November 2014

Introduction to quantum field theory II

lectures by
Horatiu Nastase

Instituto de Física Teórica, UNESP

São Paulo 01140-070, SP, Brazil

Contents

1 Lecture 1: Review of path integral and operator formalism and the Feynman diagram expansion

In this first lecture, I will review some important material from Quantum Field Theory I, just to set up the notation.

Conventions.

I will use theorist's conventions throughout, with $\hbar = c = 1$, which means that, e.g. $[E] = [1/x] = 1$. I will also use the *mostly plus* metric, for instance in 3+1 dimensions with signature $- + ++$.

Path integrals.

In quantum field theory, the classical field $\phi(x)$ is replaced by the VEV of a quantum operator $\hat{\phi}$,

$$\langle 0|\hat{\phi}(x)|0\rangle = \int \mathcal{D}\phi e^{iS[\phi]}\phi(x). \tag{1.1}$$

Here the path integral measure is defined as integration over the points of a discretized path, in the limit of infinite number of points on the path,

$$\mathcal{D}\phi \equiv \lim_{N\to\infty}\prod_{i=1}^{N}\int d\phi(x_i). \tag{1.2}$$

Scalar field.

For a scalar field, the action is typically of the type

$$\begin{aligned} S &= \int d^4x\mathcal{L} = \int d^4x\left[-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - V(\phi)\right] \\ &= \int d^4x\left[\frac{1}{2}\dot{\phi}^2 - \frac{1}{2}|\vec{\nabla}\phi|^2 - \frac{1}{2}m^2\phi^2 - V(\phi)\right], \end{aligned} \tag{1.3}$$

though the kinetic term can be more complicated also. Note that the mass term was written separately, since it is quadratic, though technically it is part of the potential $V(\phi)$.

One can define objects more general than the field VEV (of which the field VEV is a particular case), that are the objects to be studied in quantum field theory, the *correlation functions*, or *Green's functions*, or n-point functions,

$$G_n(x_1, ..., x_n) = \langle 0|T\{\hat{\phi}(x_1)...\hat{\phi}(x_n)\}|0\rangle = \int \mathcal{D}\phi e^{iS[\phi]}\phi(x_1)...\phi(x_n). \tag{1.4}$$

Here T denotes time ordering. Note that e^{iS} is a highly oscillatory phase, so this Green's function is hard to define rigorously, since it is not well behaved at infinity.

It is much easier to go to Euclidean space, by doing a Wick rotation, $t = -it_E$.

In quantum mechanics, one obtains the *Feynman-Kac formula* relating the transition amplitude in Euclidean space, expressed as a path integral, with the usual statistical mechanics partition function,

$$Z(\beta) = \mathrm{Tr}\{e^{-\beta\hat{H}}\} = \int dq\sum_n|\phi_n(q)|^2e^{-\beta E_n} = \int dq\langle q,\beta|q,0\rangle$$

$$= \int \mathcal{D}q e^{-S_E[q]}\Big|_{q(t_E+\beta)=q(t_E)} . \quad (1.5)$$

Here as usual $\beta = 1/k_B T$ and the Euclidean action S_E is defined by $iS_M = -S_E$.

In quantum field theory, where we have a space ($\vec{x}$) dependence, and one usually is interested in the vacuum functional, i.e. the transition between asymptotic vacuum states, we consider the limit of infinite periodicity, $\phi(\vec{x}, t_E + \beta) = \phi(\vec{x}, t_E)$ (or zero temperature $T = 1/\beta$), where $\beta \to \infty$. The Euclidean action for a scalar field is

$$S_E[\phi] = \int d^4x \left[\frac{1}{2}\partial_\mu\phi\partial_\mu\phi + \frac{1}{2}m^2\phi^2 + V(\phi)\right] . \quad (1.6)$$

The Euclidean space correlation functions are defined in a similar manner,

$$G_n^{(E)}(x_1, ..., x_n) = \int \mathcal{D}\phi e^{-S_E[\phi]}\phi(x_1)...\phi(x_n) , \quad (1.7)$$

the advantage being that now instead of a highly oscillatory phase, we have a highly decaying weight e^{-S_E}, sharply peaked on the classical action $S_{\rm cl}$. The generating functional of the correlation functions is called the partition function and is given by

$$Z^{(E)}[J] = \int \mathcal{D}\phi e^{-S_E[\phi]+J\cdot\phi} \equiv {}_J\langle 0|0\rangle_J , \quad (1.8)$$

where we have defined

$$J \cdot \phi \equiv \int d^dx J(x)\phi(x). \quad (1.9)$$

The correlation functions are obtained from their generating functional as usual,

$$\begin{aligned} G_n^{(E)}(x_1, ..., x_n) &= \frac{\delta}{\delta J(x_1)}...\frac{\delta}{\delta J(x_n)}\int \mathcal{D}\phi e^{-S_E[\phi]+J\cdot\phi}\Big|_{J=0} \\ &= \frac{\delta}{\delta J(x_1)}...\frac{\delta}{\delta J(x_n)} Z[J]\Big|_{J=0} . \end{aligned} \quad (1.10)$$

Note that we can define the partition function at finite temperature (finite periodicity β),

$$Z^E[\beta, J] = {\rm Tr}[e^{-\beta\hat{H}_J}] = \int \mathcal{D}\phi e^{-S_E[\phi]+J\cdot\phi}\Big|_{\phi(\vec{x},t_E+\beta)=\phi(\vec{x},t_E)} . \quad (1.11)$$

From it we can start to define quantum field theory at finite temperature, but we will not do it here.

Canonical quantization and operator formalism.

To canonically quantize a real scalar field, one expands it and its canonical conjugate momentum π in Fourier modes,

$$\phi(\vec{x}, t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}(a(\vec{p}, t)e^{i\vec{p}\cdot\vec{x}} + a^\dagger(\vec{p}, t)e^{-i\vec{p}\cdot\vec{x}})$$

$$\pi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\left(-i\sqrt{\frac{\omega_p}{2}}\right)\left(a(\vec{p},t)e^{i\vec{p}\cdot\vec{x}} - a^\dagger(\vec{p},t)e^{-i\vec{p}\cdot\vec{x}}\right) , \tag{1.12}$$

where $\omega_p = \sqrt{\vec{p}^2 + m^2}$. From the KG equation of motion for the scalar field, one finds that $a(\vec{p},t) = a_{\vec{p}} e^{-i\omega_p t}$.

Canonical quantization is achieved through the equal time commutation relations between ϕ and its canonical conjugate momentum,

$$[\phi(\vec{x},t),\pi(\vec{x}',t)] = i\hbar\delta^{(3)}(\vec{x}-\vec{x}'); \quad [\phi(\vec{x},t),\phi(\vec{x}',t)] = [\pi(\vec{x},t),\pi(\vec{x}',t)] = 0. \tag{1.13}$$

From them, we obtain the usual algebra for the creation and anihilation operator coefficients,

$$[a(\vec{p},t), a^\dagger(\vec{p}',t)] = (2\pi)^3\delta^{(3)}(\vec{p}-\vec{p}'). \tag{1.14}$$

For free scalars, one uses Heisenberg picture operators,

$$\phi_H(\vec{x},t) = e^{iHt}\phi(\vec{x})e^{-iHt} , \tag{1.15}$$

where $\phi(\vec{x})$ is a Schrödinger picture operator. As a reminder, in the Schrödinger picture, operators are time-independent and the states evolve in time with the Hamiltonian, whereas in the Heisenberg picture it is the opposite: operators evolve in time with the Hamiltonian, and the states are time-independent.

For interacting scalars however, the useful representation is the *interaction (Dirac) picture*. One splits the Hamiltonian into a free (quadratic) part and an interaction part,

$$\hat{H} = \hat{H}_0 + \hat{H}_1. \tag{1.16}$$

Then the interaction picture operators are obtained by a canonical transformation with the free part $\hat{H}_0$,

$$\phi_I(\vec{x},t) = e^{i\hat{H}_0(t-t_0)}\phi(\vec{x},t_0)e^{-i\hat{H}_0(t-t_0)}. \tag{1.17}$$

Now states evolve with the interaction Hamiltonian, and operators with the free Hamiltonian,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle &= \hat{H}_{1,I}|\psi_I(t)\rangle \\ i\hbar\frac{\partial}{\partial t}\hat{A}_I(t) &= [\hat{A}_I(t),\hat{H}_0] , \end{aligned} \tag{1.18}$$

where $H_{1,I}$ is the interacting Hamiltonian operator in the interaction picture.

In the interacting quantum field theory, $\phi(\vec{x},t)$ denotes the Heisenberg operator $\phi_H(\vec{x},t)$, and is the object we are interested in, together with the true vacuum of the full theory $|\Omega\rangle$. On the other hand, for calculational purposes, we use the vacuum $|0\rangle$ of the free theory, and the interaction picture fields, since we find that

$$\begin{aligned} \phi_I(\vec{x},t) &= e^{i\hat{H}_0(t-t_0)}\phi(\vec{x},t_0)e^{-i\hat{H}_0(t-t_0)} \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_p}}(a_{\vec{p}}e^{ip\cdot x} + a_{\vec{p}}^\dagger e^{-ip\cdot x})\Big|_{x^0=t-t_0;p^0=E_p} . \end{aligned} \tag{1.19}$$

The vacuum of the free theory satisfies that $a_{\vec{p}}|0\rangle = 0$ for all $\vec{p}$. On the other hand, the objects we want to calculate are Green's functions of the full theory, like the 2-point function (full propagator), $\langle\Omega|T\{\phi(x)\phi(y)\}|\Omega\rangle$.

We find them using perturbation theory, in terms of interaction picture objects, like the propagator. The propagator depends on a contour of complex integration, but the most commonly used one in quantum field theory is the *Feynman propagator*,

$$D_F(x-y) = \langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2+m^2-i\epsilon}e^{ip\cdot(x-y)}. \tag{1.20}$$

Other examples of propagators are the retarded and advanced propagators. But the Feynman propagator arises as the natural analytical continuation from Euclidean space. In Euclidean space, the propagator is uniquely defined, since there are no poles to be avoided using complex integration and contours, and it is given by

$$\Delta(x,y) = \int \frac{d^dp}{(2\pi)^d}\frac{e^{ip\cdot(x-y)}}{p^2+m^2}. \tag{1.21}$$

The relation between the full correlation functions and the interaction picture objects is given by the *Feynman theorem*, written for generality in terms of operators $\mathcal{O}$ that can specialize to the usual case of $\phi(x)$, as

$$\langle\Omega|T\{\mathcal{O}_H(x_1)....\mathcal{O}_H(x_n)\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)}\frac{\langle 0|T\{\mathcal{O}_I(x_1)...\mathcal{O}_I(x_n)\exp\left[-i\int_{-T}^T dtH_{1,I}(t)\right]\}|0\rangle}{\langle 0|T\{\exp\left[-i\int_{-T}^T dtH_{1,I}(t)\right]\}|0\rangle}. \tag{1.22}$$

By expanding the right hand side, we can calculate perturbatively in terms of the free vacuum and operators with time evolution given by $\hat{H}_0$, i.e. interaction picture operators.

For calculations, we use the Wick theorem, which relates the time ordering with the normal ordering,

$$T\{\phi_I(x_1)...\phi_I(x_n)\} = N\{\phi_I(x_1)...\phi_I(x_n) + \text{all possible contractions}\}. \tag{1.23}$$

Since in between $\langle 0|$ and $|0\rangle$, the normal ordering of a nontrivial interaction picture operator (written in terms of a's and $a^\dagger$'s) gives zero, the only nonzero result is given by the full contraction, which is just a c-number.

In the path integral formalism, there is an equivalent of the Wick theorem. One first writes the Dyson formula, which reads

$$\begin{aligned} Z[J] &= \langle 0|e^{-S_I[\hat{\phi}]}e^{\int d^dxJ(x)\hat{\phi}(x)}|0\rangle \\ &= \int \mathcal{D}\phi e^{-S_0[\phi]+J\cdot\phi}e^{-S_I[\phi]}. \end{aligned} \tag{1.24}$$

Note that the vacuum of the free theory is used, but inside the VEV we have the missing terms that allow to make up the usual $e^{-S+J\cdot\phi}$ in the path integral, using the weight e^{-S_0}.

Then we can easily calculate the partition function of the free theory,

$$Z_0[J] = e^{\frac{1}{2}J\cdot\Delta\cdot J}\langle 0|0\rangle_0 \ , \tag{1.25}$$

where $\langle 0|0\rangle_0$ is an irrelevant normalization constant. Then the Wick theorem for path integrals is written in two forms. The first is

$$Z[J] = e^{-\int d^dx V\left(\frac{\delta}{\delta J(x)}\right)} Z_0[J] = e^{-\int d^dx V\left(\frac{\delta}{\delta J(x)}\right)} e^{\frac{1}{2}J\cdot\Delta\cdot J} \ , \tag{1.26}$$

and the second is

$$\begin{aligned} Z[J] &= e^{\frac{1}{2}\frac{\delta}{\delta\phi}\cdot\Delta\cdot\frac{\delta}{\delta\phi}} \left\{ e^{-\int d^dx V(\phi) + J\cdot\phi} \right\}\Big|_{\phi=0} \\ &= \exp\left[\frac{1}{2}\int d^dx \ d^dy \Delta(x-y)\frac{\delta}{\delta\phi(x)}\frac{\delta}{\delta\phi(y)}\right] \left\{ e^{-\int d^dx V(\phi)+J\cdot\phi}\right\}\Big|_{\phi=0} . \end{aligned} \tag{1.27}$$

While it would seem that by these formulas we have solved quantum field theory, since we have a closed form expression for the partition function, it is not so, since the expression is formal: it is understood only as a formal perturbative expansion for the exponentials. There are divergences in the terms of the expansion that need to be regulated, etc.

To find the correlation functions of the *full* (interacting) theory, we need to get rid of the vacuum bubbles. As seen from the denominator of the Feynman theorem, this is done by dividing with the partition function (zero point function). Diagramatically (to be defined soon), one obtains only the connected diagrams, giving

$$\frac{1}{Z[J]}\frac{\delta Z[J]}{\delta J(x)} = \frac{\delta(-W[J])}{\delta J(x)} \ , \tag{1.28}$$

solved by

$$Z[J] = \mathcal{N} e^{-W[J]}. \tag{1.29}$$

Here $W[J]$ is called the free energy, from the analogy with condensed matter, and as we see is the generating functional of the connected diagrams. $\mathcal{N}$ is an irrelevant normalization constant (it cancels out).

Feynman rules in x space (Euclidean).

The Feynman rules in Euclidean coordinate space are as follows:

0. Draw all Feynman diagrams.

1. A line between x and y corresponds to the Euclidean Feynman propagator $\Delta(x, y)$.

2. For an interaction potential $V = \lambda\phi^p$, we have p-legged vertices. For each vertex, we have a factor of $(-\lambda)$ and an integration over the position of the vertex, $\int d^dx$.

3. Then we obtain the value for the Feynman diagram D with external points $x_1, ..., x_n$ nand internal points $y_1, ..., y_N$, integrated over,

$$F_D^{(N)} = I_D(x_1, ..., x_n; y_1, ..., y_N) \ , \tag{1.30}$$

and the correlation function is given as a sum over diagrams and number of vertices,

$$G_n(x_1, ..., x_n) = \sum_{N\geq 0}\frac{1}{N!}\sum_{\text{diags.D}} F_D^{(N)}(x_1, ..., x_n). \tag{1.31}$$

Simplified rules.

There is a simplified version of the Feynman rules that is usually used. We rewrite the potential as $V = \lambda_p \phi^p/p!$ (i.e., $\lambda = \lambda_p/p!$). Then the vertex is $\int d^d x(-\lambda_p)$, but we write only the *topologically inequivalent diagrams* and divide by the *statistical weight factor* or *symmetry factor* S,

$$S = \frac{N!(p!)^N}{\#\text{of equiv. diags.}} = \#\text{of symmetries of diagram.} \tag{1.32}$$

Feynman rules in p space.

Because of translational invariance of the theory, which implies momentum conservation, the Fourier transform $\tilde{G}$ of the correlation function can be redefined by an overall momentum conservation delta function,

$$\tilde{G}_n(p_1, ..., p_n) = \int \prod_i d^d x_i e^{i\sum_j p_j x_j} G(x_1, ..., x_n) = (2\pi)^d \delta^{(d)}(p_1 + ... + p_n) G_n(p_1, ..., p_n). \tag{1.33}$$

The simplest case is the free 2-point function, i.e. the propagator. Corresponding to the Feynman propagator $\Delta(x-y)$, we have the Euclidean space propagator

$$\Delta(p) = \frac{1}{p^2 + m^2} \ , \tag{1.34}$$

corresponding to G_2 above.

The rules in momentum space have:

-propagator $\Delta(p)$.

-vertex $(-\lambda)$.

-external line $e^{-ip\cdot x}$.

-Then one needs to impose momentum conservation, i.e. multiply by $(2\pi)^d \delta^{(d)}(\sum_j p_j)$, at each vertex, and integrate, $\int d^d p/(2\pi)^d$, over all *internal* momenta.

-divide by the symmetry factor.

Simplified momentum space rules.

We can write simplified momentum space rules, to get rid of the momentum conservation delta functions. We only introduce *independent loop momenta* $l_1, ..., l_L$ (integration variables). If it is not entirely obvious what is the number of loops of the diagrams, one can use the formula $L = V(p/2 - 1) - E/2 + 1$ to calculate it. Here V is the number of vertices, p the order of vertices, E the number of external lines. We write the momenta on each internal line in terms of the external momenta and the loop momenta l_i using momentum conservation at the vertices.

-We must integrate over these loop momenta, $\int d^d l_1/(2\pi)^d ... \int d^d l_L/(2\pi)^d$.

-External lines have momenta p_i, and internal lines dependent momenta q_j, depending on p_i and l_i.

-An external line gives then $1/(p_i^2 + m^2)$, and an internal line $1/(q_j^2 + m^2)$.

-As before, the vertex is $(-\lambda)$, we divide by the symmetry factor S and sum over diagrams.

Classical field.

As we mentioned at the beginning of the lecture, the classical field is replaced by the VEV of the quantum operator. More precisely, considering the field in the presence of a source J, one defines the *classical field* $\phi_{\rm cl}$ by the VEV in the $|0\rangle_J$ vacuum, and divides by the normalization,

$$\phi_{\rm cl} \equiv \phi(x;J) = \frac{{}_J\langle 0|\hat{\phi}|0\rangle_J}{{}_J\langle 0|0\rangle_J}. \tag{1.35}$$

Therefore one obtains

$$\phi_{\rm cl} = \frac{1}{Z[J]}\int \mathcal{D}\phi e^{-S[\phi]+J\cdot\phi}\phi(x) = \frac{\delta}{\delta J(x)}\ln Z[J] = -\frac{\delta}{\delta J(x)}W[J]. \tag{1.36}$$

Quantum effective action.

Whereas classically, we have the classical action $S[\phi]$ depending on the classical field ϕ, quantum mechanically we have the *quantum effective action* that includes quantum corrections, and depends on the classical field $\phi_{\rm cl}$. It is the Legendre transform of the free energy $W[J]$,

$$\Gamma[\phi_{\rm cl}] = W[J] + \int d^d x J(x)\phi_{\rm cl}(x). \tag{1.37}$$

In the same way as classically, $\phi(x)$ satisfies the classical equation of motion with a source,

$$\frac{\delta S[\phi]}{\delta\phi(x)} = J(x)\ , \tag{1.38}$$

quantum mechanically we have a precise analog of it, namely

$$\frac{\delta\Gamma[\phi_{\rm cl}]}{\delta\phi_{\rm cl}(x)} = J(x). \tag{1.39}$$

The quantum effective action Γ is the generator of the 1PI (one particle irreducible) diagrams, except for the 2-point function, where there is an extra term.

Therefore the partition function $Z[J]$ generates all diagrams, the free energy $W[J]$ generates the connected diagrams, the effective action $\Gamma[\phi_{\rm cl}]$ generates the 1PI diagrams except at 2-points, and we also have that the classical action S generates the tree diagrams.

S-matrix.

To get contact with experiments, we need to define objects that can be related to measurable quantities. Such an object is the S-matrix.

We define Heisenberg picture states (time independent) that have well isolated wavepackets at $t=-\infty$ (but are interacting, or mixed, at $t=+\infty$), $|\{\vec{k}_j\}\rangle_{\rm in}$, and states that have well isolated wavepackets at $t=+\infty$ (but are interacting, or mixed, at $t=-\infty$), $|\{\vec{p}_i\}\rangle_{\rm out}$. We also define Schrödinger picture states (time dependent) with the same momenta, $\langle\{\vec{p}_i\}|$ and $|\{\vec{k}_j\}\rangle$. Then the S matrix between the Schrödinger picture states is defined as

$$\langle\{\vec{p}_i\}|S|\{\vec{k}_j\}\rangle = {}_{\rm out}\langle\{\vec{p}_i\}|\{\vec{k}_j\}\rangle_{\rm in}. \tag{1.40}$$

Reduction formula (LSZ).

Until now we have worked with correlation functions. But these objects are rather abstract, and we saw above that physical objects we want to calculate are S-matrices. Fortunately, there is a relation between them, given by the *LSZ formula*. One first defines the Fourier transform of the correlation functions of the full (interacting) theory,

$$\tilde{G}_{n+m}(p_i^\mu, k_j^\mu) = \prod_{i=1}^{n} \int d^4x_i e^{-ip_i x_i} \prod_{j=1}^{m} \int d^4y_j e^{ik_j y_j} \langle \Omega | T\{\phi(x_1)...\phi(x_n)\phi(y_1)...\phi(y_m)\}|\Omega\rangle. \tag{1.41}$$

Then the S-matrix is obtained from the correlation function $\tilde{G}_{n+m}$ by going near on-shell for the external lines and dividing with the full inverse propagators for the external lines, except for the factor of Z being replaced by $\sqrt{Z}$,

$$_{\rm out}\langle\{\vec{p}_i\}|\{\vec{k}_j\}\rangle_{\rm out} = \lim_{p_i^2\to -m^2; k_j^2\to -m^2} \prod_{i=1}^{n} \frac{p_j^2+m^2-i\epsilon}{-i\sqrt{Z_i}} \prod_{j=1}^{m} \frac{k_j^2+m^2-i\epsilon}{-i\sqrt{Z_j}} \tilde{G}_{n+m}(p_i^\mu, k_j^\mu). \tag{1.42}$$

Diagramatically, the S-matrix minus the trivial one (identity) is given by the sum of the connected, amputated Feynman diagrams, times a $\sqrt{Z}$ factor for each external leg, i.e.

$$\langle\{\vec{p}_i\}|S-1|\{\vec{k}_j\}\rangle = \left(\sum \text{connected, amputated Feynman diagrams}\right) \times (\sqrt{Z})^{n+m}. \tag{1.43}$$

Fermions

Dirac fermions ψ_α are understood as representations of the Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\,\mathbb{1}\;, \tag{1.44}$$

i.e. column vector objects on which the gamma matrices act. We can also understand them as spinorial representations of the Lorentz (or Poincaré) algebra. But the Dirac fermion representation of the Lorentz algebra is not irreducible. In 3+1 dimensions, the irreducible representations are either of the Weyl or Majorana type.

One first defines the γ_5 object as the product of the gamma matrices up to a fixed phase, in my conventions

$$\gamma_5 = -i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{1.45}$$

Note that there are various conventions for the phase, but generally one chooses to have $\gamma_5^2 = 1$. Indeed, then we can define projectors $P_{L/R} = (1 \pm \gamma_5)/2$ onto the irreducible Weyl representations, defined as

$$\psi_{L/R} = \frac{1 \pm \gamma_5}{2}\psi_D. \tag{1.46}$$

Note that in the Weyl representation for the gamma matrices, $\gamma_5 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$, so the irreps are the two upper components of ψ_D and the two lower components of ψ_D.

One defines then the conjugate representation $\bar{\psi}$ as $\bar{\psi} = \psi^\dagger i\gamma^0$. Note again that various conventions in the literature for $\bar{\psi}$ differ by a phase. The action in Minkowski space is then

$$S_\psi = -\int d^4x \bar{\psi}(\gamma^\mu\partial_\mu + m)\psi\;, \tag{1.47}$$

and in Euclidean space without the overall minus.

The other type of irreducible representation is a real-type representation, the Majorana representation, obtained by imposing a reality constraint. Defining a C-matrix with certain properties, satisfied in 3+1 dimensions by the *choice* $C = -i\gamma^0\gamma^2$, the Majorana condition relates the Dirac conjugate $\bar{\psi}$ with the Majorana conjugate $\psi^C = \psi^T C$, i.e.

$$\bar{\psi} = \psi^T C. \tag{1.48}$$

Fermionic path integrals are written in terms of *Grassmann variables*, i.e. anticommuting objects ψ with $\{\psi, \psi\} = 0$. The Gaussian integral on Grassmann space gives

$$\int d^n x e^{x^T A x} = 2^{n/2}\sqrt{\det A}\ , \tag{1.49}$$

where for commuting objects it gives $\propto 1/\sqrt{\det A}$.

The addition of fermions means new propagators and vertices in the Feynman rules.

The fermionic Euclidean space propagator is

$$\frac{1}{i\not{p} + m} = \frac{i}{-\not{p} + im} = \frac{-i\not{p} + m}{p^2 + m^2}\ , \tag{1.50}$$

whereas in Minkowski space the Feynman propagator is

$$-i\frac{-i\not{p} + m}{p^2 + m^2 - i\epsilon}. \tag{1.51}$$

For a Yukawa interaction

$$\int d^d x g\bar{\psi}\psi\phi\ , \tag{1.52}$$

the vertex is $(-g)$ in the Euclidean case, and $(-ig)$ in the Minkowski case.

A fermion loop adds a minus sign to the Feynman rules.

Gauge fields

In covariant quantization (Gupta-Bleuler), we write the expansion of the gauge field in a similar manner with the case of the scalar field,

$$A_\mu(x) = \int \frac{d^3k}{(2\pi)^3\sqrt{2E_k}} \sum_{\lambda=0}^{3} \epsilon_\mu^{(\lambda)}(k)[e^{ikx}a^{(\lambda)}(k) + a^{\dagger(\lambda)}(k)e^{-ikx}]\ , \tag{1.53}$$

where $\epsilon_\mu^{(\lambda)}(k)$ are 4 polarizations. The 0 and 3 polarizations are unphysical, and the 1 and 2 are physical, transverse polarizations, satisfying $k^\mu\epsilon_\mu^{(1,2)} = 0$.

The gauge condition is imposed on the part of positive frequency of the operator, i.e. the annihilation part of A_μ, acting on physical states,

$$\partial^\mu A_\mu^{(+)}(x)|\psi\rangle = 0\ , \tag{1.54}$$

which defines the physical states $|\psi\rangle$. Since $k^\mu \epsilon_\mu^{(1,2)} = 0$, on the momentum modes we obtain the condition

$$[a^{(0)}(k) + a^{(3)}(k)]|\psi\rangle = 0. \tag{1.55}$$

To define a photon propagator, we must add a gauge fixing term to the action. If not, the action has zero modes (due to the gauge invariance), which makes it impossible to invert the kinetic operator. The gauge-fixed Lagrangean in Minkowski space is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - \frac{1}{2\alpha}(\partial^\mu A_\mu)^2. \tag{1.56}$$

From it, we obtain the propagator

$$G_{\mu\nu}^{(0)}(k) = \frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right). \tag{1.57}$$

In the so-called "Feynman gauge" (it is not really a gauge, just a choice) $\alpha = 1$, the propagator becomes just the scalar propagator $G_{\mu\nu}^{(0)}(k) = \delta_{\mu\nu}/k^2$.

QED S-matrix Feynman rules.

The Euclidean Lagrangean for QED is

$$\mathcal{L}^{(E)} = \frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}(\not{D} + m)\psi\ , \tag{1.58}$$

where $D_\mu = \partial_\mu - ieA_\mu$. The Minkowski Lagrangean is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - \bar{\psi}(\not{D} + m)\psi. \tag{1.59}$$

-The $e^- e^+ \gamma$ vertex, with respective indices α, β and μ is $+ie(\gamma_\mu)_{\alpha\beta}$ in Euclidean space, and $+e(\gamma_\mu)_{\alpha\beta}$ in Minkowski space.

-The incoming photon external line is $\overbrace{A_\mu|\vec{p}}\rangle = \epsilon_\mu(p)$.

-The outgoing photon external line is $\langle\overbrace{\vec{p}|A_\mu} = \epsilon^*_\mu(p)$.

-The incoming electron external line is $\overbrace{\psi|\vec{p}}, s\rangle = u^s(p)$.

-The outgoing electron external line is $\langle\overbrace{\vec{p}, s|\bar{\psi}} = \bar{u}^s(p)$.

-The incoming positron external line is $\overbrace{\bar{\psi}|\vec{k}}, s\rangle = \bar{v}^s(k)$.

-The outgoing positron external line is $\langle\overbrace{\vec{k}, s|\psi} = v^s(k)$.

-The fermion propagator is

$$-\frac{(\not{p} + im)}{p^2 + m^2 - i\epsilon}. \tag{1.60}$$

-The photon propagator is

$$-\frac{i}{k^2 - i\epsilon}\left(g_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2 - i\epsilon}\right). \tag{1.61}$$

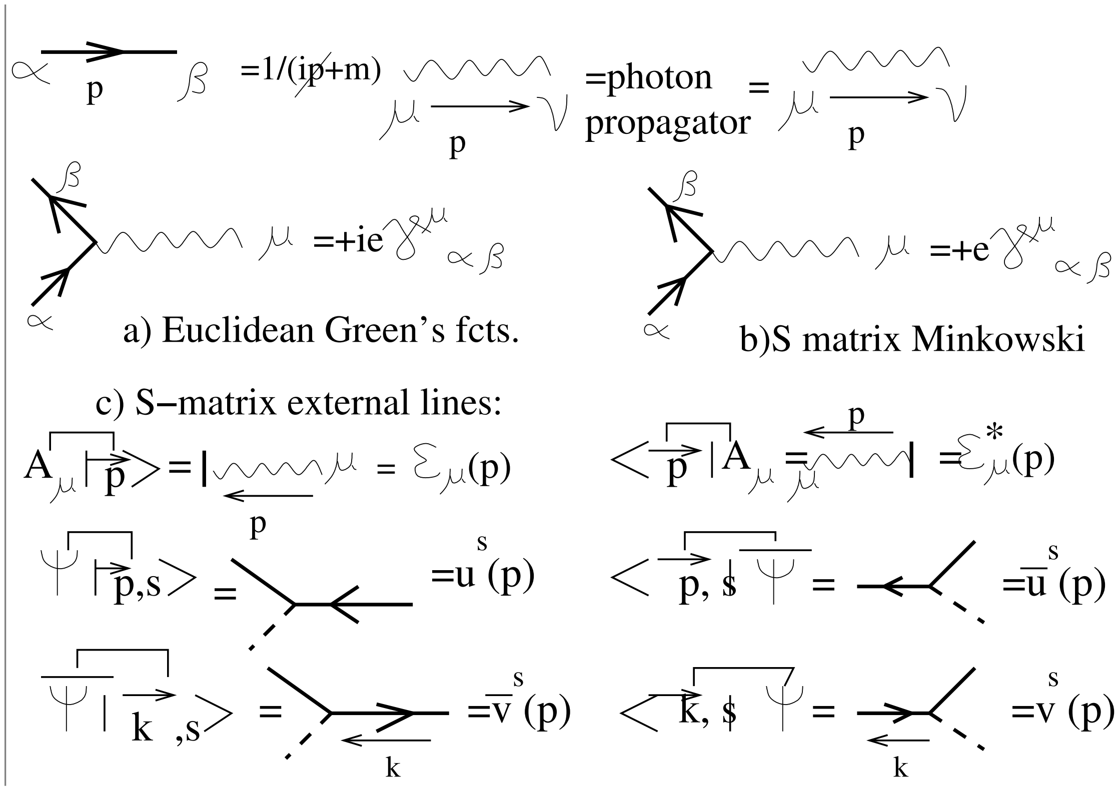

Figure 1: Relevant Feynman rules for Green's functions and S-matrices.

-There is a minus sign for a fermion loop.
The rules are summarized in Fig.1.

Important concepts to remember

- All of them, since it is a review...

Further reading: See my lecture notes for QFT I [1].

Exercises, Lecture 1

1) Use the Feynman rules to write down the integral expression for the following diagram (Fig.2) for scattering of photons in QED (the diagram is a fermion loop, cut into two loops by a photon line, and with 4 external photon lines).

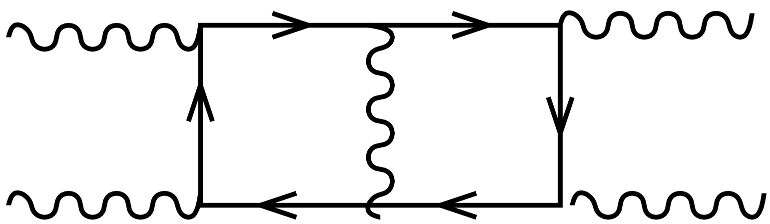

Figure 2: Two-loop QED Feynman diagram.

2) Write down the x-space and the p-space Feynman rules coming from the 4 dimensional Lagrangean in Euclidean space:

$$\mathcal{L} = +\frac{1}{2}\sum_{I=1}^{N}[(\partial_\mu\phi^I)^2 + m^2(\phi^I)^2] + \left(1 + \sum_{I=1}^{N}\left(\frac{\phi^I}{M}\right)^2\right)\frac{F_{\mu\nu}F^{\mu\nu}}{4} + \frac{\lambda}{M^2}\sum_{I,J=1}^{N}\phi^I\phi^I\partial_\mu\phi^J\partial^\mu\phi^J. \tag{1.62}$$

2 Lecture 2. One-loop divergences, renormalizability and power counting

In this lecture we will analyze possible divergences in loop integrals, in particular we will look at one-loop, and how to determine if a theory contains divergences using power counting. Finally we will define power counting renormalizability and distinguish between theories based on it.

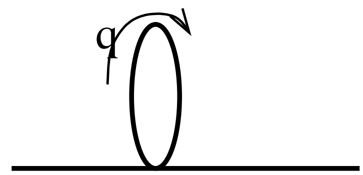

Figure 3: One-loop divergence in ϕ^4 theory for the 2-point function.

We have seen in the first semester hints of the fact that loop integrals can be divergent, giving infinite values for Feynman diagrams.

For example, in $\lambda\phi^4$ theory in Euclidean space, consider the unique one-loop $\mathcal{O}(\lambda)$ diagram, a loop connected to the free line (propagator) at a point, see Fig.3. It is given by

$$-\lambda \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2+m^2}. \tag{2.1}$$

Since the integral is

$$= -\lambda \frac{\Omega_{D-1}}{(2\pi)^D} \int q^{D-1} dq \frac{1}{q^2+m^2} \sim \int dq\, q^{D-3}\ , \tag{2.2}$$

it is divergent in $D \geq 2$ and convergent only for $D < 2$. In particular, in $D = 4$ it is quadratically divergent

$$\sim \int^{\Lambda} q dq \sim \Lambda^2. \tag{2.3}$$

We call this kind of divergence "ultraviolet", or UV divergence, from the fact that it is at large energies (4-momenta), or large frequencies.

Note also that we had one more type of divergence for loop integrals that was easily dealt with, the fact that when integrating over loop momenta in Minkowski space, the propagators can go on-shell, leading to a pole, which needed to be regulated. But the $i\epsilon$ prescription dealt with that. Otherwise, we can work in Euclidean space and then analytically continue to Minkowski space at the end of the calculation. This issue did not appear at tree level, when the propagators have fixed momenta, and are not on-shell, but it appears in loop integrals.

Let us now consider an example of a diagram that has all types of divergences, a one-loop diagram for the $2k$-point function in $\lambda\phi^{k+2}$ theory with k momenta in, then two propagators

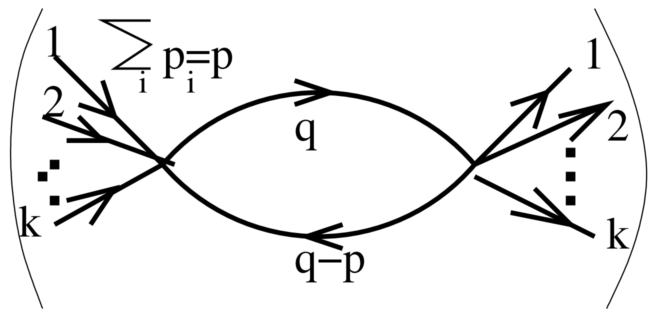

Figure 4: One-loop diagram in ϕ^{k+2} theory, for the 2k-point function.

forming a loop, and then k lines out, as in Fig.4. The incoming momenta are called $p_1, ..., p_k$ and sum to $p = \sum_{i=1}^{k} p_i$. Then the two propagators have momenta q (loop variable) and $q - p$, giving for the diagram

$$\frac{\lambda^2}{2} \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 + m^2)((q-p)^2 + m^2)}. \tag{2.4}$$

Again, we can see that at large q, it behaves as

$$\int \frac{q^{D-1} dq}{q^4} , \tag{2.5}$$

so is convergent only for $D < 4$. In particular, in $D = 4$ it is (log) divergent, and this again is an UV divergence. From this example we can see that various diagrams are divergent in various dimensions.

As mentioned, the poles in the propagators when we go to Minkowski space mean that a priori there are these divergences as well, but they are regulated by the Feynman $i\epsilon$ prescription.

But this diagram has also another type of divergence, namely at low q ($q \to 0$). This divergence appears only *if we have* $m^2 = 0$ *AND* $p^2 = 0$. Thus only if we have massless particles, and all the particles that are incoming on the same vertex sum up to something on-shell (in general, the sum of on-shell momenta is not on-shell). Then the integral is

$$\sim \int d\Omega \int \frac{q^3 dq}{q^2(q^2 - 2q \cdot p)} , \tag{2.6}$$

and in the integral over angles, there will be a point where the unit vector on q, $\hat{q}$, satisfies $\hat{q} \cdot \hat{p} = 0$ with respect to the (constant) unit vector on p, $\hat{p}$. Then we obtain

$$\int \frac{dq}{q} , \tag{2.7}$$

i.e., log divergent. We call this kind of divergences "infrared" or IR divergences, since they occur at low energies (low 4-momenta), i.e. low frequencies.

Thus we have two kinds of potential divergences, UV and IR divergences. The UV divergences are an artifact of perturbation theory, i.e. of the fact that we were forced to introduce asymptotic states as states of the free theory, and calculate using Feynman diagrams. As such, they can be removed by redefining the parameters of the theory (like masses, couplings, etc.), a process known as *renormalization*, which will be studied next in this course. But these UV divergences are a characteristic of the theory, hence their presence tells us we need to do something to define better the theory.

A nonperturbative definition is not in general available, in particular for scattering processes it isn't. But for things like masses and couplings of bound states (like the proton mass in QCD, for instance), one can define the theory nonperturbatively, for instance on the lattice, and then we always obtain finite results. The infinities of perturbation theory manifest themselves only in something called the renormalization group, which will also be studied later in this course.

By contrast, the IR divergences are genuine divergences from the point of view of the Feynman diagram (can't be reabsorbed by redefining the parameters). But they arise because the Feynman diagram we are interested in, in the case of a theory with massless external states, and with external states that are on-shell at the vertex, are not quantities that can be experimentally measured. Indeed, for a massless external state ($m = 0$), of energy E, experimentally we cannot distinguish between the process with this external state, or with it and another emitted "soft and/or collinear particle", namely one of $m = 0$ and $E \simeq 0$ and/or parallel to the first. If we include the tree level process for that second process at the same order in the coupling constant (order λ^2 for the diagram under study), and sum it together with the first (loop level), we obtain a finite differential cross section (which can be experimentally measured), for a given cut-off in energy and/or angle of resolution between two particles. Therefore this divergence arises because the quantity we study is not a physical one; a physical measurement always has a minimal resolution for the energy (minimal detectable energy) and the angle between emitted particles. Only by summing over processes that cannot be distinguished by the physical detector do we get a finite quantity.

Thus the physical processes are always finite, in spite of the infinities in the Feynman diagram.

Analytical continuation

A question which one could have already asked is: Is Wick rotation of the final result the same with Wick rotation of the integral to Minkowski space, followed by evaluation?

Let us look at the simplest one-loop diagram in Euclidean space (in $\lambda\phi^4$, already discussed above)

$$\int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 + m^2}. \tag{2.8}$$

In Minkowski space it becomes

$$-i \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 + m^2 - i\epsilon} = +i \int \frac{d^{D-1} q}{(2\pi)^{D-1}} \int \frac{dq_0}{2\pi} \frac{1}{q_0^2 - \vec{q}^2 - m^2 + i\epsilon}\ , \tag{2.9}$$

where now $q^2 = -q_0^2 + \vec{q}^2$.

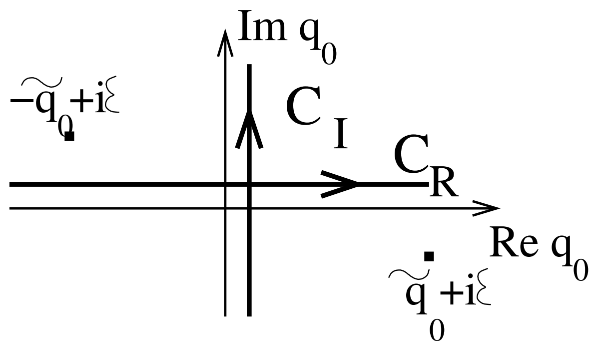

Figure 5: Wick rotation of the integration contour.

Then the poles are at $\tilde{q}_0 - i\epsilon$ and $-\tilde{q}_0 + i\epsilon$, where $\tilde{q}_0 = \sqrt{\vec{q}^2 + m^2}$. The Minkowski space integration contour is along the real axis in the q_0 plane, in the increasing direction, called C_R. On the other hand, the Euclidean space integration contour C_I is along the imaginary axis, in the increasing direction, see Fig.5. As there are no poles in betweeen C_R and C_I (in the quadrants I and III of the complex plane, the poles are in the quadrants II and IV), the integral along C_I is equal to the integral along C_R (since we can close the contour at infinity, with no poles inside). Therefore, along C_I, $q_0 = iq_D$, with q_D real and increasing, and therefore $dq_0 = idq_D$, so

$$\int_{C_R} dq_0(...) = \int_{C_I} dq_0(...) = \int \frac{d^{D-1}q}{(2\pi)^D}(-i)i \int \frac{dq_D}{\vec{q}^2 + (q_D)^2 + m^2 - i\epsilon} , \tag{2.10}$$

which gives the same result as the Euclidean space integral, after we drop the (now unnecessary) $-i\epsilon$.

However, in general it is not true that we can easily analytically continue. Instead, we must *define the Euclidean space integral and Wick rotate the final result*, since in general this will be seemingly different than the continuation of the Minkowski space integral (rather, it means that the Wick rotation of the integrals is subtle). But the quantum field theory perturbation in Euclidean space is well-defined, unlike the Minkowski space one, as we already saw, so is a good starting point.

Let's see an example of this situation. Consider the second integral we analyzed, now in Minkowski space

$$-\frac{\lambda^2}{2} \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 + m^2 - i\epsilon} \frac{1}{(q-p)^2 + m^2 - i\epsilon}. \tag{2.11}$$

We deal with the two different propagators in the loop integral using the Feynman trick. We will study it in more detail next class, but this time we will just use the result for two propagators. The Feynman trick for this case is the observation that

$$\frac{1}{AB} = \int_0^1 dx[xA + (1-x)B]^{-2} , \tag{2.12}$$

(which can be easily checked) which allows one to turn the two propagators with different momenta into a single propagator squared. Indeed, now we can write

$$-\frac{\lambda^2}{2}\int\frac{d^Dq}{(2\pi)^D}\int_0^1 dx[x(q-p)^2+(1-x)q^2+(x+1-x)(m^2-i\epsilon)]^{-2}. \tag{2.13}$$

The square bracket equals $q^2+xp^2-2xq\cdot p+m^2-i\epsilon$. Changing variables to $q'^\mu=q^\mu-xp^\mu$ allows us to get rid of the term linear in q. We can change the integration variable to q', since the Jacobian for the transformation is 1, and then the square bracket becomes $q'^2+x(1-x)p^2+m^2-i\epsilon$. Finally, the integral is

$$\frac{(-i\lambda)^2}{2}\int\frac{d^Dq'}{(2\pi)^D}\int_0^1 dx[q'^2+x(1-x)p^2+m^2-i\epsilon]^{-2}\ , \tag{2.14}$$

which has poles at

$$\tilde{q}_0^2=\vec{q}^2+m^2-i\epsilon+x(1-x)p^2. \tag{2.15}$$

If $p^2>0$, this is the same as the in the previous example, we just redefine m^2: the poles are outside quadrants I and III, so we can make the Wick rotation of the integral without problem. However, if $p^2<0$ and sufficiently large in absolute value, we can have $q_0^2<0$, so the poles are now in quadrants I and III, and we cannot simply rotate the contour C_R to the contour C_I, since we encounter poles along the way. So in this case, the Wick rotation is more subtle: apparently, the Minkowski space integral gives a different result from the Euclidean space result, Wick rotated. However, the latter is better defined, so we can use it.

Power counting

We now want to understand how we can figure out if a diagram, and more generally a theory, contains UV divergences. We do this by *power counting.* We consider here scalar $\lambda_n\phi^n$ theories.

Consider first just a (Euclidean space) diagram, with L loops and E external lines, and I internal lines and V vertices. The loop integral will be

$$I_D(p_1,...,p_E;m)=\int\prod_{\alpha=1}^{L}\frac{d^dq_\alpha}{(2\pi)^d}\prod_{j=1}^{I}\frac{1}{q_j^2+m^2}\ , \tag{2.16}$$

where $q_j=q_j(p_i,q_\alpha)$ are the momenta of the internal lines (which have propagators $1/(q_j^2+m^2)$). More precisely, they are linear combinations of the loop momenta and external momenta,

$$q_j=\sum_{\alpha=1}^{L}c_{j\alpha}q_\alpha+\sum_{i=1}^{E}c_{ji}p_i. \tag{2.17}$$

Like we already mentioned in Lecture 11 of QFT I, $L=I-V+1$, since there are I momentum variables, constrained by V delta functions (one at each vertex), but one of the delta functions is the overall (external) momentum conservation.

If we scale the momenta and masses by the same multiplicative factor t, we can also change the integration variables (loop momenta q_α) by the same factor t, getting $\prod_{\alpha=1}^L d^dq\to t^{LD}\prod_{\alpha=1}^L d^dq$, as well as $q_j\to tq_j$, and $q^2+m^2\to t^2(q^2+m^2)$, giving finally

$$I_D(tp_i;tm)=t^{\omega(D)}I_D(p_i;m)\ , \tag{2.18}$$

where

$$\omega(D) = dL - 2I \tag{2.19}$$

is called the *superficial degree of divergence* of the diagram D, since it is the overall dimension for the scaling above.

Theorem This gives rise to the following theorem: $\omega(D) < 0$ is *necessary* for the convergence of I_D. (Note: but is not sufficient!)

Proof: We have

$$\prod_{i=1}^{I}(q_i^2 + m^2) \leq (\sum_{i=1}^{I} q_i^2 + m^2)^I. \tag{2.20}$$

Then for large enough q_α, there is a constant C such that

$$\sum_{i=1}^{I}(q_i^2 + m^2) = \sum_{i=1}^{I}\left[\left(\sum_{\alpha=1}^{L} c_{i\alpha}q_\alpha + \sum_{j=1}^{E} c_{ij}p_j\right)^2 + m^2\right] \leq C\sum_{\alpha=1}^{L} q_\alpha^2 \ , \tag{2.21}$$

as we can easily see. Then we have

$$I_D > \frac{1}{C^I}\int_{\sum q_\alpha^2 > \Lambda^2}\prod_{\alpha=1}^{L}\frac{d^dq}{(2\pi)^d}\frac{1}{(\sum_{\alpha=1}^{L} q_\alpha^2)^I} > \int_{r>\Lambda}\frac{r^{DL-1}dr}{r^{2I}} \ , \tag{2.22}$$

where we used the fact that $\sum_{\alpha=1}^{L} q_\alpha^2 \equiv \sum_{M=1}^{dL} q_M^2$ is a sum of dL terms, which we can consider as a dL-dimensional space, and the condition $\sum_\alpha q_\alpha^2 > \Lambda^2$, stated before as q_α being large enough, now becomes the fact that the modulus of the dL-dimensional q_M is bounded from below. We finally see that if $\omega(D) = dL - 2I > 0$, I_D is divergent. The opposite statement is that if I_D is convergent, then $\omega(D) < 0$, i.e. $\omega(D) < 0$ is a necessary condition for convergence. q.e.d.

As we said, the condition is necessary, but not sufficient. Indeed, we can have subdiagrams that are superficially divergent ($\omega(D_s) \geq 0$), therefore divergent, then the full diagram is also divergent, in spite of having $\omega(D) < 0$.

We can take an example in $\lambda\phi^3$ theory in $D = 4$ the one in Fig.6: a circle with 3 external lines connected with it, with a subdiagram D_s connected to the inside of the circle: a propagator line that has a loop made out of two propagators in the middle. The diagram has $I_D = 9, V_D = 7$, therefore $L_D = I_D - V_D + 1 = 9 - 7 + 1 = 3$, and then $\omega(D) = dL_D - 2I_D = 4\cdot 3 - 2\cdot 9 = -6 < 0$. However, the subdiagram has $I_{D_s} = 2$, $V_{D_s} = 2$, therefore $L_{D_s} = 2 - 2 + 1 = 1$, and then $\omega(D_s) = 4\cdot 1 - 2\cdot 2 = 0$, therefore we have a logarithmically divergent subdiagram, and therefore the full diagram is also logarithmically divergent.

We can then guess that we have the following

Theorem (which we will not prove here) $\omega(D_s) < 0$, $\forall D_s$ 1PI subdiagrams of $D \Leftrightarrow I_D(p_1, ..., p_E)$ is an absolutely convergent integral.

We note that the $\Rightarrow$ implication should obviously be true, and moreover is valid for any field theory. But the $\Leftarrow$ implication is true only for scalar theories. If there are spin 1/2 and spin 1 fields, then $\omega(D) < 0$ is not even necessary, since there can be cancellations between different spins, giving a zero result for a (sum of) superficially divergent diagram(s)

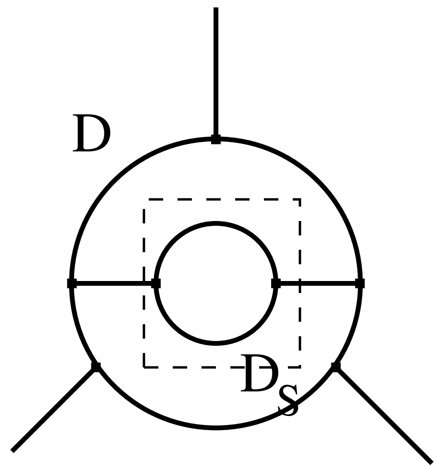

Figure 6: Power counting example: diagram is power counting convergent, but subdiagram is actually divergent.

(hence the name *superficial* degree of divergence, it is not necessarily the *actual* degree of divergence).

We can now write a formula from which we derive a condition on the type of divergencies we can have.

We note that each internal line connects to two vertices, and each external line connects only one vertex. In a theory with $\sum_n \lambda_n \phi^n/n!$ interactions, we can have a number $n = n_v$ of legs at each vertex v, meaning we have

$$2I + E = \sum_{v=1}^{V} n_v. \tag{2.23}$$

Then the superficial degree of divergence is

$$\omega(D) \equiv dL - 2I = (d-2)I - dV + d = d - \frac{d-2}{2}E + \sum_{v=1}^{V}\left(\frac{d-2}{2}n_v - d\right) , \tag{2.24}$$

where in the second equality we used $L = I - V + 1$ and in the last equality we have used $2I + E = \sum_v n_v$.

Since the kinetic term for a scalar is $-\int d^d x (\partial_\mu \phi)^2/2$, and it has to be dimensionless, we need the dimension of ϕ to be $[\phi] = (d-2)/2$. Then since the interaction term is $-\int d^d x \lambda_n \phi^n/n!$, we have

$$[\lambda_{n_v}] = d - n_v[\phi] = d - \frac{d-2}{2}n_v , \tag{2.25}$$

meaning that finally

$$\omega(D) = d - \frac{d-2}{2}E - \sum_{v=1}^{V}[\lambda_v]. \tag{2.26}$$

Thus we find that if $[\lambda_v] \geq 0$, there are only a finite number (and very small) of divergent n-point functions (where $n = E$ is the number of external lines). If $[\lambda_v] > 0$, then there are also a finite number of diagrams. Indeed, first, we note that by increasing the number of external lines, we get to $\omega(D) < 0$. Then, if $[\lambda_v] > 0$ we get to $\omega(D) < 0$ also by increasing V, the number of vertices.

As an example, consider the limiting case of $[\lambda_v] = 0$, and $d = 4$. Then $\omega(D) = 4 - E$, and only $E = 0, 1, 2, 3, 4$ give divergent results, irrespective of V. Since $E = 0, 1$ are not physical ($E = 0$ is a vacuum bubble, and $E = 1$ should be zero in a good theory), we have only $E = 2, 3, 4$, corresponding to 3 physical parameters (physical parameters are defined by the number of external lines, which define physical objects like n-point functions). For $[\lambda_v] > 0$, any vertices lower $\omega(D)$, so we could have even a smaller number of E's for divergent n-point functions, since we need at least a vertex for a loop diagram. In higher dimensions, we will have a slightly higher number of divergent n-point functions, but otherwise the same idea applies.

Such theories, where there are a finite number of n-point functions that have divergent diagrams, is called *renormalizable*, since we can absorb the infinities in the redefinition of the (finite number of) parameters of the theory, the parameters being in one to one correspondence with the *1PI* n-point functions.

Therefore we have 3 possibilities:

A) If there is some $[\lambda_v] < 0$, we can make divergent diagrams for any n-point function (any E) with λ_v vertices just by increasing V. Therefore there are an infinite number of divergent 1PI n-point functions that would need redefinition, so we can't make this by redefining the parameters of the theory. Such a theory is called *nonrenormalizable*. Note that a nonrenormalizable theory can be so only in perturbation theory, there exist examples of theories that are perturbatively nonrenormalizable, but the nonperturbative theory is well-defined. Also note that we can work with nonrenormalizable theories in perturbation theory, just by introducing new parameters at each loop order. Therefore we can compute quantum corrections, though the degree of complexity of the theory quickly increases with the loop order.

B) If there is some $[\lambda_v] = 0$, and the rest of $[\lambda_v] > 0$, that means that there is a finite number of 1PI divergent n-point functions, which means that the theory is renormalizable.

C) If all $[\lambda_v] > 0$, there are only a finite number of *diagrams* that are divergent, which means that we can actually fully renormalize the theory in practice, not just in principle. The theory is then called super-renormalizable.

Examples.

Consider scalar field theories with all power laws, $\sum_{n\geq 1} \lambda_n \phi^n/n!$. Then the condition for the theory to be renormalizable is $[\lambda_n] \geq 0$, which gives

$$n \leq \frac{2d}{d-2}. \tag{2.27}$$

d=2. Then in 2 dimensions all n are renormalizable, in fact super-renormalizable, since then $[\lambda_n] = 2$, and $\omega(D) = 2 - 2V$ becomes < 0 as V is increased.

d=3. The above condition gives $n \leq 6$, with equality for $n = 6$, which means that ϕ^3, ϕ^4, ϕ^5 are super-renormalizable and ϕ^6 is just renormalizable.

d=4. The condition above gives $n \leq 4$, with equality for $n = 4$, which means that ϕ^3 is super-renormalizable and ϕ^4 is just renormalizable.

d=5. The condition above gives $n \leq 10/3$, which means that only ϕ^3 is renormalizable, and actually is super-renormalizable.

d=6. The condition gives $n \leq 3$, which means that only ϕ^3 is renormalizable.

d>6. There are no renormalizable interactions.

Divergent ϕ^4 1PI diagrams in various dimensions.

In $d = 2$, $[\lambda_v] = 2$ and $\omega(D) = 2 - 2V$, which means that only the $V = 1$ diagram is divergent (the 2-point one-loop diagram).

In $d = 3$, $[\lambda_v] = 1$ and $\omega(D) = 3 - E/2 - V$, which means that the only 1PI diagrams are the $V = 1, E = 2$ diagram (one-loop) and the $V = 2, E = 2$ diagram (2-loops).

In $d = 4$, $[\lambda_v] = 0$, so $\omega(D) = 4 - E$, so all the diagrams of the 2,3 and 1PI 4-point functions are divergent.

Important concepts to remember

- Loop diagrams can contain UV divergences (at high momenta), divergent in diagram-dependent dimensions, and IR divergences, which appear only for massless theories and for on-shell total external momenta at vertices.

- UV divergences can be absorbed in a redefinition of the parameters of the theory (renormalization), and IR divergences can be cancelled by adding the tree diagrams for emission of low momentum ($E \simeq 0$) particles, perhaps parallel to the original external particle.

- Wick rotation of the result of the Euclidean integrals can in general not be the same as Wick rotation of the Euclidean integral, since there can be poles in between the Minkowskian and Euclidean contours for the loop energy integration. We can work in Euclidean space and continue the final result, since the Euclidean theory is better defined.

- Power counting gives the superficial degree of divergence of a diagram as $\omega(D) = dL - 2I$.

- In a scalar theory, $\omega(D) < 0$ is necessary for convergence of the integral I_D, but in general is not sufficient.

- In a scalar theory, $\omega(D_s) < 0$ for any 1PI subdiagram D_s of a diagram $D \Leftrightarrow I_D$ is absolutely convergent.

- Theories with couplings satisfying $[\lambda_v] \geq 0$ are renormalizable, i.e. one can absorb the infinities in redefinitions of the parameters of the theory, while theories with $[\lambda_v] < 0$ are nonrenormalizable, since we can't (there are an infinite number of different infinities to be absorbed).

- If all the $[\lambda_v] > 0$, the theory is super-renormalizable: it has only a finite number of divergent 1PI diagrams.

Further reading: See chapters 5.1,5.2 in [5], 9.1 in [2] and 4.2 in [4].

Exercises, Lecture 2

1) Consider the one-loop diagram below, for arbitrary masses of the various lines. Check whether there are any divergences.

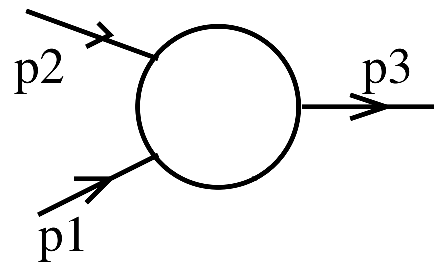

Figure 7: One loop Feynman diagram. Check for divergences.

2) Check whether there are any UV divergences in the $D = 3$ diagram in $\lambda\phi^4$ theory in Fig.8.

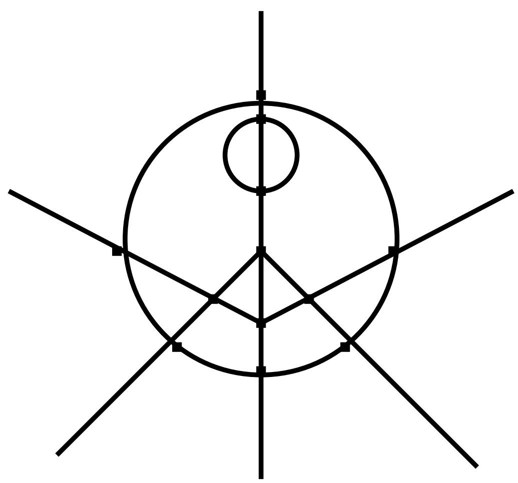

Figure 8: Check for UV divergences in this Feynman diagram.

3) Consider the Lagrangeans

$$\mathcal{L}_1 = +\frac{1}{2}(\partial_\mu\phi)^2 + g\phi^2(\partial_\mu\phi)^2 \tag{2.28}$$

in 4 Euclidean dimensions, and

$$\mathcal{L}_2 = \frac{1}{2\kappa_N^2}\sqrt{g}R + \frac{1}{2}\sqrt{g}\partial_\mu\phi\partial_\nu\phi g^{\mu\nu} + \sqrt{g}\frac{m^2}{2}e^{\alpha^2\phi^2} \tag{2.29}$$

in 2 Euclidean dimensions, where R is the Ricci scalar for gravity.

Are these Lagrangeans renormalizable, superrenormalizable, or non-renormalizable? If they are renormalizable or superrenormalizable, write down the superficially divergent diagrams.

3 Lecture 3. Regularization, definitions: cut-off, Pauli-Villars, dimensional regularization, general Feynman parametrization

.

In this lecture we will describe methods of regularization, that will be used for renormalization. We already saw a few methods of *regularization*, i.e. making finite the integrals.

Cut-off regularization

The simplest is *cut-off regularization*, which means just putting upper and lower bounds on the integral over the modulus of the momenta, i.e. a $|p|_{max} = \Lambda$ for the UV divergence, and an $|p|_{min} = \epsilon$ for the IR divergence. It has to be over the modulus only (the integral over angles is not divergent), and then the procedure works best in Euclidean space (since then we don't need to worry about the fact that $-(p_0)^2 + \vec{p}^2 = \Lambda^2$ has a continuum of solutions for Minkowski space). Note that having a $|p|_{max} = \Lambda$ is more or less the same as considering a lattice of size Λ^{-1} in Euclidean space, which breaks Euclidean ("Lorentz") invariance (since translational and rotational invariance are thus broken). For this reason, very seldom we consider the cut-off regularization.

There are many regularizations possible, and in general we want to consider a regularization that preserves all symmetries that play an important role at the quantum level. If there are several that preserve the symmetries we want, all of them can in principle be used (we could even consider cut-off regularization, just that then we would have a hard time showing that our results are consistent with the symmetries we want to preserve).

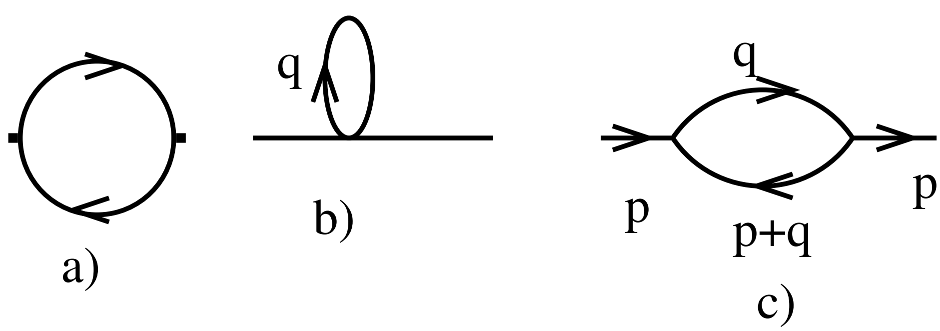

Figure 9: Examples of one-loop divergent diagrams.

Let's see the effect of cut-off regularization on the simplest diagram we can write, a loop for a massless field, with no external momentum, but two external points, i.e. two equal loop propagators inside (see Fig.9a):

$$\int \frac{d^4p}{(2\pi)^4} \frac{1}{(p^2)^2} = \frac{\Omega_3}{(2\pi)^4} \int_\epsilon^\Lambda \frac{p^3 dp}{p^4} = \frac{1}{8\pi^2} \ln \frac{\Lambda}{\epsilon}. \tag{3.1}$$

As we said, we see that this has both UV (for $\Lambda \to \infty$) and IR (for $\epsilon \to 0$) divergences.

Infinite sums

Sometimes, we can turn integrals into infinite sums, like in the case of the zero point energy, where a first regularization, putting the system in a box of size L, allows us to get a sum instead of an integral: $\sum_n \hbar\omega_n/2$. The sum however is still divergent. When one calculates the one-loop fluctuations around some classical background, the action is approximated by the classical part plus a quadratic fluctuation,

$$S \simeq S_{\rm cl} + \frac{1}{2}\int \delta\phi \left[-\partial^\mu\partial_\mu + U''(\phi_0)\right]\delta\phi \;, \tag{3.2}$$

where $\delta\phi = \phi - \phi_0$, ϕ_0 is the classical value for the field, giving the on-shell action $S_{\rm cl}$, and $U(\phi)$ is the potential. Then the partition function is found to be

$$Z[J] = \mathcal{N} \overset{-1/2}{\det}\left[(\partial_\mu\partial^\mu + U''(\phi_0))\delta_{12}\right] = ... = \sum_n e^{-\frac{iE_nT}{\hbar}} \;, \tag{3.3}$$

where the constant $\mathcal{N}$ includes $e^{-S_{\rm cl}}$, and the ... involve some calculations that will not be reproduced here. The zero-point energy E_0 is given by

$$E_0 = \frac{\hbar}{2}\sum_n \omega_n \;, \tag{3.4}$$

where ω_n are given by the eigenvalues of the spatial part of the quadratic operator,

$$(-\vec{\nabla}^2 + U''(\phi_0))\eta_r(\vec{x}) = \omega_r^2\eta_r(\vec{x}). \tag{3.5}$$

Then there were various ways of regularizing the sum. We can deduce that if the result depends on the method of regularization, there are two possibilities:

- perhaps the result is unphysical (can't be measured). This is what happens to the *full* zero-point energy. Certainly the infinite constant (which for instance in an $e^{-a\omega_n}$ regularization is something depending only on L) is not measurable.
- perhaps some of the methods of regularization are unphysical, so we have to choose the physical one.

An example of a physical calculation is the *difference* of two infinite sums. This is what usually happens. In the example of the Casimir effect (the attractive force between two infinite conducting plates due to the vacuum energy), the difference between two geometries (two values for the distance d between two parallel plates) gives a physical, measurable quantity. Another example is the difference between two vacua, say a vacuum with a soliton and a trivial vacuum, with no soliton. In general then, we will have

$$\sum_n \frac{\hbar\omega_n^{(1)}}{2} - \sum_n \frac{\hbar\omega_n^{(2)}}{2}. \tag{3.6}$$

For the case of the quantum corrections to masses of solitons, one considers the difference between quantum fluctuations ($\sum_n \hbar\omega_n/2$) in the presence of the soliton, and in the vacuum (without the soliton), and this gives the physical calculation of the quantum correction to the mass of the soliton.

Note that it would seem that we could write

$$\sum_n \left(\frac{\hbar\omega_n^{(1)} - \hbar\omega_n^{(2)}}{2} \right) \tag{3.7}$$

and calculate this, but this amounts to a choice of regularization, called mode number (n) regularization. Indeed, we have now $\infty - \infty$, so unlike the case of finite integrals, now giving the $\sum_n$ operator as a common factor is only possible if we choose the *same number N as the upper bound* in both sums (if one is N, and the other $N + a$ say, with $a \sim \mathcal{O}(1)$, then obviously we obtain a different result for the difference of two sums).

This may seem natural, but there is more than one other way to calculate: for instance, we can turn the sums into integrals in the usual way, and then take the *same upper limit in the integral (i.e., in energy)*, obtaining *energy/momentum cut-off regularization.* The difference of the two integrals gives a result differing from mode number cut-off by a finite piece.

For $\sum_n \hbar\omega_n/2$, there are other regularizations as well: for instance zeta function regularization, heat kernel regularization, and $\sum_n \omega_n \to \sum_n \omega_n e^{-a\omega_n}$.

Pauli-Villars regularization.

Returning to the loop integrals appearing in Feynman diagrams, we have other ways to regulate. One of the oldest ways used is *Pauli-Villars regularization*, and its generalizations. These fall under the category of modifications to the propagator that cut off the high momentum modes in a smoother way than the hard cut-off $|p|_{max} = \Lambda$. The generalized case corresponds to making the substitution

$$\frac{1}{q^2+m^2} \to \frac{1}{q^2+m^2} - \sum_{i=1}^{N} \frac{c_i(\Lambda; m^2)}{q^2+\Lambda_i^2}\ , \tag{3.8}$$

and we can adjust the c_i such that at large momentum q, the redefined propagator behaves as

$$\sim \frac{\Lambda^{2N}}{q^{2N+2}}. \tag{3.9}$$

In other words, if in a loop integral, we need to have the propagator at large momentum behave as $1/q^{2N+2}$ in order for the integral to become finite, we choose a redefinition with the desired N, and with c_i's chosen for the required high momentum behaviour.

In particular, the original Pauli-Villars regularization is

$$\frac{1}{q^2+m^2} \to \frac{1}{q^2+m^2} - \frac{1}{q^2+\Lambda^2} = \frac{\Lambda^2 - m^2}{(q^2+m^2)(q^2+\Lambda^2)} \sim \frac{\Lambda^2}{q^4}. \tag{3.10}$$

We can easily see that it cannot be obtained from a normal modification of the action, because of the minus sign, however it corresponds to subtracting the contribution of a very heavy particle. Indeed, physically it is clear that a heavy particle cannot modify anything physical (for instance, a Planck mass particle cannot influence Standard Model physics). But it is equally obvious that subtracting its contribution will cancel heavy momentum modes in the loop integral, cancelling the unphysical infinities of the loop.

However, there is a simple modification that has the same result as the above Pauli-Villars subtraction at high momentum, and has a simple physical interpretation as the effect of a higher derivative term in the action. Specifically, consider the replacement of the propagator

$$\frac{1}{q^2+m^2} \to \frac{1}{q^2+m^2+q^4/\Lambda^2}. \tag{3.11}$$

The usual propagator comes from

$$\int d^4x \frac{1}{2}(\partial_\mu\phi)^2 = \int \frac{d^4p}{(2\pi)^4}\frac{1}{2}\phi(p)p^2\phi(-p) = \int \frac{d^4p}{(2\pi)^4}\frac{1}{2}\phi(p)\Delta^{-1}(p)\phi(-p) \ , \tag{3.12}$$

so the above replacement is obtained by adding to the action a higher derivative term:

$$\int d^4x \frac{1}{2}\left[(\partial_\mu\phi)^2 + \frac{(\partial^2\phi)^2}{\Lambda^2}\right] = \int \frac{d^4p}{(2\pi)^4}\phi(p)\left[p^2+\frac{(p^2)^2}{\Lambda^2}\right]\phi(-p). \tag{3.13}$$

Now consider a non-Pauli-Villars, but similar modification of the the loop integral, that is strictly speaking not a modification of the propagator, but of its square. Consider the same simplest loop integral, with two equal propagators, i.e.

$$I = \int \frac{d^4p}{(2\pi)^4}\frac{1}{(p^2+m^2)^2} \to \int \frac{d^4p}{(2\pi)^4}\left[\left(\frac{1}{p^2+m^2}\right)^2 - \left(\frac{1}{p^2+\Lambda^2}\right)^2\right] = I(m^2) - I(\Lambda^2). \tag{3.14}$$

The new object in the square brackets is

$$\frac{2p^2(\Lambda^2-m^2)+\Lambda^4-m^4}{(p^2+m^2)^2(p^2+\Lambda^2)^2} \sim \frac{2\Lambda^2}{p^6} \ , \tag{3.15}$$

so is now UV convergent. Since the object is UV convergent, we can use any method to calculate it. In particular, we can take a derivative $\partial/\partial m^2$ of it, and since $I(\Lambda^2)$ doesn't contribute, we get for the integral

$$\frac{\partial}{\partial m^2}[I(m^2)-I(\Lambda^2)] = \int \frac{d^4p}{(2\pi)^4}\frac{-2}{(p^2+m^2)^3} = -2\frac{\Omega_3}{(2\pi)^4}\int_0^\infty \frac{p^3dp}{(p^2+m^2)^3} \ , \tag{3.16}$$

where $\Omega_3 = 2\pi^2$ is the volume of the 3 dimensional unit sphere. Considering $p^2+m^2 = x$, so $p^3dp = (x-m^2)dx/2$, we get

$$\frac{\partial}{\partial m^2}I(m^2,\Lambda^2) = -\frac{1}{8\pi^2}\int_{m^2}^\infty \frac{(x-m^2)dx}{x^3} = -\frac{1}{8\pi^2}\frac{1}{2m^2}. \tag{3.17}$$

Integrating this, we obtain $I(m^2)$, then

$$I(m^2,\Lambda^2) = I(m^2) - I(\Lambda^2) = \frac{1}{16\pi^2}\ln\frac{\Lambda^2}{m^2}. \tag{3.18}$$

This object is UV divergent, as $\Lambda \to \infty$, and also divergent as $m \to 0$ (IR divergent).

Derivative regularization

However, note that in the way we calculated, we really introduced another type of regularization. It was implicit, since we first found a finite result by subtracting the contribution with $m \to \Lambda$, and then calculated this finite result using what was a simple trick.

However, if we keep the original integral and do the same derivative on it, after the derivative we obtain a finite result,

$$\frac{\partial}{\partial m^2} \int \frac{d^4p}{(2\pi)^4} \frac{1}{(p^2+m^2)^2} = -2 \int \frac{d^4p}{(2\pi)^4} \frac{1}{(p^2+m^2)^3} = -\frac{1}{16\pi^2 m^2} = \text{finite.} \tag{3.19}$$

and now the integral (and its result) is UV convergent, despite the integral before the derivative being UV divergent. Hence the derivative with respect to the parameter m^2 is indeed a regularization. Both the initial and final results are however still IR divergent as $m^2 \to 0$. Now integrating, we obtain

$$\int \frac{d^4p}{(2\pi)^4} \frac{1}{(p^2+m^2)^2} = -\frac{1}{16\pi^2} \ln \frac{m^2}{\epsilon} + \text{constant} , \tag{3.20}$$

which is still IR divergent as $\epsilon \to 0$. The UV divergence is now hidden as a possible infinite integration constant.

However, all the regularizations we analyzed until now don't respect a very important invariance, namely gauge invariance. Therefore, 't Hooft and Veltman introduced a new regularization to deal with the spontaneously broken gauge theories, namely *dimensional regularization*, rather late (early seventies), since it is a rather strange concept.

Dimensional regularization

Dimensional regularization means that we analytically continue in D, the dimension of spacetime, results calculated for arbitrary D. This seems like a strange thing to do, given that the dimension of spacetime is an integer, so it is not clear what can physically mean a real dimension, but we nevertheless choose $D = 4 \pm \epsilon$. The sign has some significance as well, but here we will just consider $D = 4 + \epsilon$.

A relevant example for us, that will not only encode the features of dimensional regularization, but will actually be the only way to obtain infinities in dimensional regularization, is the case of the Euler gamma function, which is an extension of the factorial, $n!$, defined for integers, to the complex plane. Again this is done by writing an integral formula,

$$\Gamma(z) = \int_0^\infty d\alpha \alpha^{z-1} e^{-\alpha}. \tag{3.21}$$

Indeed, one easily shows that $\Gamma(n) = (n-1)!$, for $n \in \mathbb{N}_*$, but the integral formula can be extended to the complex plane, defining the Euler gamma function. The gamma function satisfies

$$z\Gamma(z) = \Gamma(z+1), \tag{3.22}$$

an extension of the factorial property. But that means that we can find the behaviour at $z = \epsilon \to 0$, which is a simple pole, since

$$\epsilon\Gamma(\epsilon) = \Gamma(1+\epsilon) \simeq \Gamma(1) = 1 \Rightarrow \Gamma(\epsilon) \simeq \frac{1}{\epsilon}. \tag{3.23}$$

We can repeat this process,

$$
\begin{aligned}
(-1+\epsilon)\Gamma(-1+\epsilon) = \Gamma(\epsilon) \Rightarrow \Gamma(-1+\epsilon) \simeq -\frac{1}{\epsilon} \\
(-2+\epsilon)\Gamma(-2+\epsilon) = \Gamma(-1+\epsilon) \Rightarrow \Gamma(-2+\epsilon) \simeq \frac{1}{2\epsilon}\ ,
\end{aligned}
\tag{3.24}
$$

etc. We see then that the gamma function has simple poles at all $\Gamma(-n)$, with $n \in \mathbb{N}$. In fact, these poles are exactly the one we obtain in dimensional regularization, as we now show.

Consider first the simplest case, the tadpole diagram, with a single loop, of momentum q, connected at a point to a propagator line, as in Fig.9b:

$$
I = \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2+m^2}. \tag{3.25}
$$

We now write

$$
\frac{1}{q^2+m^2} = \int_0^\infty d\alpha e^{-\alpha(q^2+m^2)}\ , \tag{3.26}
$$

and then

$$
\begin{aligned}
I &= \int_0^\infty d\alpha e^{-\alpha m^2} \int \frac{d^D q}{(2\pi)^D} e^{-\alpha q^2} = \int_0^\infty d\alpha e^{-\alpha m^2} \frac{\Omega_{D-1}}{(2\pi)^D} \int_0^\infty dq q^{D-1} e^{-\alpha q^2} \\
&= \int_0^\infty d\alpha e^{-\alpha m^2} \frac{\Omega_{D-1}}{(2\pi)^D} \frac{1}{2\alpha^{D/2}} \int_0^\infty dx x^{\frac{D}{2}-1} e^{-x}\ ,
\end{aligned}
\tag{3.27}
$$

and we use the fact that $\int_0^\infty dx x^{D/2-1} e^{-x} = \Gamma(D/2)$, and that the volume of the D-dimensional sphere is

$$
\Omega_D = \frac{2\pi^{\frac{D+1}{2}}}{\Gamma\left(\frac{D+1}{2}\right)}, \tag{3.28}
$$

which we can easily test on a few examples, $\Omega_1 = 2\pi/\Gamma(1) = 2\pi$, $\Omega_2 = 2\pi^{3/2}/\Gamma(3/2) = 2\pi^{3/2}/(\sqrt{\pi}/2) = 4\pi$, $\Omega_3 = 2\pi^2/\Gamma(2) = 2\pi^2$. Then we have $\Omega_{D-1}\Gamma(D/2) = 2\pi^{D/2}$, so

$$
I = \int_0^\infty d\alpha e^{-\alpha m^2} (4\pi\alpha)^{-\frac{D}{2}} = \frac{(m^2)^{\frac{D}{2}-1}}{(4\pi)^{\frac{D}{2}}} \Gamma\left(1-\frac{D}{2}\right). \tag{3.29}
$$

Taking derivatives $(\partial/\partial m^2)^{n-1}$ on both sides (both the definition of I and the result), we obtain in general

$$
\int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2+m^2)^n} = \frac{\Gamma(n-d/2)}{(4\pi)^n \Gamma(n)} \left(\frac{m^2}{4\pi}\right)^{\frac{D}{2}-n}. \tag{3.30}
$$

We see that in $D = 4$, this formula has a pole at $n = 1, 2$, as expected from the integral form. In these cases, the divergent part is contained in the gamma function, namely $\Gamma(-1)$ and $\Gamma(0)$.

Feynman parametrization with two propagators

We now move to a more complicated integral, which we will solve with Feynman parametrization, cited last lecture. Specifically, we consider the diagram for a one-loop correction to the propagator in ϕ^3 theory, with momentum p on the propagator and q and $p+q$ in the loop, as in Fig.9c, i.e.

$$\int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2+m^2} \frac{1}{(q+p)^2+m^2}. \tag{3.31}$$

We now prove the Feynman parametrization in this case of two propagators. We do the trick used in the first integral (tadpole) twice, obtaining

$$\frac{1}{\Delta_1\Delta_2} = \int_0^\infty d\alpha_1 \int_0^\infty d\alpha_2 e^{-(\alpha_1\Delta_1+\alpha_2\Delta_2)}. \tag{3.32}$$

We then change variables in the integral as $\alpha_1 = t(1-\alpha)$, $\alpha_2 = t\alpha$, with Jacobian $\left|\begin{pmatrix} 1-\alpha & \alpha \\ -t & t \end{pmatrix}\right| = t$, so

$$\frac{1}{\Delta_1\Delta_2} = \int_0^1 d\alpha \int_0^\infty dt\; te^{-t[(1-\alpha)\Delta_1+\alpha\Delta_2]} = \int_0^1 d\alpha \frac{1}{[(1-\alpha)\Delta_1+\alpha\Delta_2]^2}. \tag{3.33}$$

We want to write the square bracket as a new propagator, so we redefine $q^\mu = \tilde{q}^\mu - \alpha p^\mu$, obtaining

$$(1-\alpha)\Delta_1 + \alpha\Delta_2 = (1-\alpha)q^2 + \alpha(q+p)^2 + m^2 = \tilde{q}^2 + m^2 + \alpha(1-\alpha)p^2. \tag{3.34}$$

Finally, we obtain for the integral

$$I = \int_0^1 d\alpha \int \frac{d^D\tilde{q}}{(2\pi)^D} \frac{1}{[\tilde{q}^2 + (\alpha(1-\alpha)p^2+m^2)]^2} = \frac{\Gamma(2-D/2)}{(4\pi)^{D/2}} \int_0^1 d\alpha [\alpha(1-\alpha)p^2+m^2]^{\frac{D}{2}-1}, \tag{3.35}$$

and again we obtained the divergence as just an overall factor coming from the simple pole of the gamma function at $D=4$, and the integral in now finite.

General one-loop integrals and Feynman parametrization.

The Feynman parametrization can in fact be generalized, to write instead of a product of propagators, a single "propagator" raised at the corresponding power, with integrals over parameters α left to do. Then we can use the formula (3.30) to calculate the momentum integral, and we are left with the integration over parameters.

Consider the general one-loop integral from Fig.10, obtained from a loop with n external lines coming out of it, with momenta $p_1, ..., p_n$. The momenta on the internal lines can be chosen to be $q + \tilde{p}_i$, with $p_i = \sum_{j=1}^i p_i$, for instance, but in any case we can always write them as $q + \tilde{p}_i$, with $\tilde{p}_i$ depending on the external momenta p_i. The general one-loop scalar integral is then

$$I(p_1, ..., p_n) = \int \frac{d^D q}{(2\pi)^D} \prod_{i=1}^n \frac{1}{(q+\tilde{p}_i)^2+m^2}. \tag{3.36}$$

We first write the product of propagators in an exponential as before, as

$$\frac{1}{\Delta_1 ... \Delta_n} = \int_0^\infty \prod_{i=1}^n d\tilde{\alpha}_i e^{-(\tilde{\alpha}_1\Delta_1 + ... + \tilde{\alpha}_n\Delta_n)}. \tag{3.37}$$

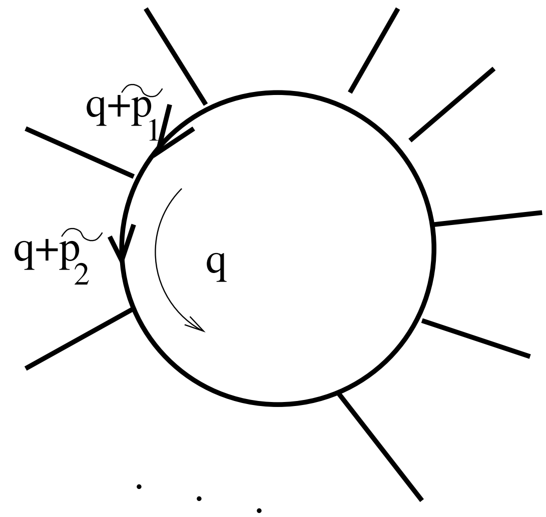

Figure 10: General One-Loop Feynman diagram.

After a transformation $\tilde{\alpha}_i = \alpha_i t$, which needs to be supplemented with the constraint $\sum_i \alpha_i = 1$ in order to still have n independent integration variables, which also means that we must have $\alpha_i \leq 1$, and whose Jacobian we can easily check is $J = t^{n-1}$ (the determinant has $n-1$ t's on the diagonal and rest zeroes, except on the last line), the product of propagators becomes

$$\begin{aligned} & \int_0^1 \prod_{i=1}^n d\alpha_i \delta\left(1 - \sum_i \alpha_i\right) \int_0^\infty dt\ t^{n-1} e^{-t(\alpha_1\Delta_1 + \ldots + \alpha_n\Delta_n)} \\ = & \int_0^1 \prod_{i=1}^n d\alpha_i \delta\left(1 - \sum_i \alpha_i\right) \frac{\Gamma(n)}{(\alpha_1\Delta_1 + \ldots + \alpha_n\Delta_n)^n}. \end{aligned} \tag{3.38}$$

Now we have replaced the product of different propagators with a single quadratic expression in momenta, raised to the n-th power, so we can use the formula (3.30) to calculate the one-loop integral, which becomes

$$I(p_1, \ldots, p_n) = \int \prod_{i=1}^n d\alpha_i \delta\left(1 - \sum_i \alpha_i\right) \int \frac{d^D q}{(2\pi)^D} \frac{\Gamma(n)}{[\sum_{i=1}^n \alpha_i (q + \tilde{p}_i)^2 + m^2]^n}. \tag{3.39}$$

But there is still one step to do, namely to shift the momenta in order to get rid of the term linear in momenta: Define $\tilde{q}^\mu = q^\mu + \sum_i \alpha_i \tilde{p}^i$, after which the quadratic form in the denominator becomes

$$\sum_{i=1}^n [\alpha_i (q + \tilde{p}_i)^2 + m^2] = \tilde{q}^2 + m^2 + \sum_i \alpha_i \tilde{p}_i^2 - \left(\sum_i \alpha_i \tilde{p}_i\right)^2. \tag{3.40}$$

Since the Jacobian of the shift is 1, finally we obtain (using (3.30)) for the Feynman parametrization of the general one-loop integral

$$I(p_1,...,p_n)=\frac{\Gamma(n-D/2)}{(4\pi)^{D/2}}\int_0^1\prod_{i=1}^n d\alpha_i\delta\left(1-\sum_i\alpha_i\right)\left[m^2+\sum_i\alpha_i\tilde{p}_i^2-\left(\sum_i\alpha_i\tilde{p}_i\right)^2\right]^{D/2-n}. \tag{3.41}$$

The α_i are called Feynman parameters. We note that the dimensional regularization procedure described above is general, and we see that the divergence always appears as the simple pole of the gamma function at $D=4$.

Alternative version of the Feynman parametrization.

In the literature, an alternative parametrization is also sometimes used, which can be found directly by making a change of variables in (3.38), but we find it useful to start at the beginning. In (3.37), we substitute the change of variables $\tilde{\alpha}_1=t(1-\alpha_1)$, $\tilde{\alpha}_2=t(\alpha_1-\alpha_2)$, $\tilde{\alpha}_3=t(\alpha_2-\alpha_3)$, ..., $\tilde{\alpha}_n=t\alpha_n$. From these definitions, we see that we must have $\alpha_1\in[0,1]$, $\alpha_2\in[0,\alpha_1]$, $\alpha_3\in[0,\alpha_2]$, ... The Jacobian of the transformation is found again to be $J=t^{n-1}$, so that we get

$$\begin{aligned}\frac{1}{\Delta_1...\Delta_n} &= \int_0^1 d\alpha_1\int_0^{\alpha_1}d\alpha_2\int_0^{\alpha_2}d\alpha_3...\int_0^{\alpha_{n-2}}d\alpha_{n-1}\int_0^\infty dt\\ &t^{n-1}e^{-t((1-\alpha_1)\Delta_1+(\alpha_2-\alpha_1)\Delta_2+...+(\alpha_{n-2}-\alpha_{n-1})\Delta_{n-1}+\alpha_{n-1}\Delta_n)}.\end{aligned} \tag{3.42}$$

Finally, we have for the product of propagators

$$\begin{aligned}\frac{1}{\Delta_1...\Delta_n} &= \Gamma(n)\int_0^1 d\alpha_1\int_0^{\alpha_1}d\alpha_2\int_0^{\alpha_2}d\alpha_3...\int_0^{\alpha_{n-2}}d\alpha_{n-1}\\ &[\Delta_1(1-\alpha_1)+\Delta_2(\alpha_1-\alpha_2)+...+\Delta_{n-1}(\alpha_{n-2}-\alpha_{n-1})+\Delta_n\alpha_{n-1}]^{-n}.\end{aligned} \tag{3.43}$$

Dimensionally continuing Lagrangeans.

So we saw that the loop integrals of the above type are OK to dimensionally continue, but is it OK to dimensionally continue the Lagrangeans?

For scalars, it is OK, but we have to be careful. The Lagrangean is

$$\mathcal{L}=\frac{1}{2}(\partial_\mu\phi)^2+\frac{m^2}{2}\phi^2+\frac{\lambda_n}{n!}\phi^n \tag{3.44}$$

and since the action $S=\int d^Dx$ must be dimensionless, the dimension of the scalar is $[\phi]=(D-2)/2$ and thus the dimension of the coupling is $[\lambda_n]=D-n(D-2)/2$. For instance, for $D=4$ and $n=4$, we have $[\lambda_4]=0$, but for $D=4+\epsilon$, we have $[\lambda_4]=-\epsilon$. That means that outside $D=4$, we must redefine the coupling with a factor μ^ϵ, where μ is some scale that appears dynamically. This process of spontaneous appearance of an arbitrary mass scale at the quantum level is called *dynamical transmutation*, and is related to the fact that we have a renormalization group, as we will see later in the course.

For higher spins however, we must be more careful. The number of components of a field depends on dimension, which is a subtle issue. We must then use dimensional continuation of various gamma matrix formulae

$$g^{\mu\nu}g_{\mu\nu}=D$$

$$\gamma_\mu \not{p} \gamma^\mu = (2-D)\not{p} \tag{3.45}$$

etc. On the other hand, the gamma matrices still satisfy the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. But the dimension of the (spinor) representation of the Clifford algebra depends on dimension in an unusual way, $n = 2^{[D/2]}$, which means it is 2 dimensional in $D = 2, 3$ and 4 dimensional in $D = 4, 5$. That means that we cannot continue dimensionally n to $D = 4 + \epsilon$. Instead, we must still consider the gamma matrices as 4×4 even in $D = 4 + \epsilon$, and thus we still have

$$\mathrm{Tr}[\gamma_\mu \gamma_\nu] = 4g_{\mu\nu} \tag{3.46}$$

This is not a problem, however there is another fact that is still a problem. The definition of γ_5 is

$$\gamma_5 = \frac{i}{4!}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma = -i\gamma^0\gamma^1\gamma^2\gamma^3 = i\gamma_0\gamma_1\gamma_2\gamma_3 \tag{3.47}$$

and that cannot be easily dimensionally continued. Since chiral fermions, i.e. fermions that are eigenvalues of the chiral projectors $P_{L,R} = (1 \pm \gamma_5)/2$, appear in the Standard Model, we would need to be able to continue dimensionally chiral fermions. But that is very difficult to do.

Therefore we can say that there are no perfect regularization procedures, always there is something that does not work easily, but for particular cases we might prefer one or another.

Next class we will see that the divergences that we have regularized in this lecture can be absorbed in a redefinition of the parameters of the theory, leaving only a finite piece giving quantum corrections.

Important concepts to remember

- We must regularize the infinities appearing in loop integrals, and the infinite sums.
- Cut-off regularization, imposing upper and lower limits on $|p|$ in Euclidean space, regulates integrals, but is not very used because it is related to breaking of Euclidean ("Lorentz") invariance, as well as breaking of gauge invariance.
- Often the difference of infinite sums is a physical observable, and then the result is regularization-dependent. In particular, we can have mode-number cut-off (giving the sum operator as a common factor), or energy cut-off (giving a resulting energy integral as a common factor). We must choose one that is more physical.
- The choice of regularization scheme for integrals is dictated by what symmetries we want to preserve. If several respect the wanted symmetries, they are equally good.
- (Generalized) Pauli-Villars regularization removes the contribution of high energy modes from the propagator, by subtracting the propagator a very massive particle from it. A related version for it is obtained from a term in the action which is higher derivative.

- By taking derivatives with respect to a parameter (e.g. m^2), we obtain derivative regularization, which also reduces the degree of divergence of integrals.

- Dimensional regularization respects gauge invariance, and it corresponds to analytically continuing the dimension, as $D = 4 + \epsilon$. It is based on the fact that we can continue $n!$ away from the integers to the Euler gamma function.

- In dimensional regularization, the divergences are the simple poles of the gamma function at $\Gamma(-n)$, and appear as a multiplicative $1/\epsilon$.

- With Feynman parametrization, we can reduce a general one-loop scalar integral to an integral over Feynman parameters only.

- For scalars, dimensional regularization of the action is OK, if we remember that couplings have extra mass dimensions away from $D = 4$. For higher spins, we must regulate the number of components, including things like $g^{\mu\nu}g_{\mu\nu}$ and gamma matrix identities.

- The dimension of gamma matrices away from $D = 4$ is still 4, so traces of gamma matrices still give a factor of 4, and the γ_5 cannot be continued away from $D = 4$, which means analytical continuation in dimension that involve chiral fermions are very hard.

Further reading: See chapters 5.3 in [5] 4.3 in [4] and 9.3 in [2].

Exercises, Lecture 3

1) Calculate

$$\int \frac{d^4p}{(2\pi)^4} \frac{1}{p^2 + m_1^2} \frac{1}{p^2 + m_2^2} \tag{3.48}$$

in Pauli-Villars regularization.

2) Calculate

$$\int \frac{d^Dq}{(2\pi)^D} \frac{1}{(q^2 + m_1^2)^2} \frac{1}{(q+p)^2 + m_2^2} \tag{3.49}$$

in dimensional regularization (it is divergent in $D = 6$).

4 Lecture 4. One-loop renormalization for scalars and counterterms in dimensional regularization.

In this lecture we will learn how to get rid of the infinities appearing in quantum field theory loops, by a process called renormalization. Basically, we will absorb the infinities in a redefinition of the parameters in the theory. We will do it in dimensional regularization, the most commonly used, though of course it can be done in any regularization we want.

We will learn how to do this in the case of one-loop corrections in scalar theories, and specifically we will consider ϕ^4 theory in $d=4$ and ϕ^3 theory in $d=6$. Of course, ϕ^3 theory is not well defined, since the potential is unbounded from below (it can have arbitrarily high negative value). But we will take it as a simple example, since the standard theory, ϕ^4 in $d=4$, has no "wave function renormalization" at one loop (we will see what that is later), but ϕ^3 in $d=6$ is the simplest example of a theory that does have it.

Note that both theories are in the critical dimension (meaning that for a dimension higher than the one considered, the theory would be non-renormalizable). We can check this, since we saw that $\omega(D) = d-(d-2)E/2 - \sum_v [\lambda_v]$ and $[\lambda_n] = d-(d-2)n/2$. That means that for ϕ^4 in $d=4$ we have $[\lambda_4]=0$ and $\omega(D)=4-E$, and for ϕ^3 in $d=6$ we have $[\lambda_3]=0$ and $\omega(D)=6-2E$.

As we saw last lecture, in dimensional regularization we need to introduce a scale μ and redefine the coupling $\tilde{\lambda}$ in dimension d in terms of the dimensionless coupling λ in the critical dimension by

$$\tilde{\lambda} = \mu^{d-\frac{(d-2)}{2}n}\lambda \equiv \lambda_d. \tag{4.1}$$

The parameter μ is a manifestation of dimensional transmutation, the spontaneous breaking of scale invariance at the quantum level, manifested through the appearance of the arbitrary scale μ. Note that in any regularization we will find such an arbitrary scale: in the cut-off regularization it is the cut-off Λ itself, in Pauli-Villars again it is the mass parameter Λ, in the PV-like higher derivative regularization again we have a parameter Λ, etc.

The statement is that the physics should be independent of this scale, and this will lead to the renormalization group equations, which will be presented in the next lecture.

We now proceed to find the divergent diagrams at one-loop and calculate the divergent parts. After that, we will learn how to get rid of them by a redefinition of parameters.

ϕ^4 **in** $d=4$.

From the above formula $\omega(D)=4-E$, we see that the $E=0,1,2,3,4$ -point functions are all superficially divergent. But we consider 1PI diagrams, so the $E=0$ case (partition function) is not physical, since it just gives the normalization which cancels out in calculations. Since we are considering 1PI diagrams, at one-loop in ϕ^4 theory we can convince ourselves that $\Gamma^{(1)}_{\rm one-loop} = \Gamma^{(3)}_{\rm one-loop} = 0$. So we are left with $\Gamma^{(2)}$ and $\Gamma^{(4)}$ to calculate.

We define $\epsilon = 4-D$ ($D=4-\epsilon$), such that $\tilde{\lambda} = \mu^\epsilon \lambda$.

For the 1PI 2-point function, the one-loop contribution (in Fig.11) is given by just one vertex with a loop on it, with result

$$\delta\Gamma^{(2)}(p) = -\frac{\tilde{\lambda}}{2}\int \frac{d^Dp}{(2\pi)^D}\frac{1}{p^2+m^2}\ , \tag{4.2}$$

where we used the fact that we have a symmetry factor $S = 2$. Using the formula (3.30) from last lecture with $n = 1$, we obtain

$$\delta\Gamma^{(2)}(p) = -\frac{\tilde{\lambda}}{2}\frac{(m^2)^{\frac{D}{2}-1}}{(4\pi)^{\frac{D}{2}}}\Gamma\left(1-\frac{D}{2}\right) = -\frac{\lambda m^2}{2(4\pi)^2}\Gamma\left(\frac{\epsilon}{2}-1\right)\left(\frac{4\pi\mu^2}{m^2}\right)^{\frac{\epsilon}{2}}. \tag{4.3}$$

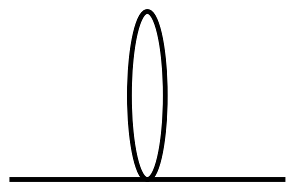

Figure 11: One-Loop Feynman diagram for the 1PI 2-point function in ϕ^4 theory.

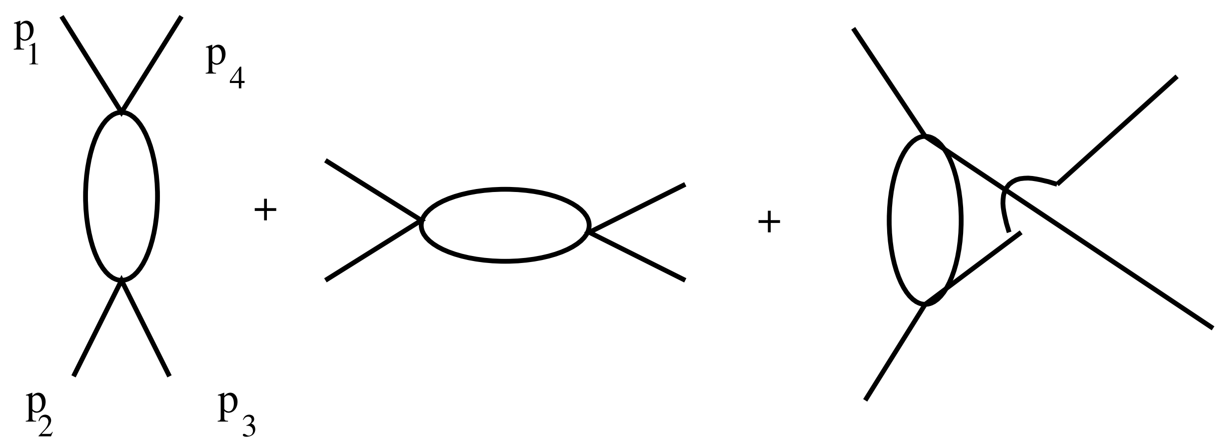

Figure 12: One-Loop Feynman diagram for the 1PI 4-point function in ϕ^4 theory.

For the one-loop contribution to $\Gamma^{(4)}$ in Fig.12, we define the 4 momenta p_1, p_2, p_3, p_4 going in, and the Mandelstam variables $s = (p_1 + p_2)^2$, $t = (p_1 + p_4)^2 = (p_2 + p_3)^2$ and $u = (p_1 + p_3)^2$, and we find 3 diagrams with 2 vertices and 2 external lines at each vertex. One diagram has $p_1 + p_4 \equiv \tilde{p}_1$ at one vertex (t diagram), the second has $p_1 + p_2$ at one vertex (s diagram) and the third has $p_1 + p_3$ at one vertex (u diagram), so that we have

$$\delta\Gamma^{(4)}(s,t,u) = \frac{\tilde{\lambda}^2}{2}[I(t) + I(s) + I(u)]. \tag{4.4}$$

Then

$$I(t) = \int \frac{d^D q}{(2\pi)^D}\frac{1}{(q^2+m^2)((q+\tilde{p}_1)^2+m^2)}. \tag{4.5}$$

Using the formula (3.41) from last lecture, for $n=2$ and $\tilde{p}_2=0$, we obtain

$$\begin{aligned}\frac{\tilde{\lambda}^2}{2}I(t) &= \frac{\tilde{\lambda}^2}{2}\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}\int_0^1 d\alpha_1 d\alpha_2\delta(1-\alpha_1-\alpha_2)[m^2+\alpha_1\tilde{p}_1^2-\alpha_1^2\tilde{p}_1^2]^{\frac{D}{2}-2}\\ &= \frac{\lambda^2\mu^{2\epsilon}}{2}\frac{\Gamma\left(\frac{\epsilon}{2}\right)}{(4\pi)^2(4\pi)^{\frac{\epsilon}{2}}}\int_0^1 d\alpha[m^2+\alpha(1-\alpha)t]^{-\frac{\epsilon}{2}}. \qquad (4.6)\end{aligned}$$

Summing over s,t,u, we obtain

$$\delta\Gamma^{(4)}(s,t,u)=\frac{\lambda^2\mu^\epsilon}{2(4\pi)^2}\Gamma\left(\frac{\epsilon}{2}\right)\left(\frac{4\pi\mu^2}{m^2}\right)^{\frac{\epsilon}{2}}\int_0^1 d\alpha\sum_{z=s,t,u}\left[1+\alpha(1-\alpha)\frac{z}{m^2}\right]^{-\frac{\epsilon}{2}}. \qquad (4.7)$$

Divergent parts.

Since we are interested in getting rid of the divergent parts, we now isolate the divergent parts of $\delta\Gamma^{(2)}(p)$ and $\delta\Gamma^{(4)}(s,t,u)$.

For the 2-point function, we first use the definition of the ψ function, $\psi(z)=d\Gamma/dz$ to write $\Gamma(1+\epsilon)\simeq 1+\epsilon\psi(1)$. Then we have also

$$\Gamma(-1+\epsilon)=\frac{\Gamma(\epsilon)}{-1+\epsilon}=\frac{\Gamma(1+\epsilon)}{-\epsilon(1-\epsilon)}\simeq-\frac{1+\epsilon\psi(1)}{\epsilon(1-\epsilon)}\simeq-\frac{1}{\epsilon}[1+\epsilon(1+\psi(1))]=-\frac{1}{\epsilon}[1+\epsilon\psi(2)]\ , \qquad (4.8)$$

where we have used $\psi(n+1)=1/n+\psi(n)$ to write $\psi(2)=1+\psi(1)$.

Then we have

$$\begin{aligned}\delta\Gamma^{(2)}(p) &\simeq -\frac{\lambda m^2}{2(4\pi)^2}\left(-\frac{2}{\epsilon}-\psi(2)\right)\left(1+\frac{\epsilon}{2}\ln\frac{4\pi\mu^2}{m^2}\right)+\mathcal{O}(\epsilon)\\ &\simeq \frac{\lambda m^2}{(4\pi)^2}\left[\frac{1}{\epsilon}+\frac{\psi(2)}{2}-\frac{1}{2}\ln\frac{m^2}{4\pi\mu^2}+\mathcal{O}(\epsilon)\right]. \qquad (4.9)\end{aligned}$$

For the 4-point function, we use

$$\Gamma(\epsilon)=\frac{\Gamma(1+\epsilon)}{\epsilon}\simeq\frac{1}{\epsilon}+\psi(1) \qquad (4.10)$$

to write the expansion

$$\begin{aligned}\delta\Gamma^{(4)}(s,t,u) &\simeq \frac{\lambda^2\mu^\epsilon}{2(4\pi)^2}\left(\frac{2}{\epsilon}+\psi(1)\right)\left(1+\frac{\epsilon}{2}\ln\frac{4\pi\mu^2}{m^2}\right)\left(3-\frac{\epsilon}{2}\int_0^1 d\alpha\sum_{z=s,t,u}\ln\left(1+\alpha(1-\alpha)\frac{z}{m^2}\right)\right)\\ &\simeq \frac{\lambda^2\mu^\epsilon}{(4\pi)^2}\left(\frac{3}{\epsilon}+\frac{3\psi(1)}{2}-\frac{3}{2}\ln\frac{m^2}{4\pi\mu^2}-\frac{1}{2}\int_0^1 d\alpha\sum_{z=s,t,u}\ln\left[1+\alpha(1-\alpha)\frac{z}{m^2}\right]+\mathcal{O}(\epsilon)\right)\\ &. \qquad (4.11)\end{aligned}$$

Note that we can do the integral over α (even though here we are not interested in the finite parts),

$$\int_0^1 d\alpha\ln\left[1+\alpha(1-\alpha)\frac{z}{m^2}\right]=-2+\sqrt{1+\frac{4m^2}{z}}\ln\left(\frac{\sqrt{1+\frac{4m^2}{z}}+1}{\sqrt{1+\frac{4m^2}{z}}-1}\right). \qquad (4.12)$$

ϕ^3 in $D = 6$.

We now repeat the same procedure for ϕ^3 in $d = 6$. Again we want to write $\tilde{\lambda} = \mu^\epsilon \lambda$, which means that now we must take $d = 6 - 2\epsilon$. From the formula $\omega(D) = 6 - 2E$, we see that the $E = 0, 1, 2, 3$ -point functions are superficially divergent, but $\Gamma^{(0)}$ is a normalization, and actually vanishes at one-loop, and $\Gamma^{(1)}$ is trivial (it gives no term in the effective action, just a constant shift of the scalar). Therefore we have only $\Gamma^{(2)}$ and $\Gamma^{(3)}$.

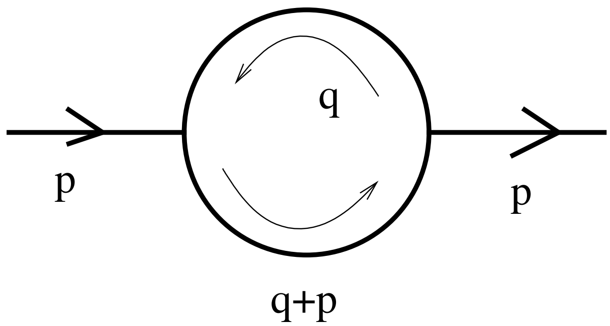

Figure 13: One-Loop Feynman diagram for the 1PI 2-point function in ϕ^3 theory.

For the 2-point function, there is only one one-loop diagram, in Fig.13, with two vertices connected by two propagators, and each having an external line. The momenta on the two internal propagators are named q and $q + p$ (p being the external momentum), and the symmetry factor is $S = 2$, giving

$$\delta\Gamma^{(2)}(p) = \frac{\tilde{\lambda}^2}{2} \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 + m^2)((q + p)^2 + m^2)}. \tag{4.13}$$

We see that it is formally the same as the $I(t)$ integral from the ϕ^4 case, with $\tilde{p}_1$ replaced by p, so we can use that result to write

$$\delta\Gamma^{(2)}(p) = \frac{\tilde{\lambda}^2}{2} \frac{\Gamma\left(2 - \frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} \int_0^1 d\alpha [m^2 + \alpha(1 - \alpha)p^2]^{\frac{D}{2} - 2}. \tag{4.14}$$

Substituting $D = 6 - 2\epsilon$ now, we obtain

$$\begin{aligned} \delta\Gamma^{(2)}(p) &= \frac{\lambda^2 m^2 \mu^\epsilon}{2(\pi)^3} \frac{\Gamma(-1 + \epsilon)}{(4\pi)^{-\epsilon}} m^{-2\epsilon} \int_0^1 d\alpha \left[1 + \alpha(1 - \alpha)\frac{p^2}{m^2}\right]^{1-\epsilon} \\ &= \frac{\lambda^2 m^2}{2(4\pi)^3} \Gamma(-1 + \epsilon) \left(\frac{4\pi\mu^2}{m^2}\right)^\epsilon \int_0^1 \left[1 + \alpha(1 - \alpha)\frac{p^2}{m^2}\right]^{1-\epsilon}. \end{aligned} \tag{4.15}$$

For the 3-point function, there is also a single one-loop diagram, in Fig.14, with 3 vertices connected pairwise by internal propagators. We denote the external lines by p_1, p_2 going in and p_3 coming out, and the internal lines by q, $q + p_1$ and $q + p_3$. Then we have

$$\delta\Gamma^{(3)}(p_i) = -\tilde{\lambda}^3 \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 + m^2)((q + p_1)^2 + m^2)((q + p_3)^2 + m^2)}. \tag{4.16}$$

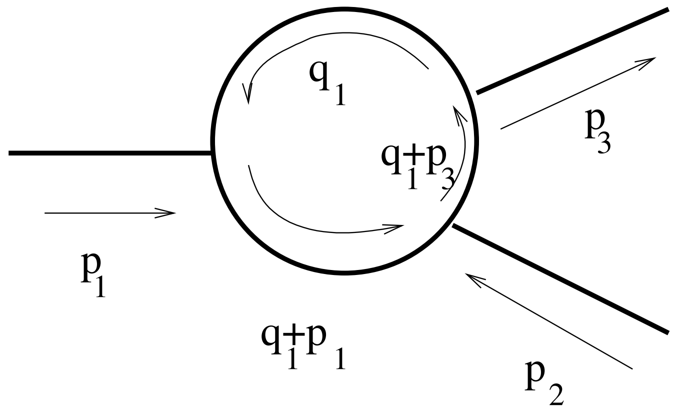

Figure 14: One-Loop Feynman diagram for the 1PI 3-point function in ϕ^3 theory.

We can use the result (3.41) from last lecture, for $n = 3$ and $\tilde{p}_1 = p_1, \tilde{p}_3 = p_3, \tilde{p}_2 = 0$, to write

$$\begin{aligned}\delta\Gamma^{(3)}(p_i) &= -\tilde{\lambda}^3 \frac{\Gamma\left(3-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} \int_0^1 d\alpha_1 d\alpha_2 d\alpha_3 \delta(1-\alpha_1-\alpha_2-\alpha_3) \\ &\quad [m^2 + \alpha_1 p_1^2 + \alpha_3 p_3^2 - (\alpha_1 p_1 + \alpha_3 p_3)^2]^{\frac{D}{2}-3} \\ &= -\frac{\lambda^3 \mu^\epsilon}{(4\pi)^3} \Gamma(\epsilon) \left(\frac{4\pi\mu^2}{m^2}\right)^\epsilon \int_0^1 d\alpha_1 d\alpha_3 \left[1 + \alpha_1 \frac{p_1^2}{m^2} + \alpha_3 \frac{p_3^2}{m^2} - \left(\frac{\alpha_1 p_1 + \alpha_3 p_3}{m}\right)^2\right]^{-\epsilon}. \end{aligned} \tag{4.17}$$

Divergent parts

We now isolate the divergent parts of $\delta\Gamma^{(2)}(p)$ and $\delta\Gamma^{(3)}(p_i)$.

For the 2-point function, we obtain

$$\begin{aligned}\delta\Gamma^{(2)}(p) &\simeq \frac{\lambda^2 m^2}{2(4\pi)^3}\left(-\frac{1}{\epsilon}\right)(1+\epsilon\psi(2))\left(1-\epsilon\ln\frac{m^2}{4\pi\mu^2}\right)\left[1+\frac{p^2}{6m^2}\right. \\ &\quad \left. -\epsilon\int_0^1 d\alpha\left[1+\alpha(1-\alpha)\frac{p^2}{m^2}\right]\ln\left[1+\alpha(1-\alpha)\frac{p^2}{m^2}\right]\right] + \mathcal{O}(\epsilon) \\ &= -\frac{\lambda^2 m^2}{2(4\pi)^3}\left(\frac{1}{\epsilon}\left(1+\frac{p^2}{6m^2}\right) + \left(1+\frac{p^2}{6m^2}\right)\left(\psi(2)-\ln\frac{m^2}{4\pi\mu^2}\right)\right. \\ &\quad \left. -\int_0^1 d\alpha\left[1+\alpha(1-\alpha)\frac{p^2}{m^2}\right]\ln\left[1+\alpha(1-\alpha)\frac{p^2}{m^2}\right]\right) + \mathcal{O}(\epsilon)\,, \end{aligned} \tag{4.18}$$

where we have used

$$\int_0^1 d\alpha\left[1+\alpha(1-\alpha)\frac{p^2}{m^2}\right] = 1+\frac{p^2}{6m^2}. \tag{4.19}$$

For the 3-point function, we obtain

$$\delta\Gamma^{(3)}(p_i) \simeq -\frac{\lambda^3\mu^\epsilon}{(4\pi)^3}\frac{1}{\epsilon}(1+\epsilon\psi(1))\left(1-\epsilon\ln\frac{m^2}{4\pi\mu^2}\right)\left[1-\epsilon\int_0^1 d\alpha_1 d\alpha_3\right.$$

$$
\begin{aligned}
& \ln\left[1+\alpha_1\frac{p_1^2}{m^2}+\alpha_3\frac{p_3^2}{m^2}-\left(\frac{\alpha_1 p_1+\alpha_3 p_3}{m}\right)^2\right]\Bigg] + \mathcal{O}(\epsilon) \\
\simeq\ & -\frac{\lambda^3\mu^\epsilon}{(4\pi)^3}\left(\frac{1}{\epsilon}+\psi(1)-\ln\frac{m^2}{4\pi\mu^2}-\int_0^1 d\alpha_1 d\alpha_3\right. \\
& \left.\ln\left[1+\alpha_1\frac{p_1^2}{m^2}+\alpha_3\frac{p_3^2}{m^2}-\left(\frac{\alpha_1 p_1+\alpha_3 p_3}{m}\right)^2\right]\right)+\mathcal{O}(\epsilon). \qquad (4.20)
\end{aligned}
$$

Counterterms

We are now finally in the position to show how to get rid of the infinities calculated above. As we can check in the two examples above, in the case of renormalizable field theories, we can hide the infinities in the redefinition of the parameters of the theory. Indeed, we see that the number and type of divergences matches the type of terms in the Lagrangean. In the ϕ^4 theory in $d=4$ we have infinities in $\Gamma^{(2)}$ and $\Gamma^{(4)}$, which are of the type of the propagators (quadratic in fields) and the interaction (ϕ^4 in fields), and in the ϕ^3 theory in $d=6$ we have infinities in $\Gamma^{(2)}$ and $\Gamma^{(3)}$, of the type of the propagators (quadratic in fields) and the interaction (ϕ^3 in fields).

Therefore we want to consider redefining the parameters by infinite factors, considering that also the original parameters were infinite, such that the physical, redefined parameters are finite quantities.

In a nonrenormalizable theory, it would happen that we have an infinite number of divergent terms, at each loop level appearing new ones. These can be cancelled only by adding new terms to the original Lagrangean. In some sense, that means that for non-renormalizable theories we would need to start with a Lagrangean with an infinite number of terms, most of them (except for a finite number) having zero coefficients, and then hide the infinities in the redefinition of all the coefficients (even the ones that are zero).

The procedure to get rid of infinities is then to add to the original Lagrangean new infinite terms called *counterterms*, giving *new infinite vertices*, such that the resulting amplitudes calculated with the total Lagrangean are finite.

One redefines the masses m_i, couplings λ_i and wave functions for the fields ϕ_i.

ϕ^4 in $d=4$.

As we argued at the beginning of the lecture, for ϕ^4 theory in $d=4$ we have no wave function renormalization, so only m and λ will be redefined. The counterterm Lagrangean to be added to the original one is such that the vertices coming from it cancel the divergences of $\Gamma^{(2)}$ and $\Gamma^{(4)}$ we found. That means that we must take the counterterm Lagrangean

$$
\mathcal{L}_{\text{c.t.}} = \left[\frac{\lambda}{16\pi^2}\frac{1}{\epsilon}\right]\frac{1}{2}m^2\phi^2+\left[\frac{\lambda^2}{16\pi^2}\frac{3}{\epsilon}\right]\mu^\epsilon\frac{\phi^4}{4!}. \qquad (4.21)
$$

Indeed, then we get two new vertices, one a 2-point vertex, denoted by a line with a cross on it, as in Fig.15a, with value (the vertex is minus the coupling in Euclidean space)

$$
-\frac{\lambda}{16\pi^2}\frac{1}{\epsilon}m^2\ , \qquad (4.22)
$$

and one a 4-point vertex, as in Fig.15b, denoted by a 4-point vertex with a circle on it (to distinguish it from the classical vertex), with value

$$-\mu^\epsilon \frac{\lambda^2}{16\pi^2}\frac{3}{\epsilon}. \tag{4.23}$$

Note that to identify $\mathcal{L}_{c.t.}$ above, we have considered the fact that a term $1/2[(\partial_\mu\phi)^2+m^2\phi^2]$ in x-space leads to a term $1/2\phi(p)[p^2+m^2]\phi(-p)$ in p-space, i.e. a "2-point vertex" $-(p^2+m^2)$. The divergent 1PI 2-point function was p-independent, depending only on m, hence there is no redefinition of the kinetic piece $(\partial_\mu\phi)^2/2$, which means no wave function renormalization (redefinition).

Also note that the 1PI n-point functions include the classical vertices, so in the redefined theory, $\mathcal{L}+\mathcal{L}_{\rm c.t.}$, we take both the loops of $\mathcal{L}$ and the classical vertices of $\mathcal{L}_{\rm c.t.}$.

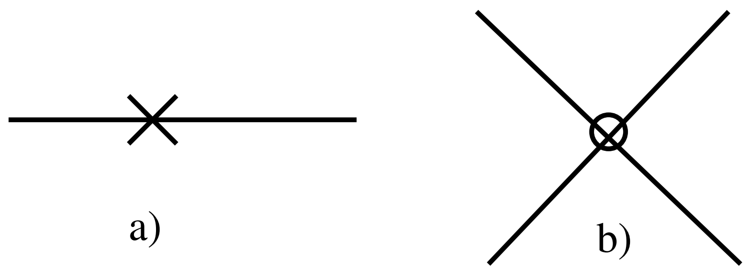

Figure 15: One-Loop Counterterm vertices in ϕ^4 theory. a) 2-point vertex. b) 4-point vertex.

Then we check that we cancel the divergences coming from $\Gamma^{(2)}_{\rm one-loop}$ and $\Gamma^{(4)}_{\rm one-loop}$ with these new vertices (the tree -"classical"- diagram from $\mathcal{L}_{\rm c.t.}$ and the one-loop diagram from the original Lagrangean are of the same order in λ, and the infinite parts cancel). Note that we have chosen $\mathcal{L}_{\rm c.t.}$ to cancel just the infinite part of the one-loop diagrams. This is called minimal subtraction, and will be discussed next lecture in more detail, but note that it is not necessary to use this, we can also consider more general subtractions, where the counterterms also contain some finite parts.

ϕ^3 in $d=6$.

In the ϕ^3 theory in $d=6$, there is a term with p^2 in $\Gamma^{(2)}(p)$, as well as a term with m^2, so now we have the countertem Lagrangean

$$\mathcal{L}_{\rm c.t.} = -\left[\frac{\lambda^2}{12(4\pi)^3}\frac{1}{\epsilon}\right]\frac{1}{2}(\partial_\mu\phi)^2 - \left[\frac{\lambda^2}{2(4\pi)^3}\frac{1}{\epsilon}\right]\frac{1}{2}m^2\phi^2 - \left[\frac{\lambda^3}{(4\pi)^3}\frac{1}{\epsilon}\right]\mu^\epsilon\frac{\phi^3}{3!}. \tag{4.24}$$

Note the sign difference with respect to ϕ^4 in $d=4$. Now the new vertices coming from this Lagrangean are:

-a 2-point vertex denoted by a line with a cross on it, with value

$$\frac{\lambda^3}{2(4\pi)^3}\frac{1}{\epsilon}\left(\frac{p^2}{6}+m^2\right), \tag{4.25}$$

-a 3-point vertex denoted by a 3-point vertex with a circle on it, with value

$$+\mu^{\epsilon}\frac{\lambda^3}{(4\pi)^3}\frac{1}{\epsilon}. \tag{4.26}$$

Therefore also now the divergences cancel between the tree diagrams coming from $\mathcal{L}_{c.t}$ and the divergent parts of the loop diagrams.

Renormalization

The general structure of $\mathcal{L}$ and $\mathcal{L}_{c.t.}$ is as follows

$$\begin{aligned}\mathcal{L} &= = \frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\mu^\epsilon\lambda}{3!}\phi^3\\ \mathcal{L}_{\text{c.t.}} &= C\frac{1}{2}(\partial_\mu\phi)^2 + B\frac{1}{2}m^2\phi^2 + A\frac{\mu^\epsilon\lambda}{3!}\phi^3.\end{aligned} \tag{4.27}$$

We now define the *renormalized Lagrangean*

$$\mathcal{L}_{\text{ren}} = \mathcal{L} + \mathcal{L}_{\text{c.t.}} = \mathcal{L}_{\text{ren}}(\phi, m, \lambda\mu^\epsilon). \tag{4.28}$$

In it, ϕ, m, λ are finite quantities as $\epsilon \to 0$.

But now we can numerically identify this with a *bare Lagrangean* written in terms of *bare quantities* ϕ_0, m_0, λ_0, which are infinite,

$$\mathcal{L}_{\text{ren.}}(\phi, m, \lambda\mu^\epsilon) = \mathcal{L}_{\text{bare}}(\phi_0, m_0, \lambda_0) = \frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m_0^2\phi_0^2 + \frac{\lambda_0}{3!}\phi_0^3\ , \tag{4.29}$$

which therefore looks like the classical Lagrangean as a function of ϕ_0, m_0, λ_0.

Sometimes one says that the renormalized Lagrangean is written in terms of bare (infinite) quantities, but that is not very satisfactory. A better interpretation is that the classical, bare Lagrangean, written in terms of infinite quantities, is reinterpreted as a renormalized Lagrangean, written as a sum of a finite, physical Lagrangean, plus a counterterm part that contains the infinities.

By comparison, we find that the relation between the bare and the renormalized quantities is given by

$$\begin{aligned}\phi_0 &= \sqrt{1+C}\phi \equiv \sqrt{Z_\phi}\phi\\ m_0^2 &= \frac{m^2(1+B)}{Z_\phi}\\ \lambda_0 &= \mu^\epsilon\lambda(1+A)Z_\phi^{n/2}\ ,\end{aligned} \tag{4.30}$$

where in the last line we wrote the coupling for a general $\mu^\epsilon\lambda\phi^n/n!$ case.

We see that the results we have found at one-loop can be written as

$$\begin{aligned}\lambda_0 &= \mu^\epsilon\left(\lambda + \frac{a_1(\lambda)}{\epsilon}\right)\\ m_0^2 &= m^2\left(1 + \frac{b_1(\lambda)}{\epsilon}\right)\end{aligned}$$

$$Z_\phi = 1+\frac{c_1(\lambda)}{\epsilon} = 1+C. \tag{4.31}$$

Of course, in the examples of ϕ^3 in $d=6$ and ϕ^4 in $d=4$, the $a_1(\lambda), b_1(\lambda), c_1(\lambda)$ are some specific powers of λ, which we will determine shortly, but in the general case of a some of $\lambda_n\phi^n/n!$ interactions, we will obtain nontrivial functions of the couplings.

In general, at general loops, we will find all higher order poles in ϵ as well (increasing order in ϵ for increasing loop number), so that the general expansion takes the form

$$\begin{aligned}
\lambda_0 &= \mu^\epsilon\left(\lambda+\sum_{k=1}^\infty \frac{a_k(\lambda)}{\epsilon^k}\right)\\
m_0^2 &= m^2\left(1+\sum_{k=1}^\infty \frac{b_k(\lambda)}{\epsilon^k}\right)\\
Z_\phi &= 1+\sum_{k=1}^\infty \frac{c_k(\lambda)}{\epsilon^k} = 1+C.
\end{aligned} \tag{4.32}$$

Examples

We now compute the a_1, b_1, c_1 coefficients in the two cases considered here.

ϕ^4 in $d=4$.

Since $C=0$, it follows that $c_1(\lambda)=0$, i.e. no wave function renormalization. Then $\lambda_0=\mu^\epsilon\lambda(1+A)$ gives

$$\lambda_0 = \lambda+\frac{\lambda^2}{16\pi^2}\frac{3}{\epsilon}\Rightarrow a_1(\lambda)=\frac{3\lambda^2}{(4\pi)^2}. \tag{4.33}$$

Then also $m_0^2=m^2(1+B)$ gives

$$m^2 = m_0^2\left(1+\frac{\lambda}{16\pi^2}\frac{1}{\epsilon}\right)\Rightarrow b_1(\lambda)=\frac{\lambda}{(4\pi)^2}. \tag{4.34}$$

ϕ^3 in $d=6$.

Now we have a wave function renormalization,

$$Z_\phi = 1+C = 1-\frac{\lambda^2}{12(4\pi)^3}\frac{1}{\epsilon}\Rightarrow c_1(\lambda)=-\frac{\lambda^2}{12(4\pi)^3}. \tag{4.35}$$

Then from

$$Z_\phi^{3/2}\lambda_0 = \mu^\epsilon\lambda(1+A)=\mu^\epsilon\left(\lambda-\frac{\lambda^3}{(4\pi)^3}\frac{1}{\epsilon}\right)\ , \tag{4.36}$$

we obtain by expanding $Z_\phi^{-3/2}$ in λ

$$a_1(\lambda) = -\frac{7}{8}\frac{\lambda^3}{(4\pi)^3}. \tag{4.37}$$

Similarly, from

$$m_0^2 Z_\phi = m^2\left(1-\frac{\lambda^2}{2(4\pi)^3}\frac{1}{\epsilon}\right)\ , \tag{4.38}$$

we obtain by expanding Z_ϕ^{-1} in λ

$$b_1(\lambda) = -\frac{5\lambda^2}{12(4\pi)^3}. \tag{4.39}$$

Note that here we have the first example of a calculation that is common in quantum field theory, though it is not very well defined mathematically. Even though the result is divergent as $\epsilon \to 0$, since the divergent term is multiplied by λ, which is considered small (perturbation theory), we can expand in λ as if the term is small. This treatment of divergences is bizarre, and it is not clear why it works from a mathematical point of view (there is probably a better underlying theory that we don't know yet), but it always does, so we will continue to use it. Moreover, note that at higher loops the pole order $1/\epsilon^k$ increases, so is even more divergent, but since it is multiplied by a higher power of λ, we consider it as an even smaller term. In fact, we treat quantum calculations order by order in λ, and at each order we make the theory finite by renormalization, and calculate finite quantities without worrying that there are worse infinities at higher orders, since we know those can be fixed as well.

Important concepts to remember

- The appearance of the arbitrary scale μ signals the quantum breaking of scale invariance, and the independence on it will lead to renormalization group equations. It is a characteristic of all regularizations.
- ϕ^4 theory in $d=4$ has one-loop infinities in $\Gamma^{(2)}(p)$ and $\Gamma^{(4)}(s,t,u)$, and ϕ^3 theory in $d=6$ has one-loop infinities in $\Gamma^{(2)}(p)$ and $\Gamma^{(3)}(p_i)$. Both theories are in the critical dimension.
- These divergences are of the same type as the terms in the Lagrangean, so can be absorbed in a redefinition of the parameters.
- To the orginal Lagrangean, we must add a counterterm Lagrangean, made up of the divergent pieces. These gives new vertices, of a higher order in λ.
- When adding to the one-loop diagrams the tree level vertices of the counterterm Lagrangean of the same order in λ, the infinities cancel.
- $\mathcal{L} + \mathcal{L}_{\text{c.t.}}$ defines the renormalized Lagrangean, written in terms of finite parameters ϕ, m, λ.
- When reinterpreted as the classical bare Lagrangean, it is written in terms of infinite bare quantities.
- The relation between bare and renormalized quantities is given by an infinite series of poles in ϵ, multiplied by functions of λ. At one-loop, we only have a single pole.

- Quantities divergent in ϵ, if multiplied by λ, are nevertheless considered small, and one can expand in them.

Further reading: See chapters 5.4 and 5.5 in [5] and 10.2 in [3].

Exercises, Lecture 4

1) Consider ϕ^3 theory in $D = 4$. Write down the divergent diagrams and calculate the divergences.

2) Write down the counterterms and renormalized Lagrangean for the above.

5 Lecture 5. Renormalization conditions and the renormalization group.

Last lecture we saw that we need to remove infinities by renormalization, which means adding (infinite) counter terms to the Lagrangean written in terms of ϕ, m, λ. The total Lagrangean, the sum of the classical one and counterterm one, when written in terms of ϕ, m, λ, called the renormalized quantities (finite), is the renormalized Lagrangean. Numerically it equals the bare Lagrangean, which formally looks the same as the classical one, but is written in terms of the bare quantities ϕ_0, m_0, λ_0, which are infinite.

The relations between the bare and renormalized quantities are written as

$$\lambda_0 = \lambda_0(\lambda, \mu, \epsilon); \quad m_0 = m_0(m, \lambda, \epsilon); \quad Z = Z(\lambda, \epsilon). \tag{5.1}$$

These can be understood as follows. We can consider λ_0, m_0, ϵ as fixed quantities, defining the theory. Then we obtain $\lambda = \lambda(\mu)$, and through it, $m = m(\lambda(\mu)) = m(\mu)$ and $Z = Z(\lambda(\mu)) = Z(\mu)$. Thus masses and couplings *run with the scale* μ. An alternative viewpoint is obtained if we fix λ and μ, obtaining $\lambda_0 = \lambda_0(\epsilon)$, $m_0 = m_0(\epsilon)$ and $Z = Z(\epsilon)$.

The first viewpoint, where physical quantities depend on scale, will lead to the renormalization group equations, which are equations for the scaling behaviour of the n-point functions.

From the fact that renormalization is a redefinition of parameters, we obtain the following scaling for the connected n-point functions:

$$G^{(n)}(p_1, ..., p_n; m\,\lambda, \mu, \epsilon) = Z_\phi^{-n/2} G_0^{(n)}(p_1, ..., p_n; m_0, \lambda_0, \epsilon). \tag{5.2}$$

Indeed, the connected n-point functions $G^{(n)}$ are obtained from the renormalized Lagrangean with sources, $\mathcal{L}^{\rm ren} - J \cdot \phi$, obtaining the partition function $Z[J]$; from it, the free energy $W[J]$, and from it, the $G^{(n)}$'s are obtained from derivation

$$G^{(n)} = -\frac{\delta^n W[J]}{\delta J_1 ... \delta J_n}. \tag{5.3}$$

But in order to have a well-defined action and partition function, we must have $J \cdot \phi = J_0 \cdot \phi_0$, which means that we need to define the bare sources as (since $\phi_0 = \sqrt{Z_\phi}\phi$)

$$J_0 = Z_\phi^{-1/2} J. \tag{5.4}$$

This, together with the fact explained above that $\mathcal{L}^{\rm ren}(\phi) = \mathcal{L}_0(\phi_0)$, means that the total action is invariant under the rescaling. The path integral measure however, transforms by

$$C_Z = \prod_{x \in \mathbb{R}^d} Z_\phi^{1/2}. \tag{5.5}$$

Then $C_Z Z[J] = Z_0[J_0]$, and since $Z = e^{-W}$, we get

$$W[J] = W_0[J_0] + \ln C_Z\ , \tag{5.6}$$

and so

$$G^{(n)} = -\frac{\delta^n W[J]}{\delta J_1 ... \delta J_n} = -Z_\phi^{-n/2} \frac{\delta^n (W_0[J_0] + \ln C_Z)}{\delta J_{01} .. \delta J_{0n}} = Z_\phi^{-n/2} G_0^{(n)}. \tag{5.7}$$

But we are actually more interested in 1PI n-point functions, relevant for the S-matrix. They are generated by the effective action, which is the Legendre transform of the free energy,

$$\begin{aligned} \Gamma[\phi^{\rm cl}] &= W[J] + J \cdot \phi^{\rm cl} \\ \Gamma_0[\phi_0^{\rm cl}] &= W_0[J_0] + J_0 \cdot \phi_0^{\rm cl}. \end{aligned} \tag{5.8}$$

Here the classical field is the connected one-point function,

$$\begin{aligned} \phi^{\rm cl} &= -\frac{\delta W[J]}{\delta J} \\ \phi_0^{\rm cl} &= -\frac{\delta W_0[J_0]}{\delta J_0} \ , \end{aligned} \tag{5.9}$$

hence we have

$$\phi_0^{\rm cl} = Z_\phi^{1/2} \phi^{\rm cl} \ , \tag{5.10}$$

which leads to a relation for Γ similar to the one for W,

$$\Gamma_0[\phi_0^{\rm cl}] = \Gamma[\phi^{\rm cl}] - \ln C_Z. \tag{5.11}$$

In turn, for the 1PI n-point functions,

$$\Gamma^{(n)} = \frac{\delta^n \Gamma}{\delta \phi_1^{\rm cl} ... \delta \phi_n^{\rm cl}} \ , \tag{5.12}$$

we also get a relation between the bare and renormalized quantities,

$$\Gamma^{(n)}(p_1, ..., p_n; m, \lambda, \mu, \epsilon) = Z_\phi^{n/2} G_0^{(n)}(p_1, ..., p_n; m_0, \lambda_0, \epsilon). \tag{5.13}$$

We note therefore, both for the case of $G^{(n)}$ and of $\Gamma^{(n)}$, the multiplicative nature of renormalization, namely that the all physical quantities are renormalized by multiplying them with the infinite factors.

Coming back to the issue we started this lecture with, the renormalized quantities λ, m, ϕ depend on the scale μ, which is understood really as being related to making measurements at a chosen scale. We will obtain that masses and couplings "run" with the scale, i.e. we have a scale dependence.

In order to do that, we need to fix λ_0, m_0, ϕ_0 when $\epsilon \to 0$. This is usually done in the UV, i.e. the infinite bare quantities are related to the UV of the theory. There is an issue of whether a theory can be defined in the IR, like we would need to in the case of QED or $\lambda\phi^4$ theory; it is believed that it cannot be done.

Subtraction schemes

In the process of renormalizing, we have subtracted only the divergences, i.e. the coefficient of the $1/\epsilon$ terms (the pole). This is a choice, and it is called *minimal subtraction scheme*, or MS scheme, but it is only a choice. We can also choose to subtract some finite parts also.

For instance, the most popular scheme is a variation of MS called $\overline{MS}$, which corresponds to subtracting the infinity, but also factors of $-\gamma + \ln(4\pi)$, where γ is the Euler constant. Indeed, we have seen that in all our example we always obtained $\psi(n) + \ln(4\pi\mu^2/m^2)$, and in the expression for $\psi(n)$ it is useful to isolate $-\gamma$, i.e. $\psi(n) = -\gamma + ...$, so $-\gamma + \ln(4\pi)$ appears naturally.

Normalization conditions

But we can also fix the parameters in a different way, that is however more physical. The tree values for the effective action are given by the 1PI n-point functions

$$\begin{aligned} \Gamma^{(2)}_{\rm tree}(p) &= p^2 + m^2 \\ \Gamma^{(4)}_{\rm tree}(p_i) &= \lambda. \end{aligned} \tag{5.14}$$

1) But we can imagine *defining* that at $p = 0$, the 1PI functions take their tree values, i.e. that

$$\begin{aligned} \Gamma^{(2)}(p=0) &= m^2; \quad \frac{d}{dp^2}\Gamma^{(2)}(p^2)\Big|_{p^2=0} = 1 \\ \Gamma^{(4)}(p_i=0) &= \lambda. \end{aligned} \tag{5.15}$$

2) The choice of $p = 0$ is nothing special, so we can in fact define that more generically, at any scale $\bar{\mu}$ the 1PI n-point functions take their tree values,

$$\begin{aligned} \Gamma^{(2)}(\bar{\mu}^2) &= \bar{\mu}^2 + m^2; \quad \frac{d}{dp^2}\Gamma^{(2)}\Big|_{p^2=\bar{\mu}^2} = 1 \\ \Gamma^{(4)}(p_i)\Big|_{p_ip_j=\bar{\mu}^2(\delta_{ij}-1/4)} &= \lambda. \end{aligned} \tag{5.16}$$

The choice of $p_ip_j = \bar{\mu}^2(\delta_{ij} - 1/4)$ was chosen such as to have the Mandelstam variables $s = t = u = \bar{\mu}^2$, but that is not even necessary, we can generalize further and require the 4-point function to be equal to its tree value at some different s, t, u related to $\bar{\mu}$ in another way.

Here 1) and 2) are different *normalization conditions*, and they are a different way to fix what we renormalize, but they should be in principle related (equivalent) to the renormalization (subtraction) schemes above. However, when we try to translate, we will see that, e.g., the MS scheme corresponds to some nontrivial normalization conditions.

Also, by the equivalence of the subtraction schemes (where different schemes would differ by the subtraction of different finite parts) to the normalization conditions, it follows that two renormalized theories differing just by different normalization conditions will differ by finite counter terms. I.e., the difference in counter terms (subtracted terms) must be finite.

A final observation at this point is that it is only in the MS scheme that the coefficient function $a_k(\lambda), b_k(\lambda)$ and $c_k(\lambda)$ depend only on λ, in general will also depend on m/μ.

Renormalization group in MS scheme

We note that we can trade the arbitrary parameter μ we have introduced when renormalizing (in dimensional regularization, it came from the fact that the coupling outside the critical dimension has a mass dimension) for the arbitrary subtraction point $\bar{\mu}$. I.e., the freedom in μ corresponds to the freedom in defining masses and couplings at some arbitrary scale.

Then the equation for the dependence on the physical scale (like $\bar{\mu}$) at which we define parameters, of observables like the 1PI Green's functions, will be the renormalization group equation.

ϕ^4 in 4 dimensions

To find it, we start with the observation that in ϕ^4 theory in $d=4$, $[\lambda_4]=0$ and so $\omega(D)=d-(d-2)E/2-\sum_v[\lambda_v]$ becomes $\omega(D)=4-E$. Thus for an 1PI function with n external lines, a *finite* diagram will give $\omega(D)=4-n$, with no necessity for renormalization, leading to the scaling law

$$\Gamma_D^{(n)}(tp_i;tm)=t^{4-n}\Gamma_D^{(n)}(p_i;m)\ , \tag{5.17}$$

where we have scaled both momenta and masses by the same factor t, to obtain the classical scaling dimension.

But since there are divergent diagrams, we need to regularize, introducing the arbitrary scale μ (or the subtraction point $\bar{\mu}$). Then, for the full *renormalized* 1PI Green's functions, which include the counterterms, to maintain the relation as it is, we need to scale also μ, i.e.

$$\Gamma^{(n)}(tp_i;tm,t\mu)=t^{4-n}\Gamma^{(n)}(p_i;m,\mu). \tag{5.18}$$

By taking td/dt onto this equation, we obtain that $\Gamma^{(n)}$ satisfies

$$t\frac{d}{dt}\Gamma^{(n)}(tp_i;tm,t\mu)=(4-n)\Gamma^{(n)}(tp_i;tm,t\mu). \tag{5.19}$$

We can rewrite this by absorbing the t multiplying m and μ and trading it for derivatives with respect to m and μ, i.e.

$$\left[t\frac{\partial}{\partial t}+m\frac{\partial}{\partial m}+\mu\frac{\partial}{\partial\mu}+n-4\right]\Gamma^{(n)}(tp_i;m,\mu)=0. \tag{5.20}$$

We can now find another equation for $\Gamma^{(n)}$ as follows. Consider the fact that Γ_0 is independent of μ,

$$\mu\frac{d}{d\mu}\Gamma_0^{(n)}(p_i;m_0,\lambda_0,\epsilon)=0\ , \tag{5.21}$$

and express it in terms of $\Gamma^{(n)}$, as

$$\mu\frac{d}{d\mu}\left[Z_\phi^{-n/2}(\lambda,\epsilon)\Gamma^{(n)}(p_i;\lambda,m,\mu,\epsilon)\right]=0. \tag{5.22}$$

Since as we discussed, we have the dependences

$$\lambda=\lambda(\mu,\epsilon;\lambda_0,m_0);\quad m=m(\lambda(\mu),\epsilon);\quad Z_\phi=Z_\phi(\lambda(\mu),\epsilon)\ , \tag{5.23}$$

when acting on $\Gamma^{(n)}(p_i,\lambda(\mu),m(\mu),\mu,\epsilon)$ with the μ derivative we have explicit and implicit dependence, leading to

$$\mu\frac{d}{d\mu}=\mu\frac{\partial}{\partial\mu}+\mu\frac{d\lambda}{d\mu}\frac{\partial}{\partial\lambda}+\mu\frac{dm}{d\mu}\frac{\partial}{\partial m}. \tag{5.24}$$

We now define the *beta function*

$$\beta(\lambda,\epsilon)\equiv \mu\frac{d\lambda}{d\mu}\bigg|_{m_0,\lambda_0,\epsilon}\;, \tag{5.25}$$

the *anomalous field dimension*

$$\gamma_d(\lambda,\epsilon)\equiv \frac{\mu}{2Z_\phi}\frac{dZ_\phi}{d\mu}\bigg|_{m_0,\lambda_0,\epsilon}\;, \tag{5.26}$$

and the *anomalous mass dimension*

$$\gamma_m(\lambda,\epsilon)\equiv \frac{\mu}{m}\frac{dm}{d\mu}\bigg|_{m_0,\lambda_0,\epsilon}\;. \tag{5.27}$$

To understand the names, note that if we have β,γ_d,γ_m constants, then the above definitions mean that

$$\lambda\sim\lambda_0+\beta\ln\mu;\quad Z_\phi\sim C\mu^{2\gamma_d}\Rightarrow\phi_0\sim\sqrt{C}\mu^{\gamma_d}\phi;\quad m\sim\tilde{C}\mu^{\gamma_m}. \tag{5.28}$$

The first means that β gives the slope of λ with $\ln\mu$, γ_d the power of μ in ϕ_0, and γ_m the power of μ in m, justifying the names.

Substituting this in (5.22), we obtain

$$\Gamma^{(n)}\mu\frac{d}{d\mu}Z_\phi^{-n/2}+Z_\phi\left(\mu\frac{\partial}{\partial\mu}+\beta\frac{\partial}{\partial\lambda}+\gamma_m m\frac{\partial}{\partial m}\right)\Gamma^{(n)}=0\Rightarrow$$
$$\left[\mu\frac{\partial}{\partial\mu}+\beta(\lambda)\frac{\partial}{\partial\lambda}+\gamma_m(\lambda)m\frac{\partial}{\partial m}-n\gamma_d(\lambda)\right]\Gamma^{(n)}(p_i,\lambda,m,\mu)=0. \tag{5.29}$$

Eliminating $\mu\partial/\partial\mu$ between (5.20) and (5.29), we obtain

$$\left[-t\frac{\partial}{\partial t}+\beta(\lambda)\frac{\partial}{\partial\lambda}+(\gamma_m(\lambda)-1)m\frac{\partial}{\partial m}-n\gamma_d(\lambda)+4-n\right]\Gamma^{(n)}(tp_i,\lambda,m,\mu)=0\;, \tag{5.30}$$

which is called the *renormalization group equation* (RGE), for the case of the ϕ^4 theory in 4 dimensions (for other theories, we have only the $4-n$ term changing to the classical scaling dimension of the 1PI n-point function).

Note that this equation gives the behaviour under scaling of the 1PI n-point functions, and is written only in terms of *physical* quantities, since we eliminated the μ dependence.

Solution. Its formal solution can be written as follows. Define $\lambda(t)$ and $\bar{m}(t)$ such that they satisfy the equations

$$\begin{aligned} t\frac{\partial}{\partial t}\bar{\lambda}(t) &= \beta(\bar{\lambda}(t))\\ t\frac{\partial}{\partial t}\bar{m}(t) &= \bar{m}(t)(\gamma_m(\bar{\lambda}(t))-1) \end{aligned} \tag{5.31}$$

with the boundary conditions

$$\bar{\lambda}(t=1)=\lambda;\quad \bar{m}(t=1)=m. \tag{5.32}$$

Then the solution of the above will be $\bar{\lambda} = \bar{\lambda}(t, \lambda)$ and $\bar{m} = \bar{m}(t, m, \lambda)$. In turn, then the formal solution to the RGE is

$$\Gamma^{(n)}(tp_i, \lambda, m; \mu) = t^{4-n} \exp\left[-n \int_1^t \frac{ds}{s} \gamma_d(\bar{\lambda}(s))\right] \Gamma^{(n)}(p_i, \bar{\lambda}(t), \bar{m}(t); \mu). \tag{5.33}$$

We will not give the proof, but one can check that this solution satisfies the equation.

It would seem that we have solved the RGE, but of course that is only in principle, not in practice. Note that in order to be able to write an explicit solution, we would need the exact formulas for $\beta(\lambda), \gamma_m(\lambda), \gamma_d(\lambda)$. We can of course plug in their values to some order in perturbation theory, but we will not obtain exact solutions.

We can compare this to the naive scaling,

$$\Gamma^{(n)}(tp_i, \lambda, m; \mu) = t^{4-n} \Gamma^{(n)}(p_i, \lambda, m/t). \tag{5.34}$$

On the other hand, if $\gamma_m(\bar{\lambda}(t))$ and $\gamma_d(\bar{\lambda}(t))$ are approximately constant as a function of t, which can happen only if $\bar{\lambda}(t)$ is approximately constant, which in turn requires the beta function to be zero, then we can easily integrate $\bar{\lambda}(t)$, $\bar{m}(t)$ to obtain $\bar{\lambda}(t) \simeq \lambda$ and $\bar{m}(t) = m/t^{1-\gamma_m}$, as well as to integrate the exponent in the solution to the RGE as

$$\exp\left[-n \int_1^t \frac{ds}{s} \gamma_d(\bar{\lambda}(s))\right] \simeq t^{-n\gamma_d} , \tag{5.35}$$

such that finally the solution to the RGE is

$$\Gamma^{(n)}(tp_i, \lambda, m, \mu) = t^{4-n(1+\gamma_d)} \Gamma^{(n)}\left(p_i, \lambda, \frac{m}{t^{1-\gamma_m}}, \mu\right). \tag{5.36}$$

We see then (by comparing with the naive scaling above) that indeed, the γ_m acts as an anomalous mass dimension (quantum correction for the scaling dimension of m with t) and γ_d as anomalous field dimension (quantum correction for the scaling dimension of the field, external to the 1PI n-point function, with t).

In general $\bar{\lambda}(t)$ is not constant (so $\lambda(\mu)$ is not constant), and we say that we have a *running coupling constant* (a coupling constant that "runs" with the energy scale), whose running is defined by the beta function, $\beta(\lambda) = \mu d\lambda/d\mu|_{m_0, \lambda_0, \epsilon}$. $\bar{\lambda}(t)$ is only approximately constant near a so-called *fixed point* of the renormalization group (RG) λ_F, where $\beta(\lambda_F) = 0$.

The running of the coupling with the energy scale means that the validity of perturbation theory can depend on scale. For instance, for QCD we know that the theory is weakly coupled in the UV, a phenomenon called asymptotic freedom, whereas it becomes strongly coupled in the IR, a phenomenon called IR slavery. Therefore QCD is perturbative only in the UV. On the other hand, for QED the situation is reversed: we have a weak coupling in the IR, and the theory becomes strongly coupled in the UV. In fact, there is a *Landau pole*, namely at some high but finite energy scale, the coupling becomes infinite in perturbation theory. Therefore QED is perturbative only in the IR. For such theories (like QED, and ϕ^4 theory in $d = 4$ which behaves in a similar manner) it is not clear that we can consistently define the theory.

If there is a well-defined perturbation theory somewhere, we have a $\lambda = 0$ point, and $\beta(\lambda = 0) = 0$ (the validity of perturbation theory means that there is a Taylor expansion for β, given by Feynman diagrams, starting at one-loop). Then $\lambda = 0$ is a universal fixed point, called the *gaussian fixed point*, since the action is free and the path integral is gaussian, at that point.

Possible behaviours for $\beta(\lambda)$.

We now analyze the most common possibilities for the behaviour of $\beta(\lambda)$. We saw that $\lambda = 0$ is the gaussian fixed point. Note that the solution to the equation $t\partial/\partial t\bar{\lambda}(t) = \beta(\bar{\lambda}(t))$ with boundary condition $\bar{\lambda}(t = 1) = \lambda$ is

$$t = \exp \int_{\lambda}^{\bar{\lambda}} \frac{d\lambda'}{\beta(\lambda')}. \tag{5.37}$$

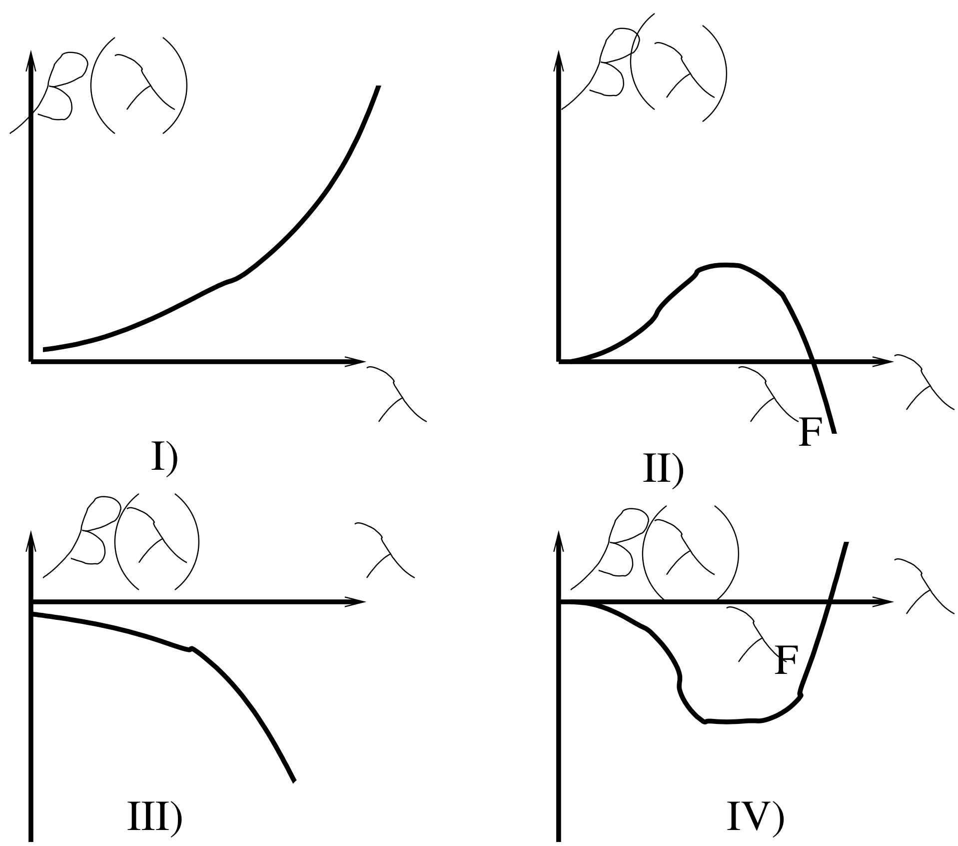

Figure 16: Possible behaviours of $\beta(\lambda)$.

I) Consider first the case when $\beta(\lambda)$ starts out positive at $\lambda = 0$ and keeps increasing, as in Fig.16I. If moreover it increases faster than λ, in (5.37), the integral $\int^{\bar{\lambda}}$ is convergent as

$\bar{\lambda} \to \infty$, and we have a Landau pole, $\bar{\lambda} \to \infty$, at some large but finite t_0,

$$t_0 = \exp \int_1^\infty \frac{d\lambda'}{\beta(\lambda')} \ , \tag{5.38}$$

i.e. at some large but finite momentum. As we mentioned, QED is of this type.

The gaussian fixed point $\lambda = 0$ is an IR stable fixed point in this case, since for $\lambda \to 0$ and $\bar{\lambda} \to \lambda \to 0$,

$$\exp \int_0^{\bar{\lambda}\to 0} \frac{d\lambda'}{\beta(\lambda')} = \exp \int_0^{\bar{\lambda}\to 0} \frac{d\lambda'}{|\beta'(0)|\lambda'} \sim \exp[+|a|\ln(0)] \to 0. \tag{5.39}$$

We can also verify that if we start with a perturbation $\lambda > 0$, then $\mu d\lambda/d\mu = \beta(\lambda) > 0$, so decreasing μ implies that λ also decreases, towards $\lambda = 0$, i.e. when going to low momenta we are driven towards $\lambda = 0$.

II) Then we can consider the case that $\beta(\lambda)$ starts out positive at $\lambda = 0$, but turns around and goes through zero, $\beta = 0$ at some $\lambda = \lambda_F$, as in Fig.16II. That means that near it we have

$$\beta(\lambda) \simeq \beta'(\lambda_F)(\lambda - \lambda_F) \ , \tag{5.40}$$

and $\beta'(\lambda_F) < 0$. In turn, that means that near λ_F, the exponent in (5.37) gives

$$\exp \int_\lambda^{\bar{\lambda}\to\lambda_F} \frac{d(\lambda' - \lambda_F)}{\beta'(\lambda_F)(\lambda' - \lambda_F)} \to +\infty. \tag{5.41}$$

That means that $t \to \infty$ as $\lambda \to \lambda_F$, i.e. for infinite momenta (infinite t) we have $\lambda \to \lambda_F$. In other words, λ_F is an UV stable fixed point. It is UV stable, since if we perturb $\lambda > \lambda_F$, $\mu d\lambda/d\mu = \beta(\lambda) < 0$, so as one increases μ, λ decreases towards λ_F, while if we perturb $\lambda < \lambda_F$, $\mu d\lambda/d\mu = \beta(\lambda) > 0$, so as one increases μ, λ increases towards λ_F. Therefore either way, as we increase μ we are driven towards λ_F.

On the other hand, in this case the gaussian fixed point $\lambda = 0$ is an IR stable fixed point, exactly as in case I, for the same reason.

III) We can consider now the mirror image of case I, i.e. that $\beta(\lambda)$ starts out negative and keeps decreasing, and moreover that it decreases faster than $-\lambda$, as in Fig.16III. Then in (5.37), the exponent behaves as

$$\exp \int_\lambda^{\bar{\lambda}\to 0} \frac{d\lambda'}{\beta(\lambda')} \simeq \exp \int_\lambda^{\bar{\lambda}\to 0} \frac{d\lambda'}{-|\beta(0)|\lambda'} = \exp[-|a|\ln(0)] \to +\infty \tag{5.42}$$

Therefore as $t \to \infty$, $\bar{\lambda}(t) \to 0$, i.e. the gaussian fixed point is an UV stable fixed point, and we are driven towards it at high momenta. We can also verify this fact since $\beta'(0) < 0$, and then as in case II for λ_F, when going at high momenta we are driven towards it, even when we initially perturb away from it. This phenomenon, of $\lambda = 0$ being UV stable is called, as we said, asymptotic freedom, and it is the behaviour QCD experiences.

On the other hand, since $\beta(\lambda)$ decreases faster than λ, for $\bar{\lambda} \to \infty$, we obtain a Landau pole, i.e. a finite value of t, since the integral is convergent at infinity,

$$t_0 = \exp \int_{\lambda}^{\bar{\lambda}\to\infty} \frac{d\lambda'}{\beta(\lambda')}. \tag{5.43}$$

Therefore we obtain a Landau pole (breakdown of perturbation theory at a finite energy scale) in the IR. That is OK, since for QCD we know that perturbation theory breaks down in the IR.

IV) Finally, we can consider the case mirror image to case II, namely when the beta function starts out negative at $\lambda = 0$, but then turns around and becomes positive again at some finite $\lambda = \lambda_F$, as in Fig.16IV. Then near λ_F we can write

$$\beta(\lambda) \simeq \beta'(\lambda_F)(\lambda - \lambda_F) \ , \tag{5.44}$$

where $\beta'(\lambda_F) > 0$.

By the same argument as in case III, the gaussian fixed point $\lambda = 0$ is UV stable. On the other hand, the fixed point λ_F is IR stable, since for $\bar{\lambda} \to \lambda_F$,

$$t = \exp \int^{\bar{\lambda}\to\lambda_F} \frac{d(\lambda' - \lambda_F)}{+|\beta'(\lambda_F)|(\lambda - \lambda_F)} \sim \exp[+|a| \ln(0)] \to 0. \tag{5.45}$$

We can also verify that, since $\beta'(\lambda_F) > 0$, the same argument as for the gaussian fixed point in case I says that if we perturb away from λ_F, when going at small μ we are driven back towards λ_F.

Perturbative beta function in dimensional regularization in MS scheme.

We now learn how to calculate the beta function in perturbation theory. We start with the formula (valid for dimensional regularization in the MS scheme) relating λ_0 and λ,

$$\lambda_0 = \mu^\epsilon \left(\lambda + \sum_{k\geq 1} \frac{a_k(\lambda)}{\epsilon^k} \right) ; \tag{5.46}$$

Since $\beta(\lambda, \epsilon) = \mu d/\partial\mu\lambda|_{m_0,\lambda_0,\epsilon}$, we take $d/d\mu|_{\lambda_0}$ on both sides of the equation above, and obtain (after dividing by μ^ϵ)

$$0 = \epsilon \left(\lambda + \sum_{k\geq 1} \frac{a_k(\lambda)}{\epsilon^k} \right) + \beta(\lambda, \epsilon) \left(1 + \sum_{k\geq 1} \frac{a_k'(\lambda)}{\epsilon^k} \right). \tag{5.47}$$

Consider now the ansatz $\beta(\lambda, \epsilon) = -\epsilon\lambda + \tilde{\beta}$, and plug it in the above equation. We can verify that then the $\mathcal{O}(\epsilon)$ terms cancel in the equation, and the $\mathcal{O}(\epsilon^0)$ terms give

$$a_1(\lambda) + \tilde{\beta} - \lambda a_1'(\lambda) = 0 \ , \tag{5.48}$$

which gives then

$$\beta(\lambda, \epsilon) = -\epsilon\lambda - a_1(\lambda) + \lambda a_1'(\lambda). \tag{5.49}$$

Taking the $\epsilon \to 0$ limit, we obtain the beta function as

$$\beta(\lambda) \equiv \lim_{\epsilon\to 0} \mu \frac{\partial \lambda}{\partial \mu} = \left(\lambda \frac{d}{d\lambda} - 1\right) a_1(\lambda). \tag{5.50}$$

This formula is exact to all orders in perturbation theory, and note that we only need to know the coefficient of the single pole, not of any of the higher order poles. Of course, $a_1(\lambda)$ receives corrections at each order in perturbation theory, so we can only calculate it to the order we know $a_1(\lambda)$.

Now we can substitute $\beta(\lambda, \epsilon) = -\epsilon\lambda + \beta(\lambda)$ in the remaining equations, and at order $\mathcal{O}(\epsilon^{-k})$ we obtain the equation

$$a_{k+1}(\lambda) + \beta(\lambda) a_k'(\lambda) - \lambda a_{k+1}'(\lambda) = 0 \ , \tag{5.51}$$

which leads to the recursion relation between the coefficients of poles

$$\left(\lambda \frac{d}{d\lambda} - 1\right) a_{k+1}(\lambda) = \beta(\lambda) \frac{d}{d\lambda} a_k(\lambda). \tag{5.52}$$

That means that we can determine (in principle) all the $a_k(\lambda)$ in terms of $a_1(\lambda)$ only.

Examples.

ϕ^4 theory in $d = 4$. We saw that one-loop, we obtained

$$a_1(\lambda) = \frac{3\lambda^2}{(4\pi)^2} + ... \tag{5.53}$$

Then from (5.50), we get

$$\beta(\lambda) = \frac{3\lambda^2}{(4\pi)^2} + \mathcal{O}(\lambda^4). \tag{5.54}$$

In turn, solving for $\bar{\lambda}(t)$, we obtain

$$\bar{\lambda}(t) = \frac{\lambda}{1 - \frac{3\lambda}{(4\pi)^2} \ln(t)}. \tag{5.55}$$

Thus it is IR free and there is a Landau pole, just like in QED.

ϕ^3 in $d = 6$. At one-loop, we had obtained

$$a_1(\lambda) = -\frac{7}{8} \frac{\lambda^3}{(4\pi)^3} + ... \ , \tag{5.56}$$

leading to the one-loop beta function

$$\beta(\lambda) = -\frac{7}{4} \frac{\lambda^3}{(4\pi)^3} + \mathcal{O}(\lambda^4). \tag{5.57}$$

This can then be solved for

$$\bar{\lambda}^2(t) = \frac{\lambda^2}{1 + \frac{7}{4} \frac{\lambda^2}{(4\pi)^3} \ln(t^2)}. \tag{5.58}$$

This is asymptotically free, just like QCD.

Perturbative calculation of γ_m and γ_d in dimensional regularization in the MS scheme.

We can now repeat the same kind of calculation to find γ_m. Start with

$$m_0^2 = m^2\left(1+\sum_{k\geq 1}\frac{b_k(\lambda)}{\epsilon^k}\right). \tag{5.59}$$

Taking $\mu d/d\mu|_{m_0,\lambda_0,\epsilon}$ on both sides and dividing by m^2, we obtain

$$0 = 2\gamma_m\left(1+\sum_{k\geq 1}\frac{b_k(\lambda)}{\epsilon^k}\right)+\beta(\lambda,\epsilon)\sum_{k\geq 1}\frac{b_k'(\lambda)}{\epsilon^k}. \tag{5.60}$$

We then substitute $\beta(\lambda,\epsilon) = -\epsilon\lambda+\beta(\lambda)$. We can check then that we cannot have a nontrivial $\gamma_m(\lambda,\epsilon)$ (a Taylor expansion in ϵ), so $\gamma_m(\lambda,\epsilon)=\gamma_m(\lambda)$. Substituting this as well, we obtain from the equation at order $\mathcal{O}(\epsilon^0)$,

$$\gamma_m(\lambda,\epsilon)=\gamma_m(\lambda)=\frac{\lambda}{2}b_1'(\lambda) \tag{5.61}$$

and substituting it back in the equation, from the order $\mathcal{O}(1/\epsilon^k)$, the recursion relation

$$\lambda b_{k+1}'(\lambda)=\beta(\lambda)b_k'(\lambda)+\lambda b_1'(\lambda)b_k(\lambda). \tag{5.62}$$

Therefore now from $b_1(\lambda)$ we can get both $\gamma_m(\lambda)$, and all the higher order poles $b_k(\lambda)$.

For γ_d, we do a similar calculation. Starting with

$$Z_\phi = 1+\sum_{k\geq 1}\frac{c_k(\lambda)}{\epsilon^k}\ , \tag{5.63}$$

and taking $\mu d/d\mu|_{m_0,\lambda_0,\epsilon}$ on both sides, we obtain

$$2\gamma_d(\lambda,\epsilon)\left(1+\sum_{k\geq 1}\frac{c_k(\lambda)}{\epsilon^k}\right)=(-\epsilon\lambda+\beta(\lambda))\sum_{k\geq 1}\frac{c_k'(\lambda)}{\epsilon^k}. \tag{5.64}$$

Again we can check that $\gamma_d(\lambda,\epsilon)$ cannot have a nontrivial Taylor expansion in ϵ, i.e. $\gamma_d(\lambda,\epsilon)=\gamma_d(\lambda)$. Then from the $\mathcal{O}(\epsilon^0)$ order in the equation, we get

$$\gamma_d(\lambda) = -\frac{\lambda}{2}c_1'(\lambda)\ , \tag{5.65}$$

and substituting it back in the equation, from the $\mathcal{O}(\epsilon^{-k})$ order we get the recursion relations

$$\lambda c_{k+1}'(\lambda)=\beta(\lambda)c_k'(\lambda)+\lambda c_1'(\lambda)c_k(\lambda). \tag{5.66}$$

Important concepts to remember

- Because of the multiplicative nature of renormalization, the relation between the bare and renormalized Green's functions (connected and 1PI) is given by multiplication with Z_ϕ factors, specifically $G^{(n)} = Z_\phi^{-n/2} G_0^{(n)}$ and $\Gamma^{(n)} = Z_\phi^{n/2} \Gamma_0^{(n)}$.

- We can think of either fixing the bare quantities λ_0, m_0, ϵ, by defining the theory usually in the UV, and then we have $\lambda(\mu)$ and implicitly $m(\mu)$ and $Z(\mu)$, meaning these physical quantities "run" with the scale, or of fixing λ, m, ϕ and then λ_0, m, ϕ are functions of ϵ.

- Renormalization is defined by a subtraction scheme, defined by removing the divergent parts by counter terms (minimal subtraction), perhaps together with some finite parts. Or equivalently by a normalization condition, whereby we fix the 1PI n-point functions to take their tree level values at some energy scale $\bar{\mu}$ (for $n \geq 2$, we need to specify also the relation of the various momenta with the scale $\bar{\mu}$).

- Two renormalized theories which differ just by normalization conditions differ by finite counter terms.

- One can trade the arbitrary parameter μ with the arbitrary subtraction point $\bar{\mu}$.

- The renormalization group equation is the equation for the variation of the renormalized 1PI n-point function under scaling, written only in terms of physical quantities (with the explicit dependence on μ solved for).

- The beta function gives the slope of scaling of λ with $\ln \mu$, the anomalous mass dimension γ_m the power of μ in m, and the anomalous field dimension the power of μ in ϕ_0. The anomalous dimensions give corresponding quantum corrections to the scaling of fields and masses in the full 1PI n-point functions.

- γ_m and γ_d are only approximately constant near a fixed point λ_F of the beta function, $\beta(\lambda_F) = 0$. If the perturbation theory is well defined, $\lambda = 0$ is an universal fixed point, the gaussian fixed point.

- QED and ϕ^4 in $d = 4$ are perturbative in the IR and has a Landau pole in the UV, QCD and ϕ^3 in $d = 6$ are asymptotically free (perturbative in the UV) and have a pole in the IR.

- $\beta(\lambda)$, as well as $a_k(\lambda)$, are given completely in terms of $a_1(\lambda)$, the coefficient of the single pole in the coupling divergences. $\gamma_m(\lambda)$, as well as $b_k(\lambda)$, are given completely in terms of $b_1(\lambda)$, and $\gamma_d(\lambda)$, as well as $c_k(\lambda)$, are given completely in terms of $c_1(\lambda)$.

Further reading: See chapters 5.6 and 5.7 in [5].

Exercises, Lecture 5

1) Calculate the beta function for ϕ^3 theory in $d = 4$ (use the results from the lecture 4's exercises).

2) Use $a_1(\lambda)$, $b_1(\lambda)$ and $c_1(\lambda)$ for ϕ^4 in $d = 4$ and ϕ^3 in $d = 6$ from class to calculate explicitly $\gamma_m(\lambda), \gamma_d(\lambda), a_2, b_2, c_2$ at one-loop. Then substitute in the RG equation for the divergent n-point functions in these respective theories. Write the explicit RG equation and its explicit solution at one-loop.

6 Lecture 6. One-loop renormalizability in QED.

In this lecture we will show the renormalizability of QED at one-loop by doing explicitly the renormalization.

We start by remembering the QED Feynman rules.

In Euclidean space, we had:

-Photon propagator between μ and ν,

$$G^{(0)}_{\mu\nu}(k) = \frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right) , \tag{6.1}$$

-Fermion (electron) propagator between α and β,

$$S^{(0)}_F(p) = \frac{1}{i\not{p} + m} = \frac{(-i\not{p} + m)_{\alpha\beta}}{p^2 + m^2} , \tag{6.2}$$

-Electron positron photon vertex, between α, β and μ,

$$+ie(\gamma^\mu)_{\alpha\beta}. \tag{6.3}$$

In Minkowski space, we had:

-Photon propagator between μ and ν,

$$G^{(0)}_{\mu\nu}(k) = \frac{-i}{k^2 - i\epsilon}\left(g_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right) , \tag{6.4}$$

-Fermion (electron) propagator between α and β,

$$S^{(0)}_F(p) = \frac{-i(-i\not{p} + m)_{\alpha\beta}}{p^2 + m^2 - i\epsilon} = \frac{-(\not{p} + im)_{\alpha\beta}}{p^2 + m^2 + i\epsilon} , \tag{6.5}$$

-Electron positron photon vertex, between α, β and μ,

$$+e(\gamma^\mu)_{\alpha\beta}. \tag{6.6}$$

We also remember that we have the Ward-Takahashi identity for the vertex,

$$p^\mu \Gamma_{\mu\alpha\beta}(p; q_2, q_1) = e(S_F^{-1}(q_2)_{\alpha\beta} - S_F^{-1}(q_1)_{\alpha\beta}). \tag{6.7}$$

and the similar one for the n-photon case,

$$k^{\mu_1}_{(1)}\Gamma^{(n)}_{\mu_1 \dots \mu_n}(k^{(1)}, ..., k^{(n)}) = 0. \tag{6.8}$$

Here of course, we can contract with any of the n momenta of the vertex. These identities are a consequence of gauge invariance.

In particular, for $n = 2$ we obtain the transversality condition for the polarization, $k^\mu \Pi_{\mu\nu}(k) = 0$, which leads to the fact that

$$\Pi_{\mu\nu}(k) = (k^2\delta_{\mu\nu} - k_\mu k_\nu)\Pi(k^2). \tag{6.9}$$

But we now note that this means that we reduce the degree of the divergence by 2, since the superficial degree of divergence $\omega(D)$ was given by the scaling dimension of the Feynman integral, but the result of the integral is constrained to have the above form, which reduces its actual (effective) degree of divergence by 2 units of mass dimension, $\omega_{\text{eff.}}(D) = \omega(D) - 2$.

In general, for the n-photon vertex $\Gamma^{(n)}$, we see that, because we have n conditions on it, given by contraction with each external momentum, the superficial degree of divergence will be reduced by n powers of the momenta, i.e. $\omega_{\text{eff.}}(D) = \omega(D) - n$, and since for $\Gamma^{(n)}$ we have $\omega(D) = 4 - n$ in 4 dimensions, we have $\omega_{\text{eff}}(D) = 4 - 2n$.

We now consider a more general analysis of the superficial degree of divergence. For the case of scalar theories, we had $\omega(D) = dL - 2I$, since there were L loop integrals in d dimensions, and I propagators of $1/(p^2 + m^2)$ for the scalars. We see that the same formula is still valid for a purely bosonic theory, with all bosons having actions with 2 derivatives, i.e. propagators with dimension -2, $\sim 1/p^2$. But when we have fermions, the fermion propagators have dimension -1, here being $1/(i\not{p} + m)$, and therefore their contribution to $\omega(D)$ is $-I_f$, for a total of

$$\omega(D) = dL - 2I_{ph} - I_f \ , \tag{6.10}$$

where I_{ph} is the number of photon internal lines and I_f is the number of fermion internal lines.

For the scalar case, we also had the formula $\sum_n nV_n = 2I + E$, since at an n-vertex we have the endpoint of n lines, but each internal line has 2 endpoints, whereas each external line has only one. In the case of QED, we have only one vertex, on which end two fermions and a photon, therefore $n = 2$ from the point of view of the fermions, and $n = 1$ from the point of view of the photons, giving

$$2I_{ph} + E_{ph} = V; \quad 2I_f + E_f = 2V. \tag{6.11}$$

We also had the relation $L = I - V + 1$, since the number of loops was given by the number of integration variables, one for each internal line, minus the number of delta function constraints, one for each vertex, except for the overall momentum conservation delta function. Now this still applies, with $I = I_{ph} + I_f$, so $L = I_{ph} + I_f - V + 1$. Finally, eliminating L, I_{ph} and I_f between the 4 relations we wrote, we obtain for the superficial degree of divergence of QED in 4 dimensions

$$\omega(D) = 4 - E_{ph} - \frac{3}{2}E_f. \tag{6.12}$$

For $\Gamma^{(n)}$, we had called E_{ph} by n before.

Then we see that a priori we have the following divergent Green's functions. We can have only photons, in which case $E_{ph} = 1, 2, 3, 4$ are superficially divergent, or only fermions, in which case $E_f = 1, 2$ are superficially divegent, or both fermions and photons, with $E_f = 1, 2$ and $E_{ph} = 1$, or $E_f = 1$ and $E_{ph} = 2$. But first, we note that we cannot have an odd number of fermions, since each fermion line must be uninterrupted (cannot just end in a vertex), so the $E_f = 1$ cases are out. Then for the pure photon case, at one-loop we have no nonzero $E_{ph} = 1$ or $E_{ph} = 3$ diagrams, whereas the $E_{ph} = 4$ case has an *effective* degree of divegence of $\omega_{\text{eff}}(D) = -4$, so is actually convergent.

We are therefore left with only 3 divergences, in $E_{ph} = 2$ (photon polarization), in $E_f = 2$ (fermion self-energy), and in $E_f = 2, E_{ph} = 1$ (vertex function). We will treat these cases separately.

Dimensional regularization of gamma matrices.

Before that, we remember some gamma matrix identities which were derived in dimension D, and are still valid in $D = 4-\epsilon$ dimensions, since the index μ on γ_μ has now D components (but the γ_μ matrix is still 4×4),

$$\begin{aligned}
\gamma_\mu\gamma^\mu &= D\,\mathbb{1}\\
\gamma_\nu\gamma_\mu\gamma^\nu &= (2-D)\gamma_\mu\\
\gamma_\nu\gamma_{\mu_1}\gamma_{\mu_2}\gamma^\nu &= 4\delta_{\mu_1\mu_2}\,\mathbb{1} + (D-4)\gamma_{\mu_1}\gamma_{\mu_2}\\
\gamma_\nu\gamma_{\mu_1}\gamma_{\mu_2}\gamma_{\mu_3}\gamma^\nu &= -2\gamma_{\mu_3}\gamma_{\mu_2}\gamma_{\mu_3} + (2-D)(\delta_{\mu_1\mu_2}\gamma_{\mu_3} + \delta_{\mu_2\mu_3}\gamma_{\mu_1} - \delta_{\mu_1\mu_3}\gamma_{\mu_2}).
\end{aligned} \tag{6.13}$$

On the other hand, we still have

$$\begin{aligned}
\mathrm{Tr}[\gamma_\mu\gamma_\nu] &= 4\delta_{\mu\nu}\\
\mathrm{Tr}[\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma] &= 4(\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\rho}\delta_{\nu\sigma} + \delta_{\mu\sigma}\delta_{\nu\rho}).
\end{aligned} \tag{6.14}$$

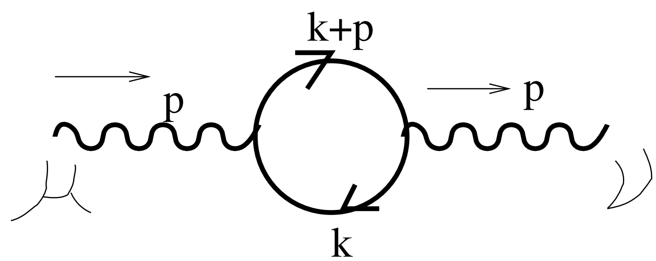

Figure 17: One loop photon polarization diagram.

1) **Photon polarization $\Pi_{\mu\nu}(p)$.**

We consider the case of an off-shell momentum p (i.e., $p^2 \neq 0$) for the photon. The diagram is a fermion loop with two external photon lines coming from it, as in Fig.17. The fermion momenta are k and $k+p$, the photon momenta are p, and the external photon indices are μ and ν. Then the value for the Feynman diagram is

$$\begin{aligned}
\Pi_{\mu\nu}(p) &= (-1)\int\frac{d^Dk}{(2\pi)^D}\mathrm{Tr}\left[(i\tilde{e}\gamma_\mu)\frac{(-i\not{k}+m)}{k^2+m^2}(i\tilde{e}\gamma_\nu)\frac{-i(\not{k}+\not{p})+m}{(k+p)^2+m^2}\right]\\
&= \tilde{e}^2\int\frac{d^Dk}{(2\pi)^D}\frac{\mathrm{Tr}[\gamma_\mu(-i\not{k}+m)\gamma_\nu(-i(\not{k}+\not{p})+m)]}{(k^2+m^2)((k+p)^2+m^2)}.
\end{aligned} \tag{6.15}$$

Here the (-1) is because of the fermion loop, and $\tilde{e} = e\mu^{\epsilon/2}$ is the coupling in $D = 4-\epsilon$ dimensions.

Using the Feynman trick for the 2 propagators $\Delta_1 = k^2 + m^2$ and $\Delta_2 = (k+p)^2 + m^2$ and shifting the momenta as $k^\mu = q^\mu - \alpha p^\mu$,

$$\frac{1}{\Delta_1 \Delta_2} = \int_0^1 \frac{1}{[q^2 + m^2 + \alpha(1-\alpha)p^2]^2} , \tag{6.16}$$

we get

$$\Pi_{\mu\nu}(p) = e^2 \mu^\epsilon \int_0^1 d\alpha \int \frac{d^D q}{(2\pi)^D} \frac{\mathrm{Tr}[\gamma_\mu(-i(\not{q} - \alpha \not{p}) + m)\gamma_\nu(-i(\not{q} + (1-\alpha)\not{p}) + m)]}{[q^2 + m^2 + \alpha(1-\alpha)p^2]^2} . \tag{6.17}$$

Rememebering that the trace of an odd number of gamma matrices is zero, the trace in the numerator becomes

$$\mathrm{Tr}[] = -\,\mathrm{Tr}[\gamma_\mu(\not{q} - \alpha\not{p})\gamma_\nu(\not{q} + (1-\alpha)\not{p})] + m^2\,\mathrm{Tr}[\gamma_\mu\gamma_\nu]. \tag{6.18}$$

By Lorentz invariance, we know that

$$\int d^D q q^\mu f(q^2) = 0 , \tag{6.19}$$

since there is no Lorentz covariant constant object with only one index. That means that the integrals of the terms in the trace with a single $\not{q}$ are zero, so

$$\mathrm{Tr}[] \to -\,\mathrm{Tr}[\gamma_\mu \not{q} \gamma_\nu \not{q}] + \alpha(1-\alpha)\,\mathrm{Tr}[\gamma_\mu \not{p} \gamma_\nu \not{p}] + 4m^2 \delta_{\mu\nu}. \tag{6.20}$$

Using (6.14), we see that

$$\mathrm{Tr}[\gamma_\mu \not{q} \gamma_\nu \not{q}] = 4(2q_\mu q_\nu - \delta_{\mu\nu} q^2) , \tag{6.21}$$

and a similar relation with p.

The integrals appearing are

$$\int \frac{d^D q}{(2\pi)^D} \frac{1}{[q^2 + \alpha(1-\alpha)p^2 + m^2]^2} = \frac{\Gamma\left(2 - \frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}} [\alpha(1-\alpha)p^2 + m^2]^{2-\frac{D}{2}}} , \tag{6.22}$$

where we have used the general result (3.30) for $m^2 \to m^2 + \alpha(1-\alpha)p^2$, and the integral with $q_\alpha q_\beta$, which by Lorentz invariance should be rewritten as an integral of $\delta_{\alpha\beta} q^2/D$, giving

$$\begin{aligned}
&\int \frac{d^D q}{(2\pi)^D} \frac{q_\alpha q_\beta}{[q^2 + \alpha(1-\alpha)p^2 + m^2]^2} = \frac{\delta_{\alpha\beta}}{D} \int \frac{d^D q}{(2\pi)^D} \frac{q^2}{[q^2 + \alpha(1-\alpha)p^2 + m^2]^2} \\
&= \frac{\delta_{\alpha\beta}}{D} \int \frac{d^D q}{(2\pi)^D} \frac{1}{[q^2 + \alpha(1-\alpha)p^2 + m^2]} - \frac{\delta_{\alpha\beta}}{D} (\alpha(1-\alpha)p^2 + m^2) \int \frac{d^D q}{(2\pi)^D} \frac{1}{[q^2 + \alpha(1-\alpha)p^2 + m^2]^2} \\
&= \frac{\delta_{\alpha\beta}}{D} \frac{\Gamma\left(1 - \frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} \frac{1}{[\alpha(1-\alpha)p^2 + m^2]^{1-\frac{D}{2}}} - \frac{\delta_{\alpha\beta}}{D} \frac{\Gamma\left(2 - \frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} \frac{1}{[\alpha(1-\alpha)p^2 + m^2]^{1-\frac{D}{2}}} \\
&= -\frac{\delta_{\alpha\beta}}{(D-2)(4\pi)^{\frac{D}{2}}} \frac{\Gamma\left(2 - \frac{D}{2}\right)}{[\alpha(1-\alpha)p^2 + m^2]^{1-\frac{D}{2}}} ,
\end{aligned} \tag{6.23}$$

where we have used $\Gamma(1-D/2)=\Gamma(2-D/2)/(1-D/2)$. Also appearing is the integral of q^2, for which we just remove the factor of $\delta_{\alpha\beta}/D$ from the above integral.

Putting all the factors together, we obtain

$$\begin{aligned}\Pi_{\mu\nu}(p) &= 4e^2\mu^\epsilon\int_0^1 d\alpha\left[-\frac{2\delta_{\mu\nu}}{(D-2)(4\pi)^{\frac{D}{2}}}\frac{\Gamma\left(2-\frac{D}{2}\right)}{[\alpha(1-\alpha)p^2+m^2]^{1-\frac{D}{2}}}\right.\\ &+2\alpha(1-\alpha)p_\mu p_\nu\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}[\alpha(1-\alpha)p^2+m^2]^{2-\frac{D}{2}}}\\ &+\frac{D\delta_{\mu\nu}}{(D-2)(4\pi)^{\frac{D}{2}}}\frac{\Gamma\left(2-\frac{D}{2}\right)}{[\alpha(1-\alpha)p^2+m^2]^{1-\frac{D}{2}}}-\alpha(1-\alpha)p^2\delta_{\mu\nu}\times\\ &\left.\times\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}[\alpha(1-\alpha)p^2+m^2]^{2-\frac{D}{2}}}+m^2\delta_{\mu\nu}\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}[\alpha(1-\alpha)p^2+m^2]^{2-\frac{D}{2}}}\right]\end{aligned} \quad (6.24)$$

We can simplify it by taking a common denominator $[\alpha(1-\alpha)p^2+m^2]^{2-\frac{D}{2}}$, and then cancelling terms that appear to finally obtain, replacing $D=4-\epsilon$,

$$\Pi_{\mu\nu}(p)=\frac{4e^2}{(4\pi)^2}(4\pi\mu^2)^{\frac{\epsilon}{2}}\Gamma\left(\frac{\epsilon}{2}\right)2(p_\mu p_\nu-p^2\delta_{\mu\nu})\int_0^1 d\alpha\frac{\alpha(1-\alpha)}{[\alpha(1-\alpha)p^2+m^2]^{\frac{\epsilon}{2}}}. \quad (6.25)$$

Expanding in epsilon (using $\Gamma(\epsilon/2)=2/\epsilon-\gamma$), we find

$$\Pi_{\mu\nu}(p)=-(\delta_{\mu\nu}p^2-p_\mu p_\nu)\frac{8e^2}{(4\pi)^2}\left(\frac{2}{\epsilon}-\gamma\right)\left[\frac{1}{6}-\frac{\epsilon}{2}\int_0^1 d\alpha\,\alpha(1-\alpha)\ln\left(\frac{\alpha(1-\alpha)p^2+m^2}{4\pi\mu^2}\right)\right]+\mathcal{O}(\epsilon)\ , \quad (6.26)$$

where we have used $\int_0^1 d\alpha\ \alpha(1-\alpha)=1/6$ and we have put the expansion of the factor $(4\pi\mu^2)^{\epsilon/2}$ together with the expansion of the power law in $\int d\alpha$. Note that the result is indeed written as $(\delta_{\mu\nu}p^2-p_\mu p_\nu)\Pi(p^2)$, as it should.

The counter term corresponds to the divergent part of $\Pi_{\mu\nu}(p)$, and since $\delta_{\mu\nu}p^2-p_\mu p_\nu$ corresponds to $-(\delta_{\mu\nu}\partial^2-\partial_\mu\partial_\nu)$, the kinetic operator coming from $1/4(\partial_\mu A_\nu-\partial_\nu A_\mu)^2$, the counter term is

$$\delta\mathcal{L}_A^{(1)}=-\frac{e^2}{12\pi^2}\frac{2}{\epsilon}\frac{1}{4}(\partial_\mu A_\nu-\partial_\nu A_\mu)^2\equiv\frac{1}{4}(Z_3-1)(\partial_\mu A_\nu-\partial_\nu A_\mu)^2\ , \quad (6.27)$$

leading to the wave function renormalization factor

$$Z_3=1-\frac{e^2}{6\pi^2\epsilon}. \quad (6.28)$$

2. Fermion self-energy $\Sigma(p)$.

It is the term appearing in the inverse propagator,

$$S_F^{-1}=i\not{p}+m+\Sigma(p). \quad (6.29)$$

The one-loop diagram for it is given by a fermion line between indices β and α, interrupted by a photon propagator starting and ending on it, starting at index μ and ending at index

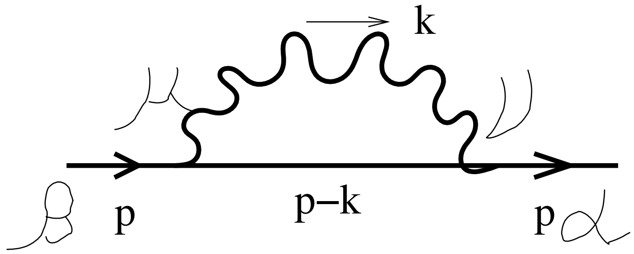

Figure 18: One loop fermion self-energy diagram.

ν, with momentum k, as in Fig.18. The fermion has external momentum p, and inside the loop has momentum $p-k$. The result of the Feynman diagram is

$$\Sigma(p)=\int\frac{d^Dk}{(2\pi)^D}\left[i\tilde{e}\gamma_\mu\frac{1}{i(\not{p}-\not{k})+m}\frac{\delta_{\mu\nu}}{k^2}i\tilde{e}\gamma_\nu\right]_{\alpha\beta}=-e^2\mu^\epsilon\int\frac{d^Dk}{(2\pi)^D}\frac{[\gamma_\mu(-i(\not{p}-\not{k})+im)\gamma^\mu]_{\alpha\beta}}{k^2((p-k)^2+m^2)}. \tag{6.30}$$

Note that the fermion propagator is gauge dependent, which means that it is not observable (it is not physical). Here that is obscured, since we have used the Feynman gauge $\alpha=1$, but the result differs in other gauges.

Again we use the Feynman trick for two propagators,

$$\frac{1}{k^2((p-k)^2+m^2)}=\int_0^1 d\alpha\frac{1}{[(1-\alpha)\Delta_1+\alpha\Delta_2]^2}=\frac{1}{[q^2+\alpha m^2+\alpha(1-\alpha)p^2]^2}\ , \tag{6.31}$$

where we have shifted $k^\mu=q^\mu+\alpha p^\mu$.

Using also $\gamma_\mu\gamma^\mu=D$ and $\gamma_\mu\not{p}\gamma^\mu=(2-D)\not{p}$ and the integral (6.22) (and again that $\int d^Dqq^\mu f(q^2)=0$), we obtain

$$\begin{aligned}\Sigma(p)_{\alpha\beta} &= -e^2\mu^\epsilon\int_0^1\int\frac{d^Dq}{(2\pi)^D}\frac{[-i\gamma_\mu((1-\alpha)\not{p}-\not{q})\gamma^\mu+m\gamma_\mu\gamma^\mu]_{\alpha\beta}}{[q^2+\alpha(1-\alpha)p^2+\alpha m^2]^2}\\ &= \frac{e^2\mu^\epsilon\Gamma\left(\frac{\epsilon}{2}\right)}{(4\pi)^{2-\frac{\epsilon}{2}}}\int_0^1 d\alpha\frac{[(2-\epsilon)(1-\alpha)(-i\not{p})-(4-\epsilon)m]_{\alpha\beta}}{[\alpha(1-\alpha)p^2+\alpha m^2]^{\epsilon/2}}.\end{aligned} \tag{6.32}$$

Expanding in ϵ, we obtain

$$\begin{aligned}\Sigma(p)_{\alpha\beta} &= \frac{e^2}{(4\pi)^2}\left(\frac{2}{\epsilon}-\gamma\right)\Bigg[(-i\not{p})\left(1-\frac{\epsilon}{2}\right)-(4-\epsilon)m\\ &\quad -\frac{\epsilon}{2}\int_0^1 d\alpha[2(1-\alpha)(-i\not{p})-4m]\ln\frac{\alpha(1-\alpha)p^2+\alpha m^2}{4\pi\mu^2}\Bigg]+\mathcal{O}(\epsilon).\end{aligned} \tag{6.33}$$

Here again we have put the expansion of $(4\pi\mu^2)^{\epsilon/2}$ in the ln, so that we form the ratio $m^2/(4\pi\mu^2)$, as before.

The divergent part of the fermion self-energy is finally

$$\Sigma(p)_{\alpha\beta} = \frac{e^2}{(4\pi)^2}\frac{2}{\epsilon}(-i\not{p} - 4m). \tag{6.34}$$

There is a term with $i\not{p} = \not{\partial}$ and a mass term, which means that we need to add a counterterm to $\bar{\psi}\not{\partial}\psi$ and one to $m\bar{\psi}\psi$,

$$\delta\mathcal{L}_{\psi}^{(1)} = -\frac{e^2}{(4\pi)^2}\frac{2}{\epsilon}\bar{\psi}\not{\partial}\psi - \frac{e^2}{(4\pi)^2}\frac{8}{\epsilon}m\bar{\psi}\psi. \tag{6.35}$$

Defining the wave function renormalization of the fermions,

$$\psi_0 = \sqrt{Z_2}\psi; \quad \bar{\psi}_0 = \sqrt{Z_2}\bar{\psi}\ , \tag{6.36}$$

from the identification of the first counterterm as $(Z_2 - 1)\bar{\psi}\not{\partial}\psi$, we obtain

$$Z_2 = 1 - \frac{e^2}{(4\pi)^2}\frac{2}{\epsilon}. \tag{6.37}$$

Then defining the renormalization of the fermion mass as

$$m_0 = \frac{Z_m}{Z_2}m\ , \tag{6.38}$$

we obtain

$$Z_m = 1 - \frac{e^2}{(4\pi)^2}\frac{8}{\epsilon}. \tag{6.39}$$

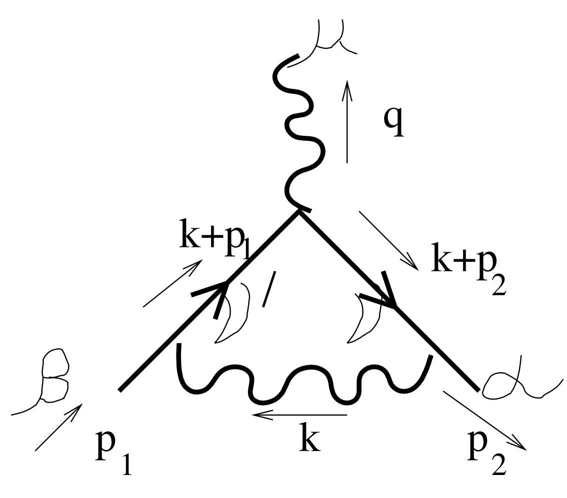

Figure 19: One loop fermions-photon vertex diagram.

3. Fermions-photon vertex $\Gamma_{\mu\alpha\beta}$.

We finally consider the one-loop correction to the fermion-fermion-photon vertex $\Gamma^{(1)}_{\mu\alpha\beta}(q;p_1,p_2)$. We consider the correction to the vertex given by a photon line connecting the two fermion lines, as in Fig.19. The external photon has index μ and (outgoing) momentum q, the external fermion lines have index β and (incoming) momentum p_1 and index α and (outgoing) momentum p_2, the internal photon line runs between ν' on p_1 and ν on p_2, modifying the fermion momenta to p_1+k and p_2+k on the internal lines.

The result of the Feynman diagram is

$$\begin{aligned}\Gamma^{(1)}_{\mu\alpha\beta}(q;p_1,p_2) &= \int\frac{d^Dk}{(2\pi)^D}\frac{\delta_{\nu\nu'}}{k^2}\left[i\tilde{e}\gamma_\nu\frac{1}{+i(\not{p}_2+\not{k})+m}i\tilde{e}\gamma_\mu\frac{1}{i(\not{p}_1+\not{k})+m}i\tilde{e}\gamma_{\nu'}\right]_{\alpha\beta}\\ &= -ie^3\mu^{3\epsilon/2}\int\frac{d^Dk}{(2\pi)^D}\frac{(\gamma_\nu[-i(\not{p}_2+\not{k})+m]\gamma_\mu[-i(\not{p}_1+\not{k}+m)]\gamma_\nu)_{\alpha\beta}}{k^2[(p_2+k)^2+m^2][(p_1+k)^2+m^2]}\end{aligned}\tag{6.40}$$

We use the Feynman parametrization for the three propagators in the denominator, giving

$$\begin{aligned}&\frac{1}{k^2[(p_2+k)^2+m^2][(p_1+k)^2+m^2]}\\ &= \int_0^1 d\alpha_1\int_0^{1-\alpha_1}d\alpha_2\frac{1}{\left([(p_1+k)^2+m^2]\alpha_1+[(p_2+k)^2+m^2]\alpha_2+k^2(1-\alpha_1-\alpha_2)\right)^3}\\ &= \int_0^1 d\alpha_1\int_0^{1-\alpha_1}d\alpha_2\frac{1}{\left(q^2+m^2(\alpha_1+\alpha_2)+\alpha_1p_1^2+\alpha_2p_2^2-(\alpha_1p_1+\alpha_2p_2)^2\right)^3}\\ &\equiv \int_0^1 d\alpha_1\int_0^{1-\alpha_1}d\alpha_2\frac{1}{\left(q^2+F(\alpha_1,\alpha_2,p_1,p_2,m)\right)^3}\,,\end{aligned}\tag{6.41}$$

where we have eliminated $\int d\alpha_3\delta(1-\alpha_1-\alpha_2-\alpha_3)$ and also used it to set the maximum value for α_2 at $1-\alpha_1$, and we have defined the function F to simplify notation.

Finally, the one-loop correction to the vertex is

$$\begin{aligned}\Gamma^{(1)}_{\mu\alpha\beta}(q;p_1,p_2) &= -ie^3\epsilon^{3\epsilon/2}\int_0^1 d\alpha_1\int_0^{1-\alpha_1}d\alpha_2\int\frac{d^Dq}{(2\pi)^D}\times\\ &\times\frac{\{\gamma_\nu[-i\not{q}-i(1-\alpha_2)\not{p}_2+i\alpha_1\not{p}_1+m]\gamma_\mu[-i\not{q}-i(1-\alpha_1)p_1+i\alpha_2\not{p}_2+m]\gamma^\nu\}_{\alpha\beta}}{(q^2+F)^3}.\end{aligned}\tag{6.42}$$

We note that, because of Lorentz invariance, the terms linear in $\not{q}$ give zero by integration, so we are left with terms quadratic in q in the numerator, generating a term that we will call $\Gamma^{(1,a)}_{\mu\alpha\beta}$, and a term with no q in the numerator, generating a term that we will call $\Gamma^{(1,b)}_{\mu\alpha\beta}$. Obviously $\Gamma^{(1b)}$ will be convergent (finite), since it is $\int d^4q/(q^2+F)^3$. On the other hand, $\Gamma^{(1a)}_{\mu\alpha\beta}$ is UV divergent, since it is $d^4qq^2/(q^2+F)^3$. But it is also IR divergent, as we will see later in the course.

We concentrate on $\Gamma^{(1a)}$, since we are interested only in the divergence. It is given by

$$\Gamma^{(1a)}_{\mu\alpha\beta} = +ie^3\mu^{3\epsilon/2}\int_0^1 d\alpha_1\int_0^{1-\alpha_1}d\alpha_2\int\frac{d^Dq}{(2\pi)^D}\frac{\{\gamma_\nu\not{q}\gamma_\mu\not{q}\gamma^\nu\}_{\alpha\beta}}{(q^2+F)^3}.\tag{6.43}$$

But from the relations (6.13), we find that

$$\gamma_\nu \not{q}\gamma_\mu \not{q}\gamma^\nu = -2\not{q}\gamma_\mu \not{q} - (D-4)\not{q}\gamma_\mu \not{q} = (2-D)\not{q}\gamma_\mu \not{q}. \tag{6.44}$$

But because $\int d^D q q^\alpha q^\beta f(q^2) = \int d^D q q^2 \delta_{\alpha\beta}/D f(q^2)$, we can replace the above with

$$\frac{(2-D)}{D} q^2 \gamma_\alpha \gamma_\mu \gamma^\alpha = \frac{(2-D)^2}{D} q^2 \gamma_\mu . \tag{6.45}$$

Then we obtain

$$\begin{aligned}
\Gamma^{(1a)}_{\mu\alpha\beta} &= +ie^3 \mu^{3\epsilon/2} \frac{(2-D)^2}{D} (\gamma_\mu)_{\alpha\beta} \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \int \frac{d^D q}{(2\pi)^D} \frac{q^2}{(q^2+F)^3} \\
&= +ie^3 \mu^{3\epsilon/2} \frac{(2-D)^2}{D} (\gamma_\mu)_{\alpha\beta} \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \left[\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} (F^2)^{\frac{D}{2}-2} - F^2 \frac{\Gamma\left(3-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} (F^2)^{\frac{D}{2}-3} \right].
\end{aligned} \tag{6.46}$$

We see that the divergence comes from the first term in the square brackets, replacing $\Gamma\left(2-\frac{D}{2}\right) = 2/\epsilon + ...$, and so we can put $D = 4$ in the rest of the integral, obtaining

$$\Gamma^{(1a)}_{\mu\alpha\beta,\text{div.}} = +ie^3 (\gamma_\mu)_{\alpha\beta} \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \frac{1}{(4\pi)^2} \frac{2}{\epsilon} = +ie(\gamma_\mu)_{\alpha\beta} \frac{e^2}{(4\pi)^2 \epsilon}. \tag{6.47}$$

That means that the counter term is

$$\delta\mathcal{L}^{(1)} = \left[-\frac{e^2}{(4\pi)^2} \frac{1}{\epsilon} \right] \left(-i\bar{\psi} \not{A} \psi \right), \tag{6.48}$$

and we can identify the coefficient in the square brackets as $Z_1 - 1$, the wavefunction renormalization of A, defining $A_0 = Z_1/Z_2 A$. Then we get

$$Z_1 = 1 - \frac{e^2}{(4\pi)^2} \frac{1}{\epsilon}. \tag{6.49}$$

But I think the correction should have been twice as large, however I could not get the factor of 2.

We finally note that all the one-loop divergence were of the form of terms in the Lagrangeans, so could be removed by renormalization and counterterms. Hence QED is one-loop renormalizable.

Important concepts to remember

- In QED, due to the Ward-Takahashi (generalized) identities, the *effective* degree of divergence of the n-photon vertex, $\Gamma^{(n)}$ is reduced by n to $\omega_{\text{eff}}(D) = 4 - 2n$.
- The superficial degree of divergence in the presence of fermions is $\omega(D) = dL - 2I_{bos} - I_f$, and for QED we obtain $\omega(D) = 4 - E_{ph} - 3E_f/2$.

- The divergent one-loop graphs in QED are in $\Pi_{\mu\nu}(p)$, $\Sigma(p)_{\alpha\beta}$ and $\Gamma_{\mu\alpha\beta}$.
- The one-loop divegences in QED can be removed by adding counterterms and renormalizing, since they have the same form as the terms in the Lagrangean.

Further reading: See chapters 6.4 in [5] and 10.3 in [3].

Exercises, Lecture 6

1) Consider $\Gamma^{(1)}_{\alpha\beta}$ from class. Calculate the general $\{\}_{\alpha\beta}$ matrix element. Calculate also the finite part of $\Gamma^{(1a)}_{\alpha\beta}$ *in the absence of IR divergences*.

2) Calculate the one-loop anomalous dimensions γ_m, γ_d for ψ, and write the explicity RG equation and its explicit solution for the one-loop $\Sigma^{(1)}(p)_{\alpha\beta}$.

7 Lecture 7. Physical applications of one-loop results 1. Vacuum polarization.

In this lecture, we will consider finally the first physical application of the quantum corrections arising from one-loop results and renormalization. But first, let us understand in a more systematic way the renormalization of QED.

Systematics of QED renormalization.

We have seen 3 renormalization factors,

$$A_0 = \sqrt{Z_3}A; \quad \psi_0 = \sqrt{Z_2}\psi; \quad m_0 = \frac{Z_m}{Z_2}m \; , \tag{7.1}$$

and a vertex renormalization with Z_1, of the $\bar\psi \not{A}\psi$ term.

The renormalized Lagrangean is then written as

$$\begin{aligned}\mathcal{L}_{\rm ren.} &= \frac{Z_3}{4}F_{\mu\nu}^2 + Z_2\bar\psi\not{\partial}\psi + Z_m\bar\psi\psi \\ &+\frac{1}{2\alpha}(\partial^\mu A_\mu)^2 + Z_1(-ie\bar\psi\not{A}\psi) \; ,\end{aligned} \tag{7.2}$$

where on the first line we have the renormalizations following from the 3 factors computed last lecture. On the second line, we have a renormalization of the gauge fixing term, considering it has no independent renormalization factor, but which, through the wave function renormalization of A_μ, implies still a renormalization

$$\alpha_0 = Z_3\alpha \; , \tag{7.3}$$

and a renormalization of the vertex function,

$$e_0 = \frac{Z_1}{Z_2\sqrt{Z_3}}e \; , \tag{7.4}$$

defined such that the overall factor is simply Z_1.

But we first note that, at least at one-loop, we have

$$Z_1^{\rm one-loop} = Z_2^{\rm one-loop}. \tag{7.5}$$

In fact, this is the result of the Ward-Takahashi identity,

$$p^\mu \Gamma_{\mu\alpha\beta}(p;q_1,q_2) = e({S_F^{-1}}_{\alpha\beta}(q_2) - {S_F^{-1}}_{\alpha\beta}(q_1)) \; , \tag{7.6}$$

since the left hand side is related to the $\bar\psi\not{A}\psi$ vertex factor Z_1, and the right hand side to the propagator term $\bar\psi\not{\partial}\psi$ with factor Z_2 (in the propagator there could be a momentum-independent contribution, related to Z_m, the coefficient of the mass term, but it cancels between the two S_F^{-1}'s of different momentum). Now the above relation actually implies only the result for the divergent parts, $Z_{1,div.} = Z_{2,div.}$, but since we are in the MS scheme, this is the whole factor. Then the result is actually exact to all loops,

$$Z_1 = Z_2. \tag{7.7}$$

This then implies for the vertex function,

$$e_0 = \frac{e}{\sqrt{Z_3}}\ , \tag{7.8}$$

which to one-loop gives

$$e_0 = e\mu^{\epsilon/2}\left(1 + \frac{e^2}{24\pi^2}\frac{2}{\epsilon} + \mathcal{O}(e^3)\right). \tag{7.9}$$

In turn, this gives for the beta function

$$\mu\frac{\partial}{\partial\mu}e \equiv \beta(e) = \frac{e^3}{12\pi^2} + \mathcal{O}(e^5). \tag{7.10}$$

This is solved by

$$\bar{e}^2(t) = \frac{e_0^2}{1 - \frac{e_0^2}{12\pi^2}\ln t^2}\ , \tag{7.11}$$

where $e_0 = e(\mu_0)$. That means that there is a Landau pole. Since $e_0^2/4\pi \simeq 1/137$ at about the eV scale (the scale of the H atom energy levels), we obtain a Landau pole at about $e^{\frac{12\pi^2}{e_0^2}} \sim e^{1370}eV$.

Vacuum polarization.

Inside a nontrivial medium, the effective action is written in terms of the electric and magnetic fields as

$$S_{\text{eff.}}(\vec{E},\vec{B}) = \frac{1}{2}\int dt\int d^3x\left[\epsilon\vec{E}^2 - \frac{\vec{B}^2}{\mu}\right]\ , \tag{7.12}$$

and the speed of light inside the medium is

$$c_{\text{medium}} = \frac{1}{\sqrt{\epsilon\mu}}. \tag{7.13}$$

The Coulomb potential associated with a static pointlike charge e is

$$eA_\mu^{\text{Coulomb}}(\vec{x}) = \frac{e^2}{4\pi\epsilon|\vec{x}|}\delta_{\mu 0}. \tag{7.14}$$

The dielectric function ϵ is in general a function, and not a constant, and is usually defined in momentum space, as the ratio of the electric induction $\vec{D}$ and the electric field $\vec{E}$,

$$\epsilon(\omega,\vec{k}) \equiv \frac{\vec{D}(\omega,\vec{k})}{\vec{E}(\omega,\vec{k})}. \tag{7.15}$$

Then really, the Coulomb potential in momentum space for a static source ($\omega = 0$) is

$$eA_\mu^{\text{Coulomb}}(\vec{k},t) = \frac{e^2}{\vec{k}^2\epsilon(0,k)}\delta_{\mu 0}. \tag{7.16}$$

Now consider the same situation in vacuum, but with nontrivial quantum corrections, i.e. with nontrivial vacuum polarization. We can again formally consider it as a nontrivial "medium" with $\epsilon, \mu \neq 1$, but unlike a real medium, now the velocity of light must be exactly equal to 1, which means

$$\epsilon(\omega, \vec{k})\mu(\omega, \vec{k}) = 1. \tag{7.17}$$

The quantum effective action starts at quadratic order,

$$\Gamma(A^{\rm cl}_\mu) = \frac{1}{2}\int \Gamma^{(2)}_{\mu\nu} A^{\rm cl}_\mu A^{\rm cl}_\nu + \mathcal{O}((A^{\rm cl}_\rho)^3)\ , \tag{7.18}$$

where the quadratic part is written as the inverse propagator, which equals the free inverse propagator plus the vacuum polarization,

$$\Gamma^{(2)}_{\mu\nu} = G^{-1}_{\mu\nu} = G^{(0)-1}_{\mu\nu} + \Pi_{\mu\nu}\ , \tag{7.19}$$

so that more precisely, (writing the vacuum polarization in terms of $\Pi(k^2)$)

$$\begin{aligned} \Gamma^{(2)}[A^{\rm cl}_\mu] &= \frac{1}{2}\int \frac{d^4k}{(2\pi)^4} A^{\rm cl}_\mu(-k) G^{-1}_{\mu\nu}(k) A^{\rm cl}_\nu(k) \\ G^{-1}_{\mu\nu} &= (k^2\delta_{\mu\nu} - k_\mu k_\nu)(1+\Pi(k^2)) + \frac{k_\mu k_\nu}{\alpha}. \end{aligned} \tag{7.20}$$

It would seem like there is some gauge depence in this effective action, due to the term with the gauge parameter α, but really there isn't, since if the classical current source is conserved, i.e. $\partial^\mu J_\mu$, in momentum space $p^\mu J_\mu(p) = 0$, then

$$k^\mu J_\mu(k) = k^\mu \frac{\delta \Gamma}{\delta A^{\rm cl}_\mu} = \frac{1}{\alpha} k^2 k_\mu A^{\rm cl}_\mu\ , \tag{7.21}$$

which in turn means that the extra term in $\Gamma^{(2)}$ vanishes, so that

$$\Gamma^{(2)}[A^{\rm cl}_\mu] = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}(1+\Pi(k^2)) A_\mu(-k)(k^2\delta_{\mu\nu} - k_\mu k_\nu) A^{\rm cl}_\nu(k). \tag{7.22}$$

We now must Wick rotate it to Minkowski space in order to be able to extract physical information, by

$$\begin{aligned} \delta_{\mu\nu} &\rightarrow g_{\mu\nu} \\ A^\mu(-k)(k^2 g_{\mu\nu} - k_\mu k_\nu) A^\nu(k) &= E_i(-k)E_i(k) - B_i(-k)B_i(k). \end{aligned} \tag{7.23}$$

Finally, the quantum corrected effective action in Minkowski space is

$$\Gamma^{(2)}[A^{\rm cl}_\mu] = \frac{1}{2}\int \frac{dk_0}{2\pi}\int \frac{d^3k}{(2\pi)^3}(1+\Pi(k^2))(|\vec{E}(k_0,\vec{k})|^2 - |\vec{B}(k_0,\vec{k})|^2). \tag{7.24}$$

From it, we can extract ϵ and μ for the vacuum,

$$\epsilon(k^2) = \frac{1}{\mu(k^2)} = 1 + \Pi(k^2). \tag{7.25}$$

Keeping only the quadratic part of Γ, and remembering that $J_\mu = \delta\Gamma/\delta A_\mu^{\rm cl}$, we obtain

$$J_\mu \simeq G^{-1}_{\mu\nu} A^{\rm cl}_\nu \ , \tag{7.26}$$

so that

$$A^{\rm cl}_\mu \simeq G_{\mu\nu} J_\nu(k) = \frac{J_\mu}{k^2(1+\Pi(k^2))} \ , \tag{7.27}$$

where we have used the explicit form of $G^{-1}_{\mu\nu}$, and in inverting it, we have considered the fact that $J_\nu k^\mu k^\nu = 0$.

The Coulomb potential of a static pointlike source of

$$J_\mu(\vec{x}, t) = e\delta^{(3)}(\vec{x})\delta_{\mu 0} \Rightarrow J_\mu(k) = 2\pi e \delta(k_0)\delta_{\mu 0} \ , \tag{7.28}$$

for Π that depends on $k^2 = \vec{k}^2$ only, is

$$e A^{\rm cl}_0(\vec{k}, k_0) \simeq \frac{e^2}{\vec{k}^2(1+\Pi(\vec{k}^2))} 2\pi\delta(k_0) \ , \tag{7.29}$$

leading to the effective coupling

$$e^2_{\rm eff} = \frac{e^2}{1+\Pi(\vec{k}^2)}. \tag{7.30}$$

Note that we have naturally $e^2_{\rm eff}(\vec{k}^2)$, but since $|\vec{k}| \sim 1/r$, we can think of this as $e^2_{\rm eff}(r)$. Then in the extreme IR, we have

$$e^2_{\rm eff}(r \to \infty) = \frac{e^2}{1+\Pi(0)}. \tag{7.31}$$

This relation is interpreted physically as *screening of the electric charge of the pointlike source*. Indeed, we see that the effectice charge is smaller than the free one. Therefore we can interpret this in the same way as interpret the screening of charge in a polarizable medium. The charge is effectively screened, since dipoles (charge pairs) orient themselves such as to screen the outside charge (opposite charge closer to the source). The difference is of course that in a material, it is only a local effect; globally, because of charge conservation, we have the same charge. But here, the charge is intepreted as a continuous coupling, and it can be made smaller (screened) by the interaction with the (polarizable) vacuum.

We also note that, since we are dealing with QED, which as we argued, is defined only in the IR, it makes sense to define the physical (observed) value of the coupling in the extreme IR, at $r \to \infty$. Defining thus $e^e_{\rm eff}(r \to \infty) = e^2$, we have the (natural) normalization condition for renormalization

$$\Pi(0) = 0. \tag{7.32}$$

Remembering that the finite part of the vacuum polarization at one-loop was given by

$$\Pi(k^2) = -\frac{e^2}{2\pi^2}\left[\frac{\psi(1)}{6} + \int_0^1 d\alpha \alpha(1-\alpha)\ln\left[\frac{k^2\alpha(1-\alpha)+m^2-i\epsilon}{4\pi\mu^2}\right]\right] \ , \tag{7.33}$$

but switching now from the MS scheme to the $\Pi(0)=0$ normalization condition, i.e. dropping the nonzero terms at $k=0$ in the above, we obtain

$$\Pi(k^2) = -\frac{e^2}{2\pi^2}\int_0^1 d\alpha\alpha(1-\alpha)\ln\frac{k^2\alpha(1-\alpha)+m^2-i\epsilon}{m^2}. \tag{7.34}$$

It is left as an exercise to prove that, if $-k^2 \leq 4m^2$, the above integral for $k^2=-\vec{k}^2$ reduces to

$$\Pi(\vec{k}^2) = -\frac{e^2}{12\pi^2}\vec{k}^2\int_{4m^2}^{\infty}\frac{dq^2}{q^2}\frac{1}{q^2+\vec{k}^2}\left(1+\frac{2m^2}{q^2}\right)\sqrt{1-\frac{4m^2}{q^2}}. \tag{7.35}$$

Since we are at small $\Pi(\vec{k}^2)$ (it is of $\mathcal{O}(e^2)$), $1/(1+\Pi)\simeq 1-\Pi$, so

$$\begin{aligned} A_0^{\rm cl}(k_0,\vec{k}) &\simeq \frac{e}{\vec{k}^2}(1-\Pi(\vec{k}^2))2\pi\delta(k_0) \Rightarrow \\ A_0(t,\vec{x}) &= e\int\frac{d^3k}{(2\pi)^3}\frac{e^{-i\vec{k}\vec{x}}}{\vec{k}^2}\left(1+\frac{e^2}{12\pi^2}\vec{k}^2\int_{4m^2}^{\infty}\frac{dq^2}{q^2}\frac{1}{q^2+\vec{k}^2}\left(1+\frac{2m^2}{q^2}\right)\sqrt{1-\frac{4m^2}{q^2}}\right) \end{aligned} \tag{7.36}$$

Using

$$\int\frac{d^3k}{(2\pi)^3}\frac{e^{-i\vec{k}\cdot\vec{x}}}{\vec{k}^2}=\frac{1}{4\pi r};\quad \int\frac{d^3k}{(2\pi)^3}\frac{e^{-i\vec{k}\cdot\vec{x}}}{q^2+\vec{k}^2}=\frac{e^{-qr}}{4\pi r} \tag{7.37}$$

in the two terms above, and defining $u^2=q^2/4m^2$, we get

$$A_0^{\rm cl}(\vec{x},t)=\frac{e}{4\pi r}\left(1+\frac{e^2}{6\pi^2}\int_1^{\infty}\frac{du}{u^2}e^{-2mru}\left(1+\frac{1}{2u^2}\right)\sqrt{u^2-1}\right). \tag{7.38}$$

We consider the extreme limits of this formula. At large distances, $mr\gg 1$, we get

$$A_0^{\rm cl}(t,\vec{x})\simeq\frac{e}{4\pi r}\left(1+\frac{e^2}{16}\frac{e^{-2mr}}{(\pi mr)^{3/2}}+\ldots\right), \tag{7.39}$$

whereas at small distances, $mr\ll 1$, we get

$$A_0^{\rm cl}(t,\vec{x})\simeq\frac{e}{4\pi r}\left(1+\frac{e^2}{12\pi^2}\ln\frac{1}{(mr)^2}+\text{constant}+\ldots\right) \tag{7.40}$$

In the large distance formula we just notice again the fact mentioned previously, that this is consistent with screening, since we obtain $e_0>e$. In the small distance formula however, we also notice another thing, namely that the effective charge *diverges* in the extreme UV, at $mr\to 0$. Therefore, the screening with respect to the UV is infinite: we have an infinite effective charge at $r=0$, but it is screened down to a finite value in the IR. This is consistent with the picture of renormalization we have advocated: have infinite quantities in the UV, which can be renormalize down to finite ones in the IR.

Pair creation rate

In the previous, we have considered the case of $-k^2\leq 4m^2$, but now we can also consider the opposite case, of $-k^2\geq 4m^2$. In this case, $k^0\geq 2m$, so we have sufficient energy to create an electron-positron pair from the vacuum. The vacuum to vacuum transition amplitude in the presence of an external source J is the partition function Z ($= e^{-W}$ in Euclidean space), written in terms of the effective action ($W = \Gamma + J \cdot A^{\rm cl}$) in Minkowski space as

$$ {}_J\langle 0|0\rangle_J \equiv Z[J] = e^{i\Gamma[A_\mu^{\rm cl}] + iJ\cdot A^{\rm cl}}. \tag{7.41}$$

If Γ has an imaginary part, we can have an absolute value different than one, $|Z|^2 = e^{-2{\rm Im}[\Gamma]}$, which is interpreted as *vacuum decay*, i.e. the probability of the vacuum to go to itself is not 1 anymore, and the difference is due, as we explained, to pair creation,

$$R = 1 - |{}_J\langle 0|0\rangle_J|^2 = 1 - e^{-2{\rm Im}[\Gamma]} \simeq 2{\rm Im}[\Gamma[A_\mu^{\rm cl}] + \mathcal{O}(e^4). \tag{7.42}$$

It remains of course to show that exactly when $-k^2 \geq 4m^2$, i.e. when we can create pairs, we do create them, i.e. $\Gamma^{(2)}$ has an imaginary part. We will see this through explicit calculation. In the formula (7.34) the log can give us an imaginary part. Indeed, if the argument of the log is negative, we will have a term $\log(-1) = -i\pi$ added. Since $k^2 \leq -4m^2 \leq 0$, this will happen if $\alpha(1-\alpha) + m^2/k^2$ is positive, therefore we obtain

$${\rm Im}\Pi(k^2) = \frac{e^2}{2\pi}\int_0^1 d\alpha\alpha(1-\alpha)\theta\left(\alpha(1-\alpha) + \frac{m^2}{k^2}\right) , \tag{7.43}$$

where θ is the Heaviside function. Its roots are at

$$\alpha(1-\alpha) = -\frac{m^2}{k^2} \Rightarrow \alpha_{1,2} = \frac{1 \pm \sqrt{1 + \frac{4m^2}{k^2}}}{2}. \tag{7.44}$$

and the positivity condition of the Heaviside function for the inverted parabola is at $\alpha_1 \leq \alpha \leq \alpha_2$. We see then that the condition for this to be nonzero is indeed the condition we advocated, $1 + 4m^2/k^2 \geq 0$, i.e. $-k^2 \geq 4m^2$. Then the integral is

$$\int_{\alpha_1}^{\alpha_2} d\alpha\alpha(1-\alpha) = \left(\frac{\alpha^2}{2} - \frac{\alpha^3}{3}\right)\Big|_{\alpha_1}^{\alpha_2} = \frac{1}{4}\sqrt{1 + \frac{4m^2}{k^2}}\left(\frac{2}{3} - \frac{m^2}{3k^2}\right). \tag{7.45}$$

Finally we obtain for the imaginary part of the vacuum polarization

$${\rm Im}\Pi(k^2) = \frac{e^2}{12\pi}\sqrt{1 + \frac{4m^2}{k^2}}\left(1 - \frac{m^2}{2k^2}\right)\theta\left(1 + \frac{4m^2}{k^2}\right). \tag{7.46}$$

Here the condition of reality of ImΠ would have been enough to show $-k^2 \geq 4m^2$, but we have put it explicitly with the Heaviside function for completeness.

The imaginary part of the effective action is

$${\rm Im}\Gamma^{(2)}[A_\mu^{\rm cl}] = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}{\rm Im}\Pi(k^2)(|\vec{E}^2(k_0, \vec{k})|^2 - |\vec{B}^2(k_0, \vec{k})|^2) , \tag{7.47}$$

and the pair creation rate is $R = 2{\rm Im}\Gamma^{(2)}$.

We observe that pair creation from the vacuum is a purely electric effect, since if $k^2 \leq -4m^2 < 0$, it means we can choose a center of mass reference frame where $\vec{k} = 0$, and then the magnetic field $\vec{B} = -i\vec{k} \times \vec{A}(k) = 0$. Another way of seeing this is that the pair creation rate is positive, so only the electric field contributes, with $+|\vec{E}|^2$, whereas the magnetic field, with $-|\vec{B}|^2$, doesn't.

Important concepts to remember

- The renormalization of the α parameter matches the one of the wave function of A, $\alpha_0 = Z_3\alpha$, whereas from the Ward-Takahashi identity, $Z_1 = Z_2$, exact to all loop orders, which means that $e_0 = e/\sqrt{Z_3}$.
- The vacuum is considered like a medium with nontrivial $\epsilon(k_0, \vec{k})$ and $\mu(k_0, \vec{k})$, just that because of relativistic invariance, signals propagate at $c = 1 = 1/\sqrt{\epsilon\mu}$, meaning that $\epsilon\mu = 1$.
- We obtain $\epsilon = 1/\mu = 1+\Pi(k^2)$, and a screening of electric charge, $e_{\rm eff}^2 = e^2/(1+\Pi(k^2))$.
- We can choose a normalization condition such that $\Pi(0) = 0$, identifying $e_{\rm eff}^2(r \to \infty)$ with e^2. Then at $r \to 0$, we obtain a divergence in $e_{\rm eff}$, meaning we have an infinite screening from the UV to the IR.
- An imaginary part of $\Pi(k^2)$ leads to vacuum decay through pair creation, happening when $-k^2 \geq 4m^2$, which is a purely electric effect (pair creation in an electric field).

Further reading: See chapter 6.5.1 in [5].

Exercises, Lecture 7

1) Prove the result assumed in the lecture, that if $-k^2 \leq 4m^2$,

$$\begin{aligned}
\Pi(k^2) &= -\frac{e^2}{2\pi^2}\int_0^1 d\alpha\alpha(1-\alpha)\ln\frac{k^2\alpha(1-\alpha)+m^2-i\epsilon}{m^2} \\
&= -\frac{e^2}{12\pi^2}\vec{k}^2\int_{4m^2}^\infty \frac{dq^2}{q^2}\frac{1}{q^2+\vec{k}^2}\left(1+\frac{2m^2}{q^2}\right)\sqrt{1-\frac{4m^2}{q^2}}.
\end{aligned} \tag{7.48}$$

2) Calculate the e^+e^- pair creation rate for $\vec{B}=0$ and an electric field

$$|\vec{E}(k)| = E_0 = \text{constant}\ , \tag{7.49}$$

as well as for an electric field

$$|\vec{E}(k)| = \delta^{(3)}(\vec{k})\theta(k_0-2M)\ , \tag{7.50}$$

where $M > m$.

8 Lecture 8. Physical applications of one-loop results 2. Anomalous magnetic moment and Lamb shift.

In this lecture we continue with the physical applications of one-loop results, describing two classic tests of radiative corrections in QED.

Anomalous magnetic moment.

The first one is related to the anomalous magnetic moment of the electron.

Classically, a particle of electric charge q, with orbital angular momentum $\vec{L}$, has a magnetic moment of

$$\vec{\mu} = \frac{q}{2m}\vec{L}. \tag{8.1}$$

But quantum mechanically, for the electron of charge e, the spin $\vec{S}$ also has a contribution to the magnetic moment,

$$\vec{\mu}_{\rm class} = \frac{e}{m}\vec{S} = g\left(\frac{e}{2m}\right)\vec{S}\ , \tag{8.2}$$

where the Landé g-factor is classically (i.e., in nonrelativistic quantum mechanics)

$$g_{\rm spin,classical} = 2. \tag{8.3}$$

But in QED one obtains quantum corrections,

$$g = 2\left(1+\frac{\alpha}{2\pi}+...\right) = 2(1+0.001159652359...)\ , \tag{8.4}$$

where we have written the numerical expression appearing from including $\mathcal{O}(\alpha^3)$ corrections and more, and the 12 digits written are all verified experimentally to be correct. This is one of the most impressive tests of QED, and we will derive here the first order term (α/π).

Consider the relativistic Dirac equation,

$$(\not{D}+m)\psi = 0 \Rightarrow (-\not{D}+m)(\not{D}+m)\psi = 0. \tag{8.5}$$

But since $[D_\mu, D_\nu] = -ieF_{\mu\nu}$, $[\gamma_\mu,\gamma_\nu] = 2\gamma_{\mu\nu} \equiv -2i\sigma_{\mu\nu}$,

$$\not{D}\not{D} = D_\mu D_\nu \gamma^\mu\gamma^\nu = D_\mu D_\nu \frac{1}{2}([\gamma_\mu,\gamma_\nu]+\{\gamma_\mu,\gamma_\nu\}) = D^2 - iD_{[\mu}D_{\nu]}\sigma^{\mu\nu} = D^2 - \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}. \tag{8.6}$$

Then the equation for ψ is

$$\left(-D^2+m^2+\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}\right)\psi = 0. \tag{8.7}$$

Thus we have an extra term $+e/2\sigma^{\mu\nu}F_{\mu\nu}$, and since $F_{ij} = \epsilon_{ijk}B_k$ and $i\sigma_{ij} = i[\sigma_i,\sigma_j]/2 = -\epsilon_{ijk}\sigma_k$,

$$\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} = -e\vec{\sigma}\cdot\vec{B} = -2e\vec{S}\cdot\vec{B}. \tag{8.8}$$

But this is a term ΔE^2 in E^2 (since $-D^2$ contains $+\partial_t^2 = -E^2$), so the difference in energy is

$$\Delta E \simeq \frac{\Delta E^2}{2E} \simeq \frac{\Delta E^2}{2m} = -\frac{e}{m}\vec{S}\cdot\vec{B} = -\vec{\mu}\cdot\vec{B} = -g\frac{e}{2m}\vec{S}\cdot\vec{B}. \tag{8.9}$$

The way to obtain $g \to g + \Delta g$ from the Lagrangean would be to add a term to the classical Lagrangean, to obtain

$$\mathcal{L} = -\bar{\psi}(\not{D} + m)\psi + \frac{\Delta g}{4}\frac{e}{2m}\bar{\psi}(\sigma^{\mu\nu}F_{\mu\nu})\psi. \tag{8.10}$$

Indeed, such a term would add to the Dirac equation

$$-\frac{\Delta g}{4}\frac{e}{2m}\sigma^{\mu\nu}F_{\mu\nu}\psi = +\Delta g\frac{e}{2m}\vec{S}\cdot\vec{B}\psi = \Delta\vec{\mu}\cdot\vec{B}\psi \to \Delta E\psi. \tag{8.11}$$

We will therefore search for such a term in the Lagrangean, generated by radiative (quantum loop) corrections, i.e. as an one-loop effective action correction.

But we saw that

$$\Gamma_{\mu,\alpha\beta}(p_1,p_2;q) = \Gamma^{(0)}_{\mu\alpha\beta} + \Gamma^{(1)}_{\mu\alpha\beta}(p_1,p_2;q); \quad \Gamma^{(1)}_{\mu\alpha\beta} = \Gamma^{(1a)}_{\mu\alpha\beta} + \Gamma^{(1b)}_{\mu\alpha\beta}\ , \tag{8.12}$$

and we saw that $\Gamma^{(1a)}$ was proportional to γ^μ (and was UV divergent), so we calculate the term

$$\bar{\psi}(p_2)\Gamma^{(1b)}_\mu\psi(p_1) \tag{8.13}$$

on-shell, i.e. when

$$\begin{array}{l} p_1^2 = -m^2;\ p_2^2 = -m^2;\ \ 0 = q^2 = (p_1 - p_2)^2 \Rightarrow p_1\cdot p_2 = -m^2 \\ (i\not{p}_1 + m)\psi(p_1) = 0; \quad (i\not{p}_2 + m)\psi(p_2) = 0. \end{array} \tag{8.14}$$

But we calculated

$$\begin{array}{l} \Gamma^{(1b)}_{\mu\alpha\beta}(p_1,p_2) = -i\frac{e^3}{(4\pi)^2}\int_0^1 d\alpha_1\int_0^{1-\alpha_1} d\alpha_2 \times \\ \times\frac{(\gamma_\nu[-i((1-\alpha_2)\not{p}_2 - \alpha_1\not{p}_1) + m]\gamma_\mu[-i((1-\alpha_1)\not{p}_1 - \alpha_2\not{p}_2) + m]\gamma_\nu)_{\alpha\beta}}{F}\ , \end{array} \tag{8.15}$$

where

$$\begin{aligned} F &= \alpha_1(1-\alpha_1)p_1^2 + \alpha_2(1-\alpha_2)p_2^2 - 2\alpha_1\alpha_2 p_1\cdot p_2 + m^2(\alpha_1+\alpha_2) \\ &= m^2[-\alpha_1(1-\alpha_1) - \alpha_2(1-\alpha_2) - 2\alpha_1\alpha_2 + \alpha_1 + \alpha_2] \\ &= -m^2(\alpha_1+\alpha_2)^2; \end{aligned} \tag{8.16}$$

Then one finds (it is left as an exercise to prove it)

$$\begin{array}{l} \bar{\psi}(p_2)\Gamma^{(1b)}_\mu(p_1,p_2)\psi(p_1) = -i\frac{e^3}{(4\pi)^2}\int_0^1 d\alpha_1\int_0^{1-\alpha_1} d\alpha_2 \times \\ \times\bar{\psi}(p_2)\frac{[m^2\gamma_\mu((\alpha_1+\alpha_2)^2 - 2(1-\alpha_1-\alpha_2)) + 8imq^\nu\sigma_{\mu\nu}(\alpha_1 - \alpha_2(\alpha_1+\alpha_2))]}{F}\psi(p_1). \end{array} \tag{8.17}$$

But we are interested only in the $\sigma_{\mu\nu}$ term, leading to

$$\bar{\psi}(p_2)\Gamma^{(1)}(p_1,p_2)\psi(p_1)\big|_{\sigma_{\mu\nu}} = \frac{e^3}{2m\pi^2}\bar{\psi}(p_2)\sigma_{\mu\nu}q^\nu\psi(p_1)\int_0^1 d\alpha_1\int_0^{1-\alpha_1} d\alpha_2\frac{\alpha_1 - \alpha_2(\alpha_1+\alpha_2)}{(\alpha_1+\alpha_2)^2}$$

$$= \frac{e^3}{16m\pi^2}\bar{\psi}(p_2)\sigma_{\mu\nu}q^\nu\psi(p_1). \tag{8.18}$$

The term in the effective action has the above multiplied by A_μ, giving

$$\frac{e^2}{8\pi^2}\left(\frac{e}{2m}\right)\bar{\psi}(p_2)\sigma^{\mu\nu}q_\nu A_\mu(q)\psi(p_1)\ , \tag{8.19}$$

but we have $F_{\mu\nu} = 2q_{[\mu}A_{\nu]}$, so

$$\frac{\Delta g}{4} = \frac{e^2}{16\pi^2} \Rightarrow \Delta g = \frac{e^2}{4\pi} = 2\frac{\alpha}{2\pi}. \tag{8.20}$$

Then,

$$g = 2\left(1 + \frac{\alpha}{2\pi} + ...\right)\ , \tag{8.21}$$

like advertised.

Lamb shift

We now move on to the Lamb shift, which was the calculation that finally convinced people of the reality of QFT radiative (loop) corrections. Indeed, this was the first example of a calculation that could not be obtained in any other way, but only through QED loops.

The Lamb shift is the lifting of the degeneracy of the $2S_{1/2}$ and $2P_{1/2}$ energy levels of the H atom.

We will only show the steps leading to the calculation of the Lamb shift, since the loop corrections themselves are difficult, and require the treatment of IR divergences, which will be dealt with later.

Step 1.

We start with the nonrelativistic Schrödinger equation analysis of the H atom, the first success of quantum mechanics. The equation

$$\left[-\frac{\Delta}{2m} + V(r)\right]\psi = E\psi \tag{8.22}$$

becomes

$$\left[-\frac{1}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{l(l+1)}{r^2}\right) - \frac{\alpha}{r}\right]\psi_{n,l}(r) = E_{n,l}\psi_{n,l}(r). \tag{8.23}$$

Here $\alpha = e^2/4\pi$ as usual, and the mass is the reduced mass of the nucleus-elctron system,

$$\frac{1}{m} = \frac{1}{m_e} + \frac{1}{m_N} \simeq \frac{1}{m_e}. \tag{8.24}$$

Then the energy levels of the H atom are

$$E_{n,l} = -\frac{m\alpha^2}{2n^2}\ , \tag{8.25}$$

so are independent of $l = 0, 1, ..., n-1$, as well as of $m_z = -l, ..., l$, giving a degeneracy of

$$\sum_{l=0}^{n-1}(2l+1) = n^2\ , \tag{8.26}$$

and the energy is given by the Rydberg constant

$$R = \frac{m\alpha^2}{2} = 13.6eV. \tag{8.27}$$

Step 2.

Next, we move on to the analysis of the Dirac equation.

As we saw, $-(\not{D} - m)(\not{D} + m)\psi = 0$ gives

$$\left(-D^2 + m^2 - \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}\right)\psi = 0. \tag{8.28}$$

We consider the Coulomb potential of a static charge,

$$eA_0 = -\frac{\alpha}{r} \Rightarrow eE_i = -\frac{\alpha\hat{r}_i}{r^2}. \tag{8.29}$$

Then, since

$$\sigma^{i0} = \frac{[\gamma^i, \gamma^0]}{2i} = i\begin{pmatrix}\sigma_i & \mathbf{0} \\ \mathbf{0} & -\sigma_i\end{pmatrix}, \tag{8.30}$$

we obtain

$$\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} = e\sigma^{i0}F_{i0} = \mp i\alpha\frac{\sigma_i\hat{r}_i}{r^2}. \tag{8.31}$$

With a stationary ansatz for the wave function in spherical coordinates,

$$\psi(x_i, t) = e^{iEt}\psi_\pm(r, \theta, \phi), \tag{8.32}$$

and using

$$D^2 = (\partial_\mu - ieA_0\delta^0_\mu)(\partial^\mu - ieA^0\delta^\mu_0) = -\partial_t^2 + \Delta - e^2A_0^2 + 2ieA_0\partial_0, \tag{8.33}$$

and considering that $\partial_0 = +iE$ on the ansatz, giving $2E\alpha/r$ for the last term, we get the equation

$$\left[-\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{\vec{L}^2 - \alpha^2 \pm i\alpha\sigma_i\hat{r}_i}{r^2} - \frac{2\alpha E}{r} - (E^2 - m^2)\right]\psi_\pm = 0. \tag{8.34}$$

Here $\vec{L}^2 = l(l+1)$ on the wavefunction, and with $\vec{J} = \vec{L} + \vec{S} = \vec{L} + \vec{\sigma}/2$, $[H, \vec{J}] = 0 = [\vec{L}^2, \vec{J}]$, we have also $\vec{J}^2 = j(j+1)$, with $J_z = m$ and $l = j \pm 1/2$.

Then we can prove that $\vec{L}^2 - \alpha^2 \pm i\alpha\sigma_i\hat{r}_i$ has eigenvalues $\lambda(\lambda+1)$, where

$$\begin{aligned}\lambda_\pm &= j \pm \frac{1}{2} - \delta_j \\ \delta_j &= j + \frac{1}{2} - \sqrt{\left(j + \frac{1}{2}\right)^2 - \alpha^2}.\end{aligned} \tag{8.35}$$

Then we note that the resulting equation is formally the same as the Schrödinger equation, just with the replacements

$$\begin{aligned}
\vec{L}^2 &\rightarrow \vec{L}^2 - \alpha^2 \pm i\alpha\sigma_i\hat{r}_i \\
l(l+1) &\rightarrow \lambda(\lambda+1) \\
\alpha &\rightarrow \alpha\frac{E}{m} \\
E &\rightarrow \frac{E^2-m^2}{2m}\ ,
\end{aligned} \tag{8.36}$$

and we also note that, because the resulting equation is an eigenvalue problem, the condition of $n-l$ to be an integer is replaced by the condition of $n-\lambda = n-l+\delta_j$ to be an integer (where $n=\sqrt{R/E}$), effectively replacing the integer n with $n-\delta_j$ in the solution to the eigenvalue problem.

Finally, we obtain the energy quantization

$$\frac{E_{nj}^2-m^2}{2m} = -\frac{m\alpha^2}{2}\frac{E_{nj}^2}{m^2}\frac{1}{(n-\delta_j)^2}\ , \tag{8.37}$$

which can be solved to give

$$E_{nj} = \frac{m}{\sqrt{1+\frac{\alpha^2}{(n-\delta_j)^2}}} = m - \frac{m\alpha^2}{2n} - \frac{m\alpha^4}{n^3(2j+1)} + \frac{3}{8}\frac{m\alpha^4}{n^4} + \mathcal{O}(\alpha^6). \tag{8.38}$$

We see that the first term is the rest mass, the second is the quantum mechanics result, and the third and forth are new, with the third lifting the degeneracy over j, i.e. giving a *fine structure*.

With the usual notation of energy levels nl_j, the degeneracy split between the $2P_{3/2}$ and the $2P_{1/2}$ levels is now

$$E(2P_{3/2}) - E(2P_{1/2}) \simeq \frac{m\alpha^4}{32} = 4.5\times 10^{-5}eV = 10.9GHz. \tag{8.39}$$

But the degneracy of over $l = j\pm 1/2$ at fixed j is not lifted, so at this point $E(2S_{1/2}) = E(2P_{1/2})$.

We need to include other effects now:

-the nucleus has a finite size, it is not a point.

-the proton recoils, since it has a finite mass, so $m\simeq m_e$ has to be corrected.

-the proton has a magnetic moment, and so it interacts with the electron spin, giving

$$\Delta E = -\frac{e}{2m}\sigma_i^{(e)}B\ , \tag{8.40}$$

where B is the magnetic field of the proton. This gives the *hyperfine splitting*,

$$\Delta E_{h.f.}(S) = 5.9\times 10^{-6}eV = 1.4GHz. \tag{8.41}$$

All these 3 effects can be treated semiclassically, but do not account for the observations.

When taken into account, there is still a *Lamb shift* of

$$E(2S_{1/2}) - E(2P_{1/2}) \simeq 1057 MHz. \tag{8.42}$$

This comes entirely from radiative corrections. The point is that the interaction vertex of the photon with two fermions changes

$$e\gamma_\mu A^\mu \to (e\gamma_\mu + \Gamma_\mu^{(1)} + \Pi_{\mu\nu}G^{\nu\lambda}\gamma_\lambda)A^\mu. \tag{8.43}$$

The first term gives the classical interaction with the Coulomb potential $A_\mu^{\rm Coulomb}$. The relevant diagrams are in Fig.20.

1. The $\Pi_{\mu\nu}$ term, the loop correction to the photon propagator connecting the vertex with the Coulomb source, was partly computed before, and gives a contribution of $-27MHz$.

2. The $\Gamma_\mu^{(1)}$ term contains the $\sigma_{\mu\nu}q^\nu$ piece partly computed for the anomalous magnetic moment (though one needs to take care of the IR divergences), which gives a contribution of $+68MHz$, and

3. The γ_μ piece, whose finite part we have not computed completely. This is difficult, since it contains IR divergences, that need to be dealt with. It gives the largest contribution, of 1010MHz.

In total these 3 contributions sum up to 1051 MHz, but by considering higher orders in α, one can go to $1057.864 \pm 0.014 MHz$, in perfect agreement with the experimental Lamb shift.

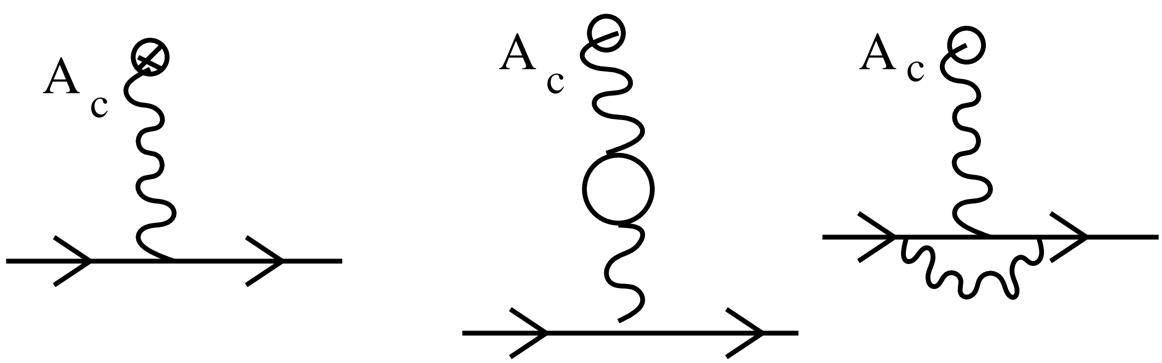

Figure 20: Contributions to the potential energy: Coulomb part; $\Pi_{\mu\nu}$ part; Γ_μ part.

Important concepts to remember

- The corrections to the anomalous magnetic moment of the electron, specifically to $g-2$, arise from $\bar\psi\sigma^{\mu\nu}F_{\mu\nu}\psi$ terms in the quantum effective action.

- The finite vertex correction $\Gamma_\mu^{(1b)}$ gives such a contribution, of $\Delta g = \alpha/\pi$.

- The Lamb shift is the lifting of the degeneracy of the energy levels $2S_{1/2}$ and $2P_{1/2}$ of the H atom.
- In nonrelativistic quantum mechanics, the energy depends only on n, but not on j or l, $E = -m\alpha^2/2n^2$.
- In relativistic quantum mechanics, i.e. the Dirac equation, the degeneracy over j is lifted at order α^4.
- The degeneracy over l is lifted only in through radiative corrections in QED.
- The Lamb shift is due to a photon propagator correction, $\Pi_{\mu\nu}$, and a vertex correction Γ_μ, splitting into a $\sigma_{\mu\nu}q^\nu$ piece and the γ_μ piece, giving the leading contribution.

Further reading: See chapter 6.5.3 in [5].

Exercises, Lecture 8

1) Fill in the omitted steps in the calculation of $\bar{\psi}(p_2)\Gamma_\mu^{(1b)}\psi(p_1)$.

2) Write down an integral expression using the Feynman rules for the 3 contributions to the Lamb shift, do the gamma matrix algebra, and isolate the γ_μ and $\sigma_{\mu\nu}q^\nu$ pieces of the Γ_μ diagram.

9 Lecture 9. Two-loop example and multiloop generalization

In this lecture, we will see how to renormalize a theory beyond one-loop, by first discussing the systematics, and then giving the example of the 2-loop 4-point function in ϕ^4 theory in 4 dimensions.

When renormalizing beyond the leading (one-loop) order, we have:

1) Divergent subdiagrams from the divergent one-loop (or, in general, $(n-1)$-loop) diagrams. They are cancelled by adding diagrams with an insertion of the corresponding one-loop (or, in general, $(n-1)$-loop) counterterm vertices.

2) Intrinsically divergent 2-loop (or, in general, n-loop) diagrams. They are cancelled by adding *new* n-loop counterterm contributions, i.e. a 2-loop (or n-loop) correction to the counterterm Lagrangean.

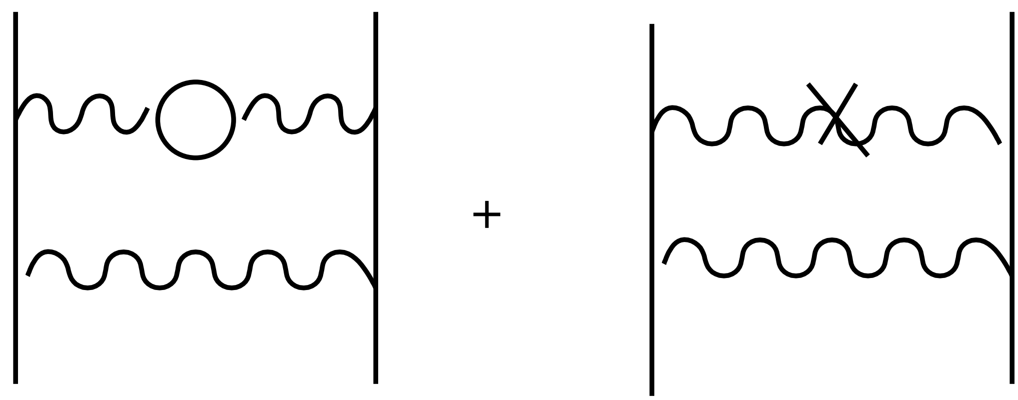

Figure 21: Two Loop independent subdiagram cancelled by a diagram with a one-loop counterterm.

But the first type admits a further subdivision, into:

1a) independent subdiagrams. For example, in QED we can have a fermion loop insertion on a photon line in the diagram, cancelled by the same diagram with the fermion loop substituted with the one-loop counterterm vertex. These divergences are polynomials in q^2 (though they can be logarithmic in the cut-off), which means that when Fourier transforming to x space, these divergences will be *local* (the Fourier transform of a power is a power, but the Fourier transform of a log, which is an infinite power series, is an infinite power series, i.e. nonlocal).

1b) *nested, or overlapping* divergences. In this case, two divergent loops share a propagator, and we will see that they correspond to nonlocal divergences (non-polynomials in q^2). For instance, in ϕ^4 theory we can have the "setting Sun" diagram for the propagator, with 2 vertices connected by 3 propagators, and for the same vertices coming out the 2 external lines, as in Fig.22. Or the diagram for the 4-point function, the same as for the setting Sun, but with one more vertex with two external lines on one of the 3 propagators. In both these

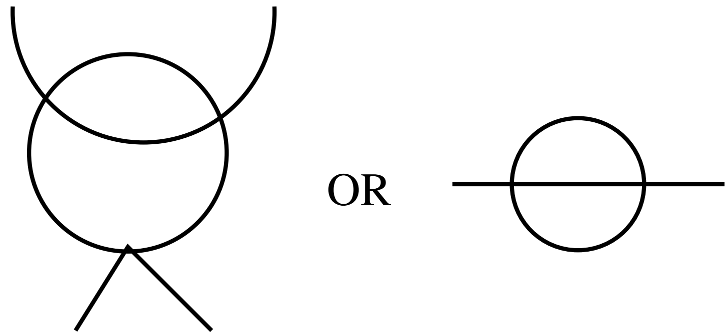

Figure 22: Two Loop Nested divergent diagrams in ϕ^4 theory.

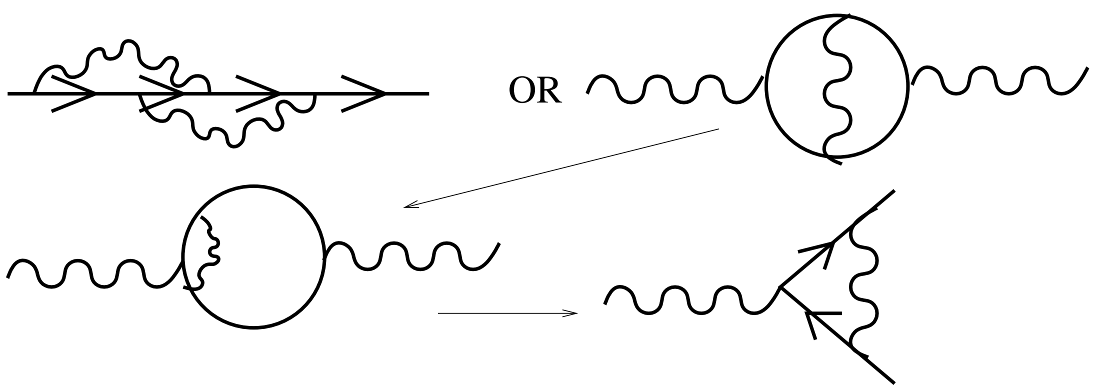

Figure 23: Two Loop Nested divergent diagrams in QED. The photon polarization diagram can be viewed as coming from a vertex correction.

cases, the divergent one-loop subdiagram is one with 2 propagators between two vertices, and two more external legs from each vertex. One of the propagators of the one-loop subdiagrams is common. In QED, we can have the fermion propagator 2-loop diagram with a photon line starting off on the fermion line and returning to it, and another photon line starting in between the endpoints of the first, and ending further on the fermion line, as in Fig.23. Or we can have a two-loop diagram for the photon propagator, with a fermion loop on the photon line, and another photon vertical line (with endpoints on both sides of the external line vertex).

Consider this last QED diagram. Its divergence includes the one-loop divergence of the photon-fermions vertex, which is of the type $(-ie\gamma_\mu)(\alpha \log \Lambda^2)$, where the first bracket isolates the classical vertex, and the second involves factors from the quantum correction. Then, when inserted inside a fermion loop correction to the photon propagator (giving in total the 2-loop diagram we are describing), the classical vertex is replaced with this quantum-corrected vertex. Since the fermion loop correction is of the type $\alpha(g^{\mu\nu}q^2 - q^\mu q^\nu)\Pi(q^2)$, and

$\Pi(q^2) \sim \log \Lambda^2 + \log q^2$, we have in total,

$$\sim \alpha(g^{\mu\nu}q^2 - q^\mu q^\nu)(\log \Lambda^2 + \log q^2)\alpha \log \Lambda^2. \quad (9.1)$$

That means that there is a part,

$$\alpha[\log q^2 \alpha \log \Lambda^2] \;, \quad (9.2)$$

which is divergent in Λ but nonpolynomial in q^2, since it comes from the divergent part of one divergence, times the finite part of the other divergence, which is non-polynomial in q^2 (there is no problem in having a non-polynomial for the finite, observable, part; the problem is for the divergent part, since all the terms in the action are local, i.e. polynomial in q^2, yet we need to remove this divergence by renormalization). This divergence is therefore non-local in x space, but it can be cancelled by adding *loop diagrams* with the 1-loop counterterm vertices, namely two one-loop fermion corrections to the photon propagator, with either one of the vertices replaced by the one-loop counterterm vertex, as in Fig.24.

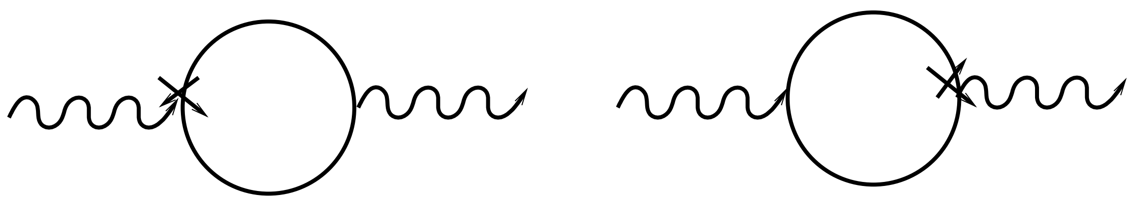

Figure 24: Two Loop level counterterm diagrams: one loop diagrams with the one-loop counterterm vertex in it.

There is of course still a part $\alpha^2(\log \Lambda^2)^2$, which however is local, so can be cancelled by the addition of a new, 2-loop, contribution to the local counterterm Lagrangean (local 2-loop counterterm vertex).

This procedure generalizes in an obvious manner to all loops.

It is a nontrivial fact that in this way, we can cancel *all* the divergences in the theory with a finite number of local counterterms (though each term with a coefficient being an infinite series expansion in loops, i.e. in powers of the coupling), if the theory is renormalizable. There is a theorem proving this fact, that any superficially divergent theory is rendered finite by the above prescription for counterterms, called the BPHZ theory, after the authors Bogoliubov and Parasiuk; Hepp; and Zimmermann.

2-loop in ϕ^4 in 4 dimensions

We consider the 2-loop contributions to the 4-point function of ϕ^4 theory in 4 dimensions, and show that we can remove all the divergence by renormalization according to the above prescription.

There are 16 diagrams corresponding to this order, 15 of which can be split into 3 groups, according to the channel, s, t and u, as in Fig.25. In the s channel, the first diagram is a "chain" made up of two one-loop "rings", and with two external lines from a vertex at each

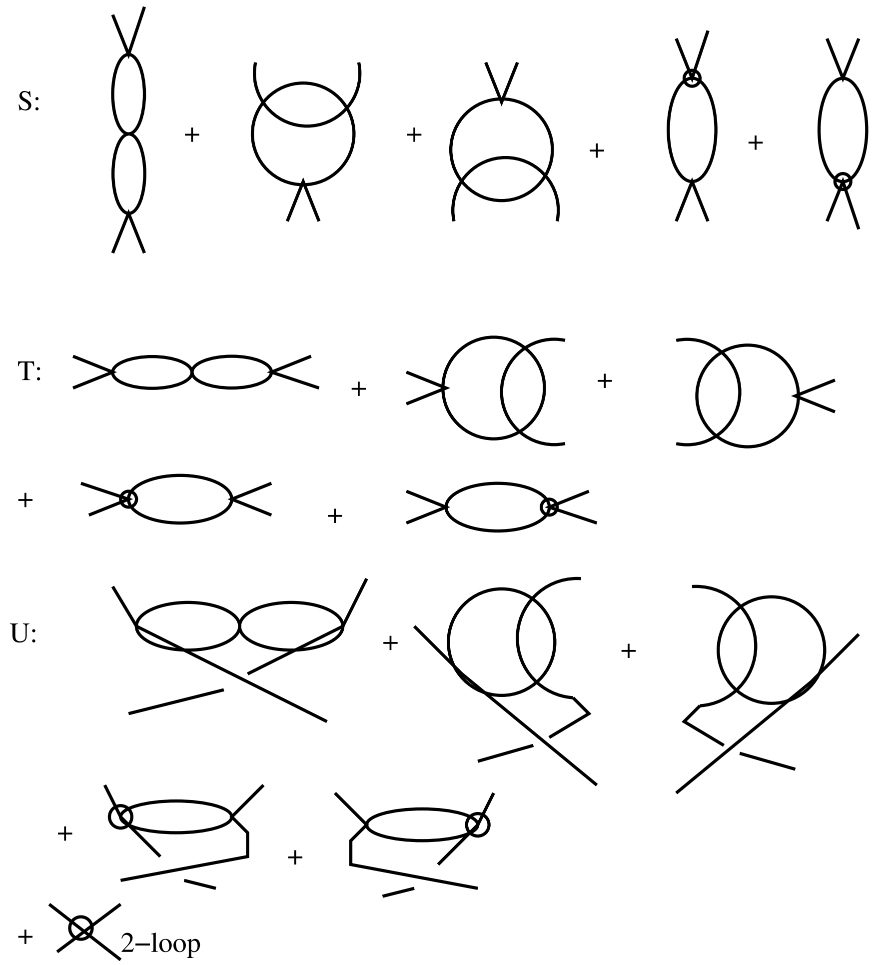

Figure 25: Two Loop diagrams in ϕ^4 theory, organized according to the s, t, u channel diagrams, plus the 2-loop counterterm vertex diagram.

end. The second is the diagram already described, the setting Sun with an extra vertex with 2 external lines on one of the propagators, coming down. The third is the up-down mirror of the first. Finally, we have the one-loop diagram (with two up and two down external lines, each pair from a vertex) with one normal vertex and one vertex being the one-loop counterterm vertex. Then there are the t and u channel version of the same 5 diagrams, which can be obtained by crossing (that exchanges s, t, u). The last diagram is of course the 2-loop counterterm vertex, which contains s, t and u pieces.

Therefore to fully renormalize the 4-point vertex at 2-loops, we need only consider the 5 s channel diagrams, and the s piece of the 2-loop counterterm vertex, as in Fig.26, and the rest will be trivially obtained by crossing.

One-loop. Let us first remember the one-loop renormalization. There is a unique s

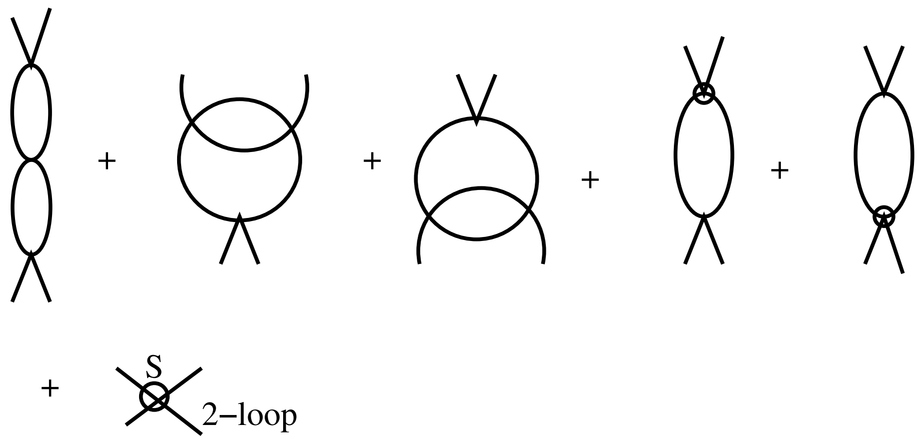

Figure 26: Two Loop independent diagrams in ϕ^4 theory. The other can be related by crossing. The 2-loop counterterm vertex diagram contains only the s piece.

channel diagram, with momentum $p = p_1 + p_2$ coming in. The result for the diagram is

$$\frac{\tilde{\lambda}^2}{2} I_2(p^2) = \frac{\tilde{\lambda}^2}{2} \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 + m^2} \frac{1}{(q+p)^2 + m^2} = \frac{\tilde{\lambda}^2}{2} \frac{\Gamma\left(2 - \frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}} \int_0^1 d\alpha [\alpha(1-\alpha)p^2 + m^2]^{\frac{D}{2}-2}. \tag{9.3}$$

We consider the normalization conditions

$$\Gamma^{(2)}(p^2) = [p^2 + m^2]^{-1}; \quad \Gamma^{(4)}(s,t,u) = -\lambda \tag{9.4}$$

at $s = 4m^2$; $t = u = 0$, where $s = -(p_1 + p_2)^2$, etc. Then we immediately obtain the one-loop counterterm vertex as

$$V^{(1)} = -\frac{\tilde{\lambda}^2}{2} [I_2(4m^2) + 2I_2(0)] \ , \tag{9.5}$$

which splits into an s piece,

$$V_s^{(1)} = -\frac{\tilde{\lambda}^2}{2} I_2(4m^2) \ , \tag{9.6}$$

and an $(t+u)$-piece,

$$V_{t+u}^{(1)} = -\frac{\tilde{\lambda}^2}{2} 2I_2(0). \tag{9.7}$$

Then, we can finally split the 6 independent s-channel diagrams into 3 groups, as in Fig.27,

I) the diagram of a chain of 2 one-loop rings, plus the one loop diagrams with one of the vertices replaced by the s part of the one-loop counterterm vertex, $V_s^{(1)}$.

II) the setting Sun with extra vertex down diagram, plus the one-loop diagram with normal vertex down, and $t+u$ part of the one-loop counterterm vertex, $V_{t+u}^{(1)}$, up.

III) the same diagrams as at II, but with up and down interchanged (mirror symmetric).

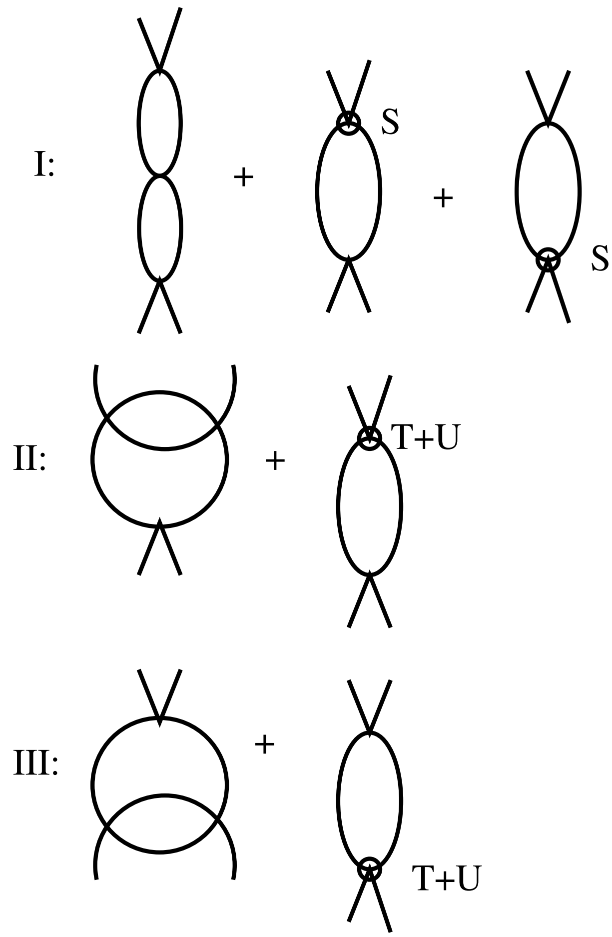

Figure 27: Two Loop independent diagrams in ϕ^4 theory divided into 3 groups. The diagrams with the one-loop counterterm vertices have been split into the s pieces and the $t+u$ pieces.

Then the *momentum-dependent part of the divergence* cancels separately in I, II and III. We also see that we only need to calculate what happens in I and II, since III is obtained from II.

I) we start with the "chain" diagram of two one-loop rings, the first one in group I, which can be split into two independent one-loop diagrams, each with momentum p coming into it, so the result of the diagram is

$$-\frac{\tilde{\lambda}^3}{4}[I_2(p^2)]^2. \tag{9.8}$$

Note that the only nontrivial part is the vertex counting, which is $(-\tilde{\lambda})^3$ (one vertex is common to the two one-loop rings), the integral (as well as the symmetry factor of 1/2 for each ring) comes from the one-loop diagrams.

Both diagrams with a $V_s^{(1)}$ insertion (up or down), i.e. the second and third diagrams in group I, equal a $V_s^{(1)}$ factor, times a vertex $(-\tilde{\lambda})$, times the one-loop integral $I_2(p^2)$, times the 1/2 for the symmetry factor, for a total of

$$-\frac{\tilde{\lambda}}{2}I_2(p^2)\left(-\frac{\tilde{\lambda}^2}{2}\right)I_2(4m^2)=\frac{\tilde{\lambda}^3}{4}I_2(p^2)I_2(4m^2). \tag{9.9}$$

The sum of the diagrams in I is therefore

$$-\frac{\tilde{\lambda}^3}{4}[(I_2(p^2))^2-2I_2(p^2)I_2(4m^2)]=-\frac{\tilde{\lambda}^3}{4}\left([I_2(p^2)-I_2(4m^2)]^2-[I_2(4m^2)]^2\right)\ , \tag{9.10}$$

and the first term is finite, because it is the same finite term appearing in the one-loop renormalization,

$$I_2(p^2)-I_2(4m^2)=\frac{2}{\epsilon}\frac{1}{(4\pi)^2}\left(-\frac{\epsilon}{2}\right)\int_0^1 d\alpha\ln\left[\frac{m^2+\alpha(1-\alpha)p^2}{m^2+\alpha(1-\alpha)4m^2}\right]. \tag{9.11}$$

Note that the divergence is constant, so cancelled between the two terms, and also the $\psi(1)=-\gamma$ term has cancelled. In the two-loop formula, we have the square of the above, for a total of

$$\frac{\tilde{\lambda}^3}{4}\left[\frac{1}{16\pi^2}\int_0^1 d\alpha\ln[...]\right]^2. \tag{9.12}$$

Therefore in I, we have cancelled a product of two *independent divergences*, and we are left with only a constant divergence,

$$+\frac{\tilde{\lambda}^3}{4}[I_2(4m^2)]^2\propto\left[\Gamma\left(2-\frac{D}{2}\right)\right]^2\propto\left(\frac{2}{\epsilon}\right)^2\ , \tag{9.13}$$

which is a double pole in ϵ. This term is cancelled by adding a 2-loop counterterm, with vertex

$$-\frac{\tilde{\lambda}^3}{4}[I_2(4m^2)]^2. \tag{9.14}$$

Some observations are in order. The first one is that at higher loops we will see higher order poles in ϵ, as seen here. But in all cases, at least the highest order will be a momentum-independent constant. The second observation related the functional form of the finite part. Note that at $p^2\to\infty$,

$$I_2(p^2)-I_2(4m^2)\sim\log\frac{p^2}{m^2}\ , \tag{9.15}$$

whereas at two-loops, because of the above, we can ignore the finite contribution of $I_2(4m^2)$ with respect to $I_2(p^2)$, and write that the 2-loop "chain" of two rings gives

$$\sim\frac{\tilde{\lambda}^3}{4}[I_2(p^2)]^2\propto\left(\log\frac{p^2}{m^2}\right)^2. \tag{9.16}$$

It is then easy to see that this generalizes to an arbitrary n-loop, where the "chain" diagram, with n one-loop rings, as in Fig.28, will give a result

$$\propto\lambda^{n+1}\left(\log\frac{p^2}{m^2}\right)^n. \tag{9.17}$$

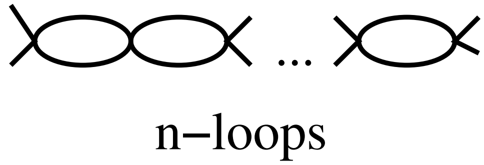

Figure 28: n-loop "chain" diagram.

II) We move on to the second set of diagrams, starting with the proper 2-loop diagram in Fig.29. Consider momenta p_1, p_2 at the classical vertex (down), connected to two internal lines of momenta k (loop momentum) and $k+p$ (where as above, $p = p_1 + p_2$), and the other external momenta being p_3 (for the vertex connected to the k propagator) and p_4. Then we can isolate on top a one-loop diagram with total incoming momentum $k+p_3$, inserted inside an up-down one-loop diagram with total momentum $p = p_1 + p_2$ coming in. The symmetry factor of the diagram is 2, giving for the amplitude

$$\mathcal{M} = -\frac{\tilde{\lambda}^3}{2} \int \frac{d^D k}{(2\pi)^D} \frac{1}{k^2 + m^2} \frac{1}{(k+p)^2 + m^2} I_2((k+p_3)^2). \tag{9.18}$$

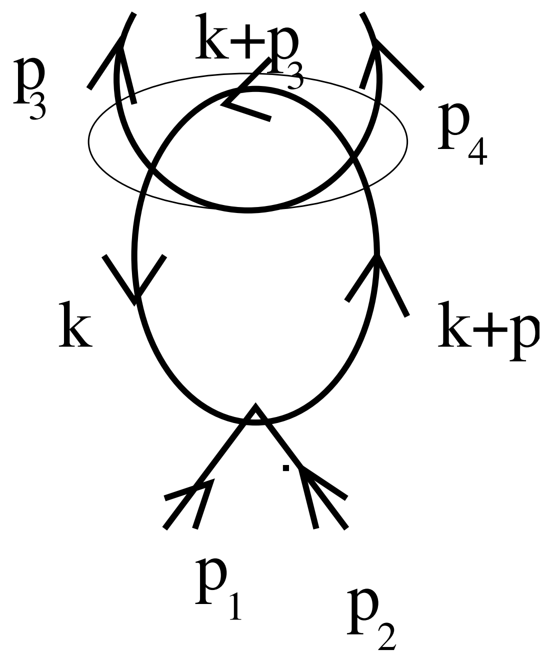

Figure 29: Two Loop Setting Sun diagram.

Replacing the two propagators with a Feynman parametrization integral over β, and writing the result of I_2 as a Feynman parametrization integral over α, we obtain

$$\mathcal{M} = -\frac{\tilde{\lambda}^3}{2}\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}\int_0^1 d\alpha\int_0^1 d\beta\frac{d^Dk}{(2\pi)^d}\frac{1}{[k^2+2\beta k\cdot p+\beta p^2+m^2]^2}\frac{1}{[\alpha(1-\alpha)(k+p_3)^2+m^2]^{2-\frac{D}{2}}} \tag{9.19}$$

where the first denominator is $\beta\Delta_2+(1-\beta)\Delta_1$. But now, since the second denominator has a non-integer power, we cannot use the usual Feynman parametrization, but rather we must use the formula

$$\frac{1}{A^\alpha B^\beta} = \int_0^1 dw\frac{w^{\alpha-1}(1-w)^{\beta-1}}{[wA+(1-w)B]^{\alpha+\beta}}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\;, \tag{9.20}$$

where the gamma factors together make the inverse of the beta function, $[B(\alpha,\beta)]^{-1}$.

The proof of this formula goes as follows. Consider the change of variables

$$\begin{aligned} z \;\equiv\; & \frac{wA}{wA+(1-w)B}\Rightarrow 1-z=\frac{(1-w)B}{wA+(1-w)B}\\ & \Rightarrow dz=\frac{ABdw}{[wA+(1-w)B]^2}. \end{aligned} \tag{9.21}$$

Then

$$\int_0^1 dw\frac{w^{\alpha-1}(1-w)^{\beta-1}}{[wA+(1-w)B]^{\alpha+\beta}} = \frac{1}{A^\alpha B^\beta}\int_0^1 dz\; z^{\alpha-1}(1-z)^{\beta-1} = \frac{1}{A^\alpha B^\beta}B(\alpha,\beta). \tag{9.22}$$

q.e.d.

Then applying it for the case of our diagram (with $\alpha=2-D/2$ and $\beta=2$), we obtain

$$\begin{aligned} \mathcal{M} \;=\; & -\frac{\tilde{\lambda}^3}{2}\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}\frac{\Gamma\left(4-\frac{D}{2}\right)}{\Gamma(2)\Gamma\left(2-\frac{D}{2}\right)}\int_0^1 d\alpha\int_0^1 d\beta\int_0^1 dw\int\times\\ & \times\frac{d^Dk}{(2\pi)^D}\frac{w^{1-\frac{D}{2}}(1-w)}{\left(w[\alpha(1-\alpha)(k+p_3)^2+m^2]+(1-w)[k^2+2k\cdot p+\beta p^2+m^2]\right)^{4-\frac{D}{2}}} \end{aligned} \tag{9.23}$$

The factor raised to $4-D/2$ in the denominator is rewritten as

$$\begin{aligned} & m^2+k^2[1-w+w\alpha(1-\alpha)]+2k\cdot[(1-w)\beta p+w\alpha(1-\alpha)p_3]+p_3^2w\alpha(1-\alpha)+p^2\beta(1-w)\\ \equiv\;& m^2+k'^2[1-w+w\alpha(1-\alpha)]+P^2\;, \end{aligned} \tag{9.24}$$

where

$$k' = k+\frac{(1-w)\beta p+w\alpha(1-\alpha)p_3}{1-w+w\alpha(1-\alpha)}\;, \tag{9.25}$$

and

$$P^2 = p_3^2w\alpha(1-\alpha)+p^2\beta(1-w)-\left(\frac{(1-w)\beta p+w\alpha(1-\alpha)p_3}{1-w+w\alpha(1-\alpha)}\right)^2. \tag{9.26}$$

For future reference, we note that

$$P^2(w\to 0) = p^2\beta-\beta^2p^2=\beta(1-\beta)p^2. \tag{9.27}$$

We now can do the integral over $\int d^Dk = \int d^Dk'$ with the formula (3.30), to obtain

$$\begin{aligned}\mathcal{M} &= -\frac{\tilde{\lambda}^3}{2}\frac{\Gamma\left(4-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}\frac{\Gamma(4-D)}{\Gamma\left(4-\frac{D}{2}\right)(4\pi)^{\frac{D}{2}}}\int_0^1 d\alpha\int_0^1 d\beta\int_0^1 dw\frac{1}{[1-w+w\alpha(1-\alpha)]^{4-\frac{D}{2}}}\times\\ &\times w^{1-\frac{D}{2}}(1-w)\left[\frac{P^2+m^2}{1-w+w\alpha(1-\alpha)}\right]^{D-4}\\ &= -\frac{\tilde{\lambda}^3}{2}\frac{\Gamma(4-D)}{(4\pi)^D}\int_0^1 d\alpha\int_0^1 d\beta\int_0^1 dw\frac{w^{1-\frac{D}{2}}(1-w)}{[1-w+w\alpha(1-\alpha)]^{\frac{D}{2}}}[P^2+m^2]^{D-4}. \qquad (9.28)\end{aligned}$$

This diagram has of course a pole coming from $\Gamma(4-D)$, but it has also a pole coming from the integral in w, specifically near $w=0$. Indeed, consider the integral

$$\int_0^1 dw\; w^{1-\frac{D}{2}}f(w) = \int_0^1 dw\; w^{1-\frac{D}{2}}f(0) + \int_0^1 dw\; w^{1-\frac{D}{2}}[f(w)-f(0)]. \qquad (9.29)$$

Then in our case we can check that the second term gives a finite integral for $D=4$, and the only pole comes from the overall $\Gamma(4-D)$. This term therefore is local (there is no p dependence at all if we set $D=4$), and its divergence can be absorbed by an $\mathcal{O}(\lambda^3)$ counterterm, i.e. a two-loop, $1/\epsilon$ term.

The first term, with $f(0)$, gives a double pole, since

$$\int_0^1 dw\; w^{1-\frac{D}{2}}f(0) = f(0)\left.\frac{w^{2-\frac{D}{2}}}{2-\frac{D}{2}}\right|_0^1 = f(0)\frac{2}{4-D}. \qquad (9.30)$$

Substituting $D=4-\epsilon$, we obtain

$$\begin{aligned}&-\frac{\tilde{\lambda}^3\Gamma(\epsilon)}{2(4\pi)^D}\frac{2}{\epsilon}\int_0^1 d\beta[P^2(0)+m^2]^{-\epsilon}\\ &= -\frac{\lambda^3\mu^\epsilon}{(4\pi)^4}\frac{1}{\epsilon}\int_0^1 d\beta\left(\frac{1}{\epsilon}-\gamma-\log\frac{\beta(1-\beta)p^2+m^2}{4\pi\mu^2}\right), \qquad (9.31)\end{aligned}$$

where in the last line we have grouped, as usual, the $(4\pi)^{-\epsilon}$ and the $\mu^{-2\epsilon}$ terms in the expansion with the log, to make the ratio of $m^2/(4\pi\mu^2)$ manifest.

Therefore we have finally the nonlocal divergence

$$\frac{\lambda^3\mu^\epsilon}{(4\pi)^4}\frac{1}{\epsilon}\int_0^1 d\beta\log\frac{\beta(1-\beta)p^2+m^2}{4\pi\mu^2}. \qquad (9.32)$$

This divergence is cancelled however by the one-loop diagram with one one-loop $t+u$ counterterm vertex $V^{(1)}_{t+u}$, given by the product of the vertex $V^{(1)}_{t+u}$ and the $(-\tilde{\lambda}^2/2I_2(p^2))$ factor (removing one $-\tilde{\lambda}$ vertex and replacing it with $V^{(1)}_{t+u}$,

$$+\frac{\tilde{\lambda}^3}{4}2I_2(0)I_2(p^2)$$

$$
\begin{aligned}
&= \frac{\tilde{\lambda}^3}{2(4\pi)^D}\int_0^1 d\alpha \frac{\left[\Gamma\left(2-\frac{D}{2}\right)\right]^2}{[\alpha(1-\alpha)p^2+m^2]^{2-\frac{D}{2}}(m^2)^{2-\frac{D}{2}}} \\
&= \frac{\lambda^3\mu^\epsilon}{2(\pi)^4}\int_0^1 d\alpha\left(\frac{2}{\epsilon}-\gamma-\log\frac{m^2}{4\pi\mu^2}\right)\left(\frac{2}{\epsilon}-\gamma-\log\frac{m^2+\alpha(1-\alpha)p^2}{4\pi\mu^2}\right). \qquad (9.33)
\end{aligned}
$$

We see now that the nonlocal divergence cancels, and we are only left with a local divergence of order $1/\epsilon^2$ (constant double pole). This will be removed by adding an $\mathcal{O}(\lambda^3)$ counterterm (local two-loop counterterm).

As we mentioned, the III diagrams are obtained by symmetry, and then the t and u channels by crossing, so it means we have indeed shown the full renormalizability of the 4-point function at 2-loops in ϕ^4 theory in 4 dimensions.

Important concepts to remember

- Intrinsically divergent 2-loop diagrams are cancelled by adding new 2-loop (n-loop) counterterms.
- Divergences from 1-loop ($(n-1)$-loop) divergent subdiagrams are cancelled by adding diagrams with the 1-loop ($(n-1)$-loop) counterterm vertices.
- Independent subdiagrams lead to local divergences, i.e. polynomial in momenta q^2, whereas nested (overlapping) divergences lead to non-local divergences, i.e. non-polynomial in momenta q^2.
- The BPHZ theorem says that any superficially renormalizable theory is rendered finite by adding the diagrams with all the n-loop counterterms, which are a finite number of local counterterms (with coefficients = a series in loop order, or λ^n).
- The finite part of the n-loop chain diagram with n one-loop rings in ϕ^4 theory at $p^2 \to \infty$ goes like $\sim \lambda^{n+1}\log^n(p^2/m^2)$.
- At higher loops in dimensional regularization, there are higher order poles (in ϵ), but at least the highest order pole is a momentum-independent constant.

Further reading: See chapter 10.5 in [3] and chapter 10.4 in [2].

Exercises, Lecture 9

1) Write down all the 2-loop divergent diagrams for $\Gamma^{(3)}$ for ϕ^3 theory in $D = 6$, paralleling ϕ^4 in $D = 4$.

2) Identify and calculate the nonlocal 2-loop divergence in the above.

10 Lecture 10. The LSZ formula.

In this lecture we return to the LSZ formula, relating correlation functions,

$$\langle\Omega|T\{\phi(x_1)...\phi(x_{n+m})\}|\Omega\rangle\ , \tag{10.1}$$

with S-matrices,

$$_{out}\langle\vec{p}_1,...,\vec{p}_n|\vec{k}_1...\vec{k}_m\rangle_{in}=\langle\vec{p}_1,...,\vec{p}_n|S|\vec{k}_1...\vec{k}_m\rangle. \tag{10.2}$$

We described it in QFTI, but now we return to a better understanding for it, using the knowledge of loops and renormalization that we have gained in the meantime.

In particular, we had mentioned that there is a wave function renormalization factor Z that appears there. We also saw that at one-loop, we defined the wave function renormalization factor as the factor in the renormalized Lagrangean that multiplies the p^2 part of the kinetic term, and that it comes from the calculation of the 2-point function. With a bit of thought, we realize that a more formal way to define it would be as follows.

Consider the 2-point function in momentum space. It will have a pole at some renormalized mass $p^2\to -m^2$. So formally, we can say that near the pole $p^2\to -m^2$,

$$\int d^4x e^{-ip\cdot x}\langle\Omega|T\{\phi(x)\phi(0)\}|\Omega\rangle\sim\frac{-iZ}{p^2+m^2-i\epsilon}. \tag{10.3}$$

To obtain the LSZ formula, we consider the Fourier transform over only one momentum of the correlation function,

$$\int d^4x e^{-ip\cdot x}\langle\Omega|T\{\phi(x)\phi(z_1)\phi(z_2)...\}|\Omega\rangle\ , \tag{10.4}$$

and we split the integral over time as

$$\int dx^0=\int_{T_+}^{+\infty}dx^0+\int_{T_-}^{T_+}dx^0+\int_{-\infty}^{T_-}dx^0\ , \tag{10.5}$$

and call the first term region I, the second region II and the third region III.

Since we are interested in the behaviour near a pole, we will ignore finite terms. But the integral over region II is finite, since the integrand is analytic in p^0 and the integration is over a finite interval. Therefore we can ignore this region.

Region I.

Consider then first the integral over region I, and let x^0 be the largest time (i.e., the time components of all the z_i's are $< T_+ < x^0$ in region I. Then we can put $\phi(x)$ to the left, and outside the time ordering operator, and we insert on its right the identity, written in terms of a complete set of states $|\lambda_{\vec{q}}\rangle$ of momentum $\vec{q}$, as

$$\mathbb{1}=\sum_\lambda\int\frac{d^3q}{(2\pi)^3}\frac{1}{2E_{\vec{q}}(\lambda)}|\lambda_{\vec{q}}\rangle\langle\lambda_{\vec{q}}|. \tag{10.6}$$

We obtain

$$\int d^3x \int_{T_+}^{+\infty} dx^0 e^{+ip^0x^0 - i\vec{p}\cdot\vec{x}} \sum_\lambda \int \frac{d^3q}{(2\pi)^3} \frac{1}{2E_{\vec{q}}(\lambda)} \langle\Omega|\phi(x)|\lambda_{\vec{q}}\rangle\langle\lambda_{\vec{q}}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle. \quad (10.7)$$

But the first matrix element can be worked out as follows, using that $\phi(x)$ is a Heisenberg operator, so

$$\langle\Omega|\phi(x)|\lambda_{\vec{q}}\rangle = \langle\Omega|e^{-i\hat{P}\cdot x}\phi(0)e^{+i\hat{P}\cdot x}|\lambda_{\vec{q}}\rangle = \langle\Omega|\phi(0)|\lambda_{\vec{q}}\rangle\; e^{+iq\cdot x}\Big|_{q^0=E_{\vec{q}}} , \quad (10.8)$$

where in the last equality we have used that $\hat{P}\cdot x|\lambda_{\vec{q}}\rangle = (q\cdot x)|_{q^0=E_{\vec{q}}}|\lambda_{\vec{q}}\rangle$ and $\langle\Omega|\hat{P}\cdot x = 0$. Now doing the integral over d^3x, $\int d^3x e^{i\vec{x}\cdot(\vec{q}-\vec{p})} = (2\pi)^3\delta^3(\vec{p}-\vec{q})$, and then doing the integral over $\int d^3q/(2\pi)^3$, we replace everywhere $\vec{q}$ with $\vec{p}$ and $E_{\vec{q}}$ with $E_{\vec{p}}$. Introducing a regularizing factor of $e^{-\epsilon x^0}$ in the usual manner, we obtain

$$\sum_\lambda \int_{T_+}^{+\infty} dx^0 \frac{1}{2E_{\vec{p}}(\lambda)} e^{+ip^0x^0}\; e^{-iq^0x^0}\Big|_{q^0=E_{\vec{q}}=E_{\vec{p}}} e^{-\epsilon x^0}\langle\Omega|\phi(0)|\lambda_{\vec{p}}\rangle\langle\lambda_{\vec{p}}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle. \quad (10.9)$$

The integral over t becomes

$$\int_{T_+}^{+\infty} e^{ix^0(p^0-E_p+i\epsilon)} = \frac{e^{ix^0(p^0-E_p+i\epsilon)}}{i(p^0-E_p+i\epsilon)}\Bigg|_{T_+}^{+\infty} = -\frac{e^{iT_+(p^0-E_p+i\epsilon)}}{i(p^0-E_p+i\epsilon)} , \quad (10.10)$$

so our region I integral is now

$$\sum_\lambda \frac{-ie^{i(p^0-E_p+i\epsilon)T_+}}{-2E_{\vec{p}}(\lambda)(p^0-E_{\vec{p}}(\lambda)+i\epsilon)}\langle\Omega|\phi(0)|\lambda_{\vec{p}}\rangle\langle\lambda_{\vec{p}}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle. \quad (10.11)$$

The denominator equals $p^2+m^2-i\epsilon$ and, near $p^0\to E_{\vec{p}}$, the exponential in the numerator becomes 1. Specializing first to the case $n=1$, when the 2-point function is supposed to be of the general form (10.3), we indeed find near on-shell for the unique momentum p,

$$\sim \sum_\lambda \frac{-i}{p^2+m^2-i\epsilon}\langle\Omega|\phi(0)|\lambda_{\vec{p}}\rangle\langle\lambda_{\vec{p}}|\phi(0)\Omega\rangle \sim \frac{-i}{p^2+m^2-i\epsilon}|\langle\Omega|\phi(0)|\vec{p}\rangle|^2 , \quad (10.12)$$

where we have implicitly assumed that there is a single one-momentum state $|\vec{p}\rangle$, and now we can identify $\langle\Omega|\phi(0)|\vec{p}\rangle$ with the factor $\sqrt{Z}$ (whose square is the wavefunction renormalization factor Z). Substituting in the general correlation function, we find that near $p^0\to E_{\vec{p}}$,

$$\int d^4x e^{-ip\cdot x}\langle\Omega|T\{\phi(x)\phi(z_1)...\phi(z_n)\}|\Omega\rangle \sim \frac{-i}{p^2+m^2-i\epsilon}\sqrt{Z}\langle\vec{p}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle. \quad (10.13)$$

We now repeat the procedure in region III, assuming that now x^0 is the smallest of all the times, i.e. that all the zero components of $z_1,...,z_n$ are larger than $T_- > x^0$, we can put x to the right inside the time ordering operator, and insert the identity to its left, obtaining

$$\int d^3x \int_{T_+}^{+\infty} dx^0 e^{+ip^0x^0 - i\vec{p}\cdot\vec{x}} \sum_\lambda \int \frac{d^3q}{(2\pi)^3} \frac{1}{2E_{\vec{q}}(\lambda)} \langle\Omega|T\{\phi(z_1)...\phi(z_n)\}|\lambda_{\vec{q}}\rangle\langle\lambda_{\vec{q}}|\phi(x)|\Omega\rangle. \quad (10.14)$$

And the matrix element is

$$\langle\lambda_{\vec{q}}|\phi(x)|\Omega\rangle = \langle\lambda_{\vec{q}}|e^{-i\hat{P}\cdot x}\phi(0)e^{+i\hat{P}\cdot x}|\Omega\rangle = \langle\lambda_{\vec{q}}|\phi(0)|\Omega\rangle\; e^{-iq\cdot x}\Big|_{q^0=E_{\vec{q}}}. \tag{10.15}$$

The integral over $\vec{x}$ is therefore $\int d^3x e^{-i\vec{x}\cdot(\vec{p}+\vec{q})} = (2\pi)^3\delta(\vec{p}+\vec{q})$, so after doing the integral over $\vec{q}$, we substitute everywhere $\vec{q}$ with $-\vec{p}$, and E_q with E_p, and get

$$\sum_\lambda \int_{-\infty}^{T_-} dx^0 \frac{1}{2E_{\vec{p}}(\lambda)} e^{+ip^0x^0}\; e^{iq^0x^0}\Big|_{q^0=E_{\vec{q}}=E_{\vec{p}}} e^{\epsilon x^0}\langle\lambda_{-\vec{p}}|\phi(0)|\Omega\rangle\langle\Omega|T\{\phi(z_1)...\phi(z_n)\}|\lambda_{-\vec{p}}\rangle. \tag{10.16}$$

The integral over x^0 gives

$$\int_{-\infty}^{T_-} e^{ix^0(p^0+E_p+i\epsilon)} = \frac{e^{ix^0(p^0+E_p-i\epsilon)}}{i(p^0+E_p-i\epsilon)}\Bigg|_{-\infty}^{T_-} = -i\frac{e^{iT_-(p^0+E_p-i\epsilon)}}{(p^0+E_p-i\epsilon)}\;, \tag{10.17}$$

so again we obtain the propagator, but now for $p^0 \to -E_p$, and so finally near $p^0 \to -E_p$,

$$\int d^4x e^{-ip\cdot x}\langle\Omega|T\{\phi(x)\phi(z_1)...\phi(z_n)\}|\Omega\rangle \sim \frac{-i}{p^2+m^2-i\epsilon}\sqrt{Z}\langle\Omega|T\{\phi(z_1)...\phi(z_n)\}|-\vec{p}\rangle. \tag{10.18}$$

We can redefine $p = -k$, so that near $k^0 = E_k$ (on-shell),

$$\int d^4x e^{ik\cdot x}\langle\Omega|T\{\phi(x)\phi(z_1)...\phi(z_n)\}|\Omega\rangle \sim \frac{-i}{k^2+m^2-i\epsilon}\sqrt{Z}\langle\Omega|T\{\phi(z_1)...\phi(z_n)\}|\vec{k}\rangle. \tag{10.19}$$

We see that in both regions, we obtain that on-shell, we relate to a correlation function with $n \to n-1$, with a vacuum state replaced by a momentum state. If we have $e^{-ip\cdot x}$, we relate to an outgoing state, and if we have $e^{ik\cdot x}$, we relate to an incoming state. We can iteratively repeat the procedure, and obtain finally that for all momenta on-shell, we relate to a product of incoming and outgoing states. The detail is that, if we have more than one momentum, the incoming states are actualy "in states", and the outgoing states are actually "out states". This is the result of the fact that we can independently consider each momentum on-shell.

But it remains to prove that when adding wavepackets instead of single momenta, nothing new happens. A wavepacket means that we replace

$$\int d^4x e^{ip^0x^0}e^{-i\vec{p}\cdot\vec{x}} \to \int\frac{d^3k}{(2\pi)^3}\int d^4x e^{ip^0x^0}e^{-i\vec{k}\cdot\vec{x}}\phi(\vec{k}). \tag{10.20}$$

Therefore the limit $\phi(\vec{k}) \to (2\pi)^3\delta^3(\vec{k}-\vec{p})$ takes us back to the original case. We can define $\tilde{p} = (p^0, \vec{k})$.

With this replacement, we obtain

$$\sum_\lambda \int \frac{d^3k}{(2\pi)^3}\phi(\vec{k})\frac{-ie^{i(p^0-E_k+i\epsilon)T_+}}{-2E_{\vec{k}}(\lambda)(p^0-E_{\vec{k}}(\lambda)+i\epsilon)}\langle\Omega|\phi(0)|\lambda_{\vec{k}}\rangle\langle\lambda_{\vec{k}}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle\;, \tag{10.21}$$

and near on-shell, $p^0 \to E_{\vec{k}}$, so $\tilde{p}^2 + m^2 \sim 0$,

$$\sim \int \frac{d^3k}{(2\pi)^3}\phi(\vec{k})\frac{-i}{\tilde{p}^2+m^2-i\epsilon}\sqrt{Z}\langle\vec{k}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle. \tag{10.22}$$

To see how this replacement works for pairs of particles, consider the scattering $n \to 2$. Then we obtain

$$\sum_\lambda \int \frac{d^3q}{(2\pi)^3}\frac{1}{2E_q}\prod_{i=1,2}\int\frac{d^3k_i}{(2\pi)^3}\int d^4x_i e^{-i\tilde{p}_i\cdot x_i}\phi_i(\vec{k}_i)\langle\Omega|T\{\phi(x_1)\phi(x_2)\}|\lambda_{\vec{q}}\rangle\langle\lambda_{\vec{q}}|T\{\phi(z_1)...\phi(z_n)\}|\Omega\rangle. \tag{10.23}$$

If the 2 outgoing particles are separated in the far future, we obtain

$$\begin{aligned}&\sum_\lambda \int \frac{d^3q}{(2\pi)^3}\frac{1}{2E_{\vec{q}}}\langle\Omega|T\{\phi(x_1)\phi(x_2)\}|\lambda_{\vec{k}}\rangle\langle\lambda_{\vec{k}}|\\ =&\sum_{\lambda_1\lambda_2}\int\frac{d^3q_1}{(2\pi)^3}\frac{1}{2E_{q_1}}\int\frac{d^3q_2}{(2\pi)^3}\frac{1}{2E_{q_2}}\langle\Omega|\phi(x_1)|\lambda_{q_1}\rangle\langle\Omega|\phi(x_2)|\lambda_{q_2}\rangle\langle\lambda_{q_1}\lambda_{q_2}|\ ,\end{aligned} \tag{10.24}$$

and then we can perform the same steps independently for each particle. The same analysis can be done in the far past as well, and we can generalize to more than 2 particles. In the limit in which the wavepackets tend to delta functions of momenta, we get the in and out states.

All in all, we obtain the LSZ formula,

$$\begin{aligned}&\prod_{i=1}^n\int d^4x_i e^{-ip_i\cdot x_i}\prod_{j=1}^m d^4y_j e^{+ik_j\cdot y_j}\langle\Omega|T\{\phi(x_1)...\phi(x_n)\phi(y_1)...\phi(y_m)\}|\Omega\rangle\\ \overset{p_i^0\to E_{p_i},k_j^0\to E_{k_j}}{\sim}&\left(\prod_{i=1}^n\frac{-i\sqrt{Z}}{p_i^2+m^2-i\epsilon}\right)\left(\prod_{j=1}^m\frac{-i\sqrt{Z}}{k_j^2+m^2-i\epsilon}\right)\langle\vec{p_1}...\vec{p_n}|S|\vec{k}_1...\vec{k}_m\rangle.\end{aligned} \tag{10.25}$$

Diagrammatic interpretation

The diagrammatic interpretation of the formula is as follows. In order to construct diagrams for the S-matrix from diagrams for the correlation fucntions, i.e. from connected diagrams, we need to perform the operation called *amputation*. We go on each external leg from the outside in until we reach the last part where it is connected with the rest of the diagram by a single leg, and cut there and excise that part, as in Fig.30. The reason is that we must divide out the connected correlation function by the full propagators for the external legs. Not quite the full propagators, of course, since we have a factor of $\sqrt{Z}$ instead of a factor of Z, but that is simply since a propagator has 2 legs instead of one, and we need to consider a factor of $\sqrt{Z}$ for each one. So for example for the 2-point function, we would have a $\sqrt{Z}$ for each leg, but only a single free propagator factor (common for both).

This amputation procedure relates diagrams in the perturbative expansion of the correlation functions to diagrams in the perturbative expansion of the S-matrices.

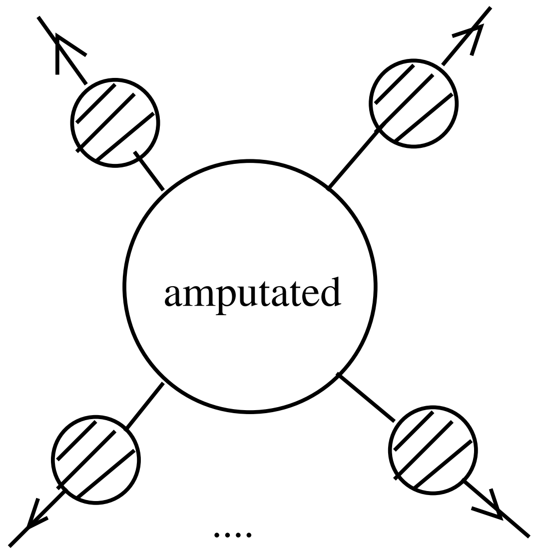

Figure 30: The diagrammatic amputation procedure.

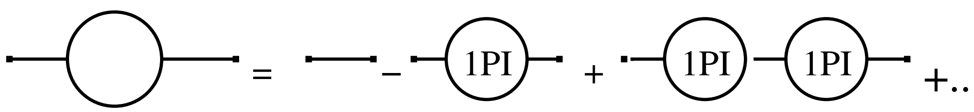

Figure 31: Diagrammatic expansion for the connected Green's function.

Finally, we want to understand the full propagator better, and what it means to be near the on-shell pole. We saw in QFTI that the connected 2-point function G^c_{ij} is related to the 1PI 2-point function Π_{ij} by

$$G^c = (1 + \Delta\Pi)^{-1}\Delta = \Delta - \Delta\Pi\Delta + ... \tag{10.26}$$

for which we can write a diagrammatic representation, as in Fig.31. But denoting $\Pi(p^2) \equiv iM^2(p^2)$, and that the free propagator has bare mass m_0, we have explicitly

$$G^c = \frac{-i}{p^2 + m_0^2} - \frac{-i}{p^2 + m_0^2} iM^2 \frac{-i}{p^2 + m_0^2} = \frac{-i}{p^2 + m_0^2 + M^2(p^2)} \ , \tag{10.27}$$

and we see that near a physical pole, we have

$$G^c \sim \frac{-iZ}{p^2 + m^2} + \text{regular}. \tag{10.28}$$

Indeed, as an example, taking $M^2(p^2) \simeq M^2 + \alpha p^2$, we would get

$$G^c = \frac{-i(1+\alpha)^{-1}}{p^2 + (m_0^2 + M^2)/(1+\alpha)} \ , \tag{10.29}$$

but if there are higher order corrections in p^2 in $M^2(p^2)$, we would get other finite terms near the physical pole. We see that the addition of $M^2(p^2)$ in general both shifts the position of the physical pole, here from $-m_0^2$ to $-m^2 = -(m_0^2 + M^2)/(1+\alpha)$, and creates a Z factor, here $(1+\alpha)^{-1}$.

Important concepts to remember

- The LSZ formula relates correlation functions to S-matrices as follows: near the on-shell physical pole for all the external legs, removing the full propagators for the external legs (with only $\sqrt{Z}$ instead of Z), we obtain the on-shell S-matrix.
- In and out states correspond to a different sign in the Fourier transform of the correlator positions.
- The diagrammmatic intepretation of going from correlators (connected diagrams) to S-matrices is of amputation, namely cut out all parts connecting external legs with the interior through a single propagator.
- For $\Pi = iM^2(p^2)$, $M^2(p^2)$ is added to the inverse propagator, shifting the physical pole and creating a Z factor.

Further reading: See chapter 7.2 in [3].

Exercises, Lecture 10

1) Write down all the 3-loop divergent diagrams for the LSZ formula at 4-points in ϕ^4 theory, and the associated diagrammatic amputation procedure.

2) Write down the LSZ formula for QED, and apply the diagrammatic procedure for the 2-loop 6-point case.

11 Lecture 11. Quantization of gauge theories I: path integrals and Fadeev-Popov

We now start the analysis of nonabelian gauge theories. In Classical Field Theory we have seen how to define them, but we will review it here.

Consider a gauge field

$$A_\mu = A_\mu^a T^a \ , \tag{11.1}$$

where T^a are the generators of a Lie algebra of a gauge group G, so A_μ is in the Lie algebra, i.e. in the adjoint representation. A general group element is a set of $N_R \times N_R$ matrices for the representation R ,

$$U = e^{\alpha^a T^a} \ , \tag{11.2}$$

where $\alpha^a \in \mathbb{R}$. The generators T^a, $a = 1, ..., N_G$ obey the Lie algebra

$$[T^a, T^b] = f^{ab}{}_c T^c \ , \tag{11.3}$$

and are normalized by the relation

$$\mathrm{Tr}[T^a T^b] = T_R \delta^{ab}. \tag{11.4}$$

A note on conventions. My conventions (that I will use unless otherwise specified, or unless left free) are of anti-hermitian generators, $(T^a)^\dagger = -T^a$, with $f^{ab}{}_c$ real and $T_R = -1/2$ in the fundamental representation of $SU(N)$. Another popular convention in the literature is with hermitian generators, $(T^a)^\dagger = T^a$, $T_R = 1$ in the adjoint of $SU(N)$ and $[T^a, T^b] = i f^{ab}{}_c T^c$.

In a representation R, the following quadratic form is a constant (proportional to the identity over a given representation),

$$\sum_a (T_R^a)^2 = C_R \, \mathbb{1} \ , \tag{11.5}$$

and is called the (quadratic) Casimir of the representation R of the group G. Then

$$T_R N_G = C_R N_R. \tag{11.6}$$

In the case of the adjoint representation, defined by

$$(T^a)_{bc} = f^a{}_{bc} \ , \tag{11.7}$$

we have

$$T_R = C_R \equiv C_2(G). \tag{11.8}$$

as well as $N_R = N_G$.

For $G = SU(N)$, we have of course $N_G = N^2 - 1$, and in the fundamental representation (for quarks) we have $N_R = N$.

My normalization $T_R = -1/2$ in the fundamental leads also to $C_R = (N^2 - 1)/2N$.

The way to introduce gauge fields is by starting with some matter action invariant under the global action of the group G,

$$\psi_i(x) \to U_{ij}\psi_j(x) \ , \tag{11.9}$$

and making the invariance local, $U_{ij} \to U_{ij}(x)$. That requires the introduction of another field, the gauge field A_μ^a, and of the minimal coupling of the matter to it, through the covariant derivative replacing the ordinary derivative,

$$\partial_\mu\psi_i(x) \to (D_\mu)_{ij}\psi_j(x); \quad (D_\mu)_{ij} \equiv \partial_\mu\delta_{ij} + g(T_R^a)_{ij}A_\mu^a(x). \tag{11.10}$$

In particular, in the adjoint representation,

$$D_\mu^{ab} = \partial_\mu\delta^{ab} + gf^{ab}{}_c A_\mu^c. \tag{11.11}$$

For an infinitesimal gauge transformation with parameter

$$U(x) = e^{g\alpha^a T^a} \simeq 1 + g\alpha^a(x)T^a + ... \ , \tag{11.12}$$

the gauge field transforms as

$$A_\mu^a(x) \to A_\mu^a(x) + D_\mu^{ab}\alpha^b(x). \tag{11.13}$$

For a finite gauge transformation, $U = e^{g\alpha^a(x)T^a} = e^{g\alpha(x)}$, the gauge field transforms as

$$A_\mu(x) \to A_\mu^U(x) = U^{-1}A_\mu(x)U(x) + \frac{1}{g}\partial_\mu U(x)U^{-1}(x). \tag{11.14}$$

The field strength is defined as

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^a{}_{bc}A_\mu^b A_\nu^c. \tag{11.15}$$

One defines also the contraction with T^a,

$$A_\mu = A_\mu^a T^a; \quad F_{\mu\nu} = F_{\mu\nu}^a T^a \ , \tag{11.16}$$

such that the field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + g[A_\mu, A_\nu] \ , \tag{11.17}$$

as well as the form notation

$$A = A_\mu dx^\mu; \quad F = \frac{1}{2}F_{\mu\nu}dx^\mu dx^\nu \ , \tag{11.18}$$

(note that in general $f_p = 1/p! f_{\mu_1...\mu_p}dx^{\mu_1} \wedge ... \wedge dx^{\mu_p}$.) leading to the field strength

$$F = dA + gA \wedge A. \tag{11.19}$$

The field strength transforms *covariantly*, i.e.

$$F_{\mu\nu} \to F'_{\mu\nu} = U^{-1}(x) F_{\mu\nu} U(x). \tag{11.20}$$

The gauge invariant action for the gauge field in Minkowski space is

$$S_M = \int d^4x \left[-\frac{1}{4} F^a_{\mu\nu} F^{a\ ,\mu\nu} \right] = +\frac{1}{2} \int d^4x \,\mathrm{Tr}[F^2_{\mu\nu}]. \tag{11.21}$$

Wick rotating to Euclidean space, $x_4 = it$, so $\partial_4 = -i\partial_t$, and the same for the gauge field, which transforms as ∂_μ under Lorentz transformations, $A_4 = -iA_0$, so

$$E_i^{\text{Eucl.}} = \frac{\partial}{\partial x_4} A_i - \frac{\partial}{\partial x_i} A_4 = -i E_i^{\text{Mink.}} \ , \tag{11.22}$$

and so the Euclidean Lagrangean is

$$\mathcal{L}^{\text{Eucl.}} = +\frac{1}{4} F^a_{\mu\nu} F^{a\ \mu\nu} = -\frac{1}{2} \mathrm{Tr}[F^2_{\mu\nu}] = \frac{1}{2}((\vec{E}^a)^2 + (\vec{B}^a)^2). \tag{11.23}$$

Correlation functions

We now write the quantum correlation functions for *gauge invariant observables* (observables in the gauge theory must be gauge invariant, by definition) in Euclidean space as path integrals,

$$\langle \mathcal{O}_1(A) ... \mathcal{O}_n(A) \rangle = \frac{\int \mathcal{D}A e^{-S[A]} \mathcal{O}_1(A) ... \mathcal{O}_n(A)}{\int \mathcal{D}A e^{-S[A]}}. \tag{11.24}$$

As in the abelian case described in QFT I, we "fix the gauge" by the **Fadeev-Popov procedure**, which amounts to dividing in the numerator and denominator by the volume of the gauge group G,

$$\prod_{x \in \mathbb{R}^d} V_x(G); \quad V(G) = \int_G dU \ , \tag{11.25}$$

where $\int dU$ is called the Haar measure. It is defined to be invariant under left and right multiplication by a fixed element of the gauge group,

$$U \to U U_0; \quad \text{and} \quad U \to U_0 U. \tag{11.26}$$

We are interested in covariant gauges like the Lorenz gauge, $\partial_\mu A^a_\mu = 0$, generalized to the form

$$\mathcal{F}^a(x) = c^a(x); \quad a = 1, ..., N. \tag{11.27}$$

We define the *orbit of the gauge field* $A_\mu(x)$ as the space of all possible gauge transformations of $A_\mu(x)$, i.e.

$$Or[A_\mu(x)] \equiv \{\tilde{A}_\mu | \exists U(x) \in G, \text{ such that } \tilde{A}_\mu(x) = {}^{U(x)} A_\mu(x)\}. \tag{11.28}$$

We also define the space of all possible gauge fields satisfying a gauge condition,

$$\mathcal{M}(A) = \{A_\mu \text{ such that } \mathcal{F}^a(A_\mu) = c^a(x)\}. \tag{11.29}$$

We assume the fact that there is a single intersection between the two, i.e. $Or[A_\mu(x)] \cap \mathcal{M} = \{\text{point}\}$, or that

$$\exists ! \tilde{A}_\mu \text{ such that } \mathcal{F}(\tilde{A}_\mu) = c^a(x). \tag{11.30}$$

Note that this assumption is only correct for infinitesimal gauge transformations, otherwise for large gauge transformations, there are *Gribov copies* that are a large distance in gauge transformation space from the identity. We will not address Gribov copies in this course. The situation above is depicted in Fig.32.

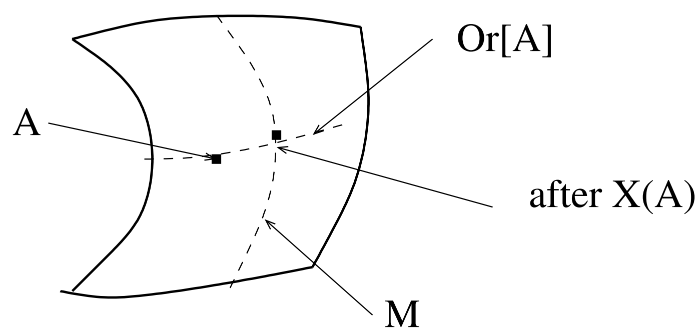

Figure 32: The gauge fixed configuration is at the intersection of the orbit of A, the gauge transformations of a gauge field configuration A, and the space $\mathcal{M}$ of all possible gauge conditions.

Then define

$$\frac{1}{\Delta_{\mathcal{F},c}[A]} \equiv \int \prod_x dU_x \prod_{y,a} \delta(\mathcal{F}^a({}^U A) - c^a). \tag{11.31}$$

By our assumption, there is a unique $U = U^{(A)}(x)$, depending on $A_\mu(x)$, such that by transforming with it we go onto the gauge condition, i.e.

$$\mathcal{F}^a({}^{U^{(A)}}A) = c^a. \tag{11.32}$$

Define the matrix

$$M^{ab}(x,y;A) \equiv \frac{\partial \mathcal{F}^a}{\partial A_\mu^c(x)} D_\mu^{cb}(x;A) \delta^{(D)}(x-y)\ , \tag{11.33}$$

where as before

$$D_\mu^{ab}(x,A) \equiv \frac{\partial}{\partial x^\mu}\delta^{ab} + g f^{ab}{}_c A_\mu^c(x). \tag{11.34}$$

The matrix M is thought of as a matrix in both the space (ab) and in the space (xy). We have the following

Lemma

1. $\Delta_{\mathcal{F},c}[A]$ is gauge invariant.
2. $\Delta_{\mathcal{F},c}[A] = \det M({}^{U^{(A)}}A)$.

Proof of 1.

Write the definition of $\Delta[A]$ for the gauge field transformed with some U_0, ${}^{U_0}A$,

$$\Delta^{-1}[{}^{U_0}A] = \int \prod_x dU_x \prod_{x,a} \delta(\mathcal{F}^a({}^U({}^{U_0}A)) - c^a). \tag{11.35}$$

By invariance of the Haar measure, $dU_x = d(UU_0)$, so we write $dU_x = d(U_xU_0) \equiv d\tilde{U}_x$, so we get

$$\Delta^{-1}[{}^{U_0}A] = \int \prod_x d\tilde{U}_x \prod_{x,a} \delta(\mathcal{F}^a({}^{\tilde{U}}A) - c^a) = \Delta^{-1}[A]. \tag{11.36}$$

q.e.d.1

Proof of 2.

We use 1. to write

$$\Delta(A) = \Delta({}^{U^{(A)}}A) \equiv \Delta(\tilde{A}) \ , \tag{11.37}$$

where by definition of $U^{(A)}$, $\mathcal{F}(\tilde{A}) = c^a$.

Then

$$\Delta^{-1}[A] = \Delta^{-1}[\tilde{A}] = \int \prod_x dU_x \delta(\mathcal{F}^a({}^U\tilde{A}) - c^a) \ , \tag{11.38}$$

For infinitesimal transformations,

$$U(x) = e^{\alpha^a(x)T^a} \simeq 1 + \alpha^a(x)T^\alpha \Rightarrow dU_x = \prod_{a=1}^N d\alpha^a(x) + \mathcal{O}(\alpha^2) \ , \tag{11.39}$$

which leads to

$$({}^U\tilde{A})^a \simeq \tilde{A}^a + D_\mu^{ab}(\tilde{A})\alpha^b + \mathcal{O}(\alpha^2) \ , \tag{11.40}$$

and in turn to

$$\mathcal{F}^a({}^U\tilde{A})(x) \simeq \mathcal{F}(\tilde{A})(x) + \frac{\partial\mathcal{F}(\tilde{A}(x))}{\partial\tilde{A}_\mu^c(x)} D_\mu^{cb}(x,\tilde{A})\alpha^b(x) + \mathcal{O}(\alpha^2). \tag{11.41}$$

Substituting in $\Delta^{-1}[\tilde{A}]$, we obtain (using that $\mathcal{F}^a(\tilde{A}_\mu) = c^a$)

$$\begin{aligned} \Delta^{-1}[A] &\simeq \int \prod_{x,a'} d\alpha_x^{a'} \prod_{y,a} \delta\left(\frac{\partial\mathcal{F}(\tilde{A}(y))}{\partial\tilde{A}_\mu^c} D_\mu^{cb}(y,\tilde{A})\alpha^b(y)\right) \\ &= \int \prod_{x,a'} d\alpha_x^{a'} \prod_{y,z} \delta\left(\int d^dz M^{ab}(y,z;\tilde{A})\alpha^b(z)\right) \equiv \left(\det M(\tilde{A})\right)^{-1}. \end{aligned} \tag{11.42}$$

In the last equality, the determinant was considered both in (ab) and in (xy) space, and we have used the generalization of the relation

$$\int \prod_{i=1}^n d\alpha_i \prod_{j=1}^n \delta(M_{ij}\alpha_j) = \frac{1}{\det M}. \tag{11.43}$$

q.e.d.2

Then we can take the determinant on the other side of the equation and write

$$1 = \int \prod_x dU_x \left[\det M({}^{U^{(A)}}A)\right] \prod_{y,a} \delta(\mathcal{F}^a({}^U A) - c^a(y)). \tag{11.44}$$

But the delta function enforces $U = U^{(A)}$ anyway, so we can replace it in the integral and write

$$1 = \int \prod_x dU_x \det M({}^U A) \prod_{y,a} \delta(\mathcal{F}^a({}^U A) - c^a(y)). \tag{11.45}$$

Now we can define

$$"1(\alpha)" \equiv \int \prod_{x,a} dc^a(x) e^{-\frac{1}{2\alpha}\int d^D x [c^a(x)]^2} , \tag{11.46}$$

which is of course a combination of π's and α's, that is irrelevant in the correlators, since would cancel in the numerator and denominator. Now substituting 1 in the form of (11.45) on the rhs of the above, and doing the integral over $dc^a(x)$, which fixes $c^a(x) = \mathcal{F}^a(x)$, we obtain

$$"1(\alpha)" = \int \prod_x dU_x \det M({}^U A) e^{-\frac{1}{2\alpha}\mathcal{F}^a \mathcal{F}^a}. \tag{11.47}$$

Inserting this "$1(\alpha)$" in the path integral, we obtain

$$\begin{aligned} \int \mathcal{D}A e^{-S[A]} &= \int \mathcal{D}[A] \int \prod_x dU_x \det M({}^U A) e^{-S[A] - \frac{1}{2\alpha}\mathcal{F}^2[{}^U A]} \\ &= \int \prod_x dU_x \int \mathcal{D}A \det M({}^U A) e^{-S[A] - \frac{1}{2\alpha}\mathcal{F}^2[{}^U A]}. \end{aligned} \tag{11.48}$$

But because of the gauge invariance of the measure $\mathcal{D}A$ and the action $S[A]$, we can replace A by ${}^U A$ in the path integral, and finally rename it A again everywhere. Note that it is exactly this step that can fail if there are gauge anomalies, but here we will assume that there aren't.

Then we finally obtain that the volume of the gauge group factorizes from the integral as

$$\int \mathcal{D}A e^{-S[A]} = \left[\int \prod_x dU_x\right] \int \mathcal{D}A \det M(A) e^{-S[A] - \frac{1}{2\alpha}\mathcal{F}^2[A]}. \tag{11.49}$$

The same thing happens for the path integral with the *gauge invariant* observables,

$$\int \mathcal{D}A e^{-S[A]} \mathcal{O}_1[A] ... \mathcal{O}_n[A] = \left[\int \prod_x dU_x\right] \int \mathcal{D}A \det M(A) e^{-S[A] - \frac{1}{2\alpha}\mathcal{F}^2[A]} \mathcal{O}_1[A] ... \mathcal{O}_n[A]. \tag{11.50}$$

That means that the volume of the gauge group cancels between the numerator and denominator in the correlation functions, and we obtain

$$\langle \prod_{i=1}^n \mathcal{O}_i(A) \rangle = \frac{\int \mathcal{D}A \prod_{i=1}^n \mathcal{O}_i(A) e^{-S_{\text{eff}}[A]}}{\int \mathcal{D}A e^{-S_{\text{eff}}}} , \tag{11.51}$$

where the $S_{\rm eff}$ is the sum of the classical action and a "gauge fixing term" and an extra term,

$$S_{\rm eff}[A]=S[A]+\frac{1}{2\alpha}\int d^Dx\mathcal{F}^2[A]-\log\det M(A). \tag{11.52}$$

The last term will turn into a "ghost action" term, as we will now see.

Ghost action

In the case of the Lorenz gauge, $\mathcal{F}^a=\partial_\mu A^a_\mu$, the matrix M becomes

$$M^{ab}(x,y;A)=\partial_\mu D^{ab}_\mu\delta^D(x-y)=(\partial^2_\mu\delta^{ab}+gf^{ab}{}_c\partial_\mu A^c_\mu(x))\delta^D(x-y)\ , \tag{11.53}$$

understood as a matrix in both (ab) and (xy) space. Since we are interested in

$$\log\det M(A)=\log\frac{\det M(A)}{\det\partial^2}+\log\det\partial^2\ , \tag{11.54}$$

we can drop the last term, which is just a constant (albeit infinite one), since it again cancels in correlators between the numerator and the denominator.

Note that $\partial^{-2}=\Delta(x,y)$ is the scalar (KG) propagator.

Therefore we consider

$$\frac{\det M(A)}{\det\partial^2}=\det\left[\frac{M(A)}{\partial^2}\right]\equiv\det(1+L)\ , \tag{11.55}$$

where

$$(1+L)^{ab}(x,y)=\delta^{ab}\delta^D(x-y)+g\int d^Dz\Delta(x-z)\frac{\partial}{\partial z^\mu}f^{ab}{}_cA^c_\mu\delta^D(z-y). \tag{11.56}$$

On the other hand, that means that

$$\begin{aligned}\det(1+L)&=e^{{\rm Tr}\log(1+L)}=\exp{\rm Tr}\left(\sum_{n\geq1}\frac{(-1)^{n+1}}{n}L^n\right)\\&=\exp\left\{g\int d^Dz_1f^a{}_{ac}\left[\Delta(x-z_1)\frac{\partial}{\partial z_1^\mu}A^c_\mu(z_1)\delta^D(z_1-x)\right]\right.\\&-\frac{g^2}{2}\int d^Dz_1d^Dz_2f^{ab}{}_c\left[\Delta(x-z_1)\frac{\partial}{\partial z_1^\mu}A^c_\mu(z_1)\delta^D(z_1-z_2)\right]f_{bad}\times\\&\times\left.\left[\Delta(z_1-z_2)\frac{\partial}{\partial z_2^\nu}A^d_\nu(z_2)\delta^D(z_2-x)\right]\right\}+...\end{aligned} \tag{11.57}$$

However, this gives an infinite number of vertices in the action (term $-\log\det(1+L)$), as we easily see. This is not very good.

Instead, we can use a representation in the action in terms of *fermions*, or rather anti-commuting variables, that will be called ghosts.

We remember that we have the formula

$$\int\mathcal{D}\Phi\mathcal{D}\bar\Phi e^{-\bar\Phi\cdot M\cdot\Phi}\propto(\det M)^{\pm1}\ , \tag{11.58}$$

where as usual the proportionality constant is not relevant, and ± 1 is for anticommuting/commuting variables, respectively.

Since we have $\det M^{+1}$, we are interested in anticommuting variables, and moreover then we have an arbitrary sign in front of the action that we can use, since

$$\int \prod_{i=1}^{N} d\eta_i d\bar{\eta}_i e^{\pm \bar{\eta}\cdot M \cdot \eta} = (\pm 1)^N \det M. \tag{11.59}$$

We choose the plus sign in the above, i.e. $e^{+\bar{\eta}\cdot M\cdot\eta}$, and as usual we choose the Lorenz gauge $\mathcal{F}^a = \partial_\mu A_\mu^a$, obtaining the gauge fixed YM action in Euclidean space

$$S_{\text{eff.}} = \int d^D x \left[\frac{1}{4}(F_{\mu\nu}^a)^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu^a)^2 - \bar{\eta}^a \partial_\mu D_\mu^{ab}\eta^b\right]. \tag{11.60}$$

Note that I wrote here $\bar{\eta}^a$ and η^b, but this is misleading, since as we know from the complex integration of anticommuting objects, the integrations are really independent. That is why one usually writes b^a for $\bar{\eta}^a$ and c^a for η^a, i.e.

$$S_{\text{eff.}} = \int d^D x \left[\frac{1}{4}(F_{\mu\nu}^a)^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu^a)^2 - b^a \partial_\mu D_\mu^{ab} c^b\right]. \tag{11.61}$$

For a general gauge condition, we have

$$S_{\text{eff.}} = \int d^D x \left[\frac{1}{4}(F_{\mu\nu}^a)^2 + \frac{1}{2\alpha}(\mathcal{F}^a(A))^2 - b^a \frac{\partial \mathcal{F}^a}{\partial A_\mu^c} D_\mu^{cb} c^b\right]. \tag{11.62}$$

Other gauge conditions, generalizing the Lorenz gauge, can be written as

$$\mathcal{F}^a = \phi_\mu^{ab} A_\mu^b. \tag{11.63}$$

In particular, axial gauges (which are however not Lorentz covariant) correspond to

$$\phi_\mu^{ab} = \eta_\mu \delta^{ab}. \tag{11.64}$$

Here η_μ is a constant 4-vector.

Important concepts to remember

- The YM field strength transforms covariantly, $F_{\mu\nu} \to U^{-1}(x) F_{\mu\nu} U(x)$.
- The YM action in Euclidean space is $+\frac{1}{4}\int d^D x F_{\mu\nu}^a F^{a\,\mu\nu}$.
- The quantum correlators of gauge invariant operators are written as ratios of path integrals with and without the operators, so we can use the Fadeev-Popov gauge fixing procedure, by factorizing and cancelling the volume of the gauge group from the two path integrals.
- The result of the gauge-fixing procedure for the gauge condition $\mathcal{F}^a(x) = c^a(x)$ is $S_{\text{eff.}}[A] = S[A] + \int d^D x \mathcal{F}^2(A)/2\alpha - \log\det M(A)$.
- The $\log\det M(A)$ can be written as a ghost action, $\int d^D x[-b^a \partial\mathcal{F}^a(A)/\partial A_\mu^c D_\mu^{cb}(A) c^b]$.

Further reading: See chapter 7.1 in [5] and 16.2 in [3].

Exercises, Lecture 11

1) Consider a solution to the self-duality equation for Yang-Mills theory in Euclidean space,

$$F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu}{}^{\rho\sigma}F_{\rho\sigma}. \tag{11.65}$$

Show that the on-shell action is bounded by a topological term (which cannot be changed by a small transformation).

2) Consider a solution to the self-duality condition that asymptotes to flat space at $x_4 = -\infty$ and to a monopole configuration at $x_4 = +\infty$.

Wick rotate the configuration to Minkowski space, and consider the path integral centered around this configuration. What is its interpretation? Putting some reasonable numbers for the Standard Model, how relevant is this now?

12 Lecture 12. Quantization of gauge theories 2. Propagators and Feynman rules.

As we have seen last lecture, to calculate correlators we must fix a gauge, and then we use instead of the classical action the gauge fixed action, with gauge fixing term and ghost term,

$$S_{\text{eff.}} = \int d^D x \left[\frac{1}{4}(F_{\mu\nu}^a)^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu^a)^2 - b^a \partial_\mu D_\mu^{ab} c^b\right]. \tag{12.1}$$

In order to calculate the Feynman rules, we split the action into a quadratic part, giving the propagators, and a cubic and quartic part, giving the interactions.

Propagators

The quadratic action is

$$S^{[2]}[A,b,c] = \int d^D x \left\{\frac{1}{2}A_\mu^a \left[\delta^{ab}\left(-\partial^2 \delta_{\mu\nu} + \partial_\mu \partial_\nu \left(1 - \frac{1}{\alpha}\right)\right)\right] A_\nu^b + b^a(-\delta^{ab}\partial^2)c^b\right\}. \tag{12.2}$$

From it, we derive the gluon propagator, which is just δ^{ab} times the abelian (photon) propagator, i.e.

$$\Delta_{\mu\nu}^{ab}(k) = \frac{\delta^{ab}}{k^2}\left[\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right], \tag{12.3}$$

and the the *ghost propagator*, which is (despite the anticommuting nature of the ghosts) just the scalar KG propagator,

$$\Delta^{ab}(k) = \frac{\delta^{ab}}{k^2}. \tag{12.4}$$

Interactions

The interaction action can be rewritten as

$$\begin{aligned} S_{\text{int}}[A,b,c] &= g f_{abc} \int d^D x \left[(\partial_\mu A_\nu^a) A_\mu^b A_\nu^c + (\partial_\mu b^a) A_\mu^b c^c\right] \\ &\quad + \frac{g^2}{4} f^a{}_{bc} f_{ade} \int d^D x A_\mu^b A_\nu^c A_\mu^d A_\nu^e. \end{aligned} \tag{12.5}$$

We can now define path integrals, in particular the free energy W is defined as usual by

$$e^{-W[J,\xi_b,\xi_c]} = \int \mathcal{D}A\mathcal{D}b\mathcal{D}c e^{-S_{\text{eff}}[A,b,c] + \int d^D x (J\cdot A + \xi_c c + b\xi_b)}. \tag{12.6}$$

Then as usual, the effective action, the generator of 1PI n-point functions, is the Legendre transform of the free energy,

$$\Gamma[A^{cl}, b^{cl}, c^{cl}] = W[J,\xi_b,\xi_c] + \int d^D x (J\cdot A^{cl} + \xi_c c^{cl} + b^{cl}\xi_b). \tag{12.7}$$

By taking derivatives of this Legendre transform relation, we obtain also

$$A_\mu^{cl}(x) = -\frac{\delta W}{\delta J(x)}; \quad c^{cl}(x) = -\frac{\delta W}{\delta \xi_c(x)}; \quad b^{cl}(x) = +\frac{\delta W}{\delta \xi_b(x)}$$

$$J_\mu(x) = +\frac{\delta\Gamma}{\delta A_\mu^{cl}(x)};\quad \xi_c(x) = -\frac{\delta\Gamma}{\delta c^{cl}(x)};\quad \xi_b(x) = +\frac{\delta\Gamma}{\delta b^{cl}(x)}. \tag{12.8}$$

Vertices

We can derive the vertices from the interaction action. The 3-gluon vertex comes from the rewriting

$$gf_{abc}\int d^Dx(\partial_\mu A_\nu^a)A_\mu^b A_\nu^c = \int \frac{d^Dkd^Dpd^Dq}{(2\pi)^6}\frac{1}{3!}A_\mu^a(k)A_\nu^b(p)A_\lambda^c(q)\tilde\Gamma^{abc}_{\mu\nu\lambda}. \tag{12.9}$$

Then we can derive the vertex, which should have the overall momentum conservation, so

$$\tilde\Gamma^{abc}_{\mu\nu\lambda}(k,p,q) = (2\pi)^D\delta^{(D)}(k+p+q)V^{abc}_{\mu\nu\lambda}. \tag{12.10}$$

The vertex should be symmetric in the external lines. We first rewrite the interaction term as

$$(gf_{abc})\int d^Dx[-A_\mu^a(\partial_\mu A_\nu^b)A_\lambda^c\delta^{\nu\lambda}] = (gf_{abc})\int d^Dx[A_\mu^a(\partial_\lambda A_\nu^b)A_\lambda^c]\delta^{\mu\nu}\ , \tag{12.11}$$

where we have redefined the indices and used the fact that $f_{bac} = -f_{abc}$ and $f_{cab} = +f_{abc}$. Then we see that we need the terms $-(ip_\mu)\delta_{\nu\lambda} + (ip_\lambda)\delta_{\mu\nu}$ among the permutations (the derivative ∂ is replaced by ip), multiplied by the usual Euclidean vertex, $-gf_{abc}$. The other terms are obtained by permuting p, k, q and the external indices. In total, we have

$$V^{abc}_{\mu\nu\lambda}(k,p,q) = (-igf_{abc})[(q-p)_\mu\delta_{\nu\lambda} + (p-k)_\lambda\delta_{\mu\nu} + (k-q)_\nu\delta_{\mu\lambda}]. \tag{12.12}$$

For the 4-gluon vertex, again taking out the delta function, the vertex $V^{abcd}_{\mu\nu\rho\sigma}$, for gluons (μa), (νb), (ρc) and (σd), is momentum-independent. There are 24 terms coming from the 4! terms in the permutations of the external lines. They will give 6 different terms and a multiplicity of 4, cancelling the 1/4 in front of the quartic interaction action. One such term is given by writing the quartic action as

$$+\frac{g^2}{4}f_{abe}f_{cd}{}^e\int d^Dx A_\mu^a A_\nu^b A_\rho^c A_\sigma^d\delta^{\mu\rho}\delta^{\nu\sigma}\ , \tag{12.13}$$

where we have relabelled the indices and used that $f^e{}_{ab}f_{ecd} = f_{abe}f^e_{cd}$. The vertex term from it is then $-g^2 f_{abe}f_{ad}{}^e\delta^{\mu\rho}\delta^{\nu\sigma}$, and the other 5 terms are found by permutations, giving in total

$$V^{abcd}_{\mu\nu\rho\sigma} = -\tilde g^2[f_{abe}f_{cd}{}^e(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\nu\rho}\delta_{\mu\sigma}) + f_{cbe}f_{ad}{}^e(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\nu}\delta_{\rho\sigma}) + f_{dbe}f_{ca}{}^e(\delta_{\rho\sigma}\delta_{\mu\nu} - \delta_{\nu\rho}\delta_{\mu\sigma})]. \tag{12.14}$$

The gluon-2-ghost vertex comes from the cubic part of the interacition action. For a ghost line from a to b, where b has momentum q, and with a gluon with (μc), we have

$$V^{abc}_\mu(q) = -\tilde g f_{abc}(iq_\mu). \tag{12.15}$$

If we also introduce fermions in a representation f with index i, ψ^i_α, so with covariant derivative

$$D_{\mu\, ij} = \partial_\mu\delta_{ij} + g(T^a_f)_{ij}A^a_\mu\ , \tag{12.16}$$

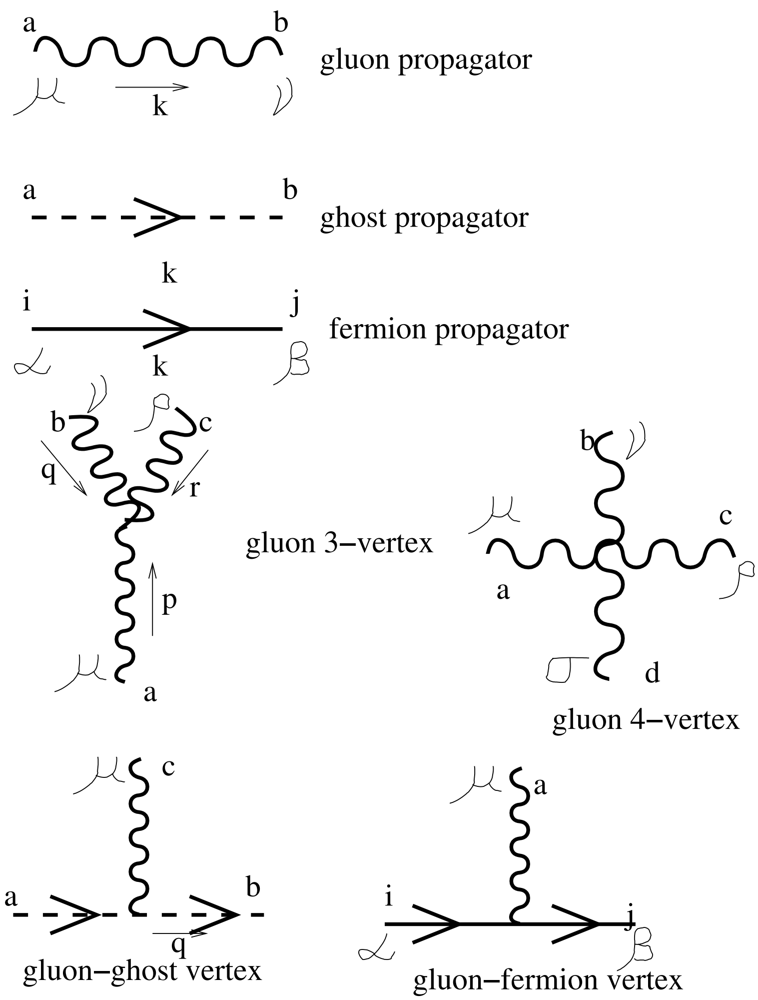

Figure 33: Feynman rules for nonabelian gauge theories (QCD).

acting on them, and $\bar{\psi}^i D_\mu^{ij} \gamma^\mu \psi^j$ kinetic term, it follows that the gluon-2-fermion vertex with the fermion going from αi to βj and the gluon with μa, is

$$-\tilde{g}(T_f^a)_{ji}(\gamma_\mu)_{\beta\alpha}. \tag{12.17}$$

Note that in all the above, as usual, $\tilde{g} = g\mu^{\epsilon/2}$, where $\epsilon = 4 - D$.

Feynman rules

All in all, we have the Feynman rules (see Fig.33):

1. Gluon propagator, represented by a wiggly line from μa to νb, with momentum k, is

$$\Delta_{\mu\nu}^{ab}(k) = \frac{\delta^{ab}}{k^2}\left[\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right]. \tag{12.18}$$

2. Ghost propagator, represented by a dashed line with an arrow from a to b, with

momentum k, is

$$\Delta^{ab}(k) = \frac{\delta^{ab}}{k^2}. \tag{12.19}$$

3. Fermion propagator, from $(i\alpha)$ to $(j\beta)$, with momentim k, is

$$S^{ij}_{F\alpha\beta}(k) = \delta_{ij}\left(\frac{1}{i\not{k}+m}\right)_{\alpha\beta}. \tag{12.20}$$

4. Gluon 3-vertex, represented by 3 wiggly lines intersecting at a point, with all momenta in, (μa) with momentum p, (νb) with momentum q and (ρc) with momentum r, is

$$V^{abc}_{\mu\nu\rho}(p,q,r) = -i\tilde{g}f_{abc}[(r-q)_\mu\delta_{\nu\rho} + (q-p)_\rho\delta_{\mu\nu} + (p-r)_\nu\delta_{\mu\rho}]. \tag{12.21}$$

5. Gluon 4-vertex, represented by 4 wiggly lines intersecting at a point, with (μa), (νb), (ρc) and (σd), is

$$V^{abcd}_{\mu\nu\rho\sigma} = -\tilde{g}^2[f_{abe}f_{cd}{}^e(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\nu\rho}\delta_{\mu\sigma}) + f_{cbe}f_{ad}{}^e(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\nu}\delta_{\rho\sigma}) + f_{dbe}f_{ca}{}^e(\delta_{\rho\sigma}\delta_{\mu\nu} - \delta_{\nu\rho}\delta_{\mu\sigma})]. \tag{12.22}$$

6. Gluon-2-ghost vertex, represented by a dashed line with an arrow from a to b (where b has momentum q), with a wiggly line coming out of it, ending on (μc), is

$$V^{abc}_\mu(q) = -\tilde{g}f_{abc}(iq_\mu). \tag{12.23}$$

7. Gluon-2-fermion vertex, represented by a continuous line with an arrow from $(i\alpha)$ to $(j\beta)$, with a wiggly line coming out of it, ending on (μc), is

$$-\tilde{g}(T^a_f)_{ji}(\gamma_\mu)_{\beta\alpha}. \tag{12.24}$$

8. The fermion loop has an extra (-1), but also the ghost loop, since the important thing is the anticommuting nature of the variables, not the kinetic term (which is KG for the ghost).

Observation. Note that we calculate Green's functions from derivatives of W or Γ, but while the action is gauge invariant under the nonabelian gauge transformation, the source term $\int d^D x J\cdot A$ is not, so the Green's functions are not physical observables, since they are not gauge invariant, as observables should be. But sums of Feynman diagrams could be gauge invariant. For example, in the case of QED, we mentioned the fact that IR divergences will mean that we need to sum loop diagrams with tree diagrams, of the same order in the coupling, but with more external massless lines, with very small momentum. Only this combination will be related to experimentally relevant quantities like the cross section, and will be gauge invariant (and IR safe). Now we can say a similar thing about YM theory, where Green's functions will be in general gauge-dependent, even at a fixed order in the coupling.

Example

As an example of calculation, we will consider the one-loop correction to the gluon propagator with 2 3-gluon vertices, as in Fig.34. It will be also gauge dependent, but we will

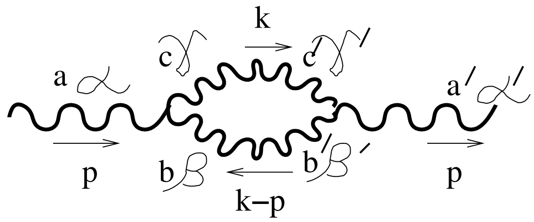

Figure 34: One loop gluon diagram with 3-vertices.

choose the Feynman gauge $\alpha = 1$. For other α, the result will change, and it will of course also change if we use instead of the Lorenz gauge some other gauge, like for instance axial gauge.

In the diagram, a photon with momentum p and indices $(a\alpha)$ comes, and out goes a photon of momentum p and indices $(a'\alpha')$. On the internal loop, one line has momentum k and indices $(c\gamma)$ on the $(a\alpha)$ side and $(c'\gamma')$ on the $(a'\alpha')$ side, and the other line has momentum $k-p$ and indices $(b\beta)$ on the $(a\alpha)$ side and $(b'\beta')$ on the $(a'\alpha')$ side. Then the two vertices are $V^{abc}_{\alpha\beta\gamma}(p, k-p, -k)$ and $V^{a'b'c'}_{\alpha'\beta'\gamma'}(-p, p-k, k)$, and using the Feynman rules above we have the expression for the amplitude

$$
\begin{aligned}
\mathcal{M}^{aa'}_{\alpha\alpha'} &= \int \frac{d^Dk}{(2\pi)^D}(-i\tilde{g}f_{abc})[(-2k+p)_\alpha\delta_{\beta\gamma} + (p+k)_\beta\delta_{\alpha\gamma} + (k-2p)_\gamma\delta_{\alpha\beta}]\frac{\delta^{cc'}}{k^2}\delta_{\gamma\gamma'}\times \\
&\quad \times(-i\tilde{g}f_{a'b'c'})[(2k-p)_{\alpha'}\delta_{\beta'\gamma'} + (-p-k)_{\beta'}\delta_{\alpha'\gamma'} + (-k+2p)_{\gamma'}\delta_{\alpha'\beta'}]\frac{\delta^{bb'}}{(k-p)^2}\delta_{\beta\beta'} \\
&= -\tilde{g}^2 f_{abc}f_{a'}{}^{bc}\int\frac{d^Dk}{(2\pi)^D}\frac{F_{\alpha\alpha'}(k,p)}{k^2(k-p)^2}\ , \qquad (12.25)
\end{aligned}
$$

where

$$
\begin{aligned}
F_{\alpha\alpha'}(k,p) &= -(2k-p)_\alpha(2k-p)_{\alpha'}D + (2k-p)_\alpha(p+k)_{\alpha'} + (2k-p)_\alpha(k-2p)_{\alpha'} \\
&\quad +(p+k)_\alpha(2k-p)_{\alpha'} - (p+k)^2\delta_{\alpha\alpha'} - (k-2p)_\alpha(p+k)_{\alpha'} \\
&\quad +(k-2p)_\alpha(2k-p)_{\alpha'} - (p+k)_{\alpha'}(k-2p)_\alpha - (k-2p)^2\delta_{\alpha\alpha'} \\
&= (-4D+6)k_\alpha k_{\alpha'} + (-D+6)p_\alpha p_{\alpha'} + (2D-3)(k_\alpha p_{\alpha'} + p_\alpha k_{\alpha'}) \\
&\quad -(2k^2+5p^2-2k\cdot p)\delta_{\alpha\alpha'}. \qquad (12.26)
\end{aligned}
$$

We note that $\tilde{g}^2 = g^2\mu^\epsilon$ and $f_{abc}f_{a'}{}^{bc} = \delta_{aa'}C_2(G)$, where by definition $\delta_{aa'}$ is the Killing metric on the group and $C_2(G)$ is the second Casimir.

The first integral that we need is one that we already calculated,

$$
\begin{aligned}
\int\frac{d^Dk}{(2\pi)^D}\frac{1}{k^2(k-p)^2} &= \frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}(p^2)^{\frac{D}{2}-2}\int_0^1 d\alpha[\alpha(1-\alpha)]^{\frac{D}{2}-2} \\
&= \frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}(p^2)^{\frac{D}{2}-2}B\left(\frac{D}{2}-1,\frac{D}{2}-1\right) \\
&\equiv I_{2,D}(p)\ , \qquad (12.27)
\end{aligned}
$$

where we have used the Euler beta function, $B(a,b) = \int_0^1 dz \ \ z^{a-1}(1-z)^{b-1} = \Gamma(a+b)/\Gamma(a)\Gamma(b)$.

Then by Lorentz invariance we have

$$\int \frac{d^Dk}{(2\pi)^D}\frac{k_\mu}{k^2(k-p)^2} = p_\mu \tilde{I}(p) \ , \tag{12.28}$$

since the integral depends on p_μ (and not just on p^2). By multiplying this relation by $2p^\mu$, we obtain

$$2p^2\tilde{I}(p) = \int \frac{d^Dk}{(2\pi)^D}\frac{-(k-p)^2+k^2+p^2}{k^2(k-p)^2} = -\int \frac{d^Dk}{(2\pi)^D}\frac{1}{k^2} + \int \frac{d^Dk}{(2\pi)^D}\frac{1}{(k-p)^2} + p^2 I_{2,D}(p) \ , \tag{12.29}$$

where we have used that the first two integrals cancel against each other, by shifting the momentum $\tilde{k} = k - p$. Actually, the integral is zero in dimensional regularization in any case, since as we saw

$$\int \frac{d^Dq}{(2\pi)^D}\frac{1}{q^2+m^2} = \frac{\Gamma\left(1-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}(m^2)^{\frac{D}{2}-1} \ , \tag{12.30}$$

whose $m \to 0$ limit gives zero for $D > 2$. Then we obtain

$$\tilde{I}(p) = \frac{I_{2,D}(p)}{2}. \tag{12.31}$$

Next, we need the integral

$$\int \frac{d^Dk}{(2\pi)^D}\frac{k_\mu k_\nu}{k^2(k-p)^2} \ , \tag{12.32}$$

which by Lorentz invariance (since the integral depends on p_μ, and the only symmetric tensor Lorentz structures available are thus $p_\mu p_\nu$ and $p^2\delta_{\mu\nu}$) equals

$$I^a(p)p_\mu p_\nu + I^b(p)p^2\delta_{\mu\nu}. \tag{12.33}$$

We need two relations to determine I^a and I^b. The first one is obtained by multiplying with $2p^\mu$, which gives

$$\begin{aligned}
2p^2p_\nu(I^a+I^p) &= \int \frac{d^Dk}{(2\pi)^D}\frac{k_\nu[-(k-p)^2+k^2+p^2]}{k^2(k-p)^2} \\
&= -\int \frac{d^Dk}{(2\pi)^D}\frac{k_\nu}{k^2} + \int \frac{d^Dk}{(2\pi)^D}\frac{k_\nu}{(k-p)^2} + p^2\frac{p_\nu}{2}I_{2,D}(p) \\
&= -\int \frac{d^Dk}{(2\pi)^D}\frac{k_\nu}{k^2} + \int \frac{d^D\tilde{k}}{(2\pi)^D}\frac{\tilde{k}_\nu}{\tilde{k}^2} + p_\nu \int \frac{d^D\tilde{k}}{(2\pi)^D}\frac{1}{\tilde{k}^2} + p^2\frac{p_\nu}{2}I_{2,D}(p) \\
&= p^2\frac{p_\nu}{2}I_{2,D}(p) \ ,
\end{aligned} \tag{12.34}$$

where in the first equality we have used (12.31) and in the last equality we have used that (12.30) vanishes. We thus obtian

$$I^a(p) + I^b(p) = \frac{I_{2,D}(p)}{4}. \tag{12.35}$$

The other relation is obtained by contracting with $\delta^{\mu\nu}$, which gives

$$p^2[I^a + DI^b] = \int \frac{d^Dk}{(2\pi)^D}\frac{1}{(k-p)^2} = 0 \ , \tag{12.36}$$

so

$$I^a = -DI^b \ , \tag{12.37}$$

so that finally

$$I^a(p) = \frac{D}{4(D-1)}I_{2,D}(p); \quad I^b(p) = -\frac{1}{4(D-1)}I_{2,D}(p). \tag{12.38}$$

Putting all the pieces of the amplitude together, we obtain

$$\begin{aligned}
\mathcal{M}^{aa'}_{\alpha\alpha'} &= -g^2\mu^\epsilon C_2(G)\delta_{aa'}I_{2,D}(p)\Big[(-4D+6)\left(\frac{D}{4(D-1)}p_\alpha p_{\alpha'} - \frac{\delta_{\alpha\alpha'}}{4(D-1)}p^2\right)\\
&\quad +(-D+6)p_\alpha p_{\alpha'} + (2D-3)p_\alpha p_{\alpha'} - \delta_{\alpha\alpha'}(5p^2-p^2)\Big]\\
&= -g^2\mu^\epsilon\delta_{aa'}C_2(G)\left[p_\alpha p_{\alpha'}\frac{D+4(D-1)(D+3)}{4(D-1)} - \delta_{\alpha\alpha'}p^2\frac{16(D-1)-4D+6}{4(D-1)}\right]\times\\
&\quad \times\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}(p^2)^{\frac{D}{2}-2}B\left(\frac{D}{2}-1,\frac{D}{2}-1\right).
\end{aligned} \tag{12.39}$$

From it, we can obtain the divergent part of the amplitude, using that $\Gamma\left(2-\frac{D}{2}\right) = \Gamma\left(\frac{\epsilon}{2}\right) \simeq 2/\epsilon$, and in the rest putting $D=4$, including $B(1,1)=1$, so that finally

$$\mathcal{M}^{aa'}_{\alpha\alpha',\text{div.}} = -\delta_{aa'}C_2(G)\frac{g^2}{(4\pi)^2}\frac{2}{\epsilon}\left(\frac{11}{3}p_\alpha p_{\alpha'} - \frac{19}{6}\delta_{\alpha\alpha'}\right). \tag{12.40}$$

Important concepts to remember

- The ghost propagator is the KG propagator (times $\delta_{aa'}$), and the gluon propagator is the photon propagator (times $\delta_{aa'}$).
- We have a 3-gluon vertex, a 4-gluon vertex and a gluon-2-ghost vertex. If we add fermions, we also have a gluon-2-fermions vertex.
- A ghost loop gives a factor of (-1), same as a fermion loop.
- Green's functions in the nonabelian gauge theory are not gauge invariant, so the cannot be directly related to observables. Sums of diagrams (for different Green's functions, even) might be gauge invariant, order by order in perturbation theory.

Further reading: See chapter 7.2,7.3 in [5] and chapter 16.1 in [3].

Exercises, Lecture 12

1) Write the integral expressions for the diagrams (Fig.35)

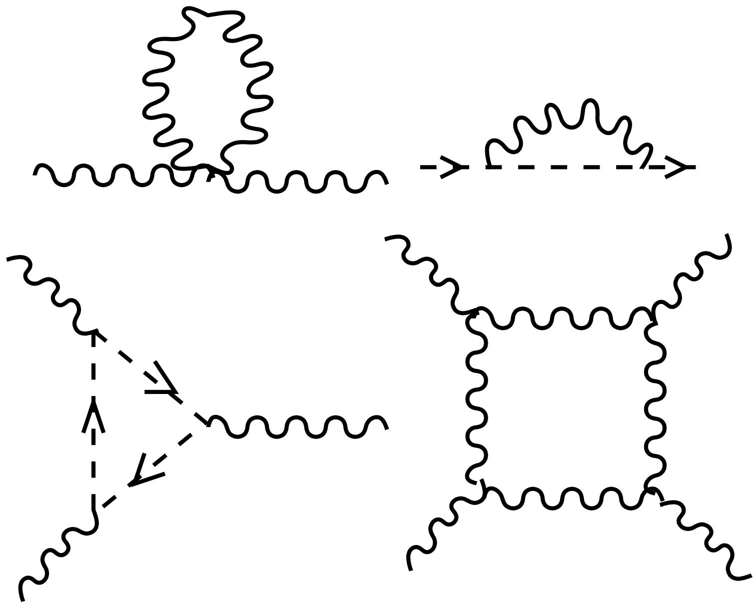

Figure 35: One loop QCD diagrams.

using the Feynman rules (without calculating them).

2) Calculate the divergent part of the diagram (Fig.36)

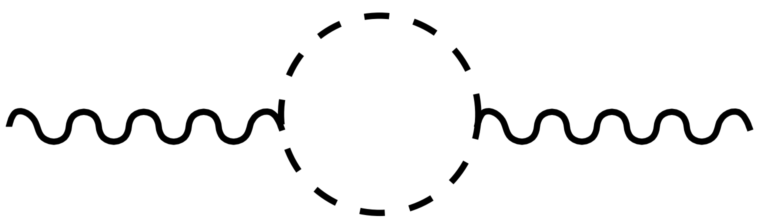

Figure 36: One loop QCD diagram with ghost loop and external gluons.

13 Lecture 13. One-loop renormalizability of gauge theories.

In this lecture we will study how to explicitly renormalize gauge theories at one-loop.

Pure gauge theory.

We will not derive them here, since they are too laborious (30 diagrams in total), but one can derive the result for the divergent parts of all the one-loop diagrams of pure gauge theory (Yang-Mills).

They fall into 5 classes, corresponding to 5 1PI n-point functions which are divergent.

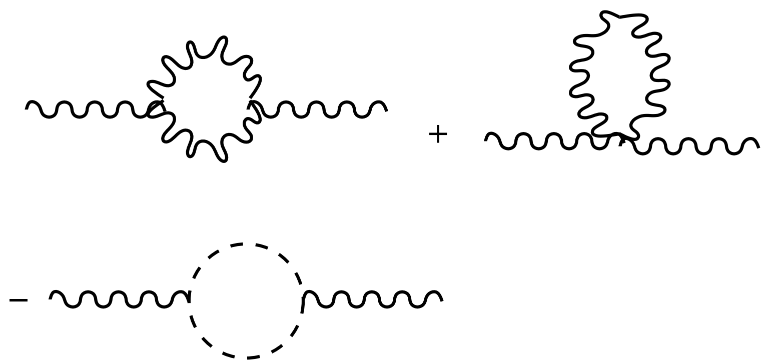

Figure 37: One loop diagrams contributing to the gauge propagator self-energy $\Sigma_{\mu\nu}(p)$.

1. The gauge propagator self-energy $\Sigma_{\mu\nu}(p)$. There are 3 relevant diagrams (see Fig.37): the gluon loop with 2 3-gluon vertices, whose divergent part we calculated last lecture (with symmetry factor 1/2); the ghost loop that was left as an exercise (with - sign for the ghost loop); the gluon loop with a single 4-gluon vertex. The divergent part of the sum of these diagrams gives

$$\Sigma_{\mu\nu;\,\text{divergent}}(p) = -\frac{g^2 C_2(G)}{16\pi^2}\left(\frac{5}{3}+\frac{1}{2}(1-\alpha)\right)\frac{2}{\epsilon}[p^2\delta_{\mu\nu}-p_\mu p_\nu]\delta^{ab}. \tag{13.1}$$

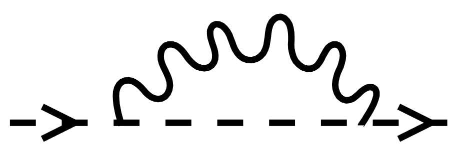

Figure 38: One loop diagram contributing to the ghost propagator self-energy $P_{ab}(p)$.

2. The ghost propagator self-energy P_{ab}, with a single one-loop diagram in Fig.38, with a gluon line starting and ending on the ghost line. Its divergent part is

$$P_{ab;\,\text{divergent}}(p) = -\frac{g^2 C_2(G)}{16\pi^2}\left(\frac{1}{2}+\frac{1}{4}(1-\alpha)\right)\frac{2}{\epsilon}(p^2\delta^{ab}). \tag{13.2}$$

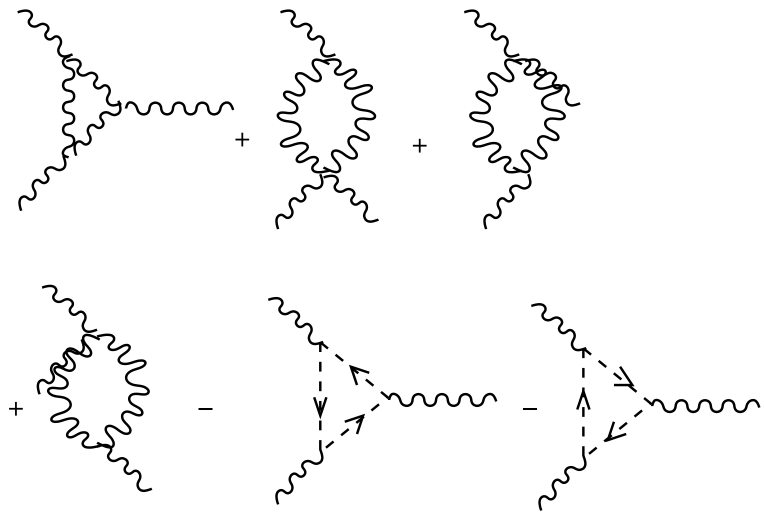

Figure 39: One loop diagrams contributing to the 3-gluon vertex.

3. The gluon 3-point vertex. There are now 6 diagrams, as in Fig.39: one with a gluon loop with 3 3-gluon vertices attached, out of which external gluons go; another 3 where we contract one propagator from the first diagram to make a gluon 4-point vertex (with symmetry factor 2 for each); 2 diagrams with a ghost loop with external gluons attached to it, and a different orientation for the ghost line differentiating between the 2 diagrams. The divergent part of the sum is

$$-\frac{g^2C_2(G)}{16\pi^2}\left[\frac{2}{3}+\frac{3}{4}(1-\alpha)\right]\frac{2}{\epsilon}V^{abc}_{\mu\nu\lambda}(k,p,q)\ , \tag{13.3}$$

where $V^{abc}_{\mu\nu\lambda}(k,p,q)$ is the classical (tree level) vertex.

4. The gluon 4-point vertex. There are 18 diagrams. The pure gluon ones are in Fig.40, and the ones with a ghost loop are in Fig.41. One is the gluon loop with 4 3-gluon vertices, together with its 2 crossed diagrams. There are 3 diagrams obtained by contracting 2 propagators of the first diagram to obtain 2 4-gluon vertices (or one diagram and 2 crossed ones). There are 6 diagrams where only one propagator has been contracted, to give one 4-gluon vertex and 2 3-gluon vertices; or 4 diagrams obtained by contraction of the first one, and 2 diagrams obtained by crossing. And finally there are 6 diagrams with a ghost loop with 4 external gluon vertices: two diagrams differing by the orientation of the ghost loop, and 2 crossed diagrams for each of these.

The divergent part of the sum of these diagrams is

$$-\frac{g^2C_2(G)}{16\pi^2}\left[-\frac{1}{3}+(1-\alpha)\right]\frac{2}{\epsilon}V^{abcd}_{\mu\nu\rho\sigma}\ , \tag{13.4}$$

where $V^{abcd}_{\mu\nu\rho\sigma}$ is the classical (tree level) vertex.

5. The ghost-gluon vertex. The loop is made of either two sides of ghost (with a gluon in the common vertex), and one gluon side; or two gluon sides (with the gluon coming out

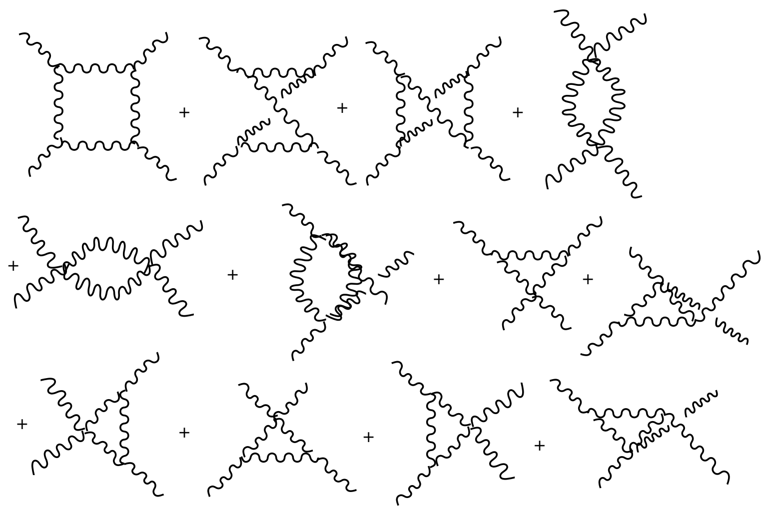

Figure 40: One loop gluon diagrams contributing to the 3-gluon vertex.

of the common vertex) and a single ghost side, as in Fig.42. The sum of these two diagrams has the divergent part

$$+\frac{g^2 C_2(G)}{16\pi^2}\frac{\alpha}{2}\frac{2}{\epsilon}V_\mu^{abc}(p)\ , \tag{13.5}$$

where again $V_\mu^{abc}(p)$ is the classical (tree level) vertex.

We see therefore that the 5 divergences all correspond to, and are proportional to, the 5 tree level objects (propagators and vertices), coming from the classical Lagrangean. Therefore renormalizability at one-loop is guaranteed, since we can absorb the divergences in the redefinition of the objects in the classical Lagrangean.

But in general, to guarantee renormalizability, one would need gauge invariance, since that generates relations betwen the Green's functions through the Ward identities. However, the formalism we use is gauge fixed, so it would seem we have a problem. As it turns out, we don't, because there is a *residual gauge symmetry* left after fixing the gauge (we know that fixing the gauge in general allows for the possibility of residual symmetries, meaning symmetry with a parameter that depends in a constrained way on spacetime). This symmetry is BRST symmetry, that will be studied in the next lecture, and we will see in the lecture after that, that it leads to relations among Green's functions similar to the Ward identities, which will allow us to prove renormalizability.

We also want to verify that the counting of the superficial degree of divergences says we have identified all the divergent 1PI n-point functions. The superficial degree of divergence in the pure gauge theory is

$$\omega(D) = 4 - E_A - \frac{3}{2}E_{b,c}\ , \tag{13.6}$$

where we have introduced also external ghost lines. Note the factor of 3/2 in front of the external ghost lines $E_{b,c}$, like in the case of fermions. The derivation is however slightly different than for the fermions. The vertex is $(\partial_\mu b^a) g f_{abc} A_\mu^b c^c$, so each b line at a vertex

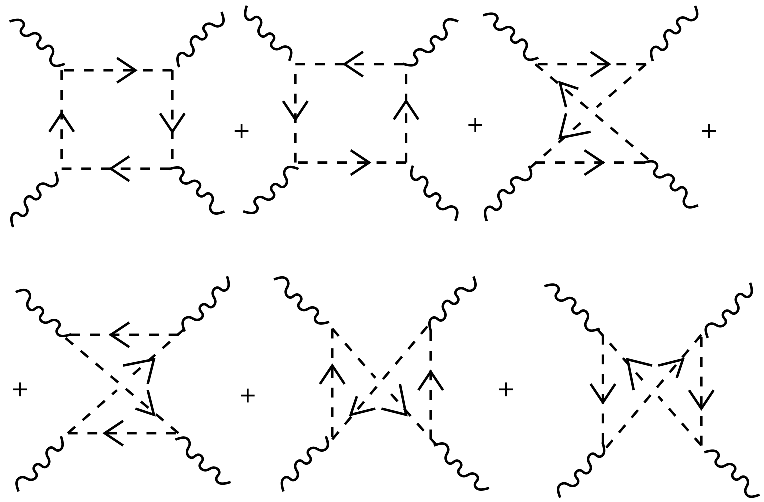

Figure 41: One loop diagrams with ghost loop contributing to the 3-gluon vertex.

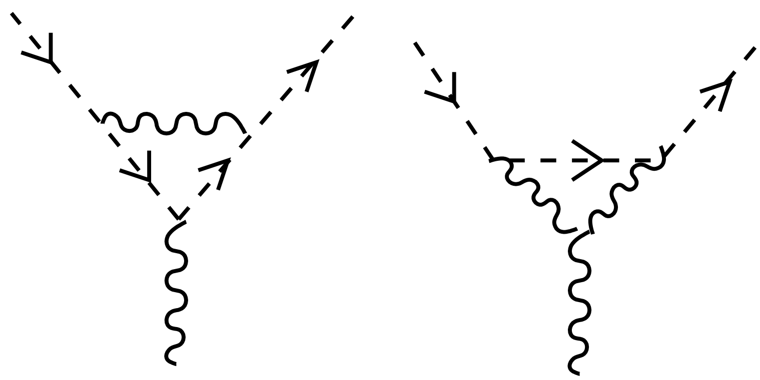

Figure 42: One loop diagrams for the ghost-gluon vertex.

comes with a factor of momentum, despite the ghost propagator being KG, $\sim 1/p^2$. Since each propagator ends in two vertices, but in one as b and in one as c, effectively we have an extra factor of p in the propagator, giving a $\sim 1/p$ propagator, like in the fermion case.

Then we can check that indeed the 5 1PI n-point functions are the divergent ones. The gluon propagator has $E_g = 2$, so $\omega(D) = 4 - 2 = 2$. The ghost propagator has $E_{bc} = 2$, so $\omega(D) = 4 - 3 \cdot 2/2 = 1$. The 3-gluon vertex has $E_g = 3$, so $\omega(D) = 4 - 3 = 1$, the 4-gluon vertex has $E_g = 4$, so $\omega(D) = 4 - 4 = 0$, and the ghost-gluon vertex has $E_g = 1, E_{bc} = 2$, so $\omega(D) = 4 - 1 - 3 \cdot 2/2 = 0$. There are no other divergent 1PI n-point functions.

Counterterms in MS scheme.

We can now write down the counterterms in the minimal subtraction scheme, as just minus the divergent terms.

1. From the gluon propagator one-loop divergence, we obtain the counterterm

$$\begin{aligned}
\delta\mathcal{L}_{A^2} &= (Z_3-1)\frac{1}{4}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2\\
Z_3^{(0+1)} &= 1+\frac{g^2C_2(G)}{16\pi^2}\left(\frac{5}{3}+\frac{1}{2}(1-\alpha)\right)\frac{2}{\epsilon}.
\end{aligned}\tag{13.7}$$

Indeed, in p space, the gluon propagator term gives $(Z_3-1)\frac{1}{2}A_\mu^a(p^2\delta_{\mu\nu}-p_\mu p_\nu)A_\nu^a$, so the counterterm is indeed minus the divergence.

2. From the ghost propagator one-loop divergence, we obtain the counterterm

$$\begin{aligned}
\delta\mathcal{L}_{bc} &= (\tilde{Z}_3-1)(-b^a\partial^2c^a)\\
\tilde{Z}_3^{(0+1)} &= 1+\frac{g^2C_2(G)}{16\pi^2}\left(\frac{1}{2}+\frac{1}{4}(1-\alpha)\right)\frac{2}{\epsilon}.
\end{aligned}\tag{13.8}$$

In p space, the gluon propagator term gives $(\tilde{Z}_3-1)(b^ap^2c^a)$, so indeed the counterterm is minus the divergence.

3. From the 3-gluon vertex divergence, we obtain the counterterm

$$\begin{aligned}
\delta\mathcal{L}_{A^3} &= (Z_1-1)gf_{abc}\partial_\mu A_\nu^a A_\mu^b A_\nu^c\\
Z_1^{(0+1)} &= 1+\frac{g^2C_2(G)}{16\pi^2}\left[\frac{2}{3}+\frac{3}{4}(1-\alpha)\right]\frac{2}{\epsilon}.
\end{aligned}\tag{13.9}$$

Since the divergence was written as a coefficient times the classical vertex, the $Z_1^{(1)}$ is just minus the coeffient, as we can check.

4. From the 4-gluon vertex divergence, we obtain the counterterm

$$\begin{aligned}
\delta\mathcal{L}_{A^4} &= (Z_4-1)\frac{g^2}{4}f_{abc}f^a{}_{de}A_\mu^bA_\nu^cA_\mu^dA_\nu^e\\
Z_4^{(0+1)} &= 1+\frac{g^2C_2(G)}{16\pi^2}\left[-\frac{1}{3}+(1-\alpha)\right]\frac{2}{\epsilon}.
\end{aligned}\tag{13.10}$$

The same comment as above applies.

5. From the ghost-gluon vertex divergence, we obtain the counterterm

$$\begin{aligned}
\delta\mathcal{L}_{bAc} &= (\tilde{Z}_1-1)gf_{abc}(\partial_\mu b^a)A_\mu^bc^c\\
\tilde{Z}_1^{(0+1)} &= 1-\frac{g^2C_2(G)}{16\pi^2}\frac{\alpha}{2}\frac{2}{\epsilon}.
\end{aligned}\tag{13.11}$$

Again the same comment applies.

As usual, renormalization means

$$(\mathcal{L}+\delta\mathcal{L})(A,b,c,g,\alpha)=\mathcal{L}_0(A_0,b_0,c_0,g_0,\alpha_0).\tag{13.12}$$

Since we have

$$\mathcal{L}+\delta\mathcal{L} = \frac{1}{4}Z_3(\partial_\mu A_\nu^a-\partial_\nu A_\mu^a)^2+\frac{1}{2\alpha}(\partial_\mu A_\mu^a)^2+gf_{abc}Z_1\partial_\mu A_\nu^aA_\mu^bA_\nu^c+\frac{g^2}{4}f_{abc}f^a{}_{de}Z_4A_\mu^bA_\nu^cA_\mu^dA_\nu^e$$

$$+\tilde{Z}_3(\partial_\mu b^a)\partial_\mu c^a + g f_{abc}\tilde{Z}_1(\partial_\mu b^a)A_\mu^b c^c \ , \tag{13.13}$$

we can obtain the renormalizations of fields and couplings.

Renormalization and consistency conditions

But we see that we have 4 objects to be renormalized (since b and c must renormalize in the same way), but 6 coefficients (for 6 terms), we will obtain also 2 consistency conditions.

From

$$Z_3(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2 = (\partial_\mu A_{0\nu}^a - \partial_\nu A_{0\mu}^a)^2 \ , \tag{13.14}$$

we get

$$A_0 = \sqrt{Z_3}A. \tag{13.15}$$

From

$$\tilde{Z}_3(\partial_\mu b^a)\partial_\mu c^a = (\partial_\mu b_0^a)\partial_\mu c_0^a \ , \tag{13.16}$$

we get

$$b_0, c_0 = \sqrt{\tilde{Z}_3} b, c. \tag{13.17}$$

From

$$\frac{1}{2\alpha}(\partial_\mu A_\mu)^2 = \frac{1}{2\alpha_0}(\partial_\mu A_{0\mu})^2 \ , \tag{13.18}$$

we get

$$\alpha_0 = Z_3\alpha. \tag{13.19}$$

From

$$g f_{abc} Z_1 \partial_\mu A_\nu^a A_\mu^b A_\nu^c = g_0 f_{abc}\partial_\mu A_{0\nu}^a A_{0\mu}^b A_{0\nu}^c \ , \tag{13.20}$$

we get

$$g_0 = g\frac{Z_1}{Z_3^{3/2}}. \tag{13.21}$$

Now we have fixed all renormalizations, but we still have two terms to check. These will give consistency conditions. The first is the gluon 4-vertex,

$$\sim g^2 Z_4 A^4 = g_0^2 A_0^4 \ , \tag{13.22}$$

from which we obtain the consistency condition

$$\frac{Z_4}{Z_1} = \frac{Z_1}{Z_3}. \tag{13.23}$$

The second is the ghost-gluon vertex,

$$\sim g\tilde{Z}_1 bAc = g_0 b_0 A_0 c_0 \ , \tag{13.24}$$

from which we obtain the second consistency condition,

$$\frac{\tilde{Z}_1}{\tilde{Z}_3} = \frac{Z_1}{Z_3}. \tag{13.25}$$

Together, these two relations form the *Slavnov-Taylor identities*.

At this point, it is not clear why they should be correct. Such relations normally appear from gauge invariance, through Ward identities that relate various 1PI n-point functions (thus their coefficients Z_i), but now we work in a gauge fixed formalism. However, as we said, there is a residual gauge symmetry called BRST symmetry, which will allow us to write relations between the n-point functions.

We can at most check it explicitly at one-loop. We remember that $Z_i = 1 + Z_i^{(1)}$ can be expanded in $Z_i^{(1)}$, so the Slavnov-Taylor identities at one-loop are

$$\begin{aligned} Z_4^{(1)} &= 2Z_1^{(1)} - Z_3^{(1)} \\ \tilde{Z}_1^{(1)} &= Z_1^{(1)} + \tilde{Z}_3^{(1)} - Z_3^{(1)}. \end{aligned} \tag{13.26}$$

The relations are indeed verified, since

$$\begin{aligned} \left[-\frac{1}{3} + (1-\alpha)\right] &= +2\left[\frac{2}{3} + \frac{3}{4}(1-\alpha)\right] - \left[\frac{5}{3} + \frac{1}{2}(1-\alpha)\right] \\ -\frac{\alpha}{2} &= \left[\frac{2}{3} + \frac{3}{4}(1-\alpha)\right] + \left[\frac{1}{2} + \frac{1}{4}(1-\alpha)\right] - \left[\frac{5}{3} + \frac{1}{2}(1-\alpha)\right]. \end{aligned} \tag{13.27}$$

Gauge theory with fermions

We can couple the gauge theory to fermions with Euclidean action

$$S_{(E)} = \int d^D x \bar{\psi}(\gamma_\mu D_\mu + m)\psi \ , \tag{13.28}$$

where the covariant derivative is

$$D_\mu^{ij} = \partial_\mu \delta^{ij} + g(T_f^a)_{ij} A_\mu^a. \tag{13.29}$$

The superficial degree of divergence in the presence of fermions is now

$$\omega(D) = 4 - E_g - \frac{3}{2}E_f - \frac{3}{2}E_{bc}. \tag{13.30}$$

Therefore, besides the previous divergent 1PI n-point functions, we also have new ones that are divergent:

-the fermion propagator $\Sigma_{\alpha\beta}(p)$, with $\omega(D) = 4 - 3\cdot 2/2 = 1$.

-the fermion-gluon vertex, $\Gamma_{\alpha\beta}^{a\mu}$, with $\omega(D) = 4 - 1 - 3\cdot 2/2 = 0$.

But before we analyze those, we will write down the new divergent contributions to the 1PI n-point functions alread considered in the pure gauge theory case, coming from a fermion loop:

-the divergent contribution to the gluon propagator, i.e. to Z_3, coming from the fermion loop with two external gluons from it in Fig.43, giving

$$-\frac{g^2}{16\pi^2}T_f\frac{4}{3}\frac{2}{\epsilon}. \tag{13.31}$$

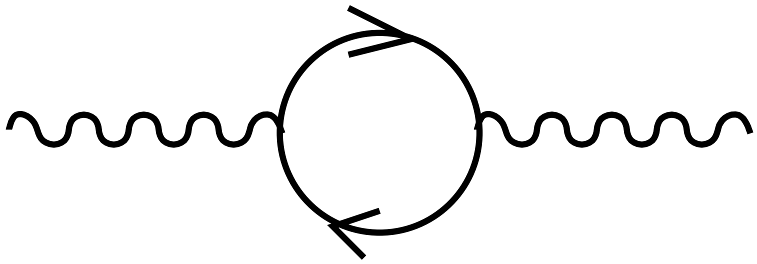

Figure 43: One fermion loop diagram for the gluon propagator.

-the divergent contribution to the 3-gluon vertex, i.e. to Z_1, coming from the two diagrams in Fig.44, with a fermion loop (with different orientation for the arrow) and 3 gluons coming out of it, giving the same

$$-\frac{g^2}{16\pi^2}T_f\frac{4}{3}\frac{2}{\epsilon}. \tag{13.32}$$

This is good, since from the second Slavnov-Taylor identity at one-loop in (13.26), we need that (since $\tilde{Z}_1$ and $\tilde{Z}_3$ are not modified by the addition of fermions) $\delta Z_1^{(1)} = \delta Z_3^{(1)}$, which is indeed true.

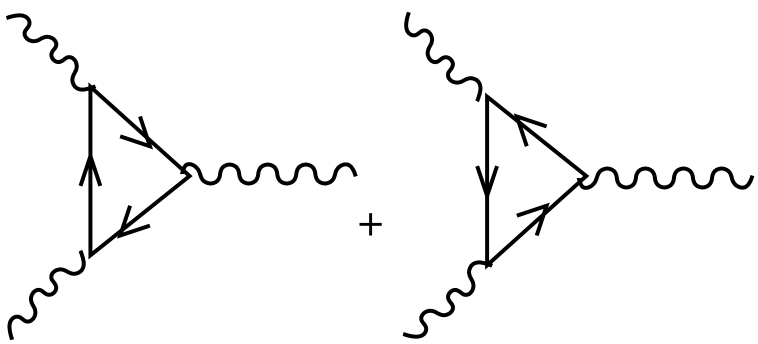

Figure 44: One fermion loop diagrams for the 3-gluon vertex.

-the divergent contribution to the 4-gluon vertex, i.e. to Z_4, again coming from two diagrams with a fermion loop (and different orientations for the arrow) and 4 gluons coming out of it, giving the same result as above for Z_1.

This is again good, since now from the first Slavnov-Taylor identity at one-loop in (13.26), we need that (since we saw already that $\delta Z_1^{(1)} = \delta Z_3^{(1)}$) $\delta Z_4^{(1)} = \delta Z_1^{(1)}$, which is verified.

Moving on to the new divergent 1PI n-point functions, these new divergences will be cancelled by the counterterms

$$\delta\mathcal{L}(A,b,c,\psi) = (Z_{f2}-1)\bar{\psi}\gamma_\mu\partial_\mu\psi + (Z_{f1}-1)g\bar{\psi}\gamma_\mu A_\mu\psi + (Z_m-1)m\bar{\psi}\psi. \tag{13.33}$$

Explicit calculations give

$$Z_{f2}^{(0+1)} = 1 - \frac{g^2 C_f}{16\pi^2}\alpha\frac{2}{\epsilon}$$

$$
\begin{aligned}
Z_{f1}^{(0+1)} &= 1 - \frac{g^2 C_f}{16\pi^2}\left[\alpha C_f + C_2(G)\left(1 - \frac{1-\alpha}{4}\right)\right]\frac{2}{\epsilon} \\
Z_m^{(0+1)} &= 1 - \frac{g^2 C_f}{16\pi^2}[4 - (1-\alpha)]\frac{2}{\epsilon}.
\end{aligned} \quad (13.34)
$$

The correction to the fermion propagator, giving both the Z_{f2} and the Z_m terms, comes from a one loop diagram with a fermion line out of which a gluon comes out and back. The correction to the fermion-gluon vertex is given by the two diagrams equivalent to the ones of the ghost-gluon vertex, namely with a triangle loop two sides fermionic and one side gluon, or two sides gluonic and one fermionic.

In the above C_f and T_f are the fermion representation case of T_R and C_R defined previously, namely T_f is the normalization of the trace, and C_f the Casimir in this representation.

For the QED case (abelian group), $C_f = 1$ and $C_2(G) = 0$, and we can check that we reproduce the results we obtained for QED.

Renormalization *of the new terms* is given as usual by

$$
\mathcal{L}(A, \bar{\psi}, g, m) + \delta\mathcal{L}(A, \bar{\psi}, \psi, g, m) = \mathcal{L}_0(A_0, \bar{\psi}_0, \psi_0, g_0, m_0)\ , \quad (13.35)
$$

which means that

$$
Z_{f2}\bar{\psi}\gamma_\mu\partial_\mu\psi + Z_{f1}g\bar{\psi}\gamma_\mu A_\mu\psi + Z_m m\bar{\psi}\psi = \bar{\psi}_0\gamma_\mu\partial_\mu\psi_0 + g_0\bar{\psi}_0\gamma_\mu A_{0\mu}\psi_0 + m_0\bar{\psi}_0\psi_0. \quad (13.36)
$$

From the first term we get the wave function renormalization

$$
\psi_0 = \sqrt{Z_{f2}}\psi; \quad \bar{\psi}_0 = \sqrt{Z_{f2}}\bar{\psi}_0\ , \quad (13.37)
$$

and from the last we get the mass renormalization

$$
Z_{f2} m_0 = m Z_m. \quad (13.38)
$$

The middle term gives a constraint, i.e. another Slavnov-Taylor identity,

$$
Z_{f1} g\bar{\psi}\gamma_\mu A_\mu\psi = g_0\bar{\psi}_0\gamma_\mu A_{0\mu}\psi_0 \Rightarrow \frac{Z_{f1}}{Z_{f2}} = \frac{Z_1}{Z_3}. \quad (13.39)
$$

We can check explicitly that it is satisfied at one-loop, as in the pure gauge case.

That means that all in all, we have the Slavnov-Taylor identities

$$
\frac{Z_{f2}}{Z_{f1}} = \frac{Z_1}{Z_3} = \frac{\tilde{Z}_1}{\tilde{Z}_3} = \frac{Z_4}{Z_1}. \quad (13.40)
$$

Important concepts to remember

- Pure gauge theory is renormalizable at one-loop, the divergences coming from the gauge propagator, ghost propagator, 3-gluon vertex, 4-gluon vertex and ghost-gluon vertex, and being of the same structure as the terms in the classical Lagrangean.

- The pure gauge theory obeys the Slavnov-Taylor identities, $Z_4/Z_1 = Z_1/Z_3 = \tilde{Z}_1/\tilde{Z}_3$.
- The gauge theory with fermions has $\omega(D) = 4 - E_g - 3E_{bc}/2 - 3E_f/2$ and introduces two more divergences, in the fermion propagator and fermion-gluon vertex, being again one-loop renormalizable.
- With fermions we have one more Slavnov-Taylor identity, $Z_{f2}/Z_{f1} = Z_1/Z_3$.

Further reading: See chapter 7.3,7.4,7.5 in [5] and chapter 16.5 in [3].

Exercises, Lecture 13

1) Calculate explicitly all the Z factors for a $SO(N)$ gauge theory with N_f fundamental fermions.

2) Calculate the divergent part of the one-loop graph with a fermion loop and two external gluon lines in Fig.43, contributing to Z_3.

14 Lecture 14. Asymptotic freedom. BRST symmetry.

Asymptotic freedom

Asymptotic freedom is a very important concept, one that has gained a Nobel prize for David Gross, Frank Wilczek and David Politzer (in 2004). We have done most of the calculations necessary for it last lecture, so now it remains to put everything together and interpret it.

We saw last lecture that renormalization of the coupling is done by $g_0 = gZ_1/Z_3^{3/2}$. However, we omitted the dimensional transmutation factor, so we actually have

$$g_0 = g\mu^{\epsilon/2}\frac{Z_1}{Z_3^{3/2}}. \tag{14.1}$$

We also saw that, in the presence of fermions we have

$$\begin{aligned} Z_1^{(0+1)} &= 1 + \frac{g^2}{16\pi^2}\left[C_2(G)\left(\frac{2}{3}+\frac{3}{4}(1-\alpha)\right)-\frac{4}{3}T_f\right]\frac{2}{\epsilon} \\ Z_3^{(0+1)} &= 1 + \frac{g^2}{16\pi^2}\left[C_2(G)\left(\frac{5}{3}+\frac{1}{2}(1-\alpha)\right)-\frac{4}{3}T_f\right]\frac{2}{\epsilon}. \end{aligned} \tag{14.2}$$

Then it follows that

$$\begin{aligned} g_0 &= g\mu^{\frac{\epsilon}{2}}\left\{1+\frac{g^2}{16\pi^2}\left[C_2(G)\left(\frac{2}{3}+\frac{3}{4}(1-\alpha)-\frac{3}{2}\left(\frac{5}{3}+\frac{1}{2}(1-\alpha)\right)\right)-\frac{4}{3}T_f\left(1-\frac{3}{2}\right)\right]\frac{2}{\epsilon}\right\} \\ &= g\mu^{\frac{\epsilon}{2}}\left\{1-\frac{g^2}{16\pi^2}\left[\frac{11}{6}C_2(G)-\frac{2}{3}T_f\right]\frac{2}{\epsilon}\right\}. \end{aligned} \tag{14.3}$$

Taking $\mu\partial/\partial\mu$ on both sides of this equation, since $\partial g_0/\partial\mu = 0$, we obtain

$$\begin{aligned} \mu\frac{\partial g_0}{\partial\mu} = 0 = \frac{\epsilon}{2}\mu^{\frac{\epsilon}{2}}\left(g-\frac{g^3}{16\pi^2}\left[\frac{11}{6}C_2(G)-\frac{2}{3}T_f\right]\frac{2}{\epsilon}\right) \\ +\mu^{\frac{\epsilon}{2}}\left(1-\frac{3g^2}{16\pi^2}\left[\frac{11}{6}C_2(G)-\frac{2}{3}T_f\right]\frac{2}{\epsilon}\right)\mu\frac{\partial g}{\partial\mu}+\mathcal{O}(g^5). \end{aligned} \tag{14.4}$$

This is solved by

$$\begin{aligned} \beta(\mu,\epsilon) &\equiv \mu\frac{\partial g}{\partial\mu} = \frac{-\frac{\epsilon}{2}g+\frac{g^3}{16\pi^2}\left[\frac{11}{6}C_2(G)-\frac{2}{3}T_f\right]+\mathcal{O}(g^5)}{1-\frac{3g^2}{16\pi^2}\left[\frac{11}{6}C_2(G)-\frac{2}{3}T_f\right]\frac{2}{\epsilon}+\mathcal{O}(g^4)} \\ &= -\frac{\epsilon}{2}g-\frac{2g^3}{16\pi^2}\left[\frac{11}{6}C_2(G)-\frac{2}{3}T_f\right]+\mathcal{O}(g^5). \end{aligned} \tag{14.5}$$

Note that above we have used the usual expansion in g^2, ignoring the fact that a higher order term might actually be divergent in ϵ, it is still considered negligible.

One also defines as usual the physical beta function as

$$\beta(g) = \beta(\mu,\epsilon\to 0) = \beta_1 g^3+\beta_3 g^5+\beta_5 g^7+\ldots, \tag{14.6}$$

which leads to

$$\beta_1 = -\frac{2}{16\pi^2}\left[\frac{11}{6}C_2(G) - \frac{2}{3}T_f\right]. \tag{14.7}$$

Again, the term in g^5 from the above would actually be divergent in ϵ, but is still considered subleading, and ignored even in the $\epsilon \to 0$ limit. Of course, the point is that the calculation is actually only to one-loop. Once we do the two-loop calculation for Z_1 and Z_3 and includes it in the above, one finds a finite β_2 as well.

We note here that β_1 and β_3 are actually gauge indepedent (universal), nonzero, and independent of renormalization scheme, while the other coefficients are not necessarily universal.

The most important observation is that the *nonabelian* gauge fields are actually the only ones with $\beta_1 < 0$! All other fields have $\beta_1 > 0$ contributions. In particular, we see that for fermions we have a positive β_1, proportional to T_f. For scalars, we get similarly a positive contribution.

For QED, we have $C_2(G) = 0$ and $T_f = 1$, so $\beta_1 = +1/12\pi^2 > 0$.

Integrating the beta function equation at one-loop,

$$\beta_1 g^3 = \mu\frac{\partial g}{\partial \mu}\ , \tag{14.8}$$

by first multiplying with $2g$, we get the solution

$$g^2(\mu) = \frac{g^2(\mu_0)}{1 - \beta_1 g^2(\mu_0)\ln\frac{\mu^2}{\mu_0^2}}\ , \tag{14.9}$$

as we can explicitly check. This relates the coupling constant at some fixed scale μ_0 with the coupling constant at the variable scale μ.

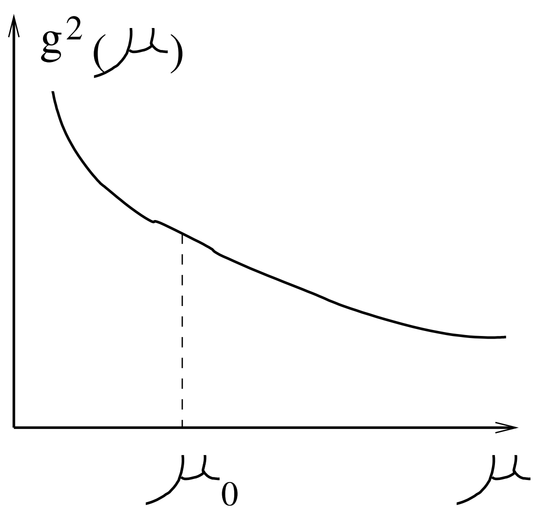

Figure 45: The coupling constant $g^2(\mu)$ is a decreasing function for QCD.

In particular, it means that, for $\beta_1 < 0$, the coupling constant $g^2(\mu)$ is a decreasing function of μ, as in Fig.45. This has two important consequences:

-IR slavery: $g^2(\mu \to 0) \to \infty$, so at large distances (small energies, i.e. IR), we have very strong coupling, leading to confinement.

-asymptotic freedom: $g^2(\mu \to \infty) \to 0$, so the theory is free in the UV.

For the gauge group $SU(N)$, we have $C_2(SU(N)) = N$ and $T_f = 1/2$, but we will consider N_f flavours (species) of fermions, leading to

$$\beta_{1,SU(N_c),N_f} = -\frac{1}{16\pi^2}\left(\frac{11}{3}N_c - \frac{2}{3}N_f\right). \tag{14.10}$$

In particular, for QCD $N_c = 3$ (3 colours) and $N_f = 6$ (6 flavours, u,d,c,s,t,b), so

$$\beta_{1,QCD} = -\frac{7}{16\pi^2} < 0\ , \tag{14.11}$$

so QCD is asymptotically free. It also exhibits IR slavery, leading to confinement of quarks and gluons, since the coupling becomes infinite at large distances, and we cannot separate the quarks and gluons from each other.

BRST symmetry.

To study the quantum properties of gauge theories, we will need a global symmetry called BRST symmetry, which is a remnant of gauge invariance (i.e., a residual gauge invariance), present once we fix the gauge.

It was found in a paper by Becchi, Rouet and Stora, and independently in a paper by Tyutin, hence the name BRST. For the ghost action, as before, we use the notation common in the BRST literature, with fields b^a and c^a. The effective gauge fixed Lagrangean in Euclidean space is

$$\mathcal{L}_{\rm eff.}(A,b,c) = +\frac{1}{4}(F^a_{\mu\nu})^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 + \partial_\mu b^a D^{ab}_\mu c^b. \tag{14.12}$$

Here b^a, c^a are anticommuting variables, and b^a is imaginary, while c^a is real, for the reality of the action. Since the dimension of the Lagrangean must be 4, we see that we need $[b^a] + [c^a] = 2$. But since a priori b^a and c^a are independent fields (remember that we found them by writing the gaussian path integral $\int \mathcal{D}b\mathcal{D}ce^{-bMc} = \det M$, where the two integrations are really independent), their dimensions need not be related. In fact, we can choose them of different dimension, as is done for instance in string theory. However, for simplicity, here we will choose $[b^a] = [c^a] = 1$.

Since the gauge invariance is

$$\delta_{{\rm gauge},\lambda} A^a_\mu = (D_\mu\lambda)^a = \partial_\mu\lambda^a + gf^a{}_{bc}A^b_\mu\lambda^c\ , \tag{14.13}$$

the BRST invariance must be similar, just with the arbitrary $\lambda^a(x)$ replaced by something. That something is $c^a(x)\Lambda$, so a given x dependence (the field c^a) times an arbitrary constant, implying a global symmetry.

We can define a *ghost number* by $N_{\rm gh}[c^a] = +1$ (since $c^a \to \eta^a$ and $b^a \to \bar{\eta}^a$, this is natural) and $N_{\rm gh}[b^a] = -1$. This is then a symmetry of $\mathcal{L}_{eff}$, i.e. the Lagrangean has ghost number zero. Since $N_{\rm gh}[\lambda]=0$, it follows that $N_{\rm gh}[\Lambda]=-1$. Moreover, since $[\lambda]=0$ because of $\delta A_\mu = D_\mu\lambda$, it follows that also $[\Lambda]=-1$. And since λ^a is commuting, it follows that Λ is anticommuting.

Since $c^a(x)\Lambda$ is a special case of λ^a, then it trivially follows that the classical Lagrangean is BRST invariant,

$$\delta_B \mathcal{L}_{class} = 0\ , \qquad (14.14)$$

under the BRST transformation

$$\delta_B A^a_\mu = (D_\mu c)^a \Lambda. \qquad (14.15)$$

Moreover, we can extend the BRST transformation to gauge fields coupled to matter. For instance, for scalars transforming under gauge transformations as

$$\delta_{\rm gauge}\phi^i = -g(T_a)^i{}_j\phi^j\lambda^a(x)\ , \qquad (14.16)$$

the BRST transformation is

$$\delta_B\phi^i = -g(T_a)^i{}_j\phi^j c^a\Lambda\ , \qquad (14.17)$$

and similarly for fermions, etc.

Now, to find the rest of the BRST transformation laws, we require that the rest of the action be invariant. Since $\delta_B A^a_\mu$ has no terms with b^a in it, and $\mathcal{L}_{eff}$ has terms with only A_μ and a term $-b_a\partial^\mu(D_\mu c)^a$, we must have $(D_\mu c)^a$ invariant, at least when multiplied by $\partial_\mu b^a$ (if not, there would be a term in the variation $-b_a\partial^\mu\delta(D_\mu c)^a$ that could not be cancelled).

We have explicitly

$$\delta_B(D_\mu c)^a = \partial_\mu(\delta_B c^a) + gf^a{}_{bc}(D_\mu c)^b\Lambda c^c + gf^a{}_{bc}A^b_\mu(\delta_B c^c). \qquad (14.18)$$

We can combine the first and the third terms into D_μ to find the equation

$$\delta_B(D_\mu c)^a = D_\mu(\delta_B c^a) + \frac{1}{2}D_\mu(gf^a{}_{bc}c^b\Lambda c^c) = 0\ , \qquad (14.19)$$

but only if we have the following identity satisfied (the terms with ∂_μ are clearly the same in the two expressions, and only the terms with A_μ need to be checked)

$$\left[\frac{1}{2}f^a{}_{pq}f^q{}_{bc} - f^a{}_{qc}f^q{}_{pb}\right]A^p_\mu c^b\Lambda c^c = 0\ , \qquad (14.20)$$

where the first term comes from $1/2D_\mu(gf^a{}_{bc}c^b\Lambda c^c)$ (in the final form for the variation) and the second term comes from $gf^a{}_{bc}(D_\mu c)^b\Lambda c^c$ (in the initial form for the variation). The above equality follows from the Jacobi identity (coming from the double commutator of the generators of the Lie algebra)

$$f^a{}_{pq}f^q{}_{bc} + f^a{}_{bq}f^q{}_{cp} + f^a{}_{cq}f^q{}_{pb} = 0\ , \qquad (14.21)$$

by noticing that the last two terms are equal when multiplied by an antisymmetric factor in (bc) like c^bc^c, and then by renaming the indices, using antisymmetry and multiplying by

1/2 we get the needed equality. The solution to (14.19) is found by first putting $A_\mu = 0$, in which case we can peel off the ∂_μ in both terms in the equation, to find

$$\delta_B c^a = \frac{1}{2} g f^a{}_{bc} c^b c^c \Lambda \;, \tag{14.22}$$

and then we can check that the A_μ terms also cancel identically.

Finally, now that we fixed the variation of c^a such that $(D_\mu c)^a$ is invariant, we have that the variation of the ghost term is

$$\delta_B \mathcal{L}_{\text{ghost}} = -(\delta_B b^a)\partial^\mu (D_\mu c)^a \;, \tag{14.23}$$

and it needs to cancel against the variation of the gauge fixing term (since the classical term is invariant)

$$\delta_B \mathcal{L}_{\text{gauge fix}} = \frac{1}{\alpha}(\partial^\rho A_\rho)\partial^\mu (D_\mu c)^a \Lambda. \tag{14.24}$$

This is achieved if

$$\delta_B b^a = -\frac{1}{\alpha}(\partial^\rho A_\rho)\Lambda. \tag{14.25}$$

In conclusion, the BRST transformation laws are

$$\begin{aligned} \delta_B A_\mu^a &= (D_\mu c)^a \Lambda \\ \delta_B c^a &= \frac{1}{2} g f^a{}_{bc} c^b c^c \Lambda \\ \delta_B b^a &= -\frac{1}{\alpha}\partial^\mu A_\mu^a \Lambda. \end{aligned} \tag{14.26}$$

For more general gauge fixing terms, written as

$$\mathcal{L}_{g.fix} = +\frac{1}{2}\gamma_{ab} F^a F^b \;, \tag{14.27}$$

where for the Lorenz gauge we have $F^a = \partial^\mu A_\mu^a$ and $\gamma_{ab} = \delta_{ab}/\alpha$, as we saw in lecture 11, we can write the ghost action as

$$-b_a \frac{\partial F^a}{\partial A_\mu^c} D_\mu^{cb} c^b = -b_a(\delta_B F^a)/\Lambda. \tag{14.28}$$

Here $/\Lambda$ means we take away Λ from the right (remember that Λ is anticommuting, so taking it away from the left or the right differs by a minus sign). Then the invariance of $\delta_B F^a/\Lambda$ fixes the same $\delta_B c^a$ as before, and in turn that implies

$$\delta_B b_a = -\gamma_{ab} F^b \Lambda \;, \tag{14.29}$$

since

$$\delta_B(+\frac{1}{2}\gamma_{ab} F^b F^a) - (\delta_B b_a)\delta_B F^a/\Lambda = \gamma_{ab} F^b \delta_B F^a = (-\gamma_{ab} F^b \Lambda)\delta_B F^a/\Lambda = 0. \tag{14.30}$$

Nilpotency of Q_B and the auxiliary field formulation.

We can define a *BRST charge* Q_B that acts by $e^{Q_B\Lambda}$, e.g. $\delta_B A^a_\mu = (Q_B A^a_\mu)\Lambda$, etc. Then Q_B can be made nilpotent, i.e.

$$Q_B^2 = 0. \tag{14.31}$$

Proof.

We saw that $\delta_B A^a_\mu = (D_\mu c)^a\Lambda$, and $\delta_B(D_\mu c)^a = 0$, so $\delta_B^2 A^a_\mu = 0$.

We also saw that $\delta c^a = \frac{1}{2} g f^a{}_{bc} c^b c^c \Lambda$, and on the other hand

$$\delta_B(f^a{}_{bc}c^b c^c) = 2f^a{}_{bc}c^b\delta c^c = g f^a{}_{bc} f^c{}_{de} c^b c^d c^e\ , \tag{14.32}$$

but this is zero by the Jacobi identity, since $c^b c^d c^e$ is totally antisymmetric in (bde), and

$$f^a{}_{[b|c|} f^c{}_{de]} = 0. \tag{14.33}$$

It follows that

$$\delta_B^2 c^a = 0. \tag{14.34}$$

On the other hand, on b^a, it is not quite nilpotent, since

$$\delta_B^2 b^a = -\frac{1}{\alpha}\partial^\mu(\delta_B A^a_\mu)\Lambda_1 = -\frac{1}{\alpha}\partial^\mu(D_\mu c)^a\Lambda_2\Lambda_1. \tag{14.35}$$

However, note that $\partial^\mu(D_\mu c)^a$ is the field equation for b^a, so $\delta_B^2 b^a = 0$ on-shell, or $Q_B^2 = 0$ on-shell.

This is a familiar situation for symmetries, in particular for supersymmetry. The algebra of charges closes only on-shell, so in order to make it close off-shell we must go to a first order formulation, through the introduction of auxiliary fields, with field equation related to the term we want to cancel.

In our case, the problematic term arises from $\delta_B(-(\partial^\mu A_\mu)/\alpha)$, which is related to the variation of the gauge fixing term. We will fix it introducing an auxiliary field called *Nakanishi-Lautrup field d_a*.

Then we write the gauge fixing term in first order formulation, as

$$\mathcal{L}_{\text{g.fix}} = -\frac{\alpha}{2}(d_a)^2 - d_a(\partial^\rho A^a_\rho). \tag{14.36}$$

The equation of motion for the auxiliary field d_a gives

$$d_a = -\frac{1}{\alpha}(\partial^\rho A^a_\rho)\ , \tag{14.37}$$

which is what we have in the variation δ_B of b^a. By replacing it in the gauge fixing term, we get back to the second order formulation.

Therefore we write the new variations of b_a and d_a as

$$\delta_B b_a = d_a\Lambda; \quad \delta_B d_a = 0. \tag{14.38}$$

Now we see explicitly that

$$\delta_B^2 b_a = 0 = \delta_B^2 d_a \Rightarrow Q_B^2 = 0. \tag{14.39}$$

Let us verify that the above variations leave the quantum action invariant. We have $\delta_B(\alpha(d_a)^2/2) = 0$, and

$$d_a\delta_B(\partial^\rho A_\rho) + (\delta_B b_a)\partial^\mu(D_\mu c)^a = 0 \; , \tag{14.40}$$

so indeed the action is invariant.

And now, as promised, $Q_B^2 = 0$ off-shell. Since δ_B acts by $Q_B\Lambda$, and $N_{gh}[\Lambda] = -1$, it follows that $N_{gh}[Q_B] = +1$. A more important observation is that the rules are now purely kinematical,

$$\begin{aligned}\delta_B A_\mu^a &= (D_\mu c)^a\Lambda; \quad \delta_B c^a = \frac{1}{2}g f^a{}_{bc}c^b c^c\\ \delta_B b_a &= d_a\Lambda; \quad \delta_B d_a = 0.\end{aligned} \tag{14.41}$$

Indeed, we see that they are completely independent of the gauge fixing term, in particular are independent of α, so they apply to any quantum gauge theory.

The gauge fixing term can be rewritten as

$$-\frac{1}{2}\gamma^{ab}d_a d_b - d_a F^a = -\frac{1}{2}\gamma^{ab}d_a\delta_B b_b/\Lambda - (\delta_B b_a)/\Lambda F^a \; , \tag{14.42}$$

and we already saw that the ghost term is

$$-b_a\delta_B F^a/\Lambda \; , \tag{14.43}$$

so all in all the quantum action can be written as

$$\mathcal{L}_{qu} = \mathcal{L}_{class} - \delta_B\left(b_a\left(F^a + \frac{\gamma^{ab}}{2}d_b\right)\right)/\Lambda. \tag{14.44}$$

This form is valid more generally, and it trivializes the BRST invariance of the quantum action, $\delta_B\mathcal{L}_{qu} = 0$, since it is now simply a result of the gauge invariance of the classical action, plus the nilpotency of Q_B, $Q_B^2 = 0$, which is as we saw true for the kinematical (model-independent) transformations (14.41).

The quantity

$$\psi \equiv b_a\left(F^a + \frac{\gamma^{ab}}{2}d_b\right) \; , \tag{14.45}$$

is called the *gauge fixing fermion*, and it can be made even more general. In terms of it, the action is

$$\mathcal{L}_{qu} = \mathcal{L}_{class} + \{Q_B, \psi\}. \tag{14.46}$$

Important concepts to remember

- Only nonabelian gauge fields have a negative contribution to β_1, fermions and scalars have a positive contribution, and abelian gauge fields zero.
- $SU(N_c)$ gauge theory with $N_f < 11N_c/2$ flavors is asymptotically free and IR enslaved.

- In particular, QCD is asymptotically free and IR enslaved, reason for confinement at large distances.

- BRST symmetry is a global remnant of gauge invariance (residual gauge symmetry) in a gauge fixed theory.

- It has an anticommuting parameter of mass dimension -1 and ghost number -1, and is defined by $\delta_B A_\mu^a = (D_\mu c)^a \Lambda$.

- The BRST charge Q_B, of ghost number $+1$, is nilpotent on-shell, $Q_B^2 = 0$, and if we introduce the Nakanishi-Lautrup auxiliary field d_a, it is nilpotent off-shell.

- With d_a, the transformation rules are purely kinematical, i.e. independent on the gauge or model, and the Lagrangean is written as the classical piece, plus the variation of a gauge fixing fermion, $\mathcal{L}_{qu} = \mathcal{L}_{class} + \{Q_B, \psi\}$, which trivializes its invariance, since $Q_B^2 = 0$.

Further reading: See chapter 7.6 in [5] and chapter 16.6,16.7 in [3].

Exercises, Lecture 14

1) Calculate explicitly the beta function for an $SO(N_c)$ gauge theory with N_f fundamental flavours.

2) **Anti-BRST invariance**. Is obtained by interchanging b and c in the transformations rules, i.e.

$$\delta_{\bar{B}} A_\mu^a = (D_\mu b)^a \zeta; \quad \delta_{\bar{B}} b^a = \frac{1}{2} g f^a{}_{bc} b^b b^c \zeta \ , \tag{14.47}$$

together with

$$\delta_{\bar{B}} c^a = -d^a \zeta + g f^a{}_{bc} b^b c^c \zeta; \quad \delta_{\bar{B}} d^a = -g f^a{}_{bc} b^b d^c \zeta. \tag{14.48}$$

Verify the nilpotency of $\delta_{\bar{B}}$ and independence of $\delta_B, \delta_{\bar{B}}$, i.e.

$$\delta_B(\Lambda_1)\delta_B(\Lambda_2) = \delta_{\bar{B}}(\zeta_1)\delta_{\bar{B}}(\zeta_2) = 0 = [\delta_B(\Lambda), \delta_{\bar{B}}(\zeta)] = 0. \tag{14.49}$$

[**Note:** A BRST-invariant and anti-BRST-invariant model is the Curci-Ferrari model, which in Minkowski space is

$$\mathcal{L} = \mathcal{L}_{class} - \frac{1}{2\alpha}(\partial_\mu A_\mu^a)^2 + \frac{1}{2} b_a (\partial^\mu D_\mu + D_\mu \partial^\mu) c^a + \frac{g^2 \alpha}{8}(f^a{}_{bc} b^b c^c)^2 + \frac{\alpha}{2}(d^a + \frac{1}{\alpha}\partial_\mu A_\mu^a - \frac{1}{2} g f^a{}_{bc} b^b c^c)^2. \tag{14.50}$$

]

15 Lecture 15. Lee-Zinn-Justin identities and the structure of divergences (formal renormalization of gauge theories)

In this lecture we will put the BRST formalism to good use, and derive from it the general structure of divergences, thus prooving the renormalizability of gauge theories at one-loop. The proof can be extended to all loops.

We start from the observation that the quantum action is invariant under BRST transformations. The integration measure is invariant as well. This statement needs to be tempered a bit. We will discuss anomalies later on in the course, but anomalies can arise from non-invariance of the integration measure. The potential anomaly coming from the Jacobian of gauge transformations is Q_B-exact, i.e.,

$$\text{Anomaly} = \delta_B \Delta S. \tag{15.1}$$

That means that any actual anomaly can be removed by a local finite counterterm ΔS (the coefficient would depend on the regulator). We will not prove it here, and instead we will simply assume the invariance of the measure.

But the source terms are not invariant under BRST. The sources themselves are invariant, meaning that the source terms transform as

$$\delta_B(J\cdot A+\beta_a c^a+b_a\cdot\gamma^a)=(J\cdot Q_BA+\beta_a\cdot Q_Bc^a+(Q_Bb_a)\cdot\gamma^a)\Lambda. \tag{15.2}$$

Then we proceed like in the proof of Ward identities, to which these are related in spirit (they both come from invariance under a -global- symmetry). We make a change of variables in the path integral from the original ones to the BRST-transformed ones. The change leaves the partition function invariant. Since S_{eff} and the measure are also invariant, we obtain

$$0=\int\mathcal{D}A\int\mathcal{D}b\int\mathcal{D}c\{\int d^dx(J\cdot Q_BA+\beta_a\cdot Q_Bc^a+Q_Bb_a\cdot\gamma^a)e^{-S_{eff}[A,b,c]+\int d^dx[J\cdot A+\beta\cdot c+b\cdot\gamma]}\}. \tag{15.3}$$

One could continue like this, but Lee and Zinn-Justin introduced a useful trick, to introduce new sources for Q_BA and Q_Bc, which are *nonlinear* in the fields. Note that Q_Bb is linear in the fields, $=-1/\alpha(\partial^\mu A_\mu)$, so it does not need an extra source. These sources are useful since we have things like

$$\langle\delta_BA^a_\mu\rangle/\Lambda\to\langle gf^a{}_{bc}A^b_\mu c^c\rangle\neq gf^a{}_{bc}\langle A^b_\mu\rangle\langle c^c\rangle\ , \tag{15.4}$$

but we would like to write equations involving products of VEVs arising from single BRST transformations, as we shall see.

Therefore we add to the exponent

$$-S_{\text{extra,source}}=\int d^dx[K^a_\mu(Q_BA)^a_\mu-L^a(Q_Bc)^a]=\int d^dx\left[K^a_\mu(D_\mu c)^a-L^a\frac{g}{2}f^a{}_{bc}c^bc^c\right]. \tag{15.5}$$

This term is still invariant, since the new sources are invariant, $Q_B K_\mu^a = Q_B L^a = 0$, and the rest is $Q_B(...)$, and $Q_B^2 = 0$. These K_μ^a and L^a are extended to objects called "antifields" in the formalism with the same name. Here K_μ^a is anticommuting, whereas L^a is commuting (and $b_a, c^a, \beta^a, \gamma_a$ are anticommuting).

Then (15.3) can be rewritten as usual, with derivatives acting on the partition function $Z = e^{-W}$, as

$$0 = \int d^d x \left[J_\mu^a \cdot \frac{\delta}{\delta K_\mu^a} - \beta_a \frac{\delta}{\delta L^a} - \frac{1}{\alpha}\partial_\mu \left(\frac{\delta}{\delta J_\mu^a} \right) \cdot \gamma^a \right] e^{-W[J,\beta,\gamma,K,L]} \,, \tag{15.6}$$

i.e. as

$$0 = \int d^d x \left[J_\mu^a \cdot \frac{\delta}{\delta K_\mu^a} - \beta_a \frac{\delta}{\delta L^a} - \frac{1}{\alpha}\partial_\mu \left(\frac{\delta}{\delta J_\mu^a} \right) \cdot \gamma^a \right] W[J,\beta,\gamma,K,L]. \tag{15.7}$$

This is the equivalent of a Ward identitiy, or a Dyson-Schwinger identity, since it is an identity on n-point functions derived from the invariance of the path integral.

Like in the Ward or Dyson-Schwinger case, we can make a Legendre transform to the 1PI case,

$$\Gamma[A_\mu^{a,cl}, b_a^{cl}, c^{a,cl}, K_\mu^a, L^a] = W[J_\mu^a, \beta_a, \gamma^a, K_\mu^a, L^a] + \int d^d x (J_\mu^a A_\mu^{a,cl} + \beta_a c^{a,cl} + b_a^{cl}\gamma^a). \tag{15.8}$$

From this Legendre transform we obtain

$$\begin{aligned} &\frac{\delta\Gamma}{\delta A_\mu^{a,cl}} = J_\mu^a; \quad \frac{\delta\Gamma}{\delta c^{a,cl}} = -\beta_a; \quad \frac{\delta\Gamma}{\delta b_a} = \gamma^a \\ &\frac{\delta\Gamma}{\delta K_\mu^a} = \frac{\delta W}{\delta K_\mu^a}; \quad \frac{\delta\Gamma}{\delta L^a} = \frac{\delta W}{\delta L^a}; \quad \frac{\delta W}{\delta J_\mu} = -A_\mu^{cl}. \end{aligned} \tag{15.9}$$

Substituting these relations in (15.7) we obtain

$$\int d^d x \left[\frac{\delta\Gamma}{\delta A_\mu^a}\frac{\delta\Gamma}{\delta K_\mu^a} + \frac{\delta\Gamma}{\delta c^{a,cl}}\frac{\delta\Gamma}{\delta L^a} + \frac{\delta\Gamma}{\delta b_a^{cl}}\frac{1}{\alpha}(\partial_\mu A_\mu^{cl}) \right] \tag{15.10}$$

Another relation is obtained in a similar manner with the first, from the fact that the measure is invariant under a shift in the b_a field, $b_a \to b_a + \delta b_a$. Changing integration variables, using the invariance of the measure and the fact that

$$\delta(\mathcal{L}_{eff} + \mathcal{L}_{\rm source} + \mathcal{L}_{\rm extra,source}) = -\delta b_a[\partial^\mu (D_\mu c)^a + \gamma^a] = -\delta b_a [\partial^\mu (Q_B A_\mu)^a + \gamma^a] \,, \tag{15.11}$$

we get

$$0 = \int \mathcal{D}A_\mu \mathcal{D}b_a \mathcal{D}c^a \int d^d x \delta b_a(x) [\partial^\mu (Q_B A_\mu)^a + \gamma^a] e^{-S_{eff}[A,b,c] + \int d^d x (J\cdot A + \beta\cdot c + b\cdot\gamma)} \tag{15.12}$$

which is true for any $\delta b_a(x)$, so we can write a *local* relation, and writing $Q_B A_\mu^a$ as $\delta/\delta K_\mu^a$ and extracting it outside the path integral, we obtain

$$0 = \left(\partial_\mu \frac{\delta}{\delta K_\mu^a} + \gamma^a \right) e^{-W[J,\beta,\gamma,K,L]} \,, \tag{15.13}$$

giving finally

$$\partial_\mu \frac{\delta}{\delta K_\mu^a} W[J,\beta,\gamma,K,L] = \gamma^a. \tag{15.14}$$

Writing this relation in terms of the effective action Γ as we did above for the first relation (15.10), we obtain

$$\partial_\mu \frac{\delta\Gamma}{\delta K_\mu^a} - \frac{\delta\Gamma}{\delta b_a^{cl}} = 0 \tag{15.15}$$

From the relations (15.10) and (15.15) we will derive the Slavnov-Taylor identities.

But before, we can simplify them a bit by removing from the effective action the classical gauge fixing term, by defining

$$\tilde{\Gamma}[A,b,c,K,L] = \Gamma[A,b,c,K,L] - \frac{1}{2\alpha}\int d^dx (\partial^\mu A_\mu^a)^2. \tag{15.16}$$

Indeed, note that $\int \mathcal{D}\text{fields} e^{-(S+\frac{1}{2\alpha}(\partial^\mu A_\mu^a)^2)} = Z = e^{-W}$ and $\Gamma = W + (...)$, so $\Gamma^{(0)}$ contains $1/2\alpha(\partial^\mu A_\mu^a)^2$.

Then (15.15) is rewrittten in the same way with $\tilde{\Gamma}$ instead of Γ, and (15.10) is rewritten, using also (15.15) without the last term. Together, the relations are now

$$\int d^dx \left[\frac{\delta\tilde{\Gamma}}{\delta A_\mu^a}\frac{\delta\tilde{\Gamma}}{\delta K_\mu^a} + \frac{\delta\tilde{\Gamma}}{\delta c^{a,cl}}\frac{\delta\tilde{\Gamma}}{\delta L^a}\right] = 0 \tag{15.17}$$

$$\partial_\mu \frac{\delta\tilde{\Gamma}}{\delta K_\mu^a} - \frac{\delta\tilde{\Gamma}}{\delta b_a^{cl}} = 0. \tag{15.18}$$

These are called the *Lee-Zinn-Justin identities.*

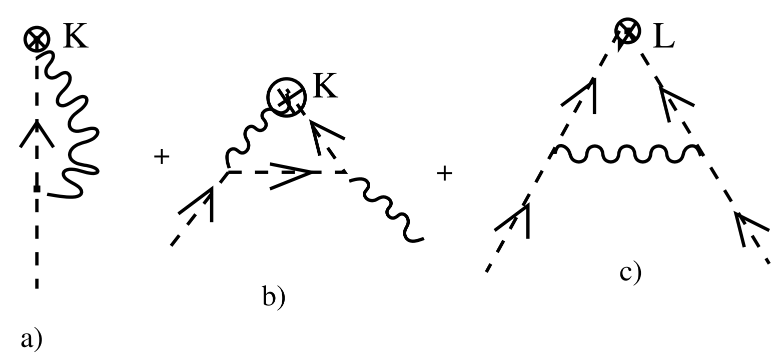

Figure 46: BRST divergences: a) For $K_\mu^a \partial_\mu c^a$. b) For $gf^a{}_{bc}K_\mu^a A_\mu^b c^c$. c) For $gf^a{}_{bc}L^a c^b c^c$.

Structure of divergences

We now use these relations to prove the renormalizability of gauge theories at one-loop (and the proof can be extended by induction to all orders). We will also derive the Slavnov-Taylor identities from the Lee-Zinn-Justin identities.

The new source terms

$$K(Q_B A) - L(Q_B c) = K_\mu^a(\partial_\mu c^a + gf^a{}_{bc}A_\mu^b c^c) + L^a \frac{g}{2} f^a{}_{bc} c^b c^c \ , \tag{15.19}$$

induce 3 new divergences in the quantum effective action, as seen in Fig.46. The sources K_μ^a and L_a are represented by crossed circles with K or L on them.

Then the one-loop divergent graph correcting the $K_\mu^a \partial_\mu c^a$ term comes is a $K_\mu^a A_\mu^b c^c$ vertex with the gluon line ending again on the gluon line. This graph is divergent, since it is the same Feynman integral (with different multiplying factors) as the ghost propagator correction at one-loop (replace the K vertex with just ghost line continuing on).

The one-loop divergent graph correcting the $gf^a{}_{bc}K_\mu^a A_\mu^b c^c$ vertex comes from the vertex itself, with a ghost line exchanged between the ghost and gluon lines, and interchanging them (turning gluon into ghost and vive versa). The graph is divergent, since it is the same Feynman integral (with different multiplying factors) as a ghost-ghost-gluon vertex correction at one loop (replace the K vertex with a ghost line continuing on).

The one-loop divergent graph correcting the $gf^a{}_{bc}L^a c^b c^c$ vertex comes from the vertex itself, with a gluon line interchanged between the two ghost lines. It is divergent, since it is the same Feynman integral as the one diagram replacing the L vertex with a gluon line continuing on, which is just the same graph as above, rotated at 90 degrees.

Therefore at one-loop, we can isolate the divergent part that depends on K_μ^a and L^a as having these 3 possible divergent structures, obtaining

$$\begin{aligned}\tilde{\Gamma}_{div}[A^{cl}, b^{cl}, c^{cl}, K, L] &= \tilde{\Gamma}_{div}[A^{cl}, b^{cl}, c^{cl}] - (\tilde{Z}_3 - 1)\int d^d x K_\mu^a \partial^\mu c^a \\ &\quad -(\tilde{Z}_1 - 1)\int d^d x g f_{abc} K_\mu^a A_\mu^b c^c + (X-1)\int d^d x L_a \left(-\frac{g}{2} f^a{}_{bc} c^b c^c\right) \ ,\end{aligned} \tag{15.20}$$

where we have written 3 renormalization factors, $\tilde{Z}_3, \tilde{Z}_1$ and X. The first two, we have anticipated a bit that they will be the same as the previously defined $\tilde{Z}_3$ and $\tilde{Z}_1$, which we will prove shortly. Otherwise, we could have called them $\tilde{Y}_2, \tilde{Y}_1$ and show that they equal $\tilde{Z}_3, \tilde{Z}_1$.

We can now use (15.17) and (15.18) to determine $\tilde{\Gamma}_{div}$, by expanding in loop order,

$$\tilde{\Gamma} = \tilde{\Gamma}^{(0)} + \hbar \tilde{\Gamma}_{div}^{(1)} + \mathcal{O}(\hbar^2) \ , \tag{15.21}$$

and using

$$\tilde{\Gamma}^{(0)} = \int d^d x \left[\frac{1}{4}(F_{\mu\nu}^a)^2 - b_a \partial^\mu D_\mu c^a - K_\mu^a Q_B A_\mu^a + L_a Q_B c^a\right] . \tag{15.22}$$

Replacing this in (15.18) and concentrating on the divergent part (to one-loop), we obtain

$$\frac{\delta \tilde{\Gamma}}{\delta b_a} = \partial_\mu \frac{\delta \tilde{\Gamma}}{\delta K_\mu^a} = -(\tilde{Z}_3 - 1)\partial^\mu \partial_\mu c^a - (\tilde{Z}_1 - 1) g f^a{}_{bc} \partial^\mu (A_\mu^b c^c) \ , \tag{15.23}$$

where in the second equality we have substituted the form of $\tilde{\Gamma}_{div}$, whose K_a dependence was fixed. We can now integrate this relation with respect to b_a and obtain

$$\tilde{\Gamma}_{div}[A,b,c] = \tilde{\Gamma}_{div}[A] - \int d^d x \left[(\tilde{Z}_3 - 1) b_a \partial^\mu \partial_\mu c^a + (\tilde{Z}_1 - 1) b_a g f^a{}_{bc} \partial^\mu (A^b_\mu c^c)\right]. \quad (15.24)$$

Thus we see that indeed $\tilde{Z}_3$ and $\tilde{Z}_1$ are the objects we have defined before, the renormaliation of the ghost propagator and ghost-ghost-gluon vertex, respectively.

We note that we have, in two short steps, fixed the dependence of the divergent piece on all fields except A^a_μ.

To fix that as well, we look at the simple pole (single divergence) of (15.17). Indeed, there we have also a double divergence, coming from the term where both $\tilde{\Gamma}$ factors are divergent, but we want to focus instead on the terms where one $\tilde{\Gamma}$ is divergent and one is finite, so we have at one-loop,

$$0 = \int d^d x \left[\left(\frac{\delta \tilde{\Gamma}_{div}}{\delta A^a_\mu}\frac{\delta \tilde{\Gamma}^{(0)}}{\delta K^a_\mu} + \frac{\delta \tilde{\Gamma}^{(0)}}{\delta A^a_\mu}\frac{\delta \tilde{\Gamma}^{div}}{\delta K^a_\mu}\right) + \left(\frac{\delta \tilde{\Gamma}^{div}}{\delta c^a}\frac{\delta \tilde{\Gamma}^{(0)}}{\delta L^a} + \frac{\delta \tilde{\Gamma}^{(0)}}{\delta c^a}\frac{\delta \tilde{\Gamma}_{div}}{\delta L^a}\right)\right]. \quad (15.25)$$

Now we must insert in this equation the expressions for $\tilde{\Gamma}_{div}$ and $\tilde{\Gamma}^{(0)}$ and identify terms. We will organize the calculation according to the types of terms in the expanded equation.

1. Terms linear in K^a_μ.

We can check that the terms coming from the second and 3rd terms in (15.25) cancel, leaving the contributions of the first and 4th terms, giving

$$0 = \int d^d x [(\tilde{Z}_1 - 1) K^a_\mu g f^a{}_{bc} c^c] D_\mu c^b - \int d^d x (X-1) \left[(-D_\mu K^a_\mu)\left(-\frac{g}{2} f^a{}_{bc} c^b c^c\right)\right], \quad (15.26)$$

which after a partial integration of D_μ and a rearranging of the terms gives

$$X = \tilde{Z}_1. \quad (15.27)$$

2. Terms linear in A and not containing K and L, and linear in c.

We write an expansion in A of $\tilde{\Gamma}_{div}[A]$,

$$\tilde{\Gamma}_{div}[A] = \tilde{\Gamma}^{(2)}_{div}[A] + \tilde{\Gamma}^{(3)}_{div}[A] + \tilde{\Gamma}^{(4)}_{div}[A]. \quad (15.28)$$

Then of course, the contribution to these terms from $\tilde{\Gamma}_{div}[A]$ is entirely from $\tilde{\Gamma}^{(2)}_{div}[A]$, to obtain linear terms in A after derivation. The variation of $1/4(F^a_{\mu\nu})^2$ with respect to A^a_ν gives $-D^\mu F^a_{\mu\nu}$, as we know (from the YM equation of motion). We obtain, from the first two terms in (15.25),

$$0 = \int d^d x \left\{-\frac{\delta \tilde{\Gamma}^{(2)}_{div}[A]}{\delta A^a_\mu}(D_\mu c^a) + (D^\mu F^a_{\mu\nu})_{\text{lin.A}}[(\tilde{Z}_3 - 1)\partial_\nu c^a + (\tilde{Z}_1 - 1) g f^a{}_{bc} A^b_\nu c^c]|_{\text{no A}}\right\}. \quad (15.29)$$

But note that by partial integration of the $(\tilde{Z}_3-1)$ term we obtain $\partial^\nu D^\mu F_{\mu\nu}|_{\text{lin.A}} = \partial^\mu\partial^\nu(\partial_\mu A_\nu - \partial_\nu A_\mu) = 0$, and the $(\tilde{Z}_1 - 1)$ term has too many A_μ's. Therefore we obtain

$$\int d^d x \frac{\delta\tilde{\Gamma}^{(2)}_{div}[A]}{\delta A^a_\mu}\partial_\mu c^a = 0, \tag{15.30}$$

which by partial integration gives

$$\partial_\mu \frac{\delta\Gamma^{(2)}_{div}}{\delta A^a_\mu} = 0 \ , \tag{15.31}$$

which means that $\Gamma^{(2)}_{div}[A]$ is transverse, which uniquely selects the form of the kinetic term, i.e. we must have

$$\tilde{\Gamma}^{(2)}_{div}[A] = (Z_3 - 1)\int d^d x \frac{1}{2}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2. \tag{15.32}$$

3. Terms quadratic in A and not containing K and L.

We continue with the expansion in A of the same terms as above, giving

$$\begin{aligned} 0 = \int d^d x \Bigg\{ &-\frac{\delta\tilde{\Gamma}^{(3)}}{\delta A^a_\mu}\partial_\mu c^a - \frac{\delta\tilde{\Gamma}^{(2)}}{\delta A^a_\mu} g f^a{}_{bc} A^b_\mu c^c + \partial^\mu(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)(\tilde{Z}_1 - 1) g f^a{}_{bc} A^b_\nu c^c \\ &+ g f^a{}_{bc} A^b_\mu(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu)(\tilde{Z}_3 - 1)\partial_\nu c^a \Bigg\} \ , \end{aligned} \tag{15.33}$$

which can be rewritten as

$$\int d^d x \left[\frac{\delta\tilde{\Gamma}^{(3)}}{\delta A^a_\mu} + (Z_3 - \tilde{Z}_1 - \tilde{Z}_3 - 1) g f^a{}_{bc}(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu) A^c_\nu\right]\partial_\mu c^a = 0. \tag{15.34}$$

Since this relation is valid for any c^a, we can put the [] bracket to zero and integrate it, to give

$$\tilde{\Gamma}^{(3)}_{div}[A] = (Z_3 + \tilde{Z}_1 - \tilde{Z}_3 - 1)\int d^d x \frac{1}{2} g f_{abc} A^a_\mu A^b_\nu(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu). \tag{15.35}$$

This is again of the type of the classical 3-gluon coupling, like in the explicit one-loop renormalization. The coefficient was usually called Z_1 at one-loop, so

$$Z_1^{(1)} = Z_3^{(1)} + \tilde{Z}_1^{(1)} - \tilde{Z}_3^{(1)} = Z_3\frac{\tilde{Z}_1}{\tilde{Z}_3} + \mathcal{O}(\hbar^2). \tag{15.36}$$

4. Terms cubic in A and not containing K and L.

We will not do this explicitly, but the analysis is similar, and the result is

$$\tilde{\Gamma}^{(4)}_{div}[A] = (Z_3 + 2\tilde{Z}_1 - 2\tilde{Z}_3 - 1)\frac{g^2}{4} f^a{}_{bc} f_{ade}\int d^d x A^b_\mu A^c_\nu A^d_\mu A^e_\nu. \tag{15.37}$$

Yet again this is of the type of the classical 4-gluon coupling, like in the explicit one-loop renormalization. The coefficient was usually called Z_4 at one-loop, so

$$Z_4^{(1)} = Z_3^{(1)} + 2Z_1^{(1)} - 2\tilde{Z}_3^{(1)} = Z_3\frac{\tilde{Z}_1^2}{\tilde{Z}_3^2} + \mathcal{O}(\hbar^2). \tag{15.38}$$

Then we see that (15.36) and (15.38) are the two Slavnov-Taylor identities at one-loop, as promised.

In conclusion, from the Lee-Zinn-Justin identities (coming from BRST invariance), we have completely fixed the form of the divergent part of the quantum effective action, and derived the Slavnov-Taylor identities.

The extra source terms are

$$\begin{aligned} K(Q_B A) &= K^a_\mu(\tilde{Z}_3\partial_\mu c^a + \tilde{Z}_1 g f^a{}_{bc} A^b_\mu c^c) \\ L(Q_B c) &= L^a X\left(\frac{g}{2} f^a{}_{bc} c^b c^c\right). \end{aligned} \tag{15.39}$$

Using the renormalizations defined at one-loop,

$$b_0, c_0 = \sqrt{\tilde{Z}_3} b, c; \quad A_0 = \sqrt{Z_3} A; \quad g_0 = g\frac{Z_1}{Z_3^{3/2}}\ , \tag{15.40}$$

we can write

$$\begin{aligned} K(Q_B A) &= K^a_\mu \sqrt{\tilde{Z}_3}(\partial_\mu c^a_0 + g_0 f^a{}_{bc} A^b_{0\mu} c^c_0) = \sqrt{\tilde{Z}_3} K(Q_B A_{0\mu}) \\ L(Q_B c) &= \sqrt{Z_3} L^a\left(\frac{g_0}{2} f^a{}_{bc} c^b_0 c^c_0\right) = \sqrt{Z_3} L(Q_B c_0)\ , \end{aligned} \tag{15.41}$$

therefore we deduce the renormalization of the sources

$$K^a_{0\mu} = \sqrt{\tilde{Z}_3} K^a_\mu; \quad L^a_0 = \sqrt{Z_3} L^a. \tag{15.42}$$

In conclusion, we have the renormalization at one-loop

$$\Gamma^{(1-loop)}[A, b, c, K, L, g, \alpha] = \Gamma_0^{(1-loop)}[A_0, b_0, c_0, K_0, L_0, g_0, \alpha_0]\ , \tag{15.43}$$

which shows that the gauge theory is one-loop renormalizable (without calculating explicitly anything).

This proof can be extended at any n-loops through induction, though we will not do it here.

One can also include fermions to the analysis. The proof of the Lee-Zinn-Justin identities in the case with fermions is left as exercise below, as is the proof of the Slavnov-Taylor identities.

Important concepts to remember

- The quantum effective action is BRST invariant, and the measure is invariant up to a possible anomaly that can be removed by a local finite counterterm.
- We derive the Lee-Zinn-Justin identities as quadratic identities on $\tilde{\Gamma}$, by adding source terms for the nonlinear terms in the BRST transformations, $K^a_\mu(Q_B A)^a_\mu$ and $-L^a(Q_B c)^a$, besides the usual source terms $\beta_a c^a + b_a \gamma^a$, and finding the equations in a similar manner to the Dyson-Schwinger identities or Ward identities.

- The introduction of K and L generates 3 new divergent structures, but the Lee-Zinn-Justin identities allow to completely fix the possible structure of divergences to be exactly the one in the bare Lagrangean, thus proving renormalizability at one-loop.
- The proof can be generalized by induction to all loop orders, as well as to include fermions, thus a general gauge theory is renormalizable.
- The renormalization of K and L is not independent (is written in terms of the other Z factors), and the Slavnov-Taylor identities are also obtained from the Lee-Zinn-Justin identities.

Further reading: See chapter 7.7.1 and 7.7.2 in [5] and chapter 16.4 in [3].

Exercises, Lecture 15

1) Introduce fermions into the gauge theory,

$$\delta_B \psi_i = -(T^a)_{ij} c^a \psi_j \Lambda \ , \tag{15.44}$$

and add to the action the fermi terms

$$S_{\text{fermi}} = \int d^d x \bar{\psi}(\not{D} + m)\psi[-\bar{H}_{i\alpha}(Q_B \psi)_{i\alpha} - (Q_B \bar{\psi})_{i\alpha} H_{i\alpha}] \tag{15.45}$$

and the fermion source terms

$$\int d^d x [\bar{\zeta}_{i\alpha}\psi_{i\alpha} + \bar{\psi}_{i\alpha}\zeta_{i\alpha} + \bar{H}_{i\alpha}(Q_B\psi)_{i\alpha} + (Q_B\bar{\psi})_{i\alpha}H_{i\alpha}]. \tag{15.46}$$

Find the generalized Lee-Zinn-Justin identities, first for $W[J, \beta, \gamma, K, L, \zeta, \bar{\zeta}, H, \bar{H}]$ and then for Γ and $\tilde{\Gamma}[A, b, c, \psi, \bar{\psi}, K, L, H, \bar{H}]$.

2) Check that we get the correct Slavnov-Taylor identities at one-loop.

16 Lecture 16. BRST quantization

We obtained BRST symmetry as a symmetry of gauge theories, and we saw that we could use it to get a proper understanding of the quantum theory, since it allowed a formal proof of renormalizability.

But more generally, we can turn it into a *quantization procedure* for any theory with a local invariance (like gauge invariance) that needs gauge fixing. *BRST quantization* builds upon the Dirac formalism for quantization of constrained systems, and is extended to the Batalin-Vilkovisky (BV, or field-antifield) formalism, that will not be treated here, but only mentioned at the end of the lecture.

So we have: Dirac quantization extends to BRST quantization, that extends to the BV formalism.

Review of the Dirac formalism

We will take a quick review of the salient points of the Dirac formalism, since we will build upon it to develop BRST quantization.

One starts with a system obeying a set of *primary constraints*

$$\phi_m(q,p) = 0; \quad m = 1, ..., M. \tag{16.1}$$

If a quantity is equal to zero only by using the constraints, we say we have a *weak equality*, and write ≈ 0, and we mean that we do all derivations, and only at the end of the calculation we put $\phi_m = 0$.

Then for some function of phase space, $g(q,p)$, the time evolution will be given in weak equality by the Poisson bracket with the *total Hamiltonian* H_T,

$$H_T = H + u_m \phi_m \ , \tag{16.2}$$

where $H = H_0$ is the classical Hamiltonian (without the constraints) i.e.,

$$\dot{g} \approx \{g, H_T\}_{P.B.}. \tag{16.3}$$

Since the time evolution must preserve the constraints, we must have $\dot{\phi}_m = 0$, which implies that we need to have the weak equality

$$\{\phi_m, H\}_{P.B.} + u_n \{\phi_m, \phi_n\}_{P.B.} \approx 0. \tag{16.4}$$

These will in general generate *secondary constraints* (we repeat the procedure of finding new constraints, then imposing that their time evolution is zero, until we don't get any more new constraints) ϕ_k, $k = M+1, ..., M+K$. The set of all constraints will be denoted by

$$\phi_j \approx 0, \quad j = 1, ..., M+K. \tag{16.5}$$

We next define a *first class function* of phase space $R(q,p)$ if its Poisson bracket with all the constraints is weakly zero, i.e.

$$\{R, \phi_j\}_{P.B.} \approx 0 \ \ (= r_{jj'} \phi_{j'}). \tag{16.6}$$

Of course, a particular case is of first class constraints ϕ_a, that must therefore satisfy a kind of algebra, $\{\phi_a, \phi_j\} = f_{aj}{}^{j'}\phi_{j'}$.

A function of phase space is second class if there is at least one j for which $\{R, \phi_j\}$ is not weakly zero.

Since the set of ϕ_j is the full set of constraints obtained from time evolution, it follows by definition that for all of them we need to have weakly zero Poisson brackets with H_T,

$$\{\phi_j, H_T\} \approx \{\phi_j, H\} + u_m\{\phi_j, \phi_m\} \approx 0. \tag{16.7}$$

This is now thought of as a set of equations for the coefficients u_m, solved by a particular solution U_m of the inhomogeneous equation (the equation above), plus linear solutions of the homogenous equation with arbitrary (numerical) coefficients, i.e.

$$u_m = U_m + v_a V_{am} \ , \tag{16.8}$$

where

$$V_{am}\{\phi_j, \phi_m\} \approx 0. \tag{16.9}$$

Then we can rewrite the total Hamiltonian

$$H_T = H + U_m\phi_m + v_a V_{am}\phi_m \ , \tag{16.10}$$

where

$$\begin{aligned} \phi_a &= v_{am}\phi_m \\ H' &= H + U_m\phi_m. \end{aligned} \tag{16.11}$$

Note that by definition, since $V_{am}\{\phi_m, \phi_j\}_{P.B.} \approx 0$, $V_{am}\phi_m = \phi_a$ are first class constraints. Also, since U_m is a particular solution of the inhomogenous equation, we have $\{\phi_j, H'\}_{P.B.} \approx 0$, so H' is also first class.

But ϕ_a are only independent first class primary constraints, and there is nothing special about primary constraints. So we define the *extended Hamiltonian* by adding also the first class secondary constraints, so as to have all the first class constraints in it,

$$H_E = H_T + v_{a'}\phi_{a'}. \tag{16.12}$$

Dirac brackets

To quantize, we need to define brackets that replace the Poisson brackets, to deal with the second class constraints.

Consider χ_s the *independent* second class constraints, and define the matrix $c_{ss'}$ as the inverse of the bracket of constraints

$$c_{ss'}\{\chi_{s'}, \chi_{s''}\} = \delta_{ss''}. \tag{16.13}$$

Then the Dirac bracket is defined as

$$[f, g]_{D.B.} = \{f, g\}_{P.B.} - \{f, \chi_s\}_{P.B.} c_{ss'}\{\chi_{s'}, g\}_{P.B.}. \tag{16.14}$$

Then the time evolution using the Dirac brackets is the same as the one using the Poisson brackets,

$$[g, H_T]_{D.B.} \approx \{g, H_T\}_{P.B.} \approx \dot{g} \ , \tag{16.15}$$

but now the Dirac brackets of any function of phase space with the second class constraints is *strongly* (identically) zero,

$$[f, \chi_{s''}]_{D.B.} = \{f, \chi_{s''}\}_{P.B.} - \{f, \chi_s\}_{P.B.} c_{ss'} \{\chi_{s'}, \chi_{s''}\}_{P.B.} = 0. \tag{16.16}$$

That means that we can impose the second class constraints operatorially on states in order to quantize,

$$\hat{\chi}_s |\psi\rangle = 0 \ , \tag{16.17}$$

and the Dirac brackets are quantized to the commutators, instead of the Poisson brackets.

BRST quantization

We saw in the Lagrangean formalism that the BRST quantum action is writen as

$$\mathcal{L}_{qu} = \mathcal{L}_{class} + Q_B \Psi \ , \tag{16.18}$$

where Ψ is the gauge fixing fermion. More precisely, since Q_B is an operator, we really have

$$\mathcal{L}_{qu} = \mathcal{L}_{class} + \{Q_B, \Psi\} \ . \tag{16.19}$$

BRST quantization is a Hamiltonian formalism, since it is based on the Dirac formalism. As such, we must use Minkowski signature. In Minkowski signature,

$$\Psi = -b_a \left(F^a + \frac{\alpha}{2} d^a \right) . \tag{16.20}$$

In a Hamiltonian formalism, one has states $|\psi\rangle$, for instance intial and final states $|i\rangle$ and $|f\rangle$, and observables are transition amplitudes $\langle f|i\rangle$. As such, they should be independent of the choice of gauge, thus independent on the change in gauge fixing function F^a. Then, under a small change δF^a, such that $\delta\Psi = b_a \delta F^a$, the variation of the transition amplitude, written as a path integral,

$$\langle f|i\rangle = \int \mathcal{D}\text{fields} e^{iS} \ , \tag{16.21}$$

is given by

$$\delta \langle f|i\rangle = \int \mathcal{D}\text{fields} e^{iS} i\delta S = i\langle f|\{Q_B, \delta\Psi\}|i\rangle \ , \tag{16.22}$$

and it must be zero. Assuming that Q_B is hermitean, $Q_B^\dagger = Q_B$, putting the two terms in the anticommutator term above to zero gives $\langle f|Q_B = 0$ and $Q_B|i\rangle = 0$, so that $Q_B|f\rangle = Q_B|i\rangle = 0$. That means that *all physical states must be BRST-invariant*,

$$Q_B|\psi\rangle = 0. \tag{16.23}$$

We say that physical states are Q_B-*closed*. On the other hand, physical states that differ by Q_B-*exact* terms, i.e. terms that are Q_B of something else, are equivalent

$$|\psi'\rangle = |\psi\rangle + Q_B|\chi\rangle \sim |\psi\rangle. \tag{16.24}$$

Note that then

$$Q_B|\psi'\rangle = Q_B|\psi\rangle + Q_B^2|\chi\rangle = 0\ , \tag{16.25}$$

so it is also physical. The equivalence means that matrix elements are identical. Consider the matrix element with another physical state $|\tilde{\psi}\rangle$, $(Q_B|\tilde{\psi}\rangle = 0)$

$$\langle\psi'|\tilde{\psi}\rangle = \langle\psi|\tilde{\psi}\rangle + \langle\chi|Q_B|\tilde{\psi}\rangle = \langle\psi|\tilde{\psi}\rangle. \tag{16.26}$$

These matrix elements must also be preserved under time evolution, which means that

$$\langle\psi'|H_{BRST}|\tilde{\psi}\rangle = \langle\psi|H_{BRST}|\tilde{\psi}\rangle + \langle\chi|H_{BRST}Q_B|\tilde{\psi}\rangle + \langle\chi|[Q_B, H_{BRST}]|\tilde{\psi}\rangle\ , \tag{16.27}$$

and this must be equal to $\langle\psi|H_{BRST}|\tilde{\psi}\rangle$. The middle term is zero since $Q_B|\tilde{\psi}\rangle = 0$, which means that we need to have (since the states $|\chi\rangle$ and $|\tilde{\psi}\rangle$ are arbitrary)

$$[Q_B, H_{BRST}] = 0. \tag{16.28}$$

In order to define the BRST Hamiltonian H_{BRST}, we must consider the Hamiltonian a la Dirac, in the presence of constraints. But moreover, we need to define Q_B and H_{BRST} together, such that $Q_B^2 = 0$ and $[Q_B, H_{BRST}] = 0$.

The quantization is done by defining a Hilbert space of states. From the above considerations, physical states are states in the Q_B-*cohomology*, since we need to have Q_B-exact states, $Q_B|\psi\rangle = 0$, modulo Q_B-exact states, i.e. $|\psi\rangle \sim |\psi\rangle + Q_B|\chi\rangle$.

In general, the cohomology of an operator that squares to zero is the coset defined as closed states, modulo exact states (equivalence classes under the equivalence by exact states),

$$\mathrm{Q_B - Cohomology} = \frac{Q_B - \text{closed}}{Q_B - \text{exact}}\ , \tag{16.29}$$

or that the Hilbert space is

$$\mathcal{H}_{BRST} = \frac{\mathcal{H}_{closed}}{\mathcal{H}_{exact}}. \tag{16.30}$$

The notion of cohomology should be familiar (except for the mathematical language), since the cohomology of the exterior derivative $d = dx^\mu\partial_\mu$ is nothing but the familiar one of physical states in electromagnetism. We have $d^2 = 0$, and pure gauge fields (in the absence of matter) satisfy $F = dA = 0$, i.e. are d-closed, and equivalent gauge field configurations differ by d-exact terms, since gauge invariance is $A \sim A + d\lambda$. Then the cohomology of d over 1-forms is the cohomology of gauge fields, i.e. the physical states for fields of vanishing field strength. The cohomology of d is naturally extended for p-forms instead of 1-forms, giving the p-th cohomology of d.

Example of BRST quantization: Electromagnetism in Lorenz gauge.

At this intermediate stage, we give a simple example for the BRST quantization formalism described so far, for abelian gauge fields, i.e. electromagnetism. The example is kind of trivial, but will show explicitly how some of the abstract ideas described above work.

The BRST transformations for the abelian gauge fields (so $f^a{}_{bc} = 0$) in Minkowski signature are

$$\delta_B A_\mu = \partial_\mu c\Lambda; \quad \delta_B b = \frac{1}{\alpha}\partial^\mu A_\mu; \quad \delta_B c = 0. \tag{16.31}$$

Consider now canonical quantization of the fields A_μ, b, c, without taking into account the gauge invariance and constraints from gauge fixing

$$\begin{aligned}
A_\mu(x) &= \int \frac{d^3p}{(2\pi)^{3/2}}\frac{1}{\sqrt{2p^0}}\left[a_\mu(\vec{p})e^{ip\cdot x}+a_\mu^\dagger(\vec{p})e^{-ip\cdot x}\right]\\
c(x) &= \int \frac{d^3p}{(2\pi)^{3/2}}\frac{1}{\sqrt{2p^0}}\left[c(\vec{p})e^{ip\cdot x}+c^\dagger(\vec{p})e^{-ip\cdot x}\right]\\
b(x) &= \int \frac{d^3p}{(2\pi)^{3/2}}\frac{1}{\sqrt{2p^0}}\left[b(\vec{p})e^{ip\cdot x}+b^\dagger(\vec{p})e^{-ip\cdot x}\right]. \qquad (16.32)
\end{aligned}$$

The BRST transformation acts by (anti)commuting with Q_B, so:

1.$\delta_B A_\mu = i[Q_B, A_\mu]\Lambda$ is compared with $\delta_B A_\mu = \partial_\mu c\Lambda$. Substituting the forms of $A_\mu(x)$ and $c(x)$, we find

$$[Q_B, a_\mu(\vec{p})] = +p_\mu c(\vec{p}); \quad [Q_B, a_\mu^\dagger(\vec{p})] = -p_\mu c^\dagger(\vec{p}). \qquad (16.33)$$

2.$\delta_B b = i\{Q_B, b\}\Lambda$ is compared with $\delta_B b = 1/\alpha\partial^\mu A_\mu$. Substituting the forms of $b(x)$ and $A_\mu(x)$, we find

$$\{Q_B, b(\vec{p})\} = \frac{1}{\alpha}p^\mu a_\mu(\vec{p}); \quad \{Q_B, b^\dagger(\vec{p})\} = -\frac{1}{\alpha}p^\mu a_\mu^\dagger(\vec{p}). \qquad (16.34)$$

3.$\delta c = -\{Q_B, c\}\Lambda$ is compared with $\delta_B c = 0$ to obtain

$$\{Q_B, c(\vec{p})\} = \{Q_B, c^\dagger(\vec{p})\} = 0. \qquad (16.35)$$

We now construct the Hilbert space of physical states. They must satisfy $Q_B|\psi\rangle = 0$. Consider the state

$$|\psi\rangle = e^\mu a_\mu^\dagger(\vec{p})|\tilde{\psi}\rangle\ , \qquad (16.36)$$

where $|\tilde{\psi}\rangle$ is a physical state, i.e. $Q_B|\tilde{\psi}\rangle = 0$. It satisfies

$$Q_B|\psi\rangle = e^\mu a_\mu^\dagger(\vec{p})Q_B|\tilde{\psi}\rangle - e^\mu p_\mu c^\dagger(\vec{p})|\tilde{\psi}\rangle = -e^\mu p_\mu c^\dagger(\vec{p})|\tilde{\psi}\rangle. \qquad (16.37)$$

Therefore $Q_B|\psi\rangle = 0 \Leftrightarrow e^\mu p_\mu = 0$, so the state with one extra photon on top of $|\tilde{\psi}\rangle$ needs to be transverse.

On the other hand, consider the state

$$Q_B b^\dagger(\vec{p})|\tilde{\psi}\rangle = -b^\dagger(\vec{p})Q_B|\tilde{\psi}\rangle - \frac{1}{\alpha}p^\mu a_\mu^\dagger(\vec{p})|\tilde{\psi}\rangle = -\frac{1}{\alpha}p^\mu a_\mu^\dagger(\vec{p})|\tilde{\psi}\rangle. \qquad (16.38)$$

It follows that

$$|\psi(e_\mu + \beta p_\mu)\rangle = |\psi(e_\mu)\rangle + \beta p^\mu a_\mu^\dagger(\vec{p})|\tilde{\psi}\rangle = |\psi(e_\mu)\rangle - \alpha\beta Q_B b^\dagger(\vec{p})|\tilde{\psi}\rangle\ , \qquad (16.39)$$

so we have equivalent states, $e_\mu \sim e_\mu + \beta p_\mu$, as it should, for massless states $p^\mu p_\mu = 0$.

On the other hand, we also learn that

$$Q_B b^\dagger(\vec{p})|\tilde{\psi}\rangle = -\frac{1}{\alpha}p^\mu a_\mu^\dagger(\vec{p})|\tilde{\psi}\rangle \neq 0\ , \qquad (16.40)$$

so there are no states with ghosts in it, i.e. with $b^\dagger(\vec{p})$, among the physical states, since any such $b^\dagger(\vec{p})$ added to a physical state turns it unphysical (Q_B on it is nonzero).

Finally, we see that

$$c^\dagger(\vec{p})|\psi\rangle = Q_B \left(\frac{-e^\mu a_\mu^\dagger(\vec{p})}{e\cdot p}\right)|\psi\rangle \ , \tag{16.41}$$

so it is equivalent with zero (the relation above is proven by commuting Q_B on the right hand side with $a_\mu^\dagger(\vec{p})$ and using $Q_B|\psi\rangle = 0$).

So there are no b ghosts in physical states, and c ghosts are equivalent with zero. Therefore the physical Hilbert space is composed only of transverse photons, as it should.

General formalism.

We now turn to the general formalism for BRST quantization.

As we mentioned, we need to define a Q_B and a H_{BRST} that satisfy $[Q_B, H_{BRST}] = 0$ and $Q_B^2 = 0$ and then we construct the Hilbert space as the states of the Q_B-cohomology. Until now $Q_B^2 = 0$ was assumed, but it is actually needed. Indeed, we want that the gauge choice does not change the commutation relation $[Q_B, H_{BRST}] = 0$, so any δH_{BRST} induced by the gauge choice should preserve it, $[Q_B, \delta H_{BRST}] = 0$. But we saw that the gauge choice in the action, and therefore in H_{BRST}, comes from the gauge fixing fermion term, $\delta H_{BRST} = -\{Q_B, b_A\delta F^A\}$. Then we have

$$0 = [Q_B, \{Q_B, b_A\delta F^A\}] = [Q_B^2, b_A\delta F^A] \Rightarrow Q_B^2 = 0. \tag{16.42}$$

Note that there is the Fradkin-Vilkovitsky theorem, that the partition function

$$Z_\Psi = \int \mathcal{D}(...)e^{iS_{qu}} \tag{16.43}$$

is Ψ-independent.

So we need to construct Q_B and H_{BRST}, which will be done by satisfying the above conditions order by order in the number of fields, but we need to start somewhere, i.e. we must define first order Q_B and H_{BRST} in the fields. We turn to that next.

Quantum action

We start by writing the quantum action, as

$$S_{qu} = \int dt \left[\dot{q}_i p^i + \dot{\lambda}^\mu \pi_\mu + \dot{\eta}^a p_a - H_{BRST} + \{\psi, Q_B\}\right]. \tag{16.44}$$

Here λ^μ are Lagrange multipliers, and π_μ their conjugate momenta, which are also in general first class constraints, so the total set of first class constraints is written as $G^a = (\phi_\alpha, \pi^\mu)$, satisfying the algebra

$$[G_a, G_b] = f_{ab}{}^c G_c. \tag{16.45}$$

We also define the ghost phase space by

$$\eta^a = (c^\alpha, -p(b)_\alpha); \quad p_a = (p(c)_\alpha, -b_\alpha). \tag{16.46}$$

Then defining Poisson brackets

$$\{c^\alpha, p(c)_\alpha\} = -\delta^\alpha_\beta; \quad \{b_\alpha, p(b)^b\} = -\delta^\beta_\alpha \tag{16.47}$$

imply that together

$$\{\eta^a, p_b\} = -\delta^a_b. \tag{16.48}$$

Observation: in 2 dimensions, $p(c) = b$ and $p(b) = c$, so the story is simpler.

Then in general, the first terms in the expansion of Q_B is

$$Q_B = \eta^a G_a - \frac{1}{2}\eta^c \eta^b f_{bc}{}^a p_a (-)^c. \tag{16.49}$$

If $f_{bc}{}^a$ are constants (note that the algebra of constraints is an algebra of functions, so in general the structure "constants" can be fields, not actual constants), then the above implies $Q_B^2 = 0$ from the Jacobi identities, which is left as an exercise to prove, and the Hamiltonian in that order is

$$H_{BRST} = H_0 + ... \tag{16.50}$$

In general however, we have

$$Q_B = \eta^{b_1}...\eta^{b_{n+1}} U^{(n)}_{b_1...b_{n+1}}{}^{a_1...a_n} p_{a_1}...p_{a_n} \tag{16.51}$$

and

$$H_{BRST} = H_0 + \eta^a V_a^b p_b + ... \tag{16.52}$$

Here the algebra of first class constraints

$$\{\phi_\alpha, \phi_b\} = f_{\alpha\beta}{}^\gamma \phi_\gamma; \quad \{H_0, \phi_\alpha\} = V_\alpha^\beta \phi_\beta \tag{16.53}$$

is generalized (extended) to

$$\{G_a, G_b\} = f_{ab}{}^c G_c; \quad \{H_0, G_a\} = V_a^b G_b. \tag{16.54}$$

The V_a^b define the first correction to the classical Hamiltonian H_0, and the $f_{ab}{}^c$ define the second term in Q_B, since as we saw above, we have

$$U_a^{(0)} = G_a; \quad U^{(1)}_{b_1 b_2}{}^{a_1} = -\frac{1}{2} f_{b_1 b_2}{}^{a_1} (-)^{b_2}. \tag{16.55}$$

Finally, note that all the general BRST formalism described until now is *classical*, quantum mechanics did not enter anywhere until now. Quantization is done by replacing the Dirac brackets with the commutator, and finding the physical Hilbert space as the space of states in the Q_B-cohomology, like we saw in the example of electromagnetism.

Example: Pure Yang-Mills.

We now turn to a more nontrivial example to explain the general BRST formalism (though not the quantization, i.e. finding the Q_B-cohomology, which is similar to the electromagnetism case), the case of (nonabelian) Yang-Mills theory.

The action of pure Yang-Mills in Minkowski space can be written as

$$S = \int d^4x \left[-\frac{1}{4}(F^a_{\mu\nu})^2\right] = \int dtd^3x \left[-\frac{1}{4}(F^a_{ij})^2 + \frac{1}{2}(\dot{A}^a_i - D_i A^a_0)^2\right] , \tag{16.56}$$

since

$$E^a_i = F^a_{0i} = \partial_0 A^a_i - \partial_i A^a_0 + g f^a{}_{bc} A^b_0 A^c_i = \dot{A}^a_i - D_i A^a_0. \tag{16.57}$$

It follows that the momentum conjugate to A^a_i is

$$(\pi^a_i \equiv) p^a_i = \dot{A}^a_i - D_i A^a_0 = F^a_{0i} = E^a_i. \tag{16.58}$$

We also define as usual

$$B^a_i = \frac{1}{2}\epsilon_{ijk} F^a_{jk}. \tag{16.59}$$

We note that there is no momentum conjugate to A_0 (there is no $\dot{A}_0$ in the action), so the primary constraint of pure YM is

$$p^a_0 = 0. \tag{16.60}$$

We can calculate the classical Hamiltonian by partially integrating a D_i acting on A^a_0 in the action (16.56) and then writing

$$S = \int dtd^3x [\dot{A}^a_i p^a_i - H_L] = \int dtd^3x [\dot{A}^a_i E^a_i - H_0 + A^a_0 (D_i E^a_i)] , \tag{16.61}$$

to obtain

$$H_0 = \int d^3x \left[\frac{1}{2}(p^a_i)^2 + \frac{1}{2}(B^a_i)^2\right] = \int d^3x \left[\frac{1}{2}(E^a_i)^2 + \frac{1}{2}(B^a_i)^2\right] \tag{16.62}$$

where $(E^a_i)^2$ should be understood as $(p^a_i)^2$ in the Hamiltonian, and $(B^a_i)^2 = (F^a_{jk})^2/2$. We can also define the *naive Hamiltonian* H_L as above, with

$$H_L = H_0 - \int d^3x A^a_0 D_i E^a_i. \tag{16.63}$$

We see that the time evolution of the primary constraint p^a_0 with the naive Hamiltonian,

$$\dot{p}^a_0 \approx \{p^a_0, H_L\} = -D_i E^a_i \approx 0 , \tag{16.64}$$

implies that the secondary constraint of the pure Yang-Mills is

$$D_i E^a_i = 0. \tag{16.65}$$

We can check there are no other secondary constraints, since the time variation of the above vanishes,

$$\{D_i E^a_i, H_L\} = 0. \tag{16.66}$$

From (16.61) we see that A^a_0 appears as a Lagrange multiplier for the secondary constraint $D_i E^a_i$, hence the definition for H_L differing from H_0.

The algebra of constraints is found easily. We have

$$\{p_0^a(x), p_0^b(y)\}_{P.B.} = 0 \tag{16.67}$$

On the other hand,

$$\{p_0^a(x), D_i E_i^b(y)\}_{P.B.} = \{p_0^a(x), \partial_i p_i^b(x) + g f^b{}_{cd} A_i^c(y) p_i^d(y)\}_{P.B.} = 0\ , \tag{16.68}$$

and

$$\begin{aligned} \{D_i E_i^a(x), D_j E_j^b(y)\}_{P.B.} &= \{\partial_i p_i^a(x) + g f^a{}_{cd} A_i^c(x) p_i^d(x), \partial_j p_j^b(y) + g f^b{}_{ef} A_j^e(y) p_j^f(y)\}_{P.B.} \\ &= g f^{ab}{}_c D_i E_i^c(x) \delta^3(x-y). \end{aligned} \tag{16.69}$$

Therefore the constraints form a closed algebra, i.e. all constraints are first class, and the structure constants are indeed constant, and are $0, 0, g f^{ab}{}_c$. Since all constraints are first class, the Dirac brackets are just the Poisson brackets.

Therefore the constraints and the phase space coordinates are

$$G^a = (D_i E_i^a, p_0^a); \quad \eta_a = (c^a, -p(b)^a); \quad p_a = (p(c)_a, -b_a). \tag{16.70}$$

Since the structure constants are actually constant, the BRST charge and Hamiltonian are given by (16.49) and (16.50), giving

$$\begin{aligned} Q_B &= \int d^3x \left[c^a (D_i E_i^a) - p(b)^a p_0^a - \frac{1}{2} g c^b c^c f_{cb}{}^a p(c)_a \right] \\ H_{BRST} &= H_0 = \int d^3x \left[\frac{1}{2}(p_i^a)^2 + \frac{1}{2}(B_i^a)^2 \right]. \end{aligned} \tag{16.71}$$

The (first class) constraints are

$$\phi_1 = p_0^a; \quad \phi_2 = D_i p_i^a. \tag{16.72}$$

The naive Hamiltonian is

$$H_L = H_0 - \int d^3x A_0^a \partial_i p_i^a. \tag{16.73}$$

Since all constraints are first class, $H_E = H_L$. Note that the constraint $p_0^a = 0$ does not appear with Lagrange multiplier since it can be absorbed in the term $p_0^a \dot{A}_0^a$.

Note that in the Dirac formalism, the classical action is written as

$$S_{cl} = \int dt [\dot{q}_i p^i - H_0 + \lambda^\alpha \phi_\alpha]\ , \tag{16.74}$$

where λ_α are Lagrange multipliers.

Now, in the BRST formalism, the quantum action (16.44) is written as

$$S_{qu} = \int d^4x \left[E_i^a \dot{A}_i^a + p_0^a \dot{A}_0^a - p(c)_a \dot{c}^a - p(b)^a \dot{b}_a - \frac{1}{2}\left\{(E_i^a)^2 + (B_i^a)^2\right\} + \{\psi, Q_H\} \right]. \tag{16.75}$$

Here the first two terms are $\dot{q}_i p^i$, the next two are $\dot{\eta}^a p_a$, there are no $\dot{\lambda}^\mu \pi_\mu$ terms, and we have substituted $H_{BRST} = H_0$.

Note that we have described the classical part of the BRST formalism, but as we said, the quantization procedure replaces Dirac brackets (which now equal Poisson brackets) with commutators, and the physical Hilbert space is the space of the Q_B-cohomology. The procedure is more involved, but again we find only transverse gluons.

Batalin-Vilkovitsky formalism (field-antifield).

Finally, a few words about the generalization of the BRST formalism, the Batalin-Vilkovitsky (BV) formalism, or field-antifield formalism.

When we have derived the Lee-Zinn-Justin identities and made a formal renormalization of gauge theories, we have introduced sources K_μ^a and L^a for the nonlinear terms $Q_B A$ and $Q_B c$, namely $S_{extra} = \int [K_\mu^a (Q_B A)_\mu^a - L^a (Q_B c)^a]$. These sources are called antifields for A_μ^a and c^a. In the BV formalism, one would introduce antifields also for b_a and d_a. As we argued there, for YM it is actually not needed to introduce such antifields, but in more general situations it is.

Then BV is a Lagrangean formalism (where BRST was a Hamiltonian formalism), and a Lorentz-covariant one. In it moreover, there is a simplicity that stems from the fact that the BRST charge Q_{BRST} equals the antifield-extended quantum action S. Then we have

$$\delta_{BRST} \phi^A = (\phi^A, S\Lambda) \; , \tag{16.76}$$

and instead of $Q_B^2 = 0$ and $[Q_B, H_{BRST}] = 0$ we have the *master equation*

$$(S, S) = 0. \tag{16.77}$$

Important concepts to remember

- BRST quantization is based on the Dirac formalism, and extends to the BV formalism.
- The extended Hamiltonian adds to a first class Hamiltonian all the first class constraints.
- Dirac brackets are introduced so as to make trivial the second class constraints, $[f, \chi_s] = 0$ strongly for any $f(q,p)$, so that we can impose χ_s on states when quantizing, $\hat{\chi}_s |\psi\rangle = 0$.
- The physical Hilbert space of BRST quantization is the Q_B-cohomology, i.e. Q_B-closed states, modulo Q_B-exact states.
- We need to construct the operators Q_B and H_{BRST} order by order in the fields, such as to have $Q_B^2 = 0$ and $[Q_B, H_{BRST}] = 0$.
- For electromagnetism, the BRST quantization selects a Hilbert space without b or c ghosts, and with only transverse photons, $e^\mu a_\mu^\dagger(\vec{p})$, with $e^\mu p_\mu = 0$.

- In general, $Q_B = \eta^a G_a - 1/2\eta^c\eta^b f_{bc}{}^a p_a(-)^c + ...$ and $H_{BRST} = H_0 + \eta^a V_a^b p_b + ...$
- In YM theory, the primary constraint is $p_0^a = 0$ and the secondary constraint is $D_i E_i^a = 0$, and both are first class. The structure constants of the algebra of constraints are constant, and are $gf_{ab}{}^c$.

Further reading: See chapter 15.7, 15.8 in vol. II of Weinberg's "The Quantum Theory of Fields". I have used however mostly various lecture notes.

Exercises, Lecture 16

1) *The bosonic string in Dirac formalism.*
Consider the action

$$S_{cl} = \int d\sigma dt \left[-\frac{1}{2}\tilde{g}^{\mu\nu}\partial_\mu \vec{X}\cdot\partial_\nu \vec{X}\right] , \tag{16.78}$$

where $\tilde{g}^{\mu\nu} = \sqrt{-g}g^{\mu\nu}$ has $\det \tilde{g}^{\mu\nu} = -1$. Then $\tilde{g}^{00}\equiv \lambda^0$ and $\tilde{g}^{01}\equiv\lambda^1$ are independent Lagrange multipliers, as we can check, and $\tilde{g}^{11}\lambda^0 - (\lambda^1)^2 = -1$. Find the primary constraints and write H_L. From the consistency conditions $\{\phi_m, H_T\} = 0$, where ϕ_m are the primary constraints, find the secondary constraints, and the "physical constraints"= linear combinations independent of λ^i. Calculate the algebra of constraints and that all are first class, and that the "classical Hamiltonian" H_0 vanishes, as in general relativity.

2) Check that $Q_B^2 = 0$ and $[Q_B, H_{BRST}] = 0$ to first order, for the general form of BRST quantization objects

$$\begin{aligned} Q_B &= \eta^a G_a - \frac{1}{2}\eta^c\eta^b f_{bc}{}^a p_a (-)^c \\ H_{BRST} &= H_0 + \eta^a V_a^b p_b + ... \end{aligned} \tag{16.79}$$

where the first class algebra

$$\{\phi_\alpha, \phi_\beta\} = f_{\alpha\beta}{}^\gamma\phi_\gamma; \quad \{H_0, \phi_\alpha\} = V_\alpha^\beta \phi_\beta , \tag{16.80}$$

is extended to

$$\{G_a, G_b\} = f_{ab}{}^c G_c; \quad \{H_0, G_a\} = V_a^b G_b. \tag{16.81}$$

17 Lecture 17. QCD: definition, deep inelastic scattering

In this lecture we start the study of QCD. We have already described gauge theories, and their perturbation theory, so there is nothing new to be said there. Instead, what is of interest is the relation to experiments, which involve nonperturbative low energy states. So in this lecture we will learn how to combine the perturbative scattering methods studied previously with non-perturbative descriptions for the low energy states.

QCD

QCD refers to the theory of quarks and gluons. The YM gauge group is $SU(3)_c$ and there are 6 flavors of quarks, divided into 3 families (generations). The quark model was introduced by Gell-Mann and Zweig. The quarks are

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} , \tag{17.1}$$

and the elements on the top line have electric charge $Q_e = +2/3$, whereas the elements on the bottom line have $Q_e = -1/3$.

We know that there are 3 generations both from particle accelerator experiments (from the decay of the Z^0 for instance) and from cosmology constraints. The 3 families of quarks go together with 3 families of leptons to make the 3 generations of fermions,

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix} , \tag{17.2}$$

where now the electric charge on the top line is -1, and on the bottom line is zero.

Low energy states are gauge-invariant, since the theory is confining at low energies. That means that there is a linear potential between free quarks (objects with color charge), $V \sim \sigma l$, that prevents one from breaking pairs apart. Only un-charged (gauge-invariant) objects are exempt from this potential.

The QCD states are called hadrons in general, and divide into mesons and baryons.

-Mesons are $\bar{q}q$ pairs, for instance the lightest objects of QCD, the pions, $\pi^\pm$ and π^0.

-Baryons are objects with 3 quarks, qqq, for instance the proton, $p = (uud)$ and the neutron, $n = (ddu)$.

In terms of color indices, the mesons are $\bar{q}^i q_i = \bar{q}^i q_j \delta_i^j$, where $i,j = 1,2,3$ $(1,2,...,N)$ so are constructed with the group invariant δ_i^j that relates the fundamental $\mathbf{N}$ representation and the antifundamental $\bar{\mathbf{N}}$ representation. The baryons are $\epsilon^{ijk} q_i q_j q_k$ objects, constructed with the invariant of $SU(3)_c$ in the fundamental representation ϵ^{ijk}.

At low energies, there is also an approximate global symmetry of the Lagrangean called isospin, which acts on (ud) by

$$\begin{pmatrix} u \\ d \end{pmatrix} \to U \begin{pmatrix} u \\ d \end{pmatrix} , \tag{17.3}$$

so exchanges u and d, and correspondingly p with n.

Deep inelastic scattering

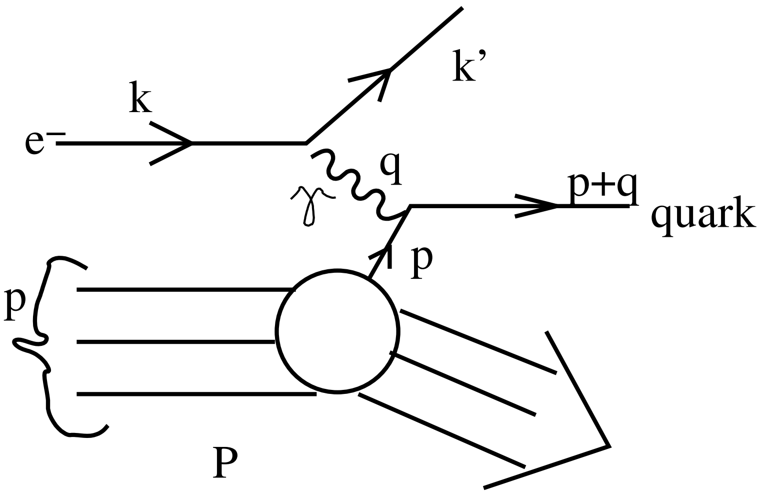

Figure 47: Deep inelastic scattering in QCD.

We begin with the most common example of scattering involving hadrons, deep inelastic scattering, see Fig.47. It is the scattering of an electron e^- off a proton p. We can ask, is it not simpler to start with two protons, for instance? But the proton has size and structure, that we want in fact to model, hence scattering two hadrons will make the problem harder. Instead we use the pointlike electron to probe the structure of the proton, by breaking it.

Therefore the scattering that we study is $e^-p \to e^-X$, where we have used p as a hadron, and after the collision we will have in general more than one hadron, a state that we called X.

Parton model.

For the nonperturbative proton, Bjorken and Feynman proposed the parton model, namely that the proton is made up of constituents called *partons*. We also assume that the partons cannot exchange large transverse momentum, i.e. large q^2 through strong interactions (gluon exchange). This is valid only to the leading order $\mathcal{O}(\alpha_s)$. Note that in the quark model, the proton is *classically* made up of *uud* quarks, but the parton model is more than that. It says that *quantum mechanically*, for the strong coupling state at low energy, the proton can be thought of as a combination of various quarks and gluons (real and virtual), with the total (overall) quantum numbers of the (*uud*) combination.

Then we consider an electron e^- with momentum k emitting a photon γ with momentum q and resulting in a momentum k' for the electron. Then the photon "extracts" a parton f with momentum p from the proton with overall momentum P and interacts with it, turning it into a parton with momentum $p' = p + q$. Since we work at leading order in α_s, we ignore gluon emission and exchange, so the parton will actually be a quark.

We will work in the center of mass frame and at total energies $\sqrt{s}$ much greater than the proton mass, which means that the proton is ultra-relativistic, i.e. almost lightlike. The parton constituents of the proton are assumed to have a finite fraction of the proton momentum, so they will also be almost lightlike, and almost collinear with the proton because the large p_T can come only from hard gluon exchange, which as we said is suppressed by α_s, so can be ignored at leading order.

Then the momentum of the parton constituent is written as

$$p = \xi P \ , \tag{17.4}$$

where ξ is called the longitudinal (momentum) fraction of the constituent.

Then the cross section for the total process is written as

$$\sigma[e^- p \to e^- X] = \int \sum_f P_f(\xi) \sigma_{e^- q_f \text{ scatt.}}(\xi) \ , \tag{17.5}$$

where the probability of finding a quark f at momentum fraction ξ is infinitesimal, and given by

$$P_f(\xi) = d\xi \ f_f(\xi) \ , \tag{17.6}$$

where $f_f(\xi)$ is the *partion distribution function* for the parton f. Here the partons (constituents of the proton) are $(q, \bar{q}, g)$. The parton distribution functions cannot be computed in perturbation theory, and have to be determined from experiments. Since ξ is a momentum fraction, it takes values in $\xi \in [0, 1]$.

Then we have

$$\sigma[e^-(k)p(P) \to e^-(k') + X] = \int_0^1 d\xi \sum_f f_f(\xi) \sigma[e^-(k) q_f(\xi P) \to e^-(k') + q_f(p')]. \tag{17.7}$$

To describe the scattering, one of the parameters we use is the quantity $Q^2 \equiv -q^2$. The Mandelstam variables for the basic scattering process $e^- q \to e^- q$ are written with hats. In particular,

$$\hat{t} = -Q^2 \tag{17.8}$$

and the other independent invariant is

$$\hat{s} = (p + k)^2 = 2p \cdot k = 2\xi P \cdot k = \xi s \ , \tag{17.9}$$

where we have used the fact that we are in the ultrarelativistic regime for the electron, parton and proton, so $p^2 = k^2 = P^2 \simeq 0$. Here s is the Mandelstam invariant for the total process, $e^- p \to e^- X$. Finally, since in general $s + t + u = \sum_i m_i^2$, we have now

$$\hat{s} + \hat{t} + \hat{u} = 0. \tag{17.10}$$

Then we also have (since also $(p + q)^2 \simeq 0$ for the final parton being ultra-relativistic)

$$0 \simeq (p + q)^2 = 2p \cdot q + q^2 = 2\xi P \cdot q - Q^2 \ , \tag{17.11}$$

which means that

$$\xi = x \equiv \frac{Q^2}{2P \cdot q}. \tag{17.12}$$

For the basic scattering process, $e^- q \to e^- q$, the formula has been derived in QFTI, so we will not repeat it here. The formula for the spin-averaged amplitude squared was (eq. (25.29) in QFTI)

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2 = \frac{8e^4 Q_f^2}{\hat{t}^2}\left(\frac{\hat{s}^2+\hat{u}^2}{4}\right) , \tag{17.13}$$

where Q_f is the electric charge of parton (quark) f.

Then, since $\alpha = e^2/4\pi$ and $\hat{u}^2 = (\hat{s}+\hat{t})^2$, we obtain for the relativistically invariant differential cross section for the basic process

$$\frac{d\sigma[e^- q_f \to e^- q_f]}{d\hat{t}} = \frac{\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2}{16\pi \hat{s}^2} = \frac{2\pi\alpha^2 Q_f^2}{\hat{s}^2}\left(\frac{\hat{s}^2+(\hat{s}+\hat{t})^2}{\hat{t}^2}\right). \tag{17.14}$$

Then for the total process, replacing $\hat{t}$ by $-Q^2$ and $\hat{s}$ by ξs, we have

$$\frac{d\sigma}{dQ^2} = \int_0^1 d\xi \sum_f f_f(\xi)\frac{2\pi\alpha^2 Q_f^2}{Q^4}\left[1+\left(1-\frac{Q^2}{\xi s}\right)^2\right]\theta(\xi s - Q^2). \tag{17.15}$$

Note that we have introduced a Heaviside function $\theta(\xi s - Q^2)$ since we need to have $\hat{s} \geq |\hat{t}|$. Normally, that would be just a constraint on external momenta, which could be put there explicitly with a θ or not, but now there are momentum fractions integrated over, so it needs to be put in explicitly.

We can also consider the second derivative of σ, taking into account that $\xi = x$, with respect to this x, obtaining

$$\frac{d^2\sigma}{dx\, dQ^2} = \left(\sum_f f_f(x) Q_f^2\right)\frac{2\pi\alpha^2}{Q^4}\left[1+\left(1-\frac{Q^2}{xs}\right)^2\right]. \tag{17.16}$$

Note that now, being a differential formula with respect to x, there is no need to put explicitly the Heaviside function, though one could.

Finally then, we obtain *Bjorken scaling*, the scaling relation that says that

$$\frac{d^2\sigma}{dx\, dQ^2}\frac{Q^4}{1+\left(1-\frac{Q^2}{xs}\right)^2} = \left(\sum_f f_f(x) Q_f^2\right) 2\pi\alpha^2 \tag{17.17}$$

is independent of Q^2, and only depends on x.

Qualitatively, the scaling says that the structure of the proton looks the same to an electromagnetic probe, no matter how hard the proton is struck (how large is Q^2). It is verified experimentally very well, but it is true *only to first order in* α_s.

Another useful variable that can be defined is

$$y \equiv \frac{2P\cdot q}{2P\cdot k} = \frac{2P\cdot q}{s}. \tag{17.18}$$

It can also be rewritten in terms of the Mandelstam variables of the basic scattering, since $(kp) \to (k'p')$ means that $\hat{s} = (p+k)^2 = 2p \cdot k$ and $\hat{u} = (p-k')^2 = -2p \cdot k'$, so

$$y = \frac{2\xi p \cdot (k - k')}{2\xi p \cdot k} = \frac{\hat{s} + \hat{u}}{\hat{s}}. \tag{17.19}$$

But since $|\hat{u}| \leq \hat{s}$, we have

$$\frac{\hat{u}}{\hat{s}} = -(1-y) \Rightarrow y \leq 1. \tag{17.20}$$

From $y = 2P \cdot q/s$ and $x = Q^2/2P \cdot q$ if follows that

$$xys = Q^2 \ , \tag{17.21}$$

which also implies

$$d\xi \ dQ^2 = dx \ dQ^2 = \frac{dQ^2}{dy} dx \ dy = xs \ dx \ dy. \tag{17.22}$$

Finally then, we have for deep inelastic scattering (DIS)

$$\frac{d^2\sigma[e^- p \to e^- X]}{dx \ dy} = \left(\sum_f x f_f(x) Q_f^2 \right) \frac{2\pi\alpha^2 s}{Q^4} [1 + (1-y)^2]. \tag{17.23}$$

This means that

$$\frac{d^2\sigma}{dx \ dy} Q^4 \tag{17.24}$$

factorizes into two factors, one depending only on x (the Bjorken scaling factor) and one depending only on y, $[1 + (1-y)^2]$, which gives the *Callan-Gross relation* for scattering of an e^- off a massless fermion (indeed, we saw that this factor originated in the calculation of the spin averaged $|\mathcal{M}|^2$ for the e^- to scatter off a massless fermion).

One more thing to note here is that the particular form of the final hadronic states X, that come from the remnant of the original proton and the "jet" that will form out of the final quark q_f, are not part of the calculation. Of course, we observe only hadronic final states, but the efect of how these final states turn into the observed hadrons, "hadronization", does not influence too much the cross section, and moreover, there are nonperturbative (lattice, etc.) methods to calculate these effects, and treat them as a "black box".

Deep inelastic neutrino scattering

We have analyzed in DIS the effect of electromagnetic probes for the proton, but we can also consider weak interaction probes, meaning consider a W exchange instead of the γ exchange, hence the probe to be considered is a neutrino. For definiteness, we consider a ν_μ.

The weak interaction, the remnant of the broken electroweak interaction, which will not be described here, is to exchange the massive $W^\pm$ vector particles. It couples to the weak doublets (i.e. quark pairs like (ud) and lepton pairs like $(e\nu_e)$ and $(\mu\nu_\mu)$), and it turns one element of the doublet into the other.

Hence the basic interaction we consider is as follows. A ν_μ comes and emits a W^-, thus turning into a μ^-. The W^+ can now interact with a d quark parton inside the proton and

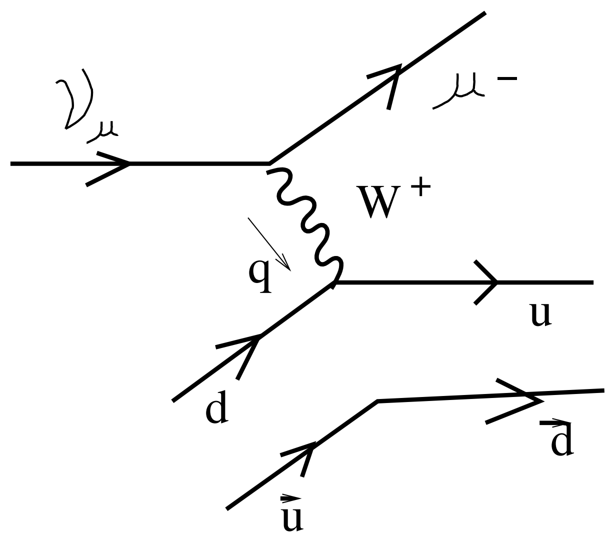

Figure 48: Deep inelastic ν scattering.

turn it into a u. But remembering that the partons are not only the classical quarks, the W^+ can also interact with a $\bar{u}$ and turn it into a $\bar{d}$.

The weak interaction through W^+ exchange can be considered at low energies $E \ll m_W$ as a *4-fermi interaction*. Fermi had introduced the 4-fermi interaction as an effective model with a vertex $G_F/\sqrt{2}$ coupling 4 fermions (2 $\bar{\psi}\psi$ lines), but it was later realized that this effective vertex comes from the approximation of the real quantum process, the exchange of the massive vector boson W^+. Indeed then, in the actual diagram in Fig.48, we would have a massive vector propagator with momentum q coupling to two fermion lines,

$$g(...)^\mu \frac{\delta_{\mu\nu}}{q^2 + m_W^2} g(...)^\nu \simeq \frac{g^2}{m_W^2}(...)^\mu(...)_\mu \; , \tag{17.25}$$

where g is the coupling of the W with the fermions, and the approximation is for $q^2 \ll m_W^2$, since $q^2 = (k-k')^2$ and $|k|, |k'| \ll m_W$. That means that we can consider an effective *Fermi coupling* of

$$\frac{G_F}{\sqrt{2}} \equiv \frac{g^2}{8m_W^2}. \tag{17.26}$$

Then, similarly with the DIS case before, we find

$$\frac{d^2\sigma}{dx\,dy}(\nu p \to \mu^- X) = \frac{G_F^2 s}{\pi}[x f_d(x) + x f_{\bar{u}}(x)(1-y)^2]. \tag{17.27}$$

The proof is left as an exercise. We also find

$$\frac{d^2\sigma}{dx\,dy}(\bar{\nu} p \to \mu^+ X) = \frac{G_F^2 s}{\pi}[x f_u(x)(1-y)^2 + x f_{\bar{d}}(x)]. \tag{17.28}$$

Normalization of the parton distribution functions

As we mentioned, classically, the proton is made up of (uud) quarks, but quantum mechanically, we expect a strongly coupled mix of states, so we will create many $q\bar{q}$ pairs, as well as gluons, inside this proton. But the overall quantum numbers of the state still have to be the numbers of the (uud) state, so it means in particular that we need to have 2 more u quarks than $\bar{u}$ quarks, and 1 more d quark than $\bar{d}$ quark. This translates into the conditions

$$\begin{aligned}\int_0^1 dx[f_u(x) - f_{\bar{u}}(x)] = 2 \\ \int_0^1 dx[f_d(x) - f_{\bar{d}}(x)] = 1.\end{aligned} \tag{17.29}$$

Of course, there can be also trace parts of other quarks, but are negligible. The next important one is the s quark, so we need the same number of s and $\bar{s}$ quarks, so

$$\int_0^1 d\xi[f_s(x) - f_{\bar{s}}(x)] = 0 \ , \tag{17.30}$$

etc.

A similar story holds for the other hadrons. For instance, for the neutron, we just exchange u and d from the proton, so

$$f_u^{(n)}(x) = f_d(x); \quad f_d^{(n)}(x) = f_u(x); \quad f_{\bar{u}}^{(n)}(x) = f_{\bar{d}}(x) \ , \tag{17.31}$$

etc. As another example, for the antiproton, we just interchange the quarks with antiquarks, so

$$f_u^{(\bar{p})}(x) = f_{\bar{u}}(x), \quad f_{\bar{u}}^{(\bar{p})}(x) = f_u(x) \ , \tag{17.32}$$

etc.

Finally, the last normalization constraint comes from the fact that the total momentum of the proton is P, so the total fraction of all the constituent partons is 1, i.e.

$$\int_0^1 dx \ x[f_u(x) + f_d(x) + f_{\bar{u}}(x) + f_{\bar{d}}(x) + f_g(x)] = 1. \tag{17.33}$$

Note that we could have introduced also $f_s(x), f_{\bar{s}}(x)$, etc., but as we said, these are negligible.

The parton distribution functions are determined experimentally, and must obey the above normalization conditions. Among the many QCD experiments, one uses some to fix the distribution functions, and then uses them to predict the other cross sections.

Hard scattering processes in hadron collisions.

Finally, a few words about hard scattering, for completeness. We can now treat the next complicated case, the case of hadron-hadron scattering, as suggested at the beginning of the lecture. The basic process in leading order will be some $q_f\bar{q}_f \to Y$ process, for instance occurring electromagnetically, through an intermediate virtual photon. A parton q_f from one proton will break off and interact with the parton $\bar{q}_f$ from the second proton. All in all, we have the total process

$$\sigma(p(P_1) + p(P_2) \to Y + X) = \int_0^1 dx_1 \int_0^1 dx_2 \sum_f f_f(x_1) f_{\bar{f}}(x_2) \sigma(q_f(x_1 P_1) + \bar{q}_f(x_2 P_2) \to Y) \ , \tag{17.34}$$

where we have as usual that the remnants of the proton will hadronize to some state X.

Important concepts to remember

- QCD is YM with the color group $SU(3)_c$ and 6 flavours, organized in 3 families.
- Physical low energy states are gauge invariant due to confinement. Mesons are $\bar{q}^i q_i$ and baryons are $\epsilon^{ijk} q_i q_j q_k$.
- In the parton model, the proton is composed at the quantum level of partons, $(q, \bar{q}, g)$.
- Deep inelastic scattering (DIS) is an electron scattering off a hadron, usually a proton, breaking out a parton from it, and interacting with it.
- The cross sections for the total process are integrals of the cross sections for the basic process involving the parton, with the distributions functions for momentum fraction ξ of the parton inside the proton, $p = \xi P$.
- Bjorken scaling gives the independence of some quantity on Q^2, and only dependence on $x = Q^2/2P \cdot q$, saying that the structure of the proton looks the same to an electromagnetic probe, no matter how hard the proton is struck (how high is Q^2).
- Deep inelastic neutrino scattering occurs through W^+ exchange, which reduces to 4-fermi interaction at low energies, and turns a d into an u.
- Parton distribution function must obey the normalization conditions, and are determined experimentally.
- Hard scattering processes (collisions of two hadrons) involve two parton distribution functions, one for each hadron.

Further reading: See chapter 17.1, 17.3 in [3].

Exercises, Lecture 17

1) Fill in the details of the calculation of $d^2\sigma/dx\,dy(\nu p \to \mu^- X)$.

2) Consider the $e^- p \to e^- X$ scattering, and assume that

$$f_u(x) = \frac{3}{2} f_d(x) = f_g(x) = \frac{1-x}{a(x+\epsilon)} \tag{17.35}$$

and

$$f_{\bar{u}}(x) = f_{\bar{d}}(x) = \frac{1-x}{3a(x+\epsilon)} \ , \tag{17.36}$$

where a and ϵ are constants and $\epsilon \ll 1$, and all possible Q_f^2 are given by their standard values ($Q_f(u) = 2/3$, etc.). From Bjorken scaling, calculate the cross section $\sigma(e^- p \to e^- X)$.

18 Lecture 18. Parton evolution and Altarelli-Parisi equation.

Last lecture we saw that we can describe the deep inelastic scattering, the collision of an electron off a hadron, via parton distribution functions $f_f(x)$, that were found to be independent of Q^2, and give Bjorken scaling. But we mentioned that this was true only to leading order in α_s. In this lecture we will consider processes subleading in α_s that will lead to violations of Bjorken scaling, through Q^2 dependence of the parton distribution functions.

The dependence on Q^2, and the subsequent *evolution* of $f_f(x)$, will be due to processes with emission of collinear quarks and gluons, also responsible for IR divergences. Therefore the parton evolution will be linked with the regularization of IR divergences, that will be the addressed after next lecture. Here however the approach will be a more practical one, so we will not deal with the formal issue of IR divergences.

We will not start with QCD, but rather with the simpler case of QED, and then see that we can import almost all the calculation to the QCD case with minimal effort.

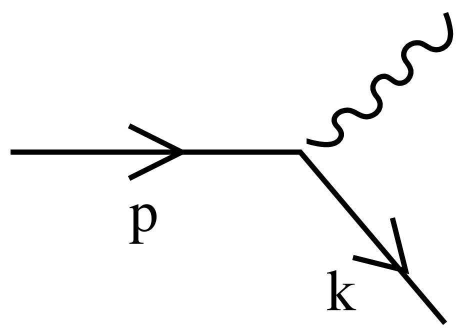

Figure 49: Emission of a photon off an electron line.

QED process

The process that most interests us is the one of photon emission from a fermion (electron) line, as in Fig.49. The initial electron has momentum p, the final one momentum k and the emitted photon momentum q. We call the momentum (energy) fraction carried by the photon z, i.e.

$$z \equiv \frac{E_\gamma}{E_{e,in}}. \tag{18.1}$$

The initial electron is ultra-relativistic, i.e. almost massless, and we choose it to be in the 3 direction, so

$$p = (p, 0, 0p). \tag{18.2}$$

The kinematics of 3-point scattering mean that one of the momenta of emitted particles has to be off-shell in the presence of transverse momentum $\vec{p}_\perp$, which is small (almost collinear emission) $p_\perp \ll p$. If the photon is massless, $q^2 = 0$, to leading order in $p_\perp/p$ we have

$$q \simeq \left(zp, \vec{p}_\perp, zp - \frac{p_\perp^2}{2zp}\right) \ , \tag{18.3}$$

and then the final electron, which is also ultrarelativistic, must be however off-shell (virtual), since momentum conservation implies

$$k \simeq \left((1-z)p, -\vec{p}_\perp, (1-z)p + \frac{p_\perp^2}{2zp} \right). \tag{18.4}$$

Then we have

$$k^2 \simeq \frac{p_\perp^2}{z} \neq 0. \tag{18.5}$$

One can also consider an on-shell final electron and a virtual photon, with

$$\begin{aligned} k &\simeq \left((1-z)p, -\vec{p}_\perp, (1-z)p - \frac{p_\perp^2}{2(1-z)p} \right) \\ q &\simeq \left(zp, \vec{p}_\perp, zp + \frac{p_\perp^2}{2(1-z)p} \right) , \end{aligned} \tag{18.6}$$

which leads to $k^2 \simeq 0$ and

$$q^2 \simeq \frac{p_\perp^2}{(1-z)}. \tag{18.7}$$

We will not do the calculation here, but it is easy to calculate the amplitude, and find $|\mathcal{M}|^2$ averaged over initial polarizations and summed over final polarizations. One obtains

$$\frac{1}{2}\sum_{\text{pol.}} |\mathcal{M}|^2 = \frac{2e^2 p_\perp^2}{z(1-z)} \left[\frac{1+(1-z)^2}{z} \right]. \tag{18.8}$$

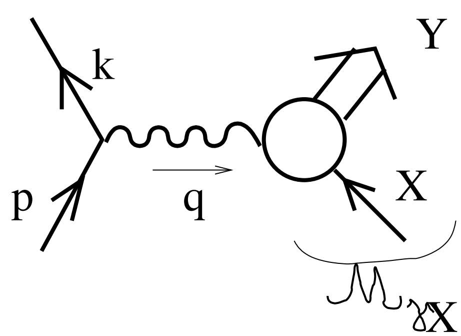

Figure 50: Equivalent photon approximation.

Equivalent photon approximation.

Now we can turn to the calculation we are interested in. We want to study QED corrections to the process of deep inelastic scattering (DIS) of an electron scattering of a hadron X to give another electron and hadron(s) Y, through interaction with an intermediate (virtual) photon γ, as in Fig.50. We can divide the process into the emission of a photon from the electron, $e^- \to e^-\gamma$, followed by the scattering $\gamma X \to Y$. For the total amplitude we find

$$\mathcal{M}^{\text{tot}} = \mathcal{M}^\mu \frac{\delta_{\mu\nu}}{q^2} \mathcal{M}^\nu_{\gamma X} , \tag{18.9}$$

and doing the sum over the photon polarizations we find

$$\frac{1}{2}\sum_{pol}|\mathcal{M}^{\rm tot}|^2 = \frac{1}{2}\sum_{\rm pol}|\mathcal{M}|^2\frac{1}{(q^2)^2}|\mathcal{M}_{\gamma X}|^2. \qquad (18.10)$$

The formula for the cross section $A+B\to\sum f$ is

$$\sigma = \frac{1}{|\vec{v}_A-\vec{v}_B|2E_A2E_B}\int\prod_f d\Pi_f|\mathcal{M}^{\rm tot}|^2\ , \qquad (18.11)$$

where $d\Pi_f$ is the phase space for the final state product f. In our case, $B=X$ is the (possibly nonrelativistic) hadron with velocity v_X and energy E_X and $A=e^-$, with $E_A=p$ and $v_A=1$ in the opposite direction to X, so

$$\sigma = \frac{1}{(1+v_X)2p2E_X}\int\frac{d^3k}{(2\pi)^3 2k^0}\int d\Pi_Y\left[\frac{1}{2}\sum_{pol}|\mathcal{M}|^2\right]\frac{1}{(q^2)^2}|\mathcal{M}_{\gamma X\to Y}|^2. \qquad (18.12)$$

Given that the energy of the photon is zp, we can form the cross section for $\gamma X\to Y$. Also using that $k^0=(1-z)p$, $(q^2)^2=p_\perp^4/(1-z)^2$ and $d^3k=dk^0d^2\vec{p}_\perp = pdz\pi dp_\perp^2$, we find

$$\begin{aligned}\sigma &= \int\frac{pdzdp_\perp^2}{16\pi^2(1-z)p}\left[\frac{1}{2}\sum_{pol}|\mathcal{M}|^2\right]\frac{(1-z)^2}{p_\perp^4}\frac{z}{(1+v_X)2zp2E_X}\int d\Pi_y|\mathcal{M}_{\gamma X\to Y}|^2\\ &= \int_0^1 dz\int\frac{dp_\perp^2}{p_\perp^2}\frac{\alpha}{2\pi}\left[\frac{1+(1-z)^2}{z}\right]\sigma(\gamma X\to Y). \qquad (18.13)\end{aligned}$$

It remains to consider the region of integration of $p_\perp^2$. $p_\perp^2$ cannot be smaller than m^2, which cuts off the potential IR divergence, and cannot be larger than the total energy squared, s, so we have $\int_{m^2}^s dp_\perp^2$. Finally, we obtain

$$\sigma(e^-X\to e^-Y) = \int_0^1 dz\frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+(1-z)^2}{z}\right]\sigma(\gamma X\to Y). \qquad (18.14)$$

This is the *Weizsacker-Williams equivalent photon approximation.* We see that the formula is of the same type as in the parton model case, with the cross section for scattering of γ integrated with a probability to find a γ inside the electron, or photon distribution function

$$f_\gamma(z) = \frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+(1-z)^2}{z}\right]\ , \qquad (18.15)$$

and the total cross section

$$\sigma(e^-X\to Y) = \int_0^1 dz f_\gamma(z)\sigma(\gamma X\to Y). \qquad (18.16)$$

Electron distribution

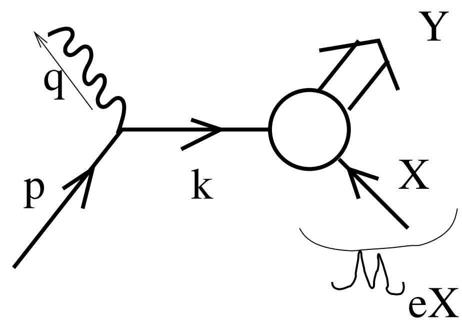

Figure 51: Electron distribution process.

We can now consider the case of photon emission, i.e. $e^- X \to \gamma Y$ scattering, proceeding through an intermediate e^-, so the γ is emitted from the e^-, and then we scatter $e^- X \to Y$. In the same was as before, we now obtain

$$\frac{1}{2}\sum_{pol}|\mathcal{M}^{\rm tot}|^2 = \frac{1}{2}\sum_{\rm pol}|\mathcal{M}|^2\frac{1}{(k^2)^2}|\mathcal{M}_{e^-X\to Y}|^2 \ , \tag{18.17}$$

so the total cross section is

$$\begin{aligned}
\sigma(e^-X\to\gamma Y) &= \frac{1}{(1+v_X)2p2E_X}\int\frac{d^3q}{(2\pi)^3 2q^0}\int d\Pi_y\left[\frac{1}{2}\sum_{pol}|\mathcal{M}|^2\right]\frac{1}{(k^2)^2}|\mathcal{M}_{e^-X\to Y}|^2\\
&= \int\frac{dzdp_\perp^2}{16\pi^2 z}\left[\frac{1}{2}\sum_{pol}|\mathcal{M}|^2\right]\frac{z^2}{p_\perp^4}(1-z)\sigma(e^-X\to Y)\\
&= \int_0^1 dz\int_{m^2}^s\frac{dp_\perp^2}{p_\perp^2}\frac{\alpha}{2\pi}\left[\frac{1+(1-z)^2}{z}\right]\sigma(e^-X\to Y)\ ,
\end{aligned} \tag{18.18}$$

where we have used that $q^0 = zp$, $d^3q = pdz\pi dp_\perp^2$ and $(k^2)^2 = p_\perp^4/z^2$. Again we can interpret this as in the parton model, through an electron distribution function $f_e^{(1)}(x)$ at momentum fraction $x = 1 - z$, with

$$f_e^{(1)}(x) = \frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+x^2}{1-x}\right] \ , \tag{18.19}$$

and

$$\sigma(e^-X\to\gamma Y) = \int_0^1 dx f_e^{(1)}(x)\sigma(e^-X\to Y). \tag{18.20}$$

However, this is interpreted as a probability to find an electron with momentum fraction x inside the electron, so we need to consider also the zeroth order process, where we already have the original electron, i.e. we need to include

$$f_e^{(0)}(x) = \delta(1-x). \tag{18.21}$$

This is however still not enough, since creating an electron with momentum fraction x means we need to subtract one from momentum fraction $x = 1$ (the initial electron), so we need to subtract a term proportional to $\delta(1-x)$, giving the normalization of $f_e^{(1)}(x)$. But for that, we need to replace $1/(1-x)$, which is divergent under the integral, with a well-behaved distribution $1/(1-x)_+$, defined by

$$\int_0^1 dx \frac{f(x)}{(1-x)_+} = \int_0^1 dx \frac{f(x)-f(1)}{1-x}. \tag{18.22}$$

Then the term we have in $f_e^{(1)}(x)$ means we must consider

$$\int_0^1 dx \frac{1+x^2}{(1-x)_+} = \int_0^1 \frac{x^2-1}{1-x} = -\int_0^1 dx(1+x) = -\frac{3}{2}\,, \tag{18.23}$$

and subtract this normalization, multiplied by $\delta(1-x)$, to obtain finally

$$f_e(x) = \delta(1-x) + \frac{\alpha}{2\pi}\log\frac{s}{m^2}\left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x)\right]. \tag{18.24}$$

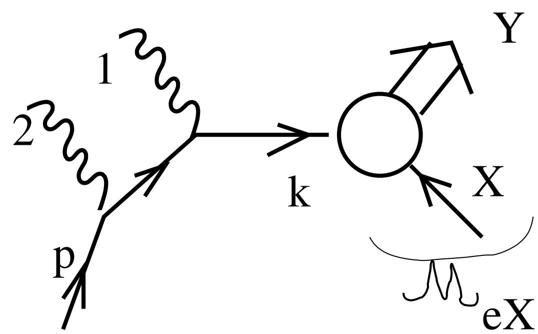

Figure 52: Multiple splittings from the electron line.

Multiple splittings

We now need to generalize the relations (18.15) and (18.24) for $f_\gamma(z)$ and $f_e(x)$ to the case of multiple splittings in Fig.52. Consider that we have 2 photons being emitted out of the electron line, with $p_{2\perp} \ll p_{1\perp}$, so the contribution is of order

$$\left(\frac{\alpha}{2\pi}\right)^2 \int_{m^2}^{s} \frac{dp_{1\perp}^2}{p_{1\perp}^2} \int_{m^2}^{p_{1\perp}^2} \frac{dp_{2\perp}^2}{p_{2\perp}^2} = \frac{1}{2}\left(\frac{\alpha}{2\pi}\right)^2 \log^2\frac{s}{m^2}. \tag{18.25}$$

Note that the integral is of the type

$$\int_a^b \frac{dx}{x}\log\frac{x}{a} = \log b \log\frac{b}{a} - \int_a^b \frac{dx}{x}\log x = \log^2\frac{b}{a} - \int_a^b \frac{dx}{x}\log\frac{x}{a} \Rightarrow \int_a^b \frac{dx}{x}\log\frac{x}{a} = \frac{1}{2}\log^2\frac{b}{a}. \tag{18.26}$$

We can generalize this to the case of multiple splittings, with $p_{1\perp} \gg p_{2\perp} \gg ... \gg p_{n\perp}$, which gives in a similar way a contribution of

$$\frac{1}{n!}\left(\frac{\alpha}{2\pi}\right)^n \log^n\frac{s}{m^2}. \tag{18.27}$$

That means we can consider the splittings as independent from each other, so we can consider a continuous process of splittings. We define distribution functions $f_\gamma(x,Q)$ and $f_e(x,Q)$ for $p_\perp < Q$. Then we consider increasing Q to $Q+\Delta Q$, which means that now the electrons can emit photons γ with $Q < p_\perp < Q+\Delta Q$. The probability of a constituent e^- to emit γs with a momentum fraction z is then obtained by differentiating (18.15). We get

$$\frac{dP}{dz} = \frac{\alpha}{2\pi}\frac{dp_\perp^2}{p_\perp^2}\frac{1+(1-z)^2}{z}. \tag{18.28}$$

But the constituent electron (parton) has a distribution function $f_e(x,p_\perp)$, so all in all we get

$$\begin{aligned} f_\gamma(x,Q+\Delta Q) &= f_\gamma(x,Q) + \int_0^1 dx' \int_0^1 dz \left[\frac{\alpha}{2\pi}\frac{\Delta Q^2}{Q^2}\frac{1+(1-z)^2}{z}\right] f_e(x',p_\perp)\delta(x-zx') \\ &= f_\gamma(x,Q) + \frac{\Delta Q}{Q}\int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}\frac{1+(1-z)^2}{z}\right] f_e\left(\frac{x}{z},p_\perp\right) , \end{aligned} \tag{18.29}$$

where in the first line we considered the fact that the momentum fraction x of the photon is the fraction z of the splitting times the original momentum fraction x' of the constituent electron, and in the second line we did the x' integral using $\delta(x-zx') = 1/z\delta(x'-x/z)$. We also used the fact that $1 \geq x' = x/z$, so $1 \geq z \geq x$.

We can then go to the continuum, and write the relation (for ΔQ infinitesimal),

$$\frac{d}{d\log Q} f_\gamma(x,Q) = \int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}\frac{1+(1-z)^2}{z}\right] f_e\left(\frac{x}{z},Q\right). \tag{18.30}$$

We can use the same logic for $f_e(x)$ and, calculating the probability to have electrons of momentum fraction x appear as a result of γ emission, it should come from (18.24) (minus the trivial delta function), times the electron distribution function itself, so

$$\frac{d}{d\log Q} f_e(x,Q) = \int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}\left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right)\right] f_e\left(\frac{x}{z},Q\right). \tag{18.31}$$

Boundary conditions

We have obtained a set of differential equations. To find a solution from them, we must give a boundary condition. The natural boundary condition is that there is only one electron, and nothing else, so $f_e = \delta(1-x)$ and $f_\gamma = 0$. Since the distribution functions f_γ and f_e in (18.15) and (18.24) have $\log s/m^2$, it means that we should define this boundary condition at $Q^2 = m^2$, so

$$f_e(x,Q)|_{Q^2=m^2} = \delta(1-x); \quad f_\gamma(x,Q)|_{Q^2=m^2} = 0. \tag{18.32}$$

With this boundary conditions and the differential equations that we found, we obtain $f_e(x,Q)$ and $f_\gamma(x,Q)$, and then the cross section for *multiple splittings* (viewed as a continuous process; note then that this does not include all possible terms, merely resums a set of diagrams) is

$$\sigma(e^-X \to e^- + n\gamma + Y) \;=\; \int_0^1 dx f_\gamma(x,Q)\sigma(\gamma X \to Y)$$

$$\sigma(e^- X \to n\gamma + Y) = \int_0^1 dx f_e(x,Q)\sigma(e^- X \to Y). \quad (18.33)$$

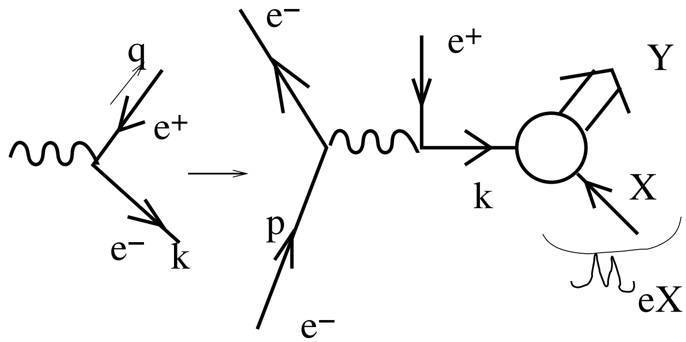

Figure 53: Photon splitting into pairs before interaction.

Photon splitting into pairs

There is one more process that we need to consider, which is a crossed diagram for the one for γ emission from an electron line, namely e^+e^- pair creation from the photon. For our DIS process, the relevant diagram is where the incoming electron emits a photon, but the photon turns into a e^+e^- pair, with e^+ emitted, and e^- interacting with X as before, as in Fig.53.

One can easily calculate the $\gamma \to e^+e^-$ process, since it is the crossed diagram to $e^- \to e^- + \gamma$, and find

$$\frac{1}{2}\sum_{pol} |\mathcal{M}|^2(\gamma \to e_L^- e_R^+) = \frac{2e^2 p_\perp^2}{z(1-z)}[z^2 + (1-z)^2]. \quad (18.34)$$

Then in a completely similar way to the previous cases, we can consider the continuous process where the virtual photon line emits a e^+ and turns into an e^-, (that emits a γ and gets out, and the γ continues on towards X and emits a pair...), etc. We easily see that the result will be a variation in the distribution function for e^-, and will be proportional to the new $[z^2 + (1-z)^2]$ bracket above and the distribution function for the γ, so

$$\frac{d}{d\log Q} f_e(x,Q)\Big|_{\text{pair cr.}} = \int_x^1 \frac{dz}{z}\left[\frac{\alpha}{\pi}(z^2 + (1-z)^2)\right] f_\gamma\left(\frac{x}{z}, Q\right). \quad (18.35)$$

But as a result, the f_γ will also be changed, since this subtracts a photon from momentum fraction $x = 1$, so as before, we must normalize, amounting to subtracting $\int_0^1 dz(z^2 + (1-z)^2) = 2/3$ times $\delta(1-z)$ in $df_\gamma/d\log Q$.

Evolution equations for QED.

Finally therefore, we obtain the evolution equations for QED, i.e. the equations giving the variation of the parton distribution functions for the electron with the scale Q, found by Gribov and Lipatov. Since we have a distribution function for the electron, we must also have (due to the process of pair creation) a distribution function for the positron, which will be given by the same formula as for the electron. Indeed, the positron can turn into a positron by emitting a γ just like the electron, and a positron can be created from a γ just like an electron (the process is e^+e^- pair creation). We must also add a term to create a γ out of an e^+, equal to the one creating a γ out of an e^-. We obtain

$$\begin{aligned}
\frac{d}{d\log Q} f_\gamma(x,Q) &= \frac{\alpha}{\pi}\int_x^1 \frac{dz}{z}\left\{P_{\gamma\leftarrow e}(z)\left[f_e\left(\frac{x}{z},Q\right)+f_{\bar e}\left(\frac{x}{z},Q\right)\right]+P_{\gamma\leftarrow\gamma}(z) f_\gamma\left(\frac{x}{z},Q\right)\right\}\\
\frac{d}{d\log Q} f_e(x,Q) &= \frac{\alpha}{\pi}\int_x^1 \frac{dz}{z}\left\{P_{e\leftarrow e}(z) f_e\left(\frac{x}{z},Q\right)+P_{e\leftarrow\gamma}(z) f_\gamma\left(\frac{x}{z},Q\right)\right\}\\
\frac{d}{d\log Q} f_{\bar e}(x,Q) &= \frac{\alpha}{\pi}\int_x^1 \frac{dz}{z}\left\{P_{e\leftarrow e}(z) f_{\bar e}\left(\frac{x}{z},Q\right)+P_{e\leftarrow\gamma}(z) f_\gamma\left(\frac{x}{z},Q\right)\right\}\ ,
\end{aligned} \tag{18.36}$$

where the *splitting functions* $P_{i\leftarrow j}(z)$ (considered as probabilities for turning parton j into parton i and calculated above) are

$$\begin{aligned}
P_{e\leftarrow e}(z) &= \frac{1+z^2}{(1-z)_+}+\frac{3}{2}\delta(1-z)\\
P_{\gamma\leftarrow e}(z) &= \frac{1+(1-z)^2}{z}\\
P_{e\leftarrow\gamma}(z) &= z^2+(1-z)^2\\
P_{\gamma\leftarrow\gamma}(z) &= -\frac{2}{3}\delta(1-z)\ ,
\end{aligned} \tag{18.37}$$

and the boundary conditions for integration of the differential equations are

$$f_e(x,Q)|_{Q^2=m^2}=\delta(1-x);\quad f_{\bar e}(x,Q)|_{Q^2=m^2}=0;\quad f_\gamma(x,Q)|_{Q^2=m^2}=0. \tag{18.38}$$

The parton distribution functions also obey the same kind of normalization conditions as in the hadron (QCD) case, namely we need to have one more electron than positron, so

$$\int_0^1 dx[f_e(x,Q)-f_{\bar e}(x,Q)]=1\ , \tag{18.39}$$

and the total momentum fraction of all the partons (e^+,e^-,γ) is 1, so

$$\int_0^1 dx\ x[f_e(x,Q)+f_{\bar e}(x,Q)+f_\gamma(x,Q)]=1. \tag{18.40}$$

It is left as an exercise to check that these normalization conditions are respected by the evolution in the Gribov-Lipatov equations.

Altarelli-Parisi equations.

We finally return to the case of QCD. Just like in the case of QED, the evolution was due to the emission of collinear ($p_\perp/p\ll 1$) photons and electrons, in the case of QCD, the evolution of the parton distribution functions is due to the emission of collinear gluons and quarks.

It will give a violation of Bjorken scaling, since now the DIS process will give

$$\frac{d^2\sigma}{dx\,dy}(e^-p \to e^-X) = \left(\sum_f f_f(x,Q)Q_f^2\right)\frac{2\pi\alpha_s^2}{Q^4}[1+(1-y)^2]\;, \tag{18.41}$$

and now the scaling is only approximate, since $f_f(x,Q)$ now depends on Q as well, not only on x.

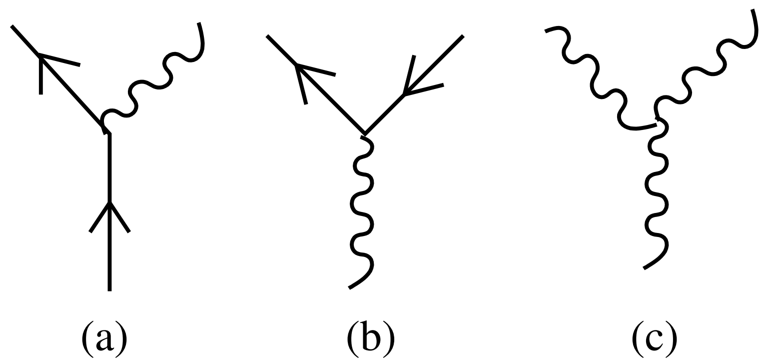

Figure 54: Diagrams relevant for the Altarelli-Parisi equation.

As in the case of QED, in QCD we have the two basic diagrams, for a gluon to be emitted from a quark line (a), and the crossed one, for a quark antiquark pair to be emitted from a gluon (b), as in Fig.54a and b. However, in the (nonabelian) QCD case, we also have the third diagram, for a gluon to split into two gluons (c), as in Fig.54c.

The calculations resulting from the diagrams (a) and (b) can be imported from QED to QCD with minimal modifications. The only difference is that we sum over final polarizations and average over initial polarizations, so now we must also sum over final colors, and average over initial colors.

The sum over initial and final colors can be easily seen to give a factor of $\mathrm{Tr}[t^a t^a]$. In diagram (a), the initial state is a quark, with 3 colors, so the averaging also gives a factor of 1/3, for a total

$$\frac{1}{3}\,\mathrm{Tr}[t^a t^a] = \frac{1}{3}\frac{1}{2}\delta^a_a = \frac{4}{3}\;, \tag{18.42}$$

since in $SU(3)$ there are 8 gluons ($N^2-1=8$ generators). In diagram (b), the initial state is a gluon, with 8 states, to the averaging also gives a factor of 1/8, for a total

$$\frac{1}{8}\,\mathrm{Tr}[t^a t^a] = \frac{1}{8}\frac{1}{2}\delta^a_a = \frac{1}{2}. \tag{18.43}$$

Diagram (c) needs to be done, since it doesn't appear in QED, but we will not do it here, and we will just quote the final result, for the gluon to gluon splitting function.

The resulting evolution equations, the *Altarelli-Parisi* evolution equations, are the same as the Gribov-Lipatov evolution equations, just replacing the electron e with quark f, $\bar{e}$ with

$\bar{f}$ and γ with g, and α with $\alpha_s(Q^2)$ (which now also runs with the scale Q^2), so

$$\begin{aligned}
\frac{d}{d\log Q}f_g(x,Q) &= \frac{\alpha_s(Q^2)}{\pi}\int_x^1\frac{dz}{z}\left\{P_{g\leftarrow q}(z)\sum_f\left[f_f\left(\frac{x}{z},Q\right)+f_{\bar{f}}\left(\frac{x}{z},Q\right)\right]+P_{g\leftarrow g}(z)f_g\left(\frac{x}{z},Q\right)\right\}\\
\frac{d}{d\log Q}f_f(x,Q) &= \frac{\alpha_s(Q^2)}{\pi}\int_x^1\frac{dz}{z}\left\{P_{q\leftarrow q}(z)f_f\left(\frac{x}{z},Q\right)+P_{q\leftarrow g}(z)f_g\left(\frac{x}{z},Q\right)\right\}\\
\frac{d}{d\log Q}f_{\bar{f}}(x,Q) &= \frac{\alpha_s(Q^2)}{\pi}\int_x^1\frac{dz}{z}\left\{P_{q\leftarrow q}(z)f_{\bar{f}}\left(\frac{x}{z},Q\right)+P_{q\leftarrow g}(z)f_g\left(\frac{x}{z},Q\right)\right\}. \qquad (18.44)
\end{aligned}$$

The only changes are in the splitting functions. The quark to quark and quark to gluon splitting functions come from diagram (a), so get an extra factor of 4/3, the gluon to quark splitting function comes from diagram (b), so get an extra factor of 1/2, and the gluon to gluon splitting function is new, as it comes from diagram (c). We then obtain

$$\begin{aligned}
P_{q\leftarrow q}(z) &= \frac{4}{3}\left[\frac{1+z^2}{(1-z)_+}+\frac{3}{2}\delta(1-z)\right]\\
P_{g\leftarrow q}(z) &= \frac{4}{3}\frac{1+(1-z)^2}{z}\\
P_{q\leftarrow g}(z) &= \frac{1}{2}\left[z^2+(1-z)^2\right]\\
P_{g\leftarrow g}(z) &= 6\left[\frac{1-z}{z}+\frac{2}{(1-z)_+}+z(1-z)+\left(\frac{11}{2}-\frac{N_f}{18}\right)\delta(1-z)\right]. \qquad (18.45)
\end{aligned}$$

Important concepts to remember

- At next to leading order in α_s, Bjorken scaling is violated, by Q^2 dependence of the parton distribution functions, now $f_f(x,Q)$.
- One can describe the situation first in QED, where the DIS, with a photon emitted from an electron, interacts with X, is understood in the Weizsacker-Williams equivalent photon approximation, as being due to a photon distribution function inside the electron, so $\sigma(e^-X\to e^-Y)=\int_0^1 dz f(z)\sigma(\gamma X\to Y)$.
- Similarly, the interaction of an electron with X via real γ emission is described as due to an electron distribution function inside the electron.
- Then the (continuous) evolution of f_γ, through multiple emissions of γ from an e^- or e^+ line, is due to the splitting function $P_{\gamma\to e}(z)$, as well as to a $\gamma\to\gamma$ transition via $P_{\gamma\leftarrow\gamma}$, the evolution of f_e, through multiple emissions of γ from an e^- line, or pair creation e^+e^- from an γ line is due to the splitting functions $P_{e\leftarrow e}$ and $P_{e\leftarrow\gamma}$, leading to the Gribov-Lipatov equations.
- The Altarelli-Parisi equations in QCD are obtained by importing the QED calculation (Gribov-Lipatov) to QCD, changing e^- with q_f, e^+ with $\bar{q}_f$, and γ with g, with some extra color factors for the splitting functions, and with an extra diagram for a gluon to split into two gluons giving a new splitting function $P_{g\leftarrow g}(z)$.

Further reading: See chapter 17.5 in [3].

Exercises, Lecture 18

1) Verify that the Gribov-Lipatov evolution equations for QED imply that the normalization conditions for the $e^+e^-\gamma$ distribution functions,

$$\begin{aligned}&\int_0^1 dx[f_e(x,Q) - f_{\bar{e}}(x,Q) = 1\\&\int_0^1 dx\; x[f_e(x,Q) + f_{\bar{e}}(x,Q) + f_\gamma(x,Q)] = 1\end{aligned} \qquad (18.46)$$

are still satisfied at all Q, for the given boundary conditions at $Q = m$.

2) Consider

$$f_u(x,Q_0) = f_g(x,Q_0) = \frac{1-x}{a(x+\epsilon)} \qquad (18.47)$$

at $Q = Q_0$ and

$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 + \log\frac{Q^2}{Q_0^2}}. \qquad (18.48)$$

Find the first correction to $f_u(x,Q)$ from the Altarelli-Parisi equations.

19 Lecture 19. The Wilson loop and the Makeenko-Migdal loop equation. Order parameters; 't Hooft loop.

In the previous 2 lectures we have learned how to deal with the fact that QCD at low energy is strongly coupled, so nonperturbative, by introducing parton distribution functions, and using perturbation theory on top of that. In this lecture however, we will learn how to probe truly nonperturbative physics. The most widely used tool for nonperturbative studies is the *Wilson loop.* It satisfies the Makeenko-Migdal loop equation, and defines an order parameter for a phase transition in QCD. These are the subjects of this lecture.

The Wilson loop is defined by external quarks, i.e. infinitely heavy ($m \to \infty$) sources for the gauge field, in a pure glue theory (YM). These are quarks that are not dynamical (i.e. are not in the path integral, or equivalently have decoupled because of the infinite mass).

We define the path ordered exponential on a path $\mathcal{P}$ from x to y, namely the *Wilson line*

$$\Phi(y,x;\mathcal{P}) = P\exp\left\{\int_x^y A_\mu(\xi)d\xi\right\} \equiv \lim_{n\to\infty}\prod_n e^{iA_\mu(\xi_n^\mu-\xi_{n+1}^\mu)}. \tag{19.1}$$

Note that $A_\mu(x) = A_\mu^a(x)T_a$ at different points do not commute in between them, so the exponential needs to be defined with an ordering along the path, i.e. as a limit $n\to\infty$ of a product of terms ordered along the path by an index n.

Abelian case

Consider first an abelian gauge fields A_μ, transforming by

$$\delta A_\mu = \partial_\mu \chi. \tag{19.2}$$

Then the objects in the definition of $\Phi(y,x;\mathcal{P})$ transforms as

$$e^{iA_\mu d\xi^\mu} \to e^{iA_\mu d\xi^\mu + i\partial_\mu\chi d\xi^\mu} = e^{iA_\mu d\xi^\mu}e^{i\chi(x+dx)-i\chi(x)}. \tag{19.3}$$

It follows that the Wilson line transforms as

$$\begin{aligned}\Phi(y,x;\mathcal{P}) = \prod_x e^{iA_\mu d\xi^\mu} &\to \prod_x\left(e^{iA_\mu d\xi^\mu}e^{i\chi(x+dx)-i\chi(x)}\right) = e^{i\chi(y)}\left(\prod_x e^{iA_\mu d\xi^\mu}\right)e^{-i\chi(x)}\\ &= e^{i\chi(y)}\Phi(y,x;\mathcal{P})e^{-i\chi(x)}.\end{aligned} \tag{19.4}$$

Then, when acting on a charged complex scalar field $\phi(x)$, transforming under the gauge transformation as

$$\phi(x) \to e^{i\chi(x)}\phi(x)\;, \tag{19.5}$$

the Wilson line gives

$$\Phi(y,x;\mathcal{P})\phi(x) \to e^{i\chi(y)}\Phi(y,x;\mathcal{P})e^{-i\chi(x)}e^{i\chi(x)}\phi(x) = e^{i\chi(y)}\left(\Phi(y,x;\mathcal{P})\phi(x)\right)\;, \tag{19.6}$$

i.e., it defines *parallel transport along the path $\mathcal{P}$ from x to y.* Parallel transport means that the properties of the object are preserved, just translated to a different point.

For a closed curve, $\mathcal{P}=\mathcal{C}$ for $y=x$, we have

$$\Phi(x,x;\mathcal{C}) \to e^{i\chi(x)}\Phi(x,x;\mathcal{C})e^{-i\chi(x)} = \Phi(x,x;\mathcal{C}) \;, \tag{19.7}$$

so the object is gauge invariant, i.e. it is a potential observable.

Nonabelian case

For a nonabelian gauge field, transforming as (for coupling $g=1$)

$$A_\mu(x) \to \Omega(x)A_\mu(x)\Omega^{-1}(x) - i(\partial_\mu\Omega)\Omega^{-1} \;, \tag{19.8}$$

with infinitesimal transformation for $\Omega(x)=e^{i\chi(x)}$,

$$\delta\Omega = D_\mu\chi \;, \tag{19.9}$$

the basic objects whose products define the Wilson line transform as

$$\begin{aligned} e^{iA_\mu d\xi^\mu} &\simeq 1 + iA_\mu d\xi^\mu \to 1 + \Omega(A_\mu d\xi^\mu)\Omega^{-1} + d\xi^\mu(\partial_\mu\Omega)\Omega^{-1} \\ &= \left[e^{i\chi(x)}(1+iA_\mu d\xi^\mu) + d\xi^\mu\partial_\mu e^{i\chi(x)}\right]e^{-i\chi(x)} \\ &\simeq e^{i\chi(x+dx)}(1+iA_\mu d\xi^\mu)e^{-i\chi(x)} \simeq e^{i\chi(x+dx)}e^{iA_\mu d\xi^\mu}e^{-i\chi(x)} + \mathcal{O}(dx^2) \;, \end{aligned} \tag{19.10}$$

where in the last line we have ignored quadratic terms in dx. We thus obtain that the Wilson line transforms as

$$\Phi(y,x;\mathcal{P}) \to e^{i\chi(y)}\Phi(y,x;\mathcal{P})e^{-i\chi(x)} \;, \tag{19.11}$$

which is formally the same as in the abelian case, just that there the order of terms did not matter, we wrote it this way to suggest the form in the nonabelian case, but in the nonabelian case the order matters.

In particular, for closed curves $y=x$, $\mathcal{P}=\mathcal{C}$, we cannot cancel the exponentials, and the Wilson line is gauge covariant, not invariant,

$$\phi(x,x;\mathcal{C}) \to e^{i\chi(x)}\Phi(x,x;\mathcal{C})e^{-i\chi(x)} \neq \Phi(x,x;\mathcal{C}). \tag{19.12}$$

But we can easily construct a gauge invariant object, the *Wilson loop*, by taking the trace (normalized with a $1/N$ since there are N terms inside the trace for $SU(N)$),

$$W[\mathcal{C}] = \frac{1}{N}\operatorname{Tr}\Phi(x,x;\mathcal{C}). \tag{19.13}$$

Note that this object is gauge invariant, and independent of the point x (there is nothing special about x in the transformation law for W).

In the abelian case, the Wilson loop can be written in a manifesly gauge invariant way through the use of the Stokes theorem, as

$$\Phi_{\mathcal{C}} = e^{i\int_{\mathcal{C}=\partial S}A_\mu dx^\mu} = e^{i\int_S F_{\mu\nu}d\Sigma^{\mu\nu}}. \tag{19.14}$$

In the nonabelian case, there are corrections to an explicitly invariant form. Consider a small square of sides a, in the $(\mu\nu)$ plane, so

$$\Phi_{\Box\mu\nu} = e^{ia^2F_{\mu\nu}} + \mathcal{O}(a^4). \tag{19.15}$$

Then the Wilson loop,

$$W_{\Box\mu\nu} = \frac{1}{N}\,\mathrm{Tr}[\Phi_{\Box\mu\nu}] \simeq 1 - \frac{a^4}{2N}\,\mathrm{Tr}[F_{\mu\nu}F_{\mu\nu}] + \mathcal{O}(a^6)\ , \tag{19.16}$$

where there is no sum over $\mu\nu$. The nontrivial object above is explictly gauge invariant (up to a^6 terms), and moreover, by summing over $\mu\nu$ we obtain the kinetic term in the action. This is an example of why the Wilson loop contains all nonperturbative information from the gauge theory: the action, that defines the theory (in the case of the pure gauge theory, we only have the kinetic term), appears in the expansion of the Wilson loop.

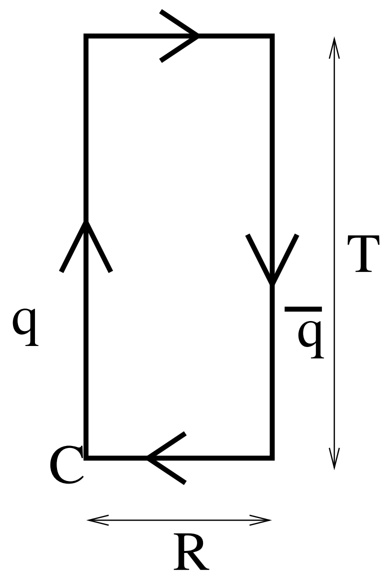

Figure 55: Wilson loop contour for the quark-antiquark potential.

The quantity that is the most studied from the Wilson loop is the *quark-antiquark potential*, measured for infinitely massive quarks (external quarks, fixed). Therefore we consider a contour $\mathcal{C}$ as in Fig.55, in the shape of a very long rectangle, made up from two long parallel lines in the time direction, of length T, one for a quark and one for an antiquark, situated at a distance R from each other, and connected by segments of length R.

Then it can be proved rigorously that the vacuum expectation value (VEV) of $W[\mathcal{C}]$ has the property that at large $T \to \infty$,

$$\langle W[\mathcal{C}]\rangle_0 \propto e^{-TV_{q\bar{q}}(R)}. \tag{19.17}$$

A simple (but not too rigorous) way to understand this is: add to the theory the infinitely heavy quarks, therefore appearing only as sources. The potential for the quark is $eA_0(x(q))$, and correspondingly for the antiquark $-eA_0(x(\bar{q}))$. Together, we obtain the source term to be added to the action, $\int d^4x j^\mu(x)A_\mu(x) = \int dt[eA_0(x(q)) - eA_0(x(\bar{q}))]$, where $j^0(x) = e\delta^{(3)}(x - x(q))$ is the quark current. This is indeed the object in the exponent of the Wilson loop, as it is $\simeq e\oint_{\mathcal{C}} A_\mu d\xi^\mu$. From the interpretation as a potential however, for a constant quark-antiquark potential, we add to e^{iS} a term $e^{iTV_{q\bar{q}}}$, so the VEV of the Wilson loop does indeed go like $e^{iTV_{q\bar{q}}(R)}$, or after a Wick rotation to Euclidean space, as $e^{-TV_{q\bar{q}}}$.

Area law and perimeter law

For a *confining gauge theory*, the potential behaves at large distances as

$$V_{q\bar{q}}(R) \sim \sigma R\ , \tag{19.18}$$

where σ is called the *QCD string tension.* A linear potential means a constant force, so we cannot pull apart the quark from the antiquark, so they are confined. Confinement however also refers to the confinement of the electric flux lines inside a flux tube between q and $\bar{q}$ of almost constant cross section, instead of spreading out all over space. The flux line density is proportional to the energy density, since $H \sim \frac{1}{2}[\vec{E}_a^2 + \vec{B}_a^2]$, and since the cross section is constant, it means there is a total energy proportional to the length.

On the other hand, for a *conformal gauge theory*, like QED, which doesn't have a mass scale, and is in fact conformal, the potential can only be of Coulomb type,

$$V_{q\bar{q}}(R) \sim \frac{\alpha}{R}\ , \tag{19.19}$$

since in that case the Wilson loop VEV scales as

$$\langle W[\mathcal{C}]\rangle_{\rm conformal} \propto e^{-TV_{q\bar{q}}(R)} \sim e^{-\frac{\alpha T}{R}}\ , \tag{19.20}$$

as it should be in a conformal theory, since the only scale invariant characterizing $\mathcal{C}$ is T/R.

In a confining theory, we obtain

$$\langle W[\mathcal{C}]\rangle_{\rm confining} \propto e^{-\sigma TR} = e^{-\sigma {\rm Area}}. \tag{19.21}$$

This is called the *area law.* But moreover, since for $C = C_1 \cup C_2$,

$$W[C = C_1 \cup C_2] = W[C_1]W[C_2]\ , \tag{19.22}$$

as we can easily check, and moreover in the large N limit for an $SU(N)$ gauge group $\langle W[C_1]W[C_2]\rangle = \langle W[C_1]\rangle\langle W[C_2]\rangle$, we can extend the area law to any smooth curve C. We can approximate its area as the sum of infinitely thin rectangular contours in the T direction as in Fig.56, thus with $T/R \gg 1$, for each of which we have the area law, and obtain the area law for the total curve.

Therefore we have in fact that

$$\langle W[\mathcal{C}]\rangle_0 \propto e^{-\sigma Area[\mathcal{C}]}\ , \tag{19.23}$$

for any contour $\mathcal{C}$ in a confining theory.

In a *Higgs phase* of a gauge theory, the quarks are screened, like in a superconductor. That is, the interaction is short range (in a superconductor, the photon becomes effectively massive through the interaction with the medium, and the range is $a \sim 1/m$), so at large distances ($R > a$), the potential is constant,

$$V_{q\bar{q}}(R) \simeq {\rm const.} \equiv \mu \tag{19.24}$$

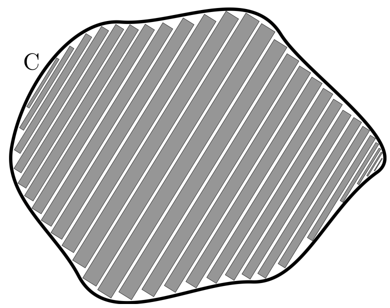

Figure 56: Wilson loop contour divided into inifnitely thin Wilson rectangles, whose orientations cancel on the neighbouring (adjacent) long lines.

Therefore the Wilson loop VEV becomes

$$\langle W[C]\rangle_{\rm Higgs} \propto e^{-\mu T} \simeq e^{-\frac{\mu}{2}L[\mathcal{C}]}\ , \tag{19.25}$$

where $L[\mathcal{C}]$ is the perimeter of the contour $\mathcal{C}$. We then obtain the *perimeter law*. Again, by the same argument as for the area law, we can extend the perimeter law for any smooth closed contour $\mathcal{C}$.

The Makeenko-Migdal loop equation.

For $SU(N)$ gauge theories at large N, the VEVs of gauge invariant operators factorize (proven by Migdal)

$$\langle \mathcal{O}_1...\mathcal{O}_n\rangle = \langle \mathcal{O}_1\rangle ...\langle \mathcal{O}_n\rangle + O\left(\frac{1}{N^2}\right). \tag{19.26}$$

That means that gauge invariant operators behave like c-numbers, not as operators, so there must be a semiclassical saddle point of the path integral that allows us to write the VEV as simply the solution at the saddle point, weighted by e^{-S} at the saddle point.

We can then infer that exists a so-called "*master field*", a colorless composite field $\Phi[A]$, with Jacobian for the gauge field

$$\left|\frac{\partial \Phi[A]}{\partial A_\mu^a}\right| \equiv e^{-N^2 J[\Phi]}. \tag{19.27}$$

In the presence of such a field, we can transform the path integral over A_μ^a into a path integral over Φ, and obtain

$$Z = \int \mathcal{D}A_\mu^a e^{-\frac{1}{4}\int d^4x (F_{\mu\nu}^a)^2} = \int \mathcal{D}\Phi \frac{1}{\left|\frac{\partial\Phi}{\partial A_\mu^a}\right|} e^{-N^2 S[\Phi]} = \int \mathcal{D}\Phi e^{-N^2(S-J)}. \tag{19.28}$$

The saddle point of this is then

$$\frac{\delta S}{\delta \Phi} = \frac{\delta J}{\delta\Phi} \to \frac{\delta S}{\delta A_\mu^a} = -(\nabla_\mu F_{\mu\nu})^a = \frac{\delta J}{\delta A_\nu^a}\ , \tag{19.29}$$

which is called the master field equation. We note that with respect to the classical field equation, we have a nonzero term on the right-hand side, coming from the variation of the Jacobian from A_μ^a to Φ.

A natural guess for the master field (which is gauge invariant), that turns out to be correct, is the Wilson loop. But moreover, one can show that we can reformulate $SU(N)$ YM at any N (and thus QCD) in terms of $W[\mathcal{C}]$. Any observable is given by a sum over paths of Wilson loops.

e.g. 1. For instance, the products of two colorless vector quark currents is written as

$$\langle\bar\psi\gamma_\mu\psi(x_1)\bar\psi\gamma_\nu\psi(x_2)\rangle = \sum_{\mathcal{C}\ni x_1,x_2} J_{\mu\nu}(\mathcal{C})\langle W[\mathcal{C}]\rangle\;, \tag{19.30}$$

where the sum is over paths that pass through x_1 and x_2, as in Fig.57.

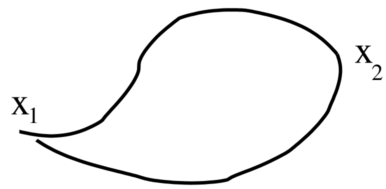

Figure 57: Paths going through x_1 and x_2 that are summed over.

e.g. 2. The connected correlator of 3 scalar quark currents,

$$\langle\bar\psi\psi(x_1)\bar\psi\psi(x_2)\bar\psi\psi(x_3)\rangle_{\text{conn.}} = \sum_{\mathcal{C}\ni x_1,x_2,x_3} J(\mathcal{C})\langle W[\mathcal{C}]\rangle\;, \tag{19.31}$$

where again the sum is over paths that pass through x_1, x_2, x_3.

If the quarks were scalars, we would have

$$J(\mathcal{C}) = e^{-\frac{m^2}{2}\tau-\frac{m^2}{2}\int_0^\tau dt\dot z_\mu^2(t)} = e^{-mL[\mathcal{C}]}\;, \tag{19.32}$$

where the thing in the exponent is Lagrangean of the quark (particle), proportional to the length of the contour.

However, for the spinor quarks, things are a bit more complicated, and we get

$$\begin{aligned}J(\mathcal{C}) &= \int \mathcal{D}K_\mu(t) P\exp\left[-\int_0^\tau (iK_\mu(t)[\dot x^\mu(t)-\gamma^\mu(t)]+m^2)\right]\\ J_{\mu\nu}(\mathcal{C}) &= \int \mathcal{D}K_\mu(t)P\left\{\gamma_\mu(t_1)\gamma_\nu(t_2)\exp\left[-\int_0^\tau (iK_\mu(t)[\dot x^\mu(t)-\gamma^\mu(t)]+m^2)\right]\right\}\end{aligned} \tag{19.33}$$

where t_1 and t_2 are times of x_1, x_2.

Path and area derivatives

To write down the loop equations, we need to define some geometric objects called the path and area derivatives.

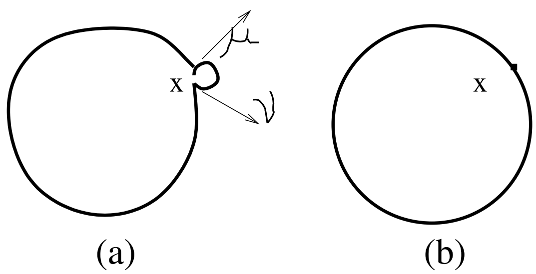

Figure 58: Diagrams for the definition of the area derivative: (a) $\mathcal{C}_{\delta\sigma_{\mu\nu}}$. (b) $\mathcal{C}_x$.

The *area derivative* of a function of a closed contour $\mathcal{C}$ is defined as follows. Consider the contour $\mathcal{C}$ with a point x singled out, $\mathcal{C}_x$, as in Fig.58b, and the same with an extra loop in the plane $(\mu\nu)$ and of area $\delta\sigma_{\mu\nu}$ at point x, $\mathcal{C}_{\delta\sigma_{\mu\nu}}$, as in Fig.58b. Then the area derivative is

$$\frac{\delta\mathcal{F}(\mathcal{C})}{\delta\sigma_{\mu\nu}(x)} \equiv \frac{1}{\delta\sigma_{\mu\nu}(x)}\left[\mathcal{F}(\mathcal{C}_{\delta\sigma_{\mu\nu}}) - \mathcal{F}(\mathcal{C}_x)\right]. \tag{19.34}$$

Here $\delta\sigma_{\mu\nu} = dx_\mu \wedge dx_\nu$. For the path derivative, consider the contour $\mathcal{C}_{\delta x_\mu}$ where at point x, the contour is shifted along δx_μ for a length δx_μ, in the μ direction, and then comes back (with zero area), as in Fig.59a, so

$$\partial^x_\mu \mathcal{F}(\mathcal{C}_x) = \frac{1}{\delta x_\mu}\left[\mathcal{F}(\mathcal{C}_{\delta x_\mu}) - \mathcal{F}(\mathcal{C}_x)\right]. \tag{19.35}$$

Note that the standard variational derivative can be written as a combination of the path and area derivatives as

$$\frac{\delta}{\delta x_\mu(\sigma)} = \dot{x}_\nu(\sigma)\frac{\delta}{\delta\sigma_{\mu\nu}(x(\sigma))} + \sum_{i=1}^m \partial^{x_i}_\mu \delta(\sigma - \sigma_i). \tag{19.36}$$

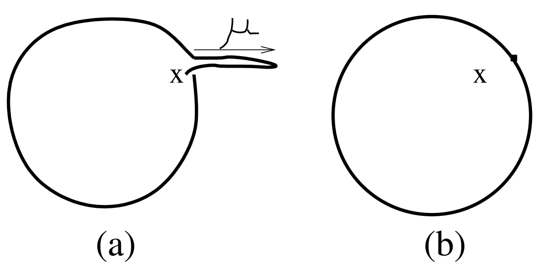

Figure 59: Diagrams for the definition of the path derivative: (a) $\mathcal{C}_{\delta x_\mu}$. (b) $\mathcal{C}_x$.

Makeenko-Migdal loop equation.
The loop equation is then written as

$$\partial^x_\mu \frac{\delta}{\delta\sigma_{\mu\nu}(x)}\langle W[\mathcal{C}]\rangle = \lambda \oint_{\mathcal{C}} dy_\nu \delta^{(D)}(x-y)\langle W[\mathcal{C}_{yx}]\rangle\langle W[\mathcal{C}_{xy}]\rangle\ , \tag{19.37}$$

for $N\to\infty$, and where $\lambda = g^2N$ is the 't Hooft coupling. Moreover, here $\mathcal{C}=\mathcal{C}_{xy}\cup\mathcal{C}_{yx}$ is closed, whereas $\mathcal{C}_{yx}$ goes from x to y (very close by points) in a long route on one side, and $\mathcal{C}_{xy}$ goes from y to x in a long route on another side, as in Fig.60. Note that this is a single equation for the Wilson loop VEV. This equation is the analogue of the Dyson-Schwinger equation.

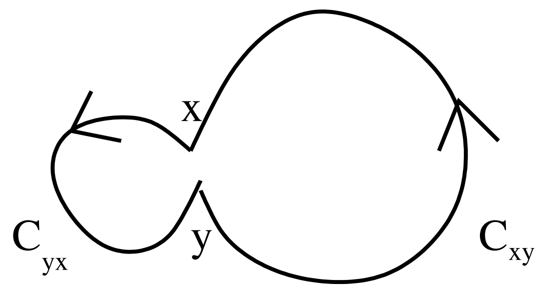

Figure 60: Diagram for the Makeenko-Migdal equation: $\mathcal{C}_{xy}$ and $\mathcal{C}_{yx}$.

At finite N, we can write a version of the above, but where now the equation does not close, i.e. it is not an equation for a single object. It is

$$\partial_\mu^x \frac{\delta}{\delta\sigma_{\mu\nu}(x)}\langle W[\mathcal{C}]\rangle = \lambda \oint_{\mathcal{C}} dy_\nu \delta^{(D)}(x-y)\left[\langle W[\mathcal{C}_{yx}]W[\mathcal{C}_{xy}]\rangle - \frac{1}{N^2}\langle W[\mathcal{C}]\rangle\right]. \qquad (19.38)$$

Therefore we see that there is an extra term with $1/N^2$, but more importantly, the right hand side at finite N does not factorize, and we get the VEV of a product of Wilson loops. Then we must write down an equation for the product of two Wilson loops, that will depend on the VEV of the product of 3 Wilson loops, etc., obtaining an infinite chain of coupled differential equations.

Order parameters.

The Wilson loop is an order parameter in the sense of Landau's theory of second order phase transitions, namely an object that has a nonzero VEV in an ordered phase, and a zero VEV in a disordered phase, since in a confining (disordered) phase,

$$\langle W[\mathcal{C}]\rangle \sim e^{-\sigma Area} \to 0\ , \qquad (19.39)$$

whereas in a Higgs (ordered) phase, we have

$$\langle W[\mathcal{C}]\rangle \sim e^{-\mu Perimeter} \neq 0. \qquad (19.40)$$

Of course, there is a slight abuse of notation, since in the limit of large contour, really both VEVs go to zero, but the confining one goes much faster. In fact, we can of course multiply everything with $e^{+\mu Perimeter}$ to make it more precise, but it is still not perfectly defined.

We can define however another type of observable, a *disorder parameter*, i.e. one that has nonzero VEV in the disordered phase and zero in the ordered phase. It is an operator dual to the Wilson loop (in the sense that its properties are opposite to the Wilson loop), called the *'t Hooft operator* $T[\mathcal{C}]$. Unlike the Wilson loop, it can only be defined in the case where all the scalars are invariant under the center $\mathbb{Z}_N$ of $SU(N)$ (the center of a group is the subgroup that commutes with all the other elements).

The 't Hooft loop is defined rather abstractly as follows. Consider a gauge transformation $\Omega^{[C]}$ that is singular along the curve C. If another curve C' winds through C with a linking number n (e.g. two links of a chain have linking number 1), and C' is parametrized by $\theta \in [0, 2\pi]$, then

$$\Omega^{[C]}(2\pi) = \Omega^{[C]}(0)e^{\frac{2\pi i n}{N}}. \tag{19.41}$$

Then the 't Hooft loop operator $T[C']$ is defined by the relation

$$W[C]T[C'] = T[C']W[C]e^{\frac{2\pi i n}{N}}. \tag{19.42}$$

As one can guess from the statement that the 't Hooft loop is dual to the Wilson loop, in a Higgs (ordered) phase we have

$$\langle T[\mathcal{C}]\rangle \sim e^{-\sigma' Area} \to 0 \;, \tag{19.43}$$

i.e., the area law, whereas in a confining (disordered) phase, we have

$$\langle T[\mathcal{C}\rangle \sim e^{-\mu Perimeter} \neq 0. \tag{19.44}$$

Note however, that 't Hooft showed that there are also *mixed phases*, where

$$\langle W[C]\rangle \sim e^{-\sigma Area}; \quad \langle T[C]\rangle \sim e^{-\sigma' Area}. \tag{19.45}$$

Polyakov loop

An important sub-case of Wilson loop, that has its own name, is called the *Polyakov loop*. It is a Wilson loop in the case of a QFT at finite temperature, i.e. described in Euclidean space by periodic time, with periodicity $\beta = 1/\text{Temperature}$. The Polyakov loop is a loop where the rectangular contour wraps once along the periodic time direction.

In this case we obtain a better understanding of why $W[C]$ is an order parameter. Indeed, in this case, the length in time $T = \beta$ is fixed and finite. That means that now, for infinite contour C, in the Higgs (ordered) phase,

$$\langle W[C]\rangle \sim e^{-\mu T} = \text{constant} \neq 0 \;, \tag{19.46}$$

and in the confining (disordered) phase,

$$\langle W[C]\rangle \sim e^{-(\sigma T)R} \to 0. \tag{19.47}$$

Important concepts to remember

- Wilson loops characterize the behaviour of external quarks (infinitely massive probe quarks, not dynamical) in gauge theories.

- In an abelian theory, $\Phi = exp[i \oint A_\mu dx^\mu]$ is gauge invariant and defines parallel transport of charged scalar fields, and in a nonabelian gauge theory, $W[C] = 1/N \,\mathrm{Tr}[P \exp\left\{i \oint A_\mu dx^\mu\right\}]$ is gauge invariant and defines parallel transport of charged scalar fields.

- The Wilson loop contains all information about observables of the gauge theory. Its first nontrivial term in the expansion on a square (plaquette) is the kinetic term of the gauge action.

- The VEV of the Wilson loop on a rectangular contour infinitely long in the time direction defines the quark-antiquark potential by $\langle W[C]\rangle \propto e^{-TV_{q\bar{q}}(R)}$ as $T \to \infty$.

- For a confining theory (a confining phase), the potential is linear $V_{q\bar{q}}(R) = \sigma R$, so we obtain the area law, whereas for a Higgs phase, the potential is constant, so we obtain the perimeter law. For a conformal theory like QED, we obtain a Coulomb potential $V_{q\bar{q}}(R) = \alpha/R$.

- VEVs of gauge invariant observables factorize in the large N limit of $SU(N)$ YM, so there is a gauge invariant, composite master field Φ. In fact the Wilson loop has its properties, and we can define (even at finite N) gauge theory observables in terms of sums over paths with the desired operator insertions.

- The Wilson loop at $N \to \infty$ satisfies the Makeenko-Migdal loop equation (the equation closes on VEVs of $W[C]$), and at finite N we obtain an infinite set of coupled equations for VEV of products of Wilson loops.

- The Wilson loop is an order parameter, and its dual, the 't Hooft loop, is a disorder parameter.

- The Polyakov loop is a Wilson loop at finite temperature, where the infinite lines in the time directions now wrap once the periodic time.

Further reading: See chapter 15.3 in [3].

Exercises, Lecture 19

1) Consider a circular Wilson loop of radius R in Euclidean space. If the theory is confining, how will the Wilson loop VEV scale with R? How about if it is conformal (like QED)?

2) Check that the confining result for the circle satisfies the Makeenko-Migdal loop equation.

20 Lecture 20. IR divergences in QED.

In this lecture we start the study of IR divergences, focusing on the example of QED.

Collinear divergences

We already saw in lecture 2 that when we have massless particles in the theory, we have IR divergences in the loop diagrams. For instance, in the one-loop diagram with two $n+2$-point vertices, two propagators and $2n$ external lines in Fig.4, with Feynman diagram

$$\frac{\lambda^2}{2}\int\frac{d^Dq}{(2\pi)^D}\frac{1}{q^2+m^2}\frac{1}{(q-P)^2+m^2}\ , \tag{20.1}$$

where $P=\sum_i p_i$ is the total external momentum at each of the two vertices, if $m^2=0$ AND $P^2=0$, the integral becomes

$$\frac{\lambda^2}{2}\int\frac{d^Dq}{(2\pi)^D}\frac{1}{q^2}\frac{1}{q^2-2q\cdot P}\ , \tag{20.2}$$

so is divergent in $D=4$, in the angular integration region where $\hat{q}\cdot\hat{P}=0$, since then

$$\sim\frac{\lambda^2}{2}\int d\Omega\int dq\frac{q}{q^2-2qP\hat{q}\cdot\hat{P}}\propto\int\frac{dq}{q}. \tag{20.3}$$

Note that this divergence appears only for massless external states, since if $P^2\neq 0$, we do not have an IR divergence. In turn, if $n>1$, that means that the external particles must be *collinear*, $p_i\propto p_j$, such as to have $P^2=0$. Note also that the divergence is due to a virtual "photon" (massless particle) of momentum q being collinear with the (set of) external particle(s) $P=\sum_i p_i$, since $q\cdot P=0\Leftrightarrow q$ is parallel with P ($P^2=0$). This type of IR divergence is then called a *collinear IR divergence*. It will be cancelled by an IR divergence in the amplitude to emit a (real) "photon" (massless particle) collinear with an external line from each it is emitted. In conclusion, a collinear IR divergence is due to a virtual or real "photon" being collinear with a massless external line. For the existence of such a divergence, we need to have a massless external state coupling to a massless internal loop, i.e. to have self-interactions of massless states, since we needed $P^2=0$, but also $q^2=0$ and $(q-P)^2=0$. This will happen in QCD, where gluons are self-interacting, but does not happen in QED.

In dimensional regularization, using the result (3.35) at $m=0$ (from lecture 3), we can calculate the diagram as

$$I=\frac{\lambda^2}{2}\frac{\Gamma\left(2-\frac{D}{2}\right)}{(4\pi)^{\frac{D}{2}}}P^{D-4}\int_0^1 d\alpha\,[\alpha(1-\alpha)]^{\frac{D}{2}-1}\propto\lambda^2\frac{2}{\epsilon}\left(\frac{P^2}{\mu^2}\right)^{-\frac{\epsilon}{2}}\propto\lambda^2\left[\frac{2}{\epsilon}-\log\frac{P^2}{\mu^2}\right]\ , \tag{20.4}$$

where as usual $D=4-\epsilon$, we have introduced a dimensional transmutation parameter μ, and finally we have expanded in ϵ to obtain a term logarithmically divergent in $\mu\to 0$ (thought of as an IR cut-off).

Soft divergences

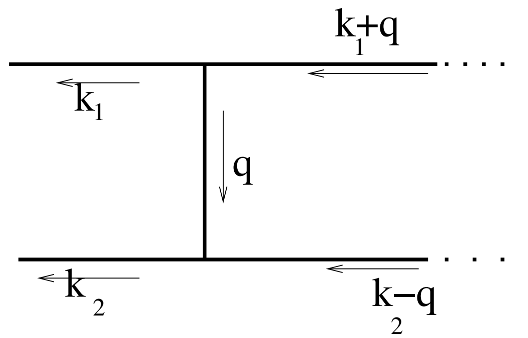

Figure 61: Soft Divergence diagram piece.

Consider a part of a larger one-loop diagram with two consecutive external states, and take them to have mass m. They are taken to be continued into the loop, so two outward extending propagators, with momenta k_1+q and k_2-q will have the same mass m. In between these two lines, there is a massless propagator with momentum q. We have here in mind the application to QED, or rather the simpler version of massive scalar QED (charged complex scalar) where the massive line would be a scalar, and the massless one a photon, but we can also consider the massless case $m=0$, and then we have in mind a theory of massless scalars. This part of the diagram, in Fig.61, will give a contribution

$$\int \frac{d^Dq}{(2\pi)^D}\frac{1}{(q+k_1)^2+m^2}\frac{1}{q^2}\frac{1}{(q-k_2)^2+m^2} = \int \frac{d^Dq}{(2\pi)^D}\frac{1}{q^2+2q\cdot k_1}\frac{1}{q^2}\frac{1}{q^2-2q\cdot k_2}\ , \quad (20.5)$$

which we see is independent of the mass m of the external states, appears at general k_1, k_2 (not necessarily collinear), and moreover the divergence is present independently of the orientation of the "photon", i.e. of whether $q\cdot k_1 = 0$ and $q\cdot k_2 = 0$ or not (which equality is true for $m^2=0$, so that $k_1^2 = k_2^2 = 0$, AND $q \ll k_1$ and $q \ll k_2$). The only source of this divergence is the fact that the "photon" q is "soft" (small energy), i.e. $q^2 \sim 0$ and moreover $|\vec{q}|^2$ small.

This divergence, for soft virtual massless particles (in a loop), and a corresponding one for soft emitted (real) massless particles, is called a *soft divergence* and is present in any theory with massless particles interacting with something, so is present both in QED and in QCD.

In nonabelian gauge theories (YM) then, we have both soft and collinear IR divergences, and correspondingly in dimensional regularization at one-loop we have a factor of $1/\epsilon$ for each: we saw that there was an $1/\epsilon$ for collinear divergences, and there is another one for soft ones. In total, in a *planar* one-loop diagram *for each pair of consecutive momenta*, k_i and k_{i+1}, with $s_{i,i+1} \equiv (k_i+k_{i+1})^2$, we have a divergent factor of

$$\sim \frac{1}{\epsilon^2}\left(\frac{-s_{i,i+1}}{\mu^2}\right)^\epsilon \simeq \frac{1}{\epsilon^2}\left[1+\epsilon\log\frac{-s_{i,i+1}}{\mu^2}+\frac{\epsilon^2}{2}\log^2\frac{-s_{i,i+1}}{\mu^2}+...\right] \quad (20.6)$$

We see that the term divergent as $\mu\to 0$ is $\sim \log^2(-s_{i,i+1}/\mu^2)$ in the nonabelian case. In the QED case (abelian), when we have only soft divergences, at one-loop we have in dimensional regularization a term $\sim 1/\epsilon(-s_{i,i+1}/\mu^2)^\epsilon \sim 1/\epsilon + \log(s_{i,i+1}/\mu^2)$, so the term divergent as $\mu \to 0$ is $\log(-s_{i,i+1}/\mu^2)$. A useful IR regularization that will be used in the following is to introduce a photon mass $\mu_{\rm ph}$. When translating dimensional regularization results into photon mass regularization, we just drop the $1/\epsilon$ terms and keep only the terms divergent for $\mu \to 0$, replacing μ with $\mu_{\rm ph}$. Therefore in the QED case, we expect an IR divergence of order $\log(q^2/\mu_{\rm ph}^2)$, and we will see that this is what we obtain.

For completeness, note that at L loops, we have in nonabelian gauge theory a leading divergent term of $\mathcal{O}(1/\epsilon^{2L})$.

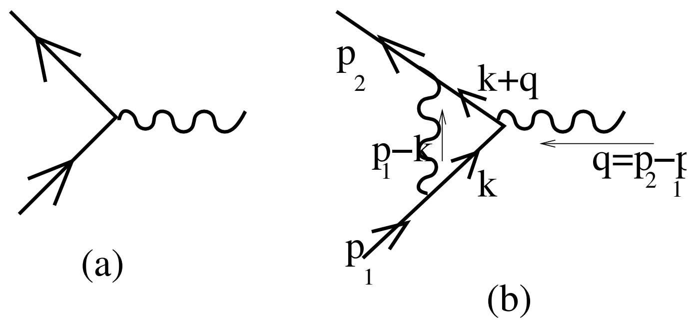

Figure 62: Soft Divergent vertex diagram. (a) Zeroth order diagram. (b) One-loop diagram.

QED vertex IR divergence

Consider the full quantum mechanical QED vertex $\Gamma^\mu_{\alpha\beta}$, with a fermion with momentum p_1 going in, a fermion with momentum p_2 coming out, and a photon with momentum $q = p_2 - p_1$ coming in, as in Fig.62a. Then, in general, by Lorentz invariance we should have

$$\Gamma^\mu = A\gamma^\mu + B(p_1^\mu + p_2^\mu) + C(p_2^\mu - p_1^\mu). \tag{20.7}$$

The vertex should satisfy the Ward-Takahashi identity,

$$q_\mu \Gamma^\mu = (p_2 - p_1)_\mu \Gamma^\mu = A(p_2 - p_1)_\mu \gamma^\mu + C(p_2 - p_1)^2 = 0\ , \tag{20.8}$$

where we have used $p_2^2 - p_1^2 = 0$ on-shell. Since the vertex also appears in between $\bar{u}(p_2)$ and $u(p_1)$, and $(p_2 - p_1)_\mu \bar{u}(p_2)\gamma^\mu u(p_1) = 0$ on-shell (by the Dirac equation for $u(p_1)$ and $\bar{u}(p_2)$), inside physical amplitudes we can ignore A, and the Ward-Takahashi identity just says that $C = 0$.

On the other hand, we have the Gordon identity, sometimes written as two equations,

$$\begin{aligned} m\bar{u}(p_2)\gamma^\mu u(p_1) &= p_1^\mu \bar{u}(p_2)u(p_1) - i\bar{u}(p_2)\sigma^{\mu\nu}p_{1\nu}u(p_1) \\ m\bar{u}(p_2)\gamma^\mu u(p_1) &= p_2^\mu \bar{u}(p_2)u(p_1) + i\bar{u}(p_2)\sigma^{\mu\nu}p_{2\nu}u(p_1)\ , \end{aligned} \tag{20.9}$$

and sometimes as the average of the two equations above, giving

$$\bar{u}(p_2)\gamma^\mu u(p_1) = \bar{u}(p_2)\left[\frac{p_2^\mu + p_1^\mu}{2m} + \frac{i\sigma^{\mu\nu}q_\nu}{2m}\right]u(p_1). \tag{20.10}$$

It means that we can swap the term with B (with $(p_1+p_2)^\mu$) for a contribution to the term with A (with γ^μ) and another with $\sigma^{\mu\nu}q_\nu$, i.e. we can always write (in between $\bar{u}(p_2)$ and $u(p_1)$)

$$\Gamma^\mu(p_2,p_1)=\gamma^\mu F_1(q^2)+\frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2(q^2)\ , \tag{20.11}$$

where we have defined the two structure functions $F_1(q^2)$ and $F_2(q^2)$.

We have treated the one-loop correction to Γ^μ, $\Gamma^{\mu(1)}$, in lecture 6 and lecture 8, but there we ignored the IR divergences. We will redo the calculation (some relevant steps) here, in Minkowski space, with a slightly different parametrization for the loop momentum, and considering the IR divergences.

The one-loop diagram has a photon with mometum p_1-k being emitted from the p_1 fermion line, turning it into a fermion with momentum k, and being reabsorbed into the other fermion line, initially with momentum $k+q$, to turn in into p_2, as in Fig.62b. Then the diagram gives

$$\begin{aligned}\bar{u}(p_2)\Gamma^{\mu(1)}(p_1,p_2)u(p_1) &= \int\frac{d^4k}{(2\pi)^4}\frac{-ig_{\nu\rho}}{(k-p_1)^2-i\epsilon}\bar{u}(p_2)(+e\gamma^\nu)\frac{-(\not{k}+\not{q}+im)}{(k+q)^2+m^2-i\epsilon}\times\\ &\quad\times\gamma^\mu\frac{-(\not{k}+im)}{k^2+m^2-i\epsilon}(+e\gamma^\rho)u(p_1)\\ &= +2ie^2\int\frac{d^4k}{(2\pi)^4}\frac{\bar{u}(p_2)[\not{k}\gamma^\mu(\not{k}+\not{p}_1)-m^2\gamma^\mu+2im(2k+q)^\mu]u(p_1)}{[(k-p_1)^2+i\epsilon][(k+q)^2+m^2-i\epsilon][k^2+m^2-i\epsilon]}.\end{aligned} \tag{20.12}$$

Doing the Feynman parametrization for the 3 propagators in the denominator, $\Delta_1,\Delta_2,\Delta_3$, with α_1 for k^2+m^2, α_2 for $(k+q)^2+m^2$ and α_3 for $(k-p_1)^2$, we get

$$\frac{1}{\Delta_1\Delta_2\Delta_3}=\int_0^1 d\alpha_1 d\alpha_2 d\alpha_3\delta(\alpha_1+\alpha_2+\alpha_3-1)\frac{1}{(\tilde{k}^2+F-i\epsilon)^3}\ , \tag{20.13}$$

where $\tilde{k}=k+\alpha_2 q-\alpha_3 p_1$. Using the formulas

$$\begin{aligned}&\int\frac{d^4\tilde{k}}{(2\pi)^4}\frac{\tilde{k}^\mu}{(\tilde{k}^2+F-i\epsilon)^3}=0\\ &\int\frac{d^4\tilde{k}}{(2\pi)^4}\frac{\tilde{k}^\mu\tilde{k}^\nu}{(\tilde{k}^2+F-i\epsilon)^3}=\int\frac{d^4\tilde{k}}{(2\pi)^4}\frac{\frac{1}{4}g^{\mu\nu}\tilde{k}^2}{(\tilde{k}^2+F-i\epsilon)^3}\ ,\end{aligned} \tag{20.14}$$

after some algebra, we obtain

$$\begin{aligned}\bar{u}(p_2)\Gamma^{\mu(1)}(p_1,p_2)u(p_1) &= 2ie^2\int\frac{d^4\tilde{k}}{(2\pi)^4}\int_0^1 d\alpha_1 d\alpha_2 d\alpha_3\delta(\alpha_1+\alpha_2+\alpha_3-1)\frac{2}{(\tilde{k}^2+F-i\epsilon)^3}\times\\ &\quad\times\bar{u}(p_2)\left[\gamma^\mu\left(-\frac{1}{2}\tilde{k}^2-(1-\alpha_1)(1-\alpha_2)q^2+(1-4\alpha_3+\alpha_3^2)m^2\right)\right.\\ &\quad\left.+\frac{i\sigma^{\mu\nu}q_\nu}{2m}(2m^2\alpha_3(1-\alpha_3))\right]u(p_1)\ ,\end{aligned} \tag{20.15}$$

where

$$F = \alpha_1\alpha_2 q^2 + (1-\alpha_3)^2 m^2. \tag{20.16}$$

One would need to UV regulate this integral, but we will ignore it here, and instead adopt a bit later on a subtraction procedure (renormalization condition). Note that the part with $\tilde{k}^2$ in the numerator, that was called $\Gamma^{(1a)}$ in lecture 6, is UV divergent, but is not of interest for us. The part without $\tilde{k}^2$ in the numerator, called $\Gamma^{(1b)}$ in lecture 8, is UV finite, and will contain the relevant IR divergences, so we are interested in it.

Performing a Wick rotation on the formula (3.30) from lecture 3, and putting $D = 4$ (for the case the integral is UV finite), giving*

$$\int \frac{d^4\tilde{k}}{(2\pi)^4}\frac{1}{(\tilde{k}^2+\Delta)^n} = \frac{i}{(4\pi)^2}\frac{1}{(n-1)(n-2)}\frac{1}{\Delta^{n-2}}\ , \tag{20.18}$$

we obtain for the IR divergent piece

$$\begin{aligned}\bar{u}(p_2)\Gamma^{\mu(1b)}u(p_1) &= \frac{\alpha}{2\pi}\int_0^1 d\alpha_1 d\alpha_2 d\alpha_3 \delta(\alpha_1+\alpha_2+\alpha_3-1)\times\\ &\times\bar{u}(p_2)\left[\gamma^\mu\frac{(1-\alpha_1)(1-\alpha_2)q^2+(1-4\alpha_3+\alpha_3^2)m^2}{F}\right.\\ &\left.+\frac{i\sigma^{\mu\nu}q_\nu}{2m}\frac{2m^2\alpha_3(1-\alpha_3)}{F}\right].\end{aligned} \tag{20.19}$$

The integral of the coefficient of γ^μ in the square brackets is $F_1^{(1b)}(q^2)$ and the integral of the coefficient of $\frac{i\sigma^{\mu\nu}q_\nu}{2m}$ is $F_2(q^2)$. F_2 will not contain IR divergences, but $F_1^{(1b)}$ will. To see that, we calculate $F_1^{(1b)}(q^2=0)$. We have

$$\begin{aligned}&\int_0^1 d\alpha_1 d\alpha_2 d\alpha_3 \delta(\alpha_1+\alpha_2+\alpha_3-1)\frac{1-4\alpha_3+\alpha_3^2}{F(q^2=0)}\\ =&\int_0^1 d\alpha_3\int_0^{1-\alpha_3} d\alpha_2\frac{-2+(1-\alpha_3)(3-\alpha_3)}{m^2(1-\alpha_3)^2}\\ =&\int^1 d\alpha_3\frac{-2}{m^2(1-\alpha_3)}+\text{finite}.\end{aligned} \tag{20.20}$$

We see that we have an IR divergence, coming from the $\alpha_3 \simeq 1$ region of integration in the last Feynman parameter. We need to regulate this IR divergence. One option would be to use dimensional regularization, and we will sketch it afterwards, but here instead we introduce a small photon mass $\mu_{\rm ph}$.

Then the photon propagator will give $[(k-p_1)^2+\mu_{\rm ph}^2-i\epsilon]^{-1}$ instead of $[(k-p_1)^2-i\epsilon]^{-1}$, and since the inverse photon propagator Δ_3 appears multiplied by α_3 in the Feynman parametrization, the effect of the regularization is to add a term $\alpha_3\mu_{\rm ph}^2$ to F.

---

*For use later on, note that the case relevant for us, of $n=3$, gives, for dimensional regularization in Minkowski space,

$$\int\frac{d^Dk}{(2\pi)^D}\frac{1}{(\tilde{k}^2+\Delta)^3} = \frac{i}{(4\pi)^{D/2}}\frac{\Gamma(3-D/2)}{\Gamma(3)}\frac{1}{\Delta^{3-\frac{D}{2}}}. \tag{20.17}$$

We need one more ingredient. For the UV regularization of the whole diagram, a simple renormalization condition is

$$\Gamma^\mu(q^2=0)=\gamma^\mu. \tag{20.21}$$

That subtracts the UV divergence in $\Gamma_\mu^{(1a)}$, contributing to $F^{(1a)}(q^2)$, but it also affects the UV-finite piece $\Gamma_\mu^{(1b)}$, specifically $F_1^{(1b)}(q^2)$, by subtracting $F_1^{(1b)}(q^2=0)$. We finally obtain

$$\begin{aligned} F_1^{(1)}(q^2)|_{\mu_{\rm ph}^2\to 0} &\simeq F_1^{(1b)}(q^2)|_{\mu_{\rm ph}^2\to 0}\simeq\frac{\alpha}{2\pi}\int_0^1 d\alpha_1 d\alpha_2 d\alpha_3\delta(\alpha_1+\alpha_2+\alpha_3-1)\times\\ &\times\left[\frac{m^2(1-4\alpha_3+\alpha_3^2)-q^2(1-\alpha_1)(1-\alpha_2)}{q^2\alpha_1\alpha_2+m^2(1-\alpha_3)^2+\mu_{\rm ph}^2\alpha_3}-\frac{m^2(1-4\alpha_3+\alpha_3^2)}{m^2(1-\alpha_3)^2+\mu_{\rm ph}^2\alpha_3}\right]. \end{aligned} \tag{20.22}$$

The IR divergence, as we saw above, comes from the $\alpha_3\simeq 1$ ($\Rightarrow \alpha_1\simeq\alpha_2\simeq 0$) region of integration in the Feynman parameters, and (since the integration in the Feynman parameters is only between 0 and 1) it appears from the denominators. Therefore we can put $\alpha_3=1$ and $\alpha_1=\alpha_2=0$ in the numerators, as well as in the regulator, so $\alpha_3\mu_{\rm ph}^2\to\mu_{\rm ph}^2$.

Therefore we have

$$\begin{aligned} F_1^{(1)}(q^2)|_{\mu_{\rm ph}^2\to 0} &\simeq \frac{\alpha}{2\pi}\int_0^1 d\alpha_3\int_0^{1-\alpha_3}d\alpha_2\left[\frac{-2m^2-q^2}{m^2(1-\alpha_3)^2+q^2\alpha_2(1-\alpha_3-\alpha_2)+\mu_{\rm ph}^2}\right.\\ &\left.-\frac{-2m^2}{m^2(1-\alpha_3)^2+\mu_{\rm ph}^2}\right]. \end{aligned} \tag{20.23}$$

With the substitution $\alpha_2=(1-\alpha_3)\xi$ and $w=1-\alpha_3$, with Jacobian w, we obtain

$$\begin{aligned} F_1^{(1)}(q^2)|_{\mu_{\rm ph}^2\to 0} &\simeq \frac{\alpha}{2\pi}\int_0^1 d\xi\frac{1}{2}\int_0^1 d(w^2)\left[\frac{-2m^2-q^2}{[m^2+q^2\xi(1-\xi)]w^2+\mu_{\rm ph}^2}-\frac{-2m^2}{m^2w^2+\mu_{\rm ph}^2}\right]\\ &= \frac{\alpha}{2\pi}\int_0^1\frac{d\xi}{24}\left[\frac{-2m^2-q^2}{m^2+q^2\xi(1-\xi)}\log\frac{m^2+q^2\xi(1-\xi)}{\mu_{\rm ph}^2}+2\log\frac{m^2}{\mu_{\rm ph}^2}\right]; \end{aligned} \tag{20.24}$$

Before we continue, we now redo this calculation in dimensional regularization. The first observation is that, with $D=4-2\epsilon$, $\epsilon>0$ gives UV regularization, but $\epsilon<0$ gives IR regularization. A typical (limit) divergence both in the UV and in the IR is a log-divergence,

$$\int\frac{d^Dk}{k^4}\propto\int_0^\infty\frac{dk}{k^{5-D}}\;, \tag{20.25}$$

and we see that the UV divergence is regulated only by $4-D=2\epsilon>0$,

$$\int^\infty\frac{dk}{k^{1+2\epsilon}}\sim\left.\frac{k^{-2\epsilon}}{-2\epsilon}\right|^\infty<\infty\;, \tag{20.26}$$

whereas the IR divergence is regulated only by $4-D=2\epsilon_{UV}=-2\epsilon_{IR}<0$,

$$\int_0\frac{dk}{k^{1-2\epsilon_{IR}}}\sim\left.\frac{k^{+2\epsilon_{IR}}}{\epsilon}\right|_0<\infty. \tag{20.27}$$

So we consider $D=4+2\epsilon$ in (20.17) and IR regularize like this instead of introducing $\mu_{\rm ph}$. Then the denominator in F_1 is now $F^{1-\epsilon}$ instead of F. But again the divergent integration region for the Feynman parameters is $\alpha_3\simeq 1\Rightarrow \alpha_1\simeq\alpha_2\simeq 0$, so we can repeat the same steps up until we do the w^2 integration, to obtain in dimensional regularization

$$\begin{aligned} F_1^{(1)}(q^2)|_{\epsilon\to 0} &\simeq \frac{\alpha}{2\pi}\int_0^1 d\xi\frac{1}{2}\int_0^1 d(w^2)\left[\frac{-2m^2-q^2}{([m^2+q^2\xi(1-\xi)]w^2)^{1-\epsilon}}-\frac{-2m^2}{[m^2w^2]^{1-\epsilon}}\right]\\ &= \frac{\alpha}{2\pi}\frac{\mu^{2\epsilon}}{\epsilon}\frac{1}{2}\int_0^1 d\xi\left[\left(\frac{m^2}{\mu^2}\right)^{\epsilon}\frac{-2-q^2/m^2}{[1+q^2\xi(1-\xi)/m^2]^{1-\epsilon}}+2\left(\frac{m^2}{\mu^2}\right)^{\epsilon}\right] \end{aligned} \quad (20.28)$$

where we have introduced a dimensional transmutation scale μ, and the overall $\mu^{2\epsilon}$ should be reabsorbed, as usual, in the redefinition of the coupling. Note that we obtain the correspodence with the photon mass regularization already suggested, since

$$\begin{aligned} \frac{1}{\epsilon}\left(\frac{m^2}{\mu^2}\right)^{\epsilon}[1+q^2\xi(1-\xi)/m^2]^{\epsilon} &\simeq \frac{1}{\epsilon}+\log\frac{m^2+q^2\xi(1-\xi)}{\mu^2}\\ \frac{1}{\epsilon}\left(\frac{m^2}{\mu^2}\right)^{\epsilon} &\simeq \frac{1}{\epsilon}+\log\frac{m^2}{\mu^2}. \end{aligned} \quad (20.29)$$

We now go back to the photon mass regularization and define

$$f_{IR}(q^2)=\int_0^1 d\xi\frac{m^2+q^2/2}{m^2+q^2\xi(1-\xi)}\ , \quad (20.30)$$

and, since generically $\log(m^2+q^2\xi(1-\xi))/\mu_{\rm ph}^2$ does not vary much, and in any case for $\mu_{\rm ph}\to 0$ it doesn't matter too much what exactly we have in the log, we can write $\log(q^2 \text{ or } m^2)/\mu_{\rm ph}^2$ and take it out of the integral in $F^{(1)}(q^2)$, and adding the tree level part $F_1^{(0)}(q^2)=1$, we finally have

$$F_1(q^2)=1-\frac{\alpha}{2\pi}f_{IR}(q^2)\log\frac{q^2\text{ or }m^2}{\mu_{\rm ph}^2}+\mathcal{O}(\alpha^2). \quad (20.31)$$

Note that in the $q^2\to\infty$ limit most of the integral in f_{IR} comes from the two endpoints, $\xi\simeq 0$ and $\xi\simeq 1$ so

$$f_{IR}(q^2)\simeq\frac{1}{2}\int_0 d\xi\frac{q^2}{q^2\xi+m^2}+\frac{1}{2}\int^1 d\xi\frac{q^2}{q^2(1-\xi)+m^2}\simeq\log\frac{q^2}{m^2}. \quad (20.32)$$

In this limit we can write $F_1^{(1)}(q^2)$ as

$$F_1^{(1)}(q^2\to\infty)\simeq-\frac{\alpha}{2\pi}\frac{1}{2}\int_0^1 d\xi\left[\frac{1}{\xi(1-\xi)+m^2/q^2}\log\frac{q^2}{\mu_{\rm ph}^2}\left[\xi(1-\xi)+\frac{m^2}{q^2}\right]-\log\frac{m^2}{\mu_{\rm ph}^2}\right]\ , \quad (20.33)$$

and after splitting $\log[(q^2/\mu_{\rm ph}^2)(\xi(1-\xi)+m^2/q^2)]$ in two logs, the second log term (with the overall minus removed),

$$\frac{\alpha}{2\pi}\frac{1}{2}\int_0^1 d\xi\frac{1}{\xi(1-\xi)+m^2/q^2}\log\left[\xi(1-\xi)+\frac{m^2}{q^2}\right]\simeq\frac{a}{2\pi}\log^2\frac{q^2}{m^2}\ll\frac{\alpha}{2\pi}\log\frac{q^2}{\mu_{\rm ph}^2}\ , \quad (20.34)$$

so can be neglected, and we obtain

$$F_1(q^2 \to \infty) \simeq 1 - \frac{\alpha}{2\pi}\log\frac{q^2}{m^2}\log\frac{q^2}{\mu_{\rm ph}^2} + \mathcal{O}(\alpha^2) = 1 - \frac{\alpha}{2\pi} f_{IR}(q^2 \to \infty)\log\frac{q^2}{\mu_{\rm ph}^2} + \mathcal{O}(\alpha^2). \quad (20.35)$$

The double log that we obtained is called a *Sudakov double log.* Note though that only one of the logs is IR divergent (as $\mu_{\rm ph}^2 \to 0$), the other one is a physical log.

In dimensional regularization, we can rewrite (20.28) as

$$F_1^{(1)}(q^2)|_{\epsilon\to 0} \simeq \frac{\alpha}{2\pi}\frac{\mu^{2\epsilon}}{\epsilon}\frac{1}{2}\int_0^1 d\xi \left[\left(\frac{q^2}{\mu^2}\right)^\epsilon \frac{-1-2m^2/q^2}{[\xi(1-\xi)+m^2/q^2]^{1-\epsilon}} + 2\left(\frac{m^2}{\mu^2}\right)^\epsilon\right] , \quad (20.36)$$

which makes it more obvious that we can always write it, after dropping the $1/\epsilon$ term, as

$$F_1^{(1)}(q^2)|_{\epsilon\to 0} \simeq -\frac{\alpha}{2\pi} f_{IR}(q^2)\log\frac{q^2}{\mu_{\rm ph}^2} , \quad (20.37)$$

where the $\log(q^2/\mu_{\rm ph}^2)$ appears from the expansion of the $(1/\epsilon)(q^2/\mu^2)^\epsilon$. Then the approximation of having $f_{IR}(q^2)$ as in (20.30) corresponds to ignoring the ϵ power of $[\xi(1-\xi)+m^2/q^2]$, since the corresponding integral is finite, so that gives an $\mathcal{O}(\epsilon)$ correction; even though by multiplication with the overall $1/\epsilon$ means we get a finite contribution also.

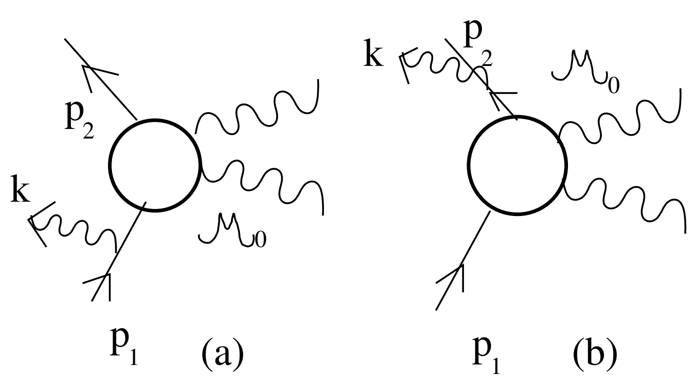

Figure 63: Photon emission diagrams cancelling the 1-loop divergences.

Cancellation of IR divergence by photon emission.

So we obtained an IR divergence of amplitudes that cannot be removed by renormalization. This would seem to be bad, but as we already explained, this is just a statement that the amplitude (or rather, the cross section obtained from it) for just this process is not something physical in a theory like QED with massless particles.

That is so since we can emit the massless particles (photons) from external lines, and if the photons have sufficiently small energies, they cannot be independently detected by any physical detector, that will always have a minimal energy cut-off. Note that in QED, we only have soft divergences, so we only need to be concerned about soft emitted photons (with energies $< E_{min}$). But in a theory like QCD, which also has collinear divergences, we also need to be concerned about emitting photons that are not soft (large energies), but instead are collinear with the particles from which they are emitted, so they cannot be distinguished from the emitted particles by a detector which has a resolution of some minimal angle θ_{min}.

Therefore we need to consider the process of emission of a soft photon off the electron lines, leading to two diagrams. Consider the tree level diagram with amplitude $\mathcal{M}_0$, in our case the tree vertex for two fermions, one with momentum p_1 (in) and one with momentum p_2 (out) and a photon with momentum $q = p_2 - p_1$. Then the first correction diagram is for a photon with momentum k to be emitted from the (initial momentum) p_1 line, turning it into $p_1 - k$, as in Fig.63a, and the second diagram is for the photon with momentum k to be emitted from the (final momentum) p_2 line, turning the line into $p_2 + k$ when it starts from $\mathcal{M}_0$, as in Fig.63b.

We also assume that the photon is soft, i.e. $|\vec{k}| \ll |\vec{p}_2 - \vec{p}_1|$, in which case the subdiagram $\mathcal{M}_0$ for the two diagrams is the same, and the same with the tree diagram, i.e.

$$\mathcal{M}_0(p_2, p_1 - k) \simeq \mathcal{M}_0(p_2 + k, p_1) \simeq \mathcal{M}_0(p_2, p_1) \equiv \mathcal{M}_0. \tag{20.38}$$

Then the amplitude for emission of a photon from one of the external lines is

$$i\mathcal{M} = e\bar{u}(p_2)\left[\mathcal{M}_0 \frac{-(\not{p}_1 - \not{k} + im)}{(p_1 - k)^2 + m^2}\gamma^\mu \epsilon^*_\mu(k) + \gamma^\mu \epsilon^*_\mu(k)\frac{-(\not{p}_2 + \not{k} + im)}{(p_2 + k)^2 + m^2}\mathcal{M}_0\right] u(p_1). \tag{20.39}$$

Because of p_1 and p_2 being on-shell, the two denominators are $-2p_1 \cdot k$ and $+2p_2 \cdot k$, respectively. Since the photon is soft, we can neglect $\not{k}$ in the numerator (small compared to $\not{p}_i$ and m). Also using the identities

$$(\not{p}_1 + im)\gamma^\mu \epsilon^*_\mu(k)u(p_1) = 2p_1^\mu \epsilon^*_\mu(k)u(p_1) + \gamma^\mu \epsilon^*_\mu(k)(-\not{p}_1 + im)u(p_1) = 2p_1^\mu \epsilon^*_\mu(k)u(p_1)\ , \tag{20.40}$$

where in the first equality we have used $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ and in the second we have used the Dirac equation written for the spinor $u(p_1)$, and the similarly proven one

$$\bar{u}(p_2)\gamma^\mu \epsilon^*_\mu(k)(\not{p}_2 + im) = \bar{u}(p_2)2p_2^\mu \epsilon^*_\mu(k)\ , \tag{20.41}$$

we finally write the amplitude as

$$i\mathcal{M} = -\bar{u}(p_2)\mathcal{M}_0(p_2, p_1)u(p_1)\ e\left[\frac{p_2 \cdot \epsilon^*(k)}{p_2 \cdot k} - \frac{p_1 \cdot \epsilon^*(k)}{p_1 \cdot k}\right]. \tag{20.42}$$

Integrating over the momentum of the photon and summing over its polarizations, we obtain the differential cross section for emission of a photon as a function of the cross section without photon emission,

$$d\sigma(p_1 \to p_2 + \gamma) = d\sigma(p_1 \to p_2)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2k}\sum_{\lambda=1,2} e^2 \left|\frac{p_2 \cdot \epsilon^{(\lambda)}(k)}{p_2 \cdot k} - \frac{p_1 \cdot \epsilon^{(\lambda)}(k)}{p_1 \cdot k}\right|^2. \tag{20.43}$$

The differential probability for a photon of momentum k is then

$$d\mathcal{P}(p_1 \to p_2 + k) = \frac{d^3k}{(2\pi)^3}\sum_\lambda \frac{e^2}{2k}\left|\vec{\epsilon}_{(\lambda)} \cdot \left(\frac{\vec{p}_2}{p_2 \cdot k} - \frac{\vec{p}_1}{p_1 \cdot k}\right)\right|^2. \tag{20.44}$$

The total probability, integrating $|k|$ between the regulator $\mu_{\rm ph}$ (photon mass) and a maximum of $|q|$ (since the condition for soft photon was for $|k| \ll |q|$, where $\vec{q} = \vec{p}_2 - \vec{p}_1$), is

$$\mathcal{P} \simeq \frac{\alpha}{\pi}\int_{\mu_{\rm ph}}^{|q|} \frac{dk}{k}\mathcal{I} = \frac{\alpha}{2\pi}\log\frac{q^2}{\mu_{\rm ph}^2}\mathcal{I}. \tag{20.45}$$

For the differential cross section summed over k, we then find

$$d\sigma(p_1 \to p_2 + k) = d\sigma(p_1 \to p_2)\frac{\alpha}{2\pi}\log\frac{q^2}{\mu_{\rm ph}^2}\mathcal{I}. \tag{20.46}$$

We will not describe here the calculation of $\mathcal{I}$ (it can be found for instance in [3]), but one finds

$$\mathcal{I}(q^2 \to \infty) \simeq 2\log\frac{q^2}{m^2}. \tag{20.47}$$

If follows then that the differential cross section for $p_1 \to p_2$, taking into account the 1-loop IR divergence, is written in terms of the tree level process as (from $|F_1^{(0+1)}(q^2)|^2$)

$$\frac{d\sigma}{d\Omega}(p_1 \to p_2) = \left(\frac{d\sigma}{d\Omega}\right)_0 (p_1 \to p_2)\left[1 - \frac{\alpha}{\pi}\log\frac{q^2}{m^2}\log\frac{q^2}{\mu_{\rm ph}^2} + \mathcal{O}(\alpha^2)\right]. \tag{20.48}$$

On the other hand, we express the differential cross section for $p_1 \to p_1 + \gamma$ in terms of the same tree level process (without γ, described by $\mathcal{M}_0$) as

$$\frac{d\sigma}{d\Omega}(p_1 \to p_2 + \gamma) = \left(\frac{d\sigma}{d\Omega}\right)_0 (p_1 \to p_2)\left[+\frac{\alpha}{\pi}\log\frac{q^2}{m^2}\log\frac{q^2}{\mu_{\rm ph}^2} + \mathcal{O}(\alpha^2)\right]. \tag{20.49}$$

That means that their sum is independent of $\mu_{\rm ph}^2$, i.e. it is IR finite! Note that this is an abstract result, because we cannot really measure the total cross section with emission of a γ of arbitrary energy.

More physically, we can consider a detector that has an energy resolution of E_{min}, i.e. it cannot detect photons of smaller energy. Then the process with emission of a photon of smaller energy is considered as part of the process without emission. We need then to integrate $\int_{\mu_{\rm ph}}^{E_{min}} dk/k = 1/2\ln E_{min}^2/\mu_{\rm ph}^2$. We also can prove (will not be done here) that

$$\mathcal{I}(q^2) = 2f_{IR}(q^2), \quad \forall q^2. \tag{20.50}$$

We thus obtain

$$\frac{d\sigma}{d\Omega}(p_1 \to p_2 + \gamma(k \leq E_{min})) = \left(\frac{d\sigma}{d\Omega}\right)_0 (p_1 \to p_2)\left[+\frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{E_{min}^2}{\mu_{\rm ph}^2} + \mathcal{O}(\alpha^2)\right], \tag{20.51}$$

and as before

$$\frac{d\sigma}{d\Omega}(p_1 \to p_2) = \left(\frac{d\sigma}{d\Omega}\right)_0 (p_1 \to p_2)\left[1 - \frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{q^2 \text{ or } m^2}{\mu_{\rm ph}^2} + \mathcal{O}(\alpha^2)\right], \tag{20.52}$$

for a total of

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{measured}} = \frac{d\sigma}{d\Omega}(p_1 \to p_2) + \frac{d\sigma}{d\Omega}(p_1 \to p_2 + \gamma(\leq E_{min}))$$
$$= \left(\frac{d\sigma}{d\Omega}\right)_0 (p_1 \to p_2)\left[1 - \frac{\alpha}{\pi} f_{IR}(q^2) \log \frac{q^2 \text{ or } m^2}{E_{min}^2} + \mathcal{O}(\alpha^2)\right]. \quad (20.53)$$

Summation of IR divergences and Sudakov factor

We now resum one-loop diagrams on the side of virtual photons (loop corrections to the vertex), of the type of exchanging photons between the two fermion lines, but planarly ("parallel", they do not cross), and correspondingly on the photon emission side we consider the process of emission of n photons, one after another, from the external lines. In a way similar to the calculation sketched in lecture 18 for the Alrarelli-Parisi evolution equation in lecture 18, we obtain that this process resums, giving a factor

$$\left[\frac{\alpha}{\pi} \log \frac{q^2}{m^2} \log \frac{q^2}{\mu_{\text{ph}}^2}\right]^n. \quad (20.54)$$

But moreover, there is a symmetry factor of $1/n!$ in front of this, because emitted photons are indistinguishable, which means that when we sum these contributions, the one-loop IR divergence exponentiates!

The exponentiation of IR divergences is rigorously proved in a theorem by Bloch and Nordsieck (1937).

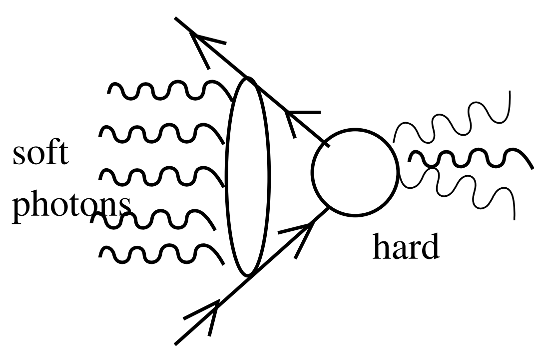

Figure 64: Factorization of the amplitude into a "hard" part and a "soft" part that governs the emission of soft photons.

Up to now, we have considered the QED vertex at one-loop, but an important property of IR divergences is *factorization*, which means that in some physical process, the "soft" and "hard" contributions factorize, as in Fig.64. We can split the process into a "hard" part (with large momenta), for scattering of 2 fermions with some other stuff, which we called $\mathcal{M}_0$ previously (note that we had assumed that $\mathcal{M}_0$ was the tree vertex, but we can easily check that it did not matter what $\mathcal{M}_0$ was, as long as it was a hard process); and a "soft" part (with small momenta) for scattering of two fermions to two fermions, possibly with soft photon emission.

Then the contribution to the soft part of soft *virtual* photons, i.e. of photons exchanged between the two fermion lines, gives a factor

$$\left[-\frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{q^2}{\mu_{\rm ph}^2}\right]^n\frac{1}{n!}\equiv\frac{X^n}{n!}. \tag{20.55}$$

On the other hand, the contribution to the soft part of the soft *real* (emitted) photons, i.e. photons emitted from the two fermion lines, gives a factor of

$$\frac{1}{n!}\left[\frac{\alpha}{\pi}\mathcal{I}\log\frac{E_{min}}{\mu_{\rm ph}}\right]^n=\left[\frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{E_{min}^2}{\mu_{\rm ph}^2}\right]^n\frac{1}{n!}. \tag{20.56}$$

That means that in the measured differential cross section, we have

$$\begin{aligned}\left(\frac{d\sigma}{d\Omega}\right)_{\rm measured} &= \left(\frac{d\sigma}{d\Omega}\right)_0(p_1\to p_2)\left(\sum_n\frac{X^n}{n!}\right)\left(\sum_m\frac{Y^m}{m!}\right)\\ &= \left(\frac{d\sigma}{d\Omega}\right)_0(p_1\to p_2)\exp\left[-\frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{q^2}{\mu_{\rm ph}^2}\right]\exp\left[\frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{E_{min}^2}{\mu_{\rm ph}^2}\right]\\ &= \left(\frac{d\sigma}{d\Omega}\right)_0(p_1\to p_2)\exp\left[-\frac{\alpha}{\pi}f_{IR}(q^2)\log\frac{q^2}{E_{min}^2}\right].\end{aligned} \tag{20.57}$$

The exponential factor is called a *Sudakov form factor*.

Once again, note that this exponential factor is only the resummation of one-loop diagrams, it is not genuinely two-loop, meaning that when expanding it in α, only the leading term is exact. In general, at each loop order we will have another power of α contribution in the exponential, i.e.

$$\exp[\alpha(...)+\alpha^2(...)+\alpha^3(...)+...] \tag{20.58}$$

Important concepts to remember

- Collinear IR divergences are due to massless particles ("photons") collinear with massless external states. For virtual particles, we have IR divergences in loops, for real particles, we have IR divergences for particle emission. It appears due to the self-interaction of massless states, so is present in QCD, but not in QED.

- Soft IR divergences are due to massless particles ("photons") being soft (very small momenta). For virtual particles, we have IR divergences in loops, for real particles, we have IR divergences for particle emission. It appears independently of the external states, and is present in both QCD and QED.

- In dimensional regularization, IR divergences appear due to factors of $1/\epsilon(P^2/\mu^2)^\epsilon$ in QED (just soft divergences) or $1/\epsilon^2(P^2/\mu^2)^\epsilon$ in QCD (soft and collinear divergences), where P^2 is some relevant invariant: in the case of nonabelian YM in the planar limit, it is $-s_{i,i+1}=-(k_i+k_{i+1})^2$ for consecutive external massless lines.

- The IR divergences can be regulated by including a mass μ for the massless particles, or by dimensional regularization, with $D = 4 - 2\epsilon$, but with $\epsilon_{UV} = -\epsilon_{IR} < 0$.

- In QED, the divergence in the form factor is $F_1 = 1 - \alpha/(2\pi) f_{IR}(q^2) \log(q^2 \text{ or } m^2)/m^2$, or at large q^2 as $F_1(q^2) = 1 - \alpha/(2\pi) \log(q^2/m^2) \log(q^2/\mu_{\rm ph}^2$, which is the Sudakov double logarithm.

- The IR divergences cancel order by order in α between processes with virtual photons (loop corrections) and processes with real photons (emission of soft or collinear photons).

- We can resum the one-loop corrections, and correspondingly the multiple emissions of photons from the external lines, and obtain that the one-loop divergences exponentiate, obtaining the Sudakov form factor.

Further reading: See chapter 12.1,12.2,12.3 in [2] and chapter 6.4, 6.5 in [3].

Exercises, Lecture 20

1) Consider scalar QED, i.e. a photon γ coupled to a massive (charged) complex scalar ϕ. Calculate $f_{IR}(q^2)$ at one-loop ($\mathcal{O}(g^2)$ correction) for the vertex $V(p_2, p_1) \equiv F_1(q^2)$ in Fig.65.

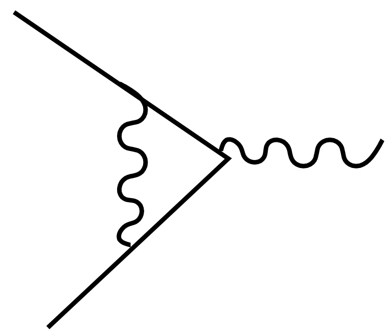

Figure 65: IR divergent diagram.

2) Calculate (using the same formulas from class, and ignoring the same calculations) the Sudakov form factor at one-loop ($\mathcal{O}(g^2)$).

21 Lecture 21. IR safety and renormalization in QCD; general IR-factorized form of amplitudes.

In this lecture, we will analyze IR divergences in QCD. To start however, we will give some more results about QED that easily generalize to QCD.

As we saw last lecture, to IR-regularize we can use dimensional regularization, but in $D = 4 - 2\epsilon$, $\epsilon > 0$ for UV divergences, whereas $\epsilon < 0$ for IR divergences. We also saw how to map between photon mass regularization and dimensional regularization. We only showed how to map the divergent terms, but the exact map actually contains some finite piece as well,

$$\ln \frac{\mu_{\rm ph}^2}{m^2} \leftrightarrow \Gamma(\epsilon)\left(\frac{4\pi\mu^2}{m^2}\right)^\epsilon = \frac{\Gamma(1+\epsilon)}{\epsilon}\left(\frac{4\pi\mu^2}{m^2}\right)^\epsilon = \frac{1}{\epsilon} - \gamma + \ln\frac{4\pi\mu^2}{m^2}. \tag{21.1}$$

This relation can be derived by computing $\Gamma^\mu_{\alpha\beta}$ in both mass and dimensional regularization, and comparing the results.

Generalizing a bit the result from last lecture, we saw that the same integral having both UV and IR divergences (at different endpoints), needs a different dimensional regularization for each, so we write

$$\frac{\Gamma(D/2)}{\pi^{D/2} i}\int \frac{d^Dq}{(q^2)^n} = \frac{1}{D/2-n} - \frac{1}{D'/2-n}\;, \tag{21.2}$$

$D = 4-2\epsilon$, $D' = 4-2\epsilon'$ for $n = 2$, and only at the end of the calculation do we set $D = D'$.

In particular, $\Gamma^\mu_{\alpha\beta}$ has still UV divergences, that is dimensionally regularized with D, whereas the IR divergence is either regularized with mass, or dimensional regularization with D'. One finds

$$\begin{aligned}
\Gamma^\mu|_{\text{mass reg.}} &= \frac{\alpha}{4\pi}\left[\gamma_\mu\left(\frac{1}{\epsilon} - \gamma + \ln\frac{4\pi\mu^2}{m^2} + \frac{v^2+1}{v}\ln\frac{v+1}{v-1}\ln\frac{\mu_{\rm ph}^2}{m^2} + F(v)\right)\right.\\
&\left. -\frac{(p_1-p_2)_\mu}{2m}\frac{v^2-1}{v}\ln\frac{v+1}{v}\right]\\
\Gamma^\mu|_{\text{dim.reg.}} &= \frac{\alpha}{4\pi}\left(\frac{4\pi\mu^2}{m^2}\right)^\epsilon\Gamma(1+\epsilon)\left[\gamma_\mu\left(\frac{1}{\epsilon'} + \frac{1}{\epsilon}\frac{v^2+1}{v}\ln\frac{v+1}{v-1} + F(v)\right)\right.\\
&\left. -\frac{(p_1-p_2)_\mu}{2m}\frac{v^2-1}{v}\ln\frac{v+1}{v}\right]\;,
\end{aligned} \tag{21.3}$$

where $v = \sqrt{1+4m^2/q^2}$ and $F(v)$ is a given function of v, of some complicated form, and finite.

We see that by expanding *in* ϵ $\left(\frac{4\pi\mu^2}{m^2}\right)^\epsilon\Gamma(1+\epsilon)$ as a coefficient of the $\ln(v+1)/(v-1)$ term, and keeping the ϵ-finite part, we have a match between the dimensional regularization and the mass regularization results only if we have (21.1).

The above was for QED, but in QCD the same calculation holds, chaging external electrons for external quarks, and photons for gluons, and then we also exchange α for $\alpha_s C_F$, where $\alpha_s = g^2/(4\pi)$.

In QED, we saw that IR divergences exponentiate and factorize. We make this a bit more formal now, to compare with QCD.

At one-loop, the IR divergent part of the vertex can be expressed in the form

$$\bar{u}(p_2)\Gamma_\mu^{(IR)}u(p_1) = e\mu^\epsilon \bar{u}(p_2)\gamma_\mu u(p_1)\alpha\Gamma(\epsilon, q^2/m^2) \;, \tag{21.4}$$

and from the calculation of the previous lecture, one can show (we will not do it here) that we can put the resulting Γ in the form

$$\alpha\Gamma(\epsilon, q^2/m^2) = -\frac{1}{2}(e\mu^\epsilon)^2 \int \frac{d^Dk}{(2\pi)^D}\frac{-i}{k^2 - i\epsilon}\left[\frac{2p_1^\alpha}{2p_1\cdot k + k^2 - i\epsilon} - \frac{2p_2^\alpha}{2p_2\cdot k + k^2 - i\epsilon}\right]^2 . \tag{21.5}$$

The approximation that we made last lecture, of considering for the leading IR divergence that k^2 can be neglected with respect to $2p_i \cdot k$ in the denominators is called *the eikonal approximation*, or *the leading divergence approximation*, so

$$\alpha\Gamma^{\text{eik}}(\epsilon) = -\frac{1}{2}(e\mu^\epsilon)^2 \int \frac{d^Dk}{(2\pi)^D}\frac{-i}{k^2 - i\epsilon}\left[\frac{2p_1^\alpha}{2p_1\cdot k - i\epsilon} - \frac{2p_2^\alpha}{2p_2\cdot k - i\epsilon}\right]^2 . \tag{21.6}$$

Note that Γ^{eik} does not depend anymore on q^2/m^2, as we can check.

We saw last lecture that IR divergences exponentiate and factorize into a hard part, containing the large momenta, and a soft part, containing the IR divergences in exponential form. Formally then

$$\Gamma_\mu(p_1, p_2) = e^{\alpha\Gamma^{\text{eik}}}\Gamma_\mu^{(H)} \;, \tag{21.7}$$

where $\Gamma_\mu^{(H)}$ is the hard part. Since however as things stand $\Gamma_\mu^{(H)}$ still has *UV divergences*, one can also write the above as

$$\Gamma_\mu(p_1, p_2) = e^{\alpha\Gamma(\epsilon, M^2/m^2)}\Gamma_\mu^{(\text{finite})} \;, \tag{21.8}$$

where $q^2 \equiv -M^2$ and

$$\Gamma_\mu^{(\text{finite})} = \Gamma_\mu^{(H)} e^{\alpha\Gamma^{\text{eik}}(\epsilon) - \alpha\Gamma(\epsilon, M^2/m^2)}. \tag{21.9}$$

IR safety in QCD

In QCD, the coupling constant depends on scale and is strong at low energy, $\alpha_s \to \infty$. This means that it is very hard to define physical quantities, since besides the IR divergences we also have to deal with infinite coupling. Quantities that are free of IR divergences are called *IR safe*. As we already saw, cross sections, that are certainly observable, are supposed to be IR safe. One way to express this is that as $m^2 \ll q^2$ (here m is the quark mass), we have

$$\sigma\left(\frac{q^2}{\mu^2}, g_s, \frac{m^2}{\mu^2}\right) = \bar{\sigma}\left(\frac{q^2}{\mu^2}, g_s\right) + \mathcal{O}\left[\left(\frac{m^2}{q^2}\right)^b\right] \;, \tag{21.10}$$

where $b > 0$, and then $\bar{\sigma}$ is completely finite. Since in asymptotically free theories the renormalized mass satisfies $m_R(\mu \to \infty) \to 0$, the zero mass limit makes sense and moreover, the high energy ($q \to \infty$) or zero mass ($m \to 0$) limits are equivalent.

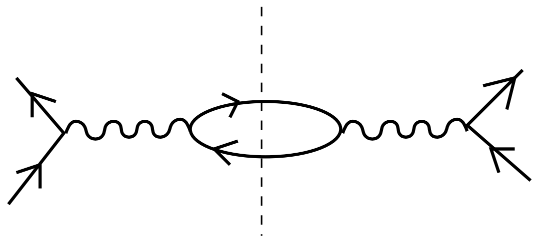

Figure 66: Diagram for the Born Cross section via unitarity cut.

Moreover, we can use cross sections to define the running of the coupling constant. We will explain this, with the example of a simple process, very relevant experimentally.

Born cross section for $e^+e^- \to (q\bar{q}) \to$ **hadrons**.

Consider the tree diagram for $e^+e^- \to q\bar{q}$ through an intermediate photon. The quarks have charge $Q_f e$. For the cross section, where we have the quantity $|\mathcal{M}|^2 = \mathcal{M}\mathcal{M}^*$, we can draw diagrams with $\mathcal{M}$ and $\mathcal{M}$ flipped (time reversed) for $\mathcal{M}^*$, forming for instance a one-loop diagram from a tree one, with the quarks in the loop being on-shell, as in Fig.66. It is related to the optical theorem, where cutting loop diagrams gives diagrams for the cross section, but here we can just think of it as a useful trick.

We will not reproduce the calculation, but one finds that

$$\frac{d\sigma}{d\cos\theta} = N\frac{\pi\alpha^2}{2s}\sum_f Q_f^2(1+\cos^2\theta)\ , \tag{21.11}$$

which integrates to ($\cos\theta$ varies between -1 and +1)

$$\sigma_{\rm tot} = N\frac{4\pi\alpha^2}{3s}\sum_f Q_f^2. \tag{21.12}$$

Experimentally, one defines the ratio of the process with hadrons in the final state to the one with leptons in the final state, namely $e^+e^- \to \mu^+\mu^-$,

$$R \equiv \frac{\sigma_{\rm tot}(e^+e^- \to hadrons)}{\sigma_{\rm tot}(e^+e^- \to \mu^+\mu^-)} = N\sum_f Q_f^2\ , \tag{21.13}$$

where in the last equality we have considered our tree level process. This quantity is measured experimentally very well, and is one of the most stringent tests of QCD: it depends on the fact of being 3 colours, and on the total number of quarks and their charges. It agrees perfectly, leaving no room for extra quarks.

We are however interested in corrections to this result from QCD processes, i.e. corrections of order $\alpha\alpha_s$, where $\alpha_s = g^2/(4\pi)$. These would be one-loop diagrams with a QCD loop, or two-loop cut diagrams for σ. There are 8 such diagrams, and the calculation is long, so we will only parametrize the result and show the final answer. Details can be found in [2].

For the *total cross section* for $e^+e^- \to$ hadrons, we need to consider cut diagrams with $e^+e^- \to \gamma \to$hadrons$\to \gamma \to e^+e^-$, and the sum is over possible intermediate hadrons. This can be parametrized in the following way

$$\sigma_{\rm tot}(q^2) = \left[\frac{e^2\mu^{2\epsilon}}{2(q^2)^3}\right](k_1^\mu k_2^\nu + k_2^\mu k_1^\nu - k_1\cdot k_2 g^{\mu\nu})H_{\mu\nu}(q^2) \; , \tag{21.14}$$

where

$$H_{\mu\nu}(q^2) = e^2\mu^{2\epsilon}\sum_n \langle 0|j_\mu(0)|n\rangle\langle n|j_\nu(0)|0\rangle(2\pi)^4\delta^4(p_n - q). \tag{21.15}$$

Here the sum over n is over hadronic states, the states in the cut loop, and j_μ is the electromagnetic current. Moreover, $H_{\mu\nu}(q)$ must be transverse, which means that

$$H_{\mu\nu}(q^2) = (q_\mu q_\nu - q^2 g_{\mu\nu})H(q^2). \tag{21.16}$$

After a long calculation, where we split the contributions to $H_{\mu\nu}$, similarly to the QED case, into contributions of real and virtual gluons (i.e. real gluons are emitted gluons, thus gluons in the hadronic state $|n\rangle$, and virtual gluons are not), one finds

$$\begin{aligned} -g^{\mu\nu}H_{\mu\nu}^{\rm real} &= +2NC_2(F)Q_f^2\frac{\alpha\alpha_s}{\pi}q^2\left(\frac{4\pi\mu^2}{q^2}\right)^{2\epsilon}\left[\frac{1-\epsilon}{\Gamma(2-2\epsilon)}\right]\left[\frac{1}{\epsilon^2}+\frac{3}{2\epsilon}-\frac{\pi^2}{2}+\frac{19}{4}+\mathcal{O}(\epsilon)\right] \\ -g^{\mu\nu}H_{\mu\nu}^{\rm virtual} &= -2NC_2(F)Q_f^2\frac{\alpha\alpha_s}{\pi}q^2\left(\frac{4\pi\mu^2}{q^2}\right)^{2\epsilon}\left[\frac{1-\epsilon}{\Gamma(2-2\epsilon)}\right]\left[\frac{1}{\epsilon^2}+\frac{3}{2\epsilon}-\frac{\pi^2}{2}+4+\mathcal{O}(\epsilon)\right]. \end{aligned} \tag{21.17}$$

We see that, while individually they are both IR divergent, in their sum the divergences cancel and, except for the overall factor, the brackets cancel except for a finite contribution $19/4 - 4 = 3/4$. Finrally we obtain for the total cross section, with the leading contribution due to $q\bar{q}$ production (and soft gluons above),

$$\sigma_{\rm tot}(q^2) = \frac{N4\pi\alpha^2}{3q^2}\sum_f Q_f^2\left[1+\frac{\alpha_s}{\pi}\frac{3}{4}C_2(F)+\mathcal{O}(\alpha^2\alpha_s^2)\right]. \tag{21.18}$$

We can now determine α_s from $\sigma_{\rm tot}$, which will give us the dependence $\alpha_s(\mu^2)$. At the first order treated above however, there is no dependence on the renormalization scheme and on μ, as we can see.

At 2-loop order however, we will get insider the square bracket a contribution $(\alpha_s/\pi)^2A_2(q^2/\mu^2)$, which is scheme dependent and μ dependent. Since however the total cross section cannot depend on μ (which is a fictitious parameter, appearing from dimensional transmutation), it means that we can infer $\alpha_s(\mu)$ from $A_2(q^2/\mu^2)$ (the total μ dependence should cancel). A more formal way to express this is that we have a renormalization group equation (RGE) for $\sigma_{\rm tot}$ saying that the total μ dependence is zero, or

$$\left[\mu\frac{\partial}{\partial\mu}+\beta(g)\frac{\partial}{\partial g}\right]\sigma_{\rm tot}(q^2,\mu^2,\alpha_s(\mu^2)) = 0 \; , \tag{21.19}$$

where as usual $\beta(g) = \mu\partial g/\partial\mu$. Since μ is an arbitrary scale, we are in principle free to choose whatever we want. A useful choice is actually $\mu^2 = q^2$, which makes the running coupling depend on the physical momentum scale in the process, $\alpha_s = \alpha_s(q^2)$.

An a priori independent definition of α_s is from the beta function, which we will denote $\bar{\alpha}_s(\mu^2)$. From the definition of the beta function above, we have

$$\mu\frac{\partial}{\partial\mu}\frac{\bar{\alpha}_s}{\pi} = \frac{\bar{g}}{2\pi^2}\beta(\bar{g}_s) = \left(\frac{\bar{\alpha}_s}{\pi}\right)^2\frac{\beta_1}{2} + \left(\frac{\bar{\alpha}_s}{\pi}\right)^3\frac{\beta_2}{8}; \quad (21.20)$$

where one defines the numerical coefficients of the beta function as

$$\beta(g) = g\left[\frac{\alpha_s}{4\pi}\beta_1 + \left(\frac{\alpha_s}{4\pi}\right)^2\beta_2 + ...\right]. \quad (21.21)$$

We can integrate the equation from μ_0 to μ, and introducing a parameter Λ (or sometimes Λ_{QCD}), the QCD scale parameter, we find $\bar{\alpha}_s(\mu^2)$ as a function of $\ln(\mu/\Lambda)$ only.

With the nontrivial choice

$$\Lambda = \mu_0 \exp\left\{\frac{2}{\beta_1}\left(\frac{\pi}{\bar{\alpha}_s(\mu_0)} - \frac{\beta_2}{4\beta_1}\ln\left[\frac{\pi}{\bar{\alpha}_s(\mu_0)}\left(1 + \frac{\beta_2}{4\beta_1}\frac{\pi}{\bar{\alpha}_s(\mu_0)}\right)\right] + \frac{\beta_2}{4\beta_1}\ln\frac{-\beta_1}{4}\right)\right\}, \quad (21.22)$$

we find the somewhat simple form

$$\frac{\bar{\alpha}_s(\mu^2)}{\pi} = \frac{-2}{\beta_1\ln(\mu/\Lambda)} - \frac{\beta_2}{\beta_1^3}\frac{\ln\ln\frac{\mu^2}{\Lambda^2}}{\ln^2\frac{\mu}{\Lambda}} + \mathcal{O}\left(\frac{1}{\ln^3\left(\frac{\mu^2}{\Lambda^2}\right)}\right). \quad (21.23)$$

The proof is left as one of the exercises.

Factorization and exponentiation of IR divergences in gauge theories.

In QCD and nonabelian gauge theories in general we still have factorization and exponentiation of IR divergences, but the result is considerably more complicated. Here we will just give a flavor for it here.

To start, we must describe the color structure of amplitudes in terms of a *color basis* $\mathcal{C}_{[i]}$. Indeed, we can convince ourselves that for all amplitudes, there are only a few color structures possible, and their coefficients are color-independent amplitudes, sometimes called color-ordered.

For instance, for amplitudes with external gluons, the external gluons are characterized only by an adjoint index a (in the case of QCD, saying which one of the 8 possible gluons it belongs to), and the propagators and vertices have only delta functions for the fundamental indices $i, j, ...$, which means that these indices are summed over, obtaining either traces of matrices T^a with a an external gluon index, or $\delta^i_i = N$ factors. It follows that the possible color structures are all possible traces of T^a matrices.

For example, in the case of the 4-point scattering of gluons in a theory like QCD, where $\text{Tr}[T^a] = 0$, for instance for any $SU(N)$, we can have either single traces,

$$\mathcal{C}_{[1]} = \text{Tr}[T^{a_1}T^{a_2}T^{a_3}T^{a_4}]$$

$$\cdots$$
$$\mathcal{C}_{[6]} = \mathrm{Tr}[T^{a_1}T^{a_4}T^{a_3}T^{a_2}] \;, \tag{21.24}$$

(by cyclicity of the trace we can always put a_1 on the first position, and then there are 6 permutations of a_2, a_3, a_4), or double traces,

$$\begin{aligned} \mathcal{C}_{[7]} &= \mathrm{Tr}[T^{a_1}T^{a_2}]\,\mathrm{Tr}[T^{a_3}T^{a_4}] \\ \mathcal{C}_{[8]} &= \mathrm{Tr}[T^{a_1}T^{a_3}]\,\mathrm{Tr}[T^{a_3}T^{a_4}] \\ \mathcal{C}_{[9]} &= \mathrm{Tr}[T^{a_1}T^{a_4}]\,\mathrm{Tr}[T^{a_2}T^{a_3}]. \end{aligned} \tag{21.25}$$

Then the 4-point amplitude for gluons can be expanded in the above basis,

$$\mathcal{A}(1234) = \sum_{i=1}^{9} A_{[i]}\mathcal{C}_{[i]} \;, \tag{21.26}$$

and this can be written as a vector with 9 components

$$|A\rangle = \begin{pmatrix} A_{[1]} \\ A_{[2]} \\ \cdots \\ A_{[9]} \end{pmatrix} \;, \tag{21.27}$$

or rather, since we have a reflection symmetry relation

$$A_n(12...n) = (-1)^n A_n(n...21) \;, \tag{21.28}$$

we can use only $A_{1234}, A_{1342}, A_{1423}$ and $A_{12;34},\ A_{13;42},\ A_{14;23}$.

In general, for the 4-point function for scattering of $2 \to 2$ partons (quarks and gluons) we have a similar story, with some color basis $\{\mathcal{C}_L\}$, and we expand in it as

$$\mathcal{A}^{[f]} = \sum_L A_L^{[f]}(\mathcal{C}_L). \tag{21.29}$$

and put the A_L coefficients in a vector $|A\rangle$.

In this case, we can write the *factorization* of the amplitude as

$$\left|A\left(\frac{s_{ij}}{\mu^2}, a(\mu^2), \epsilon\right)\right\rangle = J\left(\frac{Q^2}{\mu^2}, a(\mu^2), \epsilon\right) \mathbb{S}\left(\frac{s_{ij}}{Q^2}, \frac{Q^2}{\mu^2}, a(\mu^2), \epsilon\right) \left|H\left(\frac{s_{ij}}{\mu^2}, \frac{Q^2}{\mu^2}, a(\mu^2), \epsilon\right)\right\rangle, \tag{21.30}$$

where $s_{ij} = (k_i + k_j)^2$ are the possible kinematic invariants (the amplitude must depend on them), $a(\mu^2)$ is the effective coupling,

$$a(\mu^2) = \frac{g^2 N}{8\pi^2}(4\pi e^{-\gamma})^\epsilon. \tag{21.31}$$

Here J is called the *jet function* and is an IR-divergent scalar factor, $\mathbb{S}$ is called the *soft function* and is an IR divergent matrix in color space, and $|H\rangle$ is called the *hard function*, and it contains only short-distance behaviour, i.e. it is IR finite (as $\epsilon \to 0$).

The difference between the soft function and the jet function is that J contains all collinear dynamics, and as a result it contains (at one-loop) all the $1/\epsilon^2$ poles (we said that one $1/\epsilon$ comes from soft divergences, and one $1/\epsilon$ from collinear divergences), whereas $\mathbb{S}$ is completely determined by the *anomalous dimension matrix* $\mathbf{\Gamma}$, which will be defined better later in the course, but essentially is a generalization of the anomalous dimension for scalar theories.

Note that we have introduced an arbitrary quantity Q, called the *factorization scale*. As we see, amplitudes are independent of Q, but its factorization into jet, soft and hard functions depends on Q.

The soft function is defined from the anomalous dimension matrix $\mathbf{\Gamma}$ with components $\mathbf{\Gamma}_{LJ}$ (such that the amplitude is written as $A_L = JS_{LI}H_I$) as the renormalization group equation

$$\frac{d}{d\ln Q}S_{LI} = -\mathbf{\Gamma}_{LJ}S_{JI}. \tag{21.32}$$

Note here that this has some similarity with the evolution equations. This is not a coincidence, and we will see in more detail next lecture that there is a connection between factorization and evolution. The solution of the above RGE (with some more input that will not be explained here) is written in the form

$$\mathbb{S}\left(\frac{s_{ij}}{Q^2},\frac{Q^2}{\mu^2},a(\mu^2),\epsilon\right) = P\exp\left\{-\frac{1}{2}\int_0^{Q^2}\frac{d\tilde{\mu}^2}{\tilde{\mu}^2}\mathbf{\Gamma}\left(\frac{s_{ij}}{Q^2},\bar{a}\left(\frac{\mu^2}{\tilde{\mu}^2},a(\mu^2),\epsilon\right)\right)\right\}\,, \tag{21.33}$$

where $\mathbf{\Gamma}$ is the anomalous dimension matrix, expanded in the coupling as

$$\mathbf{\Gamma}\left(\frac{s_{ij}}{Q^2},a(\mu^2)\right) = \sum_{l=1}^{\infty}a(\mu^2)^l\mathbf{\Gamma}^{(l)}\left(\frac{s_{ij}}{Q^2}\right)\,, \tag{21.34}$$

and at leading order (one-loop), the coupling $\bar{a}$ is given by

$$\bar{a}\left(\frac{\mu^2}{\tilde{\mu}^2},a(\mu^2),\epsilon\right) = a(\mu^2)\left(\frac{\mu^2}{\tilde{\mu}^2}\right)^{\epsilon}\sum_{n=0}^{\infty}\left[\frac{\beta_1}{4\pi}\left(\left(\frac{\mu^2}{\tilde{\mu}^2}\right)^{\epsilon}-1\right)a(\mu^2)\right]^n. \tag{21.35}$$

The IR divergent structure of gauge theories is given by:

1) $\gamma(a)$, the *cusp anomalous dimension*, or *Wilson line anomalous dimension*, or *soft anomalous dimension*. As the name suggests, this characterizes soft divergences, and also divergences arising from "cusps" (angles) in the Wilson line.

2) $\mathcal{G}_0(a)$, the *collinear anomalous dimension*, which characterizes the collinear divergences.

These functions are expanded in the coupling as

$$\begin{aligned}\gamma(a) &= \sum_{l=1}^{\infty}a^l\gamma^{(l)}\\ \mathcal{G}_0(a) &= \sum_{l=1}^{\infty}a^l\mathcal{G}_0^{(l)}.\end{aligned} \tag{21.36}$$

In QCD, there is also the anomalous dimension matrix $\mathbf{\Gamma}$, which is independent of the two above functions.

However, in a gauge theory called $\mathcal{N} = 4$ SYM, where many things can be computed exactly (we will not explain what $\mathcal{N} = 4$ SYM is, we use it just for purposes of illustration), there is no nontrivial $\mathbf{\Gamma}$ matrix, and in that case we can write down the jet function just in terms of $\gamma(a)$ and $\mathcal{G}_0(a)$, as

$$J\left(\frac{Q^2}{\mu^2}, a, \epsilon\right) = \exp\left[-\frac{1}{2}\sum_{l=1}^{\infty} a^l \left(\frac{\mu^2}{Q^2}\right)^{l\epsilon}\left(\frac{\gamma^{(l)}}{(l\epsilon)^2} + \frac{2\mathcal{G}_0^{(l)}}{l\epsilon}\right)\right]. \tag{21.37}$$

Important concepts to remember

- The QED vertex is written as the exponent of the eikonal approximation vertex times a hard vertex.
- Cross sections in QCD are IR safe.
- We can define the running $\alpha_s(\mu^2)$ from the cross section, e.g. for e^+e^- into hadrons, by imposing that there is no dependence of μ for σ_{tot}.
- One can choose $\mu^2 = q^2$, and thus find $\alpha_s(q^2)$, but it is not necessary.
- QCD amplitudes can be decomposed into a color basis.
- QCD amplitudes factorizes into a scalar jet function, a matrix soft function, and a vector hard function.
- The split (but not the amplitude) depends on an arbitrary factorization scale.
- The soft function is determined by the anomalous dimension matrix alone.
- The jet function contains all collinear dynamics, thus all double poles in ϵ, and is characterized by the cusp anomalous dimension and the collinear anomalous dimension, together with the anomalous dimension matrix.

Further reading: See chapter 12.3 and 12.4 in [2] and book by Taizo Muta, "Foundation of Quantum Chromodynamics", chapter 6.1 to 6.3.

Exercises, Lecture 21

1) Prove that

$$\frac{\bar{\alpha}_s(\mu^2)}{\pi} = \frac{-2}{\beta_1 \ln(\mu/\Lambda)} - \frac{\beta_2}{\beta_1^3}\frac{\ln\ln\frac{\mu^2}{\Lambda^2}}{\ln^2\frac{\mu}{\Lambda}} + \mathcal{O}\left(\frac{1}{\ln^3\left(\frac{\mu^2}{\Lambda^2}\right)}\right). \tag{21.38}$$

from

$$\mu\frac{\partial}{\partial\mu}\frac{\bar{\alpha}_s}{\pi} = \left(\frac{\bar{\alpha}_s}{\pi}\right)^2\frac{\beta_1}{2} + \left(\frac{\bar{\alpha}_s}{\pi}\right)^3\frac{\beta_2}{8}; \tag{21.39}$$

for the choice of

$$\Lambda = \mu_0 \exp\left\{\frac{2}{\beta_1}\left(\frac{\pi}{\bar{\alpha}_s(\mu_0)} - \frac{\beta_2}{4\beta_1}\ln\left[\frac{\pi}{\bar{\alpha}_s(\mu_0)}\left(1+\frac{\beta_2}{4\beta_1}\frac{\bar{\alpha}_s(\mu_0)}{\pi}\right)\right] + \frac{\beta_2}{4\beta_1}\ln\frac{-\beta_1}{4}\right)\right\}\ , \tag{21.40}$$

2) Consider the decomposition of a 4-point gauge theory amplitude in coefficients c_{ijkl} related to box diagrams,

$$\mathcal{A}(1234) = c_{1234}A_{1234} + c_{1342}A_{1342} + c_{1423}A_{1423}\ , \tag{21.41}$$

where

$$c_{1234} = \tilde{f}^{ea_1b}\tilde{f}^{ba_2c}\tilde{f}^{ca_3d}\tilde{f}^{da_4e} \tag{21.42}$$

is the color structure of the box diagram in Fig.67, where

$$\tilde{f}^{abc} = \mathrm{Tr}([T^a, T^b]T^c). \tag{21.43}$$

Express $c_{1234}, c_{1342}, c_{1423}$ in terms of $\mathcal{C}_{[i]}$, $i = 1, ..., 9$.

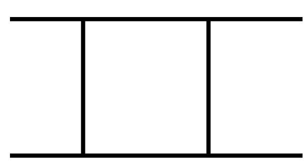

Figure 67: Box diagram.

22 Lecture 22. Factorization and the Kinoshita-Lee-Nauenberg theorem.

In this lecture, we will first give a rather general theorem about the cancellation of IR divergences in any physical transition probability, which happens when summing over initial and final states, and then we will say some words about the relation between factorization and evolution.

The KLN theorem

The theorem about finiteness of transition probabilities, when summing over initial and final states is due to Kinoshita (1962) and independently by TD Lee and Nauenberg (1964). It is somewhat more general than the specific case we are interested in, and it applies to a general quantum mechanical system.

The formalism we will use is of quantum mechanics in the interaction picture. There, the time evolution of states is

$$i\partial_t|\psi(t)\rangle = g\hat{H}_{Ii}(t)|\psi(t)\rangle \ , \tag{22.1}$$

and $\hat{H}_{Ii}$ is the interaction part of the Hamiltonian, in the interaction picture, i.e. $H = H_0 + H_I$, and

$$\hat{H}_{Ii} = e^{iH_0 t}\hat{H}_{I,S}e^{-iH_0 t}. \tag{22.2}$$

The time evolution of states is with the evolution operator U,

$$|\psi(t)\rangle = U(t,t')|\psi(t')\rangle \ , \tag{22.3}$$

where

$$U(t,t') = T\exp\left[-i\int_t^{t'} dt'' g\hat{H}_{Ii}(t'')\right]. \tag{22.4}$$

The S-matrix is given by the matrix elements of

$$S = U(+\infty,-\infty) = U(0,+\infty)^\dagger U(0,-\infty). \tag{22.5}$$

Then the transition probability between state $|a\rangle$ and state $|b\rangle$ is written as

$$|\langle b|S|a\rangle|^2 = \sum_{ij}(R^+_{bij})^* R^-_{aij} \ , \tag{22.6}$$

where

$$\begin{aligned} R^\pm_{aij} &\equiv \langle i|U(0,\pm\infty)|a\rangle^*\langle j|U(0,\pm\infty)|a\rangle \\ &= \langle j|U(0,\pm\infty)|a\rangle\langle a|U^\dagger(0,\pm\infty)|i\rangle \ , \end{aligned} \tag{22.7}$$

so that

$$|\langle b|S|a\rangle|^2 = \langle j|U(0,-\infty)|a\rangle\langle a|U^\dagger(0,-\infty)|i\rangle\langle i|U(0,+\infty)|b\rangle\langle b|U^\dagger(0,+\infty)|j\rangle. \tag{22.8}$$

We now calculate in perturbation theory (keeping only the first order in g)

$$\begin{aligned}\langle i|U(0,\pm\infty)|j\rangle &= \langle i|j\rangle - ig\int_0^{\pm\infty} dt''\langle i|e^{iH_0t''}\hat{H}_{I,S}e^{-iH_0t''}|j\rangle \\ &= \delta_{ij} - ig\int_0^{\pm\infty} dt'' e^{i(E_i-E_j)t''}H_{I,ij} \\ &= \delta_{ij} - \frac{gH_{I,ij}}{E_i - E_j \pm i\epsilon}. \end{aligned} \tag{22.9}$$

Here we have used that $|i\rangle$ are eigenstates of H_0 with energy E_i at leading order, i.e. $H_0|i\rangle = E_i|i\rangle$, we defined

$$H_{I,ij} = \langle i|H_{I,S}|j\rangle\ , \tag{22.10}$$

and we have introduced a regulator in the exponential to make the term at $\pm\infty$ decay to zero, $e^{i(E_i-E_j\pm i\epsilon)(\pm\infty)} = 0$.

Therefore now to order g we obtain also

$$R^{\pm}_{aij} = \delta_{ia}\delta_{ja} - \frac{gH^*_{Iia}}{E_i - E_a \mp i\epsilon}\delta_{ja} - \frac{gH_{ja}}{E_j - E_a \pm i\epsilon}\delta_{ia} + \mathcal{O}(g^2). \tag{22.11}$$

We now see the origin of the divergences we want to get rid of, corresponding to IR divergences. If there is a degeneracy between states $|a\rangle$ (external) and $|i\rangle$ or $|j\rangle$ (internal), i.e. $E_a = E_i$, corresponding to a degeneracy between real or virtual states, like e^- and $e^- + \gamma$ with γ having close to zero energy, then we obtain a divergence in $R^{\pm}_{aij}$.

The divergence will be eliminated by summing over initial and final states.

Let $D(E)$ be the set of all states with energy E. Then, we want to show that

$$\sum_{a\in D(E)}\sum_{b\in D(E)} |\langle b|S|a\rangle|^2 \tag{22.12}$$

is free of (IR) divergences. From the decomposition above in terms of $R^{\pm}$, the statement is completely equivalent to the statement that

$$R^{\pm}_{ij}(E) \equiv \sum_{a\in D(E)} R^{\pm}_{aij} \tag{22.13}$$

has no (IR) divergences.

Proof. To show it, we explicitly calculate $R^{\pm}_{ij}(E)$ case by case. We obtain

$$\begin{aligned} R^{\pm}_{ij}(E) &= 0, \quad \text{for} \quad i,j \notin D(E) \\ &= -\frac{gH^*_{I,ij}}{E_i - E \mp i\epsilon}, \quad \text{for} \quad j\in D(E), i\notin D(E) \Rightarrow E_i - E \neq 0 \\ &= -\frac{gH_{I,ji}}{E_j - E \pm i\epsilon}, \quad \text{for} \quad i\in D(E), j\notin D(E) \Rightarrow E_j - E \neq 0 \\ &= \delta_{ij} \quad \text{for} \quad i,j\in D(E). \end{aligned} \tag{22.14}$$

As we see from the above, it is finite in all these cases. *q.e.d.*

Next we want to generalize the proof to all orders. We will proceed by induction, since we already proved it at first order.

We first diagonalize the total Hamiltonian H by using $U(0,\pm\infty)$:

$$U^\dagger H U = \hat{H}_0 \tag{22.15}$$

is diagonal (but is different from H_0, the free Hamiltonian). Then

$$[U,\hat{H}_0] = U\hat{H}_0 - \hat{H}_0 U = (H-\hat{H}_0)U = (gH_I+\Delta)U\ , \tag{22.16}$$

where we have used that $H = H_0 + gH_I$ and defined

$$\Delta \equiv H_0 - \hat{H}_0\ , \tag{22.17}$$

which is a diagonal operator.

We expand in perturbation theory

$$\Delta = \sum_n g^n \Delta_n; \quad U = \sum_n g^n U_n; \quad R = \sum_n g^n R_n. \tag{22.18}$$

Then we calculate

$$\begin{aligned} R^\pm_{ij}(E) &= \sum_{a\in D(E)} R^\pm_{aij}(E) = \sum_{r,s} g^{r+s} \sum_{a\in D(E)} \langle i|U_r(0,\pm\infty)|a\rangle^* \langle j|U_s(0,\pm\infty)|a\rangle \\ &\equiv \sum_n g^n R^\pm_{n,ij}(E)\ , \end{aligned} \tag{22.19}$$

so that we get

$$R^\pm_{n,ij}(E) = \sum_r \sum_{a\in D(E)} \langle i|U_r(0,\pm\infty)|a\rangle^* \langle j|U_{n-r}(0,\pm\infty)|a\rangle. \tag{22.20}$$

Then we have reduced the KLN theorem to the induction step for $R^\pm_{n,ij}(E)$, i.e. to the following

Lemma.

If $R^\pm_{n,ij}(E)$ is free from IR divergences for $n \leq N$, then $R^\pm_{n,ij}$ is free from IR divergences for $n \leq N+1$.

Proof.

We prove it case by case.

i) $i \notin D(E)$.

Consider (22.16) in between $\langle i|$ and $|a\rangle$,

$$\langle i|[U,\hat{H}_0]|a\rangle = \langle i|(gH_I+\Delta)U|a\rangle\ , \tag{22.21}$$

and consider now the energies of the total diagonalized Hamiltonian $\hat{H}_0|i\rangle = E_i|i\rangle$. Then applying $\hat{H}_0$ on the left hand side in the above on the states, and on the right hand side introducing a complete set $\sum_k |k\rangle\langle k|$ in between U and $(gH_I+\Delta)$, and defining $\langle i|\Delta|i\rangle = \Delta_i$, we obtain

$$(E_a - E_i)U_{ia} = gH_{I,ik}U_{ka} + \Delta_i U_{ia}. \tag{22.22}$$

Since $a \in D(E)$, but $i \notin D(E)$, $E_i - E_a \neq 0$, so we can divide by it and write, after expanding in powers of g both sides of the equation

$$U_{r,ia} = \frac{1}{E_a - E_i}\left[\sum_k H_{I,ik}U_{r-1,ka} + \sum_s \Delta_{s,i}U_{r-s,ia}\right]. \tag{22.23}$$

Then we obtain

$$\begin{aligned} R^{\pm}_{n,ij}(E) &= \sum_r \sum_{a\in D(E)} U^*_{r,ia}U_{n-r,ja} \\ &= \frac{1}{E-E_i}\left[\sum_r \sum_{a\in D(E)}\left(\sum_k H^*_{I,ik}U^*_{r-1,ka} + \sum_s \Delta^*_{si}U^*_{r-s,ia}\right)U_{n-r,ja}\right] \\ &= \frac{1}{E-E_i}\left[\sum_k H^*_{I,ik}R^{\pm}_{n-1,kj}(E) + \sum_{s=1}^{n-1}\Delta^*_{si}R^{\pm}_{n-s,ij}(E)\right]. \end{aligned} \tag{22.24}$$

Here, after the definition in the first line, we have substituted the expansion for $U^*_{r,ia}$ and then re-formed $R^{\pm}_{ij}(E)$ coeficients, so that in the final form we have $R^{\pm}_{n,ij}$ written in terms of $R^{\pm}_{m,ij}(E)$ with $m \leq n-1$, as well as $\delta_{s,i}$ and $H_{I,ik}$.

But we assume that $H_{I,ik}$ are finite (the matrix elements of the interaction Hamiltonian cannot be infinite because of unitarity), and that Δ_i is also finite (or at least is at each order in g, i.e. $\Delta_{r,i}$ is finite), since those are the differences between the free and interacting energies, and they must be finite (or rather, since we have chosen finite E_a, E_i, and $\Delta_i = E_i - E_i^{(0)}$, the corresponding free values for the energies, $E_i^{(0)}$, must be positive, so Δ_i must be finite).

Then indeed it follows that if all $R^{\pm}_{m,ij}(E)$ for $m \leq n$ are finite, so is $R^{\pm}_{n+1,ij}(E)$. *q.e.d.*

ii) $j \notin D(E)$. This case is the same with case i), because

$$(R^{\pm}_{n,ij}(E))^* = R^{\pm}_{n,ji}(E). \tag{22.25}$$

iii) $i, j \in D(E)$. In this case, we cannot use the same equations. Instead, from unitarity,

$$UU^\dagger = 1 \Rightarrow \sum_r U_{n-r}U_r^\dagger = 0, n \neq 0\ , \tag{22.26}$$

and by sandwiching it between $\langle j|$ and $|i\rangle$, and inserting in the middle a complete set

$$\left(\sum_{a\in D(E)} + \sum_{a\notin D(E)}\right)|a\rangle\langle a|\ , \tag{22.27}$$

we get

$$\sum_r \sum_{a\in D(E)} U^*_{r,ia}U_{n-r,ja} + \sum_r \sum_{a\notin D(E)} U^*_{r,ia}U_{n-r,ja} = 0\ , \tag{22.28}$$

and since on the left we form $R^{\pm}_{n,ij}(E)$, we get

$$R^{\pm}_{n,ij}(E) = -\sum_{r=0}^{n}\sum_{a\notin D(E)} U^*_{r,ia}U_{n-r,ja} \ , \tag{22.29}$$

so it is IR finite, since all the matrix elements of an unitary operator (the evolution operator) are finite.

q.e.d. Lemma, thus *q.e.d.* KLN theorem.

Thus we have proved generally that IR divergences disappear in physical quantities, transition probabilities summed over all states of a given energy, which justifies the use of QFT despite the presence of IR divergences.

Factorization and evolution

We now give some more general remarks about factorization and evolution.

The general statement of *factorization* is that the calculable, short-distance physics factorizes from incalculable, long distance one.

The general statement of *evolution* in momentum transfer Q is that physical quantities that characterize the long-distance (IR) behaviour have an evolution in momentum transfer due to emission of soft gluons, thus related to the IR divergences.

Factorization theorem

Factorization takes the form of a theorem, but needs to be defined for a specific case. We will consider the case of hadronic structure functions $F_a^{(h)}(x,Q^2)$, $Q^2=-q^2$, where q is the momentum transferred, defined as follows. The hadronic tensor

$$\begin{aligned} W^{(i)}_{\mu\nu}(p,q) &= \frac{1}{8\pi}\sum_{\sigma,n}\langle h(p,\sigma),in|J^{(i)\dagger}_{\mu}(0)|n,out\rangle\langle n,out|J^{(i)}_{\nu}(0)|h(p,\sigma),in\rangle(2\pi)^4\delta^4(p_n-q-p)\\ &= \frac{1}{8\pi}\sum_{\sigma}\int d^4x e^{iq\cdot x}\langle h(p,\sigma),in|J^{(i)\dagger}_{\mu}(x)J^{(i)}_{\nu}(0)|h(p,\sigma),in\rangle \end{aligned} \tag{22.30}$$

is written in terms of the electromagnetic currents, interacting with the hadronic states, and is relevant for example for DIS (with the leptonic part of the amplitude taken out). It is written as given tensor structures times structure functions,

$$\begin{aligned} W_{\mu\nu}(p,q) &= -\left(g_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)W_1(x,Q^2)+\left[p_\mu-q_\mu\frac{p\cdot q}{q^2}\right]\left[p_\nu-q_\nu\frac{p\cdot q}{q^2}\right]W_2(x,Q^2)\\ &\quad -i\epsilon_{\mu\nu}{}^{\rho\sigma}\frac{q_\nu p_\sigma}{2m}W_3(x,Q^2) \ , \end{aligned} \tag{22.31}$$

and

$$F_1(x,Q^2)=W_1(x,Q^2);\quad F_2(x,Q^2)=p\cdot qW_2(x,Q^2). \tag{22.32}$$

Then the factorization theorem for the structure functions (so really for the amplitudes) is

$$F_1^{(h)}(x,Q^2) = \sum_i\int_0^1\frac{d\xi}{\xi}C_1\left(\frac{x}{\xi},\frac{Q^2}{\mu^2},\alpha_s(\mu^2)\right)\phi_{i/h}\left(\xi,\epsilon,\alpha_s(\mu^2)\right)$$

$$F_2^{(h)}(x,Q^2) \;=\; \sum_i \int_0^1 d\xi\, C_2\left(\frac{x}{\xi},\frac{Q^2}{\mu^2},\alpha_s(\mu^2)\right)\phi_{i/h}\left(\xi,\epsilon,\alpha_s(\mu^2)\right). \tag{22.33}$$

Here $\phi_{i/h}$ are the distribution functions of partons i in the hadron h and contain all long-distance dependence (including all ϵ IR depencence), and C_i are IR safe functions independent of the external hadrons h called coefficient functions, and contain all the short-distance (Q^2) behaviour. μ is like a renormalization scale, more precisely a factorization scale, such that the split depends on it, but not $F_i(x,Q^2)$.

We can define also *valence distributions* as the difference between the quark and antiquark distributions,

$$\phi_{f/h}^{(\mathrm{val})}(\xi,\epsilon,\alpha_s(\mu^2)) = \phi_{f/h}(\xi,\epsilon,\alpha_s(\mu^2) - \phi_{\bar f/h}(\xi,\epsilon,\alpha_s(\mu^2))\;, \tag{22.34}$$

which obey an evolution equation (similar to the Altarelli-Parisi equation for the full quark and gluon distributions, but for evolution in μ),

$$\frac{d}{d\log\mu}\phi_{f/h}^{(\mathrm{val})}(x,\epsilon,\alpha_s(\mu^2)) = \int_x^1 \frac{d\xi}{\xi}P_f\left(\frac{x}{\xi},\alpha_s(\mu^2)\right)\phi_{f/h}^{(\mathrm{val})}(\xi,\epsilon,\alpha_s(\mu^2))\;, \tag{22.35}$$

where P_f are related to the quark- quark splitting functions P_{qq}.

Of course, factorization and evolution and more general concepts than in this particular case, but it was shown in order to see the general principles involved.

Important concepts to remember

- The KLN theorem states that the transition probabilities summed over all the initial and final states of given energy is free of (IR) divergences.
- Factorization means in general that calculable short distance physics factorizes from incalculable long-distance one.
- Evolution means in general that physical quantities evolve in momentum transfer due to the emission of soft particles, related to IR divergences.
- The factorization theorem says that structure functions (scalar functions parametrizing amplitudes) factorize in distribution functions characterizing long distance physics, and IR safe coefficient functions that contain all short distance (Q^2) behaviour, and are independent of the hadron.
- The factorization depends on an arbitrary renormalization (or rather, factorization) scale μ.
- Valence distributions obey evolution equation in μ.

Further reading: See chapter 14.3 and 14.1 in [2] for factorization and evolution, and book by Taizo Muta, "Foundations of Quantum Chromodynamics", chapter 6.3.3 for the KLN theorem.

Exercises, Lecture 22

1) Consider the $1/N$ expansion of the IR-divergent amplitudes at L-loop, $|A^{(L)}(\epsilon)\rangle$,

$$\begin{aligned}
|A^{(1)}(\epsilon)\rangle &= \frac{1}{N}I^{(1)}(\epsilon)|A^{(0)}\rangle + |A^{(1f)}(\epsilon)\rangle \\
|A^{(2)}(\epsilon)\rangle &= \frac{1}{N^2}I^{(2)}(\epsilon)|A^{(0)}\rangle + \frac{1}{N}I^{(1)}(\epsilon)|A^{(1)}(\epsilon)\rangle + |A^{(2f)}(\epsilon)\rangle \\
|A^{(3)}(\epsilon)\rangle &= \frac{1}{N^3}I^{(3)}(\epsilon)|A^{(0)}\rangle + \frac{1}{N^2}I^{(2)}(\epsilon)|A^{(1)}(\epsilon)\rangle + \frac{1}{N}I^{(1)}(\epsilon)|A^{(2)}(\epsilon)\rangle + |A^{(3f)}(\epsilon)\rangle
\end{aligned} \tag{22.36}$$

where $I^{(L)}(\epsilon)$ are divergent and $|A^{(Lf)}(\epsilon)\rangle$ are finite.

For 4-points in $\mathcal{N}=4$ SYM, we have

$$\begin{aligned}
I^{(1)}(\epsilon) &= \frac{1}{2\epsilon}\sum_{i=1}^{4}\sum_{j\neq i}^{a} T_i\cdot T_j\left(\frac{\mu^2}{-s_{ij}}\right)^{\epsilon} \\
T_i\cdot T_j &= T_i^aT_j^a; \quad s_{ij}=(k_i+k_j)^2; \quad T_i^a=(T^a)_{c_ib_i}=if_{c_iab_i} \\
I^{(2)}(\epsilon) &= -\frac{1}{2}\left[I^{(1)}(\epsilon)\right]^2 - N\zeta_2 c(\epsilon)I^{(1)}(2\epsilon) + \frac{c(\epsilon)}{4\epsilon}\left[-\frac{N\zeta_3}{2}\sum_{i=1}^{4}\sum_{j\neq i}^{4}T_i\cdot T_j\left(\frac{\mu^2}{-s_{ij}}\right)^{2\epsilon} + \hat{H}^{(2)}\right] \\
c(\epsilon) &= 1+\frac{\pi^2}{12}\epsilon^2 + \mathcal{O}(\epsilon^3) \\
\hat{H}^{(2)} &= -4[T_1\cdot T_2, T_2\cdot T_3]\log\left(\frac{s}{t}\right)\log\left(\frac{t}{u}\right)\log\left(\frac{u}{s}\right).
\end{aligned} \tag{22.37}$$

Note that T_i^a is a matrix that acts on, e.g., $(b_1b_2b_3b_4)$, to change b_i to c_i. Prove

$$\begin{aligned}
|A^{(2)}(\epsilon)\rangle &= \frac{1}{2N}I^{(1)}(\epsilon)|A^{(1)}(\epsilon)\rangle - \frac{1}{N}(\zeta_2+\epsilon\zeta_3)c(\epsilon)I^{(1)}(2\epsilon)|A^{(0)}\rangle \\
&\quad + \frac{1}{4N^2}\frac{c(\epsilon)}{\epsilon}\hat{H}^{(2)}|A^{(0)}\rangle + \frac{1}{2N}I^{(1)}(\epsilon)|A^{(1f)}(\epsilon)\rangle + |A^{(2f)}\rangle.
\end{aligned} \tag{22.38}$$

2) In the basis $|A_{[i]}\rangle$, $i=1,...,6$ (single trace) and $i=7,8,9$ (double trace), expand

$$I^{(1)}(\epsilon) = -\frac{1}{\epsilon^2}\begin{pmatrix} N\alpha_\epsilon & \beta_\epsilon \\ \gamma_\epsilon & N\delta_\epsilon \end{pmatrix}. \tag{22.39}$$

Find the $\mathcal{O}(1)$ matrices $\alpha_\epsilon(6\times 6), \beta_\epsilon(6\times 3), \gamma_\epsilon(3\times 6), \delta_\epsilon(3\times 3)$ in terms of

$$S=\left(\frac{\mu^2}{-s}\right)^{\epsilon}; \quad T=\left(\frac{\mu^2}{-t}\right)^{\epsilon}; \quad U=\left(\frac{\mu^2}{-u}\right)^{\epsilon}. \tag{22.40}$$

23 Lecture 23. Perturbatives anomalies: chiral and gauge.

In this lecture we start the analysis of perturbative anomalies in global and local symmetries.

Chiral invariance

For a linear symmetry

$$\phi^i(x) \to \phi^i(x) + \alpha^a (T^a)^i{}_j \phi^j(x) \ , \tag{23.1}$$

such that the Lagrangean changes only by a boundary term

$$\mathcal{L} \to \mathcal{L} + \alpha^a \partial^\mu J^a_\mu \ , \tag{23.2}$$

the Noether current is

$$j^a_\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial^\mu \phi)} (T^a)^i{}_j \phi^j - J^a_\mu(x) \ , \tag{23.3}$$

and is conserved, i.e.

$$\partial^\mu j^a_\mu = 0. \tag{23.4}$$

The action for a massless fermion in Euclidean space,

$$\mathcal{L} = \bar{\psi}\gamma^\mu D_\mu \psi \ , \tag{23.5}$$

where $\bar{\psi} = \psi^\dagger \gamma^4$ $(= \psi^\dagger i \gamma^0)$, is invariant under the chiral symmetry

$$\begin{array}{rcl} \psi(x) & \to & e^{i\alpha\gamma_5}\psi \simeq (1 + i\alpha\gamma_5)\psi \\ \bar{\psi}(x) & \to & \bar{\psi} e^{i\alpha\gamma_5} \simeq \bar{\psi}(1 + i\alpha\gamma_5). \end{array} \tag{23.6}$$

That means that a mass term breaks the symmetry, since

$$m\bar{\psi}\psi \to m\bar{\psi} e^{2i\alpha\gamma_g}\psi \neq m\bar{\psi}\psi. \tag{23.7}$$

Then we have the conserved chiral current

$$j^5_\mu = \bar{\psi}\gamma_\mu\gamma_5\psi; \quad \partial^\mu j^5_\mu = 0. \tag{23.8}$$

In QFT, we expect the conservation to occur for VEVs, i.e. the Ward identity

$$\langle \partial^\mu j^5_\mu(x) \rangle = 0. \tag{23.9}$$

We review the derivation: For a general global invariance $\phi \to \phi'$, renaming the field in the partition function from ϕ to ϕ',

$$\int \mathcal{D}\phi' e^{-S[\phi']} = \int \mathcal{D}\phi e^{-S[\phi]} \ , \tag{23.10}$$

IF the Jacobian for the transformation is one, i.e. the measure is invariant, $\mathcal{D}\phi = \mathcal{D}\phi'$, then

$$0 = \int \mathcal{D}\phi \left[e^{-S[\phi']} - e^{-S[\phi]} \right] = -\int \mathcal{D}\phi \delta S[\phi] e^{-S[\phi]}. \tag{23.11}$$

But if under the global symmetry $\delta S = 0$, under a *local* version of the global symmetry, the variation of the action is

$$\delta S = \sum_a \int d^4x (\partial^\mu \epsilon^a(x)) j^a_\mu(x) = -\sum_a \int d^4x \epsilon^a(x)(\partial^\mu j^a_\mu(x)) \ , \tag{23.12}$$

and substituting this in the above, we get

$$0 = \int d^4x \epsilon^a(x) \int \mathcal{D}\phi e^{-S[\phi]} \partial^\mu j^a_\mu(x). \tag{23.13}$$

Since $\epsilon^a(x)$ is arbitrary, we can take out the integral and write

$$0 = \int \mathcal{D}\phi e^{-S[\phi]} \partial^\mu j^a_\mu(x) \to \langle \partial^\mu j^a_\mu(x) \rangle = 0. \tag{23.14}$$

We can also repeat the same process for the partition function with insertions of $\phi(x)$, and obtain various Ward identities.

In conclusion, we obtained that the quantum non-conservation of the current, i.e. an anomaly, $\langle \partial^\mu j^a_\mu \rangle \neq 0$ appears when there is a nontrivial Jacobian for the measure.

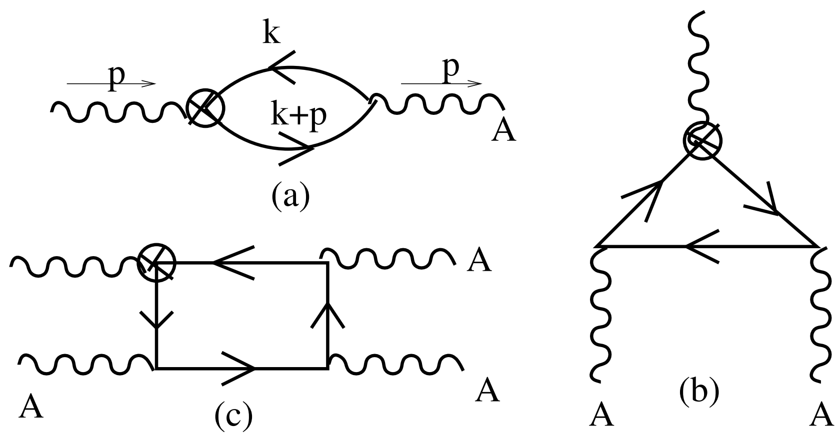

Figure 68: Anomalous diagrams. The crossed vertex connects to an outside current divergence, $\partial^\mu j_\mu(x)$, and the other ones to external gauge fields. Chiral fields run in the loop. (a) Anomalous bubble diagram in 2 dimensions. (b) Anomalous triangle diagram in 4 dimensions. (c) Anomalous box diagram in 6 dimensions.

Chiral anomaly

A fact that will be explained a bit better further on is that anomalies are *one-loop exact*, i.e. there are no perturbative or non-perturbative corrections to it, and they arise in even dimensions $d = 2n$ from one-loop diagrams with $n + 1 = d/2 + 1$ vertices, one of them being the symmetry current, or more precisely $\partial^\mu j^a_\mu$, and the other n being *gauge* currents coupling to external gauge fields A^{ext}_ν.

That means that in $d=2$, the anomalous diagram is a bubble with two vertices, in $d=4$ it is a triangle, in $d=6$ it is a box, in $d=8$ a pentagon and in $d=10$ a hexagon, see Fig.68.

d=2 Euclidean dimensions

We will calculate everything explicitly only in $d=2$ Euclidean dimensions, since it is easier, and there is nothing new appearing in higher dimensions other than longer calculations.

The Lagrangean is

$$\mathcal{L} = +\bar{\psi}\gamma_\mu D_\mu \psi + \frac{1}{4}F_{\mu\nu}^2 + \text{ghosts} \quad + \text{gauge fix} , \tag{23.15}$$

but we are interested only in the fermion part, since the gauge fields are external.

Then

$$\begin{aligned} S_0 &= \int d^2x \bar{\psi} \partial\!\!\!/ \psi \\ S_{\rm int} &= -ie \int d^2x \bar{\psi} A\!\!\!/ \psi. \end{aligned} \tag{23.16}$$

The anomaly is

$$\begin{aligned} \langle \partial^\mu j_\mu^5 \rangle &= \frac{\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \partial^\mu j_\mu^5 e^{+ie\int d^2x \bar{\psi} A\!\!\!/ \psi} e^{-S_0}}{\int \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{-S_0}} \\ &\simeq \langle \partial^\mu j_\mu^5(x)\rangle_0 + ie\int d^2y A_\nu^{ext}(y) \langle \bar{\psi}(y)\gamma_\nu \psi(y) \partial^\mu j_\mu^5(x)\rangle_0 + \mathcal{O}(e^2) \ , \end{aligned} \tag{23.17}$$

but the first term is zero, since

$$\langle \partial^\mu j_\mu^5(x)\rangle_0 = \partial_\mu \langle \bar{\psi}(x)\gamma^\mu\gamma_5\psi(x)\rangle \propto \text{Tr}[\gamma^\mu\gamma_5] = 0 \ , \tag{23.18}$$

which means that we have

$$\begin{aligned} \langle \partial^\mu j_\mu^5(x)\rangle &= ie\partial_{x^\mu}\int d^2y A_\nu^{ext}(y) \langle \psi_\beta(x)\bar{\psi}_{\alpha'}(y)\rangle_0 (\gamma_\nu)_{\alpha'\beta'} \langle \psi_{\beta'}(y)\bar{\psi}_\alpha(x)\rangle_0 (\gamma_\mu\gamma_5)_{\alpha\beta} \\ &= ie\frac{\partial}{\partial x^\mu}\int d^2y A_\nu^{ext}(y)\, \text{Tr}[S_F^0(x-y)\gamma_\nu S_F^0(y-x)\gamma_\mu\gamma_5]. \end{aligned} \tag{23.19}$$

The diagrammatic interpretation for this result is as a bubble diagram, with x^μ and ∂_{x^μ} at one vertex, and y^ν at the other, with the external gauge field there, i.e., exactly the diagram we said would contribute.

When going to p space, the diagram has now p_μ at both ends ($-p_\mu$ at the other, with the "all in" convention), and k and $k+p$ on the two lines of the loop, as in Fig.68a, i.e.

$$\langle p_\mu j_\mu^5 \rangle = e p_\mu A_\nu^{ext}(-p) \int \frac{d^2k}{(2\pi)^2} \text{Tr}[S_F^0(k)\gamma_\nu S_F^0(k+p)\gamma_\mu\gamma_5] = e p_\mu A_\nu^{ext}(-p) T_{\mu\nu}(p). \tag{23.20}$$

We now regularize using dimensional regularization, introducing the parameter μ, and using that for the massless fermion $S_F^0 = -i/p\!\!\!/ = -ip\!\!\!//p^2$, we get

$$T_{\mu\nu}(p) \ = \ -\mu^{2-d}\int \frac{d^dk}{(2\pi)^d} \frac{\text{Tr}[(\gamma_\alpha k_\alpha)\gamma_\nu (k+p)_\beta \gamma_\beta \gamma_\mu \gamma_5]}{k^2(k+p)^2}$$

$$= \quad -\mu^{2-d}\,\mathrm{Tr}[\gamma_\mu\gamma_5\gamma_\alpha\gamma_\nu\gamma_\beta]I. \tag{23.21}$$

Then the integral is

$$\begin{aligned} I &= \int \frac{d^dk}{(2\pi)^d}\frac{k_\alpha(k+p)_\beta}{k^2(k+p)^2} = \int_0^1 d\alpha \int \frac{d^dk}{(2\pi)^d}\frac{k_\alpha(k+p)_\beta}{(k^2+2\alpha k\cdot p+\alpha p^2)^2} \\ &= \int_0^1 d\alpha \int \frac{d^dk'}{(2\pi)^d}\frac{(k'-\alpha p)_\alpha(k'+(1-\alpha)p)_\beta}{[k'^2+p^2\alpha(1-\alpha)]^2} \\ &= \int_0^1 d\alpha \int \frac{d^dk'}{(2\pi)^d}\left[\frac{k'_\alpha k'_\beta}{[k'^2+p^2\alpha(1-\alpha)]^2} - \alpha(1-\alpha)p_\alpha p_\beta \frac{1}{[k'^2+p^2\alpha(1-\alpha)]^2}\right] \end{aligned} \tag{23.22}$$

where we first wrote $k'_\mu = k_\mu + \alpha p_\mu$ and then used Lorentz invariance to put to zero the integral with a single k'_α in the numerator.

Using the results from lecture 6 for the integrals, we get

$$\begin{aligned} I &= \int_0^1 d\alpha \left[\frac{\delta_{\alpha\beta}}{(2-d)(4\pi)^{d/2}}\frac{\Gamma(2-d/2)}{[\alpha(1-\alpha)p^2]^{1-d/2}} - \frac{\alpha(1-\alpha)p_\alpha p_\beta}{(4\pi)^{d/2}}\frac{\Gamma(2-d/2)}{[\alpha(1-\alpha)p^2]^{2-d/2}}\right] \\ &= \frac{1}{(4\pi)^{d/2}(p^2)^{1-d/2}}\left[\frac{\Gamma(2-d/2)}{2-d}\delta_{\alpha\beta} - \frac{p_\alpha p_\beta}{p^2}\Gamma(2-d/2)\right]\int_0^1 d\alpha[\alpha(1-\alpha)]^{d/2-1}. \end{aligned} \tag{23.23}$$

We note then that the α integral is 1 in $d=2$, and in the square bracket, the first term is divergent, and the second is finite, and proportional to $p_\alpha p_\beta$. But I is multiplied by a trace, and in $d=2$ we have

$$\mathrm{Tr}[\gamma_\mu\gamma_5\gamma_\alpha\gamma_\nu\gamma_\beta]p_\alpha p_\beta = 0\ , \tag{23.24}$$

so we can drop the last term, and we dimensionally regularize the trace, to obtain $\propto 1/(2-d)$ from the integral and $\propto (d-2)$ from the trace.

Indeed, using that $\gamma_\alpha\gamma_\nu\gamma_\alpha = -\gamma_\nu(\gamma_\alpha)^2 + 2\delta_{\alpha\nu}\gamma_\alpha = (2-d)\gamma_\nu$, we have

$$\delta^{\alpha\beta}\,\mathrm{Tr}[\gamma_\mu\gamma_5\gamma_\alpha\gamma_\nu\gamma_\beta] = (2-d)\,\mathrm{Tr}[\gamma_\nu\gamma_\mu\gamma_5] = -2i(2-d)\epsilon_{\mu\nu}\ , \tag{23.25}$$

where we have used (since "γ_5" $= \gamma_3 = -i\gamma_1\gamma_2$)

$$\gamma_\mu\gamma_\nu = i\epsilon_{\mu\nu}\gamma_5 + \delta_{\mu\nu}. \tag{23.26}$$

Then

$$T_{\mu\nu}(p) = +\mu^{2-d}(2-d)2i\epsilon_{\mu\nu}\frac{\Gamma(2-d/2)}{(2-d)(4\pi)^{d/2}(p^2)^{1-d/2}}\int_0^1 d\alpha[\alpha(1-\alpha)]^{\frac{d}{2}-1}\ , \tag{23.27}$$

so that when $d=2$ we obtain

$$T_{\mu\nu}(p) \to \frac{i}{2\pi}\epsilon_{\mu\nu}. \tag{23.28}$$

Finally then,

$$\langle p_\mu j^5_\mu(p)\rangle = \frac{iep_\mu}{2\pi}\epsilon_{\mu\nu}A^{ext}_\nu(-p)\ , \tag{23.29}$$

or going back to position space

$$\langle \partial^\mu j_\mu^5 \rangle = \frac{ie}{2\pi}\epsilon_{\mu\nu}\frac{(\partial_\mu A_\nu^{ext} - \partial_\nu A_\mu^{ext})}{2} = \frac{ie}{2\pi}\tilde{F}^{ext} , \tag{23.30}$$

where

$$\tilde{F}^{ext} \equiv \frac{1}{2}\epsilon^{\mu\nu}F_{\mu\nu}^{ext}. \tag{23.31}$$

Going back to Minkowski signature,

$$\langle \partial^\mu j_\mu^5 \rangle = \frac{e}{2\pi}\tilde{F}^{ext}. \tag{23.32}$$

d=4.

In 4 Euclidean dimensions, we have a triangle anomaly, and the calculation is similar, though longer. A similar expression is obtained in the end,

$$\langle \partial^\mu j_\mu^5(x) \rangle = -\frac{ie^2}{16\pi^2}\tilde{F}_{\mu\nu}^{ext}F_{\mu\nu}^{ext} , \tag{23.33}$$

where

$$\tilde{F}_{\mu\nu}^{ext} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}. \tag{23.34}$$

When going to Minkowski space,

$$\langle \partial^\mu j_\mu^5(x) \rangle = \frac{e^2}{16\pi^2}\tilde{F}_{\mu\nu}^{ext}F^{\mu\nu,ext} = \frac{e^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}^{ext}F_{\rho\sigma}^{ext}. \tag{23.35}$$

Properties of anomaly

There is a theorem, called the Adler-Bardeen theorem, that the anomaly is one-loop only (and only comes from the given polygonal graphs). It can be proven rigorously, and in the path integral we will see it in the next lecture, but here we give just some plausibility arguments. Note that sometimes the anomaly is called the Adler-Bell-Jackiw anomaly.

We note that in $d = 2$, the anomaly was the result of the trace of 4 gammas and a γ_5, being proportional to $(d-2)\epsilon_{\mu\nu}$, times a log divergent diagram with two massless fermion propagators, $\int d^2k(1/k)(1/k) \propto 1/(d-2)$. We can see that at higher loops we will not have this near cancellation anymore. Also at higher points we will not have it anyrmore.

In $d = 4$, we also have a trace with 6 gammas and a γ_5, being proportional to $(d-4)\epsilon_{\mu\nu\rho\sigma}$, and a log divergent diagram with 3 massless fermion propagators, $\int d^4k(k_\alpha/k^2)(k'_\beta/k^2)(k''_\gamma/k^2) \propto p_\alpha \int d^4k/k^4 \propto 1/(d-4)$ (since there are no Lorentz structures with 3 indices, this is the only possibility). Again at higher loops or points, we don't have this cancellation anymore.

Note that we have only considered massless fermions, massive fermions, with mass term allowed by the symmetry, do not contribute to the anomaly.

The anomaly cannot be removed by a local counterterm, so it is genuine.

What we can do in 4 dimension is write the anomalous contribution as a boundary term,

$$F_{\mu\nu}\tilde{F}^{\mu\nu} = 4\partial_\mu(\epsilon^{\mu\nu\rho\sigma}A_\nu\partial_\rho A_\sigma) , \tag{23.36}$$

which allows us to subtract it from the current, defining

$$\tilde{j}^5_\mu = j^5_\mu - \frac{e^2}{4\pi^2}\epsilon_{\mu\nu\rho\sigma}A^\nu\partial^\rho A^\sigma \ , \tag{23.37}$$

which is conserved, and the conserved charge

$$\tilde{Q}_5 = \int d^3x \tilde{j}^5_0. \tag{23.38}$$

But the new current is not gauge invariant, since under a gauge transformation $\delta A_\mu = \partial_\mu \alpha$, the current changes by

$$\delta\tilde{j}^5_\mu = -\frac{e^2}{4\pi^2}\epsilon_{\mu\nu\rho\sigma}\partial^\rho A^\sigma\partial^\nu\alpha. \tag{23.39}$$

Yet the charge,

$$\tilde{Q}_5 = Q_5 - S_{CS}[A_i] \ , \tag{23.40}$$

where S_{CS} is called the *Chern-Simons term* or Chern-Simons action,

$$S_{CS}[A_i] = \frac{e^2}{4\pi^2}\int d^3x \epsilon^{ijk}A_i\partial_j A_k \ , \tag{23.41}$$

is invariant, since the variation of the current is a total derivative.

Chiral anomaly in nonabelian gauge theories.

We can embed the chiral symmetry in a nonabelian theory of massless fermions. The Euclidean action

$$S = \int d^4x \bar{\psi}\gamma^\mu D_\mu\psi \ , \tag{23.42}$$

where $D_\mu = \partial_\mu + gA^a_\mu T^a$, still has an abelian chiral symmetry. Not surprisingly, it is just a trace, i.e.

$$\partial^\mu j^5_\mu = \frac{g^2}{16\pi^2}\,\mathrm{Tr}[F^{\mu\nu}\tilde{F}_{\mu\nu}] \ , \tag{23.43}$$

and it is usually called also chiral anomaly, or *singlet anomaly.*

Note that

$$\mathrm{Tr}[F^{\mu\nu}\tilde{F}_{\mu\nu}] = \partial_\mu\left[4\epsilon^{\mu\nu\rho\sigma}\,\mathrm{Tr}\left(A_\nu\partial_\rho A_\sigma + \frac{2}{3}A_\nu A_\rho A_\sigma\right)\right] \ , \tag{23.44}$$

so that the redefined current is

$$\tilde{j}^5_\mu = j^5_\mu - \frac{g^2}{4\pi^2}\epsilon^{\mu\nu\rho\sigma}\,\mathrm{Tr}\left(A_\nu\partial_\rho A_\sigma + \frac{2}{3}A_\nu A_\rho A_\sigma\right). \tag{23.45}$$

The redefined chiral charge is

$$\tilde{Q}_5 = Q_5 - S_{CS}[A_i] = Q_5 - \frac{g^2}{4\pi^2}\int d^3x \epsilon^{ijk}\,\mathrm{Tr}\left[A_i\partial_j A_k + \frac{2}{3}A_i A_j A_k\right]. \tag{23.46}$$

But now, unlike the abelian case, under a gauge transformation

$$A_i \to UA_iU^{-1} + \frac{1}{g}\partial_i UU^{-1} \ , \tag{23.47}$$

the Chern-Simons term transforms by (2π times) an integer m,

$$S_{CS}[A_i] \to S_{CS}[A_i] + \frac{1}{4\pi^2}\int d^3x \epsilon^{ijk}\,\mathrm{Tr}[\partial_i U U^{-1}\partial_j U U^{-1}\partial_k U U^{-1}]\;, \tag{23.48}$$

the extra term being a topological quantity characterizing the 3 dimensional gauge transformation called a *winding number m*.

Therefore now the Chern-Simons term is not invariant anymore under "large gauge transformations" which are gauge transformations of nontrivial winding number, therefore not connected smoothly with the identity. That means that now $\tilde{Q}_5$ is not gauge invariant anymore, so the global symmetry is explicitly broken by the anomaly.

The term appearing in the anomaly gives an important topological quantity characterizing the 4 dimensional gauge configuration,

$$n = \frac{g^2}{16\pi^2}\int d^4x\,\mathrm{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}] \tag{23.49}$$

is called the *Pontryagin index* or *instanton number.*

Gauge anomalies

Up to now we have discussed anomalies in global symmetries, which are good anomalies, that have physical implications, and can be measured.

But there is another type of anomalies, in gauge symmetries, that are bad anomalies, and signal the breakdown (inconsistency) of the quantum theory, so must be cancelled to have a good theory.

If we have chiral fermions, $\psi_{R,L} = (1\pm\gamma_5)/2\psi$ coupled to gauge fields, we have a potential anomaly in gauge invariance. For instance, for a ψ_L, with Euclidean action

$$S = \int d^4x \bar{\psi}\frac{1+\gamma_5}{2}\not{D}\,\psi\;, \tag{23.50}$$

we have a Noether current for the gauge symmetry

$$j^a_\mu = -\bar{\psi}\frac{1+\gamma_5}{2}\gamma_\mu T^a\psi\;, \tag{23.51}$$

that is covariantly conserved, $(D^\mu j_\mu)^a = 0$.

In this case, similarly to the global case, at the quantum level there is a potential anomaly that can come from a Jacobian for the transformation of the measure, or otherwise from a triangle graph, with $(D^\mu j_\mu)^a$ in one vertex, and A_ν, A_ρ in the other, similar to the case in Fig.68b.

One obtains the triangle anomaly

$$\langle D^\mu j^a_\mu\rangle = \partial_\mu\left\{\frac{1}{24\pi^2}\epsilon^{\mu\nu\rho\sigma}\,\mathrm{Tr}\left[T^a\left(A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma\right)\right]\right\}. \tag{23.52}$$

We see that now we have a T^a inside the trace, and the anomaly is proportional to

$$d_{abc} \equiv \mathrm{Tr}[T^a(T^bT^c + T^cT^b)]. \tag{23.53}$$

However, as we said, this anomaly must cancel, so we must add up represenations of fields such that the total anomaly is zero for a good theory. This is what happens for example in the Standard Model, as we will see in lecture 25.

We also notice that the coefficients of the quadratic and cubic terms inside the trace are different from the singlet anomaly case. We will explain this next lecture.

Finally, why do we need to cancel the anomaly? Since now gauge symmetry kills degrees of freedom (in 4 dimensions, a vector boson has 3 degrees of freedom, but using gauge invariance we reduce them to two), so if gauge symmetry would be broken at the quantum level, it would mean that there are a different number of degrees of freedom at the classical and quantum levels, which is nonsensical.

Important concepts to remember

- Anomalies mean non-conservation of a symmetry current at the quantum level, due to the non-invariance of the path integral measure.
- Anomalies are one-loop exact and arise in even dimensions $d = 2n$ from polygon graphs with one $\partial^\mu j_\mu$ and n gauge currents coupled to external gauge fields.
- In 2 Minkowski dimensions, the anomaly is $e/(4\pi)\epsilon^{\mu\nu}F^{ext}_{\mu\nu}$, and in 4 dimensions it is $e^2/(16\pi^2)\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$.
- The anomaly is the derivative of the Chern-Simons current.
- The Chern-Simons action changes by a winding number under large gauge transformations.
- The anomaly is given by the topological invariant Pontryagin index or instanton number.
- Gauge anomalies are bad, and must be cancelled for the consistency of the theory.

Further reading: See chapter 8.1, 8.2, 8.4 in [5], chapter 19.1, 19.2, 19.4 in [3], 22.3 and 22.4 in Weinberg vol.II.

Exercises, Lecture 23

1) Write down (up to a coefficient, which would be calculated from the one-loop diagram) the abelian chiral anomaly $\langle \partial^\mu j^5_\mu(x) \rangle$ in $d = 6$ dimensions, and the corresponding $\tilde{j}^5_\mu$ and $\tilde{Q}_5$, justifying your work based on one-loop diagrams and symmetry arguments.

2) Consider 2 dimensional conformal invariance, symmetry of flat space field theory under the complex holomorphic transformation $z' = f(z)$, $\bar{z}' = f(\bar{z})$. Is an anomaly for it allowed, or not, and why?

24 Lecture 24. Anomalies in path integrals- the Fujikawa method; consistent vs. covariant anomalies and descent equations.

In this lecture we will consider anomalies from path integrals following Fujikawa, and we will explain the various nonabelian anomalies and how they are related through the descent equations.

Chiral basis vs. V-A basis

But first, some equivalent way to express anomalies. In $d = 4$, one can consider together the anomalies by defining an axial vector current

$$j_\mu^{5,S} = \bar{\psi}\gamma_\mu\gamma_5 S\psi. \tag{24.1}$$

For $S = \mathbb{1}$, we have the chiral current, or singlet current, and for $S = T^a$, we have the nonabelian current.

The anomaly comes from triangle graphs with one $\mathcal{O} = D^\mu j_\mu^{5,S}$ and two A's, so

$$\langle D^\mu j_\mu^{5,S}\rangle \propto \epsilon^{\mu\nu\rho\sigma}\partial_\mu \operatorname{Tr}[S(\#A_\nu\partial_\rho A_\sigma + \#A_\mu A_\rho A_\sigma)]. \tag{24.2}$$

The coefficients will be fixed in the second part of the lecture, from the descent equations.

Sometimes one considers, instead of the chiral basis $\psi_{L,R}$ for the spinors, Dirac (nonchiral) fermions, coupled to *vector* and *axial vector* gauge fields,

$$\mathcal{L} = \bar{\psi}_i\gamma^\mu(\partial_\mu + V_\mu^a T_a(V) + A_\mu^a T_a(A)\gamma_5)\psi^i + \frac{1}{4g_V^2}(F_{\mu\nu}^a(V))^2 + \frac{1}{4g_A^2}(F_{\mu\nu}^a(A))^2. \tag{24.3}$$

It can be rewritten in terms of the chiral basis as

$$\mathcal{L} = \bar{\psi}_{i,L}\gamma^\mu D_{\mu,L}\psi^i + \bar{\psi}_{i,R}\gamma^\mu D_{\mu,R}\psi^i + \frac{1}{4g_L^2}(F_{\mu\nu}^{(L)a})^2 + \frac{1}{4g_R^2}(F_{\mu\nu}^{(R)a})^2\ , \tag{24.4}$$

where

$$D_{\mu L,R} = \partial_\mu + A_{\mu L,R}^a T_a \tag{24.5}$$

and

$$V_\mu = \frac{A_\mu^L + A_\mu^R}{2}; \quad A_\mu = \frac{A_\mu^L - A_\mu^R}{2}. \tag{24.6}$$

In the V,A basis for instance, we can write the chiral (singlet) anomaly ($S = 1$) as an AVV piece (diagram) and a AAA piece (diagram), with

$$\epsilon^{\mu\nu\rho\sigma}\left[F_{\mu\nu}^{lin}(V)F_{\rho\sigma}^{lin}(V) + \frac{1}{3}F_{\mu\nu}^{lin}(A)F_{\rho\sigma}^{lin}(A)\right] e^j \operatorname{Tr}[T_b^j T_c^j] \tag{24.7}$$

where e^j is the abelian charge of fermion j.

Anomaly in the path integral - Fujikawa method.

As we already mentioned, the anomaly in the path integral appears because of the non-invariance of the path integral measure. We want therefore to expand the ψ, $\bar{\psi}$ fields in eigenfunctions of $i\not{D}$ and see how it transforms under the chiral invariance.

We write therefore

$$
\begin{aligned}
\psi(x) &= \sum_n a_n \phi_n(x) \\
\bar{\psi}(x) &= \sum_n \phi_n^\dagger(x) \bar{b}_n \ ,
\end{aligned}
\tag{24.8}
$$

where

$$
i\gamma^\mu D_\mu \phi_n(x) = \lambda_n \phi_n(x) \ , \tag{24.9}
$$

and the eigenfunctions are orthonormal, i.e.

$$
\int dx \phi_n^\dagger(x)\phi_m(x) = \delta_{nm}. \tag{24.10}
$$

Then the path integral measure is written as

$$
\mathcal{D}\bar{\psi}\mathcal{D}\psi = \prod_n d\bar{b}_n \prod_m da_m. \tag{24.11}
$$

Under a local chiral transformation on a_n and $\bar{b}_n$, we get

$$
\psi(x) \to \sum_n a_n' \phi_n(x) = \psi'(x) = e^{i\alpha(x)\gamma_5}\psi(x) = \sum_n a_n e^{i\alpha(x)\gamma_5}\phi_n(x) \ , \tag{24.12}
$$

so, by multiplication with $\phi_n(x)$ and integration, we get

$$
\begin{aligned}
a_n' &= \sum_m \int dx \phi_n^\dagger(x) e^{i\alpha(x)\gamma_5}\phi_m(x) a_m \equiv \sum_m C_{nm} a_m \\
\bar{b}_n' &= \sum_m \bar{b}_m \int dx \phi_m^\dagger(x) e^{i\alpha(x)\gamma_5}\phi_n(x) \equiv \sum_m \bar{b}_m C_{mn}.
\end{aligned}
\tag{24.13}
$$

Then the path integral measure transforms as

$$
\prod_n d\bar{b}_n' \prod_m da_m' = (\det C)^2 \prod_n d\bar{b}_n \prod_m da_m \ , \tag{24.14}
$$

meaning that the Jacobian is given by (considering an infinitesimal chiral transformation)

$$
\begin{aligned}
C &= \mathbb{1} + \hat{\alpha} + \mathcal{O}(\alpha^2) \\
\hat{\alpha}_{nm} &\equiv \int dx \phi_n^\dagger(x)\gamma_5\phi_m(x) \ ,
\end{aligned}
\tag{24.15}
$$

so that

$$
(\det C)^{-1} = e^{-\operatorname{Tr}\log C} = e^{-\operatorname{Tr}\hat{\alpha} + \mathcal{O}(\alpha^2)} = 1 - \int dx \alpha \sum_n \phi_n^\dagger(x)\gamma_5\phi_n(x) + \mathcal{O}(\alpha^2). \tag{24.16}
$$

This result is formally divergent, so it needs to be regularized, with a regulator that maintains gauge invariance.

One possibility is to use *zeta function regularization* (see QFT I), by turning the integral into

$$\sum_n \frac{1}{\lambda_n^s}\phi_n^\dagger(x)\gamma_5\phi_n(x)\ , \tag{24.17}$$

and then taking $s \to 0$. This is done by analogy with Riemann's zeta function,

$$\zeta(s) = \sum_{n\geq 1}\frac{1}{n^s}. \tag{24.18}$$

But here we will instead follow the regularization used by Fujikawa, which is

$$\sum_n \phi_n^\dagger(x)\gamma_5 e^{-(\lambda_n/M)^2}\phi_n(x)\ , \tag{24.19}$$

and then take $M \to \infty$. Note that the above sum is understood (since $i\not{D}\,\phi_n = \lambda_n\phi_n$) as

$$\begin{aligned}\sum_n \phi_n^\dagger(x)\gamma_5 e^{\frac{\not{D}^2}{M^2}}\phi_n(x) &= \sum_n \langle n|x\rangle\gamma_5 e^{\frac{\not{D}^2}{M^2}}\langle x|n\rangle \\ = \ \mathrm{Tr}\left[|x\rangle\gamma_5 e^{\frac{\not{D}^2}{M^2}}\langle x|\right] &= \mathrm{Tr}_\alpha\int\frac{d^dk}{(2\pi)^d}e^{-ik\cdot x}\gamma_5 e^{\frac{\not{D}^2}{M^2}}e^{ik\cdot x}\ ,\end{aligned} \tag{24.20}$$

where in the last form we have expressed the trace in the momentum basis, and the remaining trace is over the spinor indices.

Lemma.

$$e^{-ik\cdot x}e^{\frac{\not{D}^2}{M^2}}e^{ik\cdot x} = \left(e^{-\frac{k^2}{M^2}}e^{\frac{\not{D}^2+2ik^\mu D_\mu}{M^2}}\right)\mathbb{1}(x)\ , \tag{24.21}$$

where

$$\not{D}^2 = D^2 - ie\sigma_{\mu\nu}F_{\mu\nu}\ , \tag{24.22}$$

and $\sigma_{\mu\nu} = i/4[\gamma_\mu,\gamma_\nu]$ (note that we are in Minkowski spacetime).

Proof. First we note that

$$\not{D}^2 = D_\mu D_\mu\gamma^\mu\gamma^\nu = D_\mu D_\nu(\delta^{\mu\nu}+\gamma^{\mu\nu}) = D^2 - i\sigma_{\mu\nu}(D_\mu D_\nu - D_\nu D_\mu) = D^2 - ie\sigma_{\mu\nu}F_{\mu\nu}. \tag{24.23}$$

Then, when acting on a function $f(x)$, we have

$$e^{-ik\cdot x}\not{D}^2 e^{ik\cdot x}f(x) = \not{D}^2 f(x) + \left(e^{-ik\cdot x}\not{D}^2 e^{+ik\cdot x}\right)f(x)\ , \tag{24.24}$$

where $\left(e^{-ik\cdot x}\not{D}^2 e^{+ik\cdot x}\right) = (-k^2 + 2ik^\mu D_\mu)$ (the second term is from one D_μ acting on $f(x)$ and one on $e^{ik\cdot x}$, and the terms with F are part of the $\not{D}^2 f(x)$ piece). That means that

$$e^{-ik\cdot x}e^{\frac{\not{D}^2}{M^2}}e^{ik\cdot x} = \left(e^{-\frac{k^2}{M^2}}e^{\frac{\not{D}^2+2ik^\mu D_\mu}{M^2}}\right)\mathbb{1}(x)\ , \tag{24.25}$$

q.e.d.lemma.

Then we can write our regulated sum as

$$\sum_n \phi_n^\dagger(x)\gamma_5 e^{\frac{\not{D}^2}{M^2}}\phi_n(x) = \int \frac{d^dk}{(2\pi)^d} e^{-\frac{k^2}{M^2}} \operatorname{Tr}\left(\gamma_5\left[1+\frac{\not{D}^2+2ik^\mu D_\mu}{M^2}+\frac{(\not{D}^2+eik^\mu D_\mu)^2}{2!M^4}+\mathcal{O}\left(\frac{1}{M^6}\right)\right]\right). \tag{24.26}$$

But the first term, though potentially divergent, is actually zero, since $\operatorname{Tr}[\gamma_5]=0$. The same applies to the term with $\operatorname{Tr}[\gamma_5(2ik^\mu D_\mu)]/M^2$.

d=2.

We now specialize to two dimensions. Then the remaining term with $\not{D}^2$ is the only nonzero one, since $\int d^2k/M^2(2\pi)^2 e^{-k^2/M^2} = \pi/(2\pi)^2 = 1/4\pi$ (double Gaussian integral) is the only finite term (the further ones are suppressed by powers of M as $M\to\infty$). But also

$$\gamma_\mu\gamma_\nu = i\gamma_5\epsilon_{\mu\nu}+\delta_{\mu\nu} \Rightarrow \operatorname{Tr}[\gamma_5\sigma_{\mu\nu}] = \frac{i}{4}\operatorname{Tr}[\gamma_\mu\gamma_\nu\gamma_5] - \nu\leftrightarrow\mu = -\frac{\epsilon_{\mu\nu}}{4}\operatorname{Tr}[\,\mathbb{1}] - (\mu\leftrightarrow\nu) = -\epsilon_{\mu\nu}\ , \tag{24.27}$$

so finally

$$\sum_n \phi_n^\dagger(x)\gamma_5 e^{\frac{\not{D}^2}{M^2}}\phi_n(x) = \int \frac{d^2k/M^2}{(2\pi)^2} e^{-\frac{k^2}{M^2}}(+ie\epsilon^{\mu\nu}F_{\mu\nu}) = \frac{e}{2\pi}\left(\frac{i}{2}\epsilon^{\mu\nu}F_{\mu\nu}\right). \tag{24.28}$$

Then the anomaly, coming from the transformation of the path integral measure, is

$$\mathcal{D}\psi\mathcal{D}\bar\psi \to \mathcal{D}\psi\mathcal{D}\bar\psi(\det C)^{-2} \simeq \mathcal{D}\psi\mathcal{D}\bar\psi\left(1-\frac{e}{\pi}\right)\int dx\alpha(x)\tilde F(x). \tag{24.29}$$

d=4.

In $d=4$, because

$$\int \frac{d^4k/M^4}{(2\pi)^4} e^{-\frac{k^2}{M^2}} = \frac{\pi^2}{(2\pi)^4}\ , \tag{24.30}$$

the only nonzero term is the one with $1/M^4$. Indeed, now also $\operatorname{Tr}[\gamma_5\sigma_{\mu\nu}] \propto \operatorname{Tr}[\epsilon_{\mu\nu\rho\sigma}\gamma^{\rho\sigma}]=0$ as well, so the only nonzero term is from

$$\int \frac{d^4k/M^4}{(2\pi)^4} e^{-\frac{k^2}{M^2}}\frac{1}{2}\operatorname{Tr}[\gamma_5(\not{D}^2)^2] = \int \frac{d^4k/M^4}{(2\pi)^4} e^{-\frac{k^2}{M^2}}\frac{-e^2}{2}\operatorname{Tr}[\gamma_5\sigma_{\mu\nu}\sigma_{\rho\sigma}]F_{\mu\nu}F_{\rho\sigma}. \tag{24.31}$$

But since

$$\operatorname{Tr}[\gamma_5\sigma_{\mu\nu}\sigma_{\rho\sigma}] = -\frac{1}{4}\operatorname{Tr}[\gamma_5\gamma_{\mu\nu}\gamma_{\rho\sigma}] = -\frac{i}{4}\epsilon_{\mu\nu\rho\sigma}\operatorname{Tr}[1] = -i\epsilon_{\mu\nu\rho\sigma}\ , \tag{24.32}$$

where we have used

$$\gamma_{\mu\nu}\gamma_{\rho\sigma} = i\epsilon_{\mu\nu\rho\sigma}\gamma_5 - (\delta_{\mu\rho}\delta_{\nu\sigma}-\delta_{\mu\sigma}\delta_{\nu\rho}) \tag{24.33}$$

(which can be proven by first putting $\mu=1,\nu=2,\rho=3,\sigma=4$ and identifying the two sides, then $\mu=\rho$, etc.), we have

$$\sum_n \phi_n^\dagger(x)\gamma_5 e^{\frac{\not{D}^2}{M^2}}\phi_n(x) = \frac{ie^2}{32\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = \frac{ie^2}{16\pi^2}F_{\mu\nu}\tilde F^{\mu\nu}\ , \tag{24.34}$$

and the anomaly is

$$\mathcal{D}\psi\mathcal{D}\bar{\psi} \rightarrow \mathcal{D}\psi\mathcal{D}\bar{\psi}\left(1 - \frac{i}{8\pi^2}\int dx\alpha(x)F_{\mu\nu}\tilde{F}^{\mu\nu}\right) , \tag{24.35}$$

so

$$\langle\partial^\mu j_\mu^5\rangle = \frac{-ie^2}{8\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}. \tag{24.36}$$

Consistent vs. covariant anomaly

The nonabelian anomaly can be found up to an overall coefficient from consistency conditions found by Wess and Zumino.

The conditions come from the fact that the anomaly must be the gauge variation of the effective action $\Gamma(A)$. We define the gauge variation of the gauge field for a left-handed fermion anomaly,

$$\Delta_{\Lambda_L}A_\mu^a = (D_\mu(A)\Lambda_L)^a = \partial_\mu\Lambda_L^a + [A_\mu, \Lambda_L]^a , \tag{24.37}$$

and we introduce the operator of gauge variation which varies the action with respect to A, and then the multiplies by the gauge variation of A,

$$\delta_{\Lambda_L} \equiv X_L(\Lambda_L) \equiv (\partial_\mu\Lambda_L + [A_\mu, \Lambda_L])^a\frac{\partial}{\partial A_\mu^a}. \tag{24.38}$$

If the anomaly G_a is the gauge variation of the effective action $\Gamma(A)$, it follows that

$$\delta_{\Lambda_L}\Gamma(A) = X_L(\Lambda_L)\Gamma(A) = \int d^4x\Lambda_L^aG_a. \tag{24.39}$$

But then the group algebra,

$$[X_L(\Lambda_L^{(1)}), X_L(\Lambda_L^{(2)})] = X_L([\Lambda_L^{(1)}, \Lambda_L^{(2)}] , \tag{24.40}$$

implies that when acting on the effective action we get

$$\int d^4x\Lambda_L^{(2)}\delta_{\Lambda_L^{(1)a}}G_a(A) - 1 \leftrightarrow 2 = \int d^4x[\Lambda_L^{(1)}, \Lambda_L^{(2)}]^aG_a , \tag{24.41}$$

which is the *Wess-Zumino consistency condition.*

The unique solution to this equation (up to a normalization constant) is

$$G_a(A) = \text{Tr}\left[T^ad\left(A\wedge dA + \frac{1}{2}A\wedge A\wedge A\right)\right]. \tag{24.42}$$

Substituting in the consistency condition, and partially integrating d onto $\Lambda_L^{(2)}$ on the left hand side, we obtain the condition

$$\int\text{Tr}\left[d\Lambda_L^{(2)}\delta_{\Lambda_L^{(1)}}\left(A\wedge dA + \frac{1}{2}A\wedge A\wedge A\right)\right] - (1\leftrightarrow 2) = \int\text{Tr}\left[\delta_{[\Lambda_L^{(1)},\Lambda_L^{(2)}]}\left(A\wedge dA + \frac{1}{2}A\wedge A\wedge A\right)\right] , \tag{24.43}$$

which is left as an exercise to verify.

Then the anomaly

$$G_a = D_\mu(A)\frac{\delta\Gamma(A)}{\delta A_\mu^a} = \frac{i}{24\pi^2}\epsilon^{\mu\nu\rho\sigma}\operatorname{Tr}\left[T^a\partial_\mu\left(A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma\right)\right] \tag{24.44}$$

satisfies the consistency conditions, so is a *consistent anomaly*, but is not covariant under gauge transformations. This anomaly has physical meaning. However, theoretically, it is better to work with covariant expressions. We can add a local counterterm to the effective action or to the current J_μ^a and find a *covariant anomaly*, which is important theoretically. Then

$$\tilde{G}_a = D_\mu J_a^\mu + D_\mu X_a^\mu = G_a + D_\mu X_a^\mu = D_\mu \tilde{J}_a^\mu. \tag{24.45}$$

However, the covariant anomaly is not consistent (i.e., it does not satisfy the Wess-Zumino consistency conditions).

Descent equations

Both the consistent and the covariant anomaly appear in the so-called (Stora-Zumino) *descent equations*, that start with the Chern form in $2n+2$ dimensions, $F^{\wedge(n+1)}$. Here as usual $F = dA + A\wedge A = \frac{1}{2}F_{\mu\nu}dx^\mu\wedge dx^\nu$.

For instance, we start with the Chern form in 6 dimension, $\omega_6 = \operatorname{Tr}[F\wedge F\wedge F] \propto \epsilon^{\mu_1\dots\mu_6}\operatorname{Tr}[F_{\mu_1\mu_2}F_{\mu_3\mu_4}F_{\mu_5\mu_6}]$. Then $d\omega_{2n+2}=0$, since

$$d(\operatorname{Tr} F^{n+1}) = (n+1)\operatorname{Tr}[(dF)F^n] = (n+1)\operatorname{Tr}[(DF)F^n] = 0\ , \tag{24.46}$$

where inside the trace we can replace dF with DF, but $DF=0$ by the Bianchi identity. That means that at least locally, though in fact globally, as we can explicitly check by exterior differentiation, $\omega_{2n+2} = d\omega_{2n+1}$, where ω_{2n+1} is called the *Chern-Simons form* in $2n+1$ dimensions.

For instance $\omega_6 = d\omega_5$, and we find

$$\omega_5 = \operatorname{Tr}\left[dA\wedge dA\wedge A + \frac{3}{2}dA\wedge A\wedge A\wedge A + \frac{3}{5}A^{\wedge 5}\right]. \tag{24.47}$$

Under a general variation δA, the Chern form varies as

$$\delta\omega_{2n+2} = (n+1)\operatorname{Tr}[(D\delta A)F^n] = (n+1)\operatorname{Tr}[D(\delta A F^n)] = (n+1)d\operatorname{Tr}[\delta A F^n]\ , \tag{24.48}$$

where we have used the Bianchi identity $DF=0$ to take out D, and when D is outside the trace it becomes d (since there are no more indices for him to act on). Finally we find

$$\delta d\omega_{2n+1} = d\delta\omega_{2n+1}\ , \tag{24.49}$$

which means that

$$\delta\omega_{2n+1} = (n+1)\operatorname{Tr}[\delta A F^n] + d(...). \tag{24.50}$$

As a consequence, the field equation of the Chern-Simons action $\int\omega_{2n+1}$ is $F^{\wedge n}=0$, so is covariant.

But under a gauge variation,

$$\delta_{\rm gauge}\int\omega_{2n+1}=(n+1)\int{\rm Tr}[D\Lambda F^{\wedge n}]=(n+1)\int d({\rm Tr}[\Lambda F^{\wedge n}])=0\ , \tag{24.51}$$

where again we have taken out D using the Bianchi identity, and outside the trace, D becomes d. That means that

$$\delta_{\rm gauge}\omega_{2n+1}=dY. \tag{24.52}$$

In fact, in our example, we find explicitly

$$\delta_{\rm gauge}\omega_5={\rm Tr}\left[(d\Lambda)d\left(A\wedge dA+\frac{1}{2}A\wedge A\wedge A\right)\right]=d\left[{\rm Tr}\,\Lambda d\left(A\wedge dA+\frac{1}{2}A\wedge A\wedge A\right)\right]. \tag{24.53}$$

So, in general, ω_{2n+2} is the singlet anomaly in $d=2n+2$ dimensions. $\delta_{\rm gauge}\omega_5$ gives the *consistent nonabelian anomaly*, and the field equation of ω_5 gives the *covariant* nonabelian anomaly. In general,

$$\begin{aligned}\omega_{2n+2}&=d\omega_{2n+1}\\ \delta_{\rm gauge}\omega_{2n+1}&=d\Lambda^a G_a(cons.)\\ {\rm Tr}\left(T^a\frac{\delta}{\delta A}\int\omega_{2n+1}\right)&=\tilde{G}_a(cov.).\end{aligned} \tag{24.54}$$

These are the *descent equations*.

d=2. As the simplest example, we consider 2 dimensions. The consistent anomaly is

$$G^a=c\partial_\mu A_\nu^a\epsilon^{\mu\nu}\ , \tag{24.55}$$

and the extra current piece is

$$X_a^\mu=c\epsilon^{\mu\nu}A_{\nu,a}\ , \tag{24.56}$$

leading to

$$\tilde{G}=D_\mu J^\mu+D_\mu X^\mu=c\partial_\mu A_\nu\epsilon^{\mu\nu}+c\epsilon^{\mu\nu}D_\mu A_\nu=c(\partial_\mu A_\nu-\partial_\nu A_\mu+[A_\mu,A_\nu])\epsilon^{\mu\nu}=cF_{\mu\nu}\epsilon^{\mu\nu}. \tag{24.57}$$

d=4. The more relevant example is of 4 dimensions. Here the consistent anomaly is

$$G_a=c\,{\rm Tr}\left[T^a d\left(A\wedge dA+\frac{1}{2}A\wedge A\wedge A\right)\right]\ , \tag{24.58}$$

and the extra current is

$$X_a=c\,{\rm Tr}\left[T_a\left(dA\wedge A+A\wedge dA+\frac{3}{2}A\wedge A\wedge A\right)\right]\ , \tag{24.59}$$

leading to

$$\tilde{G}_a=G_a+DX_a=c\,{\rm Tr}[T_a 3F\wedge F]. \tag{24.60}$$

Important concepts to remember

- We can write the anomaly in the chiral basis or the V-A basis.
- In the path integral, the anomaly arises because of the non-invariance of the measure. The Fujikawa method regularizes the sum over eigenfunctions of $i\not{D}$ with $e^{\not{D}^2/M^2}$.
- In the Fujikawa method, it is obvious that the anomaly is one-loop exact.
- The nonabelian anomaly must satisfy the Wess-Zumino consistency condition, leading to the consistent anomaly, but is not covariant. It has physical significance.
- By adding a local counterterm to the effective action, or to the current, we can construct a covariant anomaly, that is however not consistent. It is theoretically useful.
- The various anomalies appear in the descent equation that starts from the Chern form in 2n+2 dimensions. The Chern form is the singlet anomaly, $\omega_{2n+2} = d\omega_{2n+1}$ and the gauge variation of the Chern-Simons form ω_{2n+1} gives the consistent anomaly. The field equation of the CS form is the covariant anomaly.

Further reading: See chapter 8.3 in [5], chapter 19.2 in [3], 22.2 in Weinberg vol.II.

Exercises, Lecture 24

1) Regularize $\sum_n \phi_n^\dagger(x)\gamma_5\gamma_{\mu\nu}\phi_n(x)$ with $e^{\not{D}^2/M^2}$ in $d=4$ and calculate it.

2) Calculate the $d=6$ anomaly in the Fujikawa method.

3) Prove the relation stated in class,
d=2

$$\begin{aligned}
G^a &= c\partial_\mu A_\nu^a \epsilon^{\mu\nu} \\
X_a^\mu &= c\epsilon^{\mu\nu} A_{\nu,a} \\
\tilde{G} &= c\epsilon^{\mu\nu} F_{\mu\nu} = G + D_\mu X^\mu = D_\mu J^\mu + D_\mu X^\mu .
\end{aligned} \tag{24.61}$$

d=4

$$\begin{aligned}
G_a &= c\,\mathrm{Tr}\left[T^a d\left(A\wedge dA + \frac{1}{2}A\wedge A\wedge A\right)\right] \\
X_a &= c\,\mathrm{Tr}\left[T_a\left(dA\wedge A + A\wedge dA + \frac{3}{2}A\wedge A\wedge A\right)\right] \\
\tilde{G}_a &= G_a + DX_a = c\,\mathrm{Tr}[T^a 3F\wedge F]\ ,
\end{aligned} \tag{24.62}$$

and find the expressions in $d=6$.

25 Lecture 25. Physical applications of anomalies: 't Hooft's UV-IR anomaly matching conditions; anomaly cancellation.

In this lecture we present physical applications of anomalies, as well as theoretical applications, for restricting the set of consistent models.

$\pi^0 \to \gamma\gamma$ **decay**

The most famous physical application of anomalies is to the decay of the neutral pion, into two photons.

The pions π^a, with $a = 1, 2, 3$, are related to the divergence of the axial vector current $\partial^\mu j_\mu^{5,a}$, i.e.

$$j_\mu^{5,a} \sim \bar{\psi}(x) \frac{\sigma^a}{2} \gamma_5 \gamma_\mu \psi(x) \;, \tag{25.1}$$

where the σ^a are Pauli matrices for a flavor $SU(2)$ group. More precisely, in quark models, where the pions are $\bar{q}q$ objects (one quark and one antiquark), we have

$$j_\mu^{5,a}(x) = 2im_q \left(\bar{q}\gamma_5 \frac{\sigma^a}{2} \gamma_\mu q \right) . \tag{25.2}$$

But as we explained, we cannot really think of hadrons as having a fixed number of partons, but as having distribution functions for partons inside the handrons, due to the strong QCD interactions ("hadronization"). The claim is that under hadronization, the above relation is replaced by

$$\partial^\mu j^{5a}_{\mu,had.} = f_\pi m_\pi^2 \pi^a \;, \tag{25.3}$$

where f_π is the pion decay constant, that can be independently measured from experiment. The above relation is called the *PCAC relation*, standing for "Partially Conserved Axial vector Current" relation.

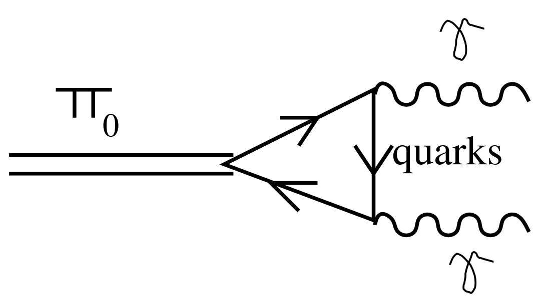

Figure 69: Anomaly for Π^0 decaying into 2 photons via a quark loop.

The relation is OK for $\pi^\pm$, corresponding to the Pauli matrices $\sigma^\pm$, but is not quite correct for π^0, corresponding to the Pauli matrix σ_3, since there is a decay of the pion into two photons. More precisely, we hace a decay of $\partial^\mu j_\mu^{5,3}$ into two photons through a triangle anomaly diagram, as in Fig.69: first the current divergence $\partial^\mu j_\mu^{5,3}$ turns (mostly) into a π^0, after which the π^0 dacays through a one-loop chiral fermion triangle diagram into two γ's.

Then, more precisely, we have

$$\begin{aligned}\partial^\mu j_\mu^{5(A),\pm} &= f_\pi m_\pi^2 \pi^\pm(x)\\ \partial^\mu j_\mu^{5(A),3} &= f_\pi m_\pi^2 \pi^3(x) + c\frac{e^2}{16\pi^2}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\ ,\end{aligned} \tag{25.4}$$

where the extra term is the anomalous contribution. The relation is correct to all orders in α_s, but only to first order in α.

The coefficient c can be calculated as follows. If it was an electron in the loop, we would have $c=1$, since we have already calculated the $U(1)$ chiral anomaly. But for quarks in the loop, we have $c = N_c/6$, since

$$\pi^0 \sim \bar\psi\gamma_5\frac{\tau_3}{2}\psi;\quad \tau_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix};\quad \psi = \begin{pmatrix}u\\ d\end{pmatrix}\ , \tag{25.5}$$

and so

$$c = \mathrm{Tr}\left[\frac{\tau_3}{2}Q^2\right] = \frac{N_c}{2}\left[\left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2\right] = \frac{N_c}{6}. \tag{25.6}$$

Note that the Q^2 is the charge squared of the quarks, appearing because of the quarks coupling to each of the external photons with the charge Q.

We now consider the $m_\pi \to 0$ limit, which is consistent, since the pion is the lightest state in QCD by approximately an order of magnitude.

We consider the matrix element of $j_\mu^{A,3}$ in between a photon state and the vacuum, i.e. the decay amplitude from vacuum to two photons, through the anomaly. The matrix is dominated by the intermediate state of a pion, i.e.

$$\langle\epsilon^1, p^1; \epsilon^2, p^2|j_\mu^{A,3}(0)|0\rangle \simeq \sum_{\vec q}\langle\epsilon^1, p^1; \epsilon^2, p^2|\pi^0, \vec q\rangle\langle\pi^0, \vec q|j_\mu^{A,3}(0)|0\rangle\ , \tag{25.7}$$

where ϵ^1, ϵ^2 are the polarizations of the two photons, and p^1, p^2 are their two momenta.

On the other hand, the matrix element of the pion operator between the vacuum and the pion state is

$$\langle\pi^0, \vec q|\pi^0(0)|0\rangle = \frac{1}{\sqrt{2\omega_q}}\ , \tag{25.8}$$

which is just the relativistic wave function normalization. But by the PCAC relation (25.4), considering that the anomaly part has no matrix element between the pion and the vacuum, we can replace $\pi^0(0)$ with $q^\mu j_\mu^{A,3}/(f_\pi m_\pi^2)$ between the states, and given that $q_\mu^2 = m_\pi^2$, we obtain

$$\langle\pi^0, \vec q|j_\mu^{A,3}|0\rangle = \frac{q_\mu f_\pi}{\sqrt{2\omega_q}}. \tag{25.9}$$

Now we consider the PCAC relation (25.4) in between the 2-photon state $\langle\epsilon^1, p^1; \epsilon^2, p^2|$ and the vacuum $|0\rangle$.

On the right hand side we obtain

$$\langle \epsilon^1, p^1; \epsilon^2, p^2 | \frac{N_c}{6} \frac{e^2}{16\pi^2} \left(\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}\right) |0\rangle = \frac{N_c}{6} \frac{e^2}{16\pi^2} \left(8\epsilon^{\mu\nu\rho\sigma} \epsilon^1_\mu \epsilon^2_\nu p^1_\rho p^2_\sigma\right) , \tag{25.10}$$

where $F_{\mu\nu}|\epsilon, p\rangle = 2\epsilon_{[\mu} p_{\nu]} |\epsilon, p\rangle$, and an extra factor of 2 comes because each F can act on both (ϵp) pair. Therefore now the (25.4) relation in between the states becomes (note that $e^2/2\pi^2 = 2\alpha/\pi$)

$$\langle \epsilon^1, p^1; \epsilon^2, p^2 | \partial^\mu j^{A,3}_\mu |0\rangle = \left(\frac{\alpha}{\pi} \frac{N_c}{3}\right) \epsilon^{\mu\nu\rho\sigma} \epsilon^1_\mu \epsilon^2_\nu p^1_\rho p^2_\sigma. \tag{25.11}$$

Finally then, using the relation (25.7) with the normalization (25.8), the above relation gives the amplitude for a pion to go to two photons,

$$\mathcal{A}(\pi^0 \to 2\gamma) \equiv \langle \epsilon^1, p^1; \epsilon^2, p^2 | \pi^0, \vec{q}\rangle = \frac{\alpha}{\pi} \frac{N_c}{3} \frac{1}{f_\pi} \epsilon^{\mu\nu\rho\sigma} \epsilon^1_\mu \epsilon^2_\nu p^1_\rho p^2_\sigma , \tag{25.12}$$

since the $\sum_{\vec{q}} 1/\sqrt{2\omega_q}$ gives the propagator $1/q^2$, which cancels the $q^\mu q_\mu$ coming from $\langle \pi^0, \vec{q} | \partial^\mu j^{A,3}_\mu |0\rangle$.

Then, using the relation between the decay probability Γ and the amplitude given in QFTI (eq. 19.49),

$$\Gamma(\pi^0 \to 2\gamma) = \frac{1}{2m_\pi} \int \frac{d^3 p_1}{(2\pi)^3} \frac{1}{2\omega_1} \int \frac{d^3 p_2}{(2\pi)^3} \frac{1}{2\omega_2} \frac{1}{2} \left(\sum_{\text{pols.}} |\mathcal{A}(\pi^0 \to 2\gamma)|^2\right) (2\pi)^4 \delta^4(q - p^1 - p^2). \tag{25.13}$$

After some calculation that will not be reproduced here, we find

$$\Gamma(\pi^0 \to 2\gamma) = \frac{1}{64\pi} \left(\frac{\alpha}{\pi} \frac{N_c}{3} \frac{1}{m_\pi}\right)^2 (m_\pi)^3. \tag{25.14}$$

This is tested experimentally to a high degree of accuracy, and one verifies that $N_c = 3$.

Nonconservation of baryon number in electroweak theory

The second important physical application of anomalies is to the non-conservation of baryon number. The gauge group of the Standard Model is $SU(3)_c \times SU(2) \times U(1)_Y$, where the electroweak gauge group is $SU(2) \times U(1)_Y$.

We first describe the fermion field content of the Standard Model. We have the lepton and quark left-handed $SU(2)$ doublets

$$L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}; \quad Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} , \tag{25.15}$$

where the left part of the electron is

$$e_L = \frac{1 - \gamma_5}{2} e^- , \tag{25.16}$$

and its right part is the full matrix

$$R = \frac{1 + \gamma_5}{2} e^- , \tag{25.17}$$

since the neutrino does not have a right-handed part in the (minimal form of the) Standard Model. A singlet right-handed neutrino is the simplest extension of the Standard Model.

We have a similar relation for the quarks, and we have 3 generations of fermions, which will be implicit in the notation.

Therefore in Euclidean space, the lepton part of the action in interaction with the electroweak gauge group (ignoring the color gauge group) is

$$S_{\rm leptons} = \int d^4x \left[\bar{R}\gamma_\mu(\partial_\mu - ig' B_\mu)R + \bar{L}\gamma_\mu \left(\partial_\mu - \frac{ig'}{2}B_\mu + \frac{ig}{2}A^a_\mu \sigma^a\right) L\right] , \quad (25.18)$$

whereas the quark part of the action is

$$S_{\rm quarks} = \int d^4x \left[\bar{Q}_L\gamma_\mu \left(\partial_\mu + \frac{ig'}{2}Y_L B_\mu + \frac{ig}{2}A^a_\mu\sigma^a\right) Q_L + \sum_{i=1,2} \bar{Q}_{R(i)}\gamma_\mu \left(\partial_\mu + \frac{ig'}{2}Y_{R(i)}B_\mu\right) Q_{R(i)}\right]. \quad (25.19)$$

Note that here B_μ is the $U(1)_Y$ gauge field, so the Y_L and Y_R are hypercharges for the L and R fields, and A^a_μ are the $SU(2)$ gauge fields.

We note that the above action has as conserved quantities the baryon number B and the lepton number L (we can consider also the independent lepton numbers, for each generation, L_e, L_μ, L_τ, but these are approximate symmetries). A quark has $B = 1/3$ such that a baryon, made up of 3 quarks, has baryon number 1 ($B(B) = 1$). On the other hand, all leptons have lepton number $L = 1$. Then B and L are classical symmetries of the above Standard Model action.

However, we have chiral fermions, so in fact B and L are anomalous: we have an abelian (singlet) anomaly in a nonabelian theory, with

$$\partial_\mu j^B_\mu = \frac{N_{\rm gen}g^2}{16\pi^2}\,{\rm Tr}[F^{\mu\nu}\tilde{F}_{\mu\nu}]. \quad (25.20)$$

Integrating over space and over time between t_1 and t_2, and using $\int d^4x \partial^\mu j_\mu = \int d^3x [j^B_0]^{t_2}_{t_1} = B(t_2) - B(t_1)$, we obtain for the difference in baryon number at times t_1 vs. t_2,

$$B(t_2) - B(t_1) = \frac{N_{\rm gen}g^2}{16\pi^2}\int_{t_1}^{t_2} dt \int d^3x\,{\rm Tr}[F^{\mu\nu}\tilde{F}_{\mu\nu}]. \quad (25.21)$$

But there is a nontrivial field configuration called an *instanton*, obeying

$$F_{\mu\nu} = \tilde{F}_{\mu\nu} \quad (25.22)$$

in Euclidean space, which means that for them, the topological number

$$n = \frac{g^2}{16\pi^2}\int d^4x\,{\rm Tr}[F^{\mu\nu}\tilde{F}_{\mu\nu}] = \frac{g^2}{4\pi^2}\frac{1}{4}\int d^4x\,{\rm Tr}[F^2_{\mu\nu}] \quad (25.23)$$

called *instanton number* is proportional to the on-shell instanton action (in Euclidean space), so

$$S^{(E)} = \int d^4x \frac{1}{4}\,{\rm Tr}[F^2_{\mu\nu}] = \frac{4\pi^2}{g^2}n , \quad (25.24)$$

and the difference in baryon number is an integer,

$$B(t_2) - B(t_1) = N_{\rm gen} n. \tag{25.25}$$

Therefore the difference in baryon number is defined by the instanton number, and the transition probability is given by the (classical) saddle point of the path integral with the given boundary condition, i.e. $e^{-S^{(E)}}$,

$$\langle B(t_2)|B(t_1)\rangle \simeq e^{-S^{(E)}}. \tag{25.26}$$

However, in the vacuum corresponding to our Universe, $e^{-\frac{4\pi}{\alpha_{\rm weak}}} \sim e^{-4\pi\cdot 30}$ is negligible for the lifetime of the Universe.

However, in the high temperature medium of the Big Bang, when the coupling is large, the probability becomes of order 1 so, by symmetry, transitions in baryon number will equalize it, resulting in $B = 0$, or an equal number of baryons and antibaryons. However, we observe in our Universe a net baryon number, which means that in the initial stages of the Big Bang there was already a *baryon asymmetry.* The question is, how is it possible, given the mechanism of wiping out an initial baryon asymmetry that we just saw? This is a very important question in theoretical physics, for which there are various models, but none is perfect. Sakharov in the 1980's already had enumerated the necessary conditions to create a baryon asymmetry, but as of yet, there is no perfect model.

The $U(1)$ problem

The last of the three important physical applications of anomalies is called the $U(1)$ problem. We will not explain all the details of its resolution, among other things because it uses information that will be given later on in the course, but rather we will sketch it.

We will see in the last lecture of the course that there is an effective symmetry $SU(2)_L \times SU(2)_R$, due to the near-masslessness of the up and down quarks (the u and d quarks are nearly massless, for them we can use a light quark effective theory, and the c, b and t are very heavy, for which we can use a heavy quark approximation; the s quark is intermediate). This symmetry is spontaneously broken to a diagonal $SU(2)$. We will also see later on in the course that whenever we break spontaneously a symmetry, there is a so-called Goldstone boson appearing, a massless scalar associated with the broken symmetry directions. In QCD, the $SU(2)_L \times SU(2)_R$ symmetry is approximate, so we have approximate Goldstone bosons, the 3 pions π^a (with masses much smaller than the masses of the other states in QCD), corresponding to the 3 broken generators (for a broken $SU(2)$).

But the actual symmetry of the QCD Lagrangean is $U(N_f) \times U(N_f)$ (N_f is the number of massless flavours, here 2), acting on the quarks as

$$\begin{array}{rcl} \psi & \rightarrow & e^{i\alpha^a T^a}\psi \\ \psi & \rightarrow & e^{i\alpha^a T^a \gamma_5}\psi. \end{array} \tag{25.27}$$

So between the actual $U(2) \times U(2)$ and the observed $SU(2)_L \times SU(2)_R$, the difference is an $U(1) \times U(1)$. One of the $U(1)$'s, acting as (the $U(1)$ is the trace of the $U(N_f)$, where we replace T^a by $1\!\!1$)

$$\psi \rightarrow e^{i\alpha}\psi\ , \tag{25.28}$$

is just the hadron number, which is conserved. But then we still have the other,

$$\psi \to e^{i\alpha\gamma_5}\psi \ , \tag{25.29}$$

which is an abelian chiral symmetry. Before the anomaly was understood, it was thought that there could only be two possibilities: the symmetry is there, but we don't see such a symmetry in the real world; or the symmetry is spontaneously broken, but then by the Goldstone theorem we should see a Goldstone boson corresponding to this broken symmetry. However, there is no fourth pion, so that is also not true. That was then the "$U(1)$ problem", and its resolution is of course, that the symmetry is anomalous, so is broken (though not spontaneously).

't Hooft's UV-IR anomaly matching conditions

We now turn to theoretical applications of anomalies, namely applications for model building. The first such application is due to 't Hooft, and is a very useful consistency condition, which simply put states that the anomaly is independent of the energy scale, so it should give the same result, for instance in the UV and in the IR.

It is useful, since we have the effective field theory approach started by Wilson, that will be studied later on in the course, which states that for a given energy range, we can use a theory in terms of some fields, without worrying if the fields are truly fundamental, as long as we include in the Lagrangean all the possible higher dimensional operators (even though they will in general not be renormalizable operators). This point of view allows us, say, to use the Standard Model, without worrying that there is at least a Planck scale (and maybe other scales, like susy scale, GUT scale, etc.) at which what we think of as the fundamental degrees of freedom will change.

But in that case, the anomaly matching conditions act as an important check of the fact that we are using the right degrees of freedom at a certain scale, given knowledge about the degrees of freedom at some other scale.

As an example, consider the anomaly of a global $U(1)$ current, and in the IR consider that we can use the nearly massless fermionic degrees of freedom that may be composite (only massless chiral fermions contribute to the anomaly, as we saw). For instance, in QCD we could consider the n and the p (which are composites of 3 quarks) as these degrees of freedom for energies higher than the m_p, m_n, but not too high so that we need to consider perturbative QCD. On the other hand, in the UV we can use the fundamental degrees of freedom (in the case of QCD, use the quarks) to calculate the anomaly. The two calculations should match.

Proof. To prove the anomaly matching condition, couple the global $U(1)$ current to a gauge field, i.e., gauge the symmetry. Then add *free* chiral fermions that only couple to the gauge field, in such a way as to cancel the anomaly. Indeed, now that we have a local symmetry, we need to cancel this local (gauge) anomaly for consistency of the quantum theory, as we said.

But local anomaly cancellation, i.e. consistency of the quantum theory, should persist both in the UV and in the IR. Now turn off the gauge coupling, $g \to 0$, going back to the global case. Subtract the anomaly of the free chiral fermions, which is independent of the scale, since now the fermions are truly free (don't couple to anything), so they don't know what an energy scale means. The result is that the global anomaly of the original system is independent of scale. *q.e.d.*

Note that the anomaly is purely one-loop, so it can be easily calculated at an energy scale using the perturbative degrees of freedom available at that energy scale.

Anomaly cancellation.

The second important theoretical application is the cancellation of gauge anomalies, which is an important consistency condition for any model we might write (without it, the quantum theory is inconsistent).

Here we study how it can happen that we have anomaly cancellation.

In $d = 4$, the anomaly is proportional to d^{abc}, which is proportional to $\mathrm{Tr}[T^a\{T^b, T^c\}]$, and in turn is related to $C_3(G)$, the third Casimir of the gauge group. That means that if the Lie algebra of the group has no $C_3(G)$, there is no d^{abc}, and thus no anomaly. Such groups are called *safe groups*.

Nearly all of the classical groups are safe. In particular, the B_n series, i.e. $SO(2n+1)$, the C_n series, i.e. $Sp(2n)$, the D_n series, i.e. $SO(2n)$, except $SO(6)$, as well as the exceptional groups G_2, F_4, E_6, E_7, E_8, are all safe. The only unsafe classical groups are the A_N series, i.e. $SU(N)$, for $N \geq 3$, and $SO(6) \simeq SU(4)$. Note that $SU(2) \simeq SO(3)$ is safe, as is $SO(4) \simeq SO(3) \times SO(3)$.

Even if a group is unsafe, in some representation R we might still have $d_R^{abc} = 0$, in which case we say the representation is safe.

Otherwise, if we have an unsafe representation of an unsafe group, to cancel the anomaly we need to combine several species of fermions such as to cancel the anomaly.

The Standard Model.

The most relevant example that we will study is the Standard Model, with gauge group $SU(3)_c \times SU(2) \times U(1)_Y$.

A potential anomaly appears for the unsafe group $SU(3)_c$, with unsafe representations. However, the $SU(3)$ anomaly in the Standard Model is cancelled in a trivial way, since both left and right fermions couple in the same way with the $SU(3)_c$ gauge field, so the total anomaly vanishes.

Then, $SU(2)$ is a safe group, so there is no anomaly with an $SU(2)$ gauge field at each of the three corners of the triangle. However, the absence of d^{abc} says nothing about combining the $SU(2)$ gauge field with another in the triangle diagrams, more precisely about the $SU(2)$ contributing to the $\partial^\mu j_\mu$ anomaly of some other gauge field. But the $SU(3)_c$ couples in the same way for left and right, so that doesn't contribute to these mixed anomalies either.

Thus we need to check only the $U(1)_Y$ (hypercharge), for the pure $U(1)$ anomaly, and the mixed anomaly with $SU(2)$. The potentially anomalous diagrams are then with a $U(1)$ gauge fields in $\partial^\mu j_\mu$, and two $SU(2)$ gauge fields in the others (the $SU(2)$ contribution to the $U(1)$ anomaly), and the one with 3 $U(1)$ gauge fields (the $U(1)$ contribution to the $U(1)$ anomaly).

Together, we can write these conditions as

$$\sum_{\text{fermion representations}} \mathrm{Tr}\left[T^a\left(A \wedge dA + \frac{1}{2} A \wedge A \wedge A\right)\right] = 0. \tag{25.30}$$

The T^a is the generator coupling to the $\partial^\mu j^a_\mu$, so for the $U(1)$ we have $T^a = 1$.

Then the $U(1) - SU(2) - SU(2)$ anomaly has $T^a = 1$, but of course we need to multiply it by the charge Y_L, and the $A \wedge dA + 1/A \wedge A \wedge A$ is proportional to $\sigma^b \sigma^c$. Then the condition becomes

$$\sum_{\text{doublets,L}} Y_L \,\text{Tr}[\sigma^b \sigma^c] \propto \delta^{bc} \sum_{\text{doublets}} Y_L = 0 , \tag{25.31}$$

giving the condition

$$\sum_{\text{doublets,L}} Y_L = 0. \tag{25.32}$$

Note that here we count doublets only (since the $SU(2)$ couples only to doublets, which doublets are only left-handed), but we also need to count color where needed.

For the $U(1)^3$ anomaly, again we have $T^a = 1$, but must be multiplied by Y, and $A \wedge dA$ is proportional to $\mathbb{1}$, but again multiplied with Y_L for each gauge field, for a total condition of

$$\sum_{\text{left-handed}} (Y_L)^3 - \sum_{\text{right-handed}} (Y_R)^3 = 0. \tag{25.33}$$

Note that here we must count each element of a doublet, and also count color.

To verify these conditions, we consider the hypercharge assignments of the Standard Model.

For *quarks*, we have

$$Y_L = 1/3; \quad Y_R(1) = 4/3; \quad Y_R(2) = -2/3 , \tag{25.34}$$

since the left quarks are doublets, but for the right quarks, we have two independent elements (not a doublet), 1 and 2.

For the *leptons*, we have

$$Y_L = -1; \quad Y_R(1) = 0; \quad Y_R(2) = -2 , \tag{25.35}$$

and the same applies, the left leptons are doublets, but the right ones are two independent elements, 1 and 2.

We now verify the two conditions. For the $U(1) - SU(2) - SU(2)$ anomaly, we get

$$N_c \times \frac{1}{3} + 1 \times (-1) = \frac{N_c}{3} - 1 , \tag{25.36}$$

and it only cancels for $N_c = 3$, as it should.

For the $U(1)^3$ anomaly, we get

$$\left[N_c \times 2 \times \left(\frac{1}{3}\right)^3 + 2 \times (-1)^3\right] - \left[N_c \times \left(\frac{4}{3}\right)^3 + N_c \times \left(-\frac{2}{3}\right)^3 + (-2)^3\right] = 0 , \tag{25.37}$$

giving

$$\frac{N_c}{3^3}(2 - 4^3 + 2^3) - 2 + 8 = -2(N_c - 3) = 0 , \tag{25.38}$$

which again only cancels for $N_c = 3$, as it should.

Important concepts to remember

- The PCAC relations relate the divergence of the axial vector $SU(2)$ current with the pions, modulo the anomaly in the σ^3 component of the current.
- The neutral pion π^0 decays into 2 photons because of the anomaly, via a diagram where $\partial^\mu j_\mu^{5,3}$ turns into a pion and then into 2 photons via a quark triangle.
- Baryon number is changed by instantons, because of the anomaly in the baryon current, and with a probability $\langle B(t_2)|B(t_1)\rangle \sim e^{-S_{\rm inst}}$.
- In the current Universe, baryon number change is irrelevant, due to the smallness of the coupling $e^{-4\pi/\alpha_{\rm weak}}$ giving a small probability. But in the initial Big Bang it is relevant, and it would wipe out any initial baryon asymmetry, hence the baryon asymmetry problem.
- The potential extra $U(1)$ does not have a Goldstone boson, since it is a chiral symmetry broken by anomalies.
- The anomaly in the UV (computed with the UV degrees of freedom) should match the anomaly in the IR (computed with the IR degrees of freedom).
- The gauge anomaly should cancel in a consistent model. There are safe groups, safe representations, and otherwise we need to combine species of fermions.
- The unsafe groups are $U(1)$, $SU(N)$, $N \geq 3$ and $SO(6) \simeq SU(4)$.

Further reading: See chapter 8.5 in [5], chapter 19.3 in [3], 22.1, 22.5, 22.6 in Weinberg vol.II.

Exercises, Lecture 25

1) Check that the $SU(5)$ GUT model, with fermionic field content (ignoring the Higgs sector):

-one (anti)fundamental representation $\bar{5}$ and one antisymmetric representation $5\times 4/2 = 10$, both left-handed, is free of gauge anomalies, given that

$$\mathrm{Tr}_5(T^a\{T^b,T^c\}) = d^{abc} = \mathrm{Tr}_{10}(T^a\{T^bT^c\}). \tag{25.39}$$

2) Calculate the $U(1)$ chiral anomaly of the $SU(5)$ GUT by using the UV-IR 't Hooft matching conditions.

26 Lecture 26. The operator product expansion, renormalization of composite operators and anomalous dimension matrices

In this lecture we will learn about renormalizing composite operators, a tool for describing that called the operator product expansion, and anomalous dimension matrices for composite operators.

Composite operators are important objects in gauge theories, since observables are gauge invariant, so need to be composite.

But introducing a composite operator at a point x, $\mathcal{O}(x)$, requires additional renormalization beyond the one in the Lagrangean, since putting several fields at a single point introduces new divergences.

Important examples of composite operators are the energy-momentum tensor $T_{\mu\nu}$, condensates $\bar{\psi}\psi$, etc.

We introduce the composite operators in the theory by adding a source term for them, $J_{\mathcal{O}} \cdot \mathcal{O}$ in the generating functional, i.e.

$$Z_{\mathcal{O}}[J_{\mathcal{O}}] = \int \mathcal{D}[\phi] e^{-S+\int J_{\mathcal{O}} \cdot \mathcal{O}(x)} \ , \tag{26.1}$$

in the same way as we did for fundamental fields in

$$Z[J] = \int \mathcal{D}\phi e^{-S+\int J \cdot \phi(x)}. \tag{26.2}$$

The resulting Green's functions for $\mathcal{O}$ can be thought of as Green's functions for sets of fields at $x_1, ..., x_n$ when we identify these points (they all converge, and are equal to x), e.g. for $\mathcal{O} = \phi_1 ... \phi_n(x)$,

$$\langle \mathcal{O}(x) \rangle \sim \langle \phi_1(x_1) ... \phi_n(x_n) \rangle |_{x_1 = x_2 = ... = x_n = x} = G^{(n)12...n}(x_1, ..., x_n)|_{x_1 = ... = x_n = x}. \tag{26.3}$$

Of course, because of the divergences we mentioned, this is not so well defined; we will use the method of the operator product expansion to make better sense of this.

One can reverse the above process and consider a method of regularization for the additional divergences appearing in composite operators called "point splitting", that means pulling the consituents apart, by having each field in $\mathcal{O}$ at a different point.

For composite operators, there is also *operator mixing*, if there are Feynman diagrams that mix the operators. For this to happen, the operators have to have the same charges under the symmetries respected by the Lagrangean, since then the interaction Lagrangean $\mathcal{L}_{\rm int}$, appearing in the Feynman diagrams, also respects the symmetries.

Therefore in general we have the renormalization of the type

$$\mathcal{O}_n[\{\phi_j\}, g, ...] = \sum_m {Z_n}^m \mathcal{O}_m^{\rm ren}[\{Z_j \phi_j^{\rm ren}\}, Z_g g^{\rm ren}, ...] \ , \tag{26.4}$$

meaning that besides the renormalization of the fields and couplings, there is an independent renormalization of the operators, that mixes them.

In particular, in a Yang-Mills theory,

$$\mathcal{O}_j[A^a_\mu, g, ...] = \sum_k Z_j{}^k \mathcal{O}_k^{\rm ren}\left[Z_3^{1/2} A_\mu^{a,ren}, \frac{Z_1}{Z_3^{3/2}} g^{\rm ren}, ...\right]. \tag{26.5}$$

The matrix of renormalization factors $\mathcal{Z}_n{}^m$ can be determined by Feynman diagrams with insertions of $\mathcal{O}_n, \mathcal{O}_m$. We will see shortly a concrete example of such a calculation.

This renormalization is also multiplicative, like the renormalization of fundamental fields. Moreover, it closes under renormalization (there are no outside operators).

In general, $Z_n{}^m$ is nontrivial, but in the particular case of conserved currents it is 1, i.e. conserved currents do not renormalize, $Z_j = 1$.

For instance, in QED, the Ward-Takahashi identity says that

$$\partial_{z^\mu}\langle 0|T[\psi_\alpha(x)\bar\psi_\beta(y)j_\mu(z)]|0\rangle = -ie\delta(x-z)\langle 0|T[\psi_\alpha(y)\bar\psi_\beta(x)]|0\rangle + ie\delta(y-z)\langle 0|T[\psi_\alpha(x)\bar\psi_\beta(y)]|0\rangle. \tag{26.6}$$

The renormalization of the field is $Z_R\psi_R = \psi_0$, and of the current is $Z_j j_R = j_0$, but as we can see there is no j, so no Z_j on the rhs, whereas there is on the lhs. By matching Z factors, we see that we must have $Z_j = 1$.

Anomalous dimension matrix

So we see that the renormalization of operators is written as

$$\mathcal{O}_j(x) = \sum_k Z_j{}^k \mathcal{O}_k^{\rm ren}(x). \tag{26.7}$$

We can therefore define for $\mathcal{O}$ also a notion of anomalous dimension, just that now it is an *anomalous dimension matrix*,

$$\Gamma_{ij} \equiv \left(Z^{-1}\cdot \Lambda\frac{\partial}{\partial\Lambda}Z\right)_{ij}\;, \tag{26.8}$$

where Z and Z^{-1} are matrices, and $\Lambda\partial/\partial\Lambda = \partial/\partial\ln\Lambda$.

Consider the *eigenvectors* $\mathcal{O}_n$ of the matrix Γ, with eigenvalue $\Delta_0+\gamma_n$, where Δ_0 is the classical (naive) dimension, and γ_n is the anomalous dimension of the operator $\mathcal{O}_n$.

Then we must have for scalar operators (no Lorentz structure)

$$\langle \mathcal{O}_n^{\rm ren}(x)\mathcal{O}_n^{\rm ren}(y)\rangle = \langle (Z\cdot\mathcal{O}_n(x)(Z\cdot\mathcal{O})_n(y)\rangle = \frac{\text{const.}}{|x-y|^{2(\Delta_0+\gamma_n)}}. \tag{26.9}$$

This is so because by translational invariance the 2-point function must depend on $|x-y|$ only, and then the dimension defines its power law. Strictly speaking, this result is only valid in a scale invariant theory, without any mass scale, since otherwise we can use the mass scale to construct dimensionless quantities together with $|x-y|$.

A nontrivial example of a composite operator in QCD is the quark mass term

$$\Delta\mathcal{L}_m = m(\bar q q)_M\;, \tag{26.10}$$

with renormalization prescription at the mass scale M.

Then for instance the Green's functions of $\bar{q}q$ with the quarks,

$$G^{(n,k)}(x_1, ..., x_n; y_1, ..., y_n; z_1, ..., z_k) = \langle q(x_1)...q(x_n)\bar{q}(y_1)...\bar{q}(y_n)\bar{q}q(z_1)...\bar{q}q(z_k)\rangle \ , \qquad (26.11)$$

obeys a renormalization group equation (RGE) that is the natural extension of the one for the Green's functions for fundamental fields, namely

$$\left[M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial g} + 2n\gamma + k\gamma_{\bar{q}q}\right] G^{(n,k)}(\{x_i\}, \{y_i\}, \{z_j\}, g, M) = 0. \qquad (26.12)$$

We can then define a mass depending on the energy scale, $\bar{m}(Q)$, in the usual way by

$$\frac{d}{d\log(Q/M)}\bar{m} = \gamma_{\bar{q}q}(\bar{g})\bar{m} \ , \qquad (26.13)$$

and the boundary (initial) conditions $\bar{m}(M) = m$. A perturbative calculation in QCD (that will not be reproduced here) finds

$$\bar{m}(Q) = m\left(1 + 8\frac{g^2}{(4\pi)^2}\log\frac{M^2}{Q^2}\right) + \mathcal{O}(g^4). \qquad (26.14)$$

Anomalous dimension calculation

We can give a simple example of an anomalous dimension calculation. Consider $\lambda\phi^4/4!$ theory, and the operator $\mathcal{O}(x) = \phi^2(x)$. The 2-point function $\langle\mathcal{O}(x)\mathcal{O}(0)\rangle$ is calculated as follows.

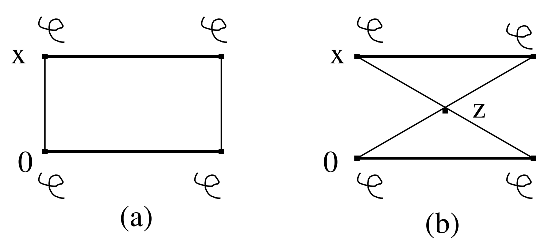

Figure 70: Diagrams for the anomalous dimension of an operator in ϕ^4 theory. (a) Tree level. (b) One loop level.

Tree level: $\mathcal{O}(1)$

At the tree level, we write a Feynman diagram where we represent the operator at 0 as a line with 2 points at its ends (for the two ϕ fields), and the same for the operators at x, for two parallel lines, as in Fig.70a. Then the free propagators connect the ϕ in the upper operator with the one in the lower operator. Since in 4d the scalar propagator is

$$\frac{1}{4\pi^2}\frac{1}{|x|^2} \ , \qquad (26.15)$$

we obtain

$$\langle\mathcal{O}(x)\mathcal{O}(0)\rangle^{\rm tree}=\frac{2}{(4\pi^2)^2}\frac{1}{|x|^4}\;, \tag{26.16}$$

for the 2 free propagators and two possible contractions.

One-loop level: $\mathcal{O}(\lambda)$

The Feynman diagram has a vertex at z connecting with the two fields in the upper operator and the two fields in the lower operator, as in Fig.70b, for a result

$$\frac{\lambda}{(4\pi^2)^4}\int d^4z\frac{1}{|x-z|^4|z|^4}\equiv\lambda\langle\mathcal{O}(x)\mathcal{O}(0)\rangle^{\rm tree}\mathcal{J}(x)\;, \tag{26.17}$$

because of the two propagators from z to x and two from z to 0. From the above definition, we have

$$\mathcal{J}(x)=\frac{|x|^4}{(4\pi^2)^2}\int d^4y\frac{1}{y^4|x-y|^4}. \tag{26.18}$$

This is UV divergent, and needs to be regularized by introducing an UV cut-off Λ, after which the integral becomes

$$\mathcal{J}(x)\sim\frac{1}{4\pi^2}\log(|x|\Lambda). \tag{26.19}$$

The proof of this statement is left as an exercise.

All in all, we can write

$$\langle\mathcal{O}(x)\mathcal{O}(0)\rangle=\langle\mathcal{O}(x)\mathcal{O}(0)\rangle^{\rm tree}\left[1+\frac{\lambda}{4\pi^2}\log(|x|\Lambda)+...\right]. \tag{26.20}$$

On the other hand, in general we can expand (since $\Delta=\Delta_0+\mathcal{O}(\lambda)$, so $\Delta-\Delta_0=\mathcal{O}(\lambda)$)

$$\begin{aligned}\langle\mathcal{O}(x)\mathcal{O}(0)\rangle &= \frac{C}{|x|^{2\Delta}}=\frac{C}{|x|^{2\Delta_0}}e^{-2(\Delta-\Delta_0)\log(|x|\Lambda)}\simeq\frac{C}{|x|^{2\Delta_0}}\left(1-2(\Delta-\Delta_0)\log(|x|\Lambda)\right)\\ &= \langle\mathcal{O}(x)\mathcal{O}(0)\rangle^{\rm tree}\left(1-2(\Delta-\Delta_0)\log(|x|\Lambda)+...\right)\end{aligned} \tag{26.21}$$

Here we have normalized the operators such that the 2-point function has $C=2/(4\pi^2)^2$ in order to reproduce the free 2-point function above.

By comparing this general result with our particular Feynman diagram calculation, we obtain first $\Delta_0=2$, and second

$$\Delta-\Delta_0=-\frac{\lambda}{8\pi^2}\Rightarrow\Delta=2-\frac{\lambda}{8\pi^2}. \tag{26.22}$$

In the same $\lambda\phi^4/4!$ theory, we can give an example of operator mixing, between the operators $\phi^2(x)$ and $\phi^4(x)$. The Feynman diagram has one free propagator coming down from $\phi^2(x)$ to $\phi^4(0)$, and one line ends in a 4-vertex, whose 3 other legs end up on the remaining fields in $\phi^4(0)$, as in Fig.71. The result for the diagram is (combinatorial factor $2\cdot 4!$)

$$\langle\phi^2(x)\phi^4(0)\rangle=\frac{2\cdot 4!\lambda}{(4\pi^2)^2|x|^4}\tilde{I}(x)\;, \tag{26.23}$$

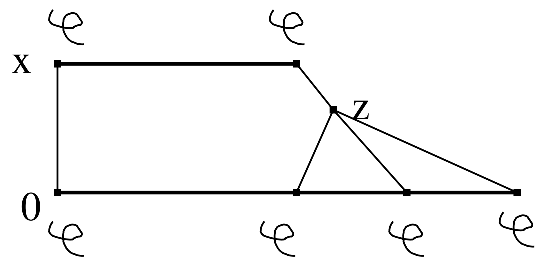

Figure 71: Operator mixing diagram in ϕ^4 theory.

where

$$\tilde{I}(x) = \frac{|x|^4}{(4\pi^2)^3}\int d^4z \frac{1}{|z|^6|x-z|^2}\ , \tag{26.24}$$

and a calculation that will not be reproduced here gives for the integral $-2\pi^2|\Lambda|^2/|x|^2$, so we obtain

$$\tilde{I}(x) = -\frac{|x|^2\Lambda^2}{32\pi^3}. \tag{26.25}$$

The operator product expansion (OPE)

In general, when 2 composite operators $\mathcal{O}_1(x)$ and $\mathcal{O}_2(0)$ go to a point, we can create disturbances (perturbations) in the vicinity of the operators, with divergent coefficients. So the result is described in terms of local operators, times divergent coefficient functions. Under this procedure, there is a complete (closed) set of operators $\{\mathcal{O}_k\}$, consistent with the symmetries. The singularity at $x \to 0$ translates into a singularity in the coefficient functions. Therefore we have

$$\mathcal{O}_1(x)\mathcal{O}_2(0) \to \sum_n C_{12}{}^n(x)\mathcal{O}_n(0). \tag{26.26}$$

By translational invariance, the rhs can only depend on the difference in the positions of the two operators, so being a bit more general we can write

$$\mathcal{O}_i(x)\mathcal{O}_j(y) \to \sum_k C_{ij}{}^k(x-y)\mathcal{O}_k(y). \tag{26.27}$$

This is the *operator product expansion* (OPE). The coefficient functions $C_{ij}{}^k(x-y)$ are c-number functions that are singular in the argument going to zero.

Not that the OPE is an operator relation, which means that it must hold on any matrix element $\langle\alpha|\quad|\beta\rangle$. Dimensional analysis suggests that in a theory with no mass scales (scale invariant),

$$C_{ij}{}^k(x-y) \to \frac{1}{|x-y|^{\Delta_i+\Delta_j-\Delta_k}}. \tag{26.28}$$

That means that, as an approximation, we can only consider the operator $\mathcal{O}_k$ of lowest dimension Δ_k, which will have the coefficient $C_{ij}{}^k$ of the highest singularity in $|x-y|$, when considering this as an expansion in $|x-y|$. That makes the OPE very useful for calculations.

The OPE is valid in all Green's functions, so for instance

$$G_{ij}(x;y;z_1,...,z_m) = \langle \mathcal{O}_i(x)\mathcal{O}_j(y)\phi(z_1)...\phi(z_m)\rangle = \sum_k C_{ij}{}^k(x-y)\langle \mathcal{O}_k(y)\phi(z_1)...\phi(z_m)\rangle. \tag{26.29}$$

That means that we can reduce the Green's functions to lower ones, and the dependence on the absorbed point, x, is now in the coefficient functions only.

In turn, that means that knowing all the OPEs solves the theory in terms of $\mathcal{O}_j(x)$'s, since we can reduce succesively the number of operators in the Green's function,

$$\begin{aligned}\langle \mathcal{O}_i(x)\mathcal{O}_j(y)\mathcal{O}_l(z)...\mathcal{O}_p(w)\rangle &= \sum_k C_{ij}{}^k(x-y)\langle \mathcal{O}_k(y)\mathcal{O}_l(z)...\mathcal{O}_p(w)\rangle\\ &= \sum_{km} C_{ij}{}^k(x-y)C_{kl}{}^m(y-z)\langle \mathcal{O}_m(z)...\mathcal{O}_p(w)\rangle = ...\\ &= \sum_{k,m,...} C_{ij}{}^k(x-y)C_{kl}{}^m(y-z)...C_{qp}{}^r(t-w)\langle \mathcal{O}_r(w)\rangle\end{aligned} \tag{26.30}$$

Then, if we have the 2-point function normalized as

$$\langle \mathcal{O}_q(x)\mathcal{O}_p(y)\rangle = \frac{C\delta_{qp}}{|x-y|^{2\Delta_q}}\ , \tag{26.31}$$

so that in particular

$$\langle \mathcal{O}_r(w)\rangle = \delta_{r1} \Rightarrow C_{qp}{}^r(t-w)\langle \mathcal{O}_r(w)\rangle = C_{qp}{}^r(t-w)\langle 1\rangle = \frac{C\delta_{qp}}{|t-w|^{2\Delta_q}}\ , \tag{26.32}$$

so we solve completely the Green's function.

QCD example

The example of OPE we are intersted in is in QCD. In particular, the currents $J^\mu(x)J^\nu(y)$ must have an OPE that has a maximal contribution from the quark current $\bar{q}\gamma^\mu q$. On the other hand, one can prove that we must expand them in the operators in irreducible representations (irreps) of the Lorentz group. That leads to the basis of gauge invariant operators for the OPE (note that we use D^μ instead of ∂^μ for gauge invariance)

$$\mathcal{O}_f^{(n)\mu_1...\mu_n} = \bar{q}_f\gamma^{\{\mu_1}D^{\mu_2}...D^{\mu_n\}}q_f - \text{traces}. \tag{26.33}$$

Here f stands as usual for flavor (the type of quark). These operators start with $\bar{q}\gamma^\mu q$ for $n=1$, have dimension $n+2$ (since there are $n-1$ D^μ's, each with dimension 1, and two quark fields, each with dimension 3/2), and spin (i.e. Lorentz irrep) n, since there are n vector indices symmetrized and with the traces subtracted.

Then in momentum space, the OPE of the currents is written as

$$i\int d^4x e^{iq\cdot x}J^\mu(x)J^\nu(0) \quad = \quad \sum_f Q_f^2\left[\sum_{n\geq 2}\frac{(2q^{\mu_1})...(2q^{\mu_{n-2}})}{(Q^2)^{n-1}}\mathcal{O}_f^{(n)\mu\nu\mu_1...\mu_{n-2}}\right.$$

$$-g^{\mu\nu}\sum_{n\geq 2}\frac{(2q^{\mu_1})...(2q^{\mu_n})}{(Q^2)^n}\mathcal{O}_f^{(n)\mu_1...\mu_n}\Bigg] + ... \qquad (26.34)$$

In x space, it becomes

$$T[J^\mu(x)J^\nu(0)] \sim \sum_{n\geq 2} C_n \frac{x^{\mu_1}...x^{\mu_{n-2}}}{|x|^2}\mathcal{O}_f^{(n)\mu\nu\mu_1...\mu_{n-2}}(0) + g^{\mu\nu}\sum_{n\geq 2}\tilde{C}_n\frac{x^{\mu_1}...x^{\mu_n}}{(|x|^2)^2}\mathcal{O}_f^{(n)\mu_1...\mu_n}(0) + ... , \qquad (26.35)$$

where C_n and $\tilde{C}_n$ are fixed numerical coefficients.

For $\mu = \nu$ and summed over, we get a simpler expression,

$$T[J^\mu(x)J_\mu(0)] = \sum_n h_n \frac{x^{\mu_1}...x^{\mu_n}}{(|x|^2)^2}\mathcal{O}_f^{(n)\mu_1...\mu_n}(0) , \qquad (26.36)$$

where $h_n = h_n(x, \mu^2, \alpha_s(\mu^2))$ is a dimensionless function.

Note then that we have 2 types of OPE expansions.

-the standard one, a short distance expansion, in $|x|$, in which we can keep only the operator $\mathcal{O}^{(n)}$ of lowest dimension.

-but we can also have a *lightcone expansion*, obtained by considering that the expansion is in $|x|^2$ around $|x|^2 = 0$ viewed as a lightcone (with x^μ finite), in which case the relevant dimension for the expansion is not the dimension of $\mathcal{O}_f^{(n)\mu_1...\mu_n}$, but rather the dimension of the full object in the numerator, i.e. the dimension of $x^{\mu_1}...x^{\mu_n}\mathcal{O}^{(n)\mu_1...\mu_n}$, equal to

$$\dim[\mathcal{O}^{(n)}] - n = \dim[\mathcal{O}^{(n)}] - \text{spin}[\mathcal{O}] \equiv \text{twist}[\mathcal{O}^{(n)}]. \qquad (26.37)$$

Here the *twist* $T = \Delta - S$. Then, in this expansion, we keep only the *leading twist operators*.

But now that there is operator mixing, $\mathcal{O}_f^{(n)\mu_1...\mu_n}$ mixing with

$$\mathcal{O}_g^{(n)\mu_1...\mu_n} = F^{\{\mu_1\nu}D^{\mu_2}...D^{\mu_{n-1}}F^{\mu_n\}}{}_\nu - \text{traces}. \qquad (26.38)$$

Indeed, these operators have also dimension $n+2$, since there are $n-2$ D's, and 2 F's, each with dimension 2. They also have spin n, since there are n vector indices symmetrized and with the traces subtracted.

That means that in the OPEs, the $\mathcal{O}_g$'s also appear, same as the $\mathcal{O}_f$'s. The mixing of the two operators is done through diagrams where the two quarks in the $\mathcal{O}_f$ emit quark and antiquark propagators, that join, by emitting two gluons from the quark line, that end on the two gluons in the $\mathcal{O}_g$ operators, as in Fig.72.

Important concepts to remember

- Composite operators needs additional renormalization, besides the one in the Lagrangean, because of extra divergences when fields come at the same point.

- Since there is also operator mixing between operators of like charges, we have in general a renormalization matrix, $\mathcal{O}_n[\{\phi_i\}, g...] = \sum_m Z_n{}^m \mathcal{O}^{\text{ren}}[\{Z_j\phi_j^{\text{ren}}\}, Z_g g, ...]$.

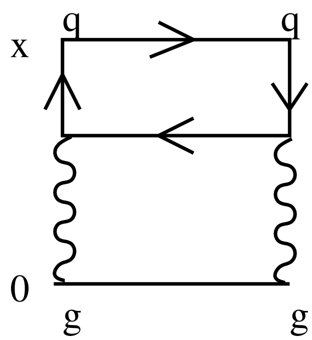

Figure 72: Operator mixing diagram in QCD.

- We also have an anomalous dimension matrix, $\Gamma_{ij} = (Z \cdot d/\partial \log \Lambda Z)_{ij}$.
- Conserved currents do not renormalize.
- When two composite operators go to the same point, we have an expansion in terms of other operators, with divergent c-number coefficient functions, $\mathcal{O}_i(x)\mathcal{O}_j(y) \to \sum_k C_{ij}{}^k(x-y)\mathcal{O}_k(y)$, called the operator product expansion, or OPE.
- Knowing the full OPE is equivalent to solving the theory in terms of the $\mathcal{O}_i$'s, since it allos us to calculate the Green's functions.
- We neeed to consider operators with both dimension and spin, and then we can have either a short distance expansion (usual OPE), staring with operators of minimal dimension, or a lightcone expansion, for $x^2 \to 0$, but with x^μ finite, starting with operators of leading twist $T = \Delta - S$.

Further reading: See chapter 18.1,18.3 and 18.5 in [3], chapter 11.2 and 14.5 in [3] and chapter 20 in Weinberg vol.II.

Exercises, Lecture 26

1) Prove that

$$\mathcal{J}(x) = \frac{|x|^4}{(4\pi^2)^2} \int \frac{d^4y}{y^4|x-y|^4} \sim \frac{1}{4\pi^2} \log(|x|\Lambda) + \text{finite}. \tag{26.39}$$

2) Consider a set of operators $\mathcal{O}_i$, complete under the OPE, and consider their OPEs. Calculate the 4-point functions of the theory.

3) Consider $\mathcal{N} = 4$ SYM, with field content: gauge fields A^a_μ, spinors $\psi^a_{I\alpha}$, scalars X^a_{IJ}, where a is an index in the adjoint of $SU(N)$, $I, J = 1, ..., 4$ are fundamental indices of a global $SU(4)$. What are the set of leading twist, gauge invariant, composite ($n \geq 2$ fields) operators?

27 Lecture 27. The Wilsonian effective action, effective field theory and applications.

In this lecture, we will describe a new view on the process of renormalization, one that will be continued in the next lecture. Through the Wilsonian effective action, we will define effective field theory, and see how to apply it.

In the real world, we can always test physics only below a maximal energy scale Λ. The question is then, can we hide our ignorance about the high energy in a consistent framework? The answer turns out to be yes, by defining the Wilsonian effective action, and through it, effective field theory.

The Wilsonian effective action.

The Wilsonian effective action is obtained in the simplest way, namely by integrating out the degrees of freedom with momenta $|k|>\Lambda$, in order to hide our ignorance about it.

ϕ^4 theory in Euclidean space

We will describe the formalism on the simplest nontrivial theory, ϕ^4 in Euclidean 4 dimensions. The (classical) action is

$$S_E=\int d^4x\left[\frac{1}{2}(\partial_\mu\phi)^2+\frac{m^2}{2}\phi^2+\frac{\lambda}{4!}\phi^4\right]\;, \tag{27.1}$$

and the partition function is

$$Z[J]=\int\mathcal{D}\phi e^{-S+\int d^4xJ(x)\phi(x)}. \tag{27.2}$$

But since we want to integrate over momenta $|k|>\Lambda$, we must go to momentum space, where

$$S_E=\frac{1}{2}\int\frac{d^4k}{(2\pi)^4}\tilde{\phi}(-k)(k^2+m^2)\tilde{\phi}(k)+\frac{\lambda}{4!}\int\frac{d^4k_1}{(2\pi)^4}...\frac{d^4k_4}{(2\pi)^4}\tilde{\phi}(k_1)...\tilde{\phi}(k_4)(2\pi)^4\delta^4(k_1+...+k_4). \tag{27.3}$$

We introduce an UV cut-off Λ. Since we want to say that we cannot access energies higher than Λ by experiments, we impose that sources are zero for these momenta, $J=0$ for $|k|>\Lambda$. Then

$$Z[J]=\int\mathcal{D}\phi_{|k|<\Lambda}e^{-S_{eff}(\phi;\Lambda)+\int J\cdot\phi}\;, \tag{27.4}$$

where

$$e^{-S_{eff}(\phi;\Lambda)}=\int\mathcal{D}\phi_{|k|>\Lambda}e^{-S_E[\phi]}. \tag{27.5}$$

Here S_{eff} is called the *Wilsonian effective action.* The Wilsonian effective Lagrangean will be then (after doing the integral)

$$\mathcal{L}_{eff}=\frac{1}{2}(\partial_\mu\phi)^2+\frac{m^2}{2}\phi^2+\frac{\lambda}{4!}\phi^4+\sum_{\Delta\geq 6}\sum_i c_{\Delta,i}\mathcal{O}_{\Delta,i}. \tag{27.6}$$

Here $\mathcal{O}_{\Delta,i}$ are all higher dimension operators (dimension higher than 4, but since the classical Lagrangean has only even powers of ϕ, so should the quantum action, so the next dimension is 6), organized by their dimension Δ and for given dimension by an index i.

However, in the above action, ϕ has now only momenta smaller than Λ, i.e.

$$\phi(x) = \int^{\Lambda} \frac{d^4k}{(2\pi)^4} e^{ik\cdot x}\tilde{\phi}(k). \tag{27.7}$$

But in reality, we have the renormalized Lagrangean,

$$\mathcal{L}_{\rm ren} = \frac{1}{2}Z_\phi(\partial_\mu\phi)^2 + \frac{Z_m m_{ph}^2}{2}\phi^2 + \frac{Z_\lambda\lambda_{ph}}{4!}\phi^4. \tag{27.8}$$

Here λ_{ph} is defined as the 1PI vertex a $p^\mu = 0$.

Then we have for the Wilsonian effective Lagrangean really

$$\mathcal{L}_{eff} = \frac{1}{2}Z(\Lambda)(\partial_\mu\phi)^2 + \frac{m^2(\Lambda)}{2}\phi^2 + \frac{\lambda(\Lambda)}{4!}\phi^4 + \sum_{\Delta\geq 6}\sum_i c_{\Delta,i}\mathcal{O}_{\Delta,i}. \tag{27.9}$$

Here all coefficients are *finite* functions of Λ, since the correlation functions of renormalized fields, calculated by $\delta/\delta J(x)$ from it, are finite.

Calculation of $c_{\Delta,i}$.

We now proceed to calculate explicitly $c_{\Delta,i}$ for the operators $\mathcal{O}_{2n,1} \equiv \phi^{2n}$ at one-loop.

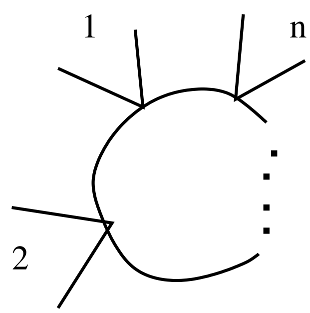

Figure 73: One loop diagram for the effective potential in the Wilson approach.

The unique one-loop diagram with $2n$ external lines is composed of a loop with n vertices on it, out of each having 2 external legs. Then in the external lines we have momenta $|k| < \Lambda$, but on the internal lines we integrate over $|k| > \Lambda$.

The symmetry factor for the diagram is

$$S = 2^n 2 \times n\ , \tag{27.10}$$

since there is a symmetry factor of 2 for the interchange of the 2 lines at each vertex, giving a total factor of 2^n, plus a rotation symmetry which means that we need to define where the vertex 1 is, giving a factor of n, plus a reflection symmetry giving another factor of 2.

There is also a $(2n)!$ factor for the ways to asign the momenta $p_1, ..., p_n$ to the external lines (permuting them).

The Feynman diagram is then identified with the effective vertex coming out as a Feynman rule out of $c_{2n,1}\mathcal{O}_{2n,1}$ in the Lagrangean, i.e. $-c_{2n,1}(2n)!$, so we have

$$-c_{2n,1}(2n)! = \frac{(2n)!}{2^n 2n}(-\lambda_{ph})^n \int_\Lambda^\infty \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m_{ph}^2)^n} + \mathcal{O}(\lambda_{ph}^{n+1}). \qquad (27.11)$$

Then for $2n \geq 6$, the integral is convergent, and is

$$\int_\Lambda^\infty \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m^2)^n} \simeq \frac{2\pi^2}{(2\pi)^4}\int_\Lambda^\infty \frac{k^3dk}{k^{2n}} = \frac{1}{8\pi^2}\frac{1}{(2n-4)\Lambda^{2n-4}}\ , \qquad (27.12)$$

where in the equality we have used the fact that $m^2 \ll \Lambda^2$, so we get for the coefficients at one-loop

$$c_{2n,1}(\Lambda) = -\frac{1}{32\pi^2}\frac{(-\lambda_{ph}/2)^n}{n(n-2)\Lambda^{2n-4}} + \mathcal{O}(\lambda_{ph}^{n+1}). \qquad (27.13)$$

2n=4

For $2n = 4$, i.e. for $\lambda(\Lambda)$, we formally have the same integral, just that we need to add the tree result to the one-loop one, and also now we wrote $\lambda(\Lambda)$ instead of $c_{4,1}4!$. Since now $(2n)!/(2^n 2n) = 4!/(2^2 4) = 3/2$, we get

$$-\lambda(\Lambda) = -Z_\lambda\lambda_{ph} + \frac{3}{2}(-\lambda_{ph})^2\int_\Lambda^\infty \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m_{ph}^2)^2} + \mathcal{O}(\lambda_{ph}^3). \qquad (27.14)$$

This result is UV divergent. However, we note that in the usual renormalization, the renormalized vertex at zero momenta, which by definition is λ_{ph}, is given by

$$-\lambda_{ph} = -V_4(0,0,0,0) = -Z_\lambda\lambda_{ph} + \frac{3}{2}(-\lambda_{ph})^2\int_0^\infty \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m_{ph}^2)^2} + \mathcal{O}(\lambda_{ph}^3). \qquad (27.15)$$

Note that the integration here is over *all* momenta, from 0 to ∞.

Then the difference of the two is finite,

$$-\lambda_{ph} + \lambda(\Lambda) = \frac{3}{2}(-\lambda_{ph})^2\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m_{ph}^2)^2} + \mathcal{O}(\lambda_{ph}^3). \qquad (27.16)$$

The integral is

$$\begin{aligned}\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m_{ph}^2)^2} &= \frac{1}{16\pi^2}\int_0^{\Lambda^2}\frac{k^2dk^2}{(k^2+m_{ph}^2)^2} = \frac{1}{16\pi^2}\int_{m_{ph}^2}^{\Lambda^2+m_{ph}^2}\frac{(\tilde{k}^2-m_{ph}^2)d\tilde{k}^2}{\tilde{k}^4}\\ &= \frac{1}{16\pi^2}\left(\ln\frac{\Lambda^2+m_{ph}^2}{m_{ph}^2} + \frac{m_{ph}^2}{\Lambda^2} - 1\right) \simeq \frac{1}{8\pi^2}\left(\ln\frac{\Lambda}{m_{ph}} - \frac{1}{2}\right);\end{aligned} \qquad (27.17)$$

thus it gives for the coupling

$$\lambda(\Lambda) = \lambda_{ph} + \frac{3}{16\pi^2}\lambda_{ph}^2\left[\ln\frac{\Lambda}{m_{ph}} - \frac{1}{2}\right] + \mathcal{O}(\lambda_{ph}^3). \qquad (27.18)$$

2n=2

For $2n = 2$, the Feynman diagram is a line with a loop on it, so the integral is independent of the external momentum p. Since the p-dependent part of the Feynman diagram gives the wave function renormalization, in our case we have no wave function renormalization at one-loop, i.e.

$$Z(\Lambda) = 1 + \mathcal{O}(\lambda_{ph}^2). \tag{27.19}$$

But there is a p-independent part, which gives the mass renormalization, so we obtain (again adding the tree contribution, and since now $(2n)!/(2^n 2n) = 2!/(2\cdot 2) = 1/2$)

$$-m^2(\Lambda) = -Z_m m_{ph}^2 + \frac{1}{2}(-\lambda_{ph})\int_\Lambda^\infty \frac{d^4k}{(2\pi)^4}\frac{1}{k^2+m_{ph}^2} + \mathcal{O}(\lambda_{ph}^2). \tag{27.20}$$

This is quadratically divergent. As before, the full renormalization gives

$$-m_{ph}^2 = -Z_m^2 m_{ph}^2 + \frac{1}{2}(-\lambda_{ph})\int_0^\infty \frac{d^4k}{(2\pi)^4}\frac{1}{k^2+m_{ph}^2} + \mathcal{O}(\lambda_{ph}^2)\ , \tag{27.21}$$

but the difference is now finite,

$$-m_{ph}^2 + m^2(\Lambda) = \frac{1}{2}(-\lambda_{ph})\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{k^2+m_{ph}^2} + \mathcal{O}(\lambda_{ph}^2). \tag{27.22}$$

The integral gives

$$\int_0^\Lambda \frac{d^4k}{(2\pi)^4}\frac{1}{k^2+m_{ph}^2} = \frac{1}{16\pi^2}\int_{m_{ph}^2}^{\Lambda^2+m_{ph}^2}\frac{(\tilde{k}^2-m_{ph}^2)d\tilde{k}^2}{\tilde{k}^2} = \frac{1}{16\pi^2}\left(\Lambda^2 - m_{ph}^2\ln\frac{\Lambda^2+m_{ph}^2}{m^2}\right)\ , \tag{27.23}$$

so the mass squared is

$$m^2(\Lambda) = m_{ph}^2 - \frac{\lambda_{ph}}{32\pi^2}\left[\Lambda^2 - m_{ph}^2\ln\frac{\Lambda^2}{m_{ph}^2}\right] + \mathcal{O}(\lambda_{ph}^2). \tag{27.24}$$

As a final observation, note that the nonrenormalizable operators $\mathcal{O}_{\Delta,i}$ have coefficients of the order

$$c_{\Delta,i} \sim \frac{1}{\Lambda^{\Delta-4}}\ , \tag{27.25}$$

which is highly suppressed for large Λ (energies much smaller than it), so it will not change too much the physics.

Effective field theory.

We consider now a new set-up, closer to the physical case. In the physical case, the theory can have a true cut-off, for instance at least the Planck scaler $m_{\rm Planck}$ can act as such, if not the susy scale, GUT scale, KK scale, etc. On top of it, we can consider also our arbitrary (variable) scale Λ below it. At this cut-off scale, the degrees of freedom of the theory change, and we have a new theory.

We also *assume* that when measured in units of the cut-off, the parameters of the theory at the cut-off scale are small,

$$\lambda(\Lambda_0) \ll 1, \quad m^2(\Lambda_0) \ll \Lambda_0^2, \quad c_{\Delta,i}(\Lambda_0) \ll \Lambda_0^{4-\Delta}. \tag{27.26}$$

So we treat the effective action as a fundamental starting point, rather than assuming a better definition of the theory in the UV.

Since the coefficients of the higher dimension operators are small, let us assume for the moment that they actually vanish, i.e.

$$c_{\Delta,i}(\Lambda_0) = 0 \ , \tag{27.27}$$

and see if that is consistent.

We now integrate over the region between the true cut-off and an arbitrary lower cut-off Λ, i.e. over $\Lambda < |k| < \Lambda_0$.

Then we have at the lower scale

$$e^{-S_{eff}[\phi;\Lambda]} = \int \mathcal{D}\phi_{\Lambda<|k|<\Lambda_0} e^{-S_{eff}(\phi;\Lambda_0)}. \tag{27.28}$$

Using the formulas already derived for $m^2(\Lambda), \lambda(\Lambda), c_{2n,1}(\Lambda)$, we can calculate the values at Λ in terms of values at Λ_0, and obtain, for Λ not too much less that Λ_0, i.e. for $\Lambda = b\Lambda_0$, $b \lesssim 1$,

$$\begin{aligned}
m^2(\Lambda) &= m^2(\Lambda_0) + \frac{1}{2}\lambda(\Lambda_0)\int_\Lambda^{\Lambda_0}\frac{d^4k}{(2\pi)^4}\frac{1}{k^2+m^2(\Lambda_0)} + ...\\
\lambda(\Lambda) &= \lambda(\Lambda_0) - \frac{3}{2}\lambda^2(\Lambda_0)\int_\Lambda^{\Lambda_0}\frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m^2(\Lambda_0))^2} + ...\\
c_{2n,1}(\Lambda) &= -\frac{(-1)^n}{2^n 2n}\lambda^n(\Lambda_0)\int_\Lambda^{\Lambda_0}\frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m^2(\Lambda_0))^n} + ...
\end{aligned} \tag{27.29}$$

Then, if Λ is not too much less than Λ_0 as before, and if also $m^2(\Lambda_0) \ll \Lambda^2$, we obtain approximately (using the integrals calculated before)

$$\begin{aligned}
m^2(\Lambda) &= m^2(\Lambda_0) + \frac{\lambda(\Lambda_0)}{32\pi^2}(\Lambda_0^2 - \Lambda^2) + ...\\
\lambda(\Lambda) &= \lambda(\Lambda_0) - \frac{3}{16\pi^2}\lambda^2(\Lambda_0)\ln\frac{\Lambda_0}{\Lambda} + ...\\
c_{2n,1}(\Lambda) &= -\frac{(-\lambda(\Lambda_0)/2)^n}{32\pi^2 n(n-2)}\left(\frac{1}{\Lambda^{2n-4}} - \frac{1}{\Lambda_0^{2n-4}}\right) + ...
\end{aligned} \tag{27.30}$$

We note therefore that Λ_0 is not very important for $c_{2n,1}$, its effect being very small, which is why it was consistent to consider $c_{2n,1}(\Lambda_0) = 0$, whereas it is very important for m^2, so arranging for $m^2(\Lambda) \ll \Lambda^2$ is a *fine-tuning problem* that is not natural.

We also note that we can now define the beta function as usual and calculate it at one-loop from the above $\lambda(\Lambda)$ as

$$\beta(\lambda) = \frac{d\lambda}{d\ln\Lambda} = \frac{3}{16\pi^2}\lambda^2(\Lambda) \ , \tag{27.31}$$

as it should be.

In conclusion, the picture of *effective field theory* is as follows. We define the theory with a cut-off, and correspondingly with higher dimension operators. Then, we consider lowering the cut-off by integrating out the intermediate degrees of freedom. When we do that, the coefficients in the (Wilsonian) effective action change. This leads to a view of renormalization as a change in cut-off scale that will be developped better in the next lecture, as Kadanoff blocking, and we will see that it leads to a connection with condensed matter theory.

Nonrenormalizable theories

This picture also leads to a way to deal with nonrenormalizable theories. We can just regard them as (Wilsonian) effective actions. We impose a cut-off Λ_0, in which case the coefficients of the higher dimension operators are

$$\frac{c_i}{\Lambda_0^{\Delta_i-4}}\,, \tag{27.32}$$

with $c_i \leq 1$. We can then use this effective action below the energy scale Λ_0 as shown above, and for energies $E \ll \Lambda_0$ the coefficients of the higher dimension operators are really small, making the theory look almost like a renormalizable one, up to powers of E/Λ_0.

Removing the cut-off.

An important question that arises then is, can we remove completely the cut-off Λ_0?

If we know an exact beta function, we can integrate

$$\frac{d\lambda}{d\ln\Lambda} = \beta(\lambda) \tag{27.33}$$

to

$$\int_{\lambda(m_{ph})}^{\lambda(\Lambda_0)} \frac{d\lambda}{\beta(\lambda)} = \ln\frac{\Lambda_0}{m_{ph}}. \tag{27.34}$$

Then, as $\Lambda_0 \to \infty$, the right hand side goes to infinity, so the left hand side should too. In that case, it means we can remove Λ_0.

But it could happen that it does not go to infinity, but instead it goes to a constant before that. Indeed, in the case that $\beta(\lambda)$ is positive and increases at infinity faster than λ, we can see that $\int d\lambda/\beta(\lambda)$ is finite at infinity, so there must be a maximum Λ, Λ_{max}, given by

$$\ln\frac{\Lambda_{max}}{m_{ph}} = \int_{\lambda(m_{ph})}^{\infty} \frac{d\lambda}{\beta(\lambda)}. \tag{27.35}$$

For instance, using the one-loop beta function $\beta_1 = 3\lambda^2/(16\pi^2)$, which indeed goes faster than λ, we find

$$\Lambda_{max} = m_{ph} e^{\frac{16\pi^2}{3\lambda_{ph}}}. \tag{27.36}$$

This is the Landau pole that we alreaday explained several times.

On the other hand, if $\beta(\lambda) > 0$, but it goes at infinity slower than λ, we *can* remove Λ_0, and we have an UV fixed point λ_*.

Indeed, then for $\lambda \to \lambda_*$, $\Lambda \to \infty$, and

$$\int_{\lambda}^{\lambda_*} \frac{d\lambda}{\beta(\lambda)} \to \infty. \tag{27.37}$$

Important concepts to remember

- In the Wilsonian effective action approach, we hide our ignorance about high energy by integrating over momenta with $|k| > \Lambda$.
- In the Wilsonian effective action, we integrate over $|k| > \Lambda$ and obtain all the higher dimension operators, with coefficients that go like $1/\Lambda^{\Delta-4}$, and we use it in quantum processes with energies smaller than Λ.
- The effective field theory approach is to consider the theory as fundamentally defined with a cut-off and a Wilsonian effective action, and then lower the cut-off by integrating over intermediate degrees of freedom, $\Lambda < |k| < \Lambda_0$.
- In this way, nonrenormalizable theories are thought of as effective field theories, and at energies $E \ll \Lambda_0$, the effect of the nonrenormalizable operators is very small.
- For $\beta(\lambda) > 0$, if we can remove completely the UV cut-off, we have an UV fixed point, if not, we have a Landau pole.

Further reading: See chapter 29 in Srednicki and chapter 12.1 in [3].

Exercises, Lecture 27

1) Consider ϕ^3 theory in $d = 6$, with

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{\lambda}{3!}\phi^3 + \frac{m^2}{2}\phi^2. \tag{27.38}$$

Calculate the coefficients $c_{n,1}(\Lambda)$ of the $\mathcal{O}_{n,1} = \phi^n$ higher dimension operators in Wilson's approach, at one-loop.

2) In the same theory, in the effective field theory approach, calculate $\lambda(\Lambda)$ from $\lambda(\Lambda_0)$ at one-loop ($\Lambda < |k| < \Lambda_0$).

28 Lecture 28. Kadanoff blocking and the renormalization group; connection with condensed matter

In this lecture we will study a way to understand the renormalization group related to the Wilsonian way from last lecture, by integrating degrees of freedom, just that the way we will do it is via a discretization, that has a natural relation to condensed matter physics, and in particular to spin systems.

Field theories as classical spin systems.

Consider a discretization of the scalar Lagrangean

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial_\mu \phi)^2 + V(\phi)(+J \cdot \phi) \tag{28.1}$$

and its path integral, on a (hiper)cubic lattice of size a, via

$$\begin{aligned} & x \to x_n; \quad \phi(x) \to \phi_n = \phi(x_n); \quad \mathcal{D}\phi \to \prod_n d\phi_n \\ & \partial_\mu \phi \to \frac{1}{a}(\phi_{n+\mu} - \phi_n) \ , \end{aligned} \tag{28.2}$$

where $n+\mu$ is the nearest neighbour on the lattice (in direction μ) to site n.

Then the discretized action is

$$S \to \sum_n a^d \left[\frac{1}{a^2} \sum_\mu \frac{1}{2}(\phi_{n+\mu} - \phi_n)^2 + V(\phi_n) + J_n \phi_n \right]. \tag{28.3}$$

For the particular case of $V(\phi) = \lambda \phi^4/4!$, we can rescale as follows to absorb the dependence on a and put the coupling outside the action:

$$\begin{aligned} \lambda &= g^{-2} a^{d-4} \\ \phi' &= g a^{\frac{d}{2}-1} \phi \\ J' &= g a^{\frac{d}{2}+1} J. \end{aligned} \tag{28.4}$$

Then the path integral becomes

$$Z[J', g] = \mathcal{N} \int \prod_n (g^{-1} d\phi'_n) e^{-\frac{S(\phi')}{g^2}} \ , \tag{28.5}$$

where the action is

$$S(\phi', J') = \sum_n \left[\frac{1}{2} \sum_\mu (\phi'_{n+\mu} - \phi_n)^2 + \frac{m^2 a^2}{2} \phi'^2 + \frac{\phi'^4}{4!} + J'_n \phi'_n \right]. \tag{28.6}$$

This has now the form of a classical spin system. Indeed, for instance a ferromagnet has the Hamiltonian

$$H(s,h) = -\sum_{n,m} v_{n,m} S_n \cdot S_m + h \sum_m s_m \ , \tag{28.7}$$

where the first term is a spin-spin interaction and the second is the interaction of the spins with an external field h. The partition function is

$$Z[h,\beta] = \int \prod_n ds_n \; \rho(s_n) e^{-\beta H(s,h)} \;, \tag{28.8}$$

where $\beta = 1/(k_B T)$ and $\rho(s_n)$ is a weight describing the spin. It should depend on s_n^2 because of relativistic invariance, and it is naturally exponential, so an effective description for this measure is given phenomenologically by

$$\rho(s_n) \propto e^{-(\kappa s_n^2 + \lambda s_n^4)}. \tag{28.9}$$

The exact form would be given by the microscopic properties of spin.

For a system with only nearest-neighbour interaction, the $\sum_{n,m}$ in the Hamiltonian turns into $\sum_n \sum_\mu$, and

$$-\sum_{n,m} v_{n,m} s_n \cdot s_m \to K \sum_n (\sum_\mu (s_{n+\mu} - s_n)^2 - 2d s_n^2). \tag{28.10}$$

The partition function then becomes

$$Z(K,\mu,\lambda,h) = \int \prod_n ds_n e^{-\beta H(s;K,\mu,\lambda,h)} \;, \tag{28.11}$$

where the effective Hamiltonian is

$$H(s;K,\mu,\lambda,h) = \sum_n \left[K(\beta) \sum_\mu (s_{n+\mu} - s_n)^2 + \mu(\beta) s_n^2 + \lambda(\beta) s_n^4 + h s_n \right]. \tag{28.12}$$

This is written in a more general form, but from the above considerations we would have $K(\beta) = K$, $\mu(\beta) = K/\beta - 2dK$ and $\lambda(\beta) = \lambda/\beta$.

As we see, $\mu(\beta)$ is governed by two opposing terms, so can be either positive or negative. When it is zero $\mu = \mu(\beta_c) = 0$, we are at a phase transition.

Indeed, for $\mu(\beta) > 0$ and no external field h, we have no magnetization, since

$$\langle s \rangle = \frac{1}{V} \sum_n \langle s_n \rangle = 0 \;, \tag{28.13}$$

whereas at $\mu(\beta) < 0$ and no external field h, the classical minimum of the Hamiltonian is at $s_n = \sqrt{-\mu/(2\lambda)}$, so the magnetization is

$$\langle s \rangle = \frac{1}{V} \sum_n \langle s_n \rangle = \sqrt{\frac{-\mu}{2\lambda}}. \tag{28.14}$$

Moreover, near $\beta = \beta_c$, assuming smoothness of the mass term, we have

$$\mu(\beta) \simeq c_0(\beta - \beta_c) \;, \tag{28.15}$$

which then gives

$$\langle s\rangle \sim \sqrt{\beta-\beta_c} \tag{28.16}$$

for $\beta > \beta_c$ (nonanalytic behaviour).

In general, the magnetization $\langle s\rangle = \langle s\rangle(h,\beta)$, and is defined by

$$\langle s\rangle = \frac{1}{V}\sum_n \langle s_n\rangle = \frac{1}{V}\sum_n \int \prod_p ds_p e^{-\beta H} s_n. \tag{28.17}$$

Then the *magnetic susceptibility* χ is

$$\chi = \frac{\partial\langle s\rangle}{\partial h} = \frac{1}{V}\sum_n \int \prod_p ds_p e^{-\beta H}\beta s_n \sum_m s_m\ , \tag{28.18}$$

and it can be rewritten as

$$\chi = \frac{1}{V}\sum_{n,m}\langle (s_n-\langle s\rangle)(s_m-\langle s\rangle)\rangle\ , \tag{28.19}$$

so it is given in terms of a 2-point function. The 2-point function at large distances behaves as

$$\langle (s_n-\langle s\rangle)(s_m-\langle s\rangle)\rangle \sim e^{-\frac{|x_n-x_m|}{\xi(h,\beta)}} \tag{28.20}$$

for $\xi(h,\beta) \ll |x_n - x_m|$, where ξ is called the *correlation length*, whereas at small distances (but still much larger than the lattice size) it goes as

$$\langle (s_n-\langle s\rangle)(s_m-\langle s\rangle)\rangle \sim |x_n-x_m|^{-(d-2+\eta)}\ , \tag{28.21}$$

for $a \ll |x_n - x_m| \ll \xi(h,\beta)$. Here η is called anomalous dimension.

We know that the susceptibility blows up at a phase transition, and we see that this is due to the 2-point function diverging.

The *scaling hypothesis* for phase transitions is that all singular behaviour near the phase transition is due to the divergence of ξ. That is, $\xi \to \infty$ at the phase transition point implies all the divergences in physical (measurable) quantities. Since $\xi \to \infty$, and this is the only relevant scale for the phase transition, it means that there are no objects with dimension at the phase transition, and thus the theory is scale invariant (fixed under a scaling transformation), and thus all diverging quantities diverge as power laws, not as exponentials (which would require a scale).

In the quadratic approximation around the minimum of the spin Hamiltonian, we can calculate (though it is also clear by dimensional analysis, since μ is the only parameter with dimension, specifically of dimension 2) that

$$\xi(h,\beta) \sim \frac{1}{\sqrt{|\mu(h,\beta)|}}\ , \tag{28.22}$$

so indeed ξ diverges at the critical point.

One can define critical exponents, for instance for the susceptibility and the correlation length,

$$\begin{aligned} \chi(\beta) &\sim |\beta-\beta_c|^{-\gamma} \\ \xi(\beta) &\sim |\beta-\beta_c|^{-\nu}. \end{aligned} \tag{28.23}$$

These critical exponents can only take few values, i.e. there is a certain *universality* for them, since a large class is independent of microscopics.

Kadanoff blocking

From now on, since we have formulated the discretized field theory like a spin system, we will talk about both on the same footing, and consider the Hamiltonian

$$H=\sum_n\left[\sum_\mu \frac{1}{2}(\phi_{n+\mu}-\phi_n)^2+\mu\phi_n^2+\lambda\phi_n^4\right]. \tag{28.24}$$

We define Kadanoff blocking as follows. We divide the lattice into blocks of size s^d, where $s\in\mathbb{N}$, and average over the blocks

$$\phi'_{n'}=s^{-d}\sum_{n\in B_s(n')}\phi_n\;, \tag{28.25}$$

after which we rescale to the original size, by

$$x_s\equiv\frac{x}{s}. \tag{28.26}$$

If we measure ξ in lattice units, it has decreased by s,

$$\xi_s=\frac{\xi}{s}. \tag{28.27}$$

Since the 2-point correlation function decays as a power law for $1\ll n\ll\xi$,

$$\langle\phi_n\phi_0\rangle\sim\frac{1}{n^{d-2+\eta}}\;, \tag{28.28}$$

it means that we need to rescale $\phi_s\equiv s^\alpha\phi$, where

$$\alpha=\frac{d-2+\eta}{2}. \tag{28.29}$$

This is a sort of *wave function renormalization* in the quantum field theory sense.

The new Hamiltonian $H'(\phi'_n)$ is found by averaging over the blocks, i.e.

$$e^{-\beta H'[\phi'_{n'}]}=\int\prod_n d\phi_n e^{-\beta H[\phi]}\prod_{n'}\delta\left(\phi'_n-s^{-d}\sum_{n\in B_s(n')}\phi_n\right). \tag{28.30}$$

After blocking, we define the rescaled Hamiltonian by

$$H_s(\phi_s(x_s)) \equiv H'(\phi'(x')). \tag{28.31}$$

But this procedure of blocking will generate new terms in the Hamiltonian, for instance

$$(\phi_{n+\mu} - \phi_n)^2\phi_n^2, \quad (\phi_{n+\mu} - 2\phi_n + \phi_{n-\mu})^2 \tag{28.32}$$

and others. This is intuitively clear from the Wilsonian picture of effective field theory from last lecture, which amounted also to blocking, though in momentum space (integrating from a physical momentum cut-off to a lower, variable, cut-off). Indeed, there we saw that the quantum averaging over a shell in momentum space naturally leads to all possible terms in the effective action (all terms allowed by symmetries), with coefficients given by the quantum averaging. It is left as an exercise to show in a bit more concrete way this statement in our case.

Then, instead of starting with a specific Hamiltonian in the UV, at the physical cut-off scale given by a, and integrating over the blocks to get new terms in the Hamiltonian, we can, like in the Wilsonian effective field theory approach from last lecture, start instead already from an effective field theory of general type, parametrizing our ignorance about the UV.

That is, start with the Hamiltonian

$$H(\phi) = \sum_\alpha K_\alpha S_\alpha(\phi)\ , \tag{28.33}$$

where $S_\alpha(\phi)$ are all possible terms in the Hamiltonian, and K_α are couplings. Then blocking amounts to just a transformation on the space of coefficients $\{K_\alpha\}$, from the original (few) ones to the final (more) ones.

That is in fact identified with the *renormalization group* (RG) transformation, since that was also a scaling transformation, just that in momentum space. Formally, the transformation $T_{RG}(s)$ acts as

$$T_{RG}(s) : \{K_\alpha\} \to \{(T_{RG}(s)K)_\alpha\}. \tag{28.34}$$

Then a fixed point of the renormalization group is one that doesn't change the couplings, i.e.

$$(T_{RG}(s)K)^*_\alpha = K^*_\alpha \tag{28.35}$$

for any α. We thus understand the renormalization group as the coarse graining procedure, i.e. blocking.

We define the *critical surface* as the set of all points attracted towards the fixed point by the RG transformation, i.e. the basin of attraction of the fixed point under RG,

$$(T^n_{RG}(s)K)_\alpha \to K^*_\alpha \tag{28.36}$$

for $n \to \infty$.

A fixed point of the RG group is then identified with the critical point, i.e. the phase transition point, in critical phenomena, since as we saw that is a scale invariant point. Then

$\xi \to \infty$ there, which means that $\xi \to \infty$ on the critical surface (since as we go towards the critical point we decrease ξ).

Then it means that the couplings of various materials $K_\alpha(\beta)$ belonging to the same critical surface ($\xi \to \infty$ at $\beta \to \beta_c$) are different, yet they lead to the same long distance physics, defined by the critical point (since the blocking means going to larger distances, and so the fixed point is the IR behaviour of the theory). This is the universality that we mentioned before.

Expansion near a critical point.

Expanding the couplings near the critical point K_α^*,

$$K_\alpha = K_\alpha^* + \delta K_\alpha \ , \tag{28.37}$$

the RG transformation is

$$(T_{RG}(s)K)_\alpha = K_\alpha^* + \sum_{\alpha'} T_{\alpha\alpha'}\delta K_{\alpha'} + ... \tag{28.38}$$

Then considering a basis $v_{a\alpha}$ of eigenfunctions of $T_{\alpha\alpha'}$ with eigenvalue λ_a,

$$\sum_{\alpha'} T_{\alpha\alpha'} v_{a\alpha'} = \lambda_a v_{a\alpha} \ , \tag{28.39}$$

we can expand the coupling variations in it as

$$\delta K_\alpha = \sum_a h_a v_{a\alpha}. \tag{28.40}$$

The Hamiltonian is then written in this basis as (remember that $H = \sum_\alpha K_\alpha S_\alpha$)

$$H = H^* + \sum_a h_a v^a \ , \tag{28.41}$$

where $H^* = \sum_\alpha K_\alpha^* S_\alpha$ and

$$v^a = \sum_\alpha v_{a\alpha} S_\alpha. \tag{28.42}$$

Then the RG transformation acts on the eigenfunctions as $v_{a\alpha} \to \lambda_a v_{a\alpha}$, thus $v^a \to \lambda_a v^a$, so after n steps

$$H \to H^* + \sum_a (\lambda_a)^n h_a v^a. \tag{28.43}$$

Therefore we can distinguish between:

- $\lambda_a < 1$, which multiplies an *irrelevant operator* v^a, since the term is suppressed after a few steps of the RG transformation. Note that then the RG transformation along the irrelevant operator takes us towards the fixed point ($H \to H^*$).

- $\lambda_a > 1$, which multiplies a *relevant operator* v^a, as the RG transformation amplifies the effect of the operator. The RG transformation along the relevant operator takes us away from the fixed point.

- $\lambda = 1$ multiplies a *marginal operator*. It means we need to consider higher orders in the perturbation (beyong the linear analysis) to see whether it is actually relevant or irrelevant. If it remains marginal to all orders, we say it is *exactly marginal*, and the RG transformation has no effect on it.

We see then that the critical surface is spanned by irrelevant operators, i.e. irrelevant deformations of the Hamiltonian.

On the other hand, the relevant deformations of the Hamiltonian take us away from the fixed point: after a few steps, the RG trajectory is dominated by them, more precisely be the relevant deformation with the largest λ_a (the largest relevant deformation), that we will call λ_{a_1}, as we can see from Fig.74.

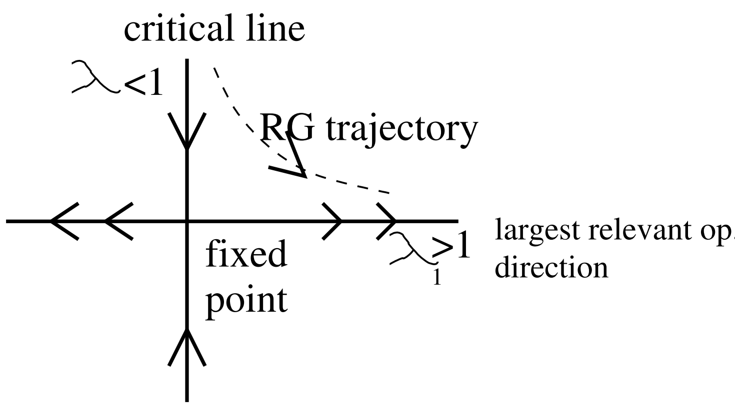

Figure 74: RG trajectory near fixed point.

Critical exponents (near the fixed point).

We then isolate this largest deformation with λ_1, writing

$$(T_{RG}(s)K)_\alpha = K^*_\alpha + \lambda_{a_1} h_{a_1}(\beta) v_{a_1\alpha} + \sum_{i\geq 2} \lambda_{a_i} h_{a_i}(\beta) v_{a_i\alpha} + \mathcal{O}(h^2). \tag{28.44}$$

We also assume that $h_{a_1}(\beta)$ is smooth, and thus can be expanded near the phase transition as

$$h_{a_1}(\beta) = (\beta - \beta_c) h^c_{a_1} + \mathcal{O}((\beta - \beta_c)^2). \tag{28.45}$$

Then, after a few steps in the RG transformation, we have

$$(T^n_{RG}(s)K)_\alpha \simeq K^*_\alpha + \lambda^n_{a_1} h_{a_1}(\beta) v_{a_1\alpha}. \tag{28.46}$$

Since $\lambda_{a_1} > 1$, we can define a $\nu > 0$ by the equality

$$\lambda_{a_1} \equiv s^{1/\nu} \,, \tag{28.47}$$

where s is the blocking size, as before.

We can moreover define a sequence $\beta_n \to \beta_c$ of β's by the relation

$$h_{a_1}(\beta_n)\lambda_{a_1}^n = h_{a_1}(\beta_n)s^{n/\nu} = 1 , \tag{28.48}$$

which implies

$$s^n \simeq \frac{1}{|(\beta_n - \beta_c)h_{a_1}^c|^\nu}. \tag{28.49}$$

It follows that

$$(T_{RG}^n(s)K)_\alpha \to K_\alpha^* + v_{a_1\alpha} \tag{28.50}$$

as $n \to \infty$, so we stay fixed away from the fixed point.

On the other hand, the 2-point correlation function goes like

$$\langle \phi_n \phi_0 \rangle \sim \frac{1}{n^{2\alpha}} , \tag{28.51}$$

which means that under blocking it transforms as

$$G(|x|, \{K_\alpha\}) = s^{-2\alpha} G(|x|/s; \{(T_{RG}(s)K)_\alpha\}) = ... = s^{-2n\alpha} G(|x|/s^n; \{(T_{RG}^n(s)K)_\alpha\}) , \tag{28.52}$$

but since we stay fixed away from the fixed point as $n \to \infty$, there is no nontrivial dependence on the couplings under scaling. Therefore at $n \to \infty$ and for large x, we have

$$G = G(|x|/s^n) , \tag{28.53}$$

but on the other hand in general we have

$$G \sim e^{-\frac{|x|}{\xi}}. \tag{28.54}$$

By comparing the two, we see that we need to have

$$\xi(\beta_n) \propto s^n , \tag{28.55}$$

i.e.

$$\xi(\beta) \propto \frac{1}{|\beta - \beta_c|^\nu}. \tag{28.56}$$

That also means that we have

$$G(|x|, \{K_\alpha(\beta)\}) = ... = \xi^{-2\alpha} G(|x|/\xi(\beta), \{K_\alpha^* + v_{a_1\alpha}\}) , \tag{28.57}$$

where $2\alpha = d - 2 + \eta$, so by Fourier transforming to momentum space we get

$$G(p, \{K_\alpha\}) = \xi^{2-\eta} G(\xi(p)b, \{K_\alpha^* + v_{a_1\alpha}\}). \tag{28.58}$$

The momentum space formulation is understood as the Wilsonian effective action formulation from last lecture.

Important concepts to remember

- A discretized scalar field theory can be written as a spin system.
- The two-point function at large distances decays exponentially with a correlation length and at intermediate distances it decays as a power law with an anomalous dimension.
- The scaling hypothesis states that all divergent behaviour in physical quantities near a critical point (phase transition) are due to the diverging correlation length.
- The magnetic susceptibility and the correlation length have (semi-)universal critical exponents.
- Kadanoff blocking is averaging over blocks of size s in each direction. It leads to new terms in the Hamiltonian.
- Kadanoff blocking can be thought of as a RG transformation on the couplings K_α of a Hamiltonian with an infinite set of terms.
- The fixed point of the RG corresponds to the critical point of a system (at the phase transition), and the critical surface is the basin of attraction of the fixed point: various materials with various K_α all have the same long distance physics.
- Irrelevant operators take us towards the fixed point, and they are suppressed after a few steps, while relevant operators take us away from the fixed point, and after a few steps the largest irrelevant operator (of largest eigenvalue) dominates.

Further reading: See chapter 9.2 and 9.3 in [5].

Exercises, Lecture 28

1) Consider the discrete model with Hamiltonian

$$H=\sum_{j=1}^{J}\frac{a_j^\dagger a_j+a_ja_j^\dagger}{2}+\frac{2\lambda}{(2\pi)^2}\sum_{j=1}^{J}(\phi_j-\phi_{j+1})^2. \tag{28.59}$$

Write down the continuum version for it, find H and then the relativistic Lagrangean, and calculate the length L of the continuum system.

2) Show that by Kadanoff blocking on

$$H=\sum_n\left[\sum_\mu\frac{1}{2}(\phi_{n+\mu}-\phi_n)^2+\mu\phi_n^2+\lambda\phi_n^4\right] \tag{28.60}$$

we generate terms like

$$(\phi_{n+\mu}-\phi_n)^2\phi_n^2,\quad (\phi_{n+\mu}-2\phi_n+\phi_{n-\mu})^2 \tag{28.61}$$

and others.

29 Lecture 29. Lattice field theory

In this lecture we will see that we can define a discretization of the field theory in a consistent way, such that we recover the continuum theory in a certain limit.

It is important to put field theories on the lattice, since then we can calculate non-perturbative quantities from computer simulations (Monte Carlo calculations for the path integral).

We will first consider the continuum limit, related to the Kadanoff blocking from last lecture.

Continuum limit

An important result is that the critical point is independent of the particular RG procedure considered (the representation of the RG group; before, Kadanoff blocking).

We had described everything in terms of lattice units, but in order to define physical quantities we need to multiply by the lattice spacing a,

$$\xi_{ph} = \xi_l a; \quad x_{ph} = x_l a; \quad p_{ph} = \frac{p_l}{a} \ , \tag{29.1}$$

where ξ_l, x_l, p_l are in lattice units.

We need to keep ξ_{ph} fixed, but we now change the lattice spacing a by $a \to a/s$ under the RG action, thus taking $a \to 0$ in the limit instead of $\beta \to \beta_c$ ($\beta_n \to \beta_{n+1}, ...$). A physical mass will be $m_{ph} = 1/\xi_{ph}$, and will be fixed as $a \to 0$.

To define continuum correlation functions, we need to make a wave function renormalization according to

$$G_{\rm cont}(p_{ph}, m_{ph}) = a^{-\eta} G(p_l; \{K_\alpha(a)\})|_{a\to 0} \ , \tag{29.2}$$

where η is the anomalous dimension, and it appears since $G \sim 1/|x|^{\Delta_{class}+\eta}$ and the extra dependence must be compensated.

Gaussian fixed point

Under the RG transformation $x \to x/s$, since the engineering (classical) dimension of a scalar field is $(d-2)/2$, we have

$$\phi_s(x/s) \simeq s^{\frac{d-2}{2}} \phi(x). \tag{29.3}$$

Then, for a coupling

$$K_n \int d^d x \phi^n(x) \ , \tag{29.4}$$

the coupling K_n transforms approximately as

$$K_n \to s^{-d_n} K_n \ , \tag{29.5}$$

where

$$d_n = [K_n] = -\frac{d-2}{2} n + d \tag{29.6}$$

is the engineering (classical) dimension of K_n.

That in turn means that the eigenvalue of the RG operator $T_{RG}(s)$ is

$$\lambda_n = s^{-d_n}. \tag{29.7}$$

We deduce then that near the Gaussian fixed point $\phi = 0$,

- relevant operators, with $\lambda_n > 1$, have $d_n = [K_n] < 0$, so make the theory super-renormalizable.

- marginal operators, with $\lambda_n = 1$, have $d_n = [K_n] = 0$, so make the theory renormalizable.

- irrelevant operators, with $\lambda_n < 1$, have $d_n = [K_n] > 0$, so make the theory non-renormalizable.

In particular, $\lambda \int \phi^4$ has $[\lambda] = 0$, so is marginal, which means that we need to go to higher orders to see whether it is relevant, irrelevant, or exactly marginal.

Then relevant (super-renormalizable) operators are found to be UV asymptotically free, and irrelevant(non-renormalizable) operators are IR asymptotically free. For the latter, at the gaussian fixed point, when we take the bare coupling to zero in the IR, the renormalized coupling goes to zero even faster, so it seems impossible to define a nontrivial theory (with nonzero physical coupling on large scales). $\lambda\phi^4$ and QED are of this type, which has generated a debate on whether these theories make sense non-perturbatively (we can define the theory order by order, but what it means non-perturbatively it is not clear). One possibility is that there are *other* fixed points, where one could define the theory.

For UV asymptotically free theories on the other hand, there is no problem, since we can keep fixed the renormalized coupling at the fixed point (in the IR), and obtain a nontrivial interacting theory.

Beta function

The general RG procedure allows us to define the beta function in an alternative way. By rescaling the correlation length by s,

$$\xi(g'^2) = s\xi(g^2) \ , \tag{29.8}$$

the couplings are changed in general, so

$$g'^2 = g^2 - \Delta g^2. \tag{29.9}$$

But we want the long distance physics to be unmodified by this rescaling procedure, so we need to rescale also a,

$$a(g'^2) = \frac{1}{s} a(g^2) \ , \tag{29.10}$$

which defines $a(g^2)$ or reversely $g(a)$.

Then, after $n \to \infty$ blocking steps, we should be on the same RG trajectory, since that defines the long distance physics, so the different blockings create various dotted lines that converge to the RG line, see Fig.75.

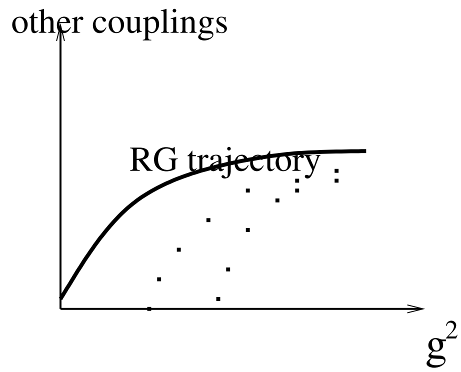

Figure 75: Kadanoff blocking near the RG trajectory.

This procedure allows us to define an alternative way to define the beta function, by

$$\beta(g) = -a\frac{d}{da}g(a). \tag{29.11}$$

Note that this definition is not identical to the usual one, however one can prove that the first two coefficients in the expansion of $\beta(g)$ are the same, and of course the fixed points are the same (since the asymptotic RG trajectory is the same).

Lattice gauge theory

We finally come to the issue of interest, namely how to put a gauge theory on the lattice. This is of interest, since QCD is a nonperturbative gauge theory in the IR, so it is hard to calculate anything at low energies, other than on the lattice.

Consider a gauge group G, so because of the local gauge invariance with G_x, we can say that we have a total group

$$G_{\rm total} = \prod_{x\in\mathbb{R}^d} G_x. \tag{29.12}$$

We saw that we can define an observable called the Wilson loop,

$$\tilde{U}_C = {\rm Tr}\, P \exp\left[i\oint_C A_\mu dx^\mu\right] , \tag{29.13}$$

where $A_\mu = A_\mu^a T_a$ and the contour C is parametrized by $t \in [0,1]$ as $C(t) : t \to x_\mu(t)$. It is the trace of a Wilson line U_C for a closed contour C, where U_C satisfies composition, i.e.

$$U_C = U_{C_n} U_{C_{n-1}} ... U_{C_1} , \tag{29.14}$$

and where

$$U_C(t) = \lim_{n\to\infty} \prod_{k=0}^{n} e^{idx_k^\mu A_\mu(x_k)} , \tag{29.15}$$

and $x_k \in dx_k$. We also saw that the Wilson line U_C transforms covariantly, but with different endpoints on the left and right, i.e.

$$U_C(t) \to V(x(t)) U_C(t) V(x(0)) , \tag{29.16}$$

for

$$A_\mu \to V(x)A_\mu(x)V(x)^{-1} - i\partial_\mu V(x)V^{-1}. \tag{29.17}$$

In the abelian case, we saw that can use Stokes's theorem to write

$$e^{i\oint_C A_\mu dx^\mu} = e^{\frac{i}{2}F_{\mu\nu}da_{\mu\nu}} , \tag{29.18}$$

where $da_{\mu\nu}$ is the (infinitesimal) surface bounded by the (infinitesimal) contour C.

In the nonabelian case, we have corrections in the exponent,

$$Pe^{i\oint_C A_\mu dx^\mu} = e^{\frac{i}{2}F_{\mu\nu}da_{\mu\nu}+\mathcal{O}(da^2)}. \tag{29.19}$$

By taking the trace and then the real part, and expanding the exponential, we obtain

$$\mathrm{Re}\,\mathrm{Tr}\,Pe^{i\oint_C A_\mu dx^\mu} = \mathrm{Tr}\,I - \frac{1}{8}\mathrm{Tr}(F_{\mu\nu}da_{\mu\nu})^2 + \mathcal{O}(da_{\mu\nu}^3). \tag{29.20}$$

We are finally ready to define the gauge theory on the lattice. On a lattice, we have sites i and links (ij) connecting them, on which naturally one can define an orientation. Then it is natural to associate an element of G to each *oriented* link, $U_{ij} \in G$, which then has the properties of a Wilson line. Indeed, reversing the orientation and composing the two elements we should get back to the identity, so we should have

$$U_{ji} = U_{ij}^{-1}. \tag{29.21}$$

Moreover, the gauge transformation of U_{ij} is the same as for the Wilson line, namely

$$U_{ij} \to U_{ij}^{(V)} = V_i U_{ij} V_j^{-1}. \tag{29.22}$$

Then it is also natural to define along a connected path C on the links the analog of a long Wilson line by composition,

$$U_C = U_{(i_n i_{n-1})}U_{(i_{n-1}i_{n-2})}...U_{(i_3 i_2)}U_{(i_2 i_1)} , \tag{29.23}$$

and then again $\tilde{U}_C = \mathrm{Tr}\,U_C$ for a closed path is gauge invariant.

Note that now the total gauge group is

$$G_{\text{total}} = \prod_{i\in a\mathbf{Z}^d} G_i. \tag{29.24}$$

We associate the link variable U_{ij} with the elementary Wilson line (for the link), which now can be written as an object on the link as

$$U_{ij} \leftrightarrow Pe^{i\int_0^1 dx_\mu(t)A_\mu(x(t))+\mathcal{O}(a^2)} , \tag{29.25}$$

where the path $x_\mu(t)$ has to be between $x_\mu(i)$ and $x_\mu(j)$, so

$$x_\mu(t) = tx_\mu(j) + (1-t)x_\mu(i) , \tag{29.26}$$

and the point at which we define the object on the link is the midpoint,

$$x_\mu = \frac{x_\mu(i) + x_\mu(j)}{2}. \tag{29.27}$$

The smallest nontrivial closed loop is called a "plaquette", and is a square with nearest neighbour vertices, $p = (ijkl)$, with area a^2. Then from (29.20), we have for the plaquette p loop

$$\mathrm{Tr}\, I - \mathrm{Re}\,\mathrm{Tr}\, U_p = \frac{1}{2} a^4\, \mathrm{Tr}\, F^2_{\mu(p)\nu(p)} + \mathcal{O}(a^6)\ , \tag{29.28}$$

where the plaquette square is in the directions $\mu(p)$ and $\nu(p)$.

Then we can define from the above the Yang-Mills action for the plaquette ($\mathrm{Tr}\, F^2_{\mu\nu}/4$), and then sum over plaquettes, to obtain the *Wilson action* for the gauge theory on the lattice,

$$S_W[U] = \beta_G \sum_p \left(1 - \frac{1}{N_G}\mathrm{Re}\,\mathrm{Tr}\, U_p\right)\ , \tag{29.29}$$

(note that we have absorbed a factor of $N_G = \mathrm{Tr}\, I$ in the definition of the coupling β_G) which goes over in the continuum limit to

$$\frac{1}{g_0^2}\int d^d x \frac{1}{4}\, \mathrm{Tr}\, F^2_{\mu\nu}\ , \tag{29.30}$$

up to terms of order a^2 that vanish, provided we identify

$$\beta_G a^{4-d} = \frac{2N_G}{g_0^2}. \tag{29.31}$$

Here g_0 is the bare coupling constant in YM, and we have used $\int d^d x \sum_{\mu\nu} = a^d \sum_p$.

Note that the Wilson action is by no means unique, there are various actions that give rise to the same continuum limit, but the Wilson action is the simplest, so it is the one used by (almost) everyone.

Now that we have defined the lattice action, to formulate the gauge theory on the lattice we only lack a definition of the path integral measure.

The measure for integration is the unique measure dU for integration over the group that is invariant under both left and right group multiplication, i.e. $U \to UU_0$ and $U \to U_0 U$, called the *Haar measure*. It also has the property that if U is close to the identity, i.e. $U = \exp(iaA)$, then

$$dU = \prod_{i=1}^{N_G} dA_i^a (1 + \mathcal{O}(a)). \tag{29.32}$$

Therefore the partition function on the lattice is

$$Z[\beta_G] = \int \prod_{l\in(ij)\ \mathrm{links}} dU_l e^{-S_W[U]}. \tag{29.33}$$

Continuum limit

The naive continuum limit of the lattice YM action would be $\beta_G \to \infty$ (or equivalently $g_0 \to 0$), since β_G is the coupling of the lattice action, and by taking it to infinity we have a small fluctuation around $\mathrm{Tr}\, U_p = 1$.

But rather we need to do it as at the beginning of the lecture, by taking $a \to 0$ while fixing some physical length ξ. Then we have a RGE-like equation,

$$a\frac{d\xi}{da} \equiv \left(a\frac{\partial}{\partial a} - \beta(g_0)\frac{\partial}{\partial g_0}\right)\xi(a, g_0) = 0. \tag{29.34}$$

Here as before

$$\beta(g_0) \equiv -a\frac{\partial}{\partial a}g_0. \tag{29.35}$$

The solution of the RGE-like equation above is

$$\xi(a, g_0) = a \exp\int_0^{g_0}\frac{dg}{\beta(g)}\ , \tag{29.36}$$

like we can easily check. This means that a divergent correlation length (and thus a fixed point) corresponds to $\beta = 0$, as it should.

In a perturbative expansion,

$$\beta(g) = -\beta_0 g^3 - \beta_1 g^5 + \mathcal{O}(g^7)\ , \tag{29.37}$$

and for $G = SU(N)$, we have

$$\beta_0 = \frac{11}{3}\frac{N}{16\pi^2};\quad \beta_1 = \frac{34}{3}\frac{N^2}{16\pi^2}\ , \tag{29.38}$$

leading to the solution for ξ of

$$\xi = a\cdot(\beta_0 g_0^2)^{\beta_1/2\beta_0}\exp\left(\frac{1}{2\beta_0 g_0^2}\right). \tag{29.39}$$

Then for fixed ξ as $a \to 0$, this defines as $g_0 \to 0 \Rightarrow \beta_G \to \infty$, $a(g_0)$, i.e. $a(\beta_G)$.

For a physical mass $m(\beta_G)$ that implies a continuum limit mass scale (for something like a glueball), $m(\beta_G)\xi(\beta_G)$ is finite as $\beta_G \to \infty$ if $a = 1$ (in lattice units) (since then the expression for $\xi(a, g_0)$ reduces to $\xi(\beta_G)$).

Adding matter.

Up to now we have seen how to describe pure YM. But for physical applications, we want to also have matter in the theory. For instance, for QCD we would need to have quarks. But unfortunately, it is difficult to put chiral fermions on the lattice while keeping the chiral symmetry. There is no perfect way to deal with chiral fermions.

So instead we will show how to put scalars on the lattice, having in mind the application to the electroweak theory, so the scalars representing the Higgs. Then the scalars $\phi(x)$ are in the fundamental representation of the gauge group.

For a scalar $\phi(x)$ defined at sites i, we have now the variables ϕ_i. The gauge invariance of the scalars is

$$\phi(x) \to V(x)\phi(x)\ , \tag{29.40}$$

so to obtain a gauge invariant observable we must form the composite object from two scalars and a Wilson line

$$\phi^\dagger(y) P\ e^{i\int_{C:x\to y} A_\mu dx^\mu}\phi(x). \tag{29.41}$$

This object is easily discretized to the lattice object

$$\phi_i^\dagger U_{ij}\phi_j. \tag{29.42}$$

Next, we want to write an action, so we need to write a discrete version of the derivative, easily seen to be

$$\partial_\mu\phi(x)\to\phi_{j+\mu}-\phi_j. \tag{29.43}$$

But we actually need the covariant derivative $D_\mu\phi$, and since $U=e^{iaA}\simeq 1+iaA$, we have

$$D_\mu\phi(x)\to\phi_{j+\mu}-U_{j+\mu,j}\phi_j. \tag{29.44}$$

Then the kinetic term of the scalar is

$$|D_\mu\phi|^2=\sum_{j,\mu}(\phi_{j+\mu}-U_{j+\mu,j}\phi_j)^2=2\sum_j\phi_j^2-2\sum_{(ij)\text{ links}}\phi_i U_{ij}\phi_j\ , \tag{29.45}$$

where (ij) are links (between nearest neighbours).

In the gauge-Higgs action we need to add a mass term and a ϕ^4 term with independent coefficients besides the Wilson action for the gauge fields, for a total action

$$S[\phi,U]=-\kappa\sum_{(ij)}\mathrm{Re}\phi_i^\dagger U_{ij}\phi_j+\mu\sum_i\phi_i^\dagger\phi_i+\lambda\sum_i(\phi_i^\dagger\phi_i)^2+\beta_G\sum_p\mathrm{Re}\,\mathrm{Tr}\,U_p. \tag{29.46}$$

Important concepts to remember

- To take the continuum limit, we take $a\to 0$, while keeping physical scales like $\xi_{ph}=\xi_l a$ fixed.
- At the gaussian fixed point, relevant interactions are super-renormalizable, marginal are renormalizable and irrelevant are non-renormalizable.
- Irrelevant operators at the gaussian fixed point could lead to a trivial theory (free on physical scales): QED and ϕ^4 are in this class.
- We can vary $\xi(g)$ by $\xi(g'^2)=s\xi(g^2)$, leading to $a(g^2)$ for an invariant long distance physics, for $a(g'^2)=a(g^2)/s$, leading to a new definition of the beta function, as $-a\,dg(a)/da$.
- In lattice gauge theory, the variables are the links $U_{ij}\in G$, acting as infinitesimal Wilson lines. The Wilson action for lattice gauge theory is written in terms of them.

- The measure on the discrete path integral is the Haar measure dU, invariant under left and right multiplications, $U \to UU_0$ and $U \to U_0 U$.
- In the continuum limit, we keep physical lengths fixed, obtaining RGE-like equations from $a\, d\xi/da = 0$.
- Fundamental scalars ϕ can be added, with covariant derivative $D_\mu \phi \to \phi_{j+\mu} - U_{j+\mu,j}\phi_j$.

Further reading: See chapter 9.5 in [5].

Exercises, Lecture 29

1) Calculate the number of sites, links and plaquettes for a symmetric hypercubic lattice with periodic boundary conditions.

2) Derive the lattice version of the Yang-Mills equations (equations of motion of Yang-Mills, or nonabelian Maxwell's equations).

30 Lecture 30. The Higgs mechanism

The subject of this lecture is spontaneous symmetry breaking, which is the situation when a theory is invariant under some local (gauge) symmetry, but the vacuum breaks it, i.e. the vacuum is not invariant under the symmetry. In this case, gauge fields become massive by the Higgs mechanism, and they "eat" a scalar degree of freedom along the symmetry direction in order to become massive. This works out, because a massive vector in 4 dimensions has 3 degrees of freedom, while a gauge field (massless vector) has 2, the difference being supplied by the eaten scalar.

Abelian case

We start with the simplest version of the mechanism, in the abelian case, that will be called "Abelian-Higgs", even though there this case is mostly relevant for superconductivity, and not for particle physics.

The Lagrangean for a complex scalar coupled to a gauge field is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - |D_\mu\phi|^2 - V(|\phi|) \ , \tag{30.1}$$

where $D_\mu = \partial_\mu - ieA_\mu$. The gauge invariance is

$$\begin{aligned} \phi(x) &\rightarrow e^{i\alpha(x)}\phi(x) \\ A_\mu(x) &\rightarrow A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x). \end{aligned} \tag{30.2}$$

We choose the most general analytic, gauge invariant, renormalizable potential with a symmetry breaking term,

$$V = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2 \ , \tag{30.3}$$

or rather, by adding a constant (that is irrelevant as long as the model is not coupled to gravity)

$$V = \frac{\lambda}{2}\left(|\phi|^2 - \frac{\mu^2}{\lambda}\right)^2. \tag{30.4}$$

Indeed, a renormalizable potential must have powers less or equal to 4, and a gauge invariant and analytic potential implies only 2nd and 4th powers. We must have $\mu^2 = -m^2 > 0$ for symmetry breaking to occur. Then there is a VEV for ϕ and the $U(1)$ is spontaneously broken, giving a vacuum

$$\langle\phi\rangle = \phi_0 \equiv \sqrt{\frac{\mu^2}{\lambda}}. \tag{30.5}$$

The various vacua are $\phi = \phi_0 e^{i\theta_0}$, with the arbitrary phase θ_0 parametrizing the vacuum. Without loss of generality, we can set the particular vacuum we are in to be real as above, by a gauge transformation. The potential is called the "Mexican hat potential" due to its shape. Spontaneous symmetry breaking means that, starting at $\phi = 0$, a small fluctuation will lead us to a random direction, choosing some particular vacuum among the continuum of vacua, and thus breaking the invariance.

Expanding around this vacuum, we obtain

$$\phi(x) = \phi_0 + \frac{\phi_1(x) + i\phi_2(x)}{\sqrt{2}} \; , \tag{30.6}$$

and the potential is

$$V(\phi) = \frac{1}{2} 2\mu^2 \phi_1^2 + \mathcal{O}(\phi_1^3). \tag{30.7}$$

That means that ϕ_1 has mass $m = \sqrt{2}\mu$ and ϕ_2 is massless and is the so-called "Goldstone boson". It is a theorem that will be proved next lecture that for every symmetry that is spontaneously broken there is an associated massless scalar called a Goldstone boson.

Expanding the scalar kinetic term, we get

$$|D_\mu \phi|^2 = \frac{1}{2}(\partial_\mu \phi_1)^2 + \frac{1}{2}(\partial_\mu \phi_2)^2 + \frac{e^2\mu^2}{\lambda} A_\mu A^\mu - \sqrt{2} e \phi_0 A_\mu \partial^\mu \phi_2 + ... \tag{30.8}$$

The third term in this expansion gives a mass term for the vector, with

$$m_A^2 = 2e^2 \phi_0^2 \; , \tag{30.9}$$

while the last term gives a mixing between the vector A_μ and the Goldstone boson ϕ_2. We can of course redefine the fields such as to get rid of this term, and in the process get rid of ϕ_2, and we will do so shortly, but before let's keep it and see that the resulting theory after the expansion still is OK quantum mechanically.

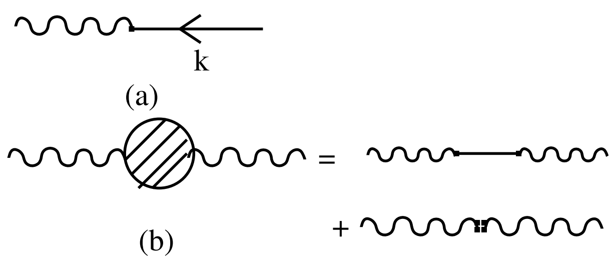

Figure 76: (a) Mixing term between the scalar and a gauge field. (b) The sum of the two diagrams with mixing gives the 1PI 2-point function, which is correctly transverse.

Consider the mixing term between the scalar of momentum k and a gauge field with index μ. The Feynman rule for it is

$$i\sqrt{2} e \phi_0 (-ik^\mu) = m_A k^\mu. \tag{30.10}$$

Considering now the modification to the photon propagator, it comes from the vector mass term, and the diagram with a mixing to scalar, scalar propagation, followed by scalar mixing again. The sum of these two Feynman diagrams gives then

$$-im_A^2 g_{\mu\nu} + (m_A k^\mu) \frac{-i}{k^2}(-m_A k^\nu) = -im_A^2 \left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right) \; , \tag{30.11}$$

which is transverse, as it should be.

Unitary gauge

The unitary, or physical gauge, is the gauge where we choose a local gauge transformation $\alpha(x)$ such as to get rid of the scalar degree of freedom in the $U(1)$ direction, i.e. we put $\phi_2 = 0$, ϕ real. Then the Lagrangean becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - (\partial_\mu\phi)^2 - e^2\phi^2 A_\mu A^\mu - V(\phi). \tag{30.12}$$

The Higgs mechanism means a mass for the gauge boson by "eating" the Goldstone boson. We can see this mechanism explicitly in the parametrization for the scalar

$$\phi = |\phi| e^{i\theta}. \tag{30.13}$$

Then the covariant derivative of the scalar is

$$D_\mu\phi = e^{i\theta}(\partial_\mu|\phi| + i\partial_\mu\theta|\phi| - ieA_\mu|\phi|). \tag{30.14}$$

By choosing the vacuum $\langle|\phi|\rangle = \phi_0 = \mu^2/\lambda$, $\theta_0 = 0$ and expanding around it as

$$|\phi| = \frac{\mu^2}{\lambda} + \frac{\delta|\phi|}{\sqrt{2}} , \tag{30.15}$$

the θ degree of freedom is removed by the redefinition

$$A_\mu \to A_\mu = A'_\mu + \frac{1}{e}\partial_\mu\theta , \tag{30.16}$$

which leaves the field strength invariant, $F'_{\mu\nu} = F_{\mu\nu}$. Then after the redefinition, the scalar kinetic term is

$$|D_\mu\phi|^2 = (\partial_\mu|\phi|)^2 + e^2 A'_\mu A'^\mu |\phi|^2 , \tag{30.17}$$

and the Lagrangean is

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}{F'_{\mu\nu}}^2 - (\partial_\mu|\phi|)^2 - e^2{A'_\mu}^2|\phi|^2 - V(|\phi|)\\ &\simeq -\frac{1}{4}{F'_{\mu\nu}}^2 - \frac{(\partial_\mu\delta|\phi|)^2}{2} - \frac{e^2\mu^2}{\lambda}{A'_\mu}^2.\end{aligned} \tag{30.18}$$

We see that we have removed the angle θ degree of freedom in $\phi = |\phi|e^{i\theta}$, which is the Goldstone boson (massless degree of freedom, since $V = V(|\phi|)$).

An important historical note is that the "Higgs mechanism" was discovered jointly by Higgs, Kibble, Guralnik, Hagen, Englert and Brout. In fact however, they explored the model, and generalized to nonabelian gauge theory, but before them the abelian model had been used in condensed matter to describe supeconductivity, specifically the Meissner effect, understood as the phenomenon of the photon becoming massive due to spontaneous symmetry breaking, and thus penetrating only a distance of $\mathcal{O}(1/m_{photon})$ inside the superconductor.

Gauge symmetry

What happened to the gauge symmetry that was present in the theory before the Higgsing? Before the Higgsing, the invariance was

$$\begin{aligned} \delta A_\mu &= \partial_\mu \alpha \\ \delta\phi &= |\phi| e^{i\theta + ie\alpha} - |\phi| e^{i\theta} \Rightarrow \delta\theta = e\alpha \ , \end{aligned} \tag{30.19}$$

and expanding around the vacuum does nothing to the symmetry, but the redefinition does. After the redefinition, we have

$$\begin{aligned} \delta A'_\mu &= \delta\left(A_\mu - \frac{1}{e}\partial_\mu\theta\right) = 0 \\ \delta|\phi| &= 0. \end{aligned} \tag{30.20}$$

So the gauge invariance is simply lost (nothing happens under it), but only after the redefinition.

Note that, while it is also common to call ϕ "the Higgs", actually the real scalar $|\phi|$ (or ϕ_1) has mass $m = \sqrt{2}\mu$, and deseves the name "the Higgs", since it is the massive boson associated with the Higgs particle.

Note that in the scalar kinetic term expansion we have a term

$$e^2\sqrt{2}\phi_1\phi_0 A_\mu A^\mu = \sqrt{2}e^2\phi_0\phi_1 A'_\mu A'^\mu + \ldots \ , \tag{30.21}$$

which leads to a vertex for 2 gluons to join and emit a Higgs, of form

$$\sqrt{2}e^2\phi_0\delta_{\mu\nu}. \tag{30.22}$$

Nonabelian case

We consider now a nonabelian gauge group G, such that the gauge transformation is

$$\phi_i \to (e^{-i\alpha^a t_a})_{ij}\phi_j = (e^{\alpha^a T_a})_{ij}\phi_j \ , \tag{30.23}$$

where we have written the transformation with Hermitian and anti-Hermitian generators. The covariant derivative is

$$D_\mu\phi_i = (\partial_\mu - igA^a_\mu t_a)_{ij}\phi_j = (\partial_\mu + gA^a_\mu T_a)_{ij}\phi_j. \tag{30.24}$$

Therefore the kinetic term for the scalar is

$$\frac{1}{2}(D_\mu\phi_i)^2 = \frac{1}{2}(\partial_\mu\phi_i)^2 + gA^a_\mu\partial^\mu\phi_i T^a_{ij}\phi_j + \frac{g^2}{2}A^a_\mu A^{b\mu}(T^a\phi)_i(T^b\phi)_i. \tag{30.25}$$

Consider the VEV

$$\langle\phi_i\rangle = (\phi_0)_i. \tag{30.26}$$

Then from the kinetic term for the scalars we obtain the mass terms

$$\Delta\mathcal{L}_{mass} = -\frac{1}{2}m^2_{ab}A^a_\mu A^{b\mu} \ , \tag{30.27}$$

where

$$m^2_{ab} = g^2(T^a\phi_0)_i(T^b\phi_0)_i \ , \tag{30.28}$$

and the diagonal elements are

$$m^2_{aa} = g^2(T^a\phi_0)^2 \geq 0. \tag{30.29}$$

Thus the vectors have mass squared positive or zero. Unlike the abelian case, when we had only nonzero masses, in the general nonabelian case we can have masses that are zero, corresponding to the existence of unbroken gauge fields.

Let us consider the general case, of a gauge symmetry with gauge group G, broken spontaneously by the vacuum to a subgroup H that leaves the vacuum invariant. (Note that in the abelian case, the vacuum was just a point, so there was no unbroken gauge group). Then the generators in the coset G/H take us from one vacuum to another (equivalent one). These generators give the Goldstone bosons then, which were the whole $U(1)$ in the abelian case.

The unbroken, massless generators, correspond to

$$m^2_A = 0 \Rightarrow (T^a\phi_0) = 0 \ , \tag{30.30}$$

and these conditions define the subgroup H. We can redefine again the vectors as

$$A'^{\,a}_\mu = A^a_\mu - \partial_\mu\phi_i(T^a\phi_0)_i \ , \tag{30.31}$$

which means that the broken vectors (with nonzero $T^a\phi_0$) eat the Goldstone bosons and become massive.

Again the existence of the mixing vertex

$$gA^a_\mu\partial_\mu\phi_i(T^a\phi_0)_i \tag{30.32}$$

implies the transversality of the theory, so the consistency of the theory at the quantum level, since the sum of the mass term diagram and the mixing, Goldstone scalar propagation, and mixing again Feynman diagram gives

$$-im^2_{ab}g_{\mu\nu} + \sum_j gk^\mu(T^a\phi_0)_j\frac{-i}{k^2}(-gk^\nu(T^b\phi_0)_j) = -im^2_{ab}\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) \ , \tag{30.33}$$

which is a transverse vacuum polarization.

$SU(2)$ **case**

The simplest nonabelian case is the $SU(2)$ case, also relevant since it is rather close to the electroweak theory, where we have $SU(2)\times U(1)$. Consider a scalar ϕ_i that is a doublet of $SU(2)$, and

$$D_\mu\phi = (\partial_\mu - igA^a_\mu\tau^a)\phi \ , \tag{30.34}$$

where $\tau_a = \sigma_a/2$ and σ_a are the Pauli matrices.

The potential is

$$V = \lambda\left(\phi^\dagger\phi - \frac{\mu^2}{\lambda}\right)^2 . \tag{30.35}$$

Choose the vacuum (as before, we can rotate it to this form by a gauge transformation)

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix}. \tag{30.36}$$

Then the scalar kinetic term is

$$|D_\mu\phi|^2 = \frac{1}{2}g^2\begin{pmatrix}0 & v\end{pmatrix}\tau^a\tau^b\begin{pmatrix}0\\v\end{pmatrix}A^a_\mu A^{b\mu}. \tag{30.37}$$

Since we have symmetry in (ab), we can replace $\tau^a\tau^b$ with $\{\tau^a,\tau^b\}/2 = \delta^{ab}/4$, and obtain finally

$$\Delta\mathcal{L}_{mass} = \frac{g^2v^2}{8}A^a_\mu A^{a\mu}\ , \tag{30.38}$$

i.e., a vector mass

$$m_A = \frac{gv}{2}. \tag{30.39}$$

Standard Model Higgs: electroweak $SU(2)\times U(1)$

In the case of the electroweak theory, the covariant derivative is

$$D_\mu\phi = \left(\partial_\mu - igA^a_\mu\tau^a - i\frac{g'}{2}B_\mu\right)\phi\ , \tag{30.40}$$

and the last $(U(1))$ term is $-ig'Y_\phi B_\mu = -ig'B_\mu/2$.

Then we have

$$gA^a_\mu\tau^a + g'\frac{B_\mu}{2} = \frac{1}{2}\begin{pmatrix}gA^3_\mu + g'B_\mu & g(A^1_\mu - iA^2_\mu)\\ g(A^1_\mu + iA^2_\mu) & -gA^3_\mu + g'B_\mu\end{pmatrix}. \tag{30.41}$$

The potential

$$V(\phi) = \frac{\lambda}{4}\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^2 \tag{30.42}$$

has a real vacuum

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix} \tag{30.43}$$

that can be set by a gauge transformation from any other vacuum.

Then the mass term is

$$\begin{aligned}\mathcal{L}_{mass} &= -\frac{v^2}{8}\begin{pmatrix}0 & 1\end{pmatrix}\left|\begin{pmatrix}gA^3_\mu + g'B_\mu & g(A^1_\mu - iA^2_\mu)\\ g(A^1_\mu + iA^2_\mu) & -gA^3_\mu + g'B_\mu\end{pmatrix}\right|^2\begin{pmatrix}0\\1\end{pmatrix}\\ &= -\frac{v^2}{8}\left[g^2(A^1_\mu)^2 + g^2(A^2_\mu)^2 + (-gA^3_\mu + g'B_\mu)^2\right]\end{aligned} \tag{30.44}$$

Defining the fields

$$W^\pm_\mu = \frac{A^1_\mu \pm iA^2_\mu}{\sqrt{2}}$$

$$
\begin{aligned}
Z^0_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gA^3_\mu - g'B_\mu)\\
A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(g'A^3_\mu + gB_\mu)\ ,
\end{aligned}
\tag{30.45}
$$

the mass terms become

$$
\Delta\mathcal{L}_{mass} = -\frac{v^2}{4}g^2 W^+_\mu W^{-\mu} - (g^2+g'^2)\frac{v^2}{8}Z^0_\mu Z^{0\mu}. \tag{30.46}
$$

Therefore we see that: $W^\pm_\mu$ have mass

$$
m_W = \frac{gv}{2}\ , \tag{30.47}
$$

the Z^0_μ has mass

$$
m_Z = \sqrt{g^2+g'^2}\frac{v}{2}\ , \tag{30.48}
$$

and the photon A_μ is massless, and corresponds to the unbroken electromagnetism.

Note that the photon is unbroken, since the corresponding field has no mass term in the Lagrangean, and $T^a\phi_0 = 0$ for it.

We now introduce the *weak mixing angle* or *Weinberg angle* θ_W, defined by

$$
\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}},\quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}. \tag{30.49}
$$

We see that in terms of it, the Z and A are just a rotation of A^3 and B, i.e.

$$
\begin{pmatrix} Z^0_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} A^3_\mu \\ B_\mu \end{pmatrix}\ , \tag{30.50}
$$

which is why we defined Z^0_μ and A_μ like this, in order to have an orthogonal rotation.

Using the further definitions

$$
\begin{aligned}
T^\pm &= T^1 \pm iT^2 = \sigma^\pm\\
e &= \frac{gg'}{\sqrt{g^2+g'^2}}\\
Q &= T^3 + Y\ ,
\end{aligned}
\tag{30.51}
$$

the covariant derivative is

$$
\begin{aligned}
D_\mu &= \partial_\mu - igA^a_\mu T_a - ig'YB_\mu\\
&= \partial_\mu - i\frac{g}{\sqrt{2}}(W^+_\mu T^+ + W^-_\mu T^-) - \frac{i}{\sqrt{g^2+g'^2}}Z_\mu(g^2T^3 - g'^2 Y) - \frac{igg'}{\sqrt{g^2+g'^2}}A_\mu(T^3+Y)\ ,
\end{aligned}
\tag{30.52}
$$

and since

$$
g^2T^3 - g'^2Y = (g^2+g'^2)T^3 - g'^2 Q\ , \tag{30.53}
$$

and $e = g \sin\theta_W$, we get

$$D_\mu = \partial_\mu - i\frac{g}{\sqrt{2}}(W_\mu^+ T^+ + W_\mu^- T^-) - i\frac{g}{\cos\theta_W} Z_\mu (T^3 - \sin^2\theta_W Q) - ieA_\mu Q. \tag{30.54}$$

We also obtain the relation between parameters

$$m_W = m_Z \cos\theta_W. \tag{30.55}$$

In unitary gauge, we write

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{30.56}$$

and then $H(x)$ is the *Higgs boson*, i.e. the field associated with the Higgs particle.

Expanding the potential in terms of it, we obtain

$$V(\phi) = \frac{\lambda v^2}{4} H^2 + \frac{\lambda v}{4} H^3 + \frac{\lambda}{16} H^4 \ , \tag{30.57}$$

so in particular

$$m_H^2 = \frac{\lambda v^2}{2}. \tag{30.58}$$

Then, the electroweak bosonic Lagrangean

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu}^a)^2 - \frac{1}{4} B_{\mu\nu}^2 - \frac{1}{2}|D_\mu \phi|^2 - V(\phi) \ , \tag{30.59}$$

where $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$, becomes (the proof is left as an exercise)

$$\begin{aligned} \mathcal{L} \;=\; & -\frac{1}{4}F_{\mu\nu}^2 - \frac{1}{4}Z_{\mu\nu}^2 - \tilde{D}^{\dagger\mu} W^{-\nu} \tilde{D}_\mu W_\nu^+ + \tilde{D}^{\dagger\mu} W^{-\nu} \tilde{D}_\nu W_\mu^+ \\ & + ie(F^{\mu\nu} + \cot\theta_W Z^{\mu\nu}) W_\mu^+ W_\nu^- - \frac{e^2/\sin^2\theta_W}{2}(W^{+\mu} W_\mu^- W^{+\nu} W_\nu^- - W^{+\mu} W_\mu^+ W^{-\nu} W_\nu^-) \\ & - \left(M_W^2 W^{+\mu} W_\mu^- + \frac{M_Z^2}{2} Z^\mu Z_\mu \right) \left(1 + \frac{M}{v} \right)^2 \\ & - \frac{1}{2}(\partial_\mu H)^2 - \frac{m_H^2}{2} H^2 - \frac{m_H^2}{2v} H^3 - \frac{m_H^2}{8v^2} H^4 \ , \end{aligned} \tag{30.60}$$

where $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$ and

$$\tilde{D}_\mu = \partial_\mu - ie(A_\mu + \cot\theta_W Z_\mu). \tag{30.61}$$

Important concepts to remember

- In spontaneous symmetry breaking, a local gauge group is (partially) broken ("spontaneously") by the choice of a vacuum.

- The vectors corresponding to the broken symmetry become massive by eating the massless scalars ("Goldstone bosons") corresponding to the directions of the broken symmetry in scalar field space.

- The Goldstone bosons make the vacuum polarization of the spontaneously broken theory transverse.

- The gauge symmetry is lost after the redefinition of the vector that eats the scalar.

- In the nonabelian case, the gauge group G is broken to H, leaving G/H broken generators that give the Goldstone bosons, and H unbroken generators of zero mass, with $T^a\phi_0 = 0$.

- In the electroweak theory, the $SU(2) \times U(1)_Y$ is broken to $U(1)_{em}$, and the Z and A are an orthogonal rotation of A^3 and B.

Further reading: See chapter 20.1in [3] and chapter 85,86,87 in Srednicki.

Exercises, Lecture 30

1) Consider the abelian-Higgs Lagrangean. Expand it up to 4th order in the perturbations around the Higgs vacuum.

2) Prove that the Standard Model (electroweak) bosonic term around the Higgs vacuum takes the form

$$\begin{aligned}
\mathcal{L} \;=\; & -\frac{1}{4}F_{\mu\nu}^2 - \frac{1}{4}Z_{\mu\nu}^2 - \tilde{D}^{\dagger\mu}W^{-\nu}\tilde{D}_\mu W_\nu^+ + \tilde{D}^{\dagger\mu}W^{-\nu}\tilde{D}_\nu W_\mu^+ \\
& + ie(F^{\mu\nu} + \cot\theta_W Z^{\mu\nu})W_\mu^+ W_\nu^- - \frac{e^2/\sin^2\theta_W}{2}(W^{+\mu}W_\mu^- W^{+\nu}W_\nu^- - W^{+\mu}W_\mu^+ W^{-\nu}W_\nu^-) \\
& - \left(M_W^2 W^{+\mu}W_\mu^- + \frac{M_Z^2}{2}Z^\mu Z_\mu\right)\left(1 + \frac{M}{v}\right)^2 \\
& - \frac{1}{2}(\partial_\mu H)^2 - \frac{m_H^2}{2}H^2 - \frac{m_H^2}{2v}H^3 - \frac{m_H^2}{8v^2}H^4 \,, \qquad (30.62)
\end{aligned}$$

where $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$ and

$$\tilde{D}_\mu = \partial_\mu - ie(A_\mu + \cot\theta_W Z_\mu). \qquad (30.63)$$

31 Lecture 31. Renormalization of spontaneously broken gauge theories 1: the Goldstone theorem and R_ξ gauges

In this lecture we will start to describe the quantization of spontaneously broken gauge theories, describing first the Goldstone theorem (at the quantum level), and then describing a gauge that is most useful for quantization of spontaneously broken gauge theories, the R_ξ gauge.

The Goldstone theorem

We already mentioned the Goldstone theorem. It states that massless states called Goldstone bosons appear whenever we break spontaneously a continuous symmetry.

Proof.

Consider a general Lagrangean composed of a kinetic part depending on $\partial_\mu \phi$ and a potential, i.e.

$$\mathcal{L} = K(\partial\phi) - V(\phi). \tag{31.1}$$

Consider the minimum of the potential, at

$$\frac{\partial}{\partial\phi^a} V\Big|_{\phi^a(x)=\phi^a_0} = 0 \ , \tag{31.2}$$

and expand the potential around it,

$$V(\phi) = V(\phi_0) + \frac{1}{2}(\phi-\phi_0)^a(\phi-\phi_0)^b\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V\right)_{\phi_0} + ... \ , \tag{31.3}$$

where

$$\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V\right)_{\phi_0} \equiv m^2_{ab} \tag{31.4}$$

is a mass matrix. Consider then the continuous symmetry

$$\phi^a \to \phi^a + \alpha\Delta^a(\phi). \tag{31.5}$$

In the particular case of constant ϕ^a, the invariance of the Lagrangean implies the invariance of the potential, i.e.

$$V(\phi^a) = V(\phi^a + \alpha\Delta^a(\phi)) \ , \tag{31.6}$$

i.e. that

$$\Delta^a(\phi)\ \frac{\partial}{\partial\phi^a}V\Big|_{\phi_0} = 0. \tag{31.7}$$

Taking a derivative $\partial/\partial\phi^b$ on the above, we get

$$0 = \left(\frac{\partial\Delta^a}{\partial\phi^b}\right)_{\phi_0}\left(\frac{\partial V}{\partial\phi^a}\right)_{\phi_0} + \Delta^a(\phi_0)\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V\right)_{\phi_0} \Rightarrow \Delta^a(\phi_0)\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V\right)_{\phi_0} = 0 \ , \tag{31.8}$$

since ϕ_0 is a minimum of V, so the first term is zero.

Then there are two possibilities:

1. $\Delta^a(\phi_0) = 0$, which means that the symmetry leaves the vacuum ϕ_0^a unchanged, which is not the situation we are interested in.

2. If there is spontaneous breaking of the symmetry, then $\Delta^a(\phi_0) \neq 0$, which means that

$$(\Delta^a(\phi_0)\cdot)\left(\frac{\partial^2}{\partial\phi^a\partial\phi^b}V\right)_{\phi_0} \equiv m_{ab}^2(\cdot\Delta^a(\phi_0)) = 0\ , \tag{31.9}$$

i.e. there is a zero eigenvalue for the mass in the direction of the symmetry, which is the Goldstone boson.

q.e.d.

But we actually have proven the theorem only at the classical level. It is more interesting to prove that quantum corrections also don't spoil this property.

In order to consider quantum corrections, we need to consider instead of the classical action, the quantum effective action Γ. But for constant fields ϕ^a, the effective action turns into the *effective potential*, or rather the effective potential times the volume of spacetime,

$$\Gamma[\phi_{\rm cl}] = -(VT)V_{\rm eff}[\phi_{\rm cl}]. \tag{31.10}$$

But if there are no quantum anomalies, the effective potential respects the same symmetries as V, so we can repeat the same argument for the effective potential.

Then we obtain in general

$$\left|\frac{\delta^2}{\delta\phi^a\delta\phi^b}\Gamma(p)\right|_{p^2\neq m^2} = 0. \tag{31.11}$$

If we consider the particular case of $p = 0$, corresponding to constant classical field, Γ turns to $V_{\rm eff}$, and we get

$$\frac{\partial^2}{\partial\phi^a\partial\phi^b}V_{\rm eff} = 0\ , \tag{31.12}$$

as expected.

R_ξ **gauges**.

Abelian case

We start with the abelian case.

In the case of gauge theory not spontaneously broken, the Minkowski Lagrangean is

$$\mathcal{L}_{g.f.} + \mathcal{L}_{gh} = -\frac{1}{2\xi}G^2 - b\frac{\delta G}{\delta\alpha}c\ , \tag{31.13}$$

where $G = 0$ is the gauge condition. The gauge transformation is

$$\delta A_\mu = -\partial_\mu\alpha; \quad \delta\phi = -ie\alpha\phi\ , \tag{31.14}$$

and in the Lorenz (covariant) gauge $G = \partial^\mu A_\mu$, we obtain

$$\frac{\delta G}{\delta\alpha} = -\partial^2. \tag{31.15}$$

In the spontaneously broken case, with scalar field expansion, we have

$$\phi = \frac{1}{\sqrt{2}}(\phi^1 + i\phi^2) = \frac{1}{\sqrt{2}}(v + h(x) + i\varphi(x)) \; , \tag{31.16}$$

where $h(x)$ is the Higgs, and $\varphi(x)$ is the Goldstone boson.

We then choose

$$G = \partial^\mu A_\mu - \xi e v \varphi \; , \tag{31.17}$$

which is called the R_ξ gauge, and we note that for $v = 0$ it reduces to the Lorenz gauge.

We compute the kinetic term for ϕ. With $D_\mu = \partial_\mu - ieA_\mu$ as usual, we get

$$D_\mu \phi = \frac{1}{\sqrt{2}}[(\partial_\mu h + e\varphi A_\mu) + i(\partial_\mu \varphi - e(v+h)A_\mu)] \; , \tag{31.18}$$

so that

$$\begin{aligned} -|D_\mu \phi|^2 &= -\frac{1}{2}(\partial_\mu h + e\varphi A_\mu)^2 - \frac{1}{2}(\partial_\mu \varphi - e(v+h)A_\mu)^2 \\ &= -\frac{1}{2}(\partial_\mu h)\partial^\mu h - \frac{1}{2}\partial_\mu \varphi \partial^\mu \varphi - \frac{e^2}{2} v^2 A_\mu A^\mu + ev A_\mu \partial^\mu \varphi \\ &\quad + eA_\mu (h \partial^\mu \varphi - \varphi \partial^\mu h) \\ &\quad - e^2 v h A_\mu A^\mu - \frac{e^2}{2}(h^2 + \varphi^2) A_\mu A^\mu. \end{aligned} \tag{31.19}$$

The potential expands as

$$V = \frac{\lambda}{2}\left(|\phi|^2 - \frac{v^2}{2}\right)^2 = \frac{\lambda v^2}{2} h^2 + \frac{\lambda v}{2} h(h^2 + \varphi^2) + \frac{\lambda}{8}(h^2 + \varphi^2)^2. \tag{31.20}$$

The gauge fixing term expands as

$$\begin{aligned} \int d^d x \mathcal{L}_{g.f.} &= \int d^d x \left[-\frac{1}{2\xi}(\partial^\mu A_\mu)\partial^\nu A_\nu - \frac{\xi e^2 v^2}{2}\varphi^2 + ev\varphi \partial^\mu A_\mu \right] \\ &= \int d^d x \left[-\frac{1}{2\xi}(\partial^\mu A_\nu)\partial^\nu A_\mu - ev A_\mu \partial^\mu \varphi - \frac{\xi e^2 v^2}{2}\varphi^2 \right] \; , \end{aligned} \tag{31.21}$$

where in the first term we have partially integrated both derivatives, and in the mixed term we have partially integrated the derivative. From the mass term for φ we have the mass

$$m_\varphi = \sqrt{\xi} e v = \sqrt{\xi} m_A. \tag{31.22}$$

We see that the mixed term cancels between the kinetic term and the gauge fixing term, which is one reason why the gauge fixing term was chosen like that.

The ghost term is found as follows. The gauge transformation around the Higgs vacuum with $\phi = (v + h + \varphi)/\sqrt{2}$ is

$$\delta A_\mu = -\partial_\mu \alpha; \quad \delta h = +e\alpha\varphi; \quad \delta\varphi = -e\alpha(v+h). \tag{31.23}$$

Then we get

$$\frac{\delta G}{\delta\alpha}=\frac{\delta}{\delta\alpha}(\partial^\mu A_\mu-\xi ev\varphi)=-\partial^2+\xi e^2v(v+h)\ , \tag{31.24}$$

leading to a ghost term

$$\begin{aligned}\int d^dx\mathcal{L}_{gh} &= -\int d^dx b[-\partial^2+\xi e^2 v(v+h)]c\\ &= -\int d^dx[\partial^\mu b\partial_\mu c+\xi e^2v^2bc+\xi e^2 vhbc].\end{aligned} \tag{31.25}$$

From the second term we see that the ghosts b and c have masses

$$m_{b,c}=\sqrt{\xi}ev=\sqrt{\xi}m_A \tag{31.26}$$

also.

The kinetic term for A_μ contains the usual kinetic term from the $-\int F_{\mu\nu}^2/4$ action, the term from the gauge fixing term, and the mass term from the scalar kinetic term, giving in total

$$\mathcal{L}_{A^2}=-\frac{1}{2}A_\mu\left[\partial^\mu\partial^\nu-g^{\mu\nu}\partial^2+m_A^2g^{\mu\nu}-\frac{\partial^\mu\partial^\nu}{\xi}\right]A_\nu\ , \tag{31.27}$$

or in momentum space

$$\mathcal{L}_{A^2}=-\frac{1}{2}\tilde{A}_\mu(-k)[(k^2+m_A^2)g^{\mu\nu}-(1-\xi^{-1})k^\mu k^\nu]\tilde{A}_\nu(k). \tag{31.28}$$

Then the kinetic matrix

$$[(k^2+m_A^2)g^{\mu\nu}-(1-\xi^{-1})k^\mu k^\nu] \tag{31.29}$$

can be written in terms of the projectors

$$P^{\mu\nu}(k)=g^{\mu\nu}-\frac{k^\mu k^\nu}{k^2};\quad K^{\mu\nu}(k)=\frac{k^\mu k^\nu}{k^2}\ , \tag{31.30}$$

satisfying the projector, orthogonality and completeness relations, i.e. schematically

$$P^2=P;\quad K^2=K;\quad P\cdot K=0;\quad P+K=1\ , \tag{31.31}$$

which can be easily checked, as

$$(k^2+m_A^2)P^{\mu\nu}(k)+\xi^{-1}(k^2+\xi m_A^2)K^{\mu\nu}(k)\ , \tag{31.32}$$

which means it can be easily inverted to give the photon propagator in R_ξ gauge,

$$\tilde{\Delta}_{\mu\nu}(k)=\frac{P_{\mu\nu}(k)}{k^2+m_A^2}+\frac{\xi K_{\mu\nu}(k)}{k^2+\xi m_A^2}. \tag{31.33}$$

We note that the transverse (physical) part of the propagator, proportional to $P_{\mu\nu}(k)$, has mass m_A, whereas the longitudinal (unphysical) part of the propagator, proportional to $K_{\mu\nu}(k)$, has mass $\sqrt{\xi}m_A$. So in total, the quartet of unphysical states, the would-be

Goldstone boson φ, the ghosts b and c and the longitudinal photon, all have the mass $\sqrt{\xi}m_A$.

We also note that for $\xi = 1$ (the equivalent of the Feynman gauge for the unbroken theory), we have

$$\tilde{\Delta}_{\mu\nu}(k)|_{\xi=1} = \frac{g_{\mu\nu}}{k^2 + m_A^2}\ , \tag{31.34}$$

which is the KG propagator (for a scalar mode). We will nevertheless continue to use arbitrary ξ in order to see ξ independence for physical quantities (test our calculation).

In conclusion, the propagators are as follows:

-for the Higgs h we have the usual scalar propagator,

$$\frac{1}{k^2 + m_h^2}\ , \tag{31.35}$$

where $m_h = \sqrt{\lambda}v$.

-for the would-be Goldstone boson (unphysical scalar φ), we have the scalar propagator

$$\frac{1}{k^2 + \xi m_A^2}\ , \tag{31.36}$$

-same as the ghosts b, c.

-the vector has the propagator $\tilde{\Delta}_{\mu\nu}(k)$, with $m_A = ev$.

Putting together all the interaction terms derived above, we find the interaction Lagrangean

$$\begin{aligned}
\mathcal{L}_{\text{int}} &= -\frac{\lambda v}{2}h(h^2+\varphi^2) - \frac{\lambda}{8}(h^2+\varphi^2)^2 \\
&\quad + eA_\mu(h\partial^\mu\varphi - \varphi\partial^\mu h) \\
&\quad - e^2 vhA_\mu A^\mu - \frac{e^2}{2}(h^2+\varphi^2)A_\mu A^\mu \\
&\quad - \xi e^2 vhbc.
\end{aligned} \tag{31.37}$$

Nonabelian case

In the nonabelian case, the gauge fixing and ghost terms in the Lagrangean are

$$\mathcal{L}_{g.f.} + \mathcal{L}_{gh} = -\frac{1}{2\xi}G_a G^a - b^a \frac{\delta G^a}{\delta \alpha^b} c^b. \tag{31.38}$$

The gauge covariant derivative is

$$D_\mu = \partial_\mu + gA_\mu^a T_a \tag{31.39}$$

and we expand around the VEV $\langle \phi_i \rangle = v_i$ as

$$\phi_i = v_i + \chi_i. \tag{31.40}$$

Then we choose the gauge condition for the R_ξ gauge as

$$G^a = \partial^\mu A_\mu^a - \xi g (T^a)_{ij} v_j \chi_i. \tag{31.41}$$

Then the gauge fixing term becomes

$$\begin{aligned}\int d^dx\mathcal{L}_{g.f.} &= \int d^dx\left[-\frac{1}{2\xi}(\partial^\mu A_\mu^a)\partial^\nu A_\nu^a + \xi g(T^a)_{ij}v_j\chi_i\partial^\mu A_\mu^a - \xi g^2((T^a)_{ik}v_k(T^a)_{jl}v_l)\chi_i\chi_j\right]\\ &= \int d^dx\left[-\frac{1}{2\xi}(\partial^\mu A_\nu^a)\partial^\nu A_\mu^a - g(T^a)_{ij}v_j\partial^\mu\chi_i A_\mu^a - \xi g^2((T^a)_{ik}v_k(T^a)_{jl}v_l)\chi_i\chi_j\right],\end{aligned} \tag{31.42}$$

where as before we have partially integrated the two derivatives in the first term and the derivative in the second. The last term, with $\chi_i\chi_j$, is a contribution to the mass term for scalars,

$$\xi M_{ij}^2 = \xi g^2((T^a)_{ik}v_k(T^a)_{jl}v_l). \tag{31.43}$$

The kinetic term for the scalars is ($(T^a)^\dagger = -T^a$)

$$\begin{aligned}-\frac{1}{2}(D^\mu\phi_i)^\dagger D^\mu\phi_i &= -\frac{1}{2}(\partial^\mu\chi_i)\partial_\mu\chi_i - \frac{g^2}{2}((T^a)_{ik}v_k(T^b)_{il}v_l)A_\mu^aA^{b\mu} + g(T^a)_{ik}v_kA_\mu^a\partial^\mu\chi_i\\ &+gA_\mu^a\chi_i(T^a)_{ij}\partial^\mu\chi_j - g^2A_\mu^aA^{b\mu}((T^a)_{ik}v_k)(T^b)_{ij}\chi_j - \frac{g^2}{2}(T^aT^b)_{ij}\chi_i\chi_jA_\mu^aA^{b\mu},\end{aligned} \tag{31.44}$$

and again the mixing term between the vector and the would-be Goldstone boson cancels with the gauge fixing term.

The ghost term is obtained by considering the gauge transformation on the fluctuation,

$$A_\mu^a \to A_\mu^a - D_\mu^{ab}\alpha^b; \quad \chi_i \to -g\alpha^a(T_a)_{ij}(v+h)_j\ , \tag{31.45}$$

leading to

$$\begin{aligned}\frac{\delta G^a}{\delta\alpha^b} &= -\partial^\mu D_\mu^{ab} + \xi g^2(T^a)_{ij}v_j(T^a)_{il}(v+\chi)_l\\ &= -\partial^\mu D_\mu^{ab} + \xi g^2(T^a)_{ij}v_j(T^b)_{il}v_l + \xi g^2((T^a)_{ij}v_j(T^b)_{il})\chi_l\ ,\end{aligned} \tag{31.46}$$

where the middle term is written as $(M_{b,c}^2)^{ab}$ and is the ghost mass term. Then the ghost term is

$$\int d^dx\mathcal{L}_{gh} = \int d^dx\left[-(\partial_\mu b^a)D_\mu^{ab}c^b - \xi(M_{b,c}^2)^{ab}b^ac^b - \xi g^2((T^a)_{ij}v_j(T^a)_{il})\chi_l b^ac^b\right]. \tag{31.47}$$

Important concepts to remember

- The Goldstone theorem says that there is a massless particle (Goldstone boson) for every spontaneously broken continuous symmetry.
- The effective potential is the effective action on constant fields, more precisely $\Gamma[\phi_{cl}] = -(VT)V_{eff}[\phi_{cl}]$.

- Quantum corrections respect the Goldstone theorem.
- The R_ξ gauge is $\partial^\mu A_\mu - \xi e v \varphi = 0$.
- In the R_ξ gauge, the quartet of unphysical states, would-be Goldstone boson φ, ghosts b and c and longitudinal gauge boson have all the mass $\sqrt{\xi} m_A$, where $m_A = ev$; the Higgs mass is $m_h = \sqrt{\lambda} v$.
- With the choice $\xi = 1$, the photon propagator is the KG propagator.

Further reading: See chapter 11.1, 21.1, 21.2 in [3].

Exercises, Lecture 31

1) Consider a theory invariant under a symmetry group G, having a spontaneously breaking vacuum invariant under $H \sup G$. How many Goldstone bosons there are? Specialize to the $SU(5) \to SU(3) \times SU(2) \times U(1)$ breaking.

2) Write down all the one-loop Feynman diagrams for the Higgs h 1PI 2-point functionin the spontaneously broken abelian theory in R_ξ gauge, and the integral expressions for them using the Feynman rules (without computing them).

32 Lecture 32. Renormalization of spontaneously broken gauge theories II: The $SU(2)$-Higgs model

In this lecture we learn how to renormalize spontaneously broken gauge theories using the R_ξ gauge, for the example of the $SU(2)$-Higgs system, which is close to the electroweak theory, without being exactly that. Also, we will not do the full renormalization, but only some of the important steps.

The theory contains $SU(2)$ gauge fields, and a complex Higgs doublet. Again, the terminology is ambigous, really we have a Higgs field h and would-be Goldstone bosons χ^a.

With respect to unbroken gauge theories, one important difference is that now there is one more Ward identity, for the fact that only the combination $v+h$ appears in the classical Lagrangean. It is of course broken by the gauge fixing term, so we need to check explicitly that this is still satisfied at the quantum level.

We parametrize the complex field slightly differently from what we had before, now considering the VEV in the upper component, instead of the lower component. So we start by parametrizing the complex doublet

$$\phi = \begin{pmatrix} \varphi^1 \\ \varphi^2 \end{pmatrix} \tag{32.1}$$

as

$$\phi = \frac{1}{\sqrt{2}}(\psi + i\chi^a\sigma^a)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \psi + i\chi^3 \\ i\chi^1 - \chi^2 \end{pmatrix}. \tag{32.2}$$

The Lagrangean is

$$\mathcal{L} = -\frac{1}{4}(F^a_{\mu\nu})^2 - (D_\mu\phi)^\dagger D^\mu\phi - V(\phi^\dagger\phi)\ , \tag{32.3}$$

where the potential is

$$V = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{32.4}$$

The VEV is

$$\langle \text{Re}\phi^1\rangle = \frac{v}{\sqrt{2}}\ , \tag{32.5}$$

so we split the scalar into VEV and fluctuations as

$$\psi = v + h. \tag{32.6}$$

The gauge-covariant derivative is

$$D_\mu\phi = \partial_\mu\phi - \frac{i}{2}gA^a_\mu\sigma^a\phi. \tag{32.7}$$

This is a particular case for the general procedure from last lecture, so we can write the Lagrangean that comes from the square of the covariant derivative as a sum of a kinetic piece, a piece linear in A and a piece quadratic in A, as

$$\mathcal{L}_{(kin)} = -\frac{1}{2}[(\partial^\mu h)\partial_\mu h + (\partial_\mu\chi^a)\partial^\mu\chi^a]$$

$$
\begin{aligned}
\mathcal{L}(A^1) &= \frac{gv}{2}(A^a_\mu \partial^\mu \chi^a) + \frac{g}{2} A^a_\mu (h \overset{\leftrightarrow}{\partial}_\mu \chi^a) + \frac{1}{2} g \epsilon_{abc}(\chi^a A^b_\mu \partial^\mu \chi^c) \\
\mathcal{L}(A^2) &= -\frac{g^2}{8} A^a_\mu A^{a\mu}[(v+h)^2 + \chi^b \chi^b].
\end{aligned} \tag{32.8}
$$

Writing the covariant derivative as

$$
D_\mu \phi = \frac{1}{\sqrt{2}}(D_\mu h + i\sigma^a D_\mu \chi^a) \;, \tag{32.9}
$$

where we have defined

$$
\begin{aligned}
D_\mu h &= \partial_\mu h + \frac{g}{2} A^a_\mu \chi_a \\
D_\mu \chi^a &= \partial_\mu \chi^a + \frac{g}{2}\epsilon_{abc} A^b_\mu \chi^c - \frac{g}{2} A^a_\mu (v+h) \;,
\end{aligned} \tag{32.10}
$$

we can rewrite the scalar kinetic term as

$$
\mathcal{L} = -\frac{1}{2}(D_\mu h)^2 - \frac{1}{2}(D_\mu \chi^a)^2 + \frac{g}{2} v A^a_\mu \partial^\mu \chi^a. \tag{32.11}
$$

To cancel the last term, we add the 't Hooft gauge fixing term (in R_ξ gauge)

$$
\mathcal{L}_{g.fix} = -\frac{1}{2\xi}\left(\partial^\mu A^a_\mu + \frac{1}{2}\xi g v \chi^a\right)^2 . \tag{32.12}
$$

Then the quadratic (kinetic) term for χ is

$$
\mathcal{L}_{\chi^2} = -\frac{1}{2}(\partial_\mu \chi)^2 - \frac{\xi}{2}\left(\frac{gv}{2}\right)^2 (\chi^a)^2 \;, \tag{32.13}
$$

so the mass of the would-be Goldstone bosons is $m_\chi = \sqrt{\xi} m_A$.

The kinetic term for the vectors is now

$$
\mathcal{L}_{YM} = -\frac{1}{4}(F^a_{\mu\nu})^2 - \frac{1}{2}\left(\frac{gv}{2}\right)^2 (A^a_\mu)^2 \;, \tag{32.14}
$$

so $m_A = gv/2$.

The ghost Lagrangean is found as usual, from $b_a \delta G^a / \delta\alpha^b c^b$, as

$$
\mathcal{L}_{gh} = b_a \partial^\mu D_\mu c^a - \xi\left(\frac{gv}{2}\right)^2 b_a c^a - \frac{\xi}{4} g^2 v b_a (h c^a + \epsilon^a{}_{bc} \chi^b c^c). \tag{32.15}
$$

As we mentioned last lecture, we see that the quartet of unphysical states, the Fadeev-Popov ghosts b_a and c^a, together with the would-be Goldstone bosons χ^a and the longitudinal part of the YM field A^a_μ have all the same mass $m = \sqrt{\xi} m_A$.

We will see why shortly, but at the quantum level we need to consider that the gauge fixing term contains two new parameters, so we will write it as

$$
\mathcal{L}_{g.fix} = -\frac{1}{2\alpha}\left(\partial^\mu A^a_\mu + \frac{1}{2}\xi g v \chi^a\right)^2 \;, \tag{32.16}
$$

where now $\alpha \neq \xi$ and are equal only at the classical level. But even in this case (at the quantum level) the quartet of unphysical states will still have the same mass.

We now split the Lagrangean for the spontaneously broken theory into matter, gauge and ghost parts,

$$\begin{aligned}
\mathcal{L}_{\text{gauge}} &= -\frac{1}{4}\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a{}_{bc} A^b_\mu A^c_\nu\right)^2 \\
\mathcal{L}_{\text{matter}} &= -\frac{1}{2}(D_\mu h)^2 - \frac{1}{2}(D_\mu \chi^a)^2 + \frac{\mu^2}{2}[(v+h)^2 + (\chi^a)^2] - \frac{\lambda}{4}[(v+h)^2 + (\chi^a)^2]^2 \\
\mathcal{L}_{\text{ghost}} &= b_a\left[\partial^\mu D_\mu c^a - \xi \frac{gv}{2}\left(\frac{1}{2}g(v+h)c^a + \frac{1}{2} g f^a{}_{bc}\chi^b c^c\right)\right]. \qquad (32.17)
\end{aligned}$$

But we still need to add a term to the Lagrangean. We saw in the unbroken case, that when we renormalize, we need to add an extra source term to the Lagrangean for the nonlinear parts of the BRST variations. We add

$$\begin{aligned}
\mathcal{L}_{\text{extra}} &= K^\mu_a Q_B A^a_\mu/\Lambda + K Q_B h/\Lambda + K_a Q_B \chi^a/\Lambda + L_a Q_B c^a/\Lambda \\
&= K^\mu_a D_\mu c^a - K\left(\frac{1}{2} g \chi_a c^a\right) + K_a\left(\frac{1}{2}g(v+h)c^a + \frac{1}{2} g f^a{}_{bc}\chi^b c^c\right) \\
&\quad + L_a\left(\frac{1}{2} g f^a{}_{bc} c^b c^c\right). \qquad (32.18)
\end{aligned}$$

It is useful to write the linear and quadratic mass terms in the scalars as

$$-\beta v h - \frac{\beta}{2}(h^2 + (\chi^a)^2)\ , \qquad (32.19)$$

where

$$\beta = -\mu^2 + 2\lambda v^2\ , \qquad (32.20)$$

though then β doesn't renormalize multiplicatively. Indeed, μ^2 and λv^2 both renormalize multiplicatively, however classically $\beta = 0$, so we require

$$\beta^{(0)}_{ren} = -\mu^2_{ren} + \lambda_{ren} v^2_{ren} = 0\ , \qquad (32.21)$$

which means that β is renormalized additively, as

$$\beta_{ren} = 0 + \Delta\beta^{(1)}_{ren} + \ldots \qquad (32.22)$$

We will therefore choose to renormalize β additively as above, rather than renormalize μ^2 multiplicatively.

Another observation is that now the matter Lagrangean depends on $v + h(x)$, but the gauge fixing and ghost terms break this, so at the quantum level we must check explicitly. (Note that we could replace v by $v + h$ in the gauge fixing term, but then it is more complicated).

We also consider that we will have $\alpha \neq \xi$, but we require $\alpha_{ren} = \xi_{ren} = 1$.

As in the unbroken case, we proceed to write the equivalent of the Ward identities for the effective action Γ in the BRST case, more precisely for

$$\hat{\Gamma} = \Gamma - \int \mathcal{L}_{fix} d^4x \ , \tag{32.23}$$

the Lee-Zinn-Justin identities, that are now written as

$$\begin{aligned}
&\int d^4x \left[\partial\hat{\Gamma}/\partial\phi^I \frac{\partial}{\partial K^I} \right] \hat{\Gamma} = 0 \\
&\left(\partial^\mu \frac{\partial}{\partial K_a^\mu} - \xi \frac{gv}{2} \frac{\partial}{\partial K_a} - \frac{\partial}{\partial b_a} \right) \hat{\Gamma} = 0.
\end{aligned} \tag{32.24}$$

Here

$$\begin{aligned}
\phi^I &= \{h, \chi^a, A_\mu^a, c^a\}; \\
K_I &= \{K, K_a, K_a^\mu, L_a\}.
\end{aligned} \tag{32.25}$$

These can be found as in the unbroken case, and we note that restricting to A_μ^a and c^a and putting $v = 0$, we find the unbroken case.

Therefore the fields in the theory are A_μ^a, b_a, c^a, h and χ^a, the sources are K, K^a, K_μ^a and L_a, and the parameters are g, v, λ, α and ξ.

Renormalization is then done as follows. The fields renormalize as

$$\begin{aligned}
A_\mu^a &= \sqrt{Z_3} A_{\mu,ren}^a \\
c^a &= \sqrt{Z_{gh}} c_{ren}^a \\
b_a &= \sqrt{Z_{gh}} b_{a,ren} \\
h &= \sqrt{Z_h} h_{ren} \\
\chi^a &= \sqrt{Z_\chi} \chi_{ren}^a \ ,
\end{aligned} \tag{32.26}$$

the sources as

$$\begin{aligned}
K &= \sqrt{\frac{Z_3 Z_{gh}}{Z_h}} K_{ren} \\
K^a &= \sqrt{\frac{Z_3 Z_{gh}}{Z_\chi}} K_{ren}^a \\
K_\mu^a &= \sqrt{Z_{gh}} K_{\mu,ren}^a \\
L^a &= \sqrt{Z_3} L_{a,ren} \ ,
\end{aligned} \tag{32.27}$$

and the parameters as

$$\begin{aligned}
g &= Z_g g_{ren} \mu^{\frac{4-d}{2}} \\
v &= \sqrt{Z_v} v_{ren} \\
\lambda &= Z_\lambda Z_h^{-2} \lambda_{ren} \mu^{4-d} \\
\alpha &= Z_3 \alpha_{ren}
\end{aligned}$$

$$\xi = \sqrt{\frac{Z_3}{Z_v Z_\chi}} Z_g^{-1} \xi_{ren}. \tag{32.28}$$

We must make several observations:

1. b_a, as well as K_a^μ, scale as c^a and L_a scales as A_μ^a. We have seen in the unbroken case that their renormalization is fixed by analyzing the possible divergent structures that solve the Lee-Zinn-Justin identities, and now a similar story holds.

But now we also see that Kh, $K_a\chi^a$, $K_a^\mu A_\mu^a$ and $L_a c^a$ all scale the same way, since the source terms are $K_\mu^a Q_B A_\mu^a$, $KQ_B h$, $K^a Q_B \chi^a$, $L_a Q_B c^a$.

Therefore the renormalization of the extra source terms is completely fixed.

2. The renormalization of α and ξ is fixed by requiring that the gauge fixing term is finite by itself. The part at $v = 0$ (from the unbroken theory) fixes in $\alpha = Z_\alpha \alpha_{ren}$ $Z_\alpha = Z_3$, whereas the v-dependent term fixes the renormalization of ξ. The result is a different renormalization factor for α and ξ.

3. This in turn means that we need separate α and ξ as advertised, and $\alpha_{ren} = \xi_{ren} = 1$.

4. A quick one-loop calculation shows that in fact also $Z_v \neq Z_h$, even though the classical Lagrangean has them equal.

One can prove renormalizability at all loops by induction. Here we will not prove the induction step, since we also didn't in the unbroken case.

5. By analyzing the solution of the Lee-Zinn-Justin identities, we find 9 possible divergent structures, but we have only 8 renormalization parameters: $Z_3, Z_{gh}, Z_h, Z_\chi, Z_g, Z_v, Z_\lambda$ and $\Delta\beta_{ren}$ (standing in for Z_μ).

This would seem like a contradiction, but as we advertised at the beginning of the lecture, there is an extra Ward identitty, coming from the fact that we only find the $v+h$ combination in the classical part (only the gauge fixing and ghost parts break it)

The classical action is the tree level part of the modified effective action, $\hat{\Gamma}^{(0)}_{ren} = S_{ren}$. Then the Ward identity is written as

$$\left(v\frac{\partial}{\partial h} - v\frac{\partial}{\partial v}\right) S_{\text{matter}} = 0 \ , \tag{32.29}$$

or in terms of $\hat{\Gamma}^{(0)}$ as

$$\left(\xi_{ren}\frac{\partial}{\partial \xi_{ren}} - v_{ren}\frac{\partial}{\partial v_{ren}} + v_{ren}\frac{\partial}{\partial h_{ren}}\right) \hat{\Gamma}^{(0)}_{ren} = 0 \ , \tag{32.30}$$

since in the ghost term and the gauge fixing term the combination ξv appears.

The above Ward identity allows the reduction of the possible divergences to 8, equal to the number of renormalization parameters.

Another useful consistency condition is found from ghost number conservation, which implies (the coefficient of each term is their ghost number)

$$\left(b_a\frac{\partial}{\partial b^a} - c^a\frac{\partial}{\partial c^a} + K\frac{\partial}{\partial K} + K_a\frac{\partial}{\partial K_a} + K_a^\mu\frac{\partial}{\partial K_a^\mu} + 2L_a\frac{\partial}{\partial L_a}\right) \hat{\Gamma}^{(0)}_{ren} = 0. \tag{32.31}$$

Important concepts to remember

- The $SU(2)$-Higgs system has a gauge field and a complex scalar doublet, and the Higgs mechanism generates a quartet of unphysical states.
- To renormalize in the R_ξ ('t Hooft) gauge, we need to add sources for the nonlinear parts of the BRST trasformations to the Lagrangean, K_a^μ, K, K_a, L_a for $Q_B A_\mu^a, Q_B h, Q_B \chi_a, Q_B c^a$.
- Instead of renormalizing μ^2 multiplicatively, we can renormalize $\beta = -\mu^2 + \lambda v^2$, which is 0 classically, additively.
- We must consider α and ξ renormalizations independently at the quantum level, for $G_a = \partial^\mu A_\mu^a + \xi g v / 2 \chi^a$ and $(G^a)^2/2\alpha$ as gauge fixing term.
- We must consider v and h renormalizing independently at the quantum level, but the Ward identity coming from having only $v + h$ dependence reduces the number of divergent structure to the number matching the number of parameters.

Further reading: See chapter 21.3 in [3].

33 Lecture 33. Pseudo-Goldstone bosons, nonlinear sigma model and chiral perturbation theory

In this lecture we will apply the method of effective field theory from lecture 27 to describe low energy QCD. We will use an approximate symmetry called *chiral symmetry*, and its perturbation theory.

We have described spontaneous symmetry breaking, and we saw that Goldstone's theorem says we obtain Goldstone bosons for the broken symmetry directions. But in reality, we never have an exact symmetry, so it is useful to know what happens when an approximate symmetry is broken. In that case, we say we have a pseudo-Goldstone boson.

In QCD, we have 6 quarks: u, d, s, c, b, t. The last 3, c, b and t, are heavy, so a different perturbation theory is used for them. The u and d quarks are nearly massless (their masses are very small), and the s quark is intermediate in mass. Therefore in low energy QCD we consider the u and d quarks, and sometimes the s quark. One can consider also source terms for various currents: vector V_μ for the vector current, axial vector A_μ for the axial vector current, s for the a scalar current and p for a pseudoscalar current, for a total Lagrangean in low energy QCD

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{4}(F^a_{\mu\nu})^2 - \bar{u}\gamma^\mu D_\mu u - \bar{d}\gamma^\mu D_\mu d[-\bar{s}\gamma^\mu D_\mu s] \\
&\quad -\sum_i m_i \bar{q}_i q^i \\
&\quad [-V_\mu \bar{q}\gamma^\mu q - A_\mu \bar{q}\gamma^\mu\gamma_5 q - s\bar{q}q - p\bar{q}\gamma_5 q] \,, \qquad (33.1)
\end{aligned}$$

where the first line is the massless QCD, the second line the mass terms, and the third line the source terms.

We will consider only the u and d quarks for most of the lecture, generalizing to the introduction of the s quark at the end of the lecture.

We will consider then the quark column vector $q = \begin{pmatrix} u \\ d \end{pmatrix}$, which is nearly massless ($m_u \simeq m_d \simeq 0$). Then the low energy Lagrangean has $U(2)_L \times U(2)_R$ symmetry, as we have described in lecture 25, composed of an $SU(2)_L \times SU(2)_R$ part and an $U(1) \times U(1)$ part.

The action of the $SU(2)_L \times SU(2)_R$ on q is

$$q = \begin{pmatrix} u \\ d \end{pmatrix} \to \exp\left[i\alpha_V^a \frac{\tau^a}{2} + i\gamma_5 \alpha_A^a \frac{\tau^a}{2}\right] \begin{pmatrix} u \\ d \end{pmatrix} \,, \qquad (33.2)$$

where τ^a are the Pauli matrices. In terms of

$$\begin{aligned}
q_{L/R} &= \left(\frac{1 \pm \gamma_5}{2}\right) q \\
g_{L/R} &= e^{i(\alpha_V \pm \alpha_A)^a \tau^a_{L/R}} \\
\tau^a_{L/R} &= \left(\frac{1 \pm \gamma_5}{2}\right) \tau^a \,, \qquad (33.3)
\end{aligned}$$

the action of $SU(2)_L \times SU(2)_R$ is

$$\begin{array}{rcl} q_{L/R} & \to & g_{L/R} q_{L/R} \\ V_\mu \pm A_\mu & \to & g_{L/R}(V_\mu \pm A_\mu) g_{L/R}^\dagger + i(\partial_\mu g_{L/R}) g_{L/R}^\dagger \\ s + ip & \to & g_L(s+ip) g_R. \end{array} \tag{33.4}$$

Then $\tau^a_{L/R}$ generate two independent $SU(2)$ algebras, i.e. $SU(2)_L \times SU(2)_R$,

$$\begin{array}{rcl} [\tau^a_L, \tau^b_L] & = & i\epsilon^{abc}\tau^c_L \\ [\tau^a_R, \tau^b_R] & = & i\epsilon^{abc}\tau^c_R \\ [\tau^a_L, \tau^b_R] & = & 0. \end{array} \tag{33.5}$$

Besides this, we have the $U(1) \times U(1)$, acting as

$$(q) \to e^{i\alpha + i\tilde{\alpha}\gamma_5}(q). \tag{33.6}$$

The action by $e^{i\alpha}$ is (3 times) the $U(1)_B$ baryon number (since the baryon number of a baryon like n and p is 1, the baryon number of the quarks u and d is 1/3).

The baryon number is conserved *in QCD!* (*in electroweak theory the baryon number is broken by anomalies through instantons* as we said in lecture 25, since the chiral fermions L and R couple differently to it, unlike the coupling of the fermions to $SU(3)_c$).

On the other hand, $e^{i\tilde{\alpha}\gamma_5}$ is an abelian global chiral symmetry, and is broken by anomalies *in QCD*, giving the solution to the $U(1)$ problem, as we explained in lecture 25.

The $SU(2)_L \times SU(2)_R$ is broken spontaneously to $SU(2)_V$, acting as

$$(q) \to e^{i\alpha^a_V \frac{\tau^a}{2}}(q). \tag{33.7}$$

But the other $SU(2)$, for

$$(q) \to e^{i\alpha^a_A \gamma_5 \frac{\tau^a}{2}}(q) , \tag{33.8}$$

called the chiral (axial) symmetry, is spontaneously broken. That issue, of *chiral symmetry breaking* is one of the most important ones of particle physics, and its exact mechanism is unknown (we don't know an exact low energy effective action that will show the spontaneous breaking). This lecture is devoted to a phenomenological description of the phenomenon.

Since we have a spontaneous breaking of an (approximate) symmetry, by the Goldstone theorem we must have (pseudo-)Goldstone bosons for the 3 generators, and transforming under the unbroken $SU(2)_V$ (isospin). There is only one candidate group, the pions π^a, which are approximately massless, since $m_\pi \ll \Lambda_{QCD}$. Here Λ_{QCD} is the spontaneously generated scale for QCD, that gives the mass of the physical states. (There are various ways to define Λ_{QCD}, but we will not try to define it here).

We note that the Wigner-Eckhart theorem, which says that states should fall under multiplets of the full symmetry group, does not apply to spontaneous symmetry breaking (SSB). Indeed, SSB can be described by the fact that the vacuum is not invariant under the symmetry, i.e. $Q|0\rangle \neq 0$, which also means that the low energy states are not in a multiplet.

So in our case, the pions are not in a multiplet of $SU(2)_A$, but are still in a multiplet of the unbroken group, $SU(2)_V$ or isospin. Indeed, they transform as the adjoint of this group.

We said that we have spontaneous symmetry breaking for $SU(2)_A$, but until now we have seen only SSB via a scalar field VEV, and now we have only fermions and gauge fields in the theory. But the point is that in such a theory, the fermions form *condensates*, i.e. composite scalars made up of the fermions have a VEV. Therefore we have a *quark condensate*,

$$\langle 0|\bar{q}_{Li}q_R^j|0\rangle = -v\delta_i^j. \tag{33.9}$$

Here i, j are flavour indices. Note that there are ways to see that in a gauge theory we have a fermion condensate at low energy, and we can show this for QCD, even if we don't know the exact mechanism.

Up to now we have described massless quarks leading to massless pions (Goldstone bosons). But of course in reality, up and down quarks are not massless, and the pions also have mass, which is much smaller however than the mass of the hadronic states p and n, made up also of only u and d quarks. So we should make an important distinction:

-For spontaneous symmetry breaking with massless pions, we would still obtain nucleons n and p (composite fermions) with a mass. In fact, m_p an m_n are approximately independent of the quark masses m_u, m_d, as could be guessed from the fact that $m_p, m_n \gg m_\pi$.

-On the other hand, a nonzero quark mass m_q is correlated with a nonzero pion mass m_π, so in the presence of quark mass, the pions (Goldstone bosons) are not massless anymore. In fact, we will see at the end of the lecture that we have the relation

$$m_\pi^2 = \frac{2(m_u + m_d)}{f_\pi^2} v = \frac{2(m_u + m_d)}{f_\pi^2} \langle 0|\bar{q}_L q_R|0\rangle. \tag{33.10}$$

The perturbation theory that we will obtain for the pion interactions will be a perturbation theory in p/f_π, and also a perturbation theory in m_π (related to it through the above), called *chiral perturbation theory*, Ch.P.T.

To model spontaneous symmetry breaking for $SO(4) \simeq SU(2)_L \times SU(2)_R \to SU(2)_V$, we can describe the $SO(4)$ symmetry through a bifundamental action on the set $(\mathbb{1}, \tau^a)$ of generators which is a complete set in the space of 2×2 matrices. So we define the matrix

$$\Sigma = \sigma \, \mathbb{1} + i\tau^a \pi^a \ , \tag{33.11}$$

which is therefore an arbitrary 2×2 matrix field (with some reality properties), and where π^a will be related to the pions, but for now is just a set of real scalars. The action of the $SO(4) \simeq SU(2)_L \times SU(2)_R$ symmetry on Σ is given by

$$\Sigma \to g_L \Sigma g_R^\dagger. \tag{33.12}$$

Then, in terms of the field Σ, we describe phenomenologically the SSB for chiral symmetry through the Lagrangean that is a simple generalization of the Higgs Lagrangean,

$$\mathcal{L}_L = -\frac{1}{4} \mathrm{Tr}[\partial_\mu \Sigma \partial^\mu \Sigma^\dagger] + \frac{\mu^2}{4} \mathrm{Tr}[\Sigma \Sigma^\dagger] - \frac{\lambda}{16} [\mathrm{Tr}(\Sigma \Sigma^\dagger)]^2. \tag{33.13}$$

This is called the *linear sigma model*. The kinetic term is the standard one (note that for instance for σ, multiplied by the identiy, the trace gives a factor of 2, so the normalization is canonical), the mass term has the spontaneous symmetry breaking sign ($m^2 = -\mu^2 < 0$), so the potential is the matrix generalization of the Higgs potential.

If we want to couple to external gauge fields A_μ and V_μ (axial vector and vector), we would do it through the covariant derivative

$$D_\mu\Sigma = \partial_\mu\Sigma - i(V_\mu + A_\mu)\Sigma + i\Sigma(V_\mu - A_\mu). \tag{33.14}$$

The theory has a spontaneously broken vacuum with VEV v, and around it the expansion of Σ is

$$\Sigma(x) = (v + s(x))U(x); \quad U(x) = e^{\frac{i\tau^a\pi'^a(x)}{v}}. \tag{33.15}$$

Here s is a scalar with zero VEV, $\langle s\rangle = 0$, v is the VEV, and $\pi'^a(x)$ are massless, so are indentified with the pions. Indeed, there is no mass term for U, so not for π'^a, while there is one for s, which therefore is the "Higgs", i.e. the massive mode.

In the Wilsonian effective action approach, one integrates momenta with $|k| \geq \Lambda$, which means in particular that one must integrate all the fields with masses $m \geq \Lambda$ (since they have always $|k| \geq \Lambda$). This is called *integrating out the massive modes* in the Lagrangean. When doing that, as we saw, we obtain higher dimensional operators in the Lagrangean, but at sufficiently low energies they are negligible, since they come with inverse powers of Λ.

So now, we integrate out the massive modes, including s, and it means that at low energies, to zeroth order, we can just drop the dependence on them (on s), since the higher dimensional operators they will generate are small.

If we do that, the Lagrangean becomes nonlinear, so we have the *nonlinear sigma model*. Indeed, the Lagrangean is now

$$\mathcal{L}_{NL} = -\frac{v^2}{4}\,\mathrm{Tr}[\partial_\mu U\partial^\mu U^\dagger]\,, \tag{33.16}$$

which looks linear, however we have to remember that now we have the constraint $U^\dagger U = 1$, or $\Sigma^\dagger\Sigma = v^2$ (whereas Σ was an arbitrary matrix before the integrating out), so if we solve the constraint, we get a nonlinear action.

To relate to QCD, we must describe who is the matrix U. If the QCD state is the state $|U\rangle$ instead of the vacuum $|0\rangle$ in (33.9), we have a generalization of the relation,

$$\langle U|\bar{q}_{Li}q_R^j|U\rangle = -vU_i^{\ j}\,, \tag{33.17}$$

where $U_i^{\ j}$ is the matrix element of the matrix U associated with the state $|U\rangle$.

The state $|U\rangle$ is a pion state so, using a new normalization that anticipates that $v = f_\pi$,

$$U(x) = \exp\left[\frac{i\pi^a(x)\tau^a}{f_\pi}\right]\,, \tag{33.18}$$

and plugging in the nonlinear sigma model action, which is now

$$\mathcal{L}_{NL} = -\frac{f_\pi^2}{4}\,\mathrm{Tr}[\partial_\mu U\partial^\mu U^\dagger]\,, \tag{33.19}$$

and expanding in $1/f_\pi$, we obtain

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\pi^a\partial^\mu\pi^a + \frac{f_\pi^{-2}}{6}(\pi^a\pi^a\partial^\mu\pi^b\partial_\mu\pi^b - \pi^a\pi^b\partial^\mu\pi^a\partial_\mu\pi^b) + \dots \,, \tag{33.20}$$

which is indeed a perturbation in p/f_π, so is chiral perturbation theory. In fact, one can use the original normalization, and obtain a Lagrangean with v, which can be compared with predictions about the pion decay, and obtain that $v = f_\pi$. We will say more on this later in the lecture.

More generally, we call a linear sigma model as set of N scalars with a symmetry and some canonical kinetic term.

The $SO(N)$ vector model

The most famous example is the $SO(N)$ model, in terms of a scalar that is a vector (fundamental representation) of $SO(N)$, with spontaneous symmetry breaking. The Lagrangean is a simple generalization of the above linear sigma model Lagrangean,

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi^i)^2 + \frac{\mu^2}{2}(\phi^i)^2 - \frac{\lambda}{4}[(\phi^i)^2]^2 \,, \tag{33.21}$$

and is invariant under $SO(N)$ transformations $\phi^i \to R^i{}_j\phi^j$. Note that our QCD case is $N = 4$, for $SO(4) \simeq SU(2) \times SU(2)$. The potential

$$V = -\frac{\mu^2}{2}(\phi^i)^2 + \frac{\lambda}{4}[(\phi^i)^2]^2 \tag{33.22}$$

has a spontaneously broken vacuum at

$$(\phi_0^i)^2 = \frac{\mu^2}{\lambda} \,, \tag{33.23}$$

and by a symmetry transformation we can orient ϕ_0 along the Nth direction,

$$\phi_0 = (0, ..., 0, v). \tag{33.24}$$

We expand the fields around it as

$$\phi^i(x) = (\pi^k(x), v + \sigma(x)) \,, \tag{33.25}$$

where $k = 1, ..., N-1$. Then the Lagrangean becomes

$$\begin{aligned}\mathcal{L} \;\; = \;\; & -\frac{1}{2}(\partial_\mu\pi^k)^2 - \frac{1}{2}(\partial_\mu\sigma)^2 - \frac{1}{2}(2\mu^2)\sigma^2 - \sqrt{\lambda}\mu\sigma^3 \\ & -\sqrt{\lambda}\mu(\pi^k)^2\sigma - \frac{\lambda}{4}\sigma^4 - \frac{\lambda}{2}(\pi^k)^2\sigma^2 - \frac{\lambda}{4}[(\pi^k)^2]^2.\end{aligned} \tag{33.26}$$

In general, a *nonlinear sigma model* is defined as any Lagrangean of the type

$$\mathcal{L} = f_{ij}(\{\phi^k\})\partial_\mu\phi^i\partial^\mu\phi^j. \tag{33.27}$$

That is, a model with a metric (depending on the scalars) on the space of scalars. For instance, a famous example is the 2 dimensional field theory on the worldsheet of a string propagating in a general spacetime, but a more relevant example for phenomenology would be a modulus scalar field in 4 dimensions.

With respect to the above $O(N)$ linear sigma model, the nonlinear sigma model is

$$\mathcal{L} = -\frac{1}{2g^2}(\partial_\mu \tilde{\phi}^i)\partial^\mu \tilde{\phi}^i \ , \tag{33.28}$$

with the constraint

$$\sum_{i=1}^{N} (\tilde{\phi}^i(x))^2 = 1 \ , \tag{33.29}$$

i.e. the fields to be on a unit sphere. It gives a phenomenological description of a system with $O(N)$ symmetry spontaneously broken by a VEV, e.g. by integrating out the massive (radial) mode in the above expansion around a VE to get $\phi^i \to v\tilde{\phi}^i(x)$, so the same thing we did in the $SU(2) \times SU(2)$ case.

We can solve the constraint by $N-1$ Goldstone bosons π^k, as

$$\tilde{\phi}^i = (\pi^1, ..., \pi^{N-1}, \sigma) \ , \tag{33.30}$$

where

$$\sigma = \sqrt{1-\vec{\pi}^2}. \tag{33.31}$$

Now the manifest $SO(N)$ symmetry turns into manifest $SO(N-1)$ symmetry. Then we find that the kinetic term becomes

$$(\partial_\mu \tilde{\phi}^i)^2 = (\partial_\mu \vec{\pi})^2 + \frac{(\vec{\pi}\cdot\partial_\mu\vec{\pi})^2}{1-\vec{\pi}^2} \ , \tag{33.32}$$

so the nonlinear sigma model action becomes

$$\begin{aligned} \mathcal{L} &= -\frac{1}{2g^2}\left[(\partial_\mu \vec{\pi})^2 + \frac{(\vec{\pi}\cdot\partial_\mu\vec{\pi})^2}{1-\vec{\pi}^2}\right] \\ &\simeq -\frac{1}{2g^2}\left[(\partial_\mu \vec{\pi})^2 + (\vec{\pi}\cdot\partial_\mu\vec{\pi})^2 + ...\right] . \end{aligned} \tag{33.33}$$

Note that here the dimension of scalars is zero, $[\phi]=0$, where the coupling has dimension -1, $[g]=-1$, and in the second line we have expanded in the scalars π^a.

To make the connection with chiral perturbation theory, we take $N=4$, so $SO(4) \simeq SU(2)\times SU(2)$, and is spontaneously broken to $SU(2) \simeq SO(3)$ acting on the π^k's. So initially, $SO(4)$ is manifest, but there is a constraint. When solving the constraint for $\tilde{\phi}^i$ in terms of π^k, only $SO(3)$ remains linearly realized, the other $SO(3)$ becomes nonlinearly realized.

The issue of nonlinear realizations follows the same pattern in general: what we usually call a symmetry is a linearly realized symmetry, i.e. a symmetry that acts linearly on the fields. When we have a nonlinearly realized symmetry, it is indicative of a case when there is a more fundamental representation that has the symmetry, like by introducing an auxiliary field, or by writing the theory with a constraint, or that we are in a spontaneously broken vacuum. But in any case, the symmetry is not *manifest* in the action.

We will see that more explicitly in the last form we will describe for the pions.

We can use yet another way to solve the constraint of the $SO(4)$ model. We can write the sigma model as a rotation R acting on the vector $(0,0,0,\sigma)$ parametrized only by the massive mode σ, as

$$\phi_i(x) = R_{i4}(x)\sigma(x) \; , \tag{33.34}$$

where the matrix R is orthogonal (in $SO(4)$), so it satisfies $RR^T = 1$. Then by squaring the above relation and using $RR^T = 1$, we get

$$\sigma(x) = \sqrt{\sum_{i=1}^{4} (\phi^i)^2}. \tag{33.35}$$

Replacing in the linear sigma model Lagrangean, we find

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - \frac{\sigma^2}{2}\sum_{i=1}^{4}\partial^\mu R_{i4}\partial_\mu R_{i4} + \frac{\mu^2}{2}\sigma^2 - \frac{\lambda}{4}\sigma^4. \tag{33.36}$$

Parametrizing the fields as

$$\begin{aligned}
\zeta_a &= \frac{\phi_a}{\phi_4 + \sigma}\\
R_{a4} &= \frac{2\zeta_a}{1+\vec{\zeta}^2} = -R_{4a}\\
R_{44} &= \frac{1-\vec{\zeta}^2}{1+\vec{\zeta}^2}\\
R_{ab} &= \delta_{ab} - \frac{2\zeta_a\zeta_b}{1+\vec{\zeta}^2} \; ,
\end{aligned} \tag{33.37}$$

we obtain the Lagrangean

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - 2\sigma^2\vec{D}_\mu\vec{D}^\mu + \frac{\mu^2}{2}\sigma^2 - \frac{\lambda}{4}\sigma^4 \; , \tag{33.38}$$

where

$$\vec{D}_\mu \equiv \frac{\partial_\mu\vec{\zeta}}{1+\vec{\zeta}^2} \tag{33.39}$$

is often called the *covariant derivative of the pion field.*

Then the transformation rules for the scalars are:

- for isospin $SU(2)_V$, that acts as a $SO(3)$ rotation of ϕ^a, thus of ζ_a, leaving ϕ^4 and σ invariant, we have

$$\delta\vec{\zeta} = \vec{\alpha}\times\vec{\zeta}; \quad \delta\sigma = 0. \tag{33.40}$$

  which is a linear transformation, since the group is unbroken.

- for the axial vector $SU(2)_A$, which is broken, from $\delta\vec{\phi} = 2\vec{\epsilon}\phi_4$ and $\delta\phi_4 = -2\vec{\epsilon}\cdot\vec{\phi}$, we get
$$\delta\vec{D}_\mu = 2(\vec{\zeta}\times\vec{\epsilon})\times\vec{D}_\mu\ , \tag{33.41}$$
which is a nonlinear transformation, i.e. broken.

The VEV of the scalar σ is $\langle\sigma\rangle = v$, and integrating out the fluctuation in σ, we obtain the nonlinear sigma model Lagrangean
$$\mathcal{L} = -2v^2\vec{D}_\mu\vec{D}^\mu. \tag{33.42}$$
Defining the pions as
$$\vec{\pi} \equiv 2v\vec{\zeta}\ , \tag{33.43}$$
we get the pion Lagrangean
$$\mathcal{L} = -\frac{1}{2}\frac{\partial_\mu\vec{\pi}\cdot\partial^\mu\vec{\pi}}{\left(1+\frac{\vec{\pi}^2}{4v^2}\right)^2}. \tag{33.44}$$
We now observe that we have written several Lagrangeans for the pions, (33.20), (33.33) and (33.44). They are all Goldstone boson Lagrangeans, meaning that the interactions all involve derivatives of the pions, and the Lagrangeans are an expansion in p/f_π, i.e. chiral perturbation theory. What does it mean?

Really, it means that we should use the *Wilsonian effective field theory* approach for the pions, and write down the most general Lagrangean consistent with all the symmetries (all the higher dimension operators consistent with the symmetry), and fix the coefficients from comparison with experiments. But this will still not account for the difference between our 3 Lagrangeans. The point is that, of course, a (possibly nonlinear) field redefinition consistent with the symmetries should not change the physics, i.e. the scattering amplitudes, for instance something like
$$\vec{\pi}' = \frac{\vec{\pi}}{1+\vec{\pi^2}} \tag{33.45}$$
should leave the physics invariant. We can use such field redefinitions to put operators we want to zero, leading to various expressions, depending on what terms we want to keep after the redefinitions.

As an example of the comparison with experiment, expanding the Lagrangean in (33.44) in $1/v$ and comparing with experiment, we see that we need $v = f_\pi$, as already stated for (33.20).

We have not described well what f_π is until now, so we will rectify this omission now. The *pion decay constant* f_π is fixed by the PCAC relation from lecture 25, which stated that
$$\partial^\mu j_\mu^{5(A)\pm} = f_\pi m_\pi^2\pi^\pm(x) \tag{33.46}$$
or equivalently
$$j_\mu^{5,\pm}(hadronic) = f_\mu\partial_\mu\pi^\pm(x)\ , \tag{33.47}$$

and another relation for π^3 that contains also the anomalous part, and from which we have derived the $\pi^0 \to \gamma\gamma$ decay. But here we are interested in the decay of $\pi^\pm$, which has no anomalous component. The normalization of π^- is given by

$$\langle 0|\pi^-|\pi^-\rangle = \frac{1}{\sqrt{2m_{\pi^-}}} , \tag{33.48}$$

so we get

$$\langle 0|j_\mu^{5-}|\pi^-\rangle = q_\mu f_\pi \frac{1}{\sqrt{2m_{\pi^-}}}. \tag{33.49}$$

Then we fix f_π from the decay $\pi^- \to \mu^- + \bar{\nu}_\mu$. The relevant electroweak interaction in the 4-fermi limit is

$$\mathcal{L}_{weak} = \frac{G_F}{\sqrt{2}} \bar{\psi}_{(\mu)} \gamma^\rho (1+\gamma_5) \psi_{\nu_\mu} (j_\rho^{V,-} + j_\rho^{A,-}) , \tag{33.50}$$

which after a calculation that will not be reproduced here gives for the decay rate

$$\Gamma = \frac{1}{8\pi} \left(\frac{m_\pi^2 - m_\mu^2}{m_\pi^2} \right)^2 (G_F m_\mu m_\pi)^2 \left(\frac{f_\pi}{m_\pi} \right)^2 \tag{33.51}$$

and then experimentally from the decay, we can fix $f_\pi \simeq 93MeV$.

Generalization.

We can also generalize the mechanism of spontaneous symmetry breaking for a chiral symmetry to any $G \to H$, though we will not do any calculation here. Then the Goldstone bosons parametrize the coset G/H and transform linearly under H and nonlinearly under G/H. Again we can find the most general form allowed by symmetries, do field redefinitions, and fix coefficients from experiment.

Generalization to $SU(3)$.

We now show how to include the s quark, which has a slightly larger mass, so its chiral perturbation theory is somewhat less useful (the corrections are large). The Goldstone boson matrix U is now written as

$$U = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta^0}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta^0}{\sqrt{6}} & K^0 \\ K^- & \bar{k}^0 & -\frac{2}{\sqrt{6}}\eta^0 \end{pmatrix} \tag{33.52}$$

otherwise we have the same idea, so the construction will not be repeated here. We just note that now the adjoint of $SU(3)$ has 8 components, and these are parametrized by $\pi^+, \pi^-, \pi^0, \eta^0, K^+, K^-, K^0, \bar{K}^0$.

As we mentioned, the light quarks are u and d, possibly together with the s quark, whereas the c, b, t quarks are heavy.

Heavy quark effective field theory

For the c, b, t quarks we have another type of effective field theory, one that takes into account the fact that for many things the quarks can be treated as nonrelativistic. So we write something like nonrelativistic quantum mechanics (NRQM), but with Lorentz indices.

We start with the quark Lagrangean

$$\bar{Q}_i i\gamma^\mu \partial_\mu Q_i - M\bar{Q}_i Q_i - \bar{Q}_i g \not{A} Q_i \tag{33.53}$$

and consider momenta

$$p_{(i)} = Mv + k_{(i)} \tag{33.54}$$

where $v \ll 1$ and k is small. The *4-vector* velocity v satisfies $v^2 = 1 \Rightarrow \not{v}^2 = 1$, and we define the projectors

$$P_\pm = \frac{1}{2}(1 \pm \not{v}) \tag{33.55}$$

and split according to it the fields into positive and negative "chirality", $\not{v}h = h$ and $\not{v}\chi = -\chi$, and throw away the "negative chirality", really negative energy, states (such as to go from field theory to NRQM) .

Thus we define

$$Q = e^{-iMv\cdot x}(h + \chi) \ , \tag{33.56}$$

where x^μ is a 4-vector like v^μ, and $h(x)$ now contains only the small variation k^μ and throw away χ, arriving at the HQET Lagrangean to zeroth order

$$\mathcal{L} = \sum_j \bar{h}_j (iv \cdot D) h_j. \tag{33.57}$$

Keeping the next order also, we get from general considerations

$$\mathcal{L} = \sum_j \left[\bar{h}_j (iv \cdot D) h + \frac{1}{2M_Q} \bar{h}_j [\alpha (i\not{D})^2 + \beta (v \cdot D)^2] h_j\right]. \tag{33.58}$$

But by reparametrization invariance we can put $\alpha = 1$, and the second term doesn't contribute to physical processes, so can be dropped.

Coupling to nucleons

Until now we have described only the pions, but from a phenomenological perspective it is even more interesting to describe how they couple to the nucleons that compose matter.

The free nucleon Lagrangean is written as

$$-\bar{\mathcal{N}} \not{\partial} \mathcal{N} - m_N \bar{\mathcal{N}} \mathcal{N}. \tag{33.59}$$

The interaction is written as follows. From Lorentz invariance it must have $\bar{\mathcal{N}}$ and $\mathcal{N}$. From the pions point of view it must be a derivative interaction, so $\partial_\mu \pi^a$, which means that we need $\gamma^\mu \tau^a$ in the action as well. The coupling is an axial coupling, so it will have an γ_5 as well, and by dimensional analysis we need a g_A/f_π, for a total interaction

$$-i\frac{g_A}{f_\pi} \partial_\mu \vec{\pi} \bar{\mathcal{N}} \gamma^\mu \gamma_5 \frac{\vec{\tau}}{2} \mathcal{N} \ , \tag{33.60}$$

where g_A is an axial vector coupling, experimentally found to be about 1.27. By partially integrating the derivative, using the free equation of motion for the nucleon to replace it with m_N, we get

$$g_{\pi NN} = \frac{m_N g_A}{f_\pi}. \tag{33.61}$$

Defining the fields

$$
\begin{aligned}
u &= \exp\left[\frac{i\pi^a\tau^a}{2f_\pi}\right] \\
A_\mu &= \frac{i}{2}(u^\dagger\partial_\mu u - u\partial_\mu u^\dagger) \\
V_\mu &= \frac{i}{2}(u^\dagger\partial_\mu u + u\partial_\mu u^\dagger)
\end{aligned}
\tag{33.62}
$$

(note that $u = \sqrt{U}$) we can complete the interaction term to the full terms

$$
\bar{\mathcal{N}}V_\mu\gamma^\mu\mathcal{N} - g_A\bar{\mathcal{N}}A_\mu\gamma^\mu\gamma_5\mathcal{N}. \tag{33.63}
$$

Then with the field redefinition

$$
\mathcal{N} = \left(u^\dagger\frac{1+\gamma_5}{2} + u\frac{1-\gamma_5}{2}\right)N\ , \tag{33.64}
$$

we can rewrite the Lagrangean as

$$
\mathcal{L} = -\bar{N}\partial\!\!\!/ N - m_N\bar{N}\left(U^\dagger\frac{1+\gamma_5}{2} + U\frac{1-\gamma_5}{2}\right)N - \frac{1}{2}(g_A-1)\bar{N}\gamma^\mu\left(U\partial_\mu U^\dagger\frac{1+\gamma_5}{2} + U^\dagger\partial_\mu U\frac{1-\gamma_5}{2}\right)N. \tag{33.65}
$$

Note that it is written in terms of U, not u now.

Mass terms

By an $SU(2)_L\times SU(2)_R$ transformation, we can put the quark mass matrix in the diagonal form

$$
M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix} e^{-i\theta/2}. \tag{33.66}
$$

Then the mass term is

$$
\mathcal{L}_{mass} = -\,\mathrm{Tr}[\bar{q}_L M q_R]\ , \tag{33.67}
$$

where the trace is over the flavor indices i,j. Since we in the vacuum we have the fermion condensate (33.9), and in the physical state we have (33.17), we replace the mass term by

$$
\mathcal{L}_{mass} = v\,\mathrm{Tr}[MU + M^\dagger U^\dagger]. \tag{33.68}
$$

Considering a real mass matrix, $M = M^\dagger$, and expanding in $1/f_\pi$, we get

$$
\mathcal{L}_{mass} = -\frac{v}{f_\pi^2}(\mathrm{Tr}\,M)\pi^a\pi^a + ...;\ , \tag{33.69}
$$

which implies

$$
m_\pi^2 = \frac{2(m_u+m_d)v}{f_\pi^2}\ , \tag{33.70}
$$

as advertised at the beginning of the lecture. This is called the *Gell-Mann-Oakes-Renner relation*.

Note that sometimes one replaces the mass term by

$$-m_\pi^2 \operatorname{Tr}[U + U^\dagger - 2] \tag{33.71}$$

considering that $m_u \simeq m_d$ and subtracting the constant term from the Lagrangean.

Important concepts to remember

- When we have an approximate symmetry spontaneously broken, we get pseudo-Goldstone bosons.
- Chiral symmetry is the $U(2) \times U(2)$ approximate symmetry of low energy QCD with just u and d quarks. The two $U(1)$'s are the conserved baryon number and an abelian chiral symmetry broken by anomalies, and the $SU(2)_L \times SU(2)_R$ is spontaneously broken to $SU(2)_V$.
- The VEV that breaks the symmetry is a fermion (quark) condensate, i.e. a composite field.
- The exact mechanism of chiral symmetry breaking is unknown, but we make models for it.
- The quark masses do not affect the nucleon masses much (they are nonzero because of confinement, not because of quark masses), but the pion mass squared is proportional to the quark masses.
- Chiral perturbation theory is the effective theory for pions, which is an expansion in p/f_π and m_π.
- The linear sigma model is a phenomenological model for the SSB of chiral symmetry, in terms of the pions π^a and the massive σ field, with canonical (linear) kinetic term. There are various descriptions for it.
- The nonlinear sigma model is the nonlinear model in terms of only the pions, obtained by integrating out the massive modes. In general it is a model with a metric on scalar field space.
- In the $SO(N)$ model, we have N scalars in a vector representation of $SO(N)$. The nonlinear sigma model corresponds to the scalars on a unit sphere.
- In general, for $G \to H$ breaking, the Goldstone bosons live in the coset G/H and transform linearly under H and nonlinearly under G/H.
- Chiral perturbation theory is undestood from the Wilsonian effective field theory approach: write the most general Lagrangean consistent with the symmetries, and use field redefinitions to put various terms to zero.

- One can include the s quark in chiral perturbation theory, and the Goldstone boson matrix includes now $K^{\pm}, K^0, \bar{K}^0$ and η^0, but is less useful since it has larger corrections.
- For the heavy quarks c, b, t, one can use heavy quark effective field theory, which is like nonrelativistic quantum mechanics with Lorentz indices.

Further reading: See chapter 83 in Srednicki, 11.1 and 13.3 in [3], chapters 19.4 and 19.5 in Weinberg vol. II.

References

[1] Horatiu Nastase, "Introduction to quantum field theory I" [HN].

[2] George Sterman, "An introduction to Quantum Field Theory" [GS]

[3] M.E. Peskin and D.V. Schroeder, "An introduction to Quantum Field Theory" [PS]

[4] Pierre Ramond, "Field theory: A modern primer" [PR]

[5] Jan Ambjorn and Jens Lyng Petersen, "Quantum field theory", Niels Bohr Institute lecture notes, 1994. [NBI]

[6] Thomas Banks, "Modern Quantum Field Theory".