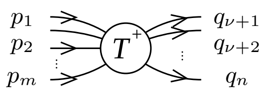


where the q–lines represent external propagators, the p–lines are amputated (for scalar fields $i(-p_i^2+m_i^2)\,\frac{i}{p_i^2-m_i^2+i\epsilon} = 1)$. T stands for the time ordered amplitude (product of Feynman propagators).

2. To the right of the cut we have anti time–ordered propagators and amplitudes.

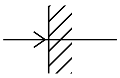
Graphically:

$$\bar\tau(-q_{\nu+1},\ldots,-q_n,\underline{p_1},\ldots,\underline{p_m}) :=$$



$\bar T$ represents the complex conjugate amplitude of the hermitian conjugate fields.

3. The **cut–lines** represent real physical (on–shell) particles



$$:= \delta_+(p_i^2-m_i^2) = 2\pi\,\Theta(p^0)\,\delta(p^2-m^2)$$

and are integrated over their n–particle phase space. Thus, for the cut lines the Feynman propagators have been replaced by positive on-shell distributions:

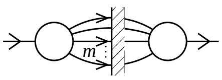
$$\frac{i}{p^2-m^2+i\epsilon} \Rightarrow 2\pi\,\Theta(p^0)\,\delta(p^2-m^2)$$

Example: <u>Propagator identity</u> $n=2$; $\nu=0,1,2$; momentum $q=q_1=-q_2$ with $q^0>0$

$$\tau(q_1,q_2)+\bar\tau(-q_1,-q_2) = \sum_m \int \prod_{i=1}^m \frac{d^4p_i}{(2\pi)^4}\tau(q_1,\underline{p_1},\ldots,\underline{p_m})\,\prod_{i=1}^m \delta_+(p_i^2-m_i^2)\bar\tau(-q_2,\underline{p_1},\ldots,\underline{p_m})$$

Graphically:

$$-2\,\tilde\Delta_F'(p) = \sum_m$$



.

The second term with

$$\tau(q_2,\underline{p_1},\ldots,\underline{p_m})\,\prod_{i=1}^m \delta_+(p_i^2-m_i^2)\bar\tau(-q_1,\underline{p_1},\ldots,\underline{p_m})$$

vanishes for $q^0>0$, which we assumed.

In terms of the full propagator $i\tilde\Delta_F'(q)$ the τ–function reads

$$\begin{aligned}\tau(q_1,q_2) &= i\tilde\Delta_F'(q_1)\,(2\pi)^4\,\delta^{(4)}(q_1+q_2)\\ \bar\tau(-q_1,-q_2) &= \tau(-q_1,-q_2)^* = -i\tilde\Delta_F'^*(-q_1)\,(2\pi)^4\,\delta^{(4)}(q_1+q_2)\end{aligned}$$

and hence

$$\begin{aligned} \tau(q_1,q_2)+\bar{\tau}(-q_1,-q_2) &= i\left\{\tilde{\Delta}'_F(q_1)\ -\tilde{\Delta}'^{*}_F(q_1)\right\}(2\pi)^4\,\delta^{(4)}(q_1+q_2) \\ &= -2\mathrm{Im}\tilde{\Delta}'_F(q_1)\,(2\pi)^4\,\delta^{(4)}(q_1+q_2)\ \ . \end{aligned}$$

Above, we have used the symmetry $\tilde{\Delta}'_F(-q_1)=\tilde{\Delta}'_F(q_1)$ valid for the propagator of scalar fields. If we factor out the four-momentum conservation δ–function ("truncation")

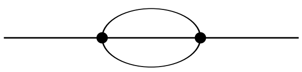
$$\tau(q_1,\ldots,q_n,\underline{p_1},\ldots,\underline{p_m})=(2\pi)^4\,\delta^{(4)}(\sum q_i-\sum p_j)\;\tau(q_1,\ldots,q_n,\underline{p_1},\ldots,\underline{p_m})^{\mathrm{trunc}}$$

we find the **absorptive part** of the propagator

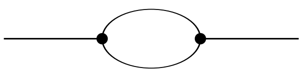
$$\begin{aligned} & \tau(q_1,q_2)|_{\text{absorptive part}} = -2\mathrm{Im}\tilde{\Delta}'_F(q) \\ = \ & \textstyle\sum_m \int \prod_{i=1}^m \frac{d^4p_i}{(2\pi)^4}(2\pi)^4\,\delta^{(4)}(q-\sum_1^m p_i)\ \prod_{i=1}^m \delta_+(p_i^2-m_i^2)|\tau(q,\underline{p_1},\cdots,\underline{p_m})^{\mathrm{trunc}}|^2 \\ \geq 0 & \end{aligned} \tag{D.6}$$

this is non–negative because $|\cdots|^2$ is integrated over the phase space.

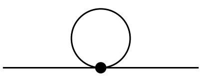
Physical interpretation: Let us assume that the field theory model under consideration has a mass gap, i.e., no massless particles and the lightest particle has mass m. Then for $q^2<4m^2$ the absorptive part vanishes, only **elastic processes** take place. For $q^2\geq 4m^2$ the right hand side of Eq. D.6 is positive if the production of two real physical particles of mass m is allowed by the dynamics, kinematically the process is allowed for $q^0\geq 2m$. The role of the dynamics may be easily illustrated for scalar self-interacting fields. In the ϕ^4–model, with unbroken discrete symmetry $\phi\to-\phi$ only the ϕ^4 interaction vertex exists and thus a transition $\phi\to\phi\phi$ is forbidden. If the symmetry is broken a ϕ^3 interaction term is present and the process is allowed. In the symmetric ϕ^4-model the threshold for the **inelastic channels** therefore is $q^0\geq 3m$ (i.e. $q^2\geq 9m^2$), in spite of the fact that two–particle production is kinematically allowed. In the symmetric ϕ^4-model the lowest non-trivial contribution to the propagator is the two–loop self-energy diagram



while in the broken case we the one–loop diagram



contributes. Note that the one–loop diagram



possible in the symmetric case is just a constant (which only can contribute to the mass counter term) and can not contribute to the imaginary part.
In general the full renormalized scalar propagator has the form

$$i\tilde{\Delta}'_F(q)\ =\ \frac{i}{q^2-m^2-\Pi(q^2)}\ \text{with}\ \Pi(q^2)=\mathrm{Re}\Pi(q^2)+i\mathrm{Im}\Pi(q^2) \tag{D.7}$$

and thus

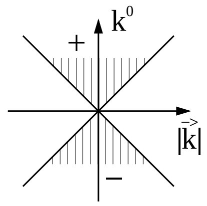
$$\mathrm{Im}\tilde{\Delta}'_F(q) \; = \mathrm{Im}\Pi(q^2) / \Big((q^2 - m^2 - \mathrm{Re}\Pi(q^2))^2 + (\mathrm{Im}\Pi(q^2)^2 \Big) \quad . \tag{D.8}$$

This may be expanded into a perturbation series for $|q^2 - m^2| \gg O(\lambda^2)$ if we denote the coupling constant with λ and note that $\mathrm{Re}\Pi(q^2), \mathrm{Im}\Pi(q^2) = O(\lambda^2)$. If there exist a physical particle , ϕ' say, with mass M larger than $2m$ and which is allowed to decay ("on it's mass-shell" $q^2 \simeq M^2$) as $\phi' \to \phi\phi$, then ϕ' is an **unstable particle** and strictly speaking can not appear as an asymptotic state in the S–matrix. If, however, $\mathrm{Im}\Pi(q^2)/M \ll M$ for $q^2 \simeq M^2$, then ϕ' is **quasi stable** and, for example, could leave a classical track in a particle detector. In this case it still might be helpful to treat the particle like a stable one. We mention that $\mathrm{Im}\Pi(M^2)/M = \Gamma$ may be used as a definition of the **width** of the instable particle $1/\Gamma = \tau$ being its **life–time**.

D.5.2 Veltman's cutting formula

Veltman's cutting formula is an important tool which allows us to write down simple proofs of unitarity and locality in perturbation theory. Let us denote by $\Delta_F(x)$ a **generalized Feynman propagator** [48] with the following properties

$$\begin{array}{lcl}
\Delta_F(x) & : & \text{generalized Feynman propagator} \\
\Delta_F(x) & = & \Theta(x^0)\,\Delta^+(x) \; - \Theta(-x^0)\,\Delta^-(x) \\
\Delta^\pm(x) & = & \frac{\pm i}{(2\pi)^{d-1}} \int d^dk \, e^{-ikx}\,\Theta(\pm k^0)\,\Theta(k^2)\,\rho(k^2) \\
\Delta^+(x) & & \text{positive frequency part} \\
\Delta^-(x) & & \text{negative frequency part} \\
\Delta^-(x) & = & (\Delta^+(x))^* \quad ; \quad \Delta^-(x) \; = \; -\,\Delta^+(-x)
\end{array} \tag{D.9}$$

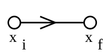


with ρ a positive spectral function.

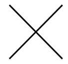
Consider Feynman integrand in configuration space: Typically

$$F(x_1, \ldots, x_n) = (-ig)^n \prod_{\ell \in L} i\Delta_{F\ell}(x_{\ell\, i} - x_{\ell\, f}) \tag{D.10}$$

where L is the set of lines of the Feynman diagram and



: $i\Delta_F(x_i - x_f)$ propagators


: $-ig$ vertices

Define:

$$\begin{array}{ccc}
F(x_1, \ldots, \check{x}_{i_1}, \ldots, \check{x}_{i_m}, \ldots, x_n & | & x_{i_1}, \ldots\ldots\ldots, x_{i_m}) \\
\text{unmarked vertices} & & \text{marked vertices} \\
\text{(unshaded part)} & & \text{(shaded part)}
\end{array} \tag{D.11}$$

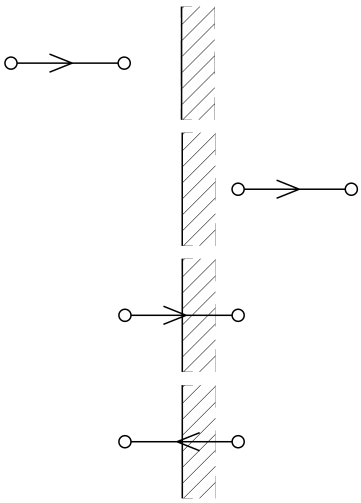
$$\begin{aligned}
&= (-ig)^{n-m}\,(ig)^m\,\textstyle\prod_{\ell \in L} f_\ell(x_{\ell\, i} - x_{\ell\, f}) \times (-1)^m \\
&= (-ig)^n\,\textstyle\prod_{\ell \in L} f_\ell(x_{\ell\, i} - x_{\ell\, f})
\end{aligned}$$

[48] This may be a free or a full propagator of any spin and in particular it may be a regularized propagator. Of course the required properties are only true for particular regularizations, like e.g. for dimensional regularization

where

$$f_\ell(x_{\ell\,i} - x_{\ell\,f}) = \begin{cases} i\Delta_{F\ell}(x_{\ell\,i} - x_{\ell\,f}) & x_{\ell\,i}, x_{\ell\,f} \in L - L^* \\ -i\Delta^*_{F\ell}(x_{\ell\,i} - x_{\ell\,f}) & x_{\ell\,i}, x_{\ell\,f} \in L^* \\ i\Delta^+_\ell(x_{\ell\,i} - x_{\ell\,f}) & x_{\ell\,i} \in L - L^*, x_{\ell\,f} \in L^* \\ -i\Delta^-_\ell(x_{\ell\,i} - x_{\ell\,f}) & x_{\ell\,i} \in L^*, x_{\ell\,f} \in L - L^* \end{cases} \tag{D.12}$$

or graphically



$: \quad i\Delta_F$

$: \quad -i\Delta^*_F$

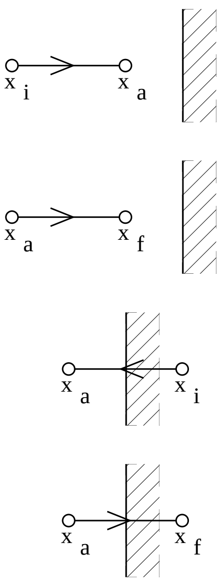
$: \quad i\Delta^+$

$: \quad -i\Delta^-$

Lemma: Let $x_a^0 \leq x_k^0$ for all $k = 1, \dots, n$ then

$$\begin{aligned} & F(x_1, \dots, \check{x}_a, \dots, \check{x}_{i_1}, \dots, \check{x}_{i_m}, \dots, x_n, x_a | x_{i_1}, \dots, x_{i_m}) \\ = \; & F(x_1, \dots, \check{x}_a, \dots, \check{x}_{i_1}, \dots, \check{x}_{i_m}, \dots, x_n | x_{i_1}, \dots, x_{i_m}, x_a) \ , \end{aligned}$$

i.e., $F(\ \ |\ \ )$ does not change if the argument with the smallest x^0 is moved from the unmarked to the marked side and vice versa.
Proof: By inspection of the left hand side:



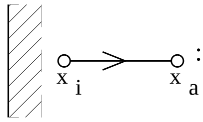
a) $: \quad +i\Delta_F(x_i - x_a) \quad = \quad +i\Delta^+(x_i - x_a)$

b) $: \quad +i\Delta_F(x_a - x_f) \quad = \quad -i\Delta^-(x_a - x_f)$

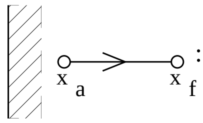
c) $: \quad -i\Delta^-(x_i - x_a)$

d) $: \quad +i\Delta^+(x_a - x_f)$

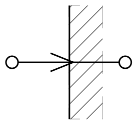
and of the right hand side:

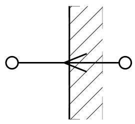
a) 

$: \quad -i\Delta_F^*(x_i - x_a) \quad = \quad -i\Delta^-(x_i - x_a)$

b) 

$: \quad -i\Delta_F^*(x_a - x_f) \quad = \quad +i\Delta^+(x_a - x_f)$

c) 
$: \quad +i\Delta^+(x_i - x_a)$

d) 
$: \quad -i\Delta^-(x_a - x_f)$

the identity is satisfied.

Theorem (an identity for integrands): Let $\{y_1, \ldots, y_p\} \subset \{x_1, \ldots, x_n\}$ and $x_a \in \{y_1, \ldots, y_p\}$ where x_a is the argument with the earliest time component,i.e., $x_a^0 \leq x_i^0 \forall i$. Furthermore, let us denote by $\{x_1, \ldots, x_{n-p}\}$ the arguments belonging to the complement of $\{y_1, \ldots, y_p\}$. This is always possible by relabeling the variables. Then we have

$$\sum_{q=0}^{p} \sum_{\{q\}} (-1)^q \; F(x_1, \ldots, x_{n-p}, y_1, \ldots, y_p | x_{i_1}, \ldots, x_{i_m}, y_{j_1}, \ldots, y_{j_q}) = 0 \tag{D.13}$$

where $\{q\}$ is a ordered choice $j_1, \ldots, j_q$ of indices out of the set $1, \ldots, p$ with $j_1 < \ldots < j_q$. Proof: The identity holds because $x_a \in \{y_1, \ldots, y_p\}$ appears always twice, once on the l.h.s and once on the l.h.s. According to the Lemma all terms cancel pairwise.

The identity stated in the previous theorem may be integrated within its domain of validity:

i) Since $x_a \in \{y_1, \ldots, y_p\}$ and $x_i^0 \leq x_a^0$ for all $i = 1, \ldots, n-p$ the variables $x_1, \ldots, x_{n-p}$ cannot be integrated (in the range $-\infty$ to $+\infty$).

ii) Therefore, all integration variables must be in the set $\{y_1, \ldots, y_p\}$, which in addition must contain the earliest non integrated variable.

This implies the following theorem, called **Veltman's cutting formula**:

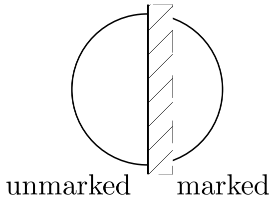
$$\int dy_{\alpha_1} \ldots dy_{\alpha_r} \; \sum_{q=0}^{p} \sum_{\{q\}} (-1)^q \; F(x_1, \ldots, x_{n-p}, y_1, \ldots, y_p | x_{i_1}, \ldots, x_{i_m}, y_{j_1}, \ldots, y_{j_q}) = 0 \tag{D.14}$$

whenever there exist an index $i \neq \{\alpha_1, \ldots, \alpha_r\}$ with $y_i^0 \leq x_i^0$ for $i = 1, \ldots, n-p$. Necessarily we must have $r \leq p - 1$. Note: For obvious reasons the vertices $\{y_{\alpha_1} \ldots y_{\alpha_r}\}$ are called **inner vertices** (the variables integrated over) while $\{x_1 \ldots x_{n-p}\}$ plus the remaining y–vertices are called

external vertices (the variables not integrated over).

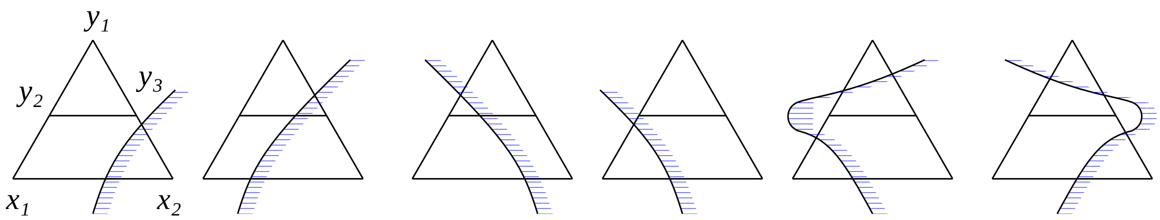
Graphically:

$$\sum_{\text{all cuts}}$$



$$= \quad 0\,.$$

Example: x_1, x_2, y_1 external vertices; y_2, y_3 internal vertices



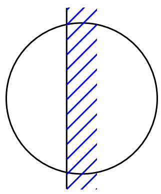
$= 0$

x_1 is unmarked, X_2 marked and $y_1^0 \leq x_1^0, x_2^0$. Using the cutting formula one can show,

- if propagators satisfy the general properties needed for the cutting formula to hold **and** couplings (and counterterms) are real then:

Unitarity: SS^{+}=1

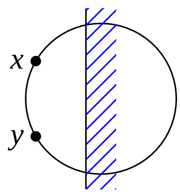
$$\bowtie \quad \sum_{\text{all cuts}}$$



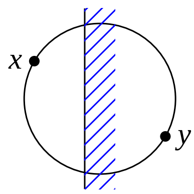
$$= \quad 0$$

Locality (causality): $[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0\ ; \quad (x-y)^2 < 0$

$$\bowtie \quad \sum_{\substack{\text{all cuts}\\ y^0 < x^0}} \Big\{$$



$+$



$$\Big\} \quad = \quad 0$$

The $\mathcal{O}_i(x)$ are vertex operators as they appear in the Lagrangian. The relative + sign of the two contributions refers to a convention where the vertex operators do not include the i, which usually is included in the Feynman rules.

As a consequence, proofs of unitarity and causality require propagators and vertices to satisfy some general requirements as specified above. In gauge theories we have a particular problem: the above proofs only work

(1) for spin 1 particles the mentioned conditions (e.g. positivity) only are satisfied in the Proca gauge form of the massive propagators;

(2) no ghosts are admitted.

This means that these properties can only be satisfied off-shell in the **unitary gauge**. The trick to control both renormalizability and unitarity is the use the renormalizable 't Hooft gauge as a one parameter family of renormalizable gauges which for S–matrix elements has a continuous limit $\xi \to \infty$ representing the unitary gauge. The key point of course is that the S–matrix is gauge inependent, such that it does not depend on ξ per se. This will be proven in the next subsection.

D.6 Gauge invariance of the S–matrix

The gauge transformation of fields is of the form

$$\begin{aligned}\hat{\delta} G_{\mu A} &= D_{\mu AB}\eta_B = \delta_{AB}\partial_\mu \eta_B + \hat{D}'_{AB}\, G_{\mu A}\eta_B \\ \hat{\delta}\varphi_j &= \hat{D}'_{jj'B}\,\varphi_{j'}\eta_B \\ \hat{\delta}\psi_i &= \hat{D}'_{ii'B}\,\psi_{i'}\eta_B\end{aligned}$$

where the $\hat{D}$ are coupling matrices, i.e. they are of higher order in the couplings. If we change the gauge fixing

$$C \to C' = C + \varepsilon R;$$

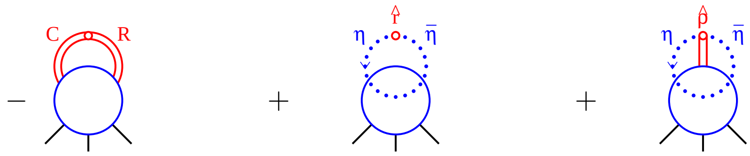
the Lagrangian changes by

$$\mathcal{L} \to \mathcal{L}' = \mathcal{L} - \varepsilon\, C\, R + \varepsilon\, \bar{\eta}\; (\hat{r} + \hat{\rho})\; \eta + O(\varepsilon^2)$$

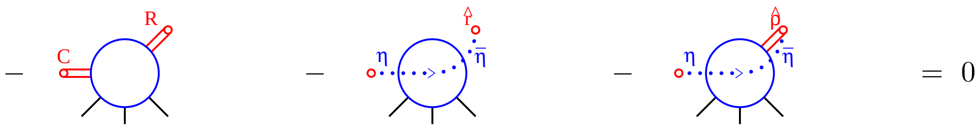
where $\hat{r}$ denotes a field independent derivative term and $\hat{\rho}$ is the field dependent part and higher order in g. The effect on a time–ordered Green function may be represented by

$$\dot{G} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\,(G_{\mathcal{L}'} - G_{\mathcal{L}})$$

and is graphically represented by

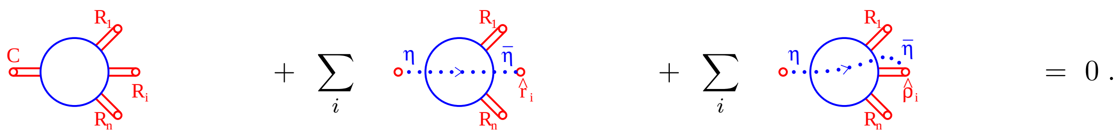


Opening up the extra vertex, taking into account the factors -1 when opening a FP loop, we have

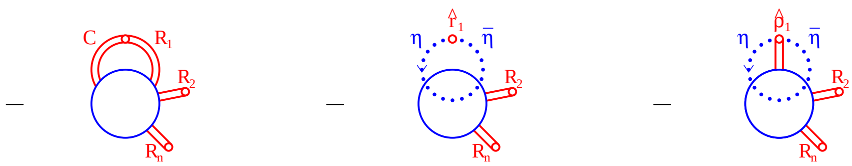


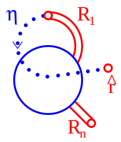
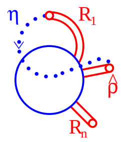
by the ST identity if all other legs are physical (on–shell).

Next we consider th generalized ST–identities:



For the dimensionally regularized theory in $d = 4 - \varepsilon$ dimensions we have:



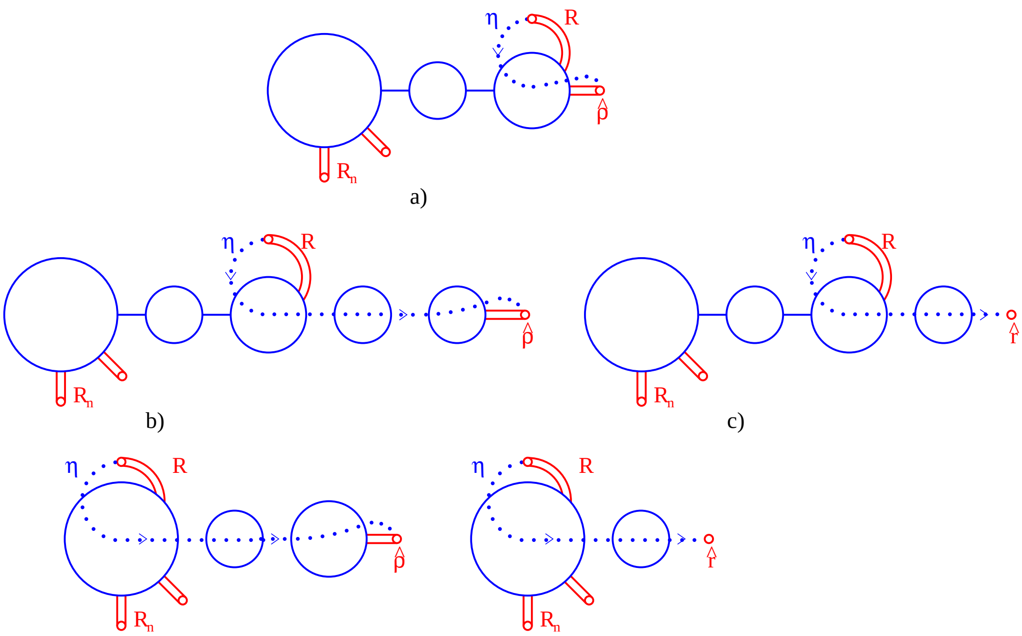
$-$  $-$  $+ \;\cdots\; = \; 0\,.$

Now the S–matrix elements we obtain when we take $R_2\,, \cdots\, R_n$ to be physical fields and we put them on–shell: $(p^2 - m^2)\, G(\cdots)|_{p^2 \to m^2}$. What happens is the following:

i) composite external vertices have no poles and thus do not contribute on–shell;

ii) except for the Feynman gauge, where the $\eta^{\pm}$ ghosts have mass M_W and the neutral η ghost has mass M_Z we have $m_{\text{ghost}} \neq m_{\text{physical}}$ and the ghost amplitudes do not contribute.

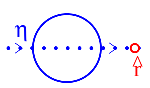
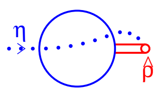
For a more detailed insight, which also covers what happens in the Feynman gauge, we need to consider the decomposition into the 1PI parts



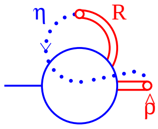
If $\xi \neq 1$ we have $m_\eta \neq m_{\text{phys}}$ such that only diagram a) is contributing. In general, the condition for a field (source) to be physical reads

$$
\begin{aligned}
&\frac{1}{\xi}\,\langle 0|T\{C_C(x)\, G_{A_1}(x_1)\,\cdots\, G_{A_n}(x_n)\,|0\rangle \\
=\;& \sum_i \langle 0|T\{G_{A_1}(x_1)\,\cdots\,\bar{\eta}_C\;(D_{A_iB}\eta_B)\;(x_i)\,\cdots\, G_{A_n}(x_n)\,|0\rangle = 0
\end{aligned}
$$

if $G_{A_i}(x_i)$ is taken on–shell, i.e.

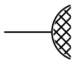
 $+$  $= \; 0\,,$

such that, b) + c) =0, as well as, d) + e) =0, if the corresponding physical field is on–shell. In contrast, diagram a) affects the matrix element: it corresponds to a change in the wave function renormalization

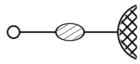


$$= \frac{\delta Z}{Z}\,.$$

Here a detailed consideration of the **LSZ–asymptotic condition** is needed in order to get a quantitative statement. The S–matrix element is given by (for simplicity we consider a physical scalar)



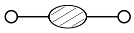
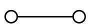
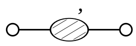
$$= \lim_{p^2 \to m^2} Z\left(p^2 - m^2\right)$$



by amputation of the normalized full propagator (extracting the 1PI part): we thus look at

$$Z\,\frac{p^2 - m^2}{p^2 - m_0^2 + \Pi(p^2)}\Bigg|_{p^2 \to m^2}$$

after mass renormalization: $m_0^2 = m^2 + \delta m^2$ with $\frac{\delta e}{e}; am^2 = \Pi(p^2 = m^2)$. The residue of the pole of the full propagator

 $=$  $+$ 

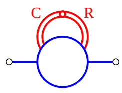
may be read off after the mass renormalization from

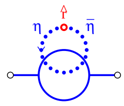
$$\begin{aligned} \frac{1}{p^2 - m_0^2 + \Pi(p^2)} &= \frac{1}{(p^2 - m^2)\left(1 + \frac{\Pi(p^2) - \Pi(m^2)}{p^2 - m^2}\right)} \\ &= \frac{1}{p^2 - m^2}\,\frac{1}{Z^2}\,, \end{aligned}$$

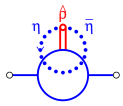
and is given by

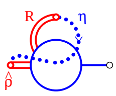
$$\frac{1}{Z^2} = \frac{1}{1 + \frac{\mathrm{d}\Pi}{\mathrm{d}p^2}(m^2)} = \frac{1}{1 + F}\,.$$

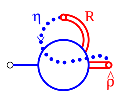
A change of gauge is graphically represented by

$\dot{G}^{(2)} \quad = \quad -$ 

$+$ 

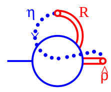
$+$ 

$+$ 

$+$ 

and the residue is changing by

$$\frac{2\alpha}{1 + F} \quad \text{with} \quad \alpha \quad =$$



.

Since

$$\delta\left(\frac{1}{1+F}\right) = \frac{2\alpha}{1+F} = \delta\left(\frac{1}{Z^2}\right) = -\frac{2}{Z^2}\,\frac{\delta Z}{Z}$$

we obtain

$$\alpha = -\frac{\delta Z}{Z}!$$

This proves that gauge invariance and unitarity of the S–matrix require particular wave–function renormalization, namely the one required by the LSZ asymptotic condition. Often it is argued that the ST identities require that under renormalization the gauge symmetry multiplets, like a lepton doublet, for example, only allow for one overall renormalization constant. In the case of the electroweak SM, where the gauge symmetry is broken by the Higgs mechanism, it is definitely so that physical fields have to be treated individually in their renormalization in order to obtain the correct S–matrix elements. The distortion of the classical form of the ST identities by individual field renormalization does not spoil those identities, but just let them look like somewhat more complicated, i.e. their form looks slightly different for different fields in a given multiplet. The path integral formalism allows this to be controlled quite easily. As the value of the integral does not depend on the choice of the integration variables (the fields) means that the choice of the wave function renormalizations only affects the external vertices, while inside diagrams effects from vertices and propagators cancel. Therefore, $\overline{MS}$ wave function renormalization as an *intermediate renormalization* procedure removing the UV singularities is the most adequate thing to do. The proper S–matrix elements are then obtained by a finite renormalization by imposing the asymptotic condition. Infrared singularities showing up in this case for the charged states have to be treaded in the usual way by the Bloch-Nordsieck prescription (Bloch and Nordsieck 1937) or improvements of it (Yennie, Frautschi and Suura 1961). Also in QED, gauge invariance of the on–shell S–matrix elements is ony obtained after gauge dependent LSZ renormalization condition is imposed (Białynicki-Birula, 1970). Bare on-shell matrix elements are unphysical and dependent on the gauge.

E Solved Problems

E.1 Exercises: Section 1

① Calculate the conversion factors using the values for c, $\hbar$ and ϵ_0 in standard units.

✎

The conversion factors fom `CGS` *to natural units (1.4) directly follow by assigning the value unity to the l.h.s. of (1.3) ans solving for* `cm`, `sec`, `gr`, *and* `Cb`. *In order to get* `gr` *we need the conversion of energy to erg = gr cm*^{2}*/sec*^{2} *as given by (1.1).* ✌

② The width of the Z boson has been measured at LEP (October 2001) to be

$$\Gamma_Z = (2.4952 \pm 0.0023)\ \mathrm{GeV}.$$

Calculate the Z lifetime $\tau_Z = \Gamma_Z^{-1}$ in seconds.

✎

The Z lifetime is

$$\begin{aligned} \tau_Z &= \frac{1}{2.4952 \pm 0.0023}\ GeV^{-1} = \frac{1}{(2.4952 \pm 0.0023)\ \times 1.5193} \times 10^{-24}\ \mathrm{sec} \\ &= (2.638 \pm 0.002)\ \times 10^{-23}\ \mathrm{sec}\ . \end{aligned}$$

This very short decay time is characteristic for unstable particles like the heavy massive gauge bosons W and Z which have widths of about 2.5 GeV. The heaviest known "particle", the top quark, with a mass of about 175 GeV has a width of about $\Gamma_t \simeq 0.1$ *GeV and thus lives about 10 times longer.* ✌

What distance (in mm) does a Z particle travel before it decays (length of track in the detector) ? Hint: Use the velocity v (in units of c), which is determined by the magnitude of the momentum $|\vec{p}\,| = \frac{vM_Z}{\sqrt{1-v^2}} = \sqrt{E_{CM}^2 - M_Z^2}$. The distance traveled in the laboratory frame is then given by (Lorentz contraction!)

$$\ell_Z(E_b) = \frac{v}{\sqrt{1-v^2}} c\tau_Z \simeq \sqrt{\frac{E_{CM}^2}{M_Z^2} - 1}\ \times\ 7.9 \times 10^{-14}\ \ \mathrm{mm}\ \ .$$

The experimental value for the Z mass is

$$M_Z = (91.1875 \pm 0.0021)\ \mathrm{GeV}.$$

The Z is produced as a real (though unstable) particle provided $E_{CM} > M_Z$. Consider typical LEP energies $E_{CM} = M_Z + n\Gamma_Z$ for $n = 1, 2, 5$.

✎

We note that $\Gamma_Z \ll M_Z$ *and hence*

$$\sqrt{\frac{E_{CM}^2}{M_Z^2} - 1} = \sqrt{(1 + n\,\frac{\Gamma_Z}{M_Z})^2 - 1} \simeq \sqrt{2n\Gamma_Z/M_Z} \simeq \sqrt{n} \times 0.234$$

and thus

$$\ell_Z(E_b) = (1.85,\ 2.61\ ,4.13)\ \times 10^{-14}\ \text{mm} \quad \text{for} \ \ n = (1,\ 2,\ 5)\ .$$

These numbers tell us that the traveling distance of a heavy gauge boson is by far too small to be resolved by any particle detector. Typical vertex detectors have resolutions of about 10 μm and allow vertex tagging for B mesons, for example.

Use the Boltzmann constant $k = 8.6173 \cdot 10^{-5}$ eV °K^{-1} to evaluate the equivalent "temperature of a Z event" at LEP.

$$T = \frac{E}{k} = \frac{(91.1875 \pm 0.0021) \cdot 10^9}{8.6173 \cdot 10^{-5}}\ {}^\circ\text{K} \simeq 1.058 \cdot 10^{15}\ {}^\circ\text{K} \simeq 1.856 \times 10^{11} \times T_\odot$$

where $T_\odot \sim 5700$ °K is the temperature at the surface of the sun.

In nature such temperatures must have existed in our universe shortly after the big bang. In the early universe the time–temperature relationship (in the radiation dominated era) is given by

$$t = \frac{2.42}{\sqrt{N(T)}} \left(\frac{1\text{MeV}}{kT}\right)^2 \ sec\ .$$

where

$$N(T) = \sum_B g_B + \frac{7}{8} \sum_F g_F$$

counts the effectively massless ($m_i \ll kT$) degrees of freedom of bosons and fermions. Calculate at what time in the history of the universe the temperature of the universe was equivalent to the mass of the Z boson.

At $kT \sim M_Z$ $N(T) \sim g_\gamma + 7/8\,(3g_\nu + 2g_l + 3\times3\times g_q) = 2 + 7/8\,(3\times2 + 2\times4 + 3\times3\times4) = 183/4$ (counting the photon, the three neutrinos, the electron, the muon and the three lightest quarks of color multiplicity three as effectively massless). g_i are the numbers of degrees of freedom per particle species i (helicity and particle–antiparticle counting). Thus we find the above temperature at times

$$t \sim 4 \times 10^{-11} \quad \text{sec after the big bang.}$$

Thus LEP experiments probe the state of matter in this age of the history of our universe.

③ The QED cross section for μ–pair production in e^+e^- annihilation at high energies ($s \gg m_\mu^2$) is given by

$$\frac{d\sigma}{d\Omega}\,(e^+e^- \to \mu^+\mu^-) = \frac{\alpha^2}{8E_b^2}\frac{1+\cos^2\theta}{2},$$

where E_b is the e^-–beam energy and θ the μ^- production angle $\sphericalangle(e^-,\mu^-)$ in the center of mass frame. Calculate σ_{total} in cm^2 for E$_b$=1 GeV. What is the event rate if the beam luminosity is $L = 10^{32}$ cm^{-2}sec^{-1}?. The luminosity measures the incoming flux of particles per cm^2 and per second.

As $d\Omega = d\varphi\, d(\cos\theta)$ we obtain

$$\begin{aligned}\sigma_{\mathrm{tot}} &= \int_0^{2\pi} d\varphi \int_{-1}^{+1} d(\cos\theta)\,\frac{d\sigma}{d\Omega} = \frac{\pi}{3}\frac{\alpha^2}{E_b^2} \\ &= 5.576\times 10^{-5}\,GeV^{-2} = 2.171\times 10^{-32}\ \mathrm{cm}^2\end{aligned}$$

and the number of events per unit time at the above luminosity is

$$\frac{dN}{dt} = 2.171\ \mathrm{sec}^{-1}$$

Note that we need high incoming particle fluxes in order to obtain reasonable counting rates. We also mention that the above form of the cross section is typical in a renormalizable quantum field theory which exhibits dimensionless coupling constants and masses which become negligible at high energies ($m_i \ll E_b$). By dimensional counting the cross section must drop like $1/E_b^2$ (at higher orders in a perturbative expansion up to powers of logarithms $\ln(E_b^2/m_i^2)$. This does not mean that physics becomes uninteresting at higher energies. In the past when going to higher energies thresholds of new particles have been found again and again, the last in 1995 when the top quark was discovered at Fermilab with a mass of about 175 GeV. The new degrees of freedom add to the coefficient in front of the asymptotic $1/E_b^2$ behavior. Thus going to higher energies the "canonical" droping of the cross section is counteracted by the opening of new channels. Nonetheless, increasing luminosity is mandatory to investigate details of the branching structures of an increasing number of channels at higher energies.

④ Range of interactions : the range of a field and the mass m of the corresponding field quantum are related by the Compton wave length

$$r_0 = \frac{\hbar}{mc},$$

where r_0 appears in the static potential (Yukawa)

$$\Phi(r) \propto \frac{e^{-r/r_0}}{r}.$$

Calculate the range of the strong, the weak and the electromagnetic interaction in `cm` under the assumption that the interactions are mediated by exchange of a pion ($m_\pi = 135$ MeV), a W boson ($M_W \simeq 80.45 \pm 0.04$ GeV) and a photon ($m_\gamma < 2\cdot 10^{-16}$ eV experimental bound), respectively. In QED, $m_\gamma = 0$. Discuss this limiting case and the role played by Gauss's law. Hint: Look for static, spherically symmetric solutions of the Klein-Gordon (KG) equation.

For static solutions of the KG–equation we have

$$(\Box + m^2)\,\Phi = (-\Delta + m^2)\,\Phi = 0$$

and since, in polar coordinates,

$$\begin{aligned}\Delta\,\Phi &= \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\varphi^2}+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right)\\ &+ \frac{1}{r^2}\frac{\partial}{\partial r}\left(\frac{r^2\partial\Phi}{\partial r}\right)\end{aligned}$$

for spherically symmetric solutions

$$\Delta\Phi(r)=\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}\right)\,\Phi(r)=m^2\,\Phi(r)\;.$$

The solution is the one given above provided $r_0 = 1/m$ (in natural units). Verify this.

The ranges of the different interactions then are:

force	range $r_0 = 1/m$				
strong	$1/m_\pi$	$=$	*1/(0.135 GeV)*	$\simeq$	1.462×10^{-13} *cm*
weak	$1/M_W$	$=$	*1/(80.45 GeV)*	$\simeq$	2.453×10^{-16} *cm*
electromagnetic	$1/m_\gamma$	$>$	*1/(*2×10^{-25} *GeV)*	$\simeq$	0.987×10^{6} *km*

Electromagnetic gauge invariance requires the photon to be strictly massless, i.e., $\Delta\Phi = 0$ and hence $\Phi \sim 1/r$ has infinite range.

Note that in case of the strong interaction we took the pion as carrier of the short ranged strong force. We know that in "reality" the strong interactions between hadrons (strongly interacting particles) derives from quantum chromodynamics (QCD). The latter is the fundamental theory of strong interactions with quarks as the matter fields and massless gluons as the force carriers (gauge fields). Why does the above arguments about the range of interaction fail for gluons?

Gauss's law

$$\oint_\Sigma \vec{E}\cdot d\vec{S}=\int_V \vec{\nabla}\cdot\vec{E}\,d^3x$$

applied to the electric field $\vec{E} = -\vec{\nabla}\Phi$ implies that the electric flux integrated over a closed surface Σ (surface element $d\vec{S}$)

$$Q=\frac{1}{4\pi}\oint \vec{E}\cdot d\vec{\Sigma}$$

yields the net charge $Q = \sum_i q_i$ which is enclosed by the surface. The l.h.s is the spacial volume integral over the charge density $\vec{\nabla}\cdot\vec{E} = -\Delta\Phi = \rho$. This is true only iff the potential is of Coulomb type. For a potential of finite range the surface integral vanishes in any case in the limit of large volumes (charge screening). ✌

E.2 Exercises: Section 2

① Determine the matrix Λ^μ_ν for

a) a rotation by an angle φ about the z-axis

b) a special Lorentz transformation of velocity $\vec{v}$ in z-direction.

Write down the operators $U(\Lambda)$ for the above transformations.

✎

a) Rotation by an angle φ about the z-axis:

A spatial rotation about the z–axis reads: which as a L–transformation takes the form

$$\Lambda^\mu{}_\nu = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & \mathbf{R} & \\ 0 & & & \end{array}\right)$$

with

$$\mathbf{R} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The corresponding operator representing a right rotation about the z–axis reads:

$$U(\Lambda) = e^{-i\varphi J_3}$$

with J_3 the 3rd component of the total angular momentum operator.

On a state space of spin 1 objects we have a 3–dimensional representation of the rotation group with

$$J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(note the Pauli matrix τ_3 in the upper left corner of the matrix). Since for $n = 1, 2, \cdots$

$$J_3^{2n} = J_3^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and $J_3^{2n+1} = J_3$ *the expansion of the exponential*

$$\begin{aligned}
\exp(-i\varphi J_3) &= \sum_n \frac{(-i\varphi)^n}{n!} J_3^n \\
&= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \cos\varphi \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} - i\sin\varphi \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \cos\varphi - i\sin\varphi & 0 & 0 \\ 0 & \cos\varphi + i\sin\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

can be easily worked out with the result given above.

b) Special Lorentz transformation of velocity $\vec{v}$ *in* z*-direction:*

Pure L–transformation in x^3*–direction (hyperbolic or pseudo rotation)*

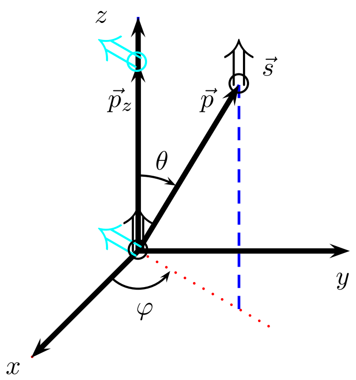
$$\begin{aligned}
x^{0\prime} &= \gamma\,(x^0 - \beta x^3) &&= x^0\cosh\chi - x^3\sinh\chi \\
x^{1\prime} &= x^1 \\
x^{2\prime} &= x^2 \\
x^{3\prime} &= \gamma\,(x^3 - \beta x^0) &&= -x^0\sinh\chi + x^3\cosh\chi
\end{aligned}$$

$$\chi = \mathrm{arctgh}\frac{v}{c}$$

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \cosh\chi & 0 & 0 & -\sinh\chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\chi & 0 & 0 & \cosh\chi \end{pmatrix}$$

Infinitesimal transformation:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & -\chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\chi & 0 & 0 & 0 \end{pmatrix} + O(\chi^2)$$



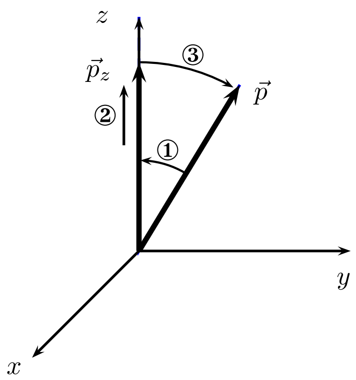


Fig. E.1: *How to get a boost: a particle of momentum $\vec{p}$ and spin $\vec{s}$ may be constructed by*

① *a rotation of $\vec{p}$ into a momentum vector $\vec{p}_z$ pointing into the z–direction (viewed as a rotation in the rest frame),*

② *a boost in z–direction by $\vec{p}_z$ and*

③ *a rotation back to $\vec{p}$.*

A short note on arbitrary boosts:
The special L–transformation $L(p)$ which transforms from a state in the rest frame $(m, \vec{0}\,)$ to a state of momentum p^μ may be written as (verify it)

$$
\begin{aligned}
L^i{}_j &= \delta^i{}_j + \hat{p}_i\hat{p}_j(\cosh\beta - 1) \\
L^i{}_0 &= L^0{}_i = \hat{p}_i \sinh\beta \\
L^0{}_0 &= \cosh\beta
\end{aligned}
$$

with $\hat{\vec{p}} = \vec{p}/|\vec{p}|$, $\cosh\beta = \omega/m$, $\sinh\beta = |\vec{p}|/m$ and $\tanh\beta = |\vec{p}|/\omega = v$ the velocity of the state. The boost operator $U(L(p),0)$ may always be represented by a combination of rotations by an angle $\pm\phi$ about the z–axis, rotations by an angle $\pm\theta$ about the y–axis and a boost in z–direction (see Fig. E.1). The angles are determined by writing the momentum $\vec{p} = |\vec{p}\,|(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ in polar coordinates. If we rotate $\vec{p}$ into the direction of the z–axis then perform a boost along the z–axis and rotate back to the original direction of $\vec{p}$ we obtain

$$U(L(p),0) = U(R_{\varphi,\theta},0)\, U(L_z(|\vec{p}\,|),0)\, U(R^{-1}_{\varphi,\theta},0)$$

with $U(L_z(|\vec{p}|),0) = e^{i\beta K_3}$, $\tanh\beta = |\vec{p}|/p^0 = v$ the velocity and $U(R_{\varphi,\theta},0) = e^{-i\varphi J_3}e^{-i\theta J_2}e^{i\varphi J_3}$. Thus the four momentum eigenstates (2.1) may be constructed by

$$|p, j, j_3\rangle = U(R_{\varphi,\theta},0)\, U(L_z(|\vec{p}\,|),0)\, U(R^{-1}_{\varphi,\theta},0)|m, j, j_3\rangle \ .$$

Note that in first place

$$U(R_{\varphi,\theta},0) = \exp(-i\vec{\omega}\vec{J})$$

where $\vec{\omega}$ is the appropriate rotation vector, i.e., it has modulus θ and rotation axis $\vec{n} = (-\sin\varphi, \cos\varphi, 0)$ in the $x-y$–plane (see Fig. E.1). Show that

$$
\begin{aligned}
U(R_{\varphi,\theta},0) &= e^{-i\theta\,(-\sin\varphi J_1 + \cos\varphi J_2)} \\
&\equiv e^{-i\varphi J_3}e^{-i\theta J_2}e^{i\varphi J_3} \ .
\end{aligned}
$$

The latter representation as a sequence of simple rotations about the y– or the z–axis is most transparent ($e^{-i\varphi J_3}$ is a right rotation about the z–axis and $e^{-i\theta J_2}$ is a right rotation about the y–axis).

② Show that for a general Poincaré transformation $U(\Lambda, a)$ the generators of the Poincaré group satisfy

$$U(\Lambda, a)P_\mu U^{-1}(\Lambda, a) = \Lambda^\nu{}_\mu P_\nu \quad , \tag{E.1}$$

and

$$U(\Lambda, a)M_{\mu\nu}U(\Lambda, a)^{-1} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \left(M_{\rho\sigma} - P_\rho a_\sigma + P_\sigma a_\rho\right) \quad . \tag{E.2}$$

While the first equation tells us that P_μ transforms as a covariant four-vector, the second proves that $M_{\mu\nu}$ is a 2nd rank tensor only with respect to homogeneous Poincaré transformation.

Proof of (E.1):

1. $U^{-1}(\Lambda, a) = U(\Lambda^{-1}, -\Lambda^{-1}a)$

since

$$\begin{aligned} U^{-1}(\Lambda, a)U(\Lambda, a) &= U(\Lambda^{-1}, -\Lambda^{-1}a)U(\Lambda, a) = \\ U(\Lambda^{-1}\Lambda, \Lambda^{-1}a - \Lambda^{-1}a) &= U(1, 0) = \mathbf{1} \end{aligned}$$

2. $U(\Lambda, a)U(1, b)U^{-1}(\Lambda, a) = U(1, \Lambda b)$

since

$$\begin{aligned} U(\Lambda, a)U(1, b)U^{-1}(\Lambda, a) &= U(\Lambda, \Lambda b + a)U^{-1}(\Lambda, a) \\ &= U(\Lambda, \Lambda b + a)U(\Lambda^{-1}, -\Lambda^{-1}a) = U(1, \Lambda b + a - a) = U(1, \Lambda b) \quad . \end{aligned}$$

Expansion for small B yields

$$\begin{aligned} U(1, b) &= \mathbf{1} + ib^\mu P_\mu + O(b^2) \\ U(1, \Lambda b) &= \mathbf{1} + i\Lambda^\mu{}_\nu b^\nu P_\mu + O(b^2) \end{aligned}$$

and hence

$$U(\Lambda, a)\left(\mathbf{1} + ib^\mu P_\mu\right) U^{-1}(\Lambda, a) = \mathbf{1} + i\Lambda^\nu{}_\mu b^\mu P_\nu + O(b^2) \quad .$$

As the unit operator drops out one may factor out ib^μ and take the limit $b \to 0$ in the remaining factor. This then proves the assertion.

Proof of (E.2):

1.

$$U(\Lambda,a)U(\Sigma,0)U^{-1}(\Lambda,a)=U(\Lambda\Sigma,a)U(\Lambda^{-1},-\Lambda^{-1}a)=U(\Lambda\Sigma\Lambda^{-1},a-\Lambda\Sigma\Lambda^{-1}a)$$

2.

$$\begin{aligned}
U(\Sigma,0) &= \mathbf{1}+i\frac{\omega^{\mu\nu}}{2}M_{\mu\nu}+O(\omega^2)\\
U(\Lambda\Sigma\Lambda^{-1},a-\Lambda\Sigma\Lambda^{-1}a) &= \mathbf{1}+i\frac{\Lambda^{\rho}{}_{\mu}\,\Lambda^{\sigma}{}_{\nu}\,\omega^{\mu\nu}}{2}\left(M_{\rho\sigma}-(a_\sigma P_\rho-a_\rho P_\sigma)\right)+O(\omega^2)
\end{aligned}$$

and then proceeding as in the case of P_μ above.
In order to obtain these expansion coefficients a few additional considerations may be helpful:
1) From the basic property of L–transformations

$$\Lambda^{\mu}{}_{\nu}\,\Lambda^{\rho}{}_{\sigma}\,g^{\nu\sigma}=g^{\mu\rho}$$

we get

$$\Lambda^{\mu}{}_{\nu}\,\Lambda^{\rho}{}_{\sigma}\,g^{\nu\sigma}g^{\rho\lambda}=g^{\mu\rho}g^{\rho\lambda}=\delta^{\mu}{}_{\lambda}$$

thus $\Lambda^{\mu}{}_{\nu}\,\Lambda_{\lambda}{}^{\nu}=\delta^{\mu}{}_{\lambda}$ *or* $\Lambda^{\mu}{}_{\nu}\,\left(\Lambda^{-1}\right)^{\nu}{}_{\lambda}=\delta^{\mu}{}_{\lambda}$ *which means*

$$\left(\Lambda^{-1}\right)^{\nu}{}_{\lambda}=\Lambda_{\lambda}{}^{\nu}$$

which means that the inverse of Λ is equal to the transposed matrix.
2) We now may work out

$$\begin{aligned}
\left(\Lambda\Sigma\Lambda^{-1}\right)^{\rho}{}_{\sigma} &= \Lambda^{\rho}{}_{\mu}\,\Sigma^{\mu}{}_{\nu}\,\left(\Lambda^{-1}\right)^{\nu}{}_{\sigma}\\
&= \Lambda^{\rho}{}_{\mu}\,\delta^{\mu}{}_{\nu}\,\left(\Lambda^{-1}\right)^{\nu}{}_{\sigma}+\Lambda^{\rho}{}_{\mu}\,\Lambda_{\sigma}{}^{\nu}\,\omega^{\mu}{}_{\nu}+\cdots\\
&= \delta^{\rho}{}_{\sigma}+\Lambda^{\rho}{}_{\mu}\,\Lambda_{\sigma}{}^{\nu}\,\omega^{\mu}{}_{\nu}+\cdots\quad .
\end{aligned}$$

Thus

$$\left(\Lambda\Sigma\Lambda^{-1}\right)^{\rho\sigma} = g^{\rho\sigma}+\Lambda^{\rho}{}_{\mu}\,\Lambda^{\sigma}{}_{\nu}\,\omega^{\mu\nu}+\cdots$$

and

$$\begin{aligned}
\left(a-\Lambda\Sigma\Lambda^{-1}a\right)^{\rho} &= a^{\rho}-a^{\rho}-\Lambda^{\rho}{}_{\mu}\,\Lambda^{\sigma}{}_{\nu}\,\omega^{\mu\nu}a_\sigma+\cdots\\
&= -\Lambda^{\rho}{}_{\mu}\,\Lambda^{\sigma}{}_{\nu}\,\omega^{\mu\nu}a_\sigma+\cdots
\end{aligned}$$

and finally

$$\Lambda^{\rho}{}_{\mu}\,\Lambda^{\sigma}{}_{\nu}\,\omega^{\mu\nu}a_\sigma P_\rho = \Lambda^{\sigma}{}_{\mu}\,\Lambda^{\rho}{}_{\nu}\,\omega^{\mu\nu}a_\rho P_\sigma=-\Lambda^{\rho}{}_{\mu}\,\Lambda^{\sigma}{}_{\nu}\,\omega^{\mu\nu}a_\rho P_\sigma$$

where we utilized the antisymmetry of $\omega^{\mu\nu}$.

③ Use the previous result to derive the Lie algebra of $\mathcal{P}_+^\uparrow$ by expanding $U(\Lambda, a)$ to first order. The result should be

$$[P_\mu, P_\nu] = 0 \ , \tag{E.3}$$

$$[P_\rho, M_{\mu\nu}] = -i\,(g_{\rho\mu}P_\nu - g_{\rho\nu}P_\mu) \ , \tag{E.4}$$

$$[M_{\mu\nu}, M_{\rho\sigma}] = i\,(g_{\mu\rho}M_{\nu\sigma} - g_{\mu\sigma}M_{\nu\rho} + g_{\nu\sigma}M_{\mu\rho} - g_{\nu\rho}M_{\mu\sigma}) \ . \tag{E.5}$$

Give a physical interpretation.

✎

The commutation relations (E.3) and (E.4) follow directly from (E.1) and (E.2), respectively, by setting $U(\Lambda, a) = \mathbf{1} + ia^\mu P_\mu$. Similarly, (E.5) is obtained from (E.2) by choosing $U(\Lambda, a) = \mathbf{1} + i\frac{\omega^{\mu\nu}}{2}M_{\mu\nu}$. These CR's characterize the Poincaré group locally in a neighborhood of unity. Wanted are the representations of this Lie–algebra, this will yield at the same time the physical interpretation of the elements of the algebra. ✌

④ Prove that

$$\frac{d^3p}{\omega_p} \quad \text{and} \quad \omega_p \delta^{(3)}(\vec{p}\,' - \vec{p}\,) \ \text{ with } \ \omega_p = \sqrt{\vec{p}^{\,2} + m^2}$$

are relativistically invariant. Hint: Use

$$\Theta(p^0)\,\delta(p^2 - m^2)\,d^4p = \frac{d^3p}{2\omega_p} \ .$$

✎

For any function $f(p)$, starting from the manifestly invariant integration measure (l.h.s. of last relation)

$$I\{f\} = \int_{-\infty}^{+\infty} dp^0 \int d^3p\,\Theta(p^0)\,\delta(p^2 - m^2)\,f(p^0, \vec{p}\,) = \int_0^{+\infty} dp^0 \int d^3p\,\delta((p^0)^2 - \omega_p^2)\,f(p^0, \vec{p}\,)$$

and a change of variable $p^0 \to (p^0)^2$ yields $dp^0 = (2p^0)^{-1}d(p^0)^2$ (since $p^0 > 0$ the mapping is unambiguous) and hence

$$I\{f\} = \int_0^{+\infty} \frac{d(p^0)^2}{2p^0} \int d^3p\,\delta((p^0)^2 - \omega_p^2)\,f(p^0, \vec{p}\,)\bigg|_{p^0 = +\sqrt{(p^0)^2}} = \int \frac{d^3p}{2\omega_p}\,f(p^0, \vec{p}\,)\bigg|_{p^0 = \omega_p}$$

we infer that $\frac{d^3p}{\omega_p}$ is an invariant integration measure. As

$$\int \frac{d^3p}{\omega_p}\,\omega_p \delta^{(3)}(\vec{p}\,' - \vec{p}\,) \ \equiv 1$$

also $\omega_p \delta^{(3)}(\vec{p}\,' - \vec{p}\,)$ must be relativistically invariant. ✌

⑤ Show that the Fourier transform of a distribution $\tilde{\phi}(p) = 2\pi\delta(p^2 - m^2)\tilde{\xi}(p)$ with support on the hyperbola $p^2 = m^2$ has the form

$$\phi(x) = \frac{1}{(2\pi)^3}\int \frac{d^3p}{2\omega_p}\left(e^{-ipx}\tilde{\xi}(\omega_p, \vec{p}\,) + e^{ipx}\tilde{\xi}(-\omega_p, -\vec{p}\,)\right)$$

which is a decomposition into a positive and a negative frequency part. Hint: Use

$$\int_{-\infty}^{+\infty} dp^0\, \tilde{f}(p^0, \vec{p}\,) = \int_0^{\infty} dp^0\, \left(\tilde{f}(p^0, \vec{p}\,) + \tilde{f}(-p^0, \vec{p}\,)\right)$$

and

$$\int_{-\infty}^{+\infty} d^3p\, e^{i\vec{p}\vec{x}}\, \tilde{g}(\vec{p}\,) = \int_{-\infty}^{+\infty} d^3p\, e^{-i\vec{p}\vec{x}}\, \tilde{g}(-\vec{p}\,) \quad .$$

Compare the form obtained with the representation of a free field in terms of creation and annihilation operators.

$$\begin{aligned}
\phi(x) &= \frac{1}{(2\pi)^4}\int d^4p e^{-ipx}\, \tilde{\phi}(p) \\
&= \frac{1}{(2\pi)^3}\int_0^{+\infty} dp^0 \int d^3p e^{i\vec{p}\vec{x}}\, \delta((p^0)^2 - \omega_p^2)\left(e^{-ip^0x^0}\tilde{\xi}(p^0, \vec{p}\,) + e^{+ip^0x^0}\tilde{\xi}(-p^0, \vec{p}\,)\right) \\
&= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2\omega_p} e^{i\vec{p}\vec{x}}\, \left(e^{-ip^0x^0}\tilde{\xi}(p^0, \vec{p}\,) + e^{+ip^0x^0}\tilde{\xi}(-p^0, \vec{p}\,)\right)\Big|_{p^0=\omega_p} \\
&= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2\omega_p}\left(e^{-ip^0x^0}e^{i\vec{p}\vec{x}}\tilde{\xi}(p^0, \vec{p}\,) + e^{+ip^0x^0}e^{-i\vec{p}\vec{x}}\tilde{\xi}(-p^0, -\vec{p}\,)\right)\Big|_{p^0=\omega_p}
\end{aligned}$$

which proves the claim.

⑥ Show that

$$\Delta(x) = \frac{i}{(2\pi)^3}\int \frac{d^3p}{2\omega_p}\left(e^{-ipx} - e^{ipx}\right)\Big|_{p_0=\omega_p=\sqrt{\vec{p}^{\,2}+m^2}} = 0 \quad \text{if} \quad x^2 < 0$$

We first check that $\Delta(x)$ solves the Lorentz (L) invariant Klein–Gordon equation: utilizing $\partial_\mu e^{-ipx} = -ip_\mu e^{-ipx}$ and hence $\Box\, e^{-ipx} = g^{\mu\nu}\partial_\mu\partial_\nu e^{-ipx} = -p^2 e^{-ipx}$ and the fact that $p^0 = \omega_p$ such that $p^2 e^{-ipx} = m^2 e^{-ipx}$ we immediately obtain

$$\Box\, \Delta(x) = -m^2\Delta(x)$$

by applying $\Box$ to the integral representation. Since $\Delta(x)$ is L–invariant we most easily evaluate the defining integral for the space–like hyperplane $x^0 = 0$, where $x^2 = -\vec{x}^2 < 0$. The integrand thus reads

$$(e^{i\vec{p}\vec{x}} - e^{-i\vec{p}\vec{x}})$$

which is an odd function in p_i $(i = 1, 2, 3)$ and thus the integral vanishes. By L–invariance it then vanishes for arbitrary space–like hyperplanes. This proves the assertion.

⑦ For a system of free particles and antiparticles the four-momentum operator can be expressed in the simple form

$$P^\mu = \sum_r \int d\mu(p) p^\mu \left\{a^+(\vec{p}, r)a(\vec{p}, r) + b^+(\vec{p}, r)b(\vec{p}, r)\right\} \quad .$$

Show that P^μ has the properties

$$\begin{aligned} P^\mu \mid 0 > &= 0 \\ P^\mu \mid \vec{p}, r, \alpha > &= p^\mu \mid \vec{p}, r, \alpha > \end{aligned}$$

and satisfies the commutation relations

$$\begin{aligned} [P^\mu, a \;\; (\vec{p}, r)] &= -p^\mu a(\vec{p}, r) \\ [P^\mu, a^+(\vec{p}, r)] &= +p^\mu a^+(\vec{p}, r) \quad \text{etc.} \end{aligned}$$

Give a physical interpretation of these properties.

✎

We first verify the last of the above relations: look at

$$P^\mu a^+(\vec{p}, r) = \sum_{r'} \int d\mu(p') p'^\mu \left\{a^+(\vec{p}\,', r')a(\vec{p}\,', r') + b^+(\vec{p}\,', r')b(\vec{p}\,', r')\right\} \; a^+(\vec{p}, r)$$

and use the (anti–)commutation relations of the creation and annihilation operators to (anti–)commute $a^+(\vec{p}, r)$ to the left of P^μ. What one obtains is $a^+(\vec{p}, r)P^\mu$ plus one term from the only non-vanishing (anti–)commutator $[a(\vec{p}', r'), a^+(\vec{p}, r)]_\pm = \delta_{r'r}(2\pi)^3\, 2\omega_p \delta^{(3)}(\vec{p}\,' - \vec{p}\,)$ which yields

$$\sum_r^{\prime} \int d\mu(p') p'^\mu \; a^+(\vec{p}\,', r')\; \delta_{r'r}(2\pi)^3\; 2\omega_p \delta^{(3)}(\vec{p}\,' - \vec{p}\,) = p^\mu \; a^+(\vec{p}, r)$$

with $p^0 = \omega_p$. Thus $[P_\mu, p^\mu\; a^+(\vec{p}, r)] = p^\mu\; a^+(\vec{p}, r)$. Taking the Hermitean conjugate we immediately get $[P_\mu, p^\mu\; a(\vec{p}, r)] = -p^\mu\; a(\vec{p}, r)$. Corresponding relations follow for b^+ and b.

The physical interpretation follows by verifying that P^μ indeed has the properties of the relativistic four–momentum operator: As each of the two terms in the representation of P^μ exhibits an annihilation operator to the right $P^\mu \mid 0 >= 0$ is trivially satisfied. Thus the vacuum is an eigenstate of P^μ with eigenvalue zero. Let $[P_\mu, p^\mu\; a^+(\vec{p}, r)] = p^\mu\; a^+(\vec{p}, r)$ act onto the vacuum to the right: $[P_\mu, p^\mu\; a^+(\vec{p}, r)] \mid 0 >= P_\mu, p^\mu\; a^+(\vec{p}, r) \mid 0 >= p^\mu\; a^+(\vec{p}, r) \mid 0 >$ which reads $P^\mu \mid \vec{p}, r >= p^\mu \mid \vec{p}, r >$. Thus the one particle states $\mid \vec{p}, r >$ are indeed eigenstates of P^μ with eigenvalues the relativistic four–momentum. This shows that P^μ acts as a four–momentum operator on the one particle states. One still has to show that P^μ acts on multi–particle states appropriately, which means the the states (2.15) are eigenstates of P^μ with eigenvalue the total four–momentum $p^\mu = \sum_{i=1}^n p_i^\mu$. In fact our considerations of the one particle states may be easily extended to the multi particle states. ✌

⑧ For a Dirac field the charge operator is given by

$$Q = \int d^3x\, j^0(x) = \int d^3x \;\, : \psi_\alpha^+(x)\, \psi_\alpha(x) :$$

where

$$j^\mu = \bar{\psi}\gamma^\mu\psi \ , \quad \partial_\mu j^\mu = 0$$

is the conserved electromagnetic current operator. Show that

$$Q = \sum_r \int d\mu(p) \ \{a^+(\vec{p},r)a(\vec{p},r) - b^+(\vec{p},r)b(\vec{p},r)\}$$

is time independent. Give a physical interpretation of Q by means of the commutation relations

$$[Q,a] = -a \ , \ \ [Q,b^+] = -b^+ \ , \ \ [Q,\psi] = -\psi \ \ .$$

⑨ Prove that for a Dirac particle

$$\begin{aligned}(\Lambda_+)_{\alpha\beta} &= \frac{1}{2m}\sum_r u_\alpha(p,r)\bar{u}_\beta(p,r) = \frac{1}{2m}(\not{p}+m)_{\alpha\beta} \\ (\Lambda_-)_{\alpha\beta} &= -\frac{1}{2m}\sum_r v_\alpha(p,r)\bar{v}_\beta(p,r) = -\frac{1}{2m}(\not{p}-m)_{\alpha\beta}\end{aligned}$$

are projection operators with property

$$\begin{aligned}\Lambda_+ u(p,s) = u(p,s) \quad &, \quad \Lambda_+ v(p,s) = 0 \\ \Lambda_- u(p,s) = 0 \qquad\quad &, \quad \Lambda_- v(p,s) = v(p,s) \quad .\end{aligned}$$

Give a physical interpretation of this result.

Note: In the space of four-spinors the usual Hermitean conjugation is replaced by going to the adjoint

$$\Gamma \to \Gamma^\dagger = \gamma^0\Gamma^+\gamma^0 \ \ .$$

Thus, the usual Hermitecity $\Gamma = \Gamma^+$ requirement is replaced be self–adjointness requirement $\Gamma = \Gamma^\dagger$, because the L–invariant scalar product between two spinors u and v is $\bar{u}v \equiv u^+\gamma^0 v$, and not u^+v. The latter is not L–invariant.

⑩ Prove that

$$\Pi_\pm = \frac{1}{2}\left(1 \pm \gamma_5 \not{n}\right)$$

for n a space-like vector orthogonal to the momentum p of a Dirac particle

$$n^2 = -1 \ \ ; \quad n \cdot p = 0$$

are projection operators with the property

$$\begin{aligned}\Pi_\pm u(p,s) &= u(p,s)\,\delta_{s,\pm} \\ \Pi_\pm v(p,s) &= v(p,s)\,\delta_{s,\pm} \quad .\end{aligned}$$

Give a physical interpretation of the latter properties.

E.3 Exercises: Section 4

① Show that the Maxwell equation $\partial_\mu F^{\mu\nu} = 0$ as a field equation for the vector potential takes the form

$$\Box\, A^\nu(x) - \partial^\nu \left(\partial_\mu A^\mu(x)\right) = 0 \ .$$

Show that $A^\nu(x)$ is not determined by this equation because the operator $\Box\, g^{\mu\nu} - \partial^\mu \partial^\nu$ has no inverse. Hint.: $\varphi_\mu = \partial_\mu \alpha(x)$, $\alpha(x)$ an arbitrary scalar function, is a solution of the above equation with eigenvalue 0.

✎

Maxwell's equation $\partial_\mu F^{\mu\nu} = 0$ with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ reads

$$\partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = 0$$

or by appropriate interchange of derivatives, relabeling of indices and using $\partial_\mu \partial^\mu = \Box$, $\partial_\nu A^\nu \equiv \partial^\nu A_\nu$ and $A^\mu = g^{\mu\nu} A_\nu$ we obtain

$$\left(\Box\, g^{\mu\nu} - \partial^\mu \partial^\nu\right)\, A_\nu = 0$$

q.e.d.

Next consider

$$\begin{aligned} \left(\Box\, g^{\mu\nu} - \partial^\mu \partial^\nu\right)\, \partial_\nu \alpha(x) &= \Box\, \partial^\mu \alpha - \partial^\mu \Box\, \alpha \\ &= \Box\, \partial^\mu \alpha - \Box\, \partial^\mu \alpha \\ &\equiv 0 \end{aligned}$$

which vanishes identically for any differentiable scalar function $\alpha(x)$. In Fourier space $\partial_\mu \to -ip_\mu$ we have

$$\left(-p^2 g^{\mu\nu} + p^\mu p^\nu\right)\, \tilde{A}_\nu(p) = 0$$

and we see that p_ν is a solution of this equation with eigenvalue 0, i.e., the (c–number) operator $(-p^2 g^{\mu\nu} + p^\mu p^\nu)$ has no inverse.

✌

② Show that for a massive spin 1 field the Proca equation

$$\left(\Box\, + m^2\right)\, A^\nu(x) - \partial^\nu \left(\partial_\mu A^\mu(x)\right) = 0$$

implies $\partial_\mu A^\mu(x) \equiv 0$ automatically. Comment on the number of degrees of freedom ! Show that the Proca equation is the Euler-Lagrange equation of the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{m^2}{2} A_\mu A^\mu \ \ ; \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ \ .$$

Discuss the invariance properties of $\mathcal{L}$ under gauge transformations.

✎

Taking the derivative ∂_ν of the Proca equation

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

we get

$$\left(\Box + m^2\right)\,\partial_\nu A^\nu - \Box\,\partial_\mu A^\mu = m^2 \partial_\nu A^\nu = 0$$

such that for $m \neq 0$ indeed $\partial_\nu A^\nu \equiv 0$. Thus the massive spin one field A^ν as a solution of the Proca equation has automatically three degrees of freedom only, which is the correct number as the field $j = 1,\ j_3 = \pm 1, 0$ should describe two transversal and one longitudinal degree of freedom.

The naive guess to require a massive spin one field to satisfy the Klein–Gordon equation

$$\left(\Box + m^2\right)\,A^\nu(x) = 0$$

would lead to unphysical degrees of freedom, namely, a non-vanishing scalar component $\partial_\nu A^\nu(x) \neq 0$.

The propagator (Green's function) is the inverse of the Proca differential operator

$$\left(\left(\Box + m^2\right) g^{\mu\nu} - \partial^\mu \partial^\nu\right)\,A_\nu(x) = 0\ ,$$

i.e.,

$$\left(\left(\Box + m^2\right) g^{\mu\nu} - \partial^\mu \partial^\nu\right)\,D_{\nu\rho}(x) = \delta^\mu{}_\rho\,\delta^{(4)}(x)\ .$$

In momentum space the differential operator becomes a c–number expression which reads

$$\left(\left(-p^2 + m^2\right) g^{\mu\nu} + p^\mu p^\nu\right)\,\tilde{D}_{\nu\rho}(p) = \delta^\mu{}_\rho$$

with $\tilde{D}_{\nu\rho}(p)$ a symmetric 2nd tensor which must have the form $\tilde{D}_{\nu\rho}(p) = A(p^2)\,g_{\nu\rho} + B(p^2)\,p_\nu p_\rho$. One easily determines A and B, with the result $A = 1/(m^2 - p^2)$ and $B = -1/(m^2(m^2 - p^2))$, and hence

$$\tilde{D}^{\mu\nu}(p) = -\left(g^{\mu\nu} - \frac{p^\mu p^\nu}{m^2}\right)\,\frac{1}{p^2 - m^2 + i\varepsilon}\ .$$

Note that the Proca propagator is transverse $p_\nu \tilde{D}^{\mu\nu}(p) = 0$ and $p_\mu \tilde{D}^{\mu\nu}(p) = 0$ on the mass-shell $p^2 = m^2$ while this would not be the case for the "Klein–Gordon propagator"

$$\tilde{D}^{\mu\nu}(p) = \frac{-g^{\mu\nu}}{p^2 - m^2 + i\varepsilon}\ .$$

Now we consider the Euler–Lagrange equations for the Proca Lagrangian. Hint: write

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4} g^{\alpha\rho} g^{\beta\sigma} F_{\alpha\beta} F_{\rho\sigma} \\ &= -\frac{1}{4} g^{\alpha\rho} g^{\beta\sigma}\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right)\left(\partial_\rho A_\sigma - \partial_\sigma A_\rho\right)\end{aligned}$$

to calculate

$$\frac{\partial \mathcal{L}}{\partial \partial_\mu A_\nu}$$

which gives

$$\frac{\partial \mathcal{L}}{\partial \partial_\mu A_\nu} = -F_{\mu\nu} \quad .$$

One easily evaluates

$$\frac{\mathcal{L}}{\partial A_\nu} = m^2 A_\nu$$

such that the Euler–Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \partial_\mu A_\nu} = \frac{\mathcal{L}}{\partial A_\nu}$$

reads

$$-\partial^\mu F_{\mu\nu} = m^2 A_\nu \quad .$$

After rewriting $F_{\mu\nu}$ as the curl of A_ν one easily identifies the last equation as the Proca equation.

What about gauge invariance? In the Proca Lagrangian the first term proportional to $F_{\mu\nu}F^{\mu\nu}$ is invariant under local transformations $A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x)$ while the mass term proportional to $A_\mu A^\mu$ evidently is not. Is this a problem? For the free massive spin 1 field it is not a problem because the Proca field describes the correct number of physical degrees of freedom from the very beginning. In the massless case (see Sec. 4) invariance under local gauge transformations was a necessary tool to disentangle the two transversal physical degrees of freedom from ghosts inherent in the four component gauge potential $A_\mu(x)$. However, for the case of interacting spin one fields the physical "Proca gauge" or unitary gauge will lead to problems with manifest renormalizability. This is an important subject which will be discussed later on.

③ Prove that under local gauge transformations

$$\psi \to e^{-ie\alpha(x)}\psi \ , \quad A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x)$$

the covariant derivative $D_\mu = \partial_\mu - ieA_\mu$ has the property: $D_\mu \psi$ transforms identical to ψ and $\bar{\psi}\Gamma D_\mu \psi$ is gauge invariant provided D_μ commutes with the 4 by 4 matrix Γ.

By direct calculation we find

$$\begin{array}{rcl} D_\mu \psi & = & (\partial_\mu - ieA_\mu)\ \psi \\ & \to & (\partial_\mu - ie(A_\mu - \partial_\mu \alpha(x)))\ e^{-ie\alpha(x)}\psi \\ & = & e^{-ie\alpha(x)} e^{+ie\alpha(x)}\ (\partial_\mu - ieA_\mu + ie\partial_\mu \alpha(x))\ e^{-ie\alpha(x)}\psi \\ & = & e^{-ie\alpha(x)}\ (\partial_\mu - ieA_\mu)\ \psi \end{array}$$

where we used $\partial_\mu e^{-ie\alpha(x)} = e^{-ie\alpha(x)}\partial_\mu - ie(\partial_\mu \alpha)(x)\, e^{-ie\alpha(x)}$. Thus

$$D_\mu e^{-ie\alpha(x)} = e^{-ie\alpha(x)} D_\mu \quad ,$$

i.e.,

$$e^{+ie\alpha(x)} D_\mu e^{-ie\alpha(x)} = D_\mu$$

q.e.d.

Gauge invariance of

$$\begin{aligned} \bar{\psi}\Gamma D_\mu \psi &\rightarrow \bar{\psi} e^{+ie\alpha} \Gamma e^{-ie\alpha} D_\mu \psi \\ &= \bar{\psi}\Gamma D_\mu \psi \end{aligned}$$

provided the group transformation commutes with Γ*:*

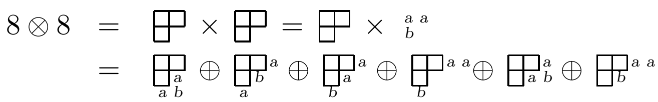
$$e^{+ie\alpha}\Gamma = \Gamma e^{+ie\alpha}$$

which for tha Abelian group is trivially satisfied for any 4 by 4 matrix in spinor space.

✌

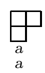
E.4 Exercises: Section 5

① Show that $8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus 10^* \oplus 27$.

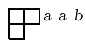


In order to append to the first octet the second one in all admissible ways which respect the (anti-) symmetrization, we replace the second one by letters with identical letters in the rows (symmetrized). The different elements of a column if they appear in the same column of the enlarged tableau must appear in the old order (anti-symmetrized). Hence append the elements from the first row in all possible ways to the first tableau, then to the such enlarged ones the elements of the second row etc. For our case one easily reads off the result.

Note that the tableau



is not allowed. Also



is not allowed, since baa is not an admissible sequence of letters.

② Discuss the iso-spin properties of the triplet of pions (π^+, π^0, π^-) .

The iso-spin symmetry of the scattering operator S not only leads to relations between matrix elements but also to selection rules: Suppose

(a) T is a generator of a symmetry transformation such that $[T, S] = 0$,

(b) $\mid \alpha >$ and $\mid \beta >$ are eigenstates of T i.e. $T \mid \alpha >= t_\alpha \mid \alpha >, T \mid \beta >= t_\beta \mid \beta >$

What does this imply for the S-matrix elements

$$S_{\beta\alpha} =< \beta \mid S \mid \alpha > \ ?$$

Find a few examples.

③ Use the Young tableaus to construct the meson states in

$$3 \otimes 3^*$$

and the baryon states in

$$3 \otimes 3 \otimes 3\,.$$

The states in the pseudo-scalar meson octet of flavor $SU(3)$ are characterized by the 3^{rd} component of iso-spin and by hypercharge $Y = B + S$ (B baryon number $B = 0$ for mesons, S strangeness $S = 0$ for pions). Display the weight diagram ($I_3 - Y$ plot) of the meson states. How are they composed of u, d and s quarks in the $SU(3)_{\text{flavor}}$ quark model ?

✎

✌

④ The structure constants c_{ikl} of a Lie-algebra $[T_i, T_k] = i c_{ikl} T_l$ satisfy the Jacobi identity.

$$c_{ikn} c_{nlm} + \text{ terms cyclic in } (ikl) = 0$$

Use this to show that $(\tilde{T}_i)_{kl} = -i c_{ikl}$ also satisfies the Lie-algebra (adjoint representation).

✎

✌

⑤ Lepton number L_e is another additive quantum number which is strictly conserved. $L_e(e^-) = 1$ by convention. Determine L_e for the other particles from the observed reactions:

1. $L_e(e^+) = -1,\ L_e(\gamma) = 0$:

$$\begin{aligned} p + e &\rightarrow p + e + \gamma \\ \gamma^* &\rightarrow e^+ + e^- \end{aligned}$$

2. $L_e(\pi^0) = L_e(\pi^\pm) = 0$:

$$\begin{aligned} \pi^0 &\rightarrow 2\gamma,\ \gamma + e^+ + e^- \\ p + \pi^- &\rightarrow n + \pi^0 \\ p + \pi^0 &\rightarrow n + \pi^+ \end{aligned}$$

3. $L_e(\bar{\nu}_e) = -1,\ L_e(\nu_e) = 1$:

$$\begin{aligned} \pi^- &\rightarrow e^- + \bar{\nu}_e \\ \pi^+ &\rightarrow e^+ + \nu_e \end{aligned}$$

From the last two reactions we learn the important result $\nu_e \neq \bar{\nu}_e$!

✎

✌

⑥ Baryon number conservation is responsible for the stability of the proton. By convention $B(p) = 1,\ B(e^-) = 0$. Determine the baryon numbers of particles from the observation of the following reactions:

a.) Baryons and mesons:

1. $B(\pi^0) = 0$:

$$p + p \quad \rightarrow \quad p + p + \pi^0$$

2. $B(n) = B(p),\ B(\pi^\pm) = B(\pi^0) = 0$:

$$\begin{array}{lcl} p + p & \rightarrow & p + n + \pi^+ \\ \pi^- + p & \rightarrow & n + \pi^0 \end{array}$$

3. $B(K^\pm) = B(K^0) = 0$:

$$\begin{array}{lcl} K^\pm & \rightarrow & \pi^\pm + \pi^0 \\ K^0 & \rightarrow & \pi^+ + \pi^-,\ \pi^+ + \pi^- + \pi^0 \end{array}$$

4. $B(\Lambda),\ B(\Sigma) = 1$:

$$\begin{array}{lcl} \pi^- + p & \rightarrow & \Lambda^0 + K^0,\ \Sigma^- + K^+ \\ \pi^+ + p & \rightarrow & \Sigma^+ + K^+,\ \Sigma^0 + \Lambda^0 \end{array}$$

5. $B(\Xi),\ B(\Omega^-) = 1$:

$$K^- + p \quad \rightarrow \quad \Xi^- + K^+,\ \Xi^0 + K^0,\ \Omega^- + K^+ + K^0$$

b.) Antibaryons:

6. $B(\bar{p}) = -1$:

$$p + p \;\rightarrow\; p + p + p + \bar{p}$$

7. $B(\bar{B}) = -1$:

$$p + \bar{p} \;\rightarrow\; \bar{n} + n,\ \bar{\Lambda}^0 + \Lambda^0,\ \bar{\Sigma}^0 + \Sigma^0,\ \bar{\Sigma}^\pm + \Sigma^\mp,\ \bar{\Xi}^+ + \Xi^-$$

c.) Photon:

8. $B(\gamma) = 0$:

$$p \;\rightarrow\; p + \gamma$$

d.) Leptons: All leptons are produced in pairs, $B(e^-) = 0$ by convention.

9. $B(e) = B(\mu) = 0$:

$$\gamma^* \;\rightarrow\; e^+ + e^-,\ \mu^+ + \mu^-$$

10. $B(\nu_e) = B(\nu_\mu) = 0$:

$$\begin{aligned} n &\rightarrow p + e^- + \bar{\nu}_e \\ \mu^- &\rightarrow e^- + \bar{\nu}_e + \nu_\mu \\ \mu^+ &\rightarrow e^+ + \nu_e + \nu_\mu^- \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu \end{aligned}$$

Index