Astroparticle physics notes

Jacopo Tissino

2020-07-02

Contents

Tuesday 2020-3-10

Introduction

Professor Antonio Masiero, `antonio.masiero@unipd.it`

There are two courses, for Astrophysics and Cosmology and for Physics, which bear the same name. The other one is by Francesco D'Eramo: it assumes a knowledge of Quantum Field Theory.

This course, instead, only requires knowledge of Quantum Mechanics. The first part of this course is devoted to an introduction about the basics of Quantum Field Theory and gauge theories.

> "By the end of the 20th century [. . .] we have a comprehensive, fundamental theory of all observed forces of nature which has been tested and might be valid from the Planck length scale of 10^{-33} cm to the edge of the universe 10^{28} cm". David Gross, 2007.

The task in APP is to be able to discuss such a fundamental theory.

First of all, we need to address the two standard models: the ΛCDM model for cosmology and the Standard Model of particle physics.

There are points of friction between the two Standard Models. There are also several questions: neutrinos' mass, what caused inflation. . .

These problems have a common denominator: the interplay between particle physics, cosmology and astrophysics. What we seek is new physics, beyond the two standard models.

Books: Peskin [Pes19], "Concepts in Elementary Particle Physics". The book is addressed to students who are not experts in QFT and particle physics, rather, it provides the fundamental knowledge for these topics.

The exam is a colloquium, an oral exam, for which we can prepare a presentation on a specific topic. There is no issue if we do not precisely remember a specific formula, it is about going deep in the concepts.

An overview of the astroparticle physics landscape

Fundamental particles: the SM of particle physics Elementary particles make up ordinary matter. Fermions have spin 1/2, and are composed of quarks:

$$\begin{bmatrix} u & c & t \\ d & s & b \end{bmatrix}, \tag{1a}$$

leptons:

$$\begin{bmatrix} \nu_e & \nu_\mu & \nu_\tau \\ e & \mu & \tau \end{bmatrix} . \tag{2a}$$

The muon and tau particles are similar to electrons, but with higher mass.

These particles' interactions are mediated by 12 vector bosons, which have spin 1: these are "radiation" (the term is outdated).

1. gluons (g) mediate the strong nuclear interaction, there are 8 of them;
2. the $W^\pm$ and Z^0 bosons mediate the weak interaction;
3. the photon (γ) mediates the electromagnetic interaction.

There was a need for a mechanism to provide mass to the weak bosons and the fermions: this is accounted for by the Higgs boson, which is a scalar (that is, it has spin 0). This realizes the electroweak symmetry breaking.

The issue is that gravity is missing. In order to describe it in this scheme we would need a way to quantize it: all of these particles are actually excitations of quantum fields.

There are two marvelous 20th century theories, but they are not compatible.

Unification of interactions In 1687 Newton unified two domains of interactions: the terrestrial phenomena and the celestial phenomena, establishing the universality of gravitational interactions.

In 1865 Sir Maxwell unified electricity and magnetism into electromagnetism.

In 1967 Glashow, Weinberg and Salam propose the Standard Model of Particle Physics, unifying the Electromagnetic and Weak interactions. This is not a true unification: it is more appropriate to say that they are "mixed together" into the Electroweak interaction.

In the Standard Model, there is a kind of "frontier" around 100 GeV: below this energy, we see two interactions: the electromagnetic and the weak interaction. They are very much different: photons are massless, so the interaction has an infinite range, while the weak bosons are massive.

How can these be unified? We will see; above 100 GeV this apparent profound difference disappears in favour of the electroweak interaction. This is a phase transition.

Is the electroweak interaction above 100 GeV massless or not? Above this energy there is still a difference between the coupling constants of the two interactions. Above this energy, the W and Z bosons are no longer massive.

Above 100 GeV the strong interaction is separated from the electroweak one. Maybe there is an energy at which the electroweak interaction is unified with the strong one? We shall explore this topic: there are theories (Grand Unified Theories, GUT) in which there is such a unification. Also, from the 1980s there started to appear string theories, in which gravity is also unified to the other interactions, so that there is only one fundamental parameter, the "string tension".

As the energy increases, the coupling of the strongest interactions becomes weaker.

The energy scale, however, is very large: around 10^{16} GeV: this is a "science fiction" energy scale, it is extremely large. This is close to the Planck mass: $M_P \sim 10^{19}$ GeV, so we might not be able to describe this energy range with vanilla SM.

The Standard Model of Cosmology The standard model in cosmology describes a **Hot Big Bang**: the ancestral temperature was very high, and gradually decreased. Now we can work backwards in our energy scale: as time progresses forward from the Big Bang, the energy of particles decreases and our particles undergo various phase transitions.

The symmetry group of the Grand Unified Theory is broken, so we get subgroups; at each transition some symmetry is broken.

The EM + weak into electroweak transition is not speculative: we have observed it at the LHC. On the other hand, the electroweak + strong into GUT transition is speculative.

When, in the expansion of the universe, we reach an energy per particle of ≈ 1 GeV we have a new transition: the quark-hadron one, when free quarks become confined into hadrons such as protons and neutrons.

Around 1 MeV we have a new transition: nucleosynthesis, where protons and neutrons become confined into nuclei.

Then, at the eV scale, we reach recombination, which is when the radiation we see as the CMB is released. This is where nuclei and free electrons form hydrogen atoms.

There must be new physics somewhere: there is no room in the SM for dark matter, the matter-antimatter asymmetry, the mass of neutrinos.

Chapter 1

The particle physics Standard Model

For this lecture, a good reference is Peskin, chapter 2 [Pes19].

Wednesday 2020-3-11, compiled 2020-07-02

We wish to describe what we described yesterday as "matter" and "radiation".

The problem is similar to the one we have in classical mechanics, an initial value problem: given the positions and velocities of the particles at a certain starting time t_0 we wish to compute their state at a later time t.

This classical description in which the particles are not wavelike fails at the microscopic level: we want to give a quantum description of such a system of particles. We will derive it from the classical description using the standard tools of quantization. We start with a refresher of the classical description.

1.1 The classical description of a system of particles

Our aim is to compute and solve the equations of motion. The usual approach is to use Hamilton's variational principle: it is the principle of least action, but it is not usually referred to as such: we are actually not *minimizing* the action but finding a *stationary point* for it. This could also be a maximum or a saddle point.

The action functional S depends on the coordinates $q_i(t)$ of the particles at at time t, on the derivatives of these positions $\dot{q}_i(t)$ which represent the velocities of the particles at a time t. We usually write $S\left[q_i(t), \dot{q}_i(t)\right]$.

If we fix $q(t_0)$ and $q(t_f)$, the positions at some initial and final time t_f, we can then trace out a path $q(t)$ and perturb it by $\delta q(t)$; we fix $\delta q(t_0) = \delta q(t_f) = 0$.

Under this perturbation of the path $q \to q + \delta q$, the action changes to $S \to S + \delta S$. We then ask that $\delta S = 0$.

S is an action: is dimensions are those of an energy times a time. In terms of the Lagrangian L, the action is defined as

$$S\left[q_i(t), \dot{q}_i(t)\right] = \int_{t_0}^{t_f} L(q_i(t), \dot{q}_i(t)) \,\mathrm{d}t \,, \tag{1.1}$$

which means that the Lagrangian must have the dimensions of an energy.

Moving on from classical mechanics to classical field theory, we will make use of a quantity called the Lagrangian density $\mathscr{L}$, such that we recover the Lagrangian as:

$$L = \int \mathscr{L}\Big(\phi(\vec{x}), \partial_\mu \phi(\vec{x})\Big)\, \mathrm{d}^3 x \,. \tag{1.2}$$

From a finite number of particles we move to considering a field $\phi(\vec{x})$: this means that, in a certain sense, we are considering an "infinite number of particles", the values of the field at each point in space.

The dependence of the Lagrangian on the q_i and $\dot{q}_i$ shifted to a dependence on the spacetime coordinates x and their 4-derivatives $\partial_\mu x$. It could depend on many fields simultaneously, we omit this dependence for simplicity. Now, this Lagrangian density has the dimensions of an energy per unit volume.

Then, the action, computed in a region Ω of 4-dimensional spacetime, is

$$S = \int_\Omega \mathrm{d}^4 x\, \mathscr{L}\Big(\phi(x), \partial_\mu(\phi(x))\Big)\,. \tag{1.3}$$

Now that we have established the notation, we can apply the action principle: we consider an infinitesimal variation of the field $\phi \to \phi + \delta\phi$. We require this variation to vanish not only at the initial and final time, but over all the boundary $\partial\Omega$:

$$\delta\phi\Big|_{\partial\Omega} = 0\,. \tag{1.4}$$

Then, it can be shown with an integration by parts that imposing $\delta S = 0$ in the region Ω is equivalent to the Euler-Lagrange equations:

$$\frac{\partial \mathscr{L}}{\partial \phi(x)} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathscr{L}}{\partial\Big(\partial_\mu \phi(x)\Big)}\right) = 0\,. \tag{1.5}$$

If we have many fields ϕ_r, then we have a set of E-L equations for each of them. This is still classical: for example, classical (relativistic) electrodynamics is formulated in this way.

In order to write the Hamiltonian formulation of the theory we need the momenta, which are

$$\pi(x) = \frac{\partial \mathscr{L}}{\partial \dot{\phi}(x)}\,, \tag{1.6}$$

where a dot denotes a time derivative. Using these we define the Hamiltonian density by

$$\mathscr{H}(x) = \pi(x)\dot{\phi}(x) - \mathscr{L}(\phi, \partial_\mu \phi)\,, \tag{1.7}$$

and similarly to the Lagrangian we have the full Hamiltonian H

$$H = \int \mathrm{d}^3 x\, \mathscr{H}\,. \tag{1.8}$$

Now, as we quantize the theory the fields (ϕ and π) go from classical fields to Heisenberg-picture **operators**. As classical fields, ϕ and π satisfy Poisson bracket relations:

$$\left\{\phi(\vec{x},t), \pi(\vec{x}',t)\right\} = \delta^{(3)}(\vec{x},\vec{x}')\left\{\phi(\vec{x},t), \phi(\vec{x}',t)\right\} \qquad = 0 = \left\{\pi(\vec{x},t), \pi(\vec{x}',t)\right\}. \tag{1.9}$$

The operatorial version, instead, will need to satisfy the classical commutation relations, which are found by substituting the Poisson bracket with a commutator divided by $i\hbar$ [Tis20, section 2.4.2]:

$$\left\{\phi, \pi\right\} \to \frac{1}{i\hbar}\left[\hat{\phi}, \hat{\pi}\right]. \tag{1.10}$$

So, we get the following:

$$\left[\phi(\vec{x},t), \pi(\vec{x}',t)\right] = i\hbar\delta(\vec{x}-\vec{x}') \tag{1.11a}$$

$$\left[\phi(\vec{x},t), \phi(\vec{x}',t)\right] = \left[\pi(\vec{x},t), \pi(\vec{x}',t)\right] = 0\,. \tag{1.11b}$$

Note that the commutation relations are computed *at equal time*.

1.1.1 A relativistic reminder

The energy-momentum four-vector is

$$p^\mu = \begin{bmatrix} E \\ \vec{p}c \end{bmatrix}, \tag{1.12a}$$

where the greek index μ can take values from 0 to 3. A four-vector is an element of the tangent bundle to the spacetime manifold; concretely speaking, under a coordinate transformation which is locally linear and represented by a Lorentz matrix $\Lambda^\mu{}_\nu$ a four-vector transforms as $p^\mu \to \Lambda^\mu{}_\nu p^\nu$.

The metric signature used here is the the mostly minus one. So, $p^\mu q_\mu = E_p E_q - \vec{p}\cdot\vec{q}$, since we raise and lower indices using the metric $\eta_{\mu\nu}$.

The square norm of the 4-momentum is $p \cdot p = p^2 = E^2 - \left|\vec{p}\right|^2 c^2$. It is Lorentz invariant.

In the rest frame of the observer, $p^\mu = \left[E_0, \vec{0}\right]$, and this E_0 is just (c^2 times) the mass of the particle: this is the *definition* of mass.

When the relation is satisfied we have

$$p^2 = E^2 - \left|p\right|^2 c^2 = \left(mc^2\right)^2 \tag{1.13a}$$

$$E = c\sqrt{\left|p\right|^2 + (mc)^2}\,. \tag{1.13b}$$

When this relation is satisfied we say we are *on shell*: for virtual particles, instead, this might not be satisfied. This is allowed because virtual particles are described in a *quantum* theory: because of the uncertainty principle, we have uncertainty in energy if we consider short times and uncertainty in momentum if we consider small position intervals. This uncertainty means that we cannot enforce the on-shell condition as an exact equality if we are considering processes on these small scales, such as those which we find in quantum field theory.

We will use natural units: $\hbar = c = 1$.

This means that we equate energies (eV) and angular velocities (Hz); also we equate times (s) and lengths (m).

The rest energy of the electron is $m_e \approx 511\,\text{keV}$. Let us consider an electron with a momentum p equal to its mass m_e: then, its uncertainty in position is of the order

$$\frac{\hbar}{pc} = \frac{1}{m_e} \approx 4 \times 10^{-11}\,\text{cm}\,. \tag{1.14}$$

The dimensions of the lagrangian density, in natural units, are those of an energy to the fourth power, or a length to the -4, or a mass to the fourth.

Another useful exercise is to calculate the coupling of the electromagnetic field:

$$V(r) = \frac{e^2}{4\pi\epsilon_0 r} = \frac{e^2}{4\pi}\frac{1}{r}\,, \tag{1.15}$$

since we set $\epsilon_0 = \mu_0 = 1$. We can introduce the electromagnetic α: this is

$$\alpha = \frac{e^2}{4\pi} \times \frac{1}{\hbar c}\,. \tag{1.16}$$

This then becomes adimensional: $\alpha \approx 1/137$. It represents the strength of the electromagnetic interaction: the strength of the coupling of the photon to the electron. The fact that it is $\sim 10^{-2}$ is important: it allows us to work in a perturbative way, in powers of α.

What is the coupling of the strong and weak interactions? This will be discussed.

Next time, we will discuss symmetries and symmetry breaking.

1.2 Symmetries and conservation laws

Tuesday 2020-3-17, compiled 2020-07-02

Our aim is to describe the fundamental constituents of matter with a Quantum Field Theory. The method used to derive the equations of motion is a variational principle: we will find a Lagrangian density for various particles, and then apply the variational principle to find their equations of motion.

A guiding principle on the description of these fundamental particles is based on using their symmetries. We have Nöether's theorem in Quantum Field Theory: from these symmetries we are able to find conserved quantities.

These symmetries are described with groups, since we can compose their application; the theory describing groups is very rich. For this lecture we will base ourselves on Peskin's chapter 2 [Pes19].

A group G is a set of elements endowed with an operation. The set of elements can be either discrete or continuous. Examples of discrete transformations are the parity transformation P: $P\vec{x} = -\vec{x}$ and the time swap T: $Tx^\mu = (-x^0, \vec{x})$. Continuous symmetries, on the other hand, are parametrized by one or more continuous-valued parameters.

We distinguish:

1. **spacetime** symmetries: groups which transform our coordinate system for spacetime, such as Lorentz and Poincaré transformations;
2. **internal** symmetries: groups which transform a certain field, or a certain property of our quantum system.

For our set to be a group, we need to be able to define an operation — we will usually call it multiplication — between the elements of the group, such that if $a, b \in G$ then $ab \in G$. Also, we must have

1. associativity: $(ab)c = a(bc)$;
2. existence of the identity $\mathbb{1}$, such that $\mathbb{1}a = a\mathbb{1} = a$;
3. existence of inverses: there exists a^{-1} such that $aa^{-1} = a^{-1}a = \mathbb{1}$.

What is of interest to us is the association of the group with a transformation which is a symmetry: this is called a *representation*, which associates to each $g \in G$ a unitary operator U_g acting on the quantum states. We ask that this representation should preserve the group structure, that is to say, $U_{gh} = U_g U_h$ and $U_{g^{-1}} = U_g^{-1}$.

We call a transformation a symmetry if, after performing the transformation, the dynamics of the system do not change.

For a quantum mechanical system, we are interested in the observables: these are described by operators, whose eigenvalues are the observations we make, and which in the Heisenberg picture evolve like

$$-i\hbar \frac{\mathrm{d}}{\mathrm{d}t} O(t) = [H, O(t)] \,; \tag{1.17}$$

if an operator commutes with the Hamiltonian, $[H, O] = 0$, then the operator's expectation value on any state is constant — which is to say, the operator is constant.

If we perform a transformation in the form

$$|\psi\rangle \to |\psi'\rangle = U |\psi\rangle \,, \tag{1.18}$$

then the operators will change by

$$O \to O' = U^\dagger O U \,. \tag{1.19}$$

Note that whether we have $U^\dagger O U$ or $U O U^\dagger$ does not matter, since we ask observables O to be Hermitian, so $O = O^\dagger$.

We know that these transformations must always be unitary, because the conservation of probability implies that we must have $\langle \psi | \psi \rangle = \text{const}$: so,

$$U^\dagger U = \mathbb{1} \,. \tag{1.20}$$

This can be also stated as $U^\dagger = U^{-1}$.

So, the function associating a unitary operator U to an element g of the group is called its *unitary representation*.

A transformation G is a symmetry if $\forall a \in G$ we have

$$[U(a), H] = 0\,, \tag{1.21}$$

that is, the unitary representation of the group element always commutes with the Hamiltonian.

If we have a state $|\psi\rangle$ with energy $H|\psi\rangle = E|\psi\rangle$, then the transformation commuting with the Hamiltonian means that $|\psi'\rangle = U|\psi\rangle$ has the same energy:

$$H(U|\psi\rangle) = HU|\psi\rangle \overset{[H,U]=0}{=} UH|\psi\rangle = UE|\psi\rangle = E\Big(u|\psi\rangle\Big)\,, \tag{1.22}$$

so the eigenvalue of $U|\psi\rangle$ is the same as that of $|\psi\rangle$.

Now, we can move to an example, taken from Peskin [Pes19, eq. 2.38 onward]. Consider the discrete group $\mathbb{Z}_2$, which only has the elements 1 and -1, with the same multiplication rules as those we would have if these elements were integers. So, the group is closed with respect to multiplication. It can be easily checked that this is indeed a group based on our definition.

In order for this to be of interest to us, we can consider a quantum mechanical system and find a unitary representation acting on its Hilbert space.

Let us suppose we have a QM system with a basis made of two states $|\pi^+\rangle$ and $|\pi^-\rangle$. Let us define the *charge conjugation* operator C, by:

$$C\left|\pi^+\right\rangle = \left|\pi^-\right\rangle \qquad \text{and} \qquad C\left|\pi^-\right\rangle = \left|\pi^+\right\rangle\,. \tag{1.23}$$

So, we can find a unitary representation of $\mathbb{Z}_2$ in this system: we need to define $U(1)$ and $U(-1)$. We define

$$U(1) = \mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad U(-1) = \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\,, \tag{1.24a}$$

where the matrices are to be interpreted as acting on vectors expressed to the basis $\big\{|\pi_+\rangle\,, |\pi_-\rangle\big\}$.

So, we can say that our unitary representation looks like

$$\mathbb{Z}_2 \to \{\mathbb{1}, C\}\,. \tag{1.25}$$

Now, if $[C, H] = 0$ (and $\mathbb{1}$ commutes with H, which is always the case) then we say that "H has the symmetry $\mathbb{Z}_2$": this implies that the energies of the two $\pi_\pm$ particles are equal.

The interesting question to determine will be whether this is actually the case for our given group.

Groups can be subdivided into abelian and non-abelian ones. A group is abelian if for every a, b in G we have $ab = ba$, or equivalently, $[a, b] = 0$. It is not if this is not the case, that is, there exist a, b such that $ab \neq ba$.

The condition on the elements directly translates to a condition on the matrices of the unitary representation. If we have commuting matrices, we can simultaneously diagonalize them: for example, in the case of $\mathbb{Z}_2$ we can go to a basis in which

$$C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\,, \tag{1.26a}$$

specifically the states on which this matrix will act will need to be

$$|\pi_1\rangle = \frac{|\pi^+\rangle + |\pi^-\rangle}{\sqrt{2}} \qquad \text{and} \qquad |\pi_2\rangle = \frac{|\pi^+\rangle - |\pi^-\rangle}{\sqrt{2}}\,, \tag{1.27}$$

since then $C\,|\pi_1\rangle = |\pi_1\rangle$ (we write $C = +1$) and $C\,|\pi_2\rangle = -\,|\pi_2\rangle$ (we write $C = -1$). We will often use this notation, confusing operator and eigenvalue.

In the case of nonabelian groups it is not in general possible to diagonalize all the matrices; we can however do a change of basis and write the matrices as a block matrix with the smallest possible blocks:

$$U_R = \begin{bmatrix} U_1 & 0 & 0 \\ 0 & U_2 & 0 \\ 0 & 0 & \dots \end{bmatrix}, \tag{1.28a}$$

where the matrices U_i are called the **irreducible unitary representations of** G. The dimension of the matrices U_i tells us the dimension of these irreducible unitary representations.

Add more details on irreps — maybe not here? They can be found in professor Rigolin's intro to groups.

Do note that some elements of a nonabelian group can commute: for example, in the rotation group we have

$$[J^i, J^j] = \epsilon^{ijk} J_k\,, \tag{1.29}$$

so if we take $i = j$, that is, we consider rotations along the same axis, they will commute since then the Kronecker symbol is equal to zero.

1.2.1 Continuous transformations: space translations

An element of the group can be written as

$$U(a) = e^{-iaP}\,, \tag{1.30}$$

where the operator P, whose eigenvalue is the momentum, is called the generator of the transformation.

If we consider a plane wave we can clearly see how this action works: if we start from

$$\langle x|p\rangle = e^{ipx}\,, \tag{1.31}$$

we can apply the operator $U(a)$ to $|p\rangle$, which will yield e^{-ipa} (since eigenvectors of an operator are also eigenvectors of its exponential): so we find

$$\langle x|\, U(a)\, |p\rangle = e^{ip(x-a)}\,, \tag{1.32}$$

which means that by acting with this operator we have effectively performed a translation with displacement a.

If our system is invariant under translations, then Nöether's theorem tells us that the momentum is conserved.

In order to be a physical observable P needs to be Hermitian: $P = P^\dagger$.

So, the adjoint of the transformation $U(a)$ is

$$U^\dagger(a) = \sum_n \left(\frac{(-iaP)^n}{n!} \right)^\dagger = \sum_n \frac{(iaP^\dagger)^n}{n!} = e^{iaP^\dagger} = e^{iaP} = U^{-1}(a)\,, \tag{1.33}$$

which confirms the fact that the transformation is unitary.

Let us suppose that the momentum operator P commutes with the Hamiltonian: $[P, H] = 0$. Then,

$$[U(a), H] = 0\,, \tag{1.34}$$

that is, the Hamiltonian is translation-invariant.

All this is to say that a constant of motion O corresponds to an operator O which commutes with the Hamiltonian. This is formalized by Nöther's theorem, which establishes the equivalence between symmetries and conservation laws:

$$[O, H] = 0 \iff [U_O, H] = 0\,. \tag{1.35}$$

As an example, take the group G of 3D rotations. They depend on a continuous parameter $\vec{\alpha}$, just like translations depended on the parameter a.

The rotation is written as

$$U(\vec{\alpha}) = e^{-i\vec{\alpha}\cdot\vec{J}}\,, \tag{1.36}$$

where the components of the angular momentum have the following commutation relations:

$$[J^i, J^j] = i\epsilon^{ijk} J^k\,. \tag{1.37}$$

We will be able to compose the representations of rotations:

$$U(\vec{\beta})U(\vec{\alpha}) = U(\vec{\gamma})\,. \tag{1.38}$$

This space of 3D rotations is called SO(3), since every rotation corresponds to a 3x3 matrix which is a rotation matrix — it is orthogonal and has determinant 1.

Now, we seek **representations** of these rotations: so, we choose the dimension d of a quantum-mechanical vector and describe how it changes upon the action of the unitary matrices found by exponentiating certain d-dimensional generators J^i, which must have the algebra discussed above.

If we look for 1D representations of the generators J^i the only option we find is $J^i = 0$, which means that we are not actually performing a rotation. This is because scalars commute with each other. Which states transform this way? These are scalar states, with spin 0.

For 2D representations, we have

$$J^i = \frac{1}{2}\sigma^i\,, \tag{1.39}$$

where the σ^i are the Pauli matrices.

We can also find 3D representations, which look like

$$J^1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \qquad J^2 = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \qquad J^1 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{1.40a}$$

and represent a spin 1 particle. In general, spin s corresponds to a $2s+1$-dimensional representation.

A rotation in 2D, represented by an element of SO(2), corresponds to a phase shift, so we can say that it is equivalent to an element of U(1). This then allows us to see that SO(2) is abelian.

In general, we write for a unitary $n \times n$ representation

$$U(n) \to e^{-i\alpha^n t^a}\,, \tag{1.41}$$

where the generators t^a are Hermitian matrices corresponding to Hermitian operators. In particular, conventionally we say that one of these is the identity: $t^0 = \mathbb{1}$ (which must always be included in the group, lest we lose closure).

So, we omit it and say that we have n^2-1 generators for the SU(n) group. We shall see that each of these generators corresponds to a particle, and for the weak interaction we will have $2^2-1=3$ particles, while for the strong one we will have $3^2-1=8$.

In general, if t^a are the generators of an abstract Lie group, we can describe the algebra of the group by

$$\left[t^a, t^b\right] = i f^{abc} t^c\,, \tag{1.42}$$

so, the commutator is decomposed into a linear combination of the generators, whose coefficients f^{abc} are called the **structure constants** of the group.

Wednesday 2020-3-18, compiled 2020-07-02

Yesterday we mentioned the fact that for group symmetries we can find corresponding conservation laws.

Invariance under spacetime translations gives us the conservation of 4-momentum p^μ. Invariance under the Lorentz group gives us conservation of angular momentum (for rotation).

1.2.2 Discrete symmetries

The symmetry in which we reverse the spatial coordinates is called parity P, if we reverse the time coordinate we have time reversal T, and later we will discuss the internal symmetry of charge conjugation C.

More precisely, if $x^\mu = \left(x^0, \vec{x}\right)$ we have

$$P x^\mu = \left(x^0, -\vec{x}\right) \tag{1.43a}$$

$$T x^\mu = \left(-x^0, \vec{x}\right). \tag{1.43b}$$

These are Lorentz transformations represented by matrices with negative determinants: $\det \Lambda_P = \det \Lambda_T = -1$. This means that we cannot obtain these transformations with a composition of proper orthochronal Lorentz transformations — those defined to be continuously connected to the identity —, since the determinant of a Lorentz transformation is always ± 1, and it is a continuous function of the Lorentz transformation.

In quantum mechanics, these must be interpreted like operators. Their eigenvalues will be ± 1.

These transformations being symmetries is an experimental question. P and T are symmetries of the strong and electromagnetic interactions.

The weak interaction, instead, does not conserve P.

We defined the charge conjugation operator C by its action on the $\left|\pi^+\right\rangle$ and $\left|\pi^-\right\rangle$, on the basis defined by them its matrix representation was σ_x. It commuted with the Hamiltonian.

In general, this swaps not only the electric charge but all the quantum numbers, including the "hypercharges" such as baryon number. The charge conjugation operator C gives us the antiparticle of a certain particle, so it must flip all of the charges. There are, however, theories in which, say, lepton number conservation is broken but baryon number is conserved. The precise meaning of this operator depends on the theory.

If we constructed our experiments with antimatter would we get the same physics?

It was experimentally determined that weak interactions violated parity, and it was thought that baryon number was conserved: however in the early universe we would expect to have equal amounts of matter and antimatter, but this is not what we see — we are made of matter, and we do not see tha gamma ray background we would expect to see if there were matter and antimatter spacetime regions.

Another important concept is that of the *intrinsic parity* of a particle: consider a state of a particle A with momentum $\vec{k}$, denoted as $\left|A(\vec{k})\right\rangle$. Then, upon application of P we can have

$$P\left|A(\vec{k})\right\rangle = \pm\left|A(-\vec{k})\right\rangle , \tag{1.44}$$

where the sign depends on the properties of the particle.

We can apply several "mirrors" to our system: this amounts to the composition of the operators. There is a theorem in Quantum Field Theory: if a QFT is consistent, then it must obey CPT symmetry. After passing through all of the three mirrors, the system has the same physical properties. The order of the three symmetries does not matter: they commute.

The conservation of the electric charge Q, for example, is connected to $\mathrm{U}(1)_{\mathrm{em}}$ symmetry. This is an *exact* symmetry, which we expect not to break down at higher energy.

On the other hand, C, P and T are not exact symmetries since there exist some interactions which violate them (separately: P and CP are violated, so T is also violated by the CPT theorem).

1.3 Relativistic wave equations

In quantum mechanics we describe the dynamics of a quantum system using the Schrödinger equation:

$$E = \frac{\vec{p}^2}{2m} + V\,, \tag{1.45}$$

where $E = i\partial_t \psi$ and $\vec{p}^2 = -\vec{\nabla}^2\psi$.

This is non (special) relativistic: if we perform a Lorentz boost the equation does not remain in the same form. This is explicit, in that we have an addition of first time derivatives and second space derivatives, while in SR time and space are on the same footing.

Also, this equation describes the dynamics of one electron. In elementary particle physics it is no longer *consistent* to only consider one particle: when the energies are of the order of the mass of the particles we can create antiparticles or other particles. So, in general there is no reason to expect that the number of particles is conserved in particle physics.

Formally, it is hard to define from the Schrödinger equation a conserved quantity connected to the probability of finding the particle. It does not account for the possibility that the particle might appear or disappear. For example, we can have matter-antimatter collisions: this is called *annihilation*, and the two incoming particles disappear completely.

The equation $E = \vec{p}^2/2m$ is intrinsically nonrelativistic; the corresponding relativistic relation is $E^2 = \vec{p}^2 + m^2$.

We can make the same canonical quantization substitutions to get

$$\left(\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + m^2\right)\phi(t, \vec{x}) = 0\,. \tag{1.46}$$

Why should we use $E^2 = \vec{p}^2 + m^2$ instead of, say, $E = \sqrt{\vec{p}^2 + m^2}$? The problem is that it is hard to see what could be meant by the square root of an operator. Anyways, later we shall use the approach of "taking the square root" in order to solve the KG equation. We can write this in a manifestly covariant way as

$$\left(\partial^\mu \partial_\mu + m^2\right)\phi(t, \vec{x}) = 0\,, \tag{1.47}$$

where $\partial_\mu = (\partial_t, \vec{\nabla})$, whose square $\partial^\mu\partial_\mu = \partial_t^2 - \vec{\nabla}^2 = \Box$ is the Dalambertian.

This is the Klein-Gordon equation. As long as ϕ is a scalar, this is a scalar Lorentz invariant equation.

It should be used to describe a spin-0 particle, since we are not accounting for spin: the only spin-0 elementary particle known is the Higgs Boson. All other spin-0 particles are composite.

A crucial fact in the KG equation is the fact that the energy can in principle be both positive and negative: $E = \pm\sqrt{\vec{p}^2 + m^2}$. A seemingly reasonable approach would be to not bother treating the "unphysical" $E < 0$ solution; however this is wrong, the negative energy solution is important.

This was a great open debate last century. We will be given the solution, but it would not have easy to figure it out.

This is the reason why people started discussing antiparticles.

An interesting question now is: can we define an action

$$S[\phi(x)] = \int \mathrm{d}^4x \, \mathscr{L}(\phi, \partial_\mu \phi) \tag{1.48}$$

whose actions of motion are the KG equation? The answer is yes, with

$$\mathscr{L} = \frac{1}{2}\left(\partial^\mu \phi \partial_\mu \phi - m^2\phi^2\right). \tag{1.49}$$

To show that this is the case is left as an exercise. Notice that we talk of "Lagrangians" but we always mean Lagrangian densities. The corresponding Hamiltonian density $\mathscr{H}$ has a vacuum state we can call $|0\rangle$, which corresponds to the absence of particles.

We will then have states describing n particles of mass m.

Next time, we will move from the KG equation with ϕ being just a wavefunction to it being an operator, which can act on the vacuum creating particles.

This is sometimes called "second quantization", to distinguish it from the first quantization, in which operators act on wavefunctions.

Tuesday 2020-3-24, compiled 2020-07-02

Last week, we moved from the usual Schrödinger equation to the Klein Gordon equation by moving from $E = p^2/2m$ to $E^2 = p^2 + m^2$. The latter is covariant under Lorentz transformations.

There are, however, two issues with the Schrödinger equations: one is that it is not relativistic, while the second one is more subtle.

The Schrödinger equation describes the evolution of a single particle in time, while when we deal with elementary particles a one-particle description is unsuitable: we must have conservation of probability, so we are unable to describe situations in which a particle disappears by decaying into other particles, or we have inelastic collisions.

This is the reason why we need a multiparticle description. An example of the issues which arise in the single-particle relativistic description is the Klein paradox.

So, we will introduce the so-called second quantization formalism. We are going to interpret ϕ as a **quantum field operator**, instead of a scalar field. Let us make this explicit. We take the KG equation from the first quantization to the KG equation describing a Quantum Field Theory.

$\phi(x)$ will be an operator which can destroy or create particles. We start from the quantization of the Hamiltonian density, $\mathscr{H}$: we consider its ground state $|0\rangle$. This corresponds to a state in which there is no particle, and is called the **vacuum** state.

We will also have higher-energy states, in which we will have one or more particles: how do we describe these? A one particle state $\left|\varphi_1(p)\right\rangle$ can be acted upon by the operator

$\phi(x)$: it is destroyed, yielding a state which is proportional to the vacuum. Formally, we have

$$\langle 0| \phi(x) |\varphi_1(p)\rangle = e^{-ipx}, \tag{1.50}$$

where p^0 is the positive energy: $p^0 = +E_p = +\sqrt{\vec{p}^2 + m^2}$. For a more in-depth discussion of the second-quantization formalism, see the Theoretical Physics notes [Tis20, section 2.4].

Now, let us consider the complex conjugate of this matrix element: we swap the bra and ket and take the adjoint of the operator, to get

$$\langle \varphi_1(p)| \phi^\dagger(x) |0\rangle = e^{ipx}. \tag{1.51}$$

If we interpret $\phi^\dagger$ as acting on the right on the vacuum, we must say that it creates a particle at the location x, with indeterminate momentum.

This means that, if the first equation describes a particle propagating with momentum p^μ, this new equation will now describe a particle propagating with momentum $-p^\mu$.

This will now have a negative energy. This was a problem historically, now we give the solution directly.

The field ϕ can be either real or complex: that is, ϕ can be either equal to $\phi^\dagger$ or not. Let us consider the complex case: we introduce a particle $|\varphi_2(p)\rangle$, such that it is destroyed by the operator $\phi^\dagger$:

$$\langle 0| \phi^\dagger(x) |\varphi_2(p)\rangle = e^{-ipx}, \tag{1.52}$$

so now this particle has the same mass m, but — as it is shown by Peskin [Pes19, sec. 3.5], this new particle $|\varphi_2\rangle$ differs from $|\varphi_1\rangle$ for the charge, which is now opposite.

This means that the complex field describes a **charged** particle. The particle described by $\phi^\dagger$ is called the **antiparticle** of that described by ϕ.

Do note that charge does not exclusively mean electric charge! We can also have other kinds of charges. For example, neutrinos have the charge of lepton number.

Particles can be their own antiparticles, if they have zero charge. So, we can interpret the negative energy solution as the presence of an antiparticle: a negative energy particle would correspond to a positive energy antiparticle.

A generic field theory is defined by its action, which is written from the density Lagrangian: a free Lagrangian is

$$\mathscr{L} = \frac{1}{2}\left(\partial^\mu \phi \partial_\mu \varphi - m^2 \varphi^2\right), \tag{1.53}$$

and if we impose $\operatorname{var}(S) = 0$ we get precisely the KG equation as our equation of motion.

1.3.1 Spin 1

We discuss 3D vectors V^i, which transform under rotations $R^i{}_j \in SO(3)$ as

$$V^i(x) \to V'^i(x') = R^i{}_j V^j(R^{-1}x). \tag{1.54}$$

What? x is controvariant...

Now, the matrix element from before reads

$$\langle 0| V^i(x) |v(p,\epsilon)\rangle = \epsilon^i e^{-ipx}\,, \tag{1.55}$$

where we need to account for the momentum p and the polarization ϵ^i. If we want to move to a relativistic description, we get the same thing, with the spatial index i being replaced by a 4-dimensional index μ:

$$\langle 0| V^\mu(x) |v(p,\epsilon)\rangle = \epsilon^\mu e^{-ipx}\,. \tag{1.56}$$

The problem is now the normalization of these states: what is the value of $\langle v|v\rangle$? This will be proportional to $\epsilon^\mu\epsilon_\mu$; but in general this is neither positive definite nor negative definite, since ϵ^μ could be timelike or spacelike *a priori*. We know that the photon has two physical helicities, which are transverse degrees of freedom. Since we know that those are physical and spacelike (with negative norm in our convention), we are tempted to say that the probability must be a positive multiple of $-\epsilon^\mu\epsilon_\mu > 0$: but then the timelike polarization has negative probability, and the longitudinal one does not belong (since it is known from classical electromagnetism that the photon only has the two transverse ones). The longitudinal polarization and the timelike one must be somehow forbidden.

This is in general a problem. It can be solved by a hack which is called the Gupta-Bleuer condition [Tis20, sec. 2.8].

The electromagnetic four-potential is defined as

$$A^\mu = \left(\varphi(x), \vec{A}\right), \tag{1.57}$$

and the EM field strength is $F^{\mu\nu} = 2\partial^{[\mu}A^{\nu]}$. So, we have

$$F^{i0} = -\nabla^i\varphi - \partial_t A^i = E^i \tag{1.58a}$$

$$F^{ij} = -2\nabla^{[i}A^{j]} = \epsilon^{ijk}B^k\,. \tag{1.58b}$$

The Maxwell equations follow from the density Lagrangian

$$\mathscr{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j^\mu A_\mu\,, \tag{1.59}$$

where j^μ is an external current. This yields $\partial_\mu F^{\mu\nu} = j^\nu$.

The photon is a massless vector boson. This description applies in general to a massless vector field. Are there massive vector fields? Yes, the weak interaction is described by a massive vector boson $W^\pm_\mu$.

1.3.2 Spin 1/2

Now we discuss spin 1/2 particles: we will need to have first derivatives on either side, as opposed to the second equations in the KG equation.

Our ansatz for what will be called the Dirac equation is:

$$i\partial_t = -i\vec{\alpha} \cdot \vec{\nabla} + \beta m \,. \tag{1.60}$$

Is it possible to find four numbers (3 encoded in the vector $\vec{\alpha}$, one more in β) so that the square of this relation is $E^2 = p^2 + m^2$ and that we still retain Lorentz invariance?

If we try to impose these conditions, we find that there are no solutions: there are no such four numbers.

We can, however, find a solution if we allow $\vec{\alpha}$ and β to be matrices: specifically, 4×4 matrices, which must obey the **anticommutation relations**:

$$\left\{\alpha_i, \alpha_j\right\} = 2\delta_{ij} \qquad \text{and} \qquad \left\{\alpha_i, \beta\right\} = 0 \,, \tag{1.61}$$

where we used the commutator: $\{a, b\} = ab + ba$, and now we define

$$\gamma^0 = \beta \qquad \text{and} \qquad \gamma^i = \beta \alpha^i \,, \tag{1.62}$$

which must be 4D, as we said, and the simplest representation is called the Dirac representation:

$$\gamma^0 = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{bmatrix} \qquad \text{and} \qquad \gamma^i = \begin{bmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{bmatrix} \,. \tag{1.63a}$$

It can be verified that then

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \,. \tag{1.64}$$

We can move between representations of these matrices using unitary transformations. Inserting these, we find:

$$\left[i\gamma^0 \frac{\partial}{\partial t} + i\vec{\gamma} \cdot \vec{\nabla} - m\mathbb{1}\right] \varphi(x) = 0 \,, \tag{1.65}$$

which means that the wavefunction must be 4-dimensional as well, since we are acting on it with a 4D operator. We are going to call such an object a **spinor**. If we denote $\gamma^\mu = \left(\gamma^0, \vec{\gamma}\right)$ we can write the Dirac equation as

$$\left(i\gamma^\mu \partial_\mu - m\right) \varphi(x) = 0 \tag{1.66a}$$

$$\left(i\not{\partial} - m\right) \varphi(x) = 0 \,, \tag{1.66b}$$

where we have defined the notation $\not{x} = \gamma^\mu x_\mu$. As we shall see tomorrow morning, this equation is extremely rich in structure.

In the Westminster abbey, this equation is inscribed as a homage to Paul Dirac.

Wednesday 2020-3-25, compiled 2020-07-02

Today we are going to examine the third case of wave equations: we discussed the KG equation for a scalar particle, then we moved to a vector particle, and now we are going to consider a spin-1/2 particle.

One has to prove that the equation $\left(i\partial\!\!\!/ - m\right)\varphi(x) = 0$ is Lorentz-covariant. Also, solutions of the Dirac equation are also solutions of the Klein-Gordon equation, therefore they satisfy $E^2 = p^2 + m^2$. The latter is relatively easy to prove: we apply the operator $i\partial\!\!\!/ + m$ to the Dirac equation and find

$$\left(i\partial\!\!\!/ + m\right)\left(i\partial\!\!\!/ - m\right)\psi(x) = 0 \tag{1.67}$$

$$\left(-\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu - m^2\right)\psi(x) = 0 \tag{1.68}$$

$$(\Box + m^2)\psi(x) = 0\,, \tag{1.69}$$

where we used the fact that, since the derivatives $\partial_\mu\partial_\nu$ commute, we can substitute $\gamma^\mu\gamma^\nu$ with half of their anticommutator, which equals $\eta^{\mu\nu}$.

Since we have proven the equivalence, we expect the Dirac equation to also have negative energy solutions.

Proving the covariance is harder, even though the equation looks covariant, since γ^μ does not transform as a vector *a priori*.

The object $\psi(x)$ is a spinor: it is neither a scalar, nor a vector. A priori we can say that its transformation law will look like

$$\psi(x) \to \psi'(x) = S(\Lambda)\psi(x)\,, \tag{1.70}$$

where $S(\Lambda)$ is a unitary transformation associated with the Lorentz transformation Λ, and by imposing the covariance of the Dirac equation we find that we must have

$$S(\Lambda)\gamma^\mu S^{-1}(\Lambda) = \left(\Lambda^{-1}\right)^\mu_{\ \nu}\gamma^\nu\,. \tag{1.71}$$

So, in order to find the explicit form of $S(\Lambda)$ we write an infinitesimal Lorentz transformation: it can be shown that we can write it as

$$\Lambda^\mu_{\ \nu} = \eta^\mu_{\ \nu} + \delta\omega^\mu_{\ \nu}\,, \tag{1.72}$$

where $\delta\omega^\mu_{\ \nu}$ is an antisymmetric matrix, if it has nonzero $0i$ components it gives boosts, if it has nonzero ij components it gives rotations. The result for the form of S is:

$$S = \mathbb{1} + \frac{1}{8}\left[\gamma_\mu, \gamma_\nu\right]\delta\omega^{\mu\nu}\,. \tag{1.73}$$

If we make a rotation, for example, we get

$$\psi(x) \to \exp\left(i\frac{i}{2}\left[\gamma_i, \gamma_j\right]\delta\omega^{ij}\right)\psi(x)\,, \tag{1.74}$$

and typically one uses the shorthand rotation

$$\frac{i}{2}\left[\gamma_i, \gamma_j\right] \stackrel{\text{def}}{=} \sigma_{ij}\,. \tag{1.75}$$

If, for example, we want to perform a rotation by an angle φ around the z axis we get

$$\psi'(x') = S(\Lambda)\psi(x) = \exp\left(\frac{i}{2}\varphi\sigma_{12}\right)\psi(x)\,, \tag{1.76}$$

where

$$\sigma_{12} = \frac{i}{2}\left[\gamma^1, \gamma^2\right] = \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix}, \tag{1.77}$$

where we used the fact that $\left[\sigma_i, \sigma_j\right] = i\epsilon_{ijk}\sigma_k$.

This means that a spinor $\psi(x)$ reacts in a peculiar way to rotations: it rotates by an angle $\varphi/2$ if we perform a Lorentz rotation of an angle φ; its periodicity is 4π.

We introduce the *chiral representation* of the gamma matrices:

$$\gamma^0 = \begin{bmatrix} 0 & -\mathbb{1} \\ -\mathbb{1} & 0 \end{bmatrix} \qquad \text{and} \qquad \vec{\gamma} = \begin{bmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{bmatrix}. \tag{1.78a}$$

Then, for Lorentz boosts we have

$$\sigma_{0i} = \frac{1}{2}[\gamma_0, \gamma_i] = -i\begin{bmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{bmatrix}, \tag{1.79a}$$

while for rotations we have

$$\sigma_{ij} = \frac{i}{2}\left[\gamma_i, \gamma_j\right] = \epsilon_{ijk}\begin{bmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{bmatrix}, \tag{1.80a}$$

which is useful since it gives us block-diagonal matrices. So, we can interpret the spinor as being made up of two components:

$$\psi(x) = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} = \begin{bmatrix} \eta \\ \omega \end{bmatrix}, \tag{1.81a}$$

on which Lorentz transformations act independently. This is relevant since, for example, if we deal with an electron, we will describe it with a 4-component spinor, however we will be able to divide it into two components e_L and e_R, which are two component spinors on which we can act independently. We have effectively divided our representation of the Lorentz group into the sum of two irreps, of dimension $(1/2, 0)$ and $(0, 1/2)$ respectively.

This will become very concrete when we will discuss how many degrees of freedom were present in the original plasma.

Since a spinor ψ also solves the KG equation, it will be able to be written as

$$\psi = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} e^{-ipx}, \tag{1.82a}$$

but what are the relations between the coefficients? We consider a simple case, that of $\vec{p} = 0$, so that we are in the rest frame of the particle.

Then, the Dirac equation reads:

$$\left(\gamma^0 E - m\right) \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0\,, \tag{1.83a}$$

so we need to choose a representation for the γ^0 in order to write this explicitly. We choose the Dirac representation, in which

$$\gamma^0 = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{bmatrix}. \tag{1.84a}$$

Then, the equation reads

$$\begin{bmatrix} E-m & 0 & 0 & 0 \\ 0 & E-m & 0 & 0 \\ 0 & 0 & -E-m & 0 \\ 0 & 0 & 0 & -E-m \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} e^{-ipx} = 0\,, \tag{1.85a}$$

so we get a solution in which $E = m$, and a solution in which $E = -m$. So, in general we write the two linearly independent solutions, respectively with positive and negative energy, as

$$\psi = \begin{bmatrix} \xi \\ 0 \end{bmatrix} e^{-imt} \qquad \text{and} \qquad \Psi = \begin{bmatrix} 0 \\ \eta \end{bmatrix} e^{+imt}. \tag{1.86a}$$

The negative energy solution, as we will see, represents the antiparticle of the Dirac fermion.

The assumption we made, $\vec{p} = 0$, does not actually mean we lose generality: we can simply boost into the rest frame of the particle. If we do this, we get the general

$$\begin{bmatrix} E-m & -\vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -E-m \end{bmatrix} \begin{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\ \begin{bmatrix} c \\ d \end{bmatrix} \end{bmatrix} e^{-imt+i\vec{p}\cdot\vec{x}} = 0\,, \tag{1.87a}$$

so we can decompose our solution into

$$\psi = u^s e^{-ipx} \qquad E > 0 \tag{1.88}$$

$$\psi = v^s e^{ipx} \qquad E < 0\,, \tag{1.89}$$

where s is an index denoting which 2D unit vector we are considering, that is, $s = 1, 2$. So, now we come to the interpretation: we introduce the existence of an antifermion, which corresponds to the solution to the Dirac equation with negative energy, but it has the opposed momentum: so, it has positive energy. Then, both of our solutions have positive energy and evolve forward in time, and both have the same mass.

Now, we make the jump to second quantization: we start interpreting $\psi(x)$ as an operator, which can destroy a one-state particle or create a particle starting from the vacuum.

We will have

$$\langle 0| \psi(x) \left| e^-(p,s) \right\rangle = u^s(p) e^{-ipx} \,. \tag{1.90}$$

Now, the one-particle state $\left|e^-(p,s)\right\rangle$ is promoted to a spinor. On the other hand, we have the creation operator $\psi^\dagger$:

$$\left\langle e^-(p,s) \right| \psi^\dagger(x) \left|0\right\rangle = u^{s\dagger} e^{-ipx} \,. \tag{1.91}$$

Now, the tricky question is to introduce the negative-energy solution. The $\psi^\dagger$ operator will destroy this state, while ψ will create it. So we will write an equation like

$$\langle 0| \psi^\dagger(x) \left| e^+(p,s) \right\rangle = v^{s\dagger}(p) e^{-ipx} \,, \tag{1.92}$$

where we would write $v^s(p)$ if we were considering the negative energy particle, instead we are looking at the antiparticle.

Now, we will be able to operate with $\psi(x)$ on the vacuum, to find

$$\left\langle e^+(p,s) \right| \psi(x) \left|0\right\rangle = v^s(p) e^{ipx} \,. \tag{1.93}$$

1.3.3 Photon-fermion coupling

Yesterday we discussed the Lagrangian of the free photon field,

$$\mathscr{L} \propto F^{\mu\nu} F_{\mu\nu} \,, \tag{1.94}$$

but as we said there can also be coupling to external currents, which we did not quantize. However, now we quantized the electron: so, can we construct the external current j_μ in the coupling term $j_\mu A^\mu$?

The first attempt would be to write something like

$$j^\mu \sim \psi^\dagger \gamma^\mu \psi \sim e^+ \gamma^\mu e^- \,, \tag{1.95}$$

but we would need to check whether it is a vector: in fact, it does not transform correctly.

Is this at least a Hermitian operator? Well, its adjoint is

$$\left(\psi^\dagger \gamma^\mu \psi\right)^\dagger = \psi^\dagger (\gamma^\mu)^\dagger \psi \,, \tag{1.96}$$

but

$$(\gamma^\mu)^\dagger = \left(\gamma^0, -\gamma^i\right) \neq \gamma^\mu \,. \tag{1.97}$$

One finds that the correct definition is to have

$$\overline{\psi} \stackrel{\text{def}}{=} \psi^\dagger \gamma^0 \,, \tag{1.98}$$

and then

$$j^\mu = \overline{\psi} \gamma^\mu \psi \tag{1.99}$$

is the correct definition. This, then, transforms as a 4-vector; it can also be shown starting from Dirac's equation (and its conjugate) that this is a conserved current: $\partial_\mu j^\mu = 0$.

Now, in order to couple the EM field to the electron we will use minimal coupling:

$$\partial_\mu \to \mathrm{D}_\mu = \partial_\mu + ieA_\mu \,, \tag{1.100}$$

so the Dirac equation will read

$$(i\not{\mathrm{D}} - m)\psi = 0 \tag{1.101a}$$

$$\Big(i\gamma^\mu \Big(\partial_\mu + ieA_\mu \Big) - m \Big) \psi = 0 \,. \tag{1.101b}$$

From which density Lagrangian can we derive the Dirac equation? It turns out to be

$$\mathscr{L} = \overline{\psi} \Big(i\gamma^\mu \mathrm{D}_\mu - m \Big) \psi \,, \tag{1.102}$$

so if we want to describe both the EM field, the electron and their interaction, we have

$$\mathscr{L}(e, A_\mu) = \underbrace{-\frac{1}{4} F^{\mu\nu} F_{\mu\nu}}_{\text{free EM field}} + \underbrace{\overline{\psi}(i\not{\partial} - m)\psi}_{\text{electron}} - \underbrace{e\overline{\psi}\not{A}\psi}_{\text{electron-EM coupling}} \tag{1.103}$$

$$= -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \overline{\psi}(i\not{\mathrm{D}} - m)\psi \,. \tag{1.104}$$

This is the density Lagrangian of Quantum ElectroDynamics. This is the first interacting QFT which was constructed, and it was extraordinarily successful.

Its predictions for the anomalous magnetic moment of the electrons were exceptional: we can solve it perturbatively to different orders in

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137} \,. \tag{1.105}$$

Tuesday 2020-3-31, compiled 2020-07-02

We have associated spin $1/2$ fermions to matter fields. Spin 1 particles, instead, are vector bosons, such as the photon γ, the gluon g and the weak $W^\pm$ and Z^0 bosons. These are the radiation fields. We will explore this in more detail.

This connection comes from the **spin-statistics** theorem: it states that particles with integer spin obey Bose-Einstein statistics, while particles with half-integer spin obey Fermi-Dirac statistics. Particles which are "matter" (electrons, quarks and such) are fermions, while particles which are "force carriers" (photons, weak-interaction W and Z particles, gluons) are bosons.

This is due to the fact that in order to consistently quantize Dirac's theory we need to use **anticommutators** instead of field commutators to replace the Poisson brackets of the classical theory.

Up until now, we have observed only particles with spins 0, 1/2 and 1.

We could have a symmetry called supersymmetry, which connects fermions and bosons.

The graviton has spin 2; in supergravity the graviton has a fermion partner called the "gravitino" with spin 3/2.

1.3.4 Scattering

How do we normalize the states in relativistic theory? Classically we did

$$\langle p_1 | p_2 \rangle = (2\pi)^3 \delta^{(3)}(\vec{p}_1 - \vec{p}_2)\,, \tag{1.106}$$

in the relativistic case instead we will do

$$\langle p_1 | p_2 \rangle = 2E_{p_1} (2\pi)^3 \delta^{(3)}(\vec{p}_1 - \vec{p}_2)\,. \tag{1.107}$$

This is a Lorentz invariant normalization (although not manifestly so — see Peskin [Pes19, sec. 3.5] for a proof).

The relativistic volume element is given by

$$\int \frac{\mathrm{d}^3 p}{(2\pi)^3} \to \int \frac{\mathrm{d}^4 p}{(2\pi)^4} (2\pi) \delta(p^2 - m^2) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_p}\,, \tag{1.108}$$

where we integrated over the variable $p^0 = E$, which removed the $\delta(p^2 - m^2) = \delta(E - \sqrt{p^2 + m^2})/2E$.

This way, we have the completeness relation

$$\int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_p} \, |p\rangle\langle p| = \mathbb{1}\,. \tag{1.109}$$

States have the dimension of an energy to the -1, field operators have the dimension of an energy.

There are two main kinds of processes which are considered in particle physics. The first is a **decay** process: we have a particle A decaying into a possibly multi-particle state f. We are interested in the decay *rate* of this process.

The probability of survival of particle A at time t is in the form $\mathbb{P}(t) = \exp(-t/\tau_A)$, so we define the decay rate

$$\Gamma_A = \frac{1}{\tau_A}\,, \tag{1.110}$$

which has the dimensions of a frequency, or equivalently an energy.

Generally the decay of a certain particle species can happen through different **channels**, that is, into different kinds of particles. We can define the branching ratios as

$$BR(A \to f) = \frac{\Gamma(A \to f)}{\Gamma_A}\,. \tag{1.111}$$

Another process of interest is a **scattering** process of $n \to m$ particles. There is no particle number conservation: we can create and destroy as many particles as we like. These types of processes are described by their *cross section*, which is an effective area corresponding to how aligned the trajectories of the incoming particles must be in order for them to interact.

This cross section allows us to compute the average time for an interaction to occur: if two particles' interaction becomes so rare that they cannot interact within a Hubble time then they are said to have *decoupled*.

We start by considering fixed-target experiments: we want to know how many events per second we will have, which will be given by

$$\frac{\text{\# events}}{\text{second}} = n_A \times v_A \times \sigma \,, \tag{1.112}$$

where σ is the cross section, v_A is the velocity of the incoming particles, while n_A is the number density of particles in the beam. The number density of the particles in the target is accounted for inside of σ. This tells us that σ has the dimensions of an area.

If we have two beams of particles coming towards each other, the term will look like $n_A n_B (v_A - v_B) \ell_B A_b \sigma$.

In general, we will be interested in the differential cross section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}^3 p_1 \, \mathrm{d}^3 p_2 \dots \mathrm{d}^3 p_n} \,. \tag{1.113}$$

This is just a definition, and it is not covariant, if we want to integrate we still need to use the covariant momentum element. We can integrate it in $\mathrm{d}^3 p_1 \dots \mathrm{d}^3 p_n$ in order to recover the total cross section, but inside it we have more information about the angular properties of the process.

For a scattering process like $A + B \to 1 + \dots + n$ we will need to compute things like

$$\langle 12 \dots n | \, T \, | AB \rangle = \mathcal{M}(A + B \to 1 + \dots + n)(2\pi)^4 \delta^{(4)}(P_A + P_B - \sum p_i) \,, \tag{1.114}$$

where T is the time evolution, and we defined the invariant scattering amplitude $\mathcal{M}$. If we want to compute the width Γ_A for a decay process, we need to define the phase space integral:

$$\int \mathrm{d}\Pi_n = \underbrace{\prod_i \int \frac{\mathrm{d}^3 p_i}{2E_i (2\pi)^3} (2\pi)^4 \delta^{(4)}(P_A - \sum_i p_i)}_{\text{phase space integral}} \,, \tag{1.115}$$

so we will have what is called Fermi's golden rule:

$$\Gamma_A = \frac{1}{2M_A} \int \mathrm{d}\Pi_n \left| \mathcal{M}(a \to f) \right|^2 , \tag{1.116}$$

since in order to get the probability we need to take the square of the amplitude.

Now, the Feynman amplitude $\mathcal{M}$'s dimensionality can be inferred from equation (1.114): the dimension of a state is $1/[M]$, the dimension of a four dimensional delta function is $1/[M]^4$, while the time evolution operator is dimensionless, so we have

$$[M]^{-n-1} = [\mathcal{M}][M]^{-4} \implies [\mathcal{M}] = [M]^{3-n}\,, \tag{1.117}$$

where n is the number of outgoing particles. Here by $[M]$ we mean the dimensions of a mass, energy, inverse length or inverse time. The dimension of the phase space element can similarly be found to be $[M]^{2n-4}$; so we can check that Fermi's golden rule is dimensionally consistent: its dimensions read

$$[\Gamma_A] = [M]^{-1}[M]^{2n-4}\Big([M]^{3-n}\Big)^2 = [M]\,, \tag{1.118}$$

which makes sense, since the decay rate is an inverse time.

The expression we gave is **polarized**, that is, by deciding on the initial and final states we are fixing the spins of the particles. This may be useful in some cases, but often we cannot control the spins of the incoming particles, and/or we cannot select only outgoing ones with a certain spin configuration. So, what is done usually is to **average** over the initial polarizations' probabilities, and to **sum** over the final ones. Note that in doing this we must add probabilities, not amplitudes: in this context we are dealing with classical mixtures.

For cross sections, Fermi's golden rule reads:

$$\sigma(A+B\to f) = \frac{1}{2E_A E_B |v_A - v_B|}\int \mathrm{d}\Pi_2 \left|\mathcal{M}(A+B\to f)\right|^2, \tag{1.119}$$

where (if we are in the COM frame) the phase space integral for the two final particles case is (nontrivially! see the TP notes [Tis20, sec. 4.2.3]) given by:

$$\int \mathrm{d}\Pi_2 = \int \frac{\mathrm{d}^3 p}{(2\pi)^3}\frac{1}{2E_1}\frac{1}{2E_2}(2\pi)\delta(E_{CM} - E_1 - E_2) \tag{1.120a}$$

$$= \frac{1}{8\pi}\frac{2p}{E_{CM}}\int\frac{\mathrm{d}\Omega}{4\pi}\,. \tag{1.120b}$$

The dimensionality check now gives $[\mathcal{M}] = [M]^{2-n}$, so

$$[\sigma] = [M]^{-2}[M]^{2n-4}\Big([M]^{2-n}\Big)^2 = [M]^{-2}\,, \tag{1.121}$$

which makes sense: it is a length squared, and lengths are inverse masses.

A kind of process which appears often is a **resonance**: something like $A+B\to X\to A+B$, where we have a certain short-lived intermediate state X, whose decay rate is Γ. The nonrelativistic Breit-Wigner formula describes this: it is given by

$$\mathcal{M} \sim \frac{1}{E - E_X + i\Gamma/2}\,, \tag{1.122}$$

where E_X is the energy of X, while E is the center-of-mass energy of the process.

If we perform a Fourier transform of this expression we get

$$\psi(x) = \int \frac{\mathrm{d}E}{2\pi} \frac{e^{-iEt}}{E - E_X + i\Gamma/2} \tag{1.123}$$

$$= 2\pi i \operatorname{Res}_{E=E_X+i\Gamma/2} \left(\frac{e^{-iEt}}{E - E_X - i\Gamma/2} \right) \tag{1.124}$$

$$= ie^{-i(E_X - i\Gamma/2)t} = ie^{-iE_X t} e^{-\Gamma t/2} , \tag{1.125}$$

so the probability of finding the resonant state, $|\psi|^2$, decays as $e^{-\Gamma t}$.

The paradigmatic process for these kinds of interactions is the process $e^+e^- \to \mu^+\mu^-$.

1.3.5 A decay example

Wednesday 2020-4-1, compiled 2020-07-02

This and next week we will finish the introduction to particle physics, then we will start discussing the open problems in cosmology and astroparticle physics.

We consider the following process:

$$e^+e^- \to \mu^-\mu^+ , \tag{1.126}$$

where the mass of the electron is around $m_e \sim 0.5\,\mathrm{MeV}$, the mass of the muon is around $m_\mu \sim 100\,\mathrm{MeV}$.

Digression: there are different families of fermions (leptons and quarks), the first encompasses e, ν_e, u, d; the second encompasses μ, ν_μ, c, s and the third encompasses τ, ν_τ, t, b. The characteristics of the four members of the family are well-known, and between families the characteristics are the same: the only thing which varies between the families is the mass.

So, Rabi famously asked "who ordered the fermions"?

Coming back to our problem: the state $|e^+e^-\rangle$ must be annihilated by the EM current $j^\mu_{EM} = \overline{\psi}_e \gamma^\mu \psi_e$; it is then converted to a photon, which however is not on mass shell — it cannot be, since its momentum must be that of the electron-positron pair, so a timelike vector. It is then called a *virtual photon*: it can exist, as long as it does so for a short time.

Then, this photon decays to a muon-antimuon pair: then we will have a term $\langle \mu^-\mu^+ | j^\mu_(\mu\rangle$. The index between parenthesis is not a Lorentz one, it just means that this is a muonic current, different from the electronic one.

Let us call p_- and p_+ the momenta of the electron and positron, and p'_- and p'_+ those of the muon and antimuon.

The momentum q of the photon cannot have $q^2 = 0$, but this is fine: it is just an excitation.

The physics of the process is all contained in the matrix element $\mathcal{M}(e^+e^- \to \mu^+\mu^-)$. How do we calculate it? We will not go into details here, but it can be directly derived from the Feynman diagram of the interaction [Tis20, sec. 4.1.2]: we have

$$\mathcal{M}(e^+e^- \to \mu^+\mu^-) = (-e) \left\langle \mu^-\mu^+ \middle| j^\mu \left| 0 \right\rangle \frac{1}{q^2} (-e) \left\langle 0 \right| j_\mu \middle| e^+e^- \right\rangle , \tag{1.127}$$

where the $-e$ factor is because of the EM coupling to the photon. The Breit-Wigner factor looks like

$$\frac{1}{P^2 - M_R^2}, \tag{1.128}$$

but for the photon we have no mass, therefore we only get a factor $1/q^2$.

The Feynman diagrams are just a way to collect the Feynman rules needed to compute the process, they are not meant to represent how the process "looks like".

Let us take the ultrarelativistic limit, in which the energy of the process is much larger than the muon's mass. If this is the case, then we can set $m_e = m_\mu = 0$.

Let us consider the Dirac equation, in the case in which the mass m is equal to zero: then we get

$$i\partial\!\!\!/\psi = 0\,. \tag{1.129}$$

Let us use the chiral representation for the γ matrices:

$$\gamma^\mu = \begin{bmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{bmatrix}, \tag{1.130a}$$

where we mean by $\sigma^0 = \mathbb{1}$, $\sigma^\mu = (\sigma^0, \sigma^i)$ and $\overline{\sigma}^\mu = (\sigma^0, -\sigma^i)$.

Let us then split the spinor ψ into

$$\psi = \begin{bmatrix} \psi_L \\ \psi_R \end{bmatrix}, \tag{1.131a}$$

where $\psi_{L,R}$ are two-component spinors. This allows us to write two two-dimensional equations:

$$i\overline{\sigma}^\mu \partial_\mu \psi_L = 0 \tag{1.132a}$$
$$i\sigma^\mu \partial_\mu \psi_R = 0\,, \tag{1.132b}$$

where we get no interaction terms between the two: if we have no mass the equations decouple.

We can do the same thing if the Dirac equation is coupled to the EM field, since the issue is with the structure of the γ^μ, it does not matter if we have $\gamma^\mu \mathrm{D}_\mu$ or $\gamma^\mu \partial_\mu$; however we will write the decoupled solution for now.

So, for the right-handed spinor we have:

$$\left(i\partial_t + i\vec{\sigma} \cdot \vec{\partial}\right)\psi_R\,, \tag{1.133}$$

which is solved by a plane wave:

$$\psi_R = u_R(p) e^{-iEt + i\vec{p}\cdot\vec{x}}\,. \tag{1.134}$$

Let us suppose the equation reads

$$\left(E - p\sigma^3\right)u_R = \begin{bmatrix} E-p & 0 \\ 0 & E+p \end{bmatrix} u_R = 0\,, \tag{1.135a}$$

so we must have two solutions: they look like

$$\psi_R = \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{-iEt+iEx_3} \qquad \text{and} \qquad \psi_R = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{+iEt+iEx_3}\,. \tag{1.136a}$$

The solution $\psi_R \sim \exp(-iEt + iEx_3)$ describes a right-handed electron with spin eigenvalue $+1/2$ along the direction of motion (this is called the *helicity* [Tis20, sec. 1.4.9])

On the other hand, the solution $\psi_R \sim \exp(iEt + iEx_3)$ describes a right-handed electron with spin eigenvalue $s = -1/2$.

Our quantum field operator ψ_R acts as on the right by destroying a right-handed electron:

$$\langle 0| \psi_R \left| e_R^-(p) \right\rangle = u_R(p) e^{-ipx}\,. \tag{1.137}$$

Then, we can have it acting on the left by destroying a left-handed positron:

$$\left\langle e_L^+(p) \right| \psi_R |0\rangle = v_L(p) e^{+ipx}\,. \tag{1.138}$$

If we were to repeat the analysis for the other spinor, we would get the specular result.

The Lagrangian can be written as

$$\mathscr{L} = \psi_R^\dagger (i\sigma \cdot \partial)\psi_R + \psi_L^\dagger (i\overline{\sigma} \cdot \partial)\psi_L - m\left(\psi_R^\dagger \psi_L + \psi_L^\dagger \psi_R\right), \tag{1.139}$$

so the coupling between the left and right handed fermions depends on the mass, if we are in a situation in which $T \gg m$ they effectively decouple.

In order to describe this spin, we introduce the helicity quantum number, which is defined as

$$h = \hat{p} \cdot \vec{s}\,, \tag{1.140}$$

the projection of the spin along the direction of motion. For ψ_R, we have the $(1, 0)$ state with helicity $h = 1/2$, while the state $(0, 1)$ is a positron with helicity $h = -1/2$.

If $m = 0$, then helicity is exactly conserved. At high energies, it is suppressed by a factor m/E.

Let us compute the cross section: we must calculate

$$\langle 0| j^\mu \left| e_R^-(p_-) e_L^+(p_+) \right\rangle\,, \tag{1.141}$$

where we have a term $j^\mu = \overline{\psi}\gamma^\mu\psi = \psi^\dagger\gamma^0\gamma^\mu\psi$; the term $\gamma^0\gamma^\mu$ reads

$$\gamma^0\gamma^\mu = \begin{bmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix} \begin{bmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{bmatrix} = \begin{bmatrix} \overline{\sigma}^\mu & 0 \\ 0 & \sigma^\mu \end{bmatrix}, \tag{1.142}$$

so we get

$$j^\mu = \psi_L^\dagger \overline{\sigma}^\mu \psi_L + \psi_R^\dagger \sigma^\mu \psi_R \,. \tag{1.143}$$

We describe the process in the center-of-mass frame, and we choose to align the axes so that the electron and positron have momenta $p^\mu = (E, 0, 0, \pm E)$ respectively (minus for the positron).

The wavefunction $\psi_R^\dagger$ annihilates the positron e_L^+ yielding a term $v_L^\dagger(p_+)$, the wavefunction ψ_R annihilates the electron e_R^- yielding a term $u_R(p_-)$.

Then, we are left with

$$\begin{aligned} v_L^\dagger(p_+)\sigma^\mu u_R(p_-) &= \sqrt{2E} \begin{bmatrix} 0 & 1 \end{bmatrix} (\mathbb{1}, \vec{\sigma}) \sqrt{2E} \begin{bmatrix} 1 \\ 0 \end{bmatrix} &(1.144)\\ &= 2E(0, 1, i, 0)^\mu \,, &(1.145) \end{aligned}$$

so if we define the vector $\vec{\epsilon}_+ = (\hat{1} + i\hat{2})/\sqrt{2}$ we can write

$$\langle 0 | j^\mu \left| e_R^-(p_-) e_L^+(p_+) \right\rangle = 2E\sqrt{2}(0, \vec{\epsilon}_+)^\mu \,, \tag{1.146}$$

while we would have

$$\langle 0 | j^\mu \left| e_L^-(p_-) e_R^+(p_+) \right\rangle = -2E\sqrt{2}(0, \vec{\epsilon}_-)^\mu \,, \tag{1.147}$$

with $\vec{\epsilon}_- = (\hat{1} + i\hat{2})/\sqrt{2}$.

On the other hand, the terms $e_R^- e_R^+$ and $e_L^- e_L^+$ do not contribute (in our $m = 0$ approximation).

The muons are massless fermions as well in our treatment, so we get analogous terms:

$$\begin{aligned} \left\langle \mu_R^-(p'_-) \mu_L^+(p'_+) \right| j^\mu \left| 0 \right\rangle &= 2E\sqrt{2}\big(0, \vec{\epsilon}_+^{\prime *}\big)^\mu &(1.148)\\ \left\langle \mu_L^-(p'_-) \mu_R^+(p'_+) \right| j^\mu \left| 0 \right\rangle &= -2E\sqrt{2}\big(0, \vec{\epsilon}_-^{\prime *}\big)^\mu \,, &(1.149) \end{aligned}$$

so in the end we find

$$\begin{aligned} \mathcal{M}(e_L^- e_R^+ \to \mu_R^- \mu_L^+) &= -\frac{e^2}{q^2} 2(2E)^2 \vec{\epsilon}_+^{\prime *} \cdot \vec{\epsilon}_+ &(1.150)\\ &= -2e^2 \vec{\epsilon}_+^{\prime *} \cdot \vec{\epsilon}_+ \,, &(1.151) \end{aligned}$$

since $q = 2E$. The scalar product here depends on the direction of emission of the muons in the center of mass frame, θ; we find the absolute value $|\mathcal{M}|^2 = e^4(1 \pm \cos\theta)^2$, depending on whether we are looking at an $LR \to LR$ process or $LR \to RL$ process. The unpolarized (spin-averaged) differential cross section comes out to be:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta} = \frac{1}{2} \frac{\pi\alpha^2}{2E_{CM}^2} \Big(1 + \cos^2\theta\Big) \,, \tag{1.152}$$

which can be integrated across the sphere to get the total cross section for the process $e^- e^+ \to \mu^- \mu^+$:

$$\sigma = \frac{4\pi}{3} \frac{\alpha^2}{E_{CM}^2} \,. \tag{1.153}$$

We could have guessed the dependence on α^2 / E_{CM}^2, but for the numerical factor we needed to do the full computation. This is because the cross section is a length square, so it must depend on the inverse square of our only energy parameter, E_{CM}.

Also, the coupling was fixed: we are working in QED, so we only have a coupling constant: e, so we will have terms $e^2/4\pi = \alpha$ inside of $\mathcal{M}$, so we will get α^2 inside of $|\mathcal{M}|^2$. The 4πs will cancel because of the phase-space angular integrals.

Once we have done this, can we generalize it? suppose we want to compute the cross section $\sigma(e^+ e^- \to \text{hadrons})$. How could we do it?

We can make a similar kind of reasoning: the process will look like $e^- e^+ \to q\bar{q}$, and while the coupling for the first vertex in the Feynman diagram will be $-e$ the one for the second vertex will look like Q_q, the charge of these quarks. We calculate the unpolarized cross section, counting all the quarks which can be produced at a fixed COM energy: we assume $E \sim 100\,\text{GeV}$, so all the quarks except for the top are candidates; so the computation goes:

$$\frac{\sigma(e^+ e^- \to q\bar{q})}{\sigma(e^+ e^- \to \mu^+ \mu^-)} \approx \sum_q Q_q^2 = \underbrace{2\frac{4}{9}}_{u,c} + \underbrace{3\frac{1}{9}}_{d,s,b} = \frac{11}{9} \,. \tag{1.154}$$

In experiments, however, we get a cross section ratio which is $11/3$, 3 times larger than expected: this indicates that we have a different type of charge, color charge, which means we have a multiplicity of 3 for each quark.

We shall see that this is related to certain kinds of internal symmetries of our field theory.

1.3.6 Gauge symmetries in QED

Tuesday 2020-4-7, compiled 2020-07-02

Last time we wrote the Lagrangian of QED:

$$\mathscr{L}_{QED} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \overline{\psi}\Big(i\gamma^\mu \mathrm{D}_\mu - m\Big)\psi \,, \tag{1.155}$$

where $\mathrm{D}_\mu = \partial_\mu + ieA_\mu$.

This possesses several symmetries: Lorentz (actually, Poincaré) invariance, P (parity), C (charge conjugation), T (time inversion).

We are also interested in the *internal* symmetries of this Lagrangian: all the aforementioned symmetries were spacetime ones.

We know that $\overline{\psi} = \psi^\dagger \gamma^0$, so if we transform $\psi \to e^{i\theta}\psi$ the Lagrangian is unchanged, since we also have $\overline{\psi} \to e^{-i\theta}\overline{\psi}$. Here, we are taking a *constant* phase angle θ: it comes out of the derivative unchanged. This is called a $U(1)$ **global** symmetry, since it is the same everywhere in space and since a phase is the same as a 1×1 unitary matrix.

The following section differs from the notes. Let us ignore the interactions of electrons with the EM fields in QED. Our world is made of electrons, and we want to describe these free propagating electrons. We take the Lagrangian

$$\mathscr{L}_{\text{Dirac}} = \overline{\psi}\Big(i\gamma^\mu \partial_\mu - m\Big)\psi\,, \tag{1.156}$$

which still has the symmetry $\psi \to e^{i\alpha}\psi$.

Is this invariant also with respect to a *local* $U(1)$ symmetry? This looks like $\psi(x) \to e^{i\alpha(x)}\psi(x)$, where $\alpha(x)$ is a continuous spacetime scalar function.

This is a "promotion" of the symmetry: why? It seems like a global symmetry is a more general thing... However, the global symmetry is a special case of the local one.

This local symmetry is called a $U(1)$ *gauge* symmetry. Properly speaking, the global symmetry is also a gauge one but it is commonly called just a global symmetry.

Substituting in, we get

$$e^{-i\alpha(x)}\overline{\psi}\Big[i\partial_\mu\gamma^\mu - m\Big]e^{i\alpha(x)}\psi = \overline{\psi}\Big[i\partial_\mu\gamma^\mu - m\Big]\psi + \overline{\psi}\psi i\gamma^\mu\partial_\mu\alpha\,, \tag{1.157}$$

so we see that the Lagrangian is *not* invariant under this gauge symmetry in general.

If we want the symmetry to hold, we need to introduce a *compensating* field to cancel out the term.

The answer is that the thing to add is a vector field A_μ. Then, if we transform

$$A_\mu \to A_\mu - \frac{1}{e}\partial_\mu\alpha \tag{1.158}$$

this compensates the change, as long as the vector is coupled to the fermion with a term $\overline{\psi}e\gamma^\mu A_\mu\psi$ in the Lagrangian. This is a profound result: if we wanted to describe a world with only electrons, and we want to have this electron be symmetric with respect to the $U(1)$ gauge symmetry $e \to e^{i\alpha(x)}e$ then we *must* have a vector coupled to it.

Then, we should also insert a term describing the propagation of the free vector A_μ, the kinetic term $\propto F^{\mu\nu}F_{\mu\nu}$.

A quote: "And she said 'let there be symmetry', and there was light".

We could have just proven that the QED Lagrangian is invariant with respect to $U(1)$ symmetry: however, this reasoning illustrates the point that the symmetry requires the insertion of photons.

One might ask: how do you know that this is the correct symmetry? The method is trial and error.

Could we have something which is more complicated than a spin 1 mediator? It is basically a guessing game, we see what works.

1.4 QCD

This section can also be followed from Peskin [Pes19, sec. II.11].

This $U(1)$ gauge symmetry is an abelian symmetry, but we also have non-abelian ones: we denote by T^a the generators of the group, their Lie algebra is defined by

$$\left[T^a, T^b\right] = i f^{abc} T^c \,. \tag{1.159}$$

The *structure coefficients* f^{abc} are manifestly antisymmetric in their first two indices. It can also be shown that they are fully antisymmetric. If the group is abelian, we have f^{abc} identically, but this need not be the case.

To say that these are generators means that any infinitesimal transformation can be written as

$$\Phi \to (1 + i\alpha^a t_R^a)\Phi \,, \tag{1.160}$$

where α^a are the parameters of the infinitesimal transformation, while t_R^a are Hermitian matrices of dimension d_R which make up the representation of the group. There are d_G of them, where d_G is the dimension of the group.

The finite unitary transformation mapping $\Phi \to U(\alpha)\Phi$ can be recovered from here by

$$U(\alpha) = e^{i\alpha^a t_R^a} \,. \tag{1.161}$$

We will be interested in Lie groups $SU(n)$ with $N \geq 2$, which are $N \times N$ unitary matrices with determinant 1.

For example, recall that $SU(2)$ has a 2-to-1 correspondence with $SO(3)$. $SU(2)$ describes the rotation of spinors, the generators of their rotation are $\sigma^i/2$.

Last time we discussed the annihilation of e^+e^- into hadrons: we can get protons and neutrons, pions, kaons…

However we can simplify by discussing only the creation of quarks.

When we compute the cross sections, our calculation seems to be wrong by a factor 3. If we multiply it by 3 we get the correct result. So there are three types of quarks: we categorize them by "color", even though it has nothing to do with colors.

We associated QED with $U(1)$: is there a group corresponding to Quantum Chromo Dynamics? Can we do this with the weak interaction as well?

Since there are three quarks, we are drawn to represent them as triplets. Since we know that unitary matrices are nice, we try $SU(3)$. We call this symmetry $SU(3)_{\text{color}}$.

We start by giving some general results for N-dimensional groups $SU(N)$: we normalize their representation by imposing

$$\mathrm{Tr}\left[t_N^a t_N^b\right] = \frac{1}{2}\delta^{ab} \,, \tag{1.162}$$

where t_N^a are the Hermitian generators of an N-dimensional unitary representation. If we generalize to an R-dimensional representation, we will have

$$\mathrm{Tr}\left[t_G^a t_G^b\right] = C(R)\delta^{ab} \,, \tag{1.163}$$

where $C(R)$ is some constant depending only on the dimension of the representation.

A special representation we can choose is the **adjoint** representation, which is the one under which the generators of the Lie algebra transform; it is defined by:

$$\left(t_G^a\right)^{bc} \stackrel{\text{def}}{=} if^{abc} . \tag{1.164}$$

By making use of the Jacobi identities we can show that this is indeed a valid representation of the group, its dimension is that of the group and we have:

$$\operatorname{Tr}\left[t_G^a t_G^b\right] = f^{acd} f^{bcd} = C(G)\delta^{ab} , \tag{1.165}$$

where the constant $C(G)$ is just the dimension N for the adjoint representation. For $SU(N)$ the dimension is N^2-1: the representation consists of N^2-1 matrices, each $N \times N$.

1.4.1 Non-abelian gauge theory: Yang-Mills

Consider the Lagrangian

$$\mathscr{L} = \overline{\psi} i\gamma^\mu \partial_\mu \psi . \tag{1.166}$$

This can describe an electron. Now, let us add an index j in the group G, according to which the particle transforms in the R-dimensional (in our case $R=3$) representation:

$$\mathscr{L} = \overline{\psi}_j i\gamma^\mu \partial_\mu \psi_j . \tag{1.167}$$

Under the local action of a group G the particle transforms like

$$\psi_j(x) \stackrel{G}{\to} \psi_j^{'}(x) = \left(1 + i\alpha^a(x) t_R^a\right)_{jk} \psi_k , \tag{1.168}$$

where a is an index going from 1 to N^2-1.

We have the same problem we had with $U(1)_{\text{em}}$: we need to cancel the term coming from the derivative of $\alpha(x)$, with a one-index object. The variation in the Lagrangian is

$$\delta\mathscr{L} = \overline{\psi}_j i\gamma^\mu \left(i\partial_\mu \alpha^a(x) t_{R,jk}^a\right) \psi_k . \tag{1.169}$$

The solution is the same: we introduce a coupling to the derivative, which takes the form

$$\mathrm{D}_\mu \to \partial_\mu - igA_\mu^a t_R^a , \tag{1.170}$$

where g is the strength of the interaction, which is also written as g_s in the case of the strong force. The index a goes from 1 to $N^2-1=8$: we must introduce a vector field A for each *generator* of the required symmetry.

The beautiful thing is the fact that starting only from the symmetry requirement we can get a full theory.

We now attribute actual *existence* to these quantum fields: they are interaction bosons. They must transform like

$$A_\mu^a(x) \to A_\mu^a(x) + \frac{1}{g}\partial_\mu \alpha^a(x) + \overbrace{A_\mu^b f^{abc} \alpha^c(x)}^{\text{new nonabelian term}} \tag{1.171}$$

$$= A_\mu^a(x) + \frac{1}{g} \mathrm{D}_\mu \alpha^a(x) \,. \tag{1.172}$$

This is derived making use of the adjoint representation definition (1.164): we are equating

$$A_\mu^b f^{abc} \alpha^c = -i A_\mu^b t_R^b \alpha^a \tag{1.173}$$

$$A_\mu^b f^{abc} \alpha^c = -i A_\mu^b (i f^{bca}) \alpha^c \tag{1.174}$$

$$A_\mu^b f^{abc} \alpha^c = + A_\mu^b f^{abc} \alpha^c \,. \tag{1.175}$$

We can do cyclic permutations of the indices of f^{abc}.

If we assign physical reality to these fields we must give them kinetic terms in the Lagrangian, which will then be

$$\boxed{\mathscr{L} = -\frac{1}{4} F^{\mu\nu a} F_{\mu\nu}^a + \overline{\psi}\Big(i\gamma^\mu \mathrm{D}_\mu - m\Big)\psi \,,} \tag{1.176}$$

which looks the same as the QED Lagrangian, but we must be careful: what is $F_{\mu\nu}^a$? We might think it is

$$F_{\mu\nu}^a = 2\partial_{[\mu} A_{\nu]}^a \,, \tag{1.177}$$

but this is not enough: inside the covariant derivative in $\mathrm{D}_\mu \alpha$ in the transformation law we also have A_μ, so we get an additional term, and the final formula looks like

$$F_{\mu\nu}^a = 2\partial_{[\mu} A_{\nu]}^a + g f^{abc} A_\mu^b A_\nu^c \,. \tag{1.178}$$

This corresponds to the fact that, while there is no photon-photon interaction since $U(1)$ is abelian, $SU(3)$ is not: so, we do have gluon-gluon interaction. Notice that this expression has abelian symmetries as a special case: the term fAA is antisymmetric in the two bosons, so if they commute it vanishes. The square of this field strength is then both Lorentz and gauge invariant.

This field strength can be interpreted as the Riemann tensor of the Lie group manifold:

$$\Big[\mathrm{D}_\mu, \mathrm{D}_\nu\Big] = -ig F_{\mu\nu}^a t_R^a \,. \tag{1.179}$$

This is crucial in very high energy situations, such as in the early universe. The dynamics of the field are now more complicated than the ones we find in electrodynamics due to the nonlinear terms.

The wavefunction ψ which appears in the Lagrangian will transform in some d-dimensional representation of $G = SU(3)_c$. It will have indices like $\psi_{\alpha,i}$: α is a four-dimensional spinorial index, while i is a three-dimensional color index.

The strong-interaction coupling constant g_s is dimensionless, we also define the parameter

$$\alpha_s = \frac{g_s^2}{4\pi} \,. \tag{1.180}$$

The coupling term of the gluons with the fields can be made explicit as

$$\overline{\psi}_{\alpha,i} \gamma^\mu_{\alpha\beta} A_\mu^a (t^a)_{ij} \psi_{\beta,j} \,, \tag{1.181}$$

where μ is a Lorentz index, i and j are color indices, while α and β are spinorial indices.

1.4.2 Quantum Chromo Dynamics

Wednesday 2020-4-8, compiled 2020-07-02

Asymptotic freedom is not a political slogan from the sixties!

With this lecture and the next we should finish discussing the standard model of particle physics.

Today we are going to explore a property which is profoundly different between EM and the strong interaction.

We introduced the quantity $\alpha_{EM} = e^2/4\pi$: how do we measure this parameter? We know that the interaction term looks like

$$\mathscr{L} \sim e j^{\mu} A_{\mu} \,, \tag{1.182}$$

so we can measure e using the cross sections of processes, such as $e^+e^- \to \mu^+\mu^-$; the free parameters are just e and the COM energy, so we are done. But how do we do it without an accelerator? We could do the Millikan drop experiment, for example.

But we are doing Quantum Field Theory: beyond physical particles we can also have virtual particles, which despite the name can influence physical processes: inside the "black box" within which the interactions occur we can have off-shell processes.

In fact, in the aforementioned decay the photon is off mass shell; but we can go beyond. The photon can create and then annihilate an e^+e^- pair.

The vacuum is a *quantum* vacuum, which does correspond to a minimal energy but it is filled by the continuous creation and annihilation of these particle-antiparticle pairs, since there is an indetermination between time and energy. It is important that after being created these pairs are indeed destroyed. This is called *vacuum polarization*.

The pairs are virtual but they have an effect. The coupling constant α has a numerical value of around $1/137$, and the loop diagrams of increasing number of loops depend on increasing powers of α. The fact that α is small allows us to work perturbatively.

This works, the predictions of QED are extremely precise and correspond to experiment. For instance, the anomalous magnetic moment of the electron: in a QFT, when an electron interacts with a photon, there can be other particles.

The Feynman diagram which governs the interaction of an external electromagnetic field with a fermion is a one-vertex one; however we can draw diagrams with loops (the first correction is shown in figure 1.1) to calculate the perturbative corrections.

The polarization of the vacuum creates a screening effect, as shown in an example scattering in figure 1.2. This is hard to compute analytically, but the result can be readily described: the vacuum acts as a dielectric medium, screening part of the electromagnetic charge and making the interaction weaker.

So, we expect the effective α_{EM} to be larger: if the vacuum is filled with a sea of screening pairs, at large distances (low energies) it is effectively uniform, while if we get at distances lower than $1/m_e$ we start to be on smaller scales than the maximum size of these dipoles, so screening is harder and we can see the "bare charge", which formally diverges at a point called the **Landau pole**.

This is an illustrative description of a mathematical process called **renormalization**: we impose the fact that the loop contributions to the theory should not diverge, and this forces

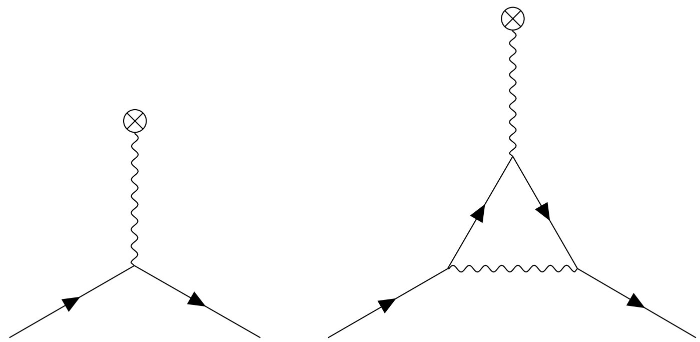

Figure 1.1: Interaction of a fermion with an external electromagnetic field, to first and third order respectively.

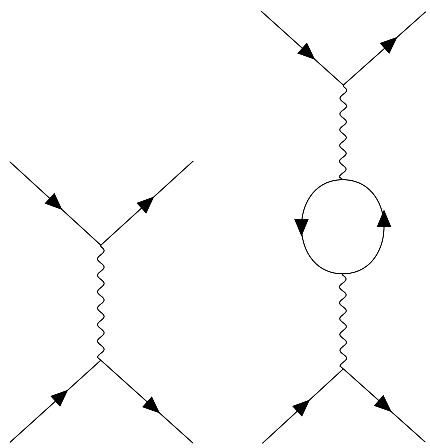

Figure 1.2: Vacuum polarization: second and fourth order contributions to $e^- e^- \to e^- e^-$ scattering.

us to have a running coupling to the theory, which corresponds to increased pair production (since the EM field is coupled more strongly with the fermions!).

We compare the Millikan experiment ($E \sim 0$), where we find $\alpha \sim 1/137$, with LEP: now $E \sim 100\,\text{GeV}$, and we find $\alpha \sim 1/128$! A figure depicting the complete relation can be found in Peskin [Pes19, fig. 11.1]. At very high energies we must also account for pair production of higher-mass fermions, such as muons.

What is the theoretical relation between α and Q, the momentum transfer of the process at hand? It comes out to be (the calculation can be found in an older book by Peskin and Schroeder [PS95, sec. 7.5])

$$\alpha(Q) = \frac{\alpha_0}{1 - \frac{2\alpha_0}{3\pi} \log\left(\frac{Q}{Q_0}\right)} \,. \tag{1.183}$$

At energies $Q \lesssim m_e$ the correction is negligible and $\alpha \sim 1/136$, it becomes significantly different at higher energies.

We also define the function β: it is a function of the coupling constant α, defined by

$$\beta(g_e) = \frac{\mathrm{d}g_e}{\mathrm{d}\log Q}. \tag{1.184}$$

What happens if we compute this for the strong interaction? We define, analogously to the electromagnetic case:

$$\alpha_s = \alpha_{\text{strong}} = \frac{g_s^2}{4\pi}. \tag{1.185}$$

The problem is that this is a large number: but our tools are perturbative! How do we compute the cross sections then? The thing is, with increasing energy this α_s becomes *smaller*! Why is this? The gluons are self interacting, because of the $A \wedge A$ term in the field strength of the strong interaction, as opposed to photons.

In the EM case, we had $\gamma \to e^+e^- \to \gamma$ as a virtual particle process; in the chromodynamics case instead we have $g \to g + g \to g$: it is the case (even though it is not easy to prove) that this loop contribution has the opposite sign! We define β in the same way as the one from QED, so we find [Pes19, eq. 11.65]:

$$\frac{\mathrm{d}g_s}{\mathrm{d}\log Q} = \beta(g_s) \qquad \text{where} \qquad \beta = -\frac{11}{3}C(G) + \frac{4}{3}n_f C(R), \tag{1.186}$$

where $C(G)$ is the Casimir[1] of the **adjoint** representation: $C(G) = 3$, while $C(R) = 1/2$ pertains to the **fundamental** representation, to which fermions belong: n_f, the number of fermions, is 6. Then, the final result is negative!

This is called asymptotic freedom: α decreases as Q increases. If the energy is high enough, the quarks and gluons are effectively decoupled. On the other hand, at low energy their coupling is very large.

So, at the low energies we look at in atomic nuclei the quarks are confined by an extremely high coupling: no one has ever observed a free quark, not even at LHC. This is called hadronization, since the quarks are always bound into hadrons. ALICE, at LHC, is looking for the phase transition between this "infrared slavery" of the quarks and their free state.

Since α_S decreases, while α_{EM} increases, there should be a point at which they cross and the electromagnetic interaction becomes stronger than the strong one.

There might be a point at which they merge, like the weak and electromagnetic interactions do. Actually, it might be possible that all of the interactions are manifestations of the same kind of interaction.

The crucial point is that the strength of an interaction depends on the energy. If we define $b_0 = 11 - 2n_f/3$ we find

$$\alpha_s(Q) = \frac{2\pi/b_0}{\log(Q/\Lambda)} \qquad\qquad \Lambda = Q_0 \exp\left(\frac{-2\pi}{b_0 \alpha_s(Q_0)}\right), \tag{1.188}$$

[1] This means that it is the constant depending on the dimension of the representation which appears in

$$\operatorname{Tr}\left[t_N^a t_N^b\right] = C(N)\delta^{ab}. \tag{1.187}$$

where we defined Λ, the QCD scale: it is the energy at which the QCD coupling becomes strong.

In QED (and more generally in abelian gauge theories) the vertex only has one photon line, so there can be no photon loops. In QCD, instead, we can have gluon loops.

1.4.3 The weak interaction

How do we formally describe whether an interaction is "strong"? We know that the time needed for hadronization is of the order $\tau_{\text{had}} \sim 10^{-23}\,\text{s}$, while the time for a pion's decay is $\tau(\pi^+) \sim 3 \times 10^{-8}\,\text{s}$. These are strong interaction-mediated processes.

The time for beta decay — a weak interaction mediated process — is $\tau(n \to p + e^- + \overline{\nu}_e) \sim 880\,\text{s}$.

This can give us a first signal for the fact that the weak interaction is *weak*.

We know that the neutron is *udd* in terms of quarks, while the proton is *uud*: so, in the beta decay one down quark must become an up quark. There is no term in QED or QCD in which this can happen: QED and QCD conserve flavour. This can be seen in the Lagrangian: there are terms which mix colors, while the flavour terms are diagonal.

But we do see beta decay: so, we must introduce a new *weak interaction*.

Beta decay is experimentally observed to be a three-body process, which is why observations of it were the first indication of the existence of neutrinos. Beta-decay is not P-symmetric, whereas QED and QCD are. This was seen in the β^- decay of ^{60}Co nuclei, they always emitted left-handed electrons.

In QED and QCD we see current-current interaction by means of the mediator (photon or gluon, respectively); will we see current-current weak interactions?

Will we?

Wednesday 2020-4-15, compiled 2020-07-02

We would like to introduce a sort of "Quantum Weak Dynamic" theory, and we will see that this is possible in the standard model, by unifying the electromagnetic and weak interactions.

1.4.4 The current-current model for the weak interaction

We have seen that fermions are described by a four-component spinor:

$$\Psi = \begin{bmatrix} \psi_L \\ \psi_R \end{bmatrix}, \tag{1.189a}$$

where $\psi_{L,R}$ are both two-component vectors. The right handed part has helicity $h = +1/2$, the left handed part has helicity $h = -1/2$.

The puzzling part of the weak interaction is that it seems like the only part of the fermion entering β-decay is the left-handed part. In the late seventies, the professor's master thesis stated that this was a mystery. Still, we do not know why this interaction breaks parity.

We need to describe the experimental results mathematically: so, we need a projector onto the left-handed components. We already introduced the Dirac γ^μ matrices. Now we introduce the **fifth** gamma matrix: in the chiral representation it reads

$$\gamma^5 = \begin{bmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{bmatrix}, \tag{1.190}$$

which has the important property that it anticommutes with all the Dirac matrices:

$$\left\{\gamma^5, \gamma^\mu\right\} = 0, \tag{1.191}$$

and actually we can write it as

$$\gamma^5 = \frac{i}{4!}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma = i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{1.192}$$

By its form we can readily see that ψ_L is an eigenstate of γ^5 with eigenvalue -1, while ψ_R has eigenvalue $+1$. So, the matrix

$$\frac{\mathbb{1} \pm \gamma^5}{2} \tag{1.193}$$

projects a state onto the left- or right-handed subspace (minus for the left handed one). If we apply a parity transformation $P\vec{x} \to -\vec{x}$ to the current $\overline{\psi}\gamma^5\psi$ we get $-\overline{\psi}\gamma^5\psi$.

The Dirac matrices can be written as

$$\gamma^\mu \begin{bmatrix} 0 & \sigma^\mu \\ -\overline{\sigma}^\mu & 0 \end{bmatrix}, \tag{1.194a}$$

where

$$\sigma^\mu = \begin{bmatrix} \mathbb{1} & \sigma^\mu \end{bmatrix} \qquad \text{and} \qquad \overline{\sigma}^\mu = \begin{bmatrix} \mathbb{1} & -\sigma^\mu \end{bmatrix}. \tag{1.195}$$

We have already seen that if we set $m = 0$ the Dirac equation for the two left- and right-handed spinors decouples:

$$i\overline{\sigma} \cdot \partial\psi_L = 0 \tag{1.196a}$$

$$i\sigma \cdot \partial\psi_R = 0. \tag{1.196b}$$

We know that neutrons are made up of ddu quarks, while protons are uud, so we must have a down quark turning into an up: the Feynman diagram for β^- decay (from far away) looks like what is shown in figure 1.3.

Only the left-handed spinors appear in the process. In order to describe this, we can introduce a current involving only the left-handed components:

$$j_L^{\mu+} = \nu_L^+\overline{\sigma}^\mu e_L + u_L^+\overline{\sigma}^\mu d_L + \dots \tag{1.197}$$

$$j_L^{\mu-} = e_L^+\overline{\sigma}^\mu \nu_L + d_L^+\overline{\sigma}^\mu u_L + \dots. \tag{1.198}$$

This current will appear in the Feynman amplitude in this fashion:

$$\mathcal{M} \sim \frac{4G_F}{\sqrt{2}} j_L^{\mu+} j_L^{\mu-}. \tag{1.199}$$

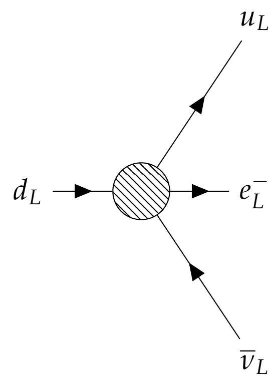

Figure 1.3: Beta decay: an unknown weak process.

These terms can be written without specifying L or R, as:

$$\overline{u}\gamma^\mu \left(\frac{1-\gamma^5}{2} \right) d = \frac{1}{2} \left(\underbrace{\overline{u}\gamma^\mu d}_{V} - \underbrace{\overline{u}\gamma^\mu \gamma^5 d}_{A} \right) = \overline{u}_L \gamma^\mu d_L \,. \tag{1.200}$$

Notice that we did not introduce here an index a as we had done for the color charge.

The interaction boson mediating the weak interaction is called a **charged current**, since it must have an electric charge.

This theory of weak interaction is called a $V - A$ theory: in the current we have distinguished the terms V — a vector current — and A, an axial current. This distinction is a group representation one: the vector term transforms like a vector, while the axial term transforms like a pseudovector.

We know that a density Lagrangian has a dimension of 4, so from the expression $\mathscr{L} = \overline{\psi} m \psi$ we get that the fermion must have dimension $3/2$.[2] So, the current $j^\mu \sim \overline{\psi}\gamma^\mu\psi$ has a dimension of 3. In the Lagrangian the current term will look like $G_F j^\mu j_\mu$, so the constant G_F must have a dimension of -2.

Experimentally, we find $G_F \sim 10^{-5}\,\text{GeV}^{-2}$.

This kind of $V - A$ theory has been successful in the description of processes such as:

$$n \to p e^- \overline{\nu}_e \tag{1.201}$$
$$\mu \to \nu_\mu e^- \overline{\nu}_e \tag{1.202}$$
$$\pi^- \to \mu^- \overline{\nu}_\mu \text{ or } e^- \overline{\nu}_e \,. \tag{1.203}$$

We can estimate the order of magnitude of the cross section of weak interaction processes without drawing any Feynman diagrams: we know that they must depend on G_F^2, which has dimension -4, while the cross section has dimension -2: what can we multiply it by to fix the dimensionality?

[2] By "dimension" we mean mass dimension: the dimension of a mass m is 1.

Our only free parameter is $s = (p_1 + p_2)^2$, the square of the center of mass energy. So, we will have

$$\sigma \sim G_F^2 s\,, \tag{1.204}$$

which means that if we increase the beam energy the cross section increases.

This is a problem: s can diverge as we raise the energy of our colliders, but σ is connected to a probability: this means that we have a unitarity violation.

The problem becomes worse if we consider n loops: then we will have $\sigma \sim G_F^{2n} s^m$, where by dimensional analysis we find $m = 2n - 1$: this increases with n!

This can be fixed with the introduction of a charged mediator W instead of a four-current vertex with a coupling G_F. This new boson will have a vertex with both of the currents. This is called **Intermediate Vector Boson** theory.

The mass of this boson is denoted as M_W; then in the amplitude we will need to insert a factor

$$\frac{1}{q^2 - M_W^2}\,, \tag{1.205}$$

which avoids the divergence. Here q^2 is the square of the four-momentum of the carrier: in simple processes it will just be given by $q^2 = s$. Then, we go from $\sigma \sim G_F^2 s$ to

$$\sigma \sim \left(\frac{1}{q^2 - M_W^2}\right)^2 s\,, \tag{1.206}$$

which scales as $\sigma \sim 1/s$ when s diverges.

A massless mediator would not work, even though it would fix the ultraviolet problems of our theory, because of the observed behaviour at low energies. The interaction we see does not have an infinite range, like the photon; instead it has a very short range. Because of the measured value of G_F, we expect this boson's mass M_W to be of the order of 100 GeV, since at low energies the Breit-Wigner factor reduces to

$$\frac{g_w^2}{q^2 - M_W^2} \sim \frac{g_w^2}{M_W^2} \sim G_F \implies M_W \sim \frac{g_w}{\sqrt{G_F}}\,, \tag{1.207}$$

and we expect the coupling g_w to be of order one. We see no q^2 dependence in G_F at low energies $\lesssim 30\,\text{GeV}$, which supports this hypothesis.

Now we need to describe this as a gauge theory. The symmetry group will be $SU(2)_L$ and we will have $2^2 - 1 =$ three generators. The photon is described by $U(1)$ symmetry, and it has no mass term $\sim m^2 A^\mu A_\mu$. If we were to insert such a term we would lose gauge symmetry. We must then solve the problem of implementing a gauge theory description of a massive boson.

Let us consider the problem kinematically: if a vector boson is massive we can go to its rest frame, so the polarization vector has *three* degrees of freedom: so, this new boson will need to have a new degree of freedom.

If we break gauge freedom, there will be no way anymore to describe our vector boson by imposing it with the Gupta-Bleuer condition: so, we will get states of negative probability.

Also, gauge invariance ensures the **renormalizability** of the theory: the possibility to be able to change the parameters off the theory with the energy so that we do not get diverging contributions from loop diagrams.

Tuesday 2020-4-21, compiled 2020-07-02

Now we have a dilemma: beauty versus pragmatism.

How do we provide a mass to the W vector boson?

As we were discussing, we could add a mass term and write

$$\mathscr{L} = \mathscr{L}_{\text{Yang-Mills}} + M^2 W_\mu^+ W^{\mu,-} , \tag{1.208}$$

but this term *brutally* breaks the symmetry. This is explicit breaking of the symmetry.

However, if we do not break the symmetry we cannot reproduce the data... or can we?

1.4.5 Spontaneous Symmetry Breaking

We introduce **Spontaneous Symmetry Breaking**. This is a phenomenon which is not exclusive to HEP: it happens in ferromagnets, for example. They start off (above the Curie temperature) with the spins pointing in uniformly distributed directions: the situation is symmetric. As we cool them, at a certain stage all the spins align. We cannot predict the direction along which they will align (since the initial state is symmetric), but once they do the symmetry is *spontaneously broken*.

Let us take the standard approach (also done in the Theoretical Physics course [Tis20, sec. 5.2]) introduce the Lagrangian of a complex scalar field ϕ:

$$\mathscr{L} = \left(\partial^\mu \phi^*(x)\right)\left(\partial_\mu \phi(x)\right) \underbrace{-\mu^2 \left|\phi(x)\right|^2 - \lambda^4 \left|\phi(x)\right|^4}_{-V(\phi)}, \tag{1.209}$$

and to find the ground state we can move to the Hamiltonian:

$$\mathscr{H} = \pi(x)\phi(x) - \mathscr{L} . \tag{1.210}$$

The parameter λ must be > 0, while μ^2 has no constraints.

It is a critical parameter: if μ^2 is positive we are in the symmetric phase of the system, and the vacuum is only $\phi = 0$: the VEV is $\langle 0| \phi |0\rangle = 0$.

If, on the other hand, μ^2 is negative we transition to a new phase which is not symmetric: the vacuum becomes a whole circle, whose VEV is $v = \sqrt{-\mu^2/2\lambda}$. The whole region parametrized as $ve^{i\theta}$, with $\theta \in \mathbb{R}$, provides equivalent vacua: a specific ground state is not symmetric under the whole symmetry of the Lagrangian.

Why do we only have terms in ϕ^2 and ϕ^4? we basically constructed the simplest potential which has the properties we want.

If $\mu^2 > 0$ we can perturb around the state without the quartic term, which is described by the KG equation.

If, instead, we want to perturb around a vacuum in the $\mu^2 < 0$ case we need to distinguish two directions, since the Hessian of the potential has two different eigenvalues, one positive and one equal to zero. We parametrize

$$\phi = \frac{1}{\sqrt{2}}\big(v + \sigma(x) + i\eta(x)\big)\,, \tag{1.211}$$

where σ and η are both real.

The Lagrangian can then be rewritten as

$$\mathscr{L} = \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma + \frac{1}{2}\partial^\mu\eta\partial_\mu\eta + 2\lambda v^2\frac{1}{2}\sigma^2 - \frac{\lambda}{4}\Big(\sigma^2+\eta^2\Big)^2 - \lambda v\sigma\Big(\sigma^2+\eta^2\Big)\,, \tag{1.212}$$

so we have two real fields, one massive (σ) and one massless (η). The massless field is called a **Goldstone boson**.

If we compute the VEV of ϕ in this configuration we get $v/\sqrt{2}$.

In general, the Golstone theorem states that for each broken symmetry generator we get a massless boson.

It seems that instead of solving the problem of our massless gauge bosons we have created another problem, predicting more massless particles; but in fact our two problems solve each other.

We must break a different symmetry, though: a gauge local symmetry, instead of a global one. If we explicitly break a $U(1)$ global symmetry, we have a Goldstone boson. If we break a $U(1)$ local symmetry, instead, we get a massless vector boson and a massless scalar. This spontaneous breaking of local gauge symmetry is the **Higgs mechanism**.

SQED Higgs mechanism

We will see this explicitly, with the breaking of the scalar-QED $U(1)$ symmetry: let us consider a Lagrangian

$$\mathscr{L} = \big(\mathrm{D}^\mu\phi\big)^*\mathrm{D}_\mu\phi - \mu^2\big|\phi\big|^2 - \lambda\big|\phi\big|^4 - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\,, \tag{1.213}$$

where $\mathrm{D}_\mu = \partial_\mu + iqA_\mu$. This is similar to a Lagrangian describing a fermion, but we substitute it with a complex scalar field with charge q.

This Lagrangian has a $U(1)$ gauge symmetry, whose action is

$$\phi \to e^{-iq\alpha(x)}\phi \tag{1.214}$$

$$\phi^* \to e^{iq\alpha(x)}\phi^* \tag{1.215}$$

$$A_\mu \to A_\mu + \partial_\mu\alpha\,. \tag{1.216}$$

If $\mu^2 > 0$ we have no issues: the only minimum is the global one at $\phi = A_\mu = 0$. This is the only configuration which is Lorentz invariant: any other vector would define a direction, breaking Lorentz symmetry.

The last point also holds if $\mu^2 < 0$; so the possible vacua are those defined as $\phi = ve^{i\theta}/\sqrt{2}$ and $A^\mu = 0$. Let us perturb as we did before: if we define

$$\phi(x) = \frac{1}{\sqrt{2}}\big(v + \sigma(x) + i\eta(x)\big)\,, \tag{1.217}$$

the Lagrangian can be rewritten as

$$\begin{aligned} \mathscr{L} &= \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma - \frac{1}{2}(2\lambda v^2)\sigma^2(x) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\ &\quad + \frac{1}{2}(qv)^2 A_\mu A^\mu + \frac{1}{2}\partial^\mu\eta\partial_\mu\eta + qvA^\mu\partial_\mu\eta + \text{higher-order interactions}\,. \end{aligned} \tag{1.218}$$

This Lagrangian describes a real scalar σ with mass $\sqrt{2\lambda v^2}$; a massless real scalar η; a massive vector boson and the interactions between them.

Let us count degrees of freedom: we started with a complex scalar ϕ (2 dof) and a massless vector (2 dof); now we have two real scalars (1 dof each) and a massive vector (3 dof): we have gone from a total of four to a total of five! So, there must be an unphysical field between these.

In fact, we can do a gauge transformation (going to the *unitary gauge*) so that ϕ is real everywhere: this removes the need for η. After this manipulation, the Lagrangian will read

$$\begin{aligned} \mathscr{L} &= \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma - \frac{1}{2}(2\lambda v^2)\sigma^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &\quad + \frac{1}{2}(qv)^2 A_\mu A^\mu + \text{higher-order interactions}\,, \end{aligned} \tag{1.219}$$

so we are left with a massive scalar σ (1 dof) and a massive vector A_μ (3 dof): the total is 4 again!

It is a "transmutation" of degrees of freedom: one dof has gone from the scalar to the vector making it massive.

This is the **Higgs mechanism**.

Wednesday 2020-4-22, compiled 2020-07-02

What would $\mu^2 < 0$ physically mean? It would be a tachyonic particle.

The field $\sigma(x)$ is called the **Higgs field**; the would-be Goldstone boson disappears yielding the longitudinal polarization of $A_\mu(x)$.

We are curing two different problems: we have found massive gauge bosons, and removed the unphysical Goldstone bosons.

A more general formulation: $SU(2)$ symmetry breaking

This section follows Peskin pretty closely [Pes19, sec. 16.2].

We can repeat this in the general case, with the group G being broken to G'. Suppose, for clarity, that this is $SU(2)$ being broken to $U(1)$.

The three vectors A^a_μ generate rotations around the axes x^a respectively, as $a = 1, 2, 3$.

The adjoint representation of $SU(2)$ is given by three real scalar fields ϕ^a; their covariant derivative is given by

$$D_\mu\phi^a = \partial_\mu\phi^a + g\epsilon^{abc}A^b_\mu\phi^c\,, \tag{1.220}$$

since the representation matrices in the adjoint representation are $(t^b_G)_{ac} = i f^{abc}$.

We want to choose a potential $V(\phi)$ which is minimized by a configuration with $\left\langle \left|\phi^a\right| \right\rangle = v$. A vacuum for this potential is, for example, given by $\phi^a = v\delta^{a3}$: it retains some rotational invariance (around the $\hat{3}$ axis), but invariance for rotations around the other two axis is broken.

A perturbative expansion around such a vacuum will then be given by

$$\phi(x) = \left(\pi^1(x), \pi^2(x), v + h(x)\right), \tag{1.221}$$

where $\pi^{1,2}$ are the would-be goldstone bosons, which will contribute to the longitudinal components of $A^{1,2}_\mu$, the bosons which become massive.

On the other hand, A^3_μ remains massless. As we go to the unitary gauge we can eliminate $\pi^{1,2}$: so we get

$$\phi(x) = (0, 0, v + h(x)) . \tag{1.222}$$

We know that

$$\begin{aligned} \mathrm{D}_\mu \phi^a &= \partial_\mu h \delta^{3a} + g\epsilon^{abc} A^b_\mu \phi^c & (1.223)\\ &= \partial_\mu h \delta^{3a} + g\epsilon^{ab3} A^b_\mu (v + h(x)) . & (1.224) \end{aligned}$$

The kinetic term of ϕ^a will then become:

$$\begin{aligned} \frac{1}{2}\left(\mathrm{D}_\mu \phi^a\right)^2 &= \frac{1}{2}\mathrm{D}_\mu \begin{bmatrix} 0 & 0 & v+h \end{bmatrix} \mathrm{D}^\mu \begin{bmatrix} 0 \\ 0 \\ v+h \end{bmatrix} & (1.225)\\ &= \frac{1}{2}\left[g\epsilon^{ab3} A^b_\mu (v+h) + \partial_\mu h \delta^{a3}\right]^2 & (1.226)\\ &= \frac{g^2 v^2}{2} \epsilon^{ab3} A^b_\mu \epsilon^{ad3} A^{d,\mu} + \mathcal{O}(h) & (1.227)\\ &= \frac{g^2 v^2}{2}\left(A^1_\mu A^{\mu 1} + A^2_\mu A^{\mu 2}\right) + \mathcal{O}(h) . & (1.228) \end{aligned}$$

Square entails contraction of both a and μ.

So, $A^{1,2}$ became massive, with mass $M^2_W = g^2 v^2$, while A^3 remained massless.

As $SU(2)$ is broken to $U(1)$, the 3 generators A^a_μ, with $a = 1, 2, 3$ are broken to give two massive vector bosons ($W^\pm$) and one massless vector boson (the photon), while the real field σ is the Higgs field.

The two W bosons are not actually the components $A^{1,2}$: instead, we choose them to be the eigenvalues of rotations around the z axis:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(A^1_\mu \mp i A^2_\mu\right), \tag{1.229}$$

since the rotation matrix around the z axis in the spin-1 representation is given by

$$J^3 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{1.230}$$

whose eigenvectors are indeed $(1, \mp i, 0)$ with eigenvalues ± 1.

This is called the **Georgi-Glashow model**. It predicts two massive bosons of the $V - A$ interaction. However, it does not work well phenomenologically.

Hypercharge: why the GG model does not work

This was a computationally easy example to clarify what this mechanism looks like, not the one we actually will use: we will have to choose another symmetry group.

The issue lies with the fact that we are associating electric charge with the generator of rotations about $\hat{3}$. We know that an electron current can be turned into an electronic neutrino current, and that the neutrino is electrically neutral. They must belong to the same isospin multiplet in order for this to happen. Isospin (I) is what we call rotation in the three-dimensional space the $SU(2)$ symmetry was initially defined in. Rotation about the $\hat{3}$ axis then corresponds to the operator I_3; the algebra of these operators is that of a rotation representation, $[I_i, I_j] = i\epsilon_{ijk} I_k$. When we say "isospin multiplet" we mean that the value of $I^2 = j(j+1)$ is fixed; then the value of I_3 will go from $-j$ to j.

We cannot have a doublet with only ν and e^- because of the fact that the neutrino is neutral: the eigenvalues of the rotation about the 3 axis would have to be 0 and 1. We need a particle with eigenvalue -1 with respect to rotations about the $\hat{3}$ axis, so with positive charge.

We then predict the existence of a heavy electron E^+, such that the triplet looks like

$$\begin{bmatrix} E^+ \\ \nu \\ e^- \end{bmatrix}, \tag{1.231}$$

where the eigenvalues of this rotation around the $\hat{3}$ axis are $-1, 0, 1$ respectively. This heavy electron was not observed; the experimental bound is at $M_{E^+} \leq 400\,\text{GeV}$ currently. We would expect to then see processes like $u + E^+ \to d + \nu$, but we do not.

Is the positron e^+ not a candidate for this? Why do we say that this particle would need to be "heavy"? Or are we already implicitly considering antiparticles maybe?

This will be a way to unify electromagnetic and weak interactions, and we will get to use a single coupling constant for our new electroweak theory.

1.4.6 Electroweak symmetry breaking

The correct symmetry group for the electroweak theory, as Glashow showed in 1961, is

$$SU(2)_L \otimes U(1)_{\text{hypercharge}}, \tag{1.232}$$

where the hypercharge is usually denoted as Y. We have three vector fields for $SU(2)$, which we call A^i_μ, and also a vector field for $U(1)$: B_μ. Note that these do not have a direct correspondence with the fields we get after SSB: hypercharge is different from electromagnetic charge.

The pair neutrino-electron is a $I = 1/2$ doublet under $SU(2) \times U(1)$, we will see that one component of this doublet has charge $+$ while the other has charge 0: its transformation will look like

$$\begin{bmatrix} \nu_L \\ e_L^- \end{bmatrix} \rightarrow \underbrace{e^{-i\vec{\alpha}\cdot\vec{\sigma}/2}}_{SU(2)} \underbrace{e^{-i\beta/2}}_{U(1)} \begin{bmatrix} \nu_L \\ e_L^- \end{bmatrix} . \tag{1.233}$$

The bosons in this theory start off as massless, and therefore the fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(A_\mu^1 \mp i A_\mu^2 \right) \tag{1.234}$$

are massless as well. However, we want massive W bosons with $M_W \sim 100\,\text{Gev}$! At the end of the sixties Weinberg and Salam introduced the Higgs mechanism in this theory. The idea is to interpret the scalar field as a $I = 1/2$ representation of rotation: it will transform under the symmetries of the theory as

$$\varphi = \begin{bmatrix} \varphi^+ \\ \varphi^0 \end{bmatrix} \rightarrow e^{i\vec{\alpha}\cdot\vec{\sigma}/2} e^{i\beta/2} \begin{bmatrix} \varphi^+ \\ \varphi^0 \end{bmatrix} . \tag{1.235}$$

So, the lepton doublet has a charge $-1/2$ with respect to the $U(1)$ symmetry, while the field has a symmetry $+1/2$.

Suppose the potential of the field reads:

$$V(\phi) = -\mu^2 \left|\phi\right|^2 + \lambda \left|\phi\right|^4 , \tag{1.236}$$

whose minimum is defined by

$$0 = -2\mu^2 \phi + 4\lambda \phi \left|\phi\right|^2 , \tag{1.237}$$

therefore the minimum is a sphere (since the two complex fields correspond to four real degrees of freedom) with $\left|\phi\right|^2 = \left|\phi^+\right|^2 + \left|\phi^0\right|^2 = \mu^2/2\lambda$; so we define $v = \mu/\sqrt{\lambda}$.

Up to a rotation, the VEV of our field will be

$$\left\langle \phi \right\rangle_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v \end{bmatrix} . \tag{1.238}$$

This vacuum breaks the $SU(2)_L \times U(1)_Y$ symmetry. We can expand around it in the following manner:

$$\phi(x) = \begin{bmatrix} \pi^+(x) \\ \frac{v + h(x) + i\pi^3(x)}{\sqrt{2}} \end{bmatrix} , \tag{1.239}$$

where $\pi^+ = (\pi^1 + i\pi^2)/\sqrt{2}$ is a complex perturbation encompassing two real degrees of freedom.

As we did before, we can gauge away the unphysical Goldstone bosons $\pi^{1,2,3}$; the real field $h(x)$ remains.

How do the bosons of this theory **couple to fermions**? The couplings can only have specific forms, fixed by gauge invariance: the covariant derivative will read:

$$\mathrm{D}_\mu \Psi = \left(\partial_\mu - igA^a_\mu I^a - ig' B_\mu Y \right) \Psi \,, \tag{1.240}$$

where g and g' are the coupling of $SU(2)$ weak isospin and $U(1)$ hypercharge respectively; I^a are the generators of $SU(2)$ in the $SU(2)$ representation acting on Ψ, and Y is the hypercharge, a scalar.

In principle, g and g' are independent, so we keep them distinct. Let us now apply D_μ to the field ϕ around the vacuum with only the h perturbation, as

$$\mathrm{D}_\mu \phi = \left(\partial_\mu - igA^a_\mu I^a - ig' B_\mu Y \right) \begin{bmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{bmatrix} \,, \tag{1.241}$$

of which we take the square norm, up to zeroth order in h:

$$\left| \mathrm{D}_\mu \phi \right|^2 = \frac{1}{2} \begin{bmatrix} 0 & v \end{bmatrix} \left(gA^a_\mu \frac{\sigma^a}{2} + g' B_\mu \frac{1}{2} \right) \left(gA^{b,\mu} \frac{\sigma^b}{2} + g' B^\mu \frac{1}{2} \right) \begin{bmatrix} 0 \\ v \end{bmatrix} (+\mathcal{O}(h)) \tag{1.242}$$

$$= \frac{1}{2} \underbrace{g^2 v^2 \frac{1}{4} \left(A^1_\mu A^{1,\mu} + A^2_\mu A^{2,\mu} \right)}_{\underbrace{(gv/2)^2}_{M_W^2} W^+_\mu W^{\mu,-}} + \frac{1}{2} \frac{v^2}{4} \left(-gA^3_\mu + g' B_\mu \right)^2 . \tag{1.243}$$

So, we have found two massive $W^\pm$ bosons, and the **combination** $-gA^3_\mu + g'B_\mu$ has also gained mass.

Let us define θ_w, the Weinberg angle, by $\tan\theta_w = g'/g$. Then we will have

$$\cos\theta_w = \frac{g}{\sqrt{g^2 + g'^2}} \qquad \text{and} \qquad \sin\theta_w = \frac{g'}{\sqrt{g^2 + g'^2}} \,, \tag{1.244}$$

and we define the electromagnetic field A_μ and the Z_μ boson by the two orthogonal combinations:

$$A_\mu = \sin\theta_w A^3_\mu + \cos\theta_w B_\mu \tag{1.245}$$

$$Z_\mu = \cos\theta_w A^3_\mu - \sin\theta_w B_\mu \,, \tag{1.246}$$

so that the kinetic term reads

$$\frac{1}{2} \frac{v^2}{4} \left(-gA^3_\mu + g' B_\mu \right)^2 = \frac{m_Z^2}{2} Z^\mu Z_\mu \,, \tag{1.247}$$

where $m_Z^2 = (g^2 + g'^2)v^2/4$, while there is no mass term for A: so, $m_A = 0$.

The residual gauge symmetry can be identified with $U(1)_{\mathrm{em}}$. When the symmetry is broken, we are left with a long-range interaction and a short range one.

The combination under which the vacuum is invariant is

$$T^3 + Y\,, \tag{1.248}$$

where T^3 is the third generator of $SU(2)$, while Y is the generator of hypercharge. This corresponds to

$$\phi \to e^{i\vec{\alpha}\cdot\vec{\sigma}/2} e^{i\beta/2} \phi\,, \tag{1.249}$$

with $\alpha^3 = \beta$: this is due to the fact that

$$\frac{1}{2}\left(\sigma^3 + \mathbb{1}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\,, \tag{1.250}$$

whose exponential leaves the second component of $\phi = (0, v/\sqrt{2})$ invariant.

In summary, the SSB has the form

$$SU(2)_L \times U(1)_Y \to U(1)_{\text{em}}\,, \tag{1.251}$$

and it yields the two $W^\pm$ bosons, each with mass $gv/2$, and the Z boson with mass $\sqrt{g^2 + g'^2}v/2$. The masses of the W and Z bosons are related by:

$$m_W = m_Z \cos\theta_w\,, \tag{1.252}$$

so if we can devise some experiment which will measure θ_w we can predict m_W/m_Z.

The full expression of the covariant derivative reads, using the notation $c_w = \cos\theta_w$, $s_w = \sin\theta_w$ and $\sigma^\pm = \left(\sigma^1 \mp i\sigma^2\right)/\sqrt{2}$:

not sure about the last one

$$\begin{aligned} D_\mu \Psi &= \left[\partial_\mu - i\frac{g}{\sqrt{2}}\left(W_\mu^+\sigma^+ + W^-\sigma^-\right) - ig\left(c_w Z_\mu + s_w A_\mu\right) I^3 - ig\left(-s_w Z_\mu + c_w A_\mu Y\right)\right]\Psi \qquad (1.253) \\ &= \left[\partial_\mu - i\frac{g}{\sqrt{2}}\left(W_\mu^+\sigma^+ + W^-\sigma^-\right) - ieA_\mu Q - i\frac{g}{c_w} Z_\mu Q_Z\right]\Psi\,, \qquad (1.254) \end{aligned}$$

where we defined $e = gs_w = g'c_w$, and

$$Q = I^3 + Y \qquad \text{and} \qquad Q_Z = I^3 - s_w^2 Q\,. \tag{1.255}$$

The next bit is not super clear.

If the mass of the fermion described by Ψ is zero we can decouple the left- and right-handed parts of the spinor. We will get interactions between the left handed components, not between the right-handed components, since the $W^\pm$ do not couple with them.

How does that come about exactly?

We recover the $V-A$ structure: the left- and right-handed spinors have different quantum numbers with respect to $SU(2)_L$ and $U(1)_Y$ but the same Q, which is the quantum number corresponding to $U(1)_{\text{em}}$.

The ψ_L are $SU(2)_L$ **doublets**, with $I = 1/2$; the ψ_R are $SU(2)_L$ **singlets**, with $I = 0$.

We then choose the values of the hypercharge such that the values of Q are compatible with experiment.

Particle	I^3	Y	Q
$\nu_{e,L}$	$+1/2$	$-1/2$	0
$\nu_{e,R}$	0	0	0
e^-_L	$-1/2$	$-1/2$	-1
e^-_R	0	-1	-1
u^-_L	$+1/2$	$1/6$	$2/3$
u^-_R	0	$2/3$	$2/3$
d^-_L	$-1/2$	$1/6$	$-1/3$
d^-_R	0	$-1/3$	$-1/3$

Figure 1.4: Electroweak quantum numbers of the fermions.

In terms of $SU(2)_L$ multiplets, we separate:

$$\begin{bmatrix} \nu_L \\ e^-_L \end{bmatrix} \qquad e^-_R \qquad \begin{bmatrix} u_L \\ d_L \end{bmatrix} \qquad u_R \qquad d_R\,. \tag{1.256}$$

This is a **generation** (or family) of fermions: the "electronic" one, and we have two more: the muonic and tauonic ones.

Is this next bit useful?

If μ^2 is a function of T, then we can get a phase transition. This is called the electroweak phase transition.

1.4.7 Neutral current interactions

Tuesday 2020-4-28, compiled 2020-07-02

We have introduced a term in our model which yields an important and testable prediction: the existence of a weak *neutral* current in addition to the charged ones. Let us make some simple examples for the three types of currents we have: the charged current with the $W^\pm$ bosons mediates interactions like

$$\nu_L + d_L \to e^-_L + u_L\,, \tag{1.257}$$

the electromagnetic current A_μ and the neutral current Z^0_μ mediate interactions like

$$e^-_L + u_L \to e^-_L + u_L\,, \tag{1.258}$$

however there is a distinction to be made: the electromagnetic current can only be coupled to charged particles, so a process like

$$\nu_L + u_L \to \nu_L + u_L \tag{1.259}$$

can only be mediated by the Z^0 boson, since the neutrino is neutral! This kind of interaction has been experimentally observed!

So, there are certain processes which can be mediated by both the Z^0 boson and the photon. Does this mean that people needed to revise their estimate of the electromagnetic coupling, since electromagnetism was not the only contributor to those scatterings?

1.5 The Standard Model of particle physics

The Lagrangian reads:

$$\boxed{\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\overline{\psi} \not{D} \psi + \psi_i y_{ij} \psi_j \phi + \text{h. c.} + \left| D_\mu \phi \right|^2 - V(\phi) \,.} \tag{1.260}$$

Let us unpack the terms:

1. $F_{\mu\nu} F^{\mu\nu}$ encompasses the kinetic and self-interaction terms of all the interaction bosons (photon, $W^\pm$, Z^0, gluons);
2. $i\overline{\psi} \not{D} \psi$ encompasses all the kinetic terms of the fermions ($\overline{\psi} \gamma^\mu \partial_\mu \psi$) and the interaction terms between the fermions and the gauge bosons: $g\overline{\psi} \gamma^\mu A_\mu \psi$;
3. $\psi_i y_{ij} \psi_j \phi$ is the interaction between the Higgs field and the fermions, which gives mass to the fermions and which we will discuss in the next section;
4. $\left| D_\mu \phi \right|^2$ is the kinetic term of the scalar (Higgs) field ϕ, also including the interactions between it and the vector bosons;
5. $V(\phi)$ is the potential for the scalar field, it must be at least quartic.

"h. c." means Hermitian Conjugate, it is needed in order to ensure that everything is self-adjoint.

The symmetry which determines this theory is

$$\text{Lorentz} \otimes SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \,. \tag{1.261}$$

Starting from the symmetry, the only other prescription needed are the quantum numbers of the various particles: we can make a table. As we have seen, the fermions are divided into left-handed isospin doublets and right-handed singlets.

	Name	T_3 (SU(2))	Y	Q
Leptons (no $SU(3)$ charge)	ν_L	$+1/2$	$-1/2$	0
	e^-	$-1/2$	$-1/2$	-1
	$\cancel{\nu_R}$	0	0	0
	e_R^-	0	-1	-1
Higgs scalar	ϕ^+	$+1/2$	$+1/2$	$+1$
	ϕ^0	$-1/2$	$+1/2$	0
Quarks ($SU(3)_c$ triplets)	u_L	$+1/2$	$1/6$	$2/3$
	d_L	$-1/2$	$1/6$	$-1/3$
	u_R	0	$2/3$	$2/3$
	d_R	0	$-1/3$	$-1/3$

Figure 1.5: Particles

Here we write only one of the generations for both leptons and quarks. The quantum numbers for the other generations are the same. An unanswered question is *why* exactly there are three families.

We are always computing the electric charge as $Q = T_3 + Y$, where T_3 is the generator of $SU(2)_L$ rotations around the $\hat{3}$ axis. Sometimes we write "1" for the right-handed quarks under T_3, this is not the eigenvalue but it is used to mean "singlet".

For the $SU(3)$ we have 8 generators (corresponding to the gluons), for $SU(2)$ we have three generators, for $U(1)$ we only have 1 generator: these four, after SSB, are three massive bosons and the photon.

The assignment of the charges cannot come from the symmetry, it must be determined experimentally.

The right-handed neutrino does not appear here: it should be invariant under any gauge transformation. This would then be the *sterile* neutrino: it would not interact with matter by any of the forces of the Standard Model.

It makes sense not to include it *a priori* then, however we can keep it in mind: it is a DM candidate.

Can we say anything about its possible mass? There are schemes in which the RH neutrino enters, and in which it can be assigned a mass. However, in general this is a free parameter: in some contexts we can give it a specific value, but we cannot say anything about its mass from basic Standard Model theory.

There are experiments at Fermilab which could point indirectly at their existence.

The rule is always:

> *Write all the terms in your Lagrangian which are invariant under your symmetry and which have dimension* 4.

Should we not include also some prescription for the couplings? No, this is included in the symmetry: if we have the coupling constants g_S, g (weak) and g' (hypercharge) we can determine everything else.

The issue is that, before including the Higgs boson, everything is massless.

Gluons and quarks have the same "destiny": there is confinement when g_S becomes large.

The charged Higgs component ϕ^+ gives mass to the $W^\pm$ vector bosons of the electroweak interaction, the ϕ^0 gives mass to the Z^0 vector boson.

1.5.1 Fermions' mass terms

We have the massless G^a and A^μ, and the massive $W^\pm$ and Z^0. These are the terms we've written so far: where do the $\overline{\psi} m \psi$ terms come from?

The kinetic term has dimension 2^2, which is fine, the wavefunction has dimension $3/2$ while the derivative has dimension 1. So, also a term $\overline{\psi} m \psi$ would be fine dimensionally. Is it symmetric? Let us consider the issue with the left- and right-handed components. If we were to write something like $\overline{\psi}_L \psi_L$ or $\overline{\psi}_R \psi_R$ we would get

$$\overline{\psi}_{L/R} \psi_{L/R} = \overline{\psi} \left(\frac{1 \mp \gamma_5}{2} \right) \left(\frac{1 \pm \gamma_5}{2} \right) \psi = 0 \,, \tag{1.262}$$

since we are computing the product of two orthogonal projectors.

So, the only terms which do not vanish look like $\overline{\psi}_L \psi_R = \overline{\psi}_R \psi_L = \overline{\psi} \psi$.

So he writes, but it does not work! We have two different components, $\overline{\psi}\psi$ and $\overline{\psi}\gamma^5\psi$!

Let us consider electrons for example: an object like $m \overline{e}_L e_R$ is a Lorentz invariant, but is it also gauge invariant?

The spinor e_L is in an $SU(2)$ doublet, while the e_R is an $SU(2)$ singlet. So, the object is not a singlet: it is not invariant under $SU(2)_L$, but each term in the Lagrangian must be a scalar with respect to the symmetry group (that is, it must be *symmetric*). In fact, this term is only $U(1)_{\text{em}}$ invariant, neither its isospin nor its hypercharge are zero.

So, we cannot include this term. Does this mean that fermions are massless? Before introducing ϕ, in the pure Yang-Mills theory, they indeed are.

There is only one kind of field which is invariant under spacetime transformations: a scalar. So, we could introduce a term like

$$\overline{\psi}^i_L \phi_i \psi_R \,, \tag{1.263}$$

where i is an index going from 1 to 2, an $SU(2)$ index: ϕ, the Higgs field, is an $SU(2)_L$ doublet! This is not a mass term, but an interaction term between the Higgs field and the electron-left and electron-right.

Then, the way the mechanism works is by the fact that the VEV of ϕ is $v \neq 0$, so at low energies we will see an effective mass term. The term we put in the Lagrangian will look like

$$y^{ij} \overline{\psi}^i_L \phi \psi^j_R \,, \tag{1.264}$$

where now the indices i and j run over the possible fermions. We will then have

$$M_\ell = y_\ell v \,, \tag{1.265}$$

and the matrix will be $y^{ij} = \delta^{ij} y^j$.

Similarly we will have mass terms for the leptons, up and down quarks. These are free parameters of the theory, and they must be different. The VEV of the Higgs field is fixed by the masses of the W and Z bosons: it is of the order 100 GeV.

But the mass of the electron is $m_e \sim 511\,\text{keV}$: so, we must have $y_e \sim 10^{-5}$.

The price we pay for this is the fact that the number of free parameters increases. This is a "dirty part" of the theory.

Why are these parameters so different? This is the *flavour problem*.

Stuff is not super clear here.

Let us write the most general set of Yukawa couplings [Pes19, eq. 18.24]:

$$\mathscr{L}_{\text{coupling}} = -y_e^{ij} L_a^{\dagger,i} \phi_a e_R^j + y_d^{ij} Q_a^{\dagger,i} \phi_a d_R^j - y_u^{ij} Q_a^{\dagger,i} \epsilon_{ab} \phi_b^* u_R^j + \text{h. c.} \tag{1.266}$$

The matrices y_f^{ij} are called the Yukawa matrices. The indices i and j label fermion generations (1, 2, 3 for electron-lik, muon-like and tau-like). They are not Hermitian in general, however we can construct

$$y_f y_f^\dagger = U_L^{(f)} Y_f U_L^{(f),\dagger} \qquad \text{and} \qquad y_f^\dagger y_f = U_R^{(f)} Y_f U_R^{(f),\dagger}\,, \tag{1.267}$$

and if the matrix $\sqrt{Y_f}$ is block-diagonal in generation space then we can write

$$y_f = U_L^{(f)} \sqrt{Y_f} U_R^{(f)\dagger}\,. \tag{1.268}$$

So, if we change variables as

$$e_R^i \to U_{R,ij}^{(e)} e_R^j \qquad \text{and} \qquad L^i \to U_{L,ij}^{(e)} L^j\,, \tag{1.269}$$

the matrices $U_{L,R}$ disappear from the couplings.

This works well for the leptons; for the quarks there is an issue, in their coupling to the W boson we are left with a term

$$u_L^\dagger (i\overline{\sigma}^\mu) d_L = u_L^\dagger U_L^{(u)\dagger} (i\sigma^\mu) U_L^{(d)} d_L \tag{1.270}$$

$$= u_L^\dagger (i\overline{\sigma}^\mu) \underbrace{U_L^{(u)\dagger} U_L^{(d)}}_{V_{CKM}} d_L\,. \tag{1.271}$$

But why does this happen? what is it about the quarks that makes those matrices not cancel?

This all has to do with the fact that mass eigenstates and flavour eigenstates are different, methinks.

In the V_{CKM} matrix we have the freedom to choose three (Euler) angles (it is a rotation matrix basically) and a phase.

1.5.2 Symmetries of the SM

Wednesday 2020-4-29, compiled 2020-07-02

Let us write the Standard Model Lagrangian *after* spontaneous symmetry breaking (so, at low energies):

$$\mathscr{L} = -\frac{1}{4}\sum_a \left(F^a_{\mu\nu}\right)^2 + m_W^2 W^+_\mu W^{-,\mu} + \frac{1}{2}m_Z^2 Z_\mu Z^\mu + \sum_f \overline{\psi}_f \left(i\not{D} - m_f\right)\psi_f + \frac{1}{2}\left(\partial_\mu h\right)^2 - V(h)\,, \tag{1.272}$$

where a runs over all our $8+3+1 = 12$ gauge bosons (8 gluons, the $W^\pm$ and Z^0, the photon), and the covariant derivative is given by:

$$D_{\mu f} = \partial_\mu - ieQ_f A_\mu - i\frac{g}{\cos\theta_w} Q_{Zf} Z_\mu - ig_s A^a_\mu t^a\,. \tag{1.273}$$

Something that should have come up before: what's up with the W mass term? It does not look like a usual mass term…

The weak force interactions are diagonal for the leptons, proportional to V_{CKM} for the quarks.

Let us discuss the remaining symmetries of the model. These are symmetries which are not imposed on the model by hand, instead they follow automatically from the terms which are allowed by the chosen gauge symmetries.

Baryon and Lepton number conservation

We define

$$B(\text{Quark}) = \frac{1}{3} \qquad \text{and} \qquad B(\text{antiQuark}) = -\frac{1}{3}\,, \tag{1.274}$$

while for any other particle we assign 0. This is conserved, but we need not impose it: the most general Lagrangian we write with our symmetries has it. This is a global symmetry: $U(1)_B$.

The way to find it is to construct a baryonic current j^μ_B, in such a way that it is conserved in our Lagrangian, then the integral of j^0_B will give the conserved charge.

This has heavy consequences: the lightest baryon is the proton. If it were to decay, this would not conserve baryon number.

If we are allowed to violate baryon number conservation we can have a decay like $p \to e^+ + \gamma$, which would occur and destabilize atoms.

We get a bound on the lifetime of the proton of

$$\tau_{\text{proton}} > 10^{32}\,\text{yr}\,. \tag{1.275}$$

We also have lepton number conservation. This is more interesting since we have neutrinos, which are hard to detect.

We will see a method to give mass to neutrinos, which will however violate lepton number conservation.

C, P and T symmetries

Strong and EM interaction conserve these three symmetries, while the coupling of the weak bosons to the fermions violate C and P maximally. If the couplings were real-valued, CP symmetry would be preserved, but the V_{CKM} matrix has a physical phase, so this is not the case.

By the CPT theorem, T is also violated.

Flavour number conservation

The only thing which couples different families of fermions together is the CKM matrix, which gives mixing terms between the quarks, as

$$\overline{u}_i \gamma^\mu (V_{CKM})_{ij} \phi_j \,, \tag{1.276}$$

and the Cabibbo angle measures how much this happens; we then expect to see decays like $u \to s + W^+$.

For leptons, instead, we expect perfect conservation of lepton number since there is no mixing. Experimentally this seems to be quite well verified: if process like $\mu \to e\gamma$ or $\tau \to e\gamma$ happen they must do so extremely rarely.

1.5.3 Neutrino masses

In the mass term we add to the Yukawa Lagrangian, the index a is an SU(2) index, and the Higgs field looks like

$$\phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} ; \tag{1.277a}$$

we have two terms like

$$\begin{bmatrix} \nu_L & e_L^- \end{bmatrix} \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} e_R \,, \tag{1.278a}$$

and also

$$\begin{bmatrix} \nu_R^c & e_R^+ \end{bmatrix} \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} e_R^- \,. \tag{1.279a}$$

This means that the neutrino is massless: the VEV of ϕ_a is $(0, v)$, so only the **second** component gains mass after SSB. So, we get Dirac mass terms for the electron but not for the neutrino.

This is equivalent to saying that the mass matrix for the charged leptons is diagonal: there is no mixing.

How can we detect whether neutrinos have mass, **experimentally**? A path is **β-decay**:

$$n \to p + e^- + \overline{\nu}_e \,. \tag{1.280}$$

We look at the so-called Kurie plot: on the horizontal axis we put the measured energy of the electrons E_e, on the vertical axis we put the square root of the number of electrons measured at that energy (in a small range around it: it is a histogram): $\sqrt{N_e}$. This plot is expected to look like a straight line if the neutrinos are massless, with some electrons reaching the maximum center-of-mass energy; on the other hand if neutrinos are massive the kinematics of the process change and the curve dips down earlier.

People have measured this to a very high degree of precision. What we have found up to now is a bound $m_{\nu_e} < 2\,\mathrm{eV}$ in these beta decay experiments.

The result that the neutrinos are massive did not come from here.

Other paths are **neutrino oscillations** and bounds from **cosmology**. From the neutrino oscillations we have the result that, at least for some neutrinos, $m_\nu \neq 0$.

However, in the Standard Model neutrinos are massless. How can we model their having mass?

Ettore Majorana proposed a way to give a mass for the neutrino (or any neutral fermion, really).

The Dirac Lagrangian, as we saw, after SSB reads:

$$\mathscr{L}_{\text{Dirac}} = \psi_R^\dagger (i\sigma \cdot \partial)\psi_R + \psi_L^\dagger (i\overline{\sigma} \cdot \partial)\psi_L - m\Big(\psi_R^\dagger \psi_L + \psi_L^\dagger \psi_R\Big)\,, \tag{1.281}$$

where $m = m_f = y_f v/\sqrt{2}$. However, in our SM we did not introduce a right-handed neutrino! So, one possibility to give mass to the left-handed neutrino is to introduce a right-handed one, with a term like

$$\Delta\mathscr{L}_{\text{Yukawa}} = -y_\nu^{ij} L_a^{\dagger,i} \epsilon_{ab} \phi_b^* \nu_R^j\,, \tag{1.282}$$

in analogy to the term we have for quarks. This yields a term $\nu_L^\dagger \phi_0 \nu_R$, which near the vacuum gives mass to the neutrinos, with a term like

$$m_\nu^{ij} = -y_\nu^{ij} \frac{v}{\sqrt{2}}\,. \tag{1.283}$$

We know that if the neutrinos do indeed have a mass it is very small, of the order of an eV or less. So, we choose a basis in which the interaction of the ν_L with the charged leptons are "diagonal": $\overline{\nu}_L^i \gamma^m \ell_L^i W_\mu$, where ℓ is a lepton (electron, muon, tauon).

As before we perform a change of variables making use of the unitary matrices U:

$$e_R^i \to U_{R,ij}^{(e)} e_R^j \qquad \text{and} \qquad L^i \to U_{L,ij}^{(e)} L^j\,. \tag{1.284}$$

Then, the fermion mass couplings will be:

$$y_\ell = U_L^{(e)} Y_\ell U_R^{(e),\dagger}\,, \tag{1.285}$$

and they can be chosen to be diagonal. However, in general y_ν is not diagonalized by these unitary matrices!

The current eigenstates $\nu_L^{(e,\mu,\tau)}$ which are produced at the charged current vertex are **different** from the mass eigenstates $\nu_{1,2,3}$ with masses $m_{\nu_{1,2,3}}$.

If we wish to diagonalize y_ν, we can do so as

$$y'_\nu = U_L^{(e),\dagger} y_\nu = U_L^{(\nu)} Y_\nu U_R^{(\nu)} . \tag{1.286}$$

The transition matrices U have three angles and one phase in freedom (which is to say, the other degrees of freedom in the matrix can be absorbed into global phases for the other fields). They are the Pontecorvo-Maki-NakagawaSakata mixing matrix.

Let us connect these musings to experiment. From cosmic ray interactions we have charged pions, which decay into

$$\pi^+ \to \mu^+ + \nu_\mu , \tag{1.287}$$

where the muonic neutrino we get is a *flavour* eigenstate, since it is produced in a weak-interaction vertex; it is a linear combination of the mass eigenstates.

Then, the muon decays into

$$\mu^+ \to e^+ + \nu_e + \overline{\nu}_\mu . \tag{1.288}$$

So, naively we expect twice as many ν_μ as ν_e.

Now, we know that for any particle $p = \sqrt{E^2 - m^2} \approx E - m^2/2E$ as long as m is small compared to E (which is definitely the case for our very light neutrinos). So, the momenta of the mass eigenstates of the neutrinos at a fixed energy E are

$$p_i \approx E - \frac{m_i^2}{2E} , \tag{1.289}$$

for $i = 1, 2, 3$. So, if the masses are different then the momenta will be slightly ($\mathcal{O}(m/E)$) different.

The wavefunction of the muonic neutrino ν_μ will look like

$$\sum_i V^{PMNS}_{(\mu),i} e^{i(E - m_i^2/2E)x} , \tag{1.290}$$

so we can see that the different components will go out of phase! This is very interesting, since even though the difference in the momenta is small we can measure at macroscopically different values of x, so that the phase difference is large!

Let us consider only two neutrino species for simplicity: ν_e and ν_μ. In order to get the mass eigenstates from the flavour ones we will have a matrix V depending on a mixing angle θ:

$$V = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} . \tag{1.291}$$

Then, the probability of observing a ν_μ given that we started with one will look like

$$\mathbb{P}(\nu_\mu \to \nu_\mu) = \left| \cos^2\theta e^{-i(m_1^2/2E)x} + \sin^2\theta e^{-i(m_2^2/2E)x} \right|^2 \tag{1.292}$$

$$= 1 - \sin^2(2\theta) \sin^2 \left(\frac{\delta m^2}{2E} x \right), \tag{1.293}$$

so we will observe an oscillation length scale of

$$L = \frac{4\pi E}{\delta m^2} \approx 2.48\,\mathrm{m} \left(\frac{E}{\mathrm{MeV}} \right) \left(\frac{\mathrm{eV}^2}{\delta m^2} \right). \tag{1.294}$$

The amplitude of this oscillation is maximal if $\theta = \pi/4$. Note that in order for oscillations to occur we need $\theta \neq 0$ and $\delta m^2 \neq 0$.

Experimentally, we see them: at a neutrino detector we distinguish downward neutrinos, which come from the sky above the detector, and upward neutrinos, which come from the other side of the Earth (through it). We see the correct ratio $\nu_\mu / \nu_e \approx 2$ for downward neutrinos, while for upward neutrinos the muonic flux is suppressed (and the electron-neutrino flux is isotropic).

So, we are seeing $\nu_\mu \leftrightarrow \nu_\tau$ flavour mixing with a length scale comparable to the Earth's diameter.

Are we sure about this? could it not be that the length scale is (much) smaller, and we are underestimating the suppression?

The parameters we get are

$$\delta m^2_{(\mu,\tau)} \approx 2.4 \times 10^{-3}\,\mathrm{eV}^2 \approx \left(5 \times 10^{-2}\,\mathrm{eV} \right)^2, \tag{1.295}$$

and $\sin^2\theta \approx 0.4$.

Also, we detect fewer neutrinos from the Sun than we'd expect from the standard Solar model; this gives us bounds with regards to e to μ, τ oscillations on the order of $|\delta m| \approx 10^{-2}\,\mathrm{eV}$ and $\sin^2\theta \approx 0.3$.

From these measurement, we get that all the neutrino masses are within about a tenth of an electronVolt of each other, two of them being close. We do not know the hierarchy: is the lone one (conventionally 3) heavier or lighter than the other two?

We recently saw some indications of CP violations in neutrino physics: this is indicated by differences in the probabilities when considering oscillations between neutrino flavours and ones between *anti*neutrino flavours.

There are two possibilities in order to give mass to the neutrinos. The first is to introduce right-handed neutrinos, so that we have a Dirac mass term: $m\bar{\nu}_L \nu_R$.

This adds two degrees of freedom to our model, since each neutrino has 2 dof.

Now, ν_L has only two degrees of freedom: denoting by $\alpha$and β a Weyl index (spinorial, from 1 to 2) we can add to the Lagrangian a term like $\nu_L^\alpha \nu_L^\beta \epsilon_{\alpha\beta}$. This is a singlet under Lorentz transformations.

This term would violate lepton number conservation, since each of the left handed neutrinos increases L by 2! However we did not include it at the start, it was an accidental outcome of our model.

However, ν_L belongs to a $SU(2)_L$ doublet (ν_L, e_L), with $T_3^{\nu_L} = +1/2$. Therefore, the term $\nu_L \nu_L$ would also be a $SU(2)_L$ triplet.

Now, the bare neutrino has nonzero hypercharge and neither do two of them.

And so. . . ?

Could we do something like a Yukawa coupling, $\nu_L \nu_L \phi$? No, since ϕ is a $SU(2)_L$ doublet, and in the product between a doublet and a triplet we do not get a singlet. Also, this term would have nonzero hypercharge, so it would not be $U(1)_Y$ invariant either.

However, if we consider $\phi\phi$ we get a doublet times a doublet, and $2 \times 2 = 1 + 3$.

Does the spin algebra really work just like that?

So, we can add to the Lagrangian a term like

$$\Delta \mathscr{L} = y_{ij}^M \left(L^i_{a\alpha} \epsilon_{ab} \phi_b \right) \left(L^j_{c\beta} \epsilon_{cd} \phi_d \right) \epsilon_{\alpha\beta} \,, \tag{1.296}$$

where i, j are flavour indices, α, β are Weyl spinorial indices (from 1 to 2), while a, b are $SU(2)_L$ indices.

The tensor y_{ij}^M contains the Yukawa couplings, and it is adimensional.

If we dimension-count we find a dimension of 5, so we must divide by a mass: this will be a new mass parameter M, so that the term will be

$$y_{ij}^M \frac{L^i \phi L^j \phi}{M} \,. \tag{1.297}$$

Then, after SSB we substitute ϕ with its VEV to get

$$M_{ij}^\nu = y_{ij}^M \frac{v^2}{2M} \,. \tag{1.298}$$

The left handed neutrino ν_L is associated with its antiparticle, $(\nu^c)_R$. If we draw a Feynman diagram for this interaction, we will see a vertex with two fermionic lines coming towards it and two scalar ones. Both of the fermionic lines will have arrows pointing towards the vertex, since neither of them is conjugated! This shows lepton number violation graphically.

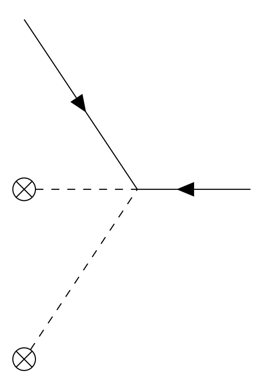

Figure 1.6:

This is a so-called **Majorana mass term**. It also entails lepton number violation.

By orders of magnitude, the Majorana mass term reads

$$M^{\nu}_{ij} \sim \frac{y^M_{ij} v^2}{M}\,, \tag{1.299}$$

and we know that $v \sim 100\,\text{GeV}$. If we assume $y \sim 10^{-1} \div 10^{-2}$ we find $M \sim 10^{13} \div 10^{14}\,\text{GeV}$. If, on the other hand, we assume $M \sim v$ then we must have $y \sim 10^{-13}$. Either y is very small or M is very large, or some combination of the two.

We also have the **seesaw mechanism**. If we add a right-handed neutrino to our model, we could also get terms like $\nu_R \nu_R$ (more precisely, $\nu_{R,\alpha}\nu_{R,\beta}\epsilon_{\alpha\beta}$): the right-handed neutrino is a singlet with respect to all the gauge symmetries (it being a singlet with respect to hypercharge follows from the fact that it is one with respect to $SU(2)_L$ and it is uncharged). This is why it is called a "sterile" neutrino, it is not affected by gauge transformations.

The term $\nu_R \nu_R$ is Lorentz and gauge invariant; like the $\nu_L \nu_L$ term it explicitly breaks lepton number conservation.

So, if we add to our Lagrangian a Majorana mass term for the right-handed neutrino:

$$\Delta \mathscr{L} = \frac{1}{2} M_{ij} \nu^i_{R,\alpha} \nu^j_{R,\beta} \epsilon_{\alpha\beta}\,, \tag{1.300}$$

BTW, why are we always antisymmetrizing the neutrinos' spinorial indices?

we have no issues related to SSB of the $SU(2)_L \times U(1)_Y$ gauge symmetry: M_{ij} can be at any energy scale it likes.

This is a sort of seesaw model: the mass matrix (on both axes we have ν_L and the ν_R) looks something like

$$\begin{bmatrix} 0 & m_{\text{Dirac}} \\ m_{\text{Dirac}} & M \end{bmatrix}. \tag{1.301a}$$

The mass of the right-handed neutrino is not "protected" by the gauge symmetry, it can be as large as it likes. This term then gives left-handed neutrinos very small masses by giving right-handed neutrinos very large ones.

It is important to find out whether the mass of the neutrino is due to a Majorana or Dirac mass term.

The experiments which can determine this are called *neutrinoless double beta decay*: a double β decay in which the two neutrinos produces would annihilate violating lepton number conservation.

If this were found, it would mean that neutrinos have a Majorana mass term. This concludes the discussion of the particle physics standard model. Next week we are going to start looking at the standard model of cosmology.

Chapter 2

Early Universe

2.1 Basics of cosmology

The classic textbook to be used for this part is Kolb-Turner [KT94]. For something more recent, we have Gorbunov and Rubakov [GR11]. The first part was also discussed in the course in Fundamentals of Astrophysics and Cosmology (whose notes are still being reviewed, they are more in-depth on these topics).

Tuesday 2020-5-5, compiled 2020-07-02

We establish some basic notions in cosmology. We use natural units $\hbar = c = k_B = 1$, so that we have the equivalence of mass, energy, temperature and angular frequency.

The Planck units are defined by also setting $G = 1$. Conventionally, we do not do so and instead define a Planck mass by $G = 1/M_P^2$, so that $M_P \approx 1.2 \times 10^{19}\,\text{GeV}$. Because of all the equivalencies we have, this also defines a length, a time, and an actual mass.

2.1.1 Homogeneous and isotropic universe

Our description of the universe begins with the simplest model, in which the metric only has one parameter, whose variation scales the size of the spatial coordinates. This is the Friedmann - Lemaître - Robertson - Walker (FLRW) metric, given by

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\gamma_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,, \tag{2.1}$$

where γ_{ij} is the metric of a unit 3-sphere, 3-hyperboloid or 3-plane. Which of these three we have determines the *spatial curvature* of the universe, described by a single parameter k which can be -1 for an open universe, 0 for a flat universe and $+1$ for a closed universe.

The Hubble parameter is given in terms of a, by $H(t) = \dot{a}(t)/a(t)$. It describes the rate of expansion of the universe. If it is written with an index 0, it means we are measuring it *now*.

Our model is general-relativistic, so we must solve the Einstein equations, which read

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}\,. \tag{2.2}$$

The pressure is assumed to linearly scale with the density, following the law $P = w\rho$, and fluids are characterized by their value of w:

1. matter has $w = 0$;
2. radiation has $w = 1/3$;
3. the vacuum has $w = -1$.

The stress-energy tensor is $T^{\mu}_{\nu} = \operatorname{diag}(\rho, -P, -P, -P)$ [KT94, eq. 3.4].[1]

The Friedmann equations are what we get if we substitute in the FRLW metric into the Einstein equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{2.3a}$$

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0 \tag{2.3b}$$

$$P = P(\rho)\,. \tag{2.3c}$$

Only two of these equations are independent; we can interpret the dynamical $\ddot{a}$ Friedmann equation as a consequence of the first two, which makes things easier since it is second order.

The density ρ which appears in the equations is the sum of all the ones which make up the universe, we must account for the w of all the fluids in the model we are considering.

Now we will discuss solutions to the Friedmann equations. Ones with $k = 0$ are *spatially flat models*, and for them the first Friedmann equation reads:

$$H^2 = \frac{8\pi G}{3}\rho\,, \tag{2.4}$$

so in this case the absolute value of a at a specific time does not matter, only ratios of values of a at different times have physical meaning. This can be seen from the fact that the equation has a conformal global symmetry: if we map $a \to Ca$ for some constant C the equation does not change. Also, the Friedmann equation is invariant under time translations.

Dust solution

If we have nonrelativistic matter, which is commonly called "dust", with $w = 0 \implies P = 0$, we get $\rho \propto a^{-3}$ and $a(t) \sim t^{2/3}$ (for time starting at a certain $t_s = 0$). So, the energy density scales like $\rho(t) \sim t^{-2}$.

The Hubble parameter, on the other hand, is given by $H(t) = 2/(3t)$. This model predicts the current age of the universe to be $t_0 = 2/(3H_0) \approx 9$ billion years, which is too short (regardless of the value we take for H_0, this does not work).

[1] Note that we must be careful with the signs of the stress-energy tensor: if we lower or raise both the indices the sign of the pressure changes. It is negative only when the tensor is $(1,1)$.

Radiation solution

For relativistic matter, we have $P = \rho/3$, so the density scales like $\rho \sim a^{-4}$, $a(t) \sim t^{1/2}$, $H = 1/(2t)$ and $\rho = \frac{3}{32\pi G t^2}$.

Let us assume that all the particles in the universe are in thermal equilibrium. Then, the density will scale like

$$\rho = \frac{\pi^2}{30} g_* T^4 \,, \tag{2.5}$$

where

$$g_* = \sum_b g_b + \frac{7}{8} \sum_f g_f \,, \tag{2.6}$$

where the two sums are over all the bosons and fermions respectively whose masses are smaller than the current temperature (so, which are still relativistic): it is the effective number of degrees of freedom. The objects being summed are the number of spin states of each boson or fermion. This relation comes from integrating thermal distributions in momentum space in the ultrarelativistic limit (so that the particles we are considering are "radiation"); the particulars of the different statistics makes it so that the contribution for fermions to the density is smaller than that of bosons by a factor 7/8.

So, the Hubble parameter scales like

$$\begin{aligned} H^2 &= \frac{8\pi G}{3} \rho \qquad (2.7)\\ &= \frac{8\pi}{3} \frac{1}{M_P^2} \frac{\pi^2}{30} g_* T^4 \qquad (2.8)\\ &= \frac{8}{90} \pi^3 g_* \frac{T^4}{M_P^2} \,. \qquad (2.9) \end{aligned}$$

Then, we can define a new reduced Planck mass M_P^*, such that $H = T^2/M_P^*$: it will need to be

$$M_P^* = \sqrt{\frac{90}{8\pi^3 g_*}} M_P \approx \frac{M_P}{1.66\sqrt{g_*}} \,. \tag{2.10}$$

Now, since we know $\rho \sim a^{-4}$ and $\rho \sim T^4$ we can write $T \sim 1/a$, keeping in mind however that the constant in front is not truly so for the whole evolution of the universe, instead it varies depending on which particles are decoupled at that point.

Vacuum

Let us consider the vacuum. In flat spacetime, its stress-energy tensor is given by

$$T_{\mu\nu} = \rho_{\text{vac}} \eta_{\mu\nu} \,, \tag{2.11}$$

so that $P_{\text{vac}} = -\rho_{\text{vac}}$.

This is equivalent to the introduction of a $-8\pi G \Lambda g_{\mu\nu}$ term to the left-hand side of the field equations: a "cosmological constant" with $\Lambda = \rho_{\text{vac}}$.

If we only had this fluid in our model universe we would see $a \sim e^{H_{\text{dS}} t}$, where

$$H_{\text{dS}} = \sqrt{\frac{8\pi G}{3} r_{\text{vac}}} \,. \tag{2.12}$$

Our spacetime will then be described by a metric like

$$\mathrm{d}s^2 = \mathrm{d}t^2 - e^{2H_{\text{dS}} t} \, \mathrm{d}x^2 \,, \tag{2.13}$$

which is called **de Sitter** spacetime. In this case we will have $\ddot{a} > 0$, there will be no initial singularity.

2.2 The ΛCDM model

The ingredients are:

1. Nonrelativistic matter: baryons, dark matter, and also neutrinos with masses $m_\nu \gtrsim -3\,\text{eV}$. Its density is denoted as ρ_M.
2. Relativistic matter: photons and neutrinos with $m_\nu \lesssim 10^{-4}\,\text{eV}$. Its density is denoted as ρ_{rad}.
3. Dark energy: is it vacuum energy? Its density is denoted as ρ_Λ.

The Hubble parameter is given by the first Friedmann equation, by accounting for all of these plus curvature:

$$H^2 = \frac{8\pi G}{3} \left(\rho_M + \rho_{\text{rad}} + \rho_\Lambda + \rho_{\text{curv}} \right) , \tag{2.14}$$

where

$$\rho_{\text{curv}} = -\frac{k}{a^2} \frac{3}{8\pi G} \,. \tag{2.15}$$

The critical density is given by

$$\rho_c = \frac{3}{8\pi G} H_0^2 \approx 5 \times 10^{-6}\,\text{GeV}/\text{cm}^3 \approx \frac{5 m_P}{\text{m}^3} \,, \tag{2.16}$$

where the numerical value is obtained by taking $h = 0.7$.

By construction, if $\rho = \rho_c$ for our universe then it is spatially flat.

We can take the ratios of $\rho_i / \rho_c = \Omega_i$, this quantifies how much of the density of the universe is in the form of a certain kind of fluid. If we also include Ω_{curv}, we get $\sum_i \Omega_i = 1$ always.

Let us give some values for them: for photons, we consider the CMB photons, which closely follow a thermal distribution at $T_0 \approx 2.7\,\text{K}$. By looking at these, we find

$$\Omega_\gamma \approx 2.5 \times 10^{-5} h^2 \approx 5 \times 10^{-5} \,. \tag{2.17}$$

Why do we only look at CMB photons? stellar fusion has an efficiency of the order of 1 %, so I'd imagine a considerable (compared to 5×10^{-5}) fraction of its mass is radiated away as high energy photons (compared to the CMB)...

Even if we had neutrinos with small masses such that they are radiation today, their Ω would be smaller than the photons', so negligible overall.

For the matter and dark energy contributions we get

$$\Omega_M = \Omega_B + \Omega_{DM} \approx 0.05 + 0.27 \approx 0.32 \qquad \text{and} \qquad \Omega_\Lambda \approx 0.68\,. \tag{2.18}$$

How do these depend on time? We have

1. $\rho_{\text{rad}} \sim a^{-4}$;
2. $\rho_M \sim a^{-3}$;
3. $\rho_{\text{curv}} \sim a^{-2}$;
4. $\rho_\Lambda \sim a^0 = \text{const.}$

Then, the Friedmann equation in the ΛCDM model looks like

$$H^2 = \frac{8\pi G}{3}\rho_c\left[\Omega_{\text{rad}}\left(\frac{a_0}{a}\right)^4 + \Omega_M\left(\frac{a_0}{a}\right)^3 + \Omega_{\text{curv}}\left(\frac{a_0}{a}\right)^2 + \Omega_\Lambda\right] \tag{2.19}$$

$$= H_0^2\left[\Omega_{\text{rad}}(1+z)^4 + \Omega_M(1+z)^3 + \Omega_{\text{curv}}(1+z)^2 + \Omega_\Lambda\right], \tag{2.20}$$

where we used $1 + z = a_0/a$, where z is the redshift, defined by

$$1 + z = \frac{\lambda_{\text{absorption}}}{\lambda_{\text{emission}}}\,. \tag{2.21}$$

For close objects (with small z) we recover Hubble's law: $z = H_0\tau$.

2.2.1 On the vacuum energy

In particle physics, we use the vacuum energy as our reference point from which to measure other energies; energies of particles are excitations from that vacuum.

In GR, instead, **the vacuum energy gravitates**: there is no such thing as defining energy up to a constant, it all contributes to the gravitational field. The vacuum energy is present everywhere always and it does not cluster, so it is a good candidate for dark energy and ρ_{vac}.

What is the vacuum energy prediction by the Standard Model, after symmetry breaking?

What remains when we take the VEV are the scalar terms, since the vector fields are zero in the vacuum. Specifically, we are left with

$$\mu^2\phi^2 + \lambda\phi^4\,. \tag{2.22}$$

The order of magnitude of the VEV of the Higgs is around 100 GeV.

The dimensions of $\rho = E/L^3$ are those of an energy to the fourth power. So, our estimate could be of $\rho \sim (100\,\text{GeV})^4$.

Alternatively, we could consider the characteristic energy scales of the various interactions: the strong interaction is around $1\,\text{GeV}$, the weak interaction is around $100\,\text{GeV}$, the gravitational interaction is around $M_P \sim 10^{19}\,\text{GeV}$. All of these have to be taken to the fourth power to recover the dimensions of an energy density.

The estimate from cosmology is $10^{-46}\,\text{GeV}^4 \approx (10^{11}\,\text{GeV})^4$: the one from particle physics is 45, 50 or even 120 orders of magnitude larger, depending on whether we are looking at the strong, electroweak or gravitational interaction.

2.3 Thermodynamics of the early universe

Wednesday 2020-5-6, compiled 2020-07-02

We start off by discussing particles in thermal equilibrium, but the interesting thing is the transition between this equilibrium and non-equilibrium.

We use the usual approximation of a dilute gas with g degrees of freedom, we take a Boltzmann distribution function in momentum space $f(p)$, from which we can compute n, ρ and P:

$$n = \frac{g}{(2\pi)^3} \int f(\vec{p})\,\mathrm{d}^3 p \tag{2.23}$$

$$\rho = \frac{g}{(2\pi)^3} \int E(\vec{p}) f(\vec{p})\,\mathrm{d}^3 p \tag{2.24}$$

$$P = \frac{g}{(2\pi)^3} \int \frac{|p|^2}{3E} f(\vec{p})\,\mathrm{d}^3 p\,. \tag{2.25}$$

In thermal equilibrium the distribution function in momentum space is

$$f(\vec{p}) = \frac{1}{\exp\left(\frac{E-\mu}{T}\right) \pm 1}\,, \tag{2.26}$$

where we have $+$ for fermions and $-$ for bosons. The energy as a function of momentum is given by

$$E(\vec{p}) = \sqrt{|\vec{p}|^2 + m^2}\,. \tag{2.27}$$

We can make some useful approximations in the relativistic limit $T \gg m$ and in the nonrelativistic limit $T \ll m$: in the **relativistic** limit we find

$$\rho = \begin{cases} \dfrac{\pi^2}{30} g T^4 & \text{bosons} \\ \dfrac{7}{8}\dfrac{\pi^2}{30} g T^4 & \text{fermions} \end{cases} \qquad \text{and} \qquad n = \begin{cases} \dfrac{\zeta(3) g T^3}{\pi^2} & \text{bosons} \\ \dfrac{3}{4}\dfrac{\zeta(3) g T^3}{\pi^2} & \text{fermions,} \end{cases} \tag{2.28}$$

where ζ is the Riemann zeta function, so that $\zeta(3) = \sum_{n\in\mathbb{N}} n^{-3} \approx 1.2$. Since we are in the ultrarelativistic approximation, we will have $P = \rho/3$.

In the **nonrelativistic** limit, instead, we get

$$n = g\left(\frac{mT}{2\pi}\right)^{3/2} e^{-(m-\mu)/T}, \tag{2.29}$$

while $\rho = mn$ and $P = nT \ll \rho$.

Add reference to Fundamentals notes, chapter 3, when it will be done.

A very useful parameter to define in this context is g_*: it is the *effective* number of degrees of freedom for different particle species, computed as

$$g_* = \sum_{i \in \text{bosons}} g_i \left(\frac{T_i}{T}\right)^4 + \frac{7}{8} \sum_{i \in \text{fermions}} g_i \left(\frac{T_i}{T}\right)^4, \tag{2.30}$$

since we need to weigh the contributions to the degrees of freedom according to whether the particles are in equilibrium or not, and when the particle decoupled. With this definition, in the relativistic limit we can simply write

$$\rho_R = \sum_i \rho_i = \frac{\pi^2}{30} g_* T^4 \qquad \text{and} \qquad P_R = \frac{\pi^2}{90} g_* T^4 . \tag{2.31}$$

Effective number of degrees of freedom examples

Let us compute g_* in a couple of examples. First, let us consider $T \ll \text{MeV}$ — but at a time in which the photons are still coupled, that is, *before* the emission of the CMB.

So, in the standard model the only decoupled particles are photons and the three neutrino species. The temperature of the neutrinos, as we will see later, is given in this time period by

$$T_\nu = \sqrt[3]{\frac{4}{11}} T_\gamma . \tag{2.32}$$

So, since: photons have two physical polarizations and so do (left-handed) neutrinos, there are three neutrino species and neutrinos are fermions we get

$$g_*(T \ll \text{MeV}) = 2 + \frac{7}{8} \times 3 \times 2 \times \left(\frac{4}{11}\right)^{4/3} \approx 3.37 . \tag{2.33}$$

If we consider a temperature around $1\,\text{MeV} < T < 100\,\text{MeV}$, we also will have to account for electrons and positrons. Now, the temperature of all these species can be taken to be equal: so, we get

$$g_*(1\,\text{MeV} < T < 100\,\text{MeV}) \approx 2 + \frac{7}{8}(3 \times 2 + 1 \times 4) \approx 10.75 . \tag{2.34}$$

For $T > 200\,\text{GeV}$, all the SM particles will be relativistic, so we will find $g_* \approx 106.75$. For a nice figure summarizing the evolution of g_* as the temperature decreases see figure 1 in a recent paper by Lars Husdal [Hus16]. The take-away is that, going from the TeV to the keV, if we plot g_* on a log scale, it has sharp drop-offs around 200 MeV (a QCD transition) and around 500 keV (electrons and positrons), while it is mostly flat elsewhere.

The radiation domination epoch

This was the period in the early universe in which ρ_R was larger than ρ_M; we will consider the early phase in which this difference was large (in order to approximate $\rho_M \approx 0$), but the inequality stayed on the radiation side for about a millennium.

In the era of radiation domination the scale factor varies like $a(t) \propto \sqrt{t}$ and the Hubble parameter scales like

$$H \approx 1.66\sqrt{g_*}\frac{T^2}{M_{\text{Pl}}}\,. \tag{2.35}$$

So, we get a relation between time and temperature: approximately it is

$$t \approx \left(\frac{T}{\text{MeV}}\right)^{-2}\text{s}\,, \tag{2.36}$$

and more precisely it is

$$tT^2 = \sqrt{\frac{90}{32\pi^2}\frac{1}{G_N g_*}}\,. \tag{2.37}$$

2.3.1 Thermal equilibrium

If the universe were not expanding, after a certain amount of time we would reach thermal equilibrium between the particle species. But this is not the case: the expansion rate of the universe is H.

So, we take the decay rate of a certain particle species i, which is denoted by Γ_i. A particle is said to reach equilibrium if its decay rate is $\Gamma_i > H$. This is a rule of thumb, an order of magnitude estimate: if $\Gamma_i \gg H$ then the particle has time to decay many times within the age of the universe, while if $\Gamma_i \ll H$ then its decay curve has barely started since the beginning of the universe. We are doing cosmology so the precise details of what happens when $\Gamma_i \approx H$ are not very important, that is the region in which a macroscopic fraction of our particle species has has time to decay but not all of it.

Note that this is a time- (or temperature-) dependent condition: a particle may be in equilibrium at a specific time and go out of equilibrium later.

Neutrino decoupling: the *freeze-out*

We want to discuss processes involving the neutrino, which are things like $e^+e^- \to \nu\overline{\nu}$ or $e\nu \to e\nu$.

This are processes mediated by the weak interaction, so it is relevant to ask whether the temperature is larger or smaller than the mediator's mass. Let us first suppose that $T > M_W$, so roughly $T > 100\,\text{GeV}$. This means that the Breit-Wigner term in the cross section can be approximated as

$$\frac{1}{T^2 - M_W^2} \sim \frac{1}{T^2}\,. \tag{2.38}$$

The rate of these processes is given by

$$\Gamma_\nu = \sigma n v\,, \tag{2.39}$$

where the physics is really given by the cross section σ: the other factors are always roughly the same, $n \sim T^3$ and $v \sim 1$. The cross section depends on the coupling and the temperature: the coupling constant for weak processes is $\alpha_w = g_w^2/4\pi$, and the cross section is given by $\sigma \sim \alpha_w^2/T^2$,[2] so the decay rate is roughly

$$\Gamma \sim \frac{\alpha_w^2}{T^2} \times T^3 \times 1 = \alpha_w^2 T\,. \tag{2.40}$$

On the other hand, the Hubble rate is roughly proportional to

$$H \sim \sqrt{g_*}\frac{T^2}{M_P}\,, \tag{2.41}$$

so if we impose $\Gamma_\nu > H$ we get

$$T < \frac{\alpha_w^2 M_P}{\sqrt{g_*}}\,. \tag{2.42}$$

So, our result is that for

$$M_W < T < \frac{\alpha_w^2 M_P}{\sqrt{g_*}} \tag{2.43}$$

neutrinos are in thermal equilibrium. In terms of actual energies, this means $10^2\,\text{GeV} < T < 10^{16}\,\text{GeV}$. The higher end of this is almost at the Planck mass, so we will not worry about it.

Now let us consider the later time at which $T < M_W$. In this case, in the Breit-Wigner the mass of the weak interaction boson will dominate, and we will have

$$\sigma \sim \mathcal{O}(10^{-2})\frac{T^2}{M_W^4} \approx 10^{-10}\frac{T^2}{\text{GeV}^4}\,. \tag{2.44}$$

What is the factor which is of the order 0.01?

Now, since we have $\sigma \sim T^2$ the rate will scale like $\Gamma \sim T^5$. Specifically,

$$\Gamma \sim 10^{-10}\frac{T^5}{\text{GeV}^4}\,. \tag{2.45}$$

We must compare to the expansion rate of the universe, which as before scales according to 2.41. Then, we find

$$10^{-10}\frac{T^5}{\text{GeV}^4} > \sqrt{g_*}\frac{T^2}{10^{19}\,\text{GeV}} \tag{2.46}$$

[2] The factor which appears in the cross section is precisely the Breit-Wigner one.

$$T > \sqrt[3]{\frac{\sqrt{g_*}10^{10}}{10^{19}}}\mathrm{GeV} = \sqrt[6]{g_*}\mathrm{MeV}\,. \tag{2.47}$$

Now, $g_*^{1/6} \sim 1$, so we find that for $1\,\mathrm{MeV} < T < M_W$ neutrinos are coupled.

The end result is $\Gamma_\nu > H$ as long as $T > 1\,\mathrm{MeV}$, roughly. So, for larger temperatures the neutrinos are coupled, for smaller temperatures they are decoupled. Below $T_D \approx 1\,\mathrm{MeV}$ they completely decouple. At this temperature the universe becomes transparent for them, they no longer significantly interact with matter. The decoupling temperature is defined by the relation given

$$\Gamma_\nu(T_D^\nu) = H(T_D^\nu)\,. \tag{2.48}$$

Now, we can define the current temperature and number density of the leftover neutrinos: they are respectively T_0^ν and n_ν. When the temperature was $T > 1\,\mathrm{MeV}$ the neutrinos and photons were in thermal equilibrium: we had $T_\nu \sim T_{\text{plasma}} \sim T_\gamma$. In this plasma there were neutrinos, photons, electrons and positrons.

At $T < 500\,\mathrm{keV}$ electron-positron annihilation was still occurring, but pair creation slowed: there was not enough energy for a process like $\nu\overline{\nu} \to e^+e^-$. The pair annihilation poured energy into photons only, since neutrinos were decoupled by that point.

So, the "relic neutrinos" are colder than the relic photons. Moreover, their number density is smaller.

The entropy density for relativistic matter with $P = \rho/3$ is given[3] by

$$s = \frac{\rho + P}{T} = \frac{4}{3}\frac{\rho}{T}\,. \tag{2.50}$$

So, for the i-th particle species we can write

$$s_i = \frac{4}{3}\frac{\rho_i}{T} = g_i\frac{2\pi^2}{45}T^3 \tag{2.51}$$

for bosons, and the same result times $7/8$ for fermions.

Now, let us define $T_>$ as the temperature of the photons, T_γ, before the disappearance of the e^+e^- pairs. This temperature will be in the range $500\,\mathrm{keV} < T_> < 1\,\mathrm{MeV}$.

Then, we can define a corresponding entropy $s_>$: it will be given by

$$s_> = \frac{4}{3}\frac{\rho_>}{T_>} = \frac{2\pi^2}{45}g_*(T_>)T_>^3\,, \tag{2.52}$$

where the effective number of degrees of freedom is given by:

$$g_*(T_>) = 2 + \frac{7}{8}(2+2) = \frac{11}{2}\,, \tag{2.53}$$

[3] To see this, start from the thermodynamic relation $T\,\mathrm{d}S = \mathrm{d}E + P\,\mathrm{d}V$, divide through by $T\,\mathrm{d}V$ to get

$$\frac{\mathrm{d}S}{\mathrm{d}V} = \frac{1}{T}\frac{\mathrm{d}E}{\mathrm{d}V} + \frac{P}{T} = \frac{\rho + P}{T}\,. \tag{2.49}$$

since we need to account only for photons, electrons and positrons — neutrinos have already decoupled.

Also, let us define another temperature, $T_<$, such that $T_< < 500\,\text{keV}$. The corresponding $g_*(T_<)$ will be of 2, since electrons and positrons will have decoupled. Then, the corresponding entropy density will read:

$$s_< = \frac{4}{3}\frac{\rho_<}{T_<} = \frac{4}{3}\frac{\pi^2}{30} g_* T_<^3 = \frac{2\pi^2}{45} g_*(T_<) T_<^3 \,. \tag{2.54}$$

Now, we impose $s_< = s_>$: in the phase transition the temperature can be discontinuous by the entropy is conserved. This yields

$$\frac{2\pi^2}{45} g_*(T_<) T_<^3 = \frac{2\pi^2}{45} g_*(T_>) T_>^3 \tag{2.55}$$

$$g_*(T_<) T_<^3 = g_*(T_>) T_>^3 \tag{2.56}$$

$$2T_<^3 = \frac{11}{2} T_>^3 \tag{2.57}$$

$$T_< = \sqrt[3]{\frac{11}{4} T^3} \,. \tag{2.58}$$

The next step is to state that $T_>/T_< = T_\gamma/T_\nu$. This is because photons can still thermalize at $T_<$, while neutrinos are "stuck" at the temperature $T_>$.

So, the temperature of relic neutrinos today can be calculated from that of the CMB to be around $T_\nu^0 = \sqrt[3]{4/11} T_\gamma^0 \approx 1.96\,\text{K}$.

We also know that the number density of relic photons today is given by

$$n_\gamma^0 = \frac{\zeta(3)}{\pi^2} 2T_\gamma^3 \approx 422\,\text{cm}^{-3} \,, \tag{2.59}$$

so we can calculate that of neutrinos using what we know from (2.28):

$$\frac{n_\nu^0}{n_\gamma^0} = \frac{3}{4}\left(\frac{T_\nu}{T_\gamma}\right)^3 = \frac{3}{11} \,. \tag{2.60}$$

So, $n_\nu^0 \approx 100\,\text{cm}^{-3}$.

Another example of a particle decoupling is given by the graviton: the rate of interactions which are only gravitational is given by

$$\Gamma = n\sigma v \approx G_N^2 T^5 \sim \frac{T^5}{M_P^4} \,, \tag{2.61}$$

so if we impose $\Gamma < H$ we get the expected result: the decoupling temperature is just the Planck mass,

$$\frac{T^5}{M_P^4} < \frac{T^2}{M_P} \implies T^3 < M_P^3 \implies T < M_P \,. \tag{2.62}$$

The g_* at that early time was around 100, so we find that the temperature of relic gravitational radiation should be around $T_{\text{grav}} \approx 0.8\,\text{K}$ and its number density should be $n_{\text{grav}} \approx 15\,\text{cm}^{-3}$

Next week we will consider nucleosynthesis.

2.4 Big Bang Nucleosynthesis

Tuesday 2020-5-12, compiled 2020-07-02

Now we move to the infrared regime, from 1 MeV to a few tens of keV. Neutrons and protons are the new protagonists: the quarks are in the infrared "slavery". This is the period in which we see the first combination of neutrons and protons into light nuclei, such as ^{4}He, D = ^{2}H, ^{3}H, ^{7}Li.

Neutron freeze-out

Neutrons and protons can turn into each other through weak processes, such as

$$\mathrm{n} + \nu_e \leftrightarrow \mathrm{p} + e^- \tag{2.63}$$

$$\mathrm{n} + e^+ \leftrightarrow \mathrm{p} + \overline{\nu}_e \tag{2.64}$$

$$\mathrm{n} \leftrightarrow \mathrm{p} + e^- + \nu_e \,. \tag{2.65}$$

We want to consider the ratio of their number densities, denoted as n/p. At equilibrium, and as long as the baryons are nonrelativistic, this is given by

$$\frac{n}{p} \approx \exp\left(-\frac{\Delta m}{T}\right), \tag{2.66}$$

where $\Delta m = m_n - m_p \approx 1.3\,\mathrm{MeV}$: for $T \gg \Delta m$ this is approximately $\exp(0) = 1$.

The proton, as far as we can tell, is stable; on the other hand the neutron is unstable: it can beta-decay into a proton.

At 1 MeV, which is around the moment of the freeze-out, the ratio n/p is around $1/6$.

For temperatures larger than an MeV nucleosynthesis cannot start: deuterium nuclei can form by

$$\mathrm{n} + \mathrm{p} \to \mathrm{D} + \gamma\,, \tag{2.67}$$

but they are readily photodissociated, since there are plenty of photons whose energy is $E_\gamma > 2.2\,\mathrm{MeV}$, the binding energy of deuterium. At this stage, the ratio of the number of photons to that of baryons is of the order of $n_\gamma / n_b \approx 10^{10}$, way too high still.

Now, what is the temperature at which neutrons decouple, T_D^{neutron}? The decay rate of neutrons is given by

$$\Gamma_n = C_n G_F^2 T^5\,, \tag{2.68}$$

where C_n is some constant which is not much different from 1. As always, in order to find out at which temperature they decouple, T_D^{neutron}, we impose $H(T_D) = \Gamma(T_D)$.

The g_* at this temperature accounts for photons, electrons, positrons and neutrinos. So, we find

$$T_D^n = \frac{1}{\left(C_n M_P^* G_F\right)^{1/3}}\,, \tag{2.69}$$

where M_P^* is the corrected Planck mass (2.10). The order-1 constant C_n can be determined experimentally from the lifetimes of neutrons: what we find is $C_n \approx 1.2$. If the number of neutrino species is $N_\nu = 3$ then we get $T_D^n \approx 1.4\,\mathrm{MeV}$, which is very close to the mass difference between protons and neutrons!

Why is this relevant? Well, let us consider the extreme behaviors:

- if $\Delta m \gg T_D$, then freezeout happens at a near-zero value of $n/p = \exp(-\Delta m/T)$;
- while if $\Delta m \ll T_D$, freezeout happens when we still have $n/p \approx 1$: what happens then is that all the neutrons and protons turn into ^{4}He and no hydrogen is left in the primordial plasma.

An exact computation yields $T_D^n \approx 0.7\,\mathrm{MeV}$, which corresponds to around $t = 1.1\,\mathrm{s}$ after the beginning.

What is the temperature at which the deuterium can stay bound? It is the temperature after which there is more formation of deuterium than photodissociation. The binding energy of deuterium is around $E_D \approx 2.2\,\mathrm{MeV}$. To get stable deuterium we must wait until

$$A = \frac{n_\gamma}{n_B} e^{-\frac{2.2\,\mathrm{MeV}}{T}} < 1\,. \tag{2.70}$$

When $A < 1$, the rate of deuterium production surpasses that of its dissociation.

We actually have to wait quite a long time for this to occur, since $n_\gamma \gg n_B$. This is since baryons are quite massive. The ratio comes out to be around one billionth: $n_\gamma/n_B \sim 10^9$.

So, we need a strong drop in temperature: we find that the necessary temperature is around $0.1\,\mathrm{MeV}$, corresponding to $t \approx 10^2\,\mathrm{s}$.

So, nucleosynthesis starts at $T < 1\,\mathrm{MeV}$, while the weak-interaction processes freeze out at $T \approx 1\,\mathrm{MeV}$. The neutron-to-proton ratio n/p is around $1/6$ at freezeout, and we will see that in the time between this and the start of nucleosynthesis it reaches $1/7$ because of neutron decay.

The lifetime of a neutron is measured experimentally to be

$$\tau_n = (10.5 \pm 0.2)\,\mathrm{min}\,. \tag{2.71}$$

After we go below the temperature $T \approx 0.1\,\mathrm{MeV}$ more processes can occur, like

$$\mathrm{n} + \mathrm{p} \leftrightarrow \mathrm{D} + \gamma \tag{2.72}$$
$$\mathrm{D} + \mathrm{p} \leftrightarrow {}^3\mathrm{He} + \gamma \tag{2.73}$$
$$\mathrm{D} + \mathrm{n} \leftrightarrow {}^3\mathrm{H} + \gamma \tag{2.74}$$
$${}^3\mathrm{He} + \mathrm{n} \leftrightarrow {}^4\mathrm{He} \tag{2.75}$$
$${}^3\mathrm{H} + \mathrm{p} \leftrightarrow {}^4\mathrm{He}\,. \tag{2.76}$$

Helium-4 is the most stable of these isotopes, and most of the free neutrons end up in it. The production of heavier nuclei is quite suppressed at this temperature, and we do not consider them.

We can now compute the number fraction of helium nuclei:

$$X_4 = \frac{N_{^4\mathrm{He}}}{N_\mathrm{H}} = \frac{\frac{1}{2}n}{p-n} = \frac{1}{2}\frac{n/p}{1-n/p}\,. \tag{2.77}$$

since the number of hydrogen nuclei is the same as the number of unpaired protons. The mass fraction of helium, on the other hand, is

$$Y_4 = \frac{M_{\mathrm{He}}}{M_{\mathrm{H}} + M_{^4\mathrm{He}}} = \frac{4N_{\mathrm{He}}}{N_{\mathrm{H}} + 4N_{^4\mathrm{He}}} = \frac{4X_4}{1+4X_4} = \frac{2(n/p)}{1+n/p}\,. \tag{2.78}$$

For $n/p \approx 1/7$ this is around $Y_4 \approx 0.25$, in agreement with observations.

To get a more precise estimate we should properly consider all the formation channels for ^{4}He, beyond the ones we have written there are others such as ones in which the product of a reaction is not only a helium nucleus but we also get another proton or neutron.

The chain keeps going beyond Helium, skipping the mass numbers $A = 5$ and $A = 8$. We can get ^{5}Li and ^{6}Li.

To summarize, there have been **three main parameters** in this calculation:

1. the ratio n_B/n_γ determines the beginning of the nucleosynthesis;
2. the lifetime of a neutron, τ_n, enters the determination of the rate of weak processes, which is connected to the weak coupling G_F

   ...and determines n/p at the beginning of nucleosynthesis, right?

3. the effective number of degrees of freedom g_* at $m < 1\,\mathrm{MeV}$ is important since $H \propto \sqrt{g_*}$.

Let us investigate the effect of the variation of these parameters.

If we increase $n_B/n_\gamma = \eta_B$ nucleosynthesis starts before, so there is less time for the neutrons to β-decay, so the ratio n/p is higher, so Y_4 will be higher.

The number density of baryons today is given by

$$n_B^0 = \frac{\rho_B^0}{m_B} = \frac{\Omega_B \rho_c}{m_B} \approx 1.13 \times 10^{-5} \Omega_B h_0^2 \mathrm{cm}^{-3}\,, \tag{2.79}$$

while the number density of photons today is

$$n_\gamma^0 = \frac{2\zeta(3)}{\pi^2} T_\gamma^3 \approx 400 \left(\frac{T_0}{2.7\,\mathrm{K}} \right)^3 \mathrm{cm}^{-3}\,. \tag{2.80}$$

The value of η today, which equals that at nucleosynthesis[4] is then given by

$$\eta \approx 2.81 \times 10^{-8} \Omega_B h_0^2 \left(\frac{2.7}{T_0} \right)^3\,. \tag{2.81}$$

Another thing which can affect helium abundance after nucleosynthesis is a variation of $\sqrt{g_*}$, that is, of the number of relativistic particles' degrees of freedom at the nucleosynthesis temperature. For example, this would be an issue if we were to add a new neutrino species.

[4] Since at that point photons were not coupled to anything which could annihilate or create them, while baryons were decoupled altogether. This line of reasoning holds as long as no later processes produce entropy: that would modify n_γ. In some models this is not the case.

In weighting these new degrees of freedom, we must take into account that the temperature corresponding to them may be lower than that of the plasma.

As $\sqrt{g_*}$ increases, Y_4 increases: in more detail the causal chain of increases goes N_ν, g_*, B, $T_{\text{freezeout}}$, n/p, Y_4.

What is meant by B is n_B, right?

In our descriptions of nucleosynthesis we do not only have ^{4}He but also other light isotopes such as deuterium, ^{3}He, ^{7}Li. For the currently estimated value of $\eta \approx 6\,\text{e-}10$ we predict

$$\frac{\text{Li}}{\text{H}} \approx 10^{-10} \qquad \text{and} \qquad \frac{\text{D}}{\text{H}} \approx 2 \div 3 \times 10^{-5}\,, \tag{2.82}$$

but we can extend our predictions for the various nuclides to different values for η. The value of Ω_B corresponding to this estimate is $\Omega_B \approx 0.044 \pm 0.005$. The estimates we get from the CMB agree with these results.

Now, in the SM we have three neutrino species, so we can ask: how many *effective* neutrino species could we accommodate in order to still reproduce data? Well, we have

$$\frac{\Delta g_*}{g_*} = \frac{(7/8) 2 \Delta N^\nu_{\text{eff}}}{g_*^{SM}} = \frac{7}{43} \Delta N^\nu_{\text{eff}}\,, \tag{2.83}$$

which we can express in terms of the (variation of the) ratio of neutron to proton number densities:

$$\frac{\Delta(n/p)}{n/p} = \frac{7}{258} \frac{\Delta m}{T_n^D} \Delta N^\nu_{\text{eff}}\,, \tag{2.84}$$

and our error bounds on this quantity give $\Delta(n/p)/(n/p) \lesssim 0.025$, so we get

$$\left| \Delta N^\nu_{\text{eff}} \right| \lesssim 0.5\,. \tag{2.85}$$

So, our DM candidate must contribute as half a neutrino species: when we propose a DM candidate, we must always ask whether it spoils nucleosynthesis.

Again, this must be weighted by their temperature which could be different from that on the plasma depending on whether and when these species are at equilibrium. If they decoupled earlier than the neutrinos they could be colder than them, and thus give a smaller contribution.

There are models in which we introduce a new particle which must decay into some other particle plus a photon. This is dangerous: how energetic are the photons which are produces? We must ensure that they do not destroy deuterium.

Comments on nucleosynthesis

The nucleosynthesis epoch, between $\sim 1\,\text{s}$ and $\sim 300\,\text{s}$, is the earliest one which has been tested observationally.

We can get estimates for relative abundances of nuclides in the early universe by looking at regions which are both far away and with low star formation rates, in which the elements' abundances are still well-preserved.

Deuterium is special: because of its low binding energy it is *not produced* in stellar nucleosynthesis, it is disassociated instead. So, any amount of it we measure gives us a *lower bound* on its primordial abundance.

As expected from BB Nucleosynthesis theory, we have variations of several orders of magnitude between the relative abundances of ^{4}He, deuterium and ^{7}Li. BBN and CMB data fit the theory for $\eta_B \approx 6 \times 10^{-10}$: this theory is very successful!

We have a problem with ^{7}Li: the prediction for its abundance for this value of η_B is [Cyb+16]:

$$\left.\frac{^7\text{Li}}{H}\right|_{\text{th}} = (4.87 \pm 0.67) \times 10^{-10}\,, \tag{2.86}$$

while the experimental observation is [CV17]:

$$\left.\frac{^7\text{Li}}{H}\right|_{\text{exp}} = (1.58 \pm 0.31) \times 10^{-10}\,. \tag{2.87}$$

Besides this problem, the theory of BBN has generally been a very succesful marriage of the particle physics and cosmology standard models.

The impact of BBN on the standard model of particle physics are the bound

$$\left|\Delta N^{\nu}_{\text{eff}}\right| < 0.5 \tag{2.88}$$

and more bounds on radiative decays of heavy particles X, in the form

$$X \to \gamma + Y\,, \tag{2.89}$$

where Y is some lighter particle species, since this would produce a photon with $E_\gamma \sim m_X/2$, which would spoil nucleosynthesis in two ways:

1. it would destruct light nuclides (particularly deuterium) with its energetic photon;
2. it would break up ^{4}He into lighter nuclides.

We get bounds on these kinds of processes, specifically on the lifetime τ_X and on the abundance of X, measured as $m_X n_X / n_\gamma$. We are safe if $\tau_X < 10^{-3} \div 10^{-4}$ s. If we assume a very long lifetime, like $\tau_X > 10^8$ s, we must have a very small X abundance, like $m_X n_X / n_\gamma < 10^{-12}$.

I think this is what he is saying, not sure though.

2.5 Dark Matter

Let us start our discussion of dark matter with the **baryon budget**: how much do baryons contribute to the total energy density of the universe? Wednesday 2020-5-13, compiled 2020-07-02

In the *high-redshift universe* we have:

1. at nucleosynthesis ($z \sim 10^9$) the baryonic matter fraction was $\Omega_B^{BBN} = 0.04 \pm 0.02$;
2. at CMB emission ($z \sim 10^3$) the baryonic matter fraction was $\Omega_B^{CMB} \approx \Omega_B^{BBN} \approx 5\,\%$;
3. at $z \sim 3$ we have observations of the Lyman α ("frest"?) of the intergalactic medium: this again yields $\Omega_B \approx 5\,\%$.

what is written there?

In the recent universe ($z \sim 0$) we have observations of galaxy clusters, in which we measure a matter fraction of $\Omega_M \approx 0.26$, of which about 15 % is made of baryons: so, we get an estimate of around $\Omega_B \approx 4\,\%$.

So, the fraction of luminous baryonic matter is $\Omega_B^{\text{lumin}} \approx 1\,\%$, with a corresponding density of $\rho_{\text{lumin}} \approx 9 \times 10^{-29}\,\text{kgm}^{-3}$.

Not really clear how we estimate this based on what is mentioned above.

On the other hand, our estimate for the total baryonic mass fraction is $\Omega_B^{\text{tot}} \approx 5\,\%$, with a corresponding density of $\rho_b \approx 4 \times 10^{-28}\,\text{kgm}^{-3}$.

We know that the total value of $\Omega_{\text{tot}} \sim 1$.

The density fraction due to dark energy is around 70 %, while $\Omega_{\text{matter}} \sim 30\,\%$: this can be deduced from the gravitational potential energy, which is inferred from galactic rotation curves, and the mass distribution in clusters. The corresponding matter density is around $\rho_M \approx 2 \times 10^{-27}\,\text{kgm}^{-3}$.

So, most (something like 80 %) of the matter is not made of baryons. We are then looking for non-baryonic matter with $\Omega \sim 25\,\%$. More precisely, the latest results give

$$\Omega_{DM} \approx 26.4\,\% = 84.4\,\%\Omega_M \,. \tag{2.90}$$

Accounting for our uncertainty in H_0, we can write $\Omega_B = 0.022h^{-2} \approx 4 \div 5\,\%$.

How sure are we that dark matter is indeed the best way to describe galactic rotation curves and such?

It makes other predictions beyond them: there were interesting studies about how the lensing of light is affected by Dark Matter. All the evidence is based on our knowledge of gravitational interaction: Are we sure that Newton's law of gravitation is right on such large scales? There are MOND theories, in which it does not. However, the hints about the existence of non-baryonic DM come from very different directions.[5] We then take the standard view: Dark Matter exists.

Could the Higgs boson be the source of Dark Matter? No. It decays too fast. We have similar problems for the Z boson. So, in the end we must have some other neutral particle.

So, neutrinos? Maybe. They are light, of course, but there are many.

We get

$$\Omega_\nu = \frac{\rho_\nu^0}{\rho_{\text{crit}}} \,, \tag{2.91}$$

[5] And I'd be remiss not to cite https://xkcd.com/1758/.

where we used the known neutrino number density today $n_\nu^0 \approx 339.5\,\mathrm{cm}^{-3}$. So, this comes out to be

$$\Omega_\nu = \frac{\sum_i m_{\nu_i}}{h^2 93.14\,\mathrm{eV}}\,. \tag{2.92}$$

Now, $h^2 \sim 1/2$, so this is very roughly

$$\Omega_\nu \sim \frac{\sum m_\nu}{50\,\mathrm{eV}}\,. \tag{2.93}$$

Recall that "electron neutrino" is a current eigenstate, not a mass eigenstate. The bound we have is of the order $m_{\nu_e} \lesssim 1\,\mathrm{eV}$.

We also have bounds for the square of the mass differences. In the end, this means that the term $\sum m_\nu$ can be at most a few eV.

Even if all the three neutrinos' masses were at the very top of their current experimental bounds at $m_\nu \sim 2\,\mathrm{eV}$ each we would have $\Omega_\nu < 0.1$, which is too small to fit observations of $\Omega_{DM} \sim 0.25$.

In spite of the fact that they are very numerous, neutrinos have too small a mass.

2.5.1 Hot & cold DM

Let us call a generic DM particle X. It will decouple at a certain temperature T_X^D. Then, **cold** DM is such that $M_X \gg T_X^D$, while for **hot** DM we have $M_X \ll T_X^D$. So, cold DM is *nonrelativistic* when it decouples, while hot DM is *relativistic* when it does.

We can also have **warm** DM, in the case where $M_X \sim T_D$.

Neutrinos would be an example of hot dark matter, since they decouple around 1 eV while their mass is smaller than that.

Whether DM is hot or cold[6] is relevant for the formation of structures: relativistic matter has a hard time clumping at small scales.

For HDM, in the period $T_D > T > M_X$ the DM is free-streaming, so it quickly fills in under-dense regions, and the density perturbations are washed out.

The warm case is at around $M_X \gtrsim 1\,\mathrm{eV}$, while HDM has $M_X < 1\,\mathrm{eV}$.

There are simulations used in order to compare these models: what is seen is that for HDM the first structures to form are the very largest — superclusters of galaxies —, which then fragment into smaller pieces. This type of evolution is in stark disagreement with our observations. With this line of reasoning we have bounded the fraction of DM which may be hot: most of it will be cold. So, neutrinos may be a part of DM, but they cannot be the dominant part of it.

Dark Matter around 1 keV in mass would be a candidate for this.

Let us consider **Thermal Dark Matter**, which has been in equilibrium with the plasma for some time in the early universe, so it is cold when it decouples. By assuming chemical equilibrium we can figure out the number density of those particles. Consider a heavy particle X, with mass larger than, say, 10 GeV, and which is stable or quasi-stable (that is, such that its lifetime is long compared to the age of the universe).

[6] https://www.youtube.com/watch?v=kTHNpusq654

Then, its equilibrium number density will be given by

$$n_X^{\mathrm{eq}} = g_X \left(\frac{M_X T}{2\pi} \right)^{3/2} \exp\left(-\frac{M_X}{T} \right). \tag{2.94}$$

The number of X in a certain comoving volume can only change by annihilation and creation: $X\overline{X} \leftrightarrow$ light particles.

If the rate Γ of this process is $> H$ at a certain temperature T, then the number density of X is indeed the equilibrium one. However, as T decreases below M_X we will reach a point at which the light particles will not be energetic enough to form X pairs. At this point, only the annihilation of $X\overline{X}$ pairs will be able to occur, so their number will diminish.

After some more time, their encounters will not be likely anymore, so we say that the abundance of X *freezes out*. From that moment onward, n_X will not change in the comoving frame.

Note that this may still be before *kinetic* equilibrium is reached: scatterings like $X + A \rightarrow X + B$ can still happen. The number of Xs is governed by the Boltzmann equation:

$$\frac{\mathrm{d}n_X}{\mathrm{d}t} + 3Hn_X = -\left\langle \sigma_{\mathrm{annihilation}} v \right\rangle \left(n_X^2 - n_{X,\mathrm{eq}}^2 \right). \tag{2.95}$$

The cross section $\sigma_{\mathrm{annihilation}}$ is averaged over all the equilibrium momentum space distribution functions of the particle species, and summed over the possible annihilation channels.

Here start many calculations to get the density of X today…

2.5.2 WIMPs

What is the window of the parameters of this particle which could give us the Ω_{DM} we observe?

If the particle is a Weakly Interacting Massive Particle, we get the “WIMP miracle”: with very simple assumptions, we get Ω between 1 and 10^{-1}. Let us show this.

Claim 2.5.1. *For a heavy relic particle the following holds:*

$$\Omega_X h^2 \approx 10^{-10} \left(\frac{\mathrm{GeV}^{-2}}{\langle \sigma v \rangle_{\mathit{eff}}} \right) \frac{1}{\sqrt{g_*(t_f)}} \log \left(\frac{g_X M_X M_P \langle \sigma v \rangle}{(2\pi)^{3/2} \sqrt{g_*}} \right). \tag{2.96}$$

We apply this to a particle with mass $M_X \sim 100\,\mathrm{GeV}$, which interacts through the weak interaction and therefore has

$$\langle \sigma v \rangle \sim \frac{\alpha_w^2}{M_X^2} \sim \frac{10^{-4}}{10^4} \mathrm{GeV}^{-2}. \tag{2.97}$$

The values of $\sqrt{g_*}$ we need are $\sqrt{g_*} \sim 10$ at $T > 200\,\mathrm{GeV}$ and $\sqrt{g_*} \sim 3$ at $100\,\mathrm{MeV}$. Plugging these in we get

$$\Omega_X h^2 \approx 2 \times 10^{-10} 10^8 \frac{1}{10} \log \left(\frac{10^2 10^{19} 10^{-7}}{10^2} \right) \tag{2.98}$$

$$\Omega_X \approx 2 \times 2 \times 10^{-3} 12 \log 10 \approx 10^{-1}\,, \tag{2.99}$$

since $h^2 \sim 1/2$.

So, a particle with $M_X \sim 100\,\mathrm{GeV}$ which is weakly coupled, $\alpha_w \sim 0.01$ could explain our dark matter observations!

So, the WIMP possibility looks good; alas, the LHC has not seen any new particles at the electroweak scale.

We have a result of Ω_X between 0.1 and 1, but the higher end of this range is not good for our purposes: we know that $\Omega_{DM} < 0.25$. So, we can put a lower limit on $\sigma_{\text{annihilation}}$ in order to avoid the over-production of dark matter (or, more accurately, we want to avoid the fact that too much of it remains). What we find is

$$|\sigma_{\text{ann}} v| > 0.3 \div 10 \times 10^{-8}\,\mathrm{GeV}^{-2}\,. \tag{2.100}$$

Now, if the mass of X were large (like, more than $10\,\mathrm{TeV}$) then we would get $\Omega_X > 10$, which would definitely over-close the universe. So, we expect DM not be super massive.

There are ways to get around this restriction: one is a strongly coupled theory, in which case X would be annihilated more; then it would be a Strongly Interacting Massive Particle but it would work.

Another possibility is for X not to have been in thermal equilibrium, so that we have non-thermal DM, so that their number is not linked to the abundance of other particle species.

But wait, what does thermal equilibrium have to do with it? We should only consider chemical equilibrium for the variation of the number of X, right?

For regular WIMPs, though, the bound $M_X < 100\,\mathrm{TeV}$ remains.

We can expect that there are more than one kind particle constituting DM. A possibility is the presence of a "Dark Sector" in the SM. It is generally assumed that there is a sort of "portal", which makes the ordinary matter sector allowing the sectors to communicate.

A very simple kind of dark sector would be constituted by right-handed neutrinos: they only interact with the Higgs boson, with terms like $\overline{\nu}_L H^0 \nu_R$.

How do we **search for WIMPs**? There are two ways, one is direct: we set up a target and wait for X to hit it. If we set up a target with $10\,\mathrm{kg}$ of nuclei with $A \sim 100$, and have $M_X \sim 100\,\mathrm{GeV}$ and $\sigma_0 \sim 10^{-10}\,\mathrm{GeV}^{-2}$, then we will have

$$\begin{aligned} \text{event rate} &\approx v_X n_X N_A \sigma_{AX} && (2.101) \\ &\approx v_X \frac{\rho_{DM}^{\text{local}}}{M_X} N_A \sigma_{AX}\,, && (2.102) \end{aligned}$$

where $N_A \approx 6 \times 10^{25}$ is the number of A nuclei, $\sigma_{AX} \sim 10 A \sigma_0$ is the cross section of the nucleus-X process, and $\rho_{DM}^{\text{local}} \approx 0.3\,\mathrm{GeV/cm^3}$.

This then yields (after multiplying by $\hbar^2 c^2$, and taking $v_X \sim 2\,\mathrm{km/s}$)

$$\text{event rate} \approx 10^{-8}\,\mathrm{s}^{-1}\,, \tag{2.103}$$

or about 1 event a year.

The other method is indirect, we look for the products of $X\overline{X}$ annihilation in our astronomical observations: photons, protons, electrons, neutrinos or their antiparticles.

2.5.3 WIMP candidates for Dark Matter

Tuesday 2020-5-19, compiled 2020-07-02

WIMP means Weakly Interacting Massive Particle, this is a so-called heavy stable relic of the early universe.

So, we require stability; this could also be a very long-lived particle, the only requirement is that it be still abundant now.

The cross section must also be small.

Let us consider a typical WIMP, with mass $M_x \sim 100\,\text{GeV}$.

Let us consider a scenario in which the particle remains in equilibrium even when the temperature drops below the mass of the particle.

We want to request that $T_f < M_X$, where T_f is the temperature of the freezeout.

The value of $\rho_X/\rho_C = \Omega_X$ is of the order $1/4$ today. If we had $n_X \sim n_\gamma$ before the freezeout, because then DM was radiation-like.

Today we have around a baryon every billion photons, and $\Omega_B \sim 5\,\%$. The statistics of the DM distribution are given by

$$(M_X T)^{3/2} \exp\left(-\frac{M_X}{T}\right) , \tag{2.104}$$

so the exponential suppression plays a very important role as the temperature decreases. We cannot over-close the universe, if we get $\Omega_X > 1$ the universe would become strongly closed.

Now, let us compute the freezeout temperature, assuming that the particle's mass is $100\,\text{GeV}$. We want to show that $T_f < M_X$.

What do we mean by "equilibrium" precisely? There are two possible meanings; one is the chemical equilibrium, and the other is the kinetic equilibrium.

The annihilation of X with its antiparticle changes the total number of particles. In discussing the freezeout temperature we are interested in the *chemical* equilibrium. So, we must ask: "what is the temperature at which the annihilation Γ_X is equal to the expansion rate of the universe H?"

The first is given by

$$\Gamma_X = \langle \sigma v \rangle \, n_X \, , \tag{2.105}$$

where

$$n_X = (M_X T)^{3/2} e^{-M_X/T} \, , \tag{2.106}$$

while $\langle \sigma v \rangle$ is model-dependent. Therein lies the physics of the problem. Without a specific model, we take something like

$$\sigma \sim \frac{\alpha_W^2}{M_w^4} \, . \tag{2.107}$$

We plug everything into the expression, using a standard cosmology for H. We can get an order-of-magnitude estimate. If we did the calculation more properly, we would be able to get the Boltzmann equation:

$$\frac{\mathrm{d} n_x}{\mathrm{d} t} + 3 H n_x = - \langle \sigma v \rangle \left(n_x^2 - \right) , \tag{2.108}$$

complete formula

which can be found in the notes. We find

$$\Omega_X h^2 \approx 10^{-10} \left(\frac{\mathrm{GeV}^{-2}}{\langle \sigma v \rangle} \right) \frac{1}{\sqrt{g(T_f)}} L\,, \tag{2.109}$$

where

$$L = \log \left(\frac{g_X M_X}{} \right) \dots\,, \tag{2.110}$$

complete formula

which is generally of the order of 20 to 30.

So in the end the temperature is given by

$$T_f \approx M_X L^{-1}\,. \tag{2.111}$$

We are, however, assuming that the annihilation goes like the weak interaction — is this accurate? If the value α_W were different, we could get the correct number for Ω_X. There are good reasons to think that there might be new physics in the electroweak scale.

How is the estimate for Ω_X derived?

Suppose we asked for $\Omega_X < 0.3$. Then, we need to reduce the number of X, n_X.

2.5.4 Axions

Wednesday 2020-5-20, compiled 2020-07-02

The search for WIMPs can be indirect or direct: direct searches look for DM interacting with a target with a large cross section, indirect searches look for the recoil of particles.

There might be a DM "wind" passing through the Earth. We would then expect a seasonal modulation of this effect, because of the Earth's orbit around the Sun.

Is there a modulation of this kind in the neutrino flux? Yes, at Gran Sasso they have shown this with very high confidence.

We have our SM particles and the "Dark Sector": there might be a "portal" between them, for instance the Higgs boson.

Suppose the Dark Sector had a $U(1)$ symmetry: then, we would have a sort of "dark photon". If this symmetry were spontaneously broken, we could then have a mixing between the two sectors.

A particular light pseudoscalar DM candidate is called the **axion**. Its interest is not only as a DM candidate, but also in other contexts: it is linked to BTSM physics below the electroweak scale.

It was introduced by particle physicists first, and then it was understood that it might be useful as DM.

Weak interactions violate CP symmetry. This is due to an effect related to the electroweak interaction.

In the QCD Lagrangian we have a kinetic term $G^a_{\mu\nu} G_a^{\mu\nu}$; then we can also add a term $G_{\mu\nu}\widetilde{G}^{\mu\nu}$, where $\widetilde{G}$ is the Hodge dual of the field strength. This would produce a CP-violating term. We can write it but people were not worried: this term can be written as a 4-divergence, so it could be removed.

However, it could be shown that due to instantons — nonperturbative quantum effects — there was an anomalous current, the term could not be neglected.

The term can still be written as a divergence, but because of the instanton the asymptotic values of the field are not zero anymore: therefore, the surface integral in the action is not zero anymore.

Then, a term $\theta_0 G\widetilde{G}$ was added. This also affected the phases of quark transitions. Then, we introduce a term

$$\overline{\theta} = \theta_0 + \arg\left(\det M\right), \tag{2.112}$$

where M is the quark mass matrix. This $\overline{\theta}$ will then multiply the $G\widetilde{G}$. This is a free parameter of the theory: the quark mass matrix depends on the Yukawa couplings, and θ_0 is completely free. This should be $\overline{\theta} < 10^{-9}$ to comply with experiment: it looks like fine-tuning!

There is no physical reason why the parameter should be small. This is weird: there must be some hidden symmetry.

We can introduce a $U(1)$ global symmetry, called the Pacci-Quinn symmetry. This is discussed by Rubakov.

Since the field H feels the symmetry, in terms like $\overline{Q}_L H d_R$. The spontaneous breaking of this symmetry yields a goldstone boson called the axion.

The mass of this pseudo-Goldstone boson depends on the scale at which this symmetry is broken, Q.

We can introduce additional scalars, which are singlets or doublets under SU(2), whose VEV is much larger than the electroweak scale.

If the photons inside a star could convert into axions, they could speed up the cooling down of a star: our stellar evolution observations then allow us to give a bound

$$f_{PQ} > 10^9\,\text{GeV}\,. \tag{2.113}$$

What happens to the axions after they are produced? The axion is light, since $m_a \sim 1/f_Q$. The mass is small, like $< 10^{-5}\,$eV. They oscillate, moving around the minimum in their potential. Most off the energy is contained in this oscillation, the mass contributes very little.

This is then a type of DM which is Cold, but also very light! In order to not over-close the universe we must also have $f_{PQ} < 10^{12}\,$GeV.

How would we detect axions? They interact with photons, so we can have an experiment in which we have photon conversion into an axion.

We have a spectacular effect which is called "light shining through a wall": we send a photon towards a wall, it will not pass through usually. Suppose that we put a source of a strong magnetic field: if there are axions, we could have photon-photon conversion into an axion, which can pass through the wall and then emerges as an axion. We then put another strong source of magnetic field which can deconvert the axion into a photon.

So far, this has not provided any evidence for the existence of the axion.

2.6 Matter-antimatter asymmetry

Wednesday 2020-6-3, compiled 2020-07-02

Sacharov conditions.

Bibliography

[CV17] Alain Coc and Elisabeth Vangioni. "Primordial Nucleosynthesis". In: *International Journal of Modern Physics E* 26.08 (July 23, 2017), p. 1741002. ISSN: 0218-3013. DOI: 10.1142/S0218301317410026. URL: https://www.worldscientific.com/doi/abs/10.1142/S0218301317410026 (visited on 2020-07-02) (cit. on p. 79).

[Cyb+16] Richard H. Cyburt et al. "Big Bang Nucleosynthesis: 2015". In: *Reviews of Modern Physics* 88.1 (Feb. 23, 2016), p. 015004. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.88.015004. arXiv: 1505.01076. URL: http://arxiv.org/abs/1505.01076 (visited on 2020-07-02) (cit. on p. 79).

[GR11] D. S. Gorbunov and V. Rubakov. *Introduction to the Theory of the Early Universe, Hot Big Bang Theory*. Singapore: World Scientific Publishing Company, 2011 (cit. on p. 64).

[Hus16] Lars Husdal. "On Effective Degrees of Freedom in the Early Universe". In: *Galaxies* 4.4 (Dec. 17, 2016), p. 78. ISSN: 2075-4434. DOI: 10.3390/galaxies4040078. arXiv: 1609.04979. URL: http://arxiv.org/abs/1609.04979 (visited on 2020-06-30) (cit. on p. 70).

[KT94] E. Kolb and M. Turner. *Early Universe*. New York: Westview Press, 1994 (cit. on pp. 64, 65).

[Pes19] M. E. Peskin. *Concepts of Elementary Particle Physics*. Oxford Master Series in Particle Physics, Astrophysics and Cosmology. Oxford University Press, 2019 (cit. on pp. 2, 5, 8, 10, 17, 25, 33, 38, 39, 46, 56).

[PS95] Micheal E. Peskin and Daniel V. Schroeder. *An Introduction To Quantum Field Theory*. Oct. 2, 1995 (cit. on p. 38).

[Tis20] Jacopo Tissino. *Theoretical Physics Notes*. Notes for the course held by Stefano Rigolin at the university of Padua in spring 2020. 2020, p. 192. URL: https://github.com/jacopok/notes/blob/master/ap_second_semester/theoretical_physics.pdf (cit. on pp. 7, 17, 18, 27, 28, 30, 44).