19.11.2022

Gauge theories

(Lecture notes - 2022/23)

Fiorenzo Bastianelli

1 Introduction

Gauge theories are building blocks of the standard model of particle physics. Gauge symmetries arise from the requirement that massless spin-one particles, which mediate some of the fundamental forces of nature, should carry only two independent polarizations even when described in terms of equations that are manifestly Lorentz invariant: the photon is conveniently described by $A_\mu(x)$ which has four components as any four-vector, but the expected physical degrees of freedoms are only two. The other two degrees of freedom are eliminated by the gauge symmetry. The principle of gauge invariance also fixes in a simple way all possible interactions mediated by massless particles of spin one in a way consistent with Lorentz invariance. Let us present this method following the construction of the QED and QCD lagrangians, the main examples of abelian and non-abelian gauge theories, respectively

2 Abelian gauge theories and QED

Let us first review how, starting from the theory of free electrons described by the free Dirac equation, one finds the complete QED lagrangian using the gauge invariance principle (and Lorentz invariance). Let us consider the free lagrangian of a Dirac field of mass m

$$\mathcal{L}_{Dirac} = -\overline{\psi}\gamma^\mu\partial_\mu\psi - m\,\overline{\psi}\psi\ . \tag{1}$$

It is invariant under symmetry transformations belonging to the group $U(1)$

$$\begin{array}{lll} \psi(x) & \rightarrow & \psi'(x) = \mathrm{e}^{i\alpha}\psi(x)\ , \qquad\qquad \mathrm{e}^{i\alpha} \in U(1) \\ \overline{\psi}(x) & \rightarrow & \overline{\psi'}(x) = \mathrm{e}^{-i\alpha}\,\overline{\psi}(x)\ . \end{array} \tag{2}$$

This is a global symmetry as the parameter α is constant (spacetime independent).

Let us now see how to extend the symmetry to a local one with arbitrary functions $\alpha(x)$

$$\psi(x) \quad\rightarrow\quad \psi'(x) = \mathrm{e}^{i\alpha(x)}\psi(x) \tag{3}$$

$$\overline{\psi}(x) \quad\rightarrow\quad \overline{\psi'}(x) = \mathrm{e}^{-i\alpha(x)}\,\overline{\psi}(x)\ . \tag{4}$$

The mass term in the lagrangian is already invariant

$$m\,\overline{\psi}\psi \quad\rightarrow\quad m\,\overline{\psi'}\psi' = m\,\overline{\psi}\,\mathrm{e}^{-i\alpha(x)}\mathrm{e}^{i\alpha(x)}\psi = m\,\overline{\psi}\psi\ , \tag{5}$$

but the term with the derivative is not

$$\overline{\psi}\gamma^\mu\partial_\mu\psi \quad\rightarrow\quad \overline{\psi'}\gamma^\mu\partial_\mu\psi' = \overline{\psi}\,\mathrm{e}^{-i\alpha(x)}\gamma^\mu\partial_\mu(\mathrm{e}^{i\alpha(x)}\psi) = \overline{\psi}\gamma^\mu\partial_\mu\psi + i\,\overline{\psi}\gamma^\mu\psi\,\partial_\mu\alpha(x)\ . \tag{6}$$

There is an extra term $i\,\overline{\psi}\gamma^\mu\psi\,\partial_\mu\alpha(x)$ that vanishes only for constant $\alpha(x)$. The lagrangian is not invariant, and one has to modify it to achieve gauge invariance, i.e. invariance for arbitrary functions $\alpha(x)$. Note that the term multiplying the derivative of $\alpha(x)$ is the Noether current associated to the global symmetry (2), namely $J^\mu = i\,\overline{\psi}\gamma^\mu\psi$.

To construct gauge invariant actions it is useful to introduce a formalism based on the definition of tensors of the gauge group and covariant derivatives. The latter are constructed in such a way as to produce tensors out of tensors.

We say that $\psi(x)$ is a tensor under the gauge group $U(1) = \{\mathrm{e}^{i\alpha(x)}\}$ if transforms as in (3). Then $\partial_\mu\psi(x)$ is not a tensor, as it transforms in a more complicated way. The covariant derivative on ψ is defined by

$$D_\mu = \partial_\mu - iA_\mu(x) \tag{7}$$

where $A_\mu(x)$ is a vector field that is required to transform under the gauge group in a suitable way, so that the "tensorial" transformation rule remains valid

$$D_\mu\psi(x) \quad \to \quad D'_\mu\psi'(x) = \mathrm{e}^{i\alpha(x)} D_\mu\psi(x) \tag{8}$$

where $D'_\mu = \partial_\mu - iA'_\mu(x)$. A short calculation shows that we must have the following rule

$$A_\mu(x) \quad \to \quad A'_\mu(x) = A_\mu(x) + \partial_\mu\alpha(x). \tag{9}$$

With covariant derivatives it is simple to obtain a gauge invariant lagrangian from (1):

$$\boxed{\mathcal{L} = -\overline{\psi}\gamma^\mu D_\mu\psi - m\overline{\psi}\psi}\,. \tag{10}$$

Comment: a "tensor" for the gauge group $U(1)$ is more generally defined as a field $\psi_q(x)$ that transforms as

$$\psi_q(x) \to \psi'_q(x) = \mathrm{e}^{iq\alpha(x)}\psi_q(x) \tag{11}$$

where the integer $q \in \mathbb{Z}$ is called the "charge" of $\psi_q(x)$. Thus, $\psi_q(x)$ is a tensor of charge q: in mathematical terms q identifies an irreducible representation of the group $U(1)$. The general definition of covariant derivative is extended to

$$D_\mu = \partial_\mu - iA_\mu(x)Q \tag{12}$$

where Q is an operator that measures the charge of the tensor on which it acts, i.e. it is the generator of the $U(1)$ group in the same representation of the tensor it acts upon. We now recognize that the transformation in (4) corresponds to that of a tensor of charge -1. The covariant derivative has the property that it does not destroy the tensorial character of the object on which it acts: it generates tensors out of tensors. Indeed, one may verify that

$$D_\mu\psi_q(x) = \partial_\mu\psi_q(x) - iqA_\mu(x)\psi_q(x) \tag{13}$$

is again a tensor of charge q (like eq. (8) for charge $q = 1$). Another property of this definition is the validity of the Leibniz rule for covariant derivatives: product of tensors are again tensors and one may verify that

$$D_\mu(\psi_{q_1}\psi_{q_2}) = (D_\mu\psi_{q_1})\psi_{q_2} + \psi_{q_1}(D_\mu\psi_{q_2}). \tag{14}$$

Thus, local invariance is achieved by introducing the gauge field $A_\mu(x)$, readily recognized as the potential of the electromagnetic field. Having introduced a new field, one has to give it a suitable dynamics by adding to the lagrangian a kinetic term for $A_\mu(x)$. This term has to be gauge invariant, because the rest of the Lagrangian already is: gauge symmetry is the guiding principle for building the action. It is useful to proceed using tensors. We can calculate the commutator of two covariant derivatives acting on the tensor ψ of charge 1

$$[D_\mu, D_\nu]\psi = -iF_{\mu\nu}\psi \tag{15}$$

that defines the quantity $F_{\mu\nu}$. Since we have only tensorial quantities on the left-hand side, the right-hand side must also be built out of tensors. We recognize that $F_{\mu\nu}$ is a tensor of charge $q = 0$, i.e. a quantity that is invariant under gauge transformations (to see this it is enough to set $q = 0$ in eq. (11)). Computing explicitly the left-hand side of (15) one finds

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{16}$$

readily recognized as the electromagnetic field tensor.

Now, it is immediate to construct a gauge invariant lagrangian with at most two derivatives on A_μ. It is enough to use as building block the field strength $F_{\mu\nu}$ which is gauge invariant. One must require also Lorentz invariance to have a relativistic theory, and one is led to the free Maxwell lagrangian, that in a standard normalization reads

$$\mathcal{L}_{Maxwell} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{17}$$

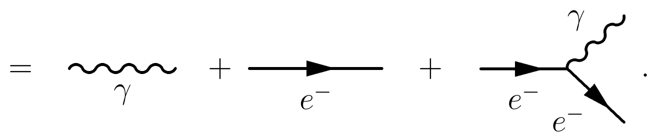
Summing together all the pieces that are separately gauge invariant (i.e. eqs. (10) and (17)) one finds the QED lagrangian

$$\boxed{\mathcal{L}_{QED} = -\frac{1}{4e^2}F_{\mu\nu}F^{\mu\nu} - \overline{\psi}\gamma^\mu D_\mu\psi - m\overline{\psi}\psi} \tag{18}$$

where a free multiplicative parameter $1/e^2$ accounts for a relative weight between the different terms that are separately gauge invariant.

Let us analyze the various terms contained in (18). It is useful to redefine $A_\mu \to eA_\mu$ (to obtain the standard nomalization of the free Maxwell action) and recognize that e is the coupling constant (now it appears in the covariant derivative $D_\mu = \partial_\mu - ieA_\mu(x)$)

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \overline{\psi}(\gamma^\mu\partial_\mu + m)\psi + ieA_\mu\overline{\psi}\gamma^\mu\psi$$



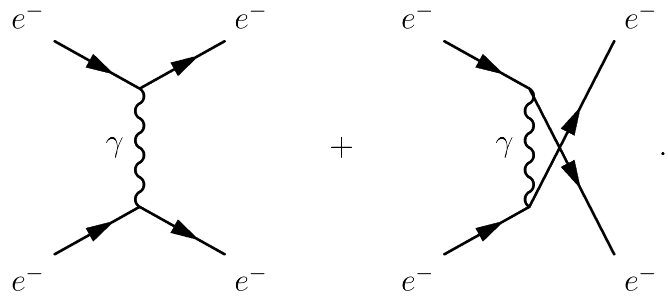
(19)

The first term describes the free propagation of photons, the second one the free propagation of electrons, and the third one the elementary interaction between photons and electrons. The constant e is the coupling constant, identified with the elementary charge of the electron: the gauge principle has allowed us to discover the interaction between fields of spin 1/2 and 1. Let us summarize again the rules of gauge transformations for the lagrangian in (19): rescaling for simplicity also the angle $\alpha(x) \to e\alpha(x)$ we have

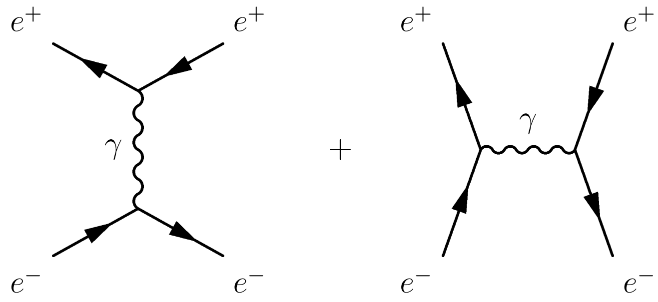
$$\psi \to \psi' = \mathrm{e}^{ie\alpha}\psi$$
$$\overline{\psi} \to \overline{\psi'} = \mathrm{e}^{-ie\alpha}\overline{\psi}$$

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \alpha. \tag{20}$$

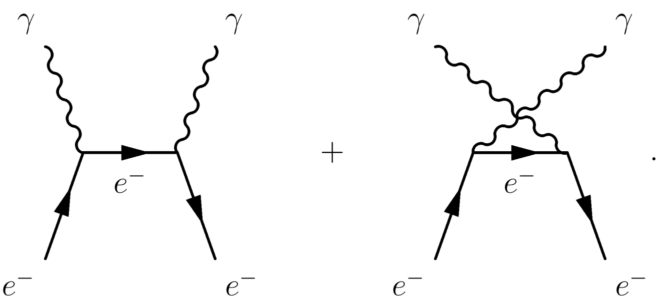
If the coupling constant e is small enough it can be treated perturbatively, and the amplitudes for the various physical processes of QED can be associated to the Feynman diagrams built with the elementary vertex in (19). For example, the electron-electron scattering (Möller scattering) at the lowest order is given by (time runs along the horizontal axis)



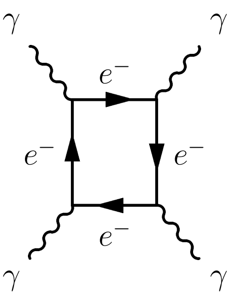
Other processes are the electron/positron scattering (Bhabha scattering)



and the electron/photon scattering (Compton scattering)



Also photon-photon scattering is possible: there is no elementary vertex and the first perturbative term that is found is given by the graph



together with similar graphs where the external photon lines are attached to the vertices with different orderings. In general, loop corrections can be divergent and must be cured by renormalization. However, for the photon-photon scattering the calculation of the Feynman graph depicted above is finite, and there is no need to renormalize it. This fact may be interpreted as a consequence of gauge invariance.

3 Non-abelian gauge theories and QCD

The construction of gauge invariant actions can be extended to compact non-abelian groups.

3.1 Lie groups

Let us briefly recall some properties of non-abelian Lie groups. We consider simple and compact Lie gropus, having in mind $SU(N)$ as the main example. An element U of a non-abelian Lie group G can be parametrized by coordinates α_a (the parameters) associated to the hermitian generators T^a. Here is a list of the main properties:

$$
\begin{aligned}
(i) &\quad U = \exp(i\alpha_a T^a) \in G \qquad a = 1, .., \dim G \\
(ii) &\quad [T^a, T^b] = i f^{ab}{}_c T^c \\
(iii) &\quad \mathrm{tr}(T^a_F T^b_{\mathrm{F}}) = \frac{1}{2}\delta^{ab} \\
(iv) &\quad f^{abc} = f^{ab}{}_d \delta^{dc} \qquad \text{antisymmetric tensor} \\
(v) &\quad [[T^a, T^b], T^c] + [[T^b, T^c], T^a] + [[T^c, T^a], T^b] = 0 \\
&\quad \Rightarrow\ f^{ab}{}_d f^{dc}{}_e + f^{bc}{}_d f^{da}{}_e + f^{ca}{}_d f^{db}{}_e = 0 \\
(iv) &\quad (T^a_{\mathrm{Adj}})^b{}_c = -i f^{ab}{}_c \ .
\end{aligned}
$$

(i) describes the exponential representation of an arbitrary element U of the group G connected to the identity, $U \in G$. The index a takes as many values as the dimension of the group. Therefore, an element of the group is parameterized by the "angles" α_a.
(ii) is the Lie algebra satisfied by the hermitian generators T^a. The real constants $f^{ab}{}_c$ are the structure constants and characterize the group G.
(iii) defines a choice for the normalization of the generators in the fundamental representation T^a_F, also called the defining representation and identifies the so-called "Killing metric". More generally, one could define $\mathrm{tr}(T^a_F T^b_{\mathrm{F}}) = \frac{1}{2}\gamma^{ab}$ with γ^{ab} the Killing metric that is proved to be positive definite for compact Lie groups (such as $SU(N)$). The normalization chosen above produces the Kronecker delta δ^{ab} as the Killing metric for the compact group G.
(iv) makes use of the Killing metric to raise an index in the structure constants. Then, the symbols f^{abc} are completely antisymmetric: this property can be deduced by multiplying the Lie algebra with an additional generator, taking the trace, and using (iii) and the cyclic property of the trace. On the other hand, the antisymmetry on the indices a and b is obvious from (ii).
(v) gives the Jacobi identities.
(vi) defines the adjoint representation. It is proven to be a representation by using the Jacobi identities.

3.2 Action with rigid $SU(N)$ symmetry

Let us now consider N free Dirac fields with identical masses m, assembled in column and raw vectors

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \cdot \\ \cdot \\ \psi^N \end{pmatrix} , \qquad \overline{\psi} = (\overline{\psi}_1, \overline{\psi}_2, ., ., \overline{\psi}_N) \tag{21}$$

so that the scalar product

$$\overline{\psi}\psi = \overline{\psi}_1\psi^1 + \overline{\psi}_2\psi^2 + \cdots + \overline{\psi}_N\psi^N \tag{22}$$

is an $SU(N)$ invariant. The free lagrangian is given by

$$\mathcal{L}_{Dirac} = -\overline{\psi}\gamma^\mu\partial_\mu\psi - m\,\overline{\psi}\psi \tag{23}$$

and is invariant under the $SU(N)$ symmetry transformations given by

$$\begin{aligned} \psi(x) \quad &\to \quad \psi'(x) = U\psi(x) \\ \overline{\psi}(x) \quad &\to \quad \overline{\psi}'(x) = \overline{\psi}(x)U^\dagger = \overline{\psi}(x)U^{-1} \end{aligned} \tag{24}$$

where $U \in SU(N)$, and $U^\dagger = U^{-1}$ since U is unitary. These are global transformations, as the α^a parameters contained in $U = U(\alpha) = \exp(i\alpha^a T^a)$ are constant (indices in α^a are raised and lowered with the Killing metric, that coincides with the identity in our conventions).

3.3 Covariant derivative

To make the $SU(N)$ symmetry local it is again convenient to introduce the concept of covariant derivatives. By definition, the covariant derivative when applied to tensors produces new tensors. To start with, we recall that a tensor $\psi(x)$ in the fundamental representation of the gauge group $SU(N)$, i.e. the representation that is sometimes indicated by its dimension N, is a field defined by the transformation

$$\psi(x) \quad \to \quad \psi'(x) = U(x)\psi(x) \tag{25}$$

where $U(x)$ is a $N \times N$ matrix of $SU(N)$ for any spacetime point x. More generally, fields transforming in any given representation $R(U(x))$ of the original matrices $U(x)$ are said to be tensors in the representation R. For example, the field $\overline{\psi}(x)$ is a tensor in the antifundamental representation, usually indicated by $\bar{N}$, and its transformation rule is

$$\overline{\psi}(x) \quad \to \quad \overline{\psi}'(x) = \overline{\psi}(x)U^{-1}(x)\ . \tag{26}$$

Evidently, the term $\overline{\psi}(x)\psi(x)$ is a scalar under the gauge transformation. As said, the covariant derivative acting on tensors produces new tensors. It is defined by

$$D_\mu = \partial_\mu + W_\mu(x) \tag{27}$$

where $W_\mu(x)$ is a matrix valued gauge field, also known as the connection (geometrically, it defines a parallel transport in a certain space). When applied to the tensor $\psi(x)$, the term with the gauge field W_μ mixes the N Dirac fermions contained in ψ, and thus is formed by $N \times N$ matrices for any μ. It performs infinitesimal group transformations (that defines a sort of parallel transport) and thus can be expanded in terms of the generators as follows

$$W_\mu(x) = -iW_\mu^a(x)T^a \ . \tag{28}$$

This relation defines the gauge fields $W_\mu^a(x)$, and there are N^2-1 of them for the gauge group $SU(N)$. From the requirement of covariance

$$\begin{aligned} \psi(x) \quad &\to \quad \psi'(x) = U(x)\psi(x) \\ D_\mu\psi(x) \quad &\to \quad D'_\mu\psi'(x) = U(x)D_\mu\psi(x) \end{aligned} \tag{29}$$

one obtains the following transformation rule for the gauge potentials

$$W_\mu(x) \quad \to \quad W'_\mu(x) = U(x)W_\mu(x)U^{-1}(x) + U(x)\partial_\mu U^{-1}(x) \ . \tag{30}$$

Indeed, requiring that $D'_\mu\psi' = UD_\mu\psi$, one computes

$$\begin{aligned} D'_\mu\psi' &\equiv (\partial_\mu + W'_\mu)\psi' \\ &= UD_\mu\psi = U\partial_\mu\psi + UW_\mu\psi = U\partial_\mu(U^{-1}U\psi) + UW_\mu U^{-1}U\psi \\ &= U\partial_\mu(U^{-1}\psi') + UW_\mu U^{-1}\psi' = \partial_\mu\psi' + [UW_\mu U^{-1} + U\partial_\mu U^{-1}]\psi' \end{aligned} \tag{31}$$

and finds the above transformation rule for W_μ. To be more precise, as ψ transforms in the fundamental representation of $SU(N)$, the T^a contained in the W_μ of eq. (31) are the generators in the fundamental representation.

Covariant derivatives do not commute. This fact allows to define the "curvature" tensor (or "field strength") the following way

$$[D_\mu, D_\nu]\psi = F_{\mu\nu}\psi \tag{32}$$

so that

$$F_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + [W_\mu, W_\nu] \ . \tag{33}$$

It is immediate to check that the field strength transform covariantly as

$$F_{\mu\nu} \quad \to \quad F'_{\mu\nu} = UF_{\mu\nu}U^{-1} \tag{34}$$

which follows from the covariance of (32). This rule corresponds to the adjoint representation.

3.4 Gauge invariant action

It is now simple to construct a gauge invariant lagrangian from (23): it is enough to substitute derivatives with gauge covariant derivatives (this is also called "minimal coupling"). One obtains

$$\mathcal{L}_1 = -\overline{\psi}(\gamma^\mu D_\mu + m)\psi \tag{35}$$

that depends on the new field W_μ contained in D_μ. Now, one must give a dynamics to W_μ by using the simplest gauge and Lorentz invariant action with at most two derivatives: the lagrangian is

$$\mathcal{L}_2 = \frac{1}{2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu}) = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu a} \tag{36}$$

where we used the generators in the fundamental representation normalized by $\mathrm{tr}\, T^a T^b = \frac{1}{2}\delta^{ab}$. Now, introducing a coupling constant g to define a relative weight between the different gauge invariant pieces, one obtains the final lagrangian

$$\boxed{\mathcal{L} = \frac{1}{2g^2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu}) - \overline{\psi}(\gamma^\mu D_\mu + m)\psi} \tag{37}$$

with gauge symmetries recapitulated as follows

$$\begin{aligned} \psi(x) &\rightarrow \psi'(x) = U(x)\psi(x) \\ \overline{\psi}(x) &\rightarrow \overline{\psi}'(x) = \overline{\psi}(x)U^{-1}(x) \\ W_\mu(x) &\rightarrow W'_\mu(x) = U(x)W_\mu(x)U^{-1}(x) + U(x)\partial_\mu U^{-1}(x)\ . \end{aligned} \tag{38}$$

Let us report the infinitesimal transformations as well. Defining the matrix $\alpha \equiv -i\alpha_a T^a$ with parameters $\alpha_a \ll 1$, one writes an infinitesimal transformation in the form $U = e^{i\alpha_a T^a} = e^{-\alpha} = 1 - \alpha + O(\alpha^2)$, so that

$$\begin{aligned} \delta\psi &= -\alpha\psi \\ \delta\overline{\psi} &= \overline{\psi}\alpha \\ \delta W_\mu &= \partial_\mu \alpha + [W_\mu, \alpha] = D_\mu \alpha \end{aligned} \tag{39}$$

where in the last line the covariant derivative acts in the adjoint representation. The infinitesimal form of the gauge transformations will be used when studying the gauge fixing procedure that is needed to quantize the theory.

We can rewrite the lagrangian by a field redefinition, $W_\mu \rightarrow \bar{W}_\mu = gW_\mu$, to get the canonical normalization for the gauge field. In components

$$\begin{aligned} W_\mu(x) &= -iW^a_\mu(x)T^a \\ F_{\mu\nu}(x) &= -iF^a_{\mu\nu}(x)T^a \end{aligned} \tag{40}$$

and we get

$$F^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + gf^{abc}W^b_\mu W^c_\nu \tag{41}$$

so that the complete lagrangian takes the form

$$\boxed{\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu a} - \overline{\psi}[\gamma^\mu(\partial_\mu - igW^a_\mu T^a) + m]\psi}\ . \tag{42}$$

The infinitesimal gauge transformations, obtained by redefining also the parameters $\alpha^a \rightarrow g\alpha^a$, now read

$$\begin{aligned} \delta\psi(x) &= ig\alpha^a(x)T^a\psi(x) \\ \delta W^a_\mu(x) &= \partial_\mu \alpha^a(x) + gf^{abc}W^b_\mu(x)\alpha^c(x) = D_\mu \alpha^a(x)\ . \end{aligned} \tag{43}$$

The first term in the action (42) describes the free propagation of the fields W^a_μ (the non-abelian spin 1 particles) along with cubic and quartic self-interactions. A positive non-definite Killing metric would result in a term with kinetic energy that is not positive-definite, and this would not be acceptable: it is necessary to consider only compact groups to satisfy this request. The second term describes the free propagation of the ψ fields (spin 1/2 particles with non-abelian charges, i.e. “color” charges) together with their interaction with the gauge field. The constant g is the coupling constant. It can be treated perturbatively if it is small enough. The "non-abelian" or "color" charge corresponds to the representation of the gauge group chosen for the ψ fields (in our case we have taken the fundamental representation, but any other representation could have been chosen as well.). The gauge principle allows to derive all the interaction vertices between fields of spin 1/2 and 1 in terms of the single coupling constant g.

As a consequence of the transformation law (43), or directly from (34), one recognizes that the field $F^a_{\mu\nu}$ transforms in the adjoint representation

$$\delta F^a_{\mu\nu} = g f^{abc} F^b_{\mu\nu} \alpha^c = ig\alpha^c (T^c_{\rm Adj})^{ab} F^b_{\mu\nu} \tag{44}$$

with the generators in the adjoint representation given by

$$(T^c_{\rm Adj})^{ab} = -if^{abc} \ . \tag{45}$$

That this defines a representation follows from the Jacobi identities.

The transformation of the field $F^a_{\mu\nu}$ may be compared with the transformation of the fermion field $\psi(x)$ in the first line of (43), which after introducing indices may be written as

$$\delta\psi^i(x) = ig\alpha^a(x)(T^a)^i{}_j\psi^j(x) \tag{46}$$

with $i, j = 1, ..., N$, and $(T^a)^i{}_j$ the generators in the fundamental representation. Similarly, the transformation rules for the Dirac conjugate field (morally, the complex conjugate field) are as follows

$$\delta\overline{\psi}_i(x) = ig\alpha^a(x)(T^a_{\bar F})_i{}^j\overline{\psi}_j(x) \tag{47}$$

where $T^a_{\bar F} = -T^{a*}_F = -T^{aT}_F$ are the generators in the complex conjugate of the fundamental represenation (the latter has generators $T^a_F = T^a$ as used above). Thus, one may appreciate the similarities of the given expressions for tensors in different representations. Let us also show explicitly that the transformation law of W^a_μ can be expressed in terms of the covariant derivative acting on a tensor in the adjoint representation

$$\delta W^a_\mu = \partial_\mu\alpha^a + gf^{abc}W^b_\mu\alpha^c = \partial_\mu\alpha^a - igW^b_\mu(T^b_{\rm Adj})^{ac}\alpha^c = D_\mu\alpha^a \ . \tag{48}$$

Note also that the non-derivative part of this transformation can be as

$$\delta W^a_\mu = ig\alpha^c(T^c_{\rm Adj})^{ab}W^b_\mu + \cdots \tag{49}$$

which highlights the tensorial character of this part of the transformation, matching (44).

Finally, one may recall that the Jacobi identity for arbitrary operators, once applied to the covariant derivatives

$$[D_\mu, [D_\nu, D_\lambda]] + [D_\nu, [D_\lambda, D_\mu]] + [D_\lambda, [D_\mu, D_\nu]] = 0 \ , \tag{50}$$

gives rise to the so-called Bianchi identities for the field strength $F_{\mu\nu}$

$$D_\mu F_{\nu\lambda} + D_\nu F_{\lambda\mu} + D_\lambda F_{\mu\nu} = 0 \ . \tag{51}$$

3.5 The action of quantum crodmodynamics (QCD)

The action of quantum chromodynamics is based on the group $SU(3)$. In addition to the gluons (the eight particles associated to the gauge field W^a_μ, which has an index in the adjoint representation, and thus belongs to the **8** of $SU(3)$), the lagrangian contains six fermion fields ψ_f corresponding to the six known flavors of quarks, $f = (u, d, c, s, t, b)$, i.e. up, down, charm, strange, top, bottom. Each quark flavor is degenerate, as it transforms in the **3** of the $SU(3)$ gauge group: the quark is said to be colored (with color red, green and blue, in the usual convention). The absence of color indicates a scalar, like the lagrangian (it correspond to the **1** of $SU(3)$). Of course, the corresponding antiparticles, the antiquarks $(\bar{u}, \bar{d}, \bar{c}, \bar{s}, \bar{t}, \bar{b})$, transform in the conjugate representation, the $\bar{\mathbf{3}}$ of $SU(3)$ (which is the representation of $\overline{\psi}_f$, the Dirac conjugate of ψ_f).

The eight infinitesimal generators of $SU(3)$ in the fundamental representation are given by the Gell-Mann matrices λ^a (which generalize the Pauli matrices σ^i of $SU(2)$)

$$T^a = \frac{\lambda^a}{2} \qquad a = 1, \dots, 8 \tag{52}$$

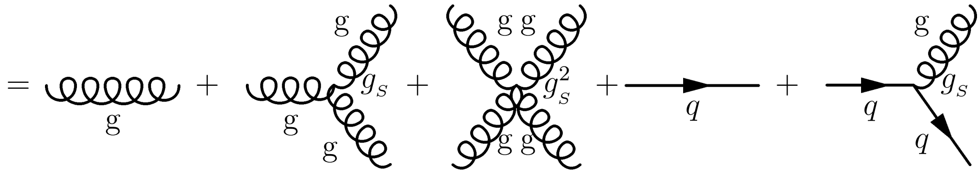
where

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda^5 = \begin{pmatrix} 0 & & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}$$
$$\lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{53}$$

These matrices are normalized according the convention $\mathrm{tr}(T^a T^b) = \frac{1}{2}\delta^{ab}$.

An arbitrary element of the $SU(3)$ group in the fundamental representation is therefore described by 3×3 matrices of the form $U = \exp(i\alpha_a \frac{\lambda^a}{2})$. By calculating the Lie algebra one finds the explicit values of the structure constants of the $SU(3)$ group. The QCD lagrangian is therefore

$$\begin{aligned}
\mathcal{L}_{\mathcal{QCD}} &= -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} - \sum_{f=1}^{6} \overline{\psi}_f \Big(\gamma^\mu D_\mu + m_f \Big) \psi_f \\
&= -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} - \sum_{f=1}^{6} \overline{\psi}_f \Big(\gamma^\mu \partial_\mu + m_f \Big) \psi_f + i \frac{g_S}{2} W^a_\mu \sum_{f=1}^{6} \overline{\psi}_f \gamma^\mu \lambda^a \psi_f
\end{aligned} \tag{54}$$



where the coupling constant is denoted by g_S, and the index $f \in (1, 2, \cdots, 6) = (u, d, c, s, t, b)$ indicates the flavor of the quark. Different flavors of quarks have different masses m_f. Note that to obtain the propagator of the gauge field from the first term, as indicated in the figure, one must implement a gauge-fixing procedure.

The QCD lagrangian possesses also additional rigid symmetries. A well-known rigid symmetry is the $U(1)$ symmetry which rotates all fields of the quarks by the same phase: the associated conserved charge is the *baryon number*. It is a symmetry that is also preserved by the other fundamental interactions.

Other $U(1)$ symmetries rotate the various fermionic fields separately. They give rise to conservation laws of the respective *fermion numbers* (e.g. *strangeness* $\mathcal{S}$, *charm* $\mathcal{C}$, etc ..). These flavor symmetries are exact only for QCD (and QED), but the weak force violates them. In total there are six $U(1)$ independent conserved charges, one for each quark flavor, and the baryon number is a particular linear combination of these six independent charges. Also the *electric charge* Q is a linear combination of them: is the one that is gauged to obtain the electromagnetic couplings.

A summary of these U(1) symmetries is given in the following table, which reports the various $U(1)$ charges with a standard normalization:

Quarks	$\mathcal{U}$	$\mathcal{D}$	$\mathcal{C}$	$\mathcal{S}$	$\mathcal{T}$	$\mathcal{B}$	B	Q
u	1	0	0	0	0	0	$\frac{1}{3}$	$\frac{2}{3}$
d	0	-1	0	0	0	0	$\frac{1}{3}$	$-\frac{1}{3}$
c	0	0	1	0	0	0	$\frac{1}{3}$	$\frac{2}{3}$
s	0	0	0	-1	0	0	$\frac{1}{3}$	$-\frac{1}{3}$
t	0	0	0	0	1	0	$\frac{1}{3}$	$\frac{2}{3}$
b	0	0	0	0	0	-1	$\frac{1}{3}$	$-\frac{1}{3}$

note that we have indicated the baryon number by B, and the bottom (or beauty) quantum number by $\mathcal{B}$. Just to be clear, for each symmetry, each quark flavour transforms with the charge indicated in the table, for example for the electric charge Q we have

$$\psi_f \to \psi'_f = e^{i\alpha Q_f}\psi_f \ . \tag{55}$$

By looking at the table, one recognizes the following relations

$$\begin{aligned} B &= \frac{1}{3}(\mathcal{U}+\mathcal{C}+\mathcal{T}) - \frac{1}{3}(\mathcal{D}+\mathcal{S}+\mathcal{B}) \\ Q &= \frac{2}{3}(\mathcal{U}+\mathcal{C}+\mathcal{T}) + \frac{1}{3}(\mathcal{D}+\mathcal{S}+\mathcal{B}) \ . \end{aligned} \tag{56}$$

There are also other approximate symmetries of the QCD lagrangian. In the limit in which some of the quark masses are taken to be identical, there is a rigid additional non-abelian symmetry. For example, assuming identical masses for the up and down quarks, $m_u = m_d$, one can rotate the fields ψ_u and ψ_d with each other with a $SU(2)$ matrix

$$\begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix} \to \begin{pmatrix} \psi'_u \\ \psi'_d \end{pmatrix} = U \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix} \qquad U \in SU(2) \ . \tag{57}$$

This rigid $SU(2)$ symmetry corresponds to the strong *isospin* $\vec{I}$, used to group hadrons into families (states of quarks bound by the strong force show the phenomenon of color confinement: the bound states are color singlets, corresponding to the mesons and baryons). Examples of these families are: (i) the isospin doublet of the nucleons (proton and nucleon) composed of three confined up and down quarks; (ii) the triplet of π mesons, the pions $\pi^\pm$ and π^0, composed of a quark and an antiquark of the up and down types.

Considering identical the masses for the quarks up, down, and strange, $m_u = m_d = m_s$, one finds an even larger symmetry group, the $SU(3)$ *flavor group*, that mixes the three flavors up, down and strange:

$$\begin{pmatrix} \psi_u \\ \psi_d \\ \psi_s \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} \psi'_u \\ \psi'_d \\ \psi'_s \end{pmatrix} = U \begin{pmatrix} \psi_u \\ \psi_d \\ \psi_s \end{pmatrix} \qquad U \in SU(3)\ . \tag{58}$$

This $SU(3)$ flavor group is the one used in the static quark model (the "eightfold way" of Gell-Mann) to take care of the similarities observed between the various hadrons. It should not be confused with the color group, also an $SU(3)$ group. As already said, color is expected to confine inside the hadrons and leave composite colorless states. Examples of multiplets of hadronic particles described by the $SU(3)$ flavor group are:
the meson octet $(\pi^\pm, \pi^0, K^\pm, K^0, \bar{K}^0, \eta)$,
the baryon octet $(p, n, \Sigma^\pm, \Sigma^0, \Xi^\pm, \Lambda)$,
the baryon decuplet $(\Delta^-, \Delta^0, \Delta^+, \Delta^{++}, \Sigma^{*\pm}, \Sigma^{*0}, \Xi^{*\pm}, \Omega^-)$.
The existence of these families is understandable from group theory: the **8** and the **10** are representations of $SU(3)$. Let us consider the mesons in more detail. They consist of a quark-antiquark pair $(q\bar{q})$. The quarks q transform in the $\mathbf{3}$ of $SU(3)$, with $\mathbf{3} \sim (u, d, s)$, while antiquarks $\bar{q}$ transforms in the $\bar{\mathbf{3}}$ of $SU(3)$, with $\bar{\mathbf{3}} \sim (\bar{u}, \bar{d}, \bar{s})$. From this, it follows that possible bound states $(q\bar{q})$ must transform in the

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$$

and therefore both singlet and octets could in principle exist for the mesons.

On the other hand, baryons are bound states of three quarks (qqq), and since under $SU(3)$

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = (\mathbf{6} \oplus \bar{\mathbf{3}}) \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$$

octets and decuplets could exist for the baryons, as indeed they do.

A Notes on group theory

A.1 Lie groups and algebras

Given a simple and compact Lie group G, we indicate its elements using the exponential parametrization $U(\alpha) = \exp(i\alpha_a T^a)$, where T^a are the infinitesimal hermitian generators that satisfy the Lie algebra

$$[T^a, T^b] = i f^{ab}{}_c T^c\ . \tag{59}$$

In general, considering an irreducible representation R of G, we get an irreducible representation of its Lie algebra with traceless hermitian matrices T_R^a

$$[T_R^a, T_R^b] = if^{ab}{}_c T_R^c \ . \tag{60}$$

The matrices T_R^a act on a vector space of dimensions $D(R)$, and thus are $D(R)\times D(R)$ matrices. $D(R)$ is called the dimension of the representation. We will mostly consider $SU(N)$, whose most used representations are:
• the fundamental (or defining) representation N, with $D(N) = N$
• its complex conjugate representation $\bar{N}$, with $D(\bar{N}) = N$
• the adjoint representation Adj, with $D(\text{Adj}) = N^2 - 1$.

Given a representation R with generators T_R^a, the generators of its complex conjugate representation $\bar{R}$ are given by

$$T_{\bar{R}}^a = -(T_R^a)^* \tag{61}$$

as seen from taking the complex conjugate of the original representation

$$(\exp(i\alpha_a T_R^a))^* = \exp(-i\alpha_a (T_R^a)^*) \equiv \exp(i\alpha_a T_{\bar{R}}^a) \ . \tag{62}$$

The generators are normalized so that in the fundamental representation one has

$$\text{tr}(T^a T^b) = \frac{1}{2}\delta^{ab} \tag{63}$$

which normalizes the so-called Killing metric $\gamma^{ab} = 2\,\text{tr}(T^a T^b)$ to $\gamma^{ab} = \delta^{ab}$. This matrix is used to define scalar products and to raise/lower the indices that label the generators. In particular, it is used to define the structure constants with all upper indices

$$f^{abc} = f^{ab}{}_d \delta^{dc} \tag{64}$$

(more generally $f^{abc} = f^{ab}{}_d \gamma^{dc}$). This is proven to be totally antisymmetric. The antisymmetry of f^{abc} is obvious on the first two indices, as seen from the definition of the Lie algebra. Then using (59) and (63) one can compute

$$\begin{aligned} \text{tr}([T^a, T^b]T^c) &= if^{ab}{}_d \,\text{tr}(T^d T^c) = \frac{i}{2} f^{abc} = \text{tr}(T^a T^b T^c) - \text{tr}(T^b T^a T^c) \\ &= \text{tr}(T^c T^a T^b) - \text{tr}(T^a T^c T^b) = -\text{tr}([T^a, T^c]T^b) = -\frac{i}{2} f^{acb} \end{aligned} \tag{65}$$

so that $f^{abc} = -f^{acb}$, which implies complete antisymmetry. In the above manipulations, we have used the cyclic property of the trace.

The structure constants can be used to define the adjoint representation 'Adj' by

$$(T_{\text{Adj}}^a)^b{}_c = -if^{ab}{}_c \tag{66}$$

since the relation

$$[T_{\text{Adj}}^a, T_{\text{Adj}}^b] = if^{ab}{}_c T_{\text{Adj}}^c \tag{67}$$

reduces to the Jacobi identity and is thus satisfied.

One defines the *index* $T(R)$ of a representation R by

$$\text{tr}(T_R^a T_R^b) = T(R)\,\delta^{ab} \ . \tag{68}$$

with the index of the fundamental representation N normalized by (63) to $T(N) = \frac{1}{2}$.

Casimir operators are operators built from the generators which commute with all the generators of the group. In particular, the quadratic Casimir operator that is constructed using the Killing metric

$$C_2 = T^a T^b \gamma_{ab} = T^a T^a \tag{69}$$

is such an operator. The proof is simple

$$\begin{aligned}[C_2, T^b] = [T^a T^a, T^b] &= T^a [T^a, T^b] + [T^a, T^b] T^a = T^a i f^{abc} T^c + i f^{abc} T^c T^a \\ &= i f^{abc} (T^a T^c + T^c T^a) = 0 \end{aligned} \tag{70}$$

that follows since the structure constants are completely antisymmetric[1]. Since C_2 commutes with all the generators, it must be proportional to the identity in any given irreducible representation. This defines the number $C(R)$, the quadratic Casimir in the irrep R, by

$$T^a_R T^a_R = C(R)\, \mathbb{1}\ . \tag{71}$$

Setting $a = b$ in (68) and summing (i.e. taking the scalar product with the Killing metric) gives the relation

$$T(R) D(\mathrm{Adj}) = C(R) D(R)\ . \tag{72}$$

For the simplest representations one finds

$$D(N) = D(\bar{N}) = N \qquad T(N) = T(\bar{N}) = \frac{1}{2} \qquad C(N) = C(\bar{N}) = \frac{N^2-1}{2N} \tag{73}$$

$$D(\mathrm{Adj}) = N^2 - 1 \qquad T(\mathrm{Adj}) = N \qquad C(\mathrm{Adj}) = N\ . \tag{74}$$

Finally, it is useful to recall the concept of *invariant tensors*. They are defined to be tensors that remain invariant after group transformations. For example, denoting by ψ^i the vectors transforming in the defining representation of $SU(N)$, so that the upper index i is transformed by the defining matrices $U^i{}_j$ of $SU(N)$, then the Kronecker symbol δ^i_j is an invariant tensor

$$\delta^i_j \ \rightarrow\ \delta'^i_j = U^i{}_k (U^{-1,T})_j{}^l \delta^k_l = U^i{}_k (U^*)_j{}^l \delta^k_l = U^i{}_k (U^*)_j{}^k = \delta^i_j\ . \tag{75}$$

It tells that in combining the representation N with $\bar{N}$ there appears a scalar

$$N \otimes \bar{N} = 1 \otimes + \cdots \tag{76}$$

i.e. one can form the scalar $\psi^i \chi_i$ out of ψ^i and χ_i. Similarly, the completely antisymmetric tensor with N upper indices, $\epsilon^{i_1 i_2 \ldots i_N}$, normalized to one, $\epsilon^{12\ldots N} = 1$, is an invariant tensor

$$\epsilon'^{i_1 i_2 \ldots i_N} = \epsilon^{i_1 i_2 \ldots i_N} \tag{77}$$

known also as the Levi-Civita symbol. Indeed, one computes

$$\epsilon^{i_1 i_2 \ldots i_N} \ \rightarrow\ \epsilon'^{i_1 i_2 \ldots i_N} = U^{i_1}{}_{j_1} U^{i_2}{}_{j_2} \ldots U^{i_N}{}_{j_N} \epsilon^{j_1 j_2 \ldots j_N} = (\det U) \epsilon^{i_1 i_2 \ldots i_N} \tag{78}$$

but $\det U = 1$ for $SU(N)$, and the invariant property follows. Same thing for $\epsilon_{i_1 i_2 \ldots i_N}$.

[1] We have used that $[AB, C] = A[B, C] + [A, C]B$ for arbitrary operators.

Other invariant tensors are the generators in any given representation R, which we may write as $(T^a_R)^\alpha{}_\beta$, where the upper index α belongs to (the vectors of) the representation R and the lower index β to the conjugate representation $\bar{R}$ (see note[2]) This statement follows from the Lie algebra (60) by recognizing that the structure constants $f^{ab}{}_c$ give rise to the generators in the adjoint representation, that transforms the index a in $(T^a_R)^\alpha{}_\beta$. This also means that

$$R \otimes \bar{R} \otimes \mathrm{Adj} = 1 \oplus \cdots \ . \tag{79}$$

Moreover, since the adjoint is a real representation (the $f^{ab}{}_c$ are real numbers and thus the group elements $e^{i\alpha_a T^a_{\mathrm{Adj}}}$ are real) one may understand that

$$\mathrm{Adj} \otimes \mathrm{Adj} = 1 \oplus \cdots \tag{80}$$

that matches with the fact that the Killing metric δ^{ab} is an invariant tensor that can be used to construct scalar products (more generally the tensor δ^α_β for the arbitrary representations R and $\bar{R}$ is an invariant tensor). Then (79) and (80) imply

$$R \otimes \bar{R} = \mathrm{Adj} \oplus \cdots \tag{81}$$

which is interpreted by saying that $(T^a_R)^\alpha{}_\beta$ are Clebsch-Gordan coefficients: they combine the tensors in the representation R with those in the representation $\bar{R}$ to produce a tensor transforming in the adjoint. Said differently, Clebsch-Gordan coefficients are invariant tensors.

Finally, let us define another invariant tensor, the d^{abc} tensor, together with the anomaly coefficients $A(R)$ by

$$A(R) d^{abc} = \frac{1}{2} \mathrm{tr} \left(T^a_R \{ T^b_R, T^c_R \} \right) \tag{82}$$

where the overall normalization may be fixed by setting $A = 1$ for the fundamental representation. It is totally symmetric and appears in the study of chiral anomalies. The only simple groups that have a non-vanishing d^{abc} tensor, and therefore a cubic Casimir operator $C_3 \sim d^{abc} T^a T^b T^c$, are $SU(N)$ for $N \geq 3$ and $SO(6)$.

A.2 Cartan-Weyl basis

It is often useful to rewrite the generators of a Lie algebra in the Cartan-Weyl basis. This is defined by first finding the maximal number of generators (or independent linear combination of generators) H_i that commute between themselves

$$[H_i, H_j] = 0 \ . \tag{83}$$

This maximal number is called the *rank* of the group. They are taken to be hermitian, and they define the Cartan subalgebra of the Lie algebra. Since they commute, they can be diagonalized simultaneously in any given representation, and the eigenvalues are called the *weights*. This definition generalizes the angular momentum generator J_3 of $SU(2)$, which is a group of rank 1. J_3 is the generator that is usually diagonalized in quantum mechanics[3]. The particular weights of the adjoint representation are called *roots*.

[2] One may recall that given a representation R, one finds that $R^{-1,T}$, R^* and $R^{-1,\dagger}$ are also representations. These four representations acts on vectors $v^\alpha, v_\alpha, v^{\dot\alpha}, v_{\dot\alpha}$ belonging to the appropriate vector space. For unitary representations $v^{\dot\alpha} \sim v_\alpha$ and $v_{\dot\alpha} \sim v^\alpha$.

[3] Recall the SU(2) algebra: $[J_3, J_\pm] = \pm J_\pm$ and $[J_+, J_-] = 2J_3$.

The remaining generators are combined in complex combinations so that they correspond to the roots α_i

$$[H_i, E_\alpha] = \alpha_i E_\alpha \tag{84}$$

which can be interpreted by saying that α_i are eigenvalues and E_α are eigenvectors (the root α is a vector with components α_i). The generators E_α cannot be hermitian, but rather one has that $E_\alpha^\dagger = E_{-\alpha}$, so that if α is a root then also $-\alpha$ is a root. They generalize the $J_\pm$ angular momentum operators of $SU(2)$. Finally, one has the remaining structure constants that appear in calculating

$$[E_\alpha, E_\beta] \ . \tag{85}$$

The Jacobi identity can be used to study them, and in particular one finds that

$$[E_\alpha, E_{-\alpha}] = \alpha_i H_i \ . \tag{86}$$

which also generalizes the $SU(2)$ case.

This basis (and a related one called the Chevalley basis) is very useful in deriving general properties of Lie algebras, in a close analogy with the theory of angular momentum in quantum mechanics. In particular, it is useful to prove the complete classification of simple Lie algebras, due to Killing and Cartan. This classification is often encoded by the Dynkin diagrams of fig. 1. The algebras depicted there correspond to the following compact groups: $A_n = SU(n+1)$, $B_n = SO(2n+1)$, $C_n = Sp(2n)$, and $D_n = SO(2n)$, where n is the rank. The remaining algebras G_2, F_4, E_6, E_7, E_8 correspond to the so-called exceptional groups.

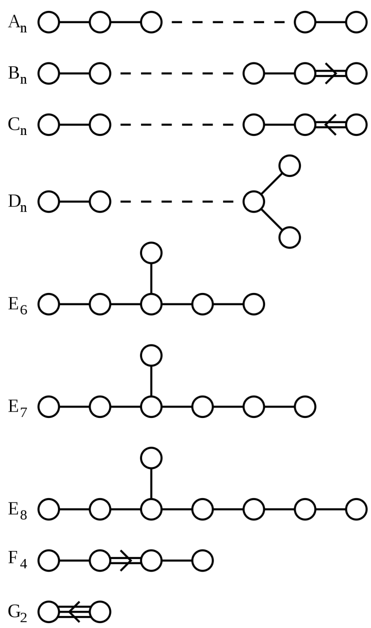


Figure 1: Dynkin diagrams