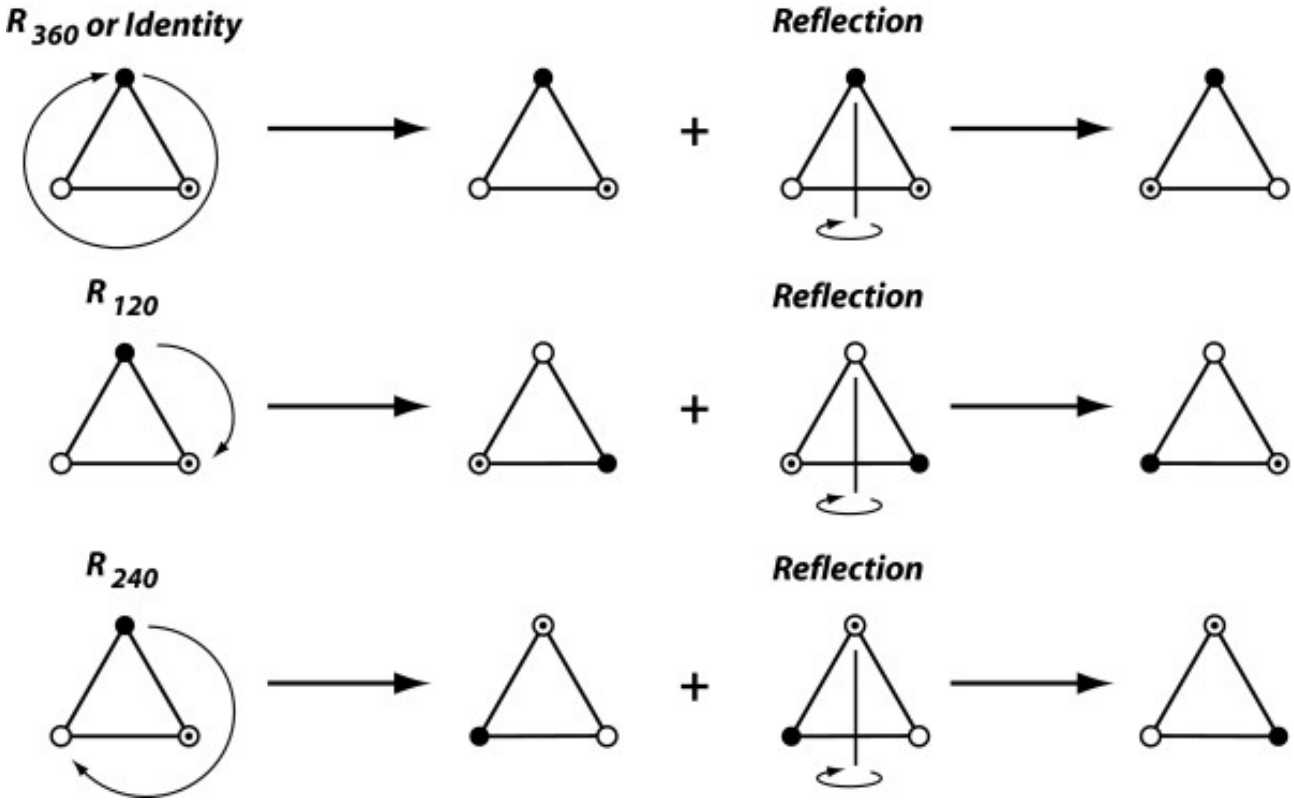
Transformations and symmetries

Panos Christakoglou[1,2]



[1] P.Christakoglou@uu.nl - Utrecht University, Leonard S. Ornsteinlaboratorium Princetonplein 1, 3584CC Utrecht, The Netherlands
[2] Panos.Christakoglou@nikhef.nl - Nikhef, Science Park 105, 1098XG Amsterdam, The Netherlands

Contents

1 Introduction to transformations, groups and symmetries 1

1.1 The Lagrangian formalism 1

1.2 Global, local, continuous, discrete transformation 2

1.3 When is a quantity conserved and how can they be found? 5

1.4 Introduction to groups 8

1.4.1 Lie groups 8

1.5 Continuous transformations 9

1.5.1 Invariance under translations 10

1.5.2 Invariance under rotations 11

1.6 Conservation of charge 12

1.6.1 Global gauge transformation 13

1.6.2 Local gauge transformation 13

1.7 Discrete transformations 14

1.7.1 Parity transformations 14

1.7.2 Charge conjugation 15

1.7.3 Baryon number 15

1.7.4 Isospin 16

1.7.5 Flavour 17

Chapter 1
Introduction to transformations, groups and symmetries

This chapter introduces some of the basic concepts that will be used later on to build the various theories that form the Standard Model of particle physics. The main concepts that we will focus on are the ones of the various types of transformations (e.g. continuous versus discrete), how they can lead to symmetries of a system and how these symmetries in turns can be translated into conservation laws.

We will start with an essential reminder of the Lagrangian formalism. We will then discuss various types of transformations focusing on their differences: global versus local, continuous versus discrete. After a very brief introduction to groups, we will then see when a quantity conserved and how it can be found. We will move to continuous transformations, examining what it means that a system is invariant under translations or under rotations in coordinate space. We will then turn our focus on a specific, discrete quantity the charge and we will introduce the notions of global and local gauge transformations. At first glace these notions will feel strange but do not be afraid! We are just planting the seeds, in the form of basic ideas and notions such that by the time we reach the chapter where we will need to apply global and local gauge invariance on a given theory, the plants will be in full blossom. We will conclude the chapter by introducing some of the basic discrete transformations that we will use at a later stage, the one of parity and charge conjugation.

1.1 The Lagrangian formalism

In classical dynamics the motion of a particle of mass m that feels a force $\mathbf{F}$ is described in terms of forces using the second Newton law:

$$\mathbf{F} = m\mathbf{a} = m\ddot{\mathbf{x}}$$

If the force is conservative, it can be expressed as the gradient of a scalar potential energy U according to

$$F = -\nabla U$$

so that the law of Newton can now be written as

$$m\ddot{\mathbf{x}} = -\nabla U$$

The same equation of motion can be obtained from the Lagrangian, defined by

$$L = T - U, \tag{1.1.1}$$

where $T = \frac{1}{2}m\dot{\mathbf{x}}^2$ and U are the system's kinetic and potential energies.

The Lagrangian is a function of the generalised coordinates q_i (e.g. $q_3 = z$) and their time derivatives $\dot{q}_i$ (e.g. $\dot{q}_3 = \dot{z} = v_z$, with v_z being the velocity of the particle along the z-axis). In the Lagrangian formalism the fundamental law of motion is given by the Euler-Lagrange equation:

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big) - \frac{\partial L}{\partial q_i} = 0 \tag{1.1.2}$$

Another way is to express the same quantities in terms of the Hamiltonian of the system. For this we need to calculate the generalised momenta given by

$$p_i(q_i, \dot{q}_i, t) = \frac{\partial L}{\partial \dot{q}_i}$$

Velocities are then calculated by inverting the previous equation. The Hamiltonian is then expressed as the Legendre transformation of L, according to

$$\mathscr{H} = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \sum_i \dot{q}_i p_i - L \tag{1.1.3}$$

1.2 Global, local, continuous, discrete transformation

In this section we are going to get familiar with different types of transformations. Towards the end of the section we are going to make the connection with symmetries. It is essential to note however that when we talk about transformations we do not necessarily imply a symmetry i.e. there are transformations that do not lead to symmetries.

Transformations are usually described by operators. These operators are usually described by matrices. An example is shown in fig. 1.1. The transformation is described by a matrix of the form

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

that transforms one reference system X into another $X^{'}$ according to:

$$x^{'} = x\cos\theta + y\sin\theta$$

and

$$y^{'} = -x\sin\theta + y\cos\theta$$

At this stage, let us open a parenthesis and discuss about some properties of matrices. Suppose we have a matrix A with elements α_{ij} then

- A^* is the complex conjugate matrix with elements α^*_{ij},
- $\tilde{A}$ is the transposed matrix with elements α_{ji},
- $A^\dagger$ is the Hermitian conjugate or transposed matrix with elements α^*_{ji}.

Let us also make another parenthesis to remind ourselves of some basic notations from the course about quantum mechanics.

- The "ket" notation of a quantum state $|\psi\rangle$ which is represented by a column vector:

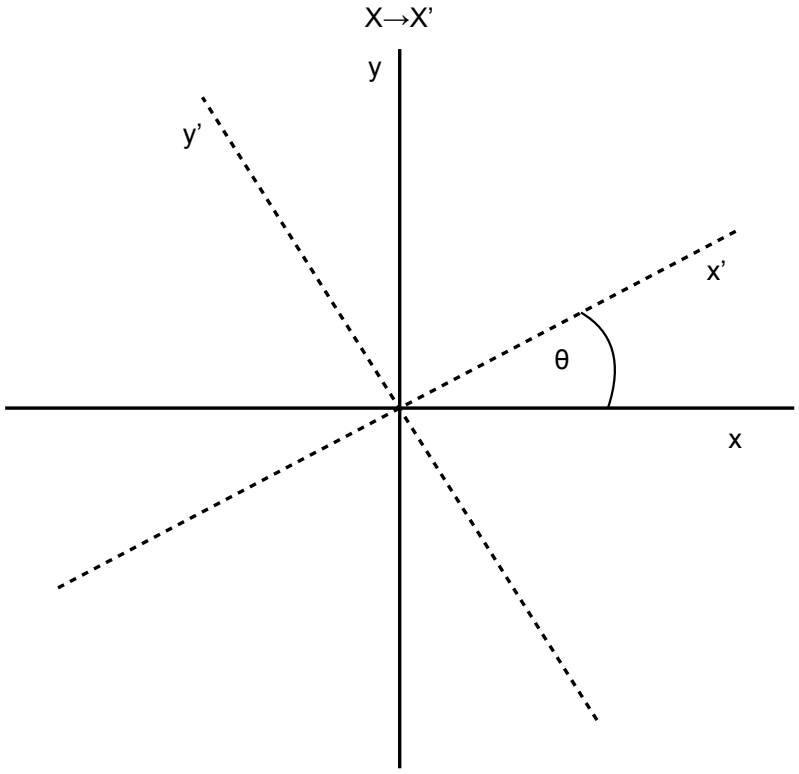


Fig. 1.1: An example of a familiar transformation: the rotation of a reference system by an angle θ.

$$|\psi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ . \\ . \\ . \\ a_n \end{pmatrix}$$

- The "bra" notation of a quantum state $\langle\psi| \equiv |\psi\rangle^\dagger$, the conjugate transpose state, which is represented by a row vector of complex conjugate elements:

$$\langle\psi| = \begin{pmatrix} a_1^* & a_2^* & a_3^* & \dots & a_n^* \end{pmatrix}$$

- The "dot" product given by $\psi_\alpha \psi_\beta^*$ which represents a number C such that $\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^\dagger = C^\dagger = C^*\langle\alpha|\beta\rangle^*$
- Suppose that we have an operator $\hat{P}$ that transforms an eigenstate $|\alpha\rangle$ into another one $|\gamma\rangle$ according to $\hat{P}|\alpha\rangle = |\gamma\rangle$. The Hermitian conjugate action gives

$$(\hat{P}|\alpha\rangle)^\dagger = \langle\alpha|\hat{P} = \langle\gamma|$$

- For an operator U that represents a transformation, we define the Hermitian operator being the one that satisfies $U^\dagger = U$,
- for an operator U that represents a transformation, we define the unitary operator being the one that satisfies $U^\dagger = U^{-1} \Rightarrow U^\dagger U = I$.

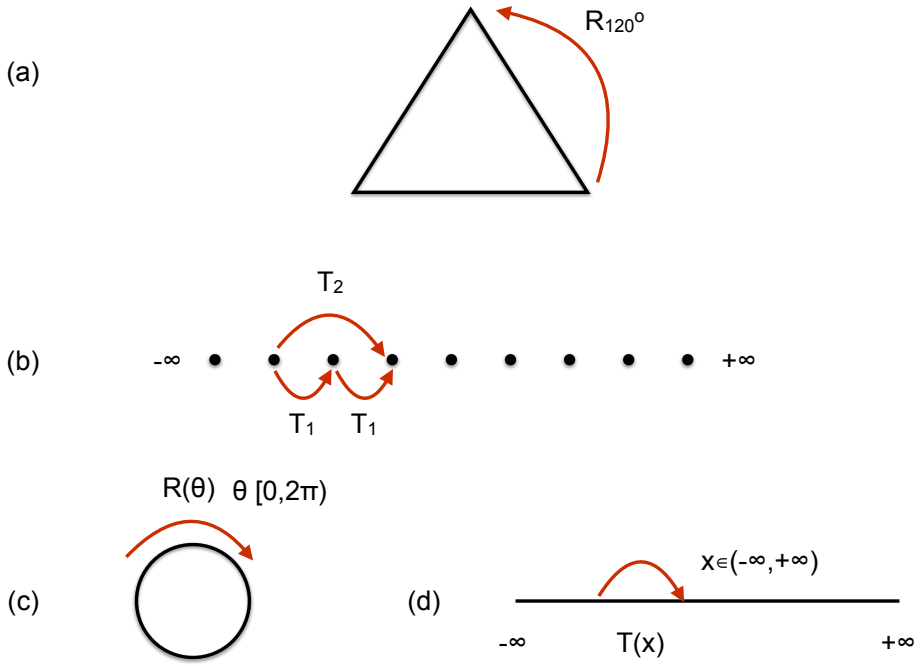


Fig. 1.2: Some of the different types of transformations: (a) finite, (b) infinite, (c) compact, (d) non-compact.

Let's now go back to transformations and review some of their basic types. There are two general types of transformations: discrete or continuous. Discrete transformations do not connect to the unit operator i.e. the one that leaves a system unchanged. An example of such transformation is the one of parity which represents a mirroring operation around the origin. On the other hand, continuous transformations are the ones that connect smoothly to the unit operator.

Two examples of discrete transformations are given in fig. 1.2–a and fig. 1.2–b. Discrete transformations are the ones that take discrete values e.g. rotation by 120 degrees in fig. 1.2–a to match one vertex position to the next one, translation that moves one point to the next or the next-to-next one in fig. 1.2–b. These two cases give also the order of the group that describes these transformations. One has to ask the question of how many independent, unique rotations can we have in fig. 1.2–a and still get the same shape? The answer is three: we can rotate by 120 or 240 degrees while the last one is the identity i.e. we rotate by 0 degrees. These type of discrete transformations whose order is a finite number are called finite transformations. On the other hand, fig. 1.2–b illustrates a discrete transformation with infinite elements: this transformation is called infinite.

Moving now to continuous transformations, that can transform a system using any value. In fig. 1.2–c one can rotate the ring by any angle between 0 and 2π. These are called compact continuous transformations. If we now take the points of fig. 1.2–b and reduce the spacing between them in a way that we recover a continuous line, we can then move any point along this line. Each point can be moved at will and the values of this shift are only limited by the length of this line. If the line is infinite, then the relevant transformation is called non–compact.

In all these previous cases, each transformation led directly to the same shape i.e. the transformation was leading to a symmetry. It does not have to be like this of course. Not all transformations lead to a symmetry. Suppose we have a system of nine points like the one of fig. 1.3. Let us move all points at the same time by a fixed amount e.g. Δx along the x–axis. This is called a global transformation. The new system that emerges (see top right in the same figure) is identical to the initial one. That means that if we apply a global transformation to a system, this leads to global symmetry. Let us now promote this transformation from global to local, that is we are allowed to move each individual point of the initial system by a different amount i.e. by Δx_i that is obviously different for each of the i–points. Unless these individual translations

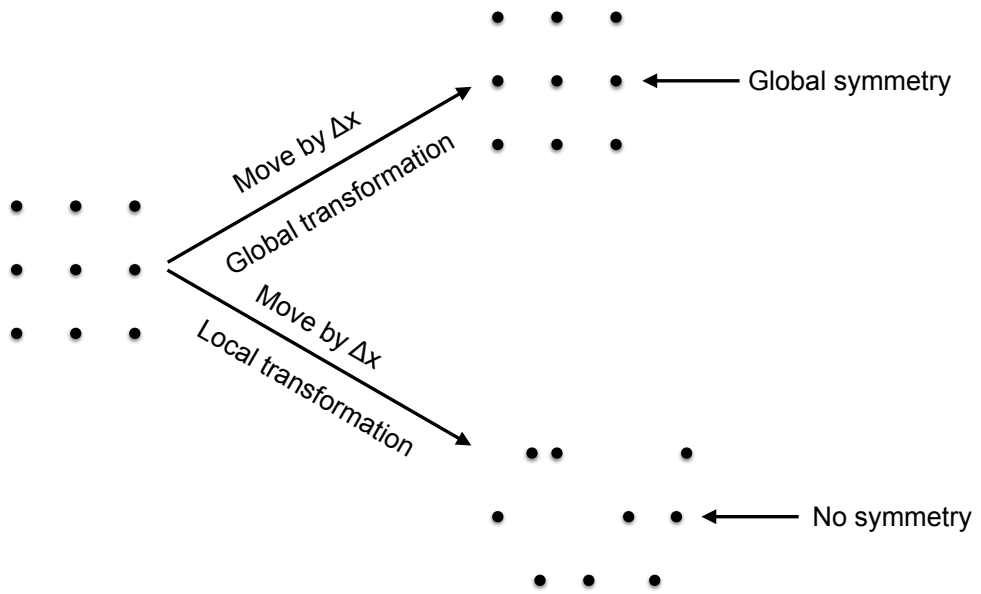


Fig. 1.3: An example of how a global transformation can lead to a global symmetry. Applying the same type of transformation locally does not necessarily lead to a local symmetry.

are identical between all point in which case we go back to a global transformation, we are led to a system that does not look like the initial one (see bottom right plot of fig. 1.3). **If a transformation leads to global symmetry, this does not imply automatically local symmetry.**

Let us now review things starting from a local transformation that leads to a local symmetry. On the left part of fig. 1.4, we can see a system of nine rings. Let us apply a local transformation which involves now a rotation by $\Delta\theta_i$ for each of the nine rings. This, as you can imagine, can lead to an identical to the initial one system as depicted in the top right part of the same figure. We applied a local transformation that led to a local symmetry. If now we apply a global transformation i.e. we rotate all rings by the same angle $\Delta\theta$, then again the final system is identical to the initial one (see the bottom right system of fig. 1.4). **If a local transformation leads to local symmetry, this in turns leads automatically to global symmetry.**

In conclusion, a symmetry under local transformations leads to global symmetry. Reversely, symmetry under global transformations does not necessarily lead to local symmetry!

1.3 When is a quantity conserved and how can they be found?

Let us now consider the time-independent Hamiltonian of a system $\mathscr{H}$. The wave function of this system satisfies the Schrodinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \mathscr{H}\Psi$$

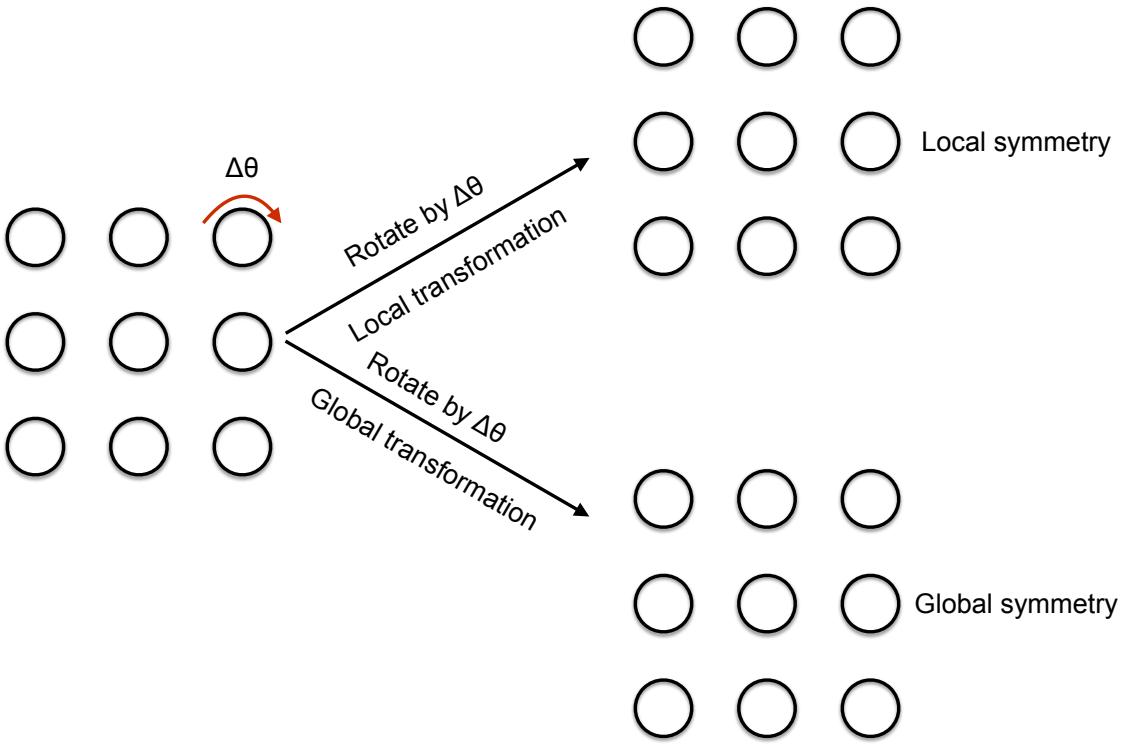


Fig. 1.4: An example of how a local transformation can lead to a local symmetry which translates automatically into a global symmetry.

An observable is usually represented by a quantum mechanical operator $\hat{F}$ whose expectation value corresponds to a measurement:

$$\langle F \rangle = \int \Psi_a^*(\mathbf{x}) \hat{F} \Psi_a(\mathbf{x}) d^3x = \langle \Psi | F | \Psi \rangle$$

Let's now consider two states $|a\rangle$ and $|b\rangle$. A quantity similar to $\langle F \rangle$ can also be constructed

$$\langle F_{ba} \rangle = \int \Psi_b^*(\mathbf{x}) \hat{F} \Psi_a(\mathbf{x}) d^3x$$

The quantity $\langle F_{ba} \rangle$ is called the matrix element that describes the transition of a system from one quantum state (i.e. $|a\rangle$) to another (i.e. $|b\rangle$). The expectation value of a state can be measured and is thus a real number. That means that it must be Hermitian ($F^\dagger = F$) i.e. it is the diagonal element of F_{ba} with $b = a$: $\langle F \rangle = F_{aa}$. The value of an observable in the state Ψ is given as we saw from $\langle F \rangle$.

Let's now see when an observable is conserved, when it is independent of time. Both Ψ and it's Hermitian conjugate $\Psi^\dagger$ satisfy the same Schrodinger equation:

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = \mathscr{H} |\Psi\rangle$$

and

$$\left(i\hbar \frac{\partial |\Psi\rangle}{\partial t} \right)^\dagger = (\mathscr{H} |\Psi\rangle)^\dagger \Rightarrow -i\hbar \frac{\partial \langle \Psi |}{\partial t} = |\Psi\rangle^\dagger \mathscr{H}^\dagger = \langle \Psi | \mathscr{H}$$

The expectation value of an observable is $\langle F \rangle = \langle \Psi | F | \Psi \rangle$ and if $\langle F \rangle$ is conserved so is $\langle F^* \rangle = \langle \Psi | F^\dagger | \Psi \rangle$:

$$\frac{\partial \langle F \rangle}{\partial t} = \frac{\partial \langle \Psi |}{\partial t} F |\Psi\rangle + \langle \Psi | F \frac{\partial |\Psi\rangle}{\partial t} = 0 \Rightarrow$$

$$-\frac{1}{i\hbar}\langle \Psi | \mathscr{H} F | \Psi \rangle + \frac{1}{i\hbar}\langle \Psi | \Psi F \mathscr{H} | \Psi \rangle = 0 \Rightarrow \frac{1}{i\hbar}\langle \Psi | F\mathscr{H} - \mathscr{H} F | \Psi \rangle = 0 \Rightarrow$$

$$F\mathscr{H} - \mathscr{H} F = 0 \Rightarrow [\mathscr{H}, F] = 0$$

An observable constant of motion F is Hermitian and commutes with the Hamiltonian of the system.

But how can we find conserved quantities? We can write down the Hamiltonian of a system and test the commutation criterium for all possible observables. As you can imagine this is not particularly feasible. The Hamiltonian of a system does not need to be fully known. We just need to find an invariance of $\mathscr{H}$ under a given transformation. This will lead to a conserved quantity.

Let us then introduce a transformation operator U that transforms the wave function as

$$\Psi'(\mathbf{x}, t) = U\Psi(\mathbf{x}, t)$$

The normalisation equation of the wave function should not change:

$$\int \Psi^* \Psi d^3x = \int (U\Psi) * (U\Psi) d^3x = \int \Psi^* (U^\dagger U) \Psi$$

It can thus be seen that the transformation operator must be unitary

$$U^\dagger U = U U^\dagger = I$$

The transformation operator U is then a symmetry operator if Ψ' satisfies the same Schrodinger equation as Ψ

$$i\hbar \frac{\partial \Psi}{\partial t} = \mathscr{H}\Psi$$

and

$$i\hbar \frac{\partial \Psi'}{\partial t} = \mathscr{H}\Psi' \Rightarrow i\hbar \frac{\partial (U\Psi)}{\partial t} = \mathscr{H} U\Psi \Rightarrow$$

$$i\hbar U \frac{\partial (\Psi)}{\partial t} = \mathscr{H} U\Psi \Rightarrow i\hbar U^{-1} U \frac{\partial (\Psi)}{\partial t} = U^{-1}\mathscr{H} U\Psi \Rightarrow$$

$$i\hbar \frac{\partial (\Psi)}{\partial t} = U^{-1}\mathscr{H} U\Psi \Rightarrow \mathscr{H} = U^{-1}\mathscr{H} U \Rightarrow U\mathscr{H} - \mathscr{H} U = 0 \Rightarrow$$

$$[\mathscr{H}, U] = 0$$

In conclusion:

- The operator $\hat{F}$ is an observable and represents a physical quantity, its expectation value is real and $\hat{F}$ is Hermitian i.e. $\hat{F} = \hat{F}^\dagger$.
- An observable constant of motion F is Hermitian and commutes with the Hamiltonian of the system.
- The transformation operator U is unitary and changes the wave function from Ψ to Ψ'.

- The operator U is a symmetry operator if it commutes with the Hamiltonian of a system.

In general, transformation operators are not always Hermitian and thus do not correspond to observables.

1.4 Introduction to groups

Group theory is a branch of mathematics that underlies the treatment of symmetry. A group G is a collection of elements (or operators), say (a_1, a_2,...,a_n), with well defined laws that describe how one can combine any two of the elements with an operator (let's indicate it with $\times$) so as to form the product such that the following four conditions are fulfilled:

- **Closure**: For every element a_i and a_j of the group G, their product (or in general an operator $\times$) $a_i \times a_j$ is also a member of the group G, such that $a_i \times a_j = a_k$.
- **Associativity**: The law of combination os associative i.e. $(a_i \times a_j) \times a_k = a_i \times (a_j \times a_k)$.
- **Identity element**: Every group G contains an identity element e such that for all a_i in G $a_i \times e = e \times a_i = a_i$.
- **Inverse element**: For all a_i of a group G there is a unique inverse element a_i^{-1} such that $a_i \times (a_i)^{-1} = (a_i)^{-1} \times a_i = e$.

The associativity requirement does not imply that the elements of the group commute. **A group for which any of the two elements commute is called Abelian, otherwise the group is call non-Abelian**. When a group contains a finite number of elements, let's say n, then the group is called **finite group of order n**.

1.4.1 Lie groups

As stated before, in these lectures we are not interested so much in discrete but rather in continuous transformations. These are described in terms of continuous groups, whose elements are labelled by a number of continuously variable real parameters, say $a_1, a_2, ..., a_n$, and by the relevant operators $g(a_1, a_2, ..., a_n) \equiv g(\mathbf{a})$. In particular we are interested in coordinate transformations that, as we will see later, are not only of space-time origin but can also be internal e.g. colour in SU(3). The following examples are indicative of these type of transformations which are usually represented by matrices:

- Rotations in three dimensions form the SO(3) group (i.e. special orthogonal group in three dimensions), whose elements are specified by three real numbers (i.e. two defining the axis of rotation and one the angle for the angle of rotation about the axis).
- Lorenz transformations form the group with six real variables i.e. three for the rotations and the other three for the velocity transformations.

By convention, parameterisations are arranged in such a way so that $g(0)$ is the identity element of the group. For a continuous group, the first group condition takes the form $g(\alpha) \times g(\beta) = g[\gamma(\alpha, \beta)]$, where the parameters γ are continuous functions of the parameters α and β. If the parameters γ are analytic functions of α and β the group is called a **Lie group**.

Let us consider a group of transformations defined by $x_i^{'} = f_i(x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_n)$, where x_i with $i = 1, 2, , ..., n$ are the coordinates on which the transformation acts and α_i are the real parameters of the transformation. By convention, $\alpha = \mathbf{0}$ is the identity transformation, such that $x_i = f_i(\mathbf{x}, \mathbf{0})$.

A transformation in the neighbourhood of the identity is then given by

$$dx_i = \sum_{\nu=1}^{r} \frac{\partial f_i}{\partial \alpha_\nu} d\alpha_\nu,$$

where the $d\alpha_\nu$ are infinitesimal parameters and the partial derivative is evaluated at the point $\mathbf{x}, \mathbf{0}$.

Let's now consider the change in a function $F(\mathbf{x})$ under the previous infinitesimal transformation. This reads:

$$F \to F + dF = F + \sum_{i=1}^{N} \frac{\partial F}{\partial x_i} dx_i$$

$$= F + \sum_{i=1}^{N} [\sum_{\nu=1}^{r} \frac{\partial f_i}{\partial \alpha_\nu} d\alpha_\nu] \frac{\partial F}{\partial x_i} \equiv [1 - \sum_{\nu=1}^{r} d\alpha_\nu i\hat{X}_\nu] F \tag{1.4.1}$$

where in Eq. 1.4.1,

$$\hat{X}_\nu \equiv i \sum_{i=1}^{N} \frac{\partial f_i}{\partial \alpha_\nu} \frac{\partial}{\partial x_i} \tag{1.4.2}$$

is a generator of infinitesimal transformations. Finite transformations are obtained by:

$$U(\alpha) = \lim_{n\to\infty} [\, 1 + i(\alpha/n) \cdot X \,]^n = \exp(i\alpha \cdot X) \tag{1.4.3}$$

where we have written $\sum_{\nu=1}^{r} \alpha_\nu \hat{X}_\nu = \alpha \hat{X}$.

There is a theorem which states that the commutator of two generators is always a linear combination of the generators

$$[\hat{X}_i, \hat{X}_j] = f_{ij}^k \hat{X}_k \qquad \text{(summation over } k \text{ implied)}$$

These commutation relations are called the algebra, and the complex numbers f_{ij}^k are called the structure constants of the group. It can be shown that these structure constants fully characterise the multiplication structure of a Lie group.

1.5 Continuous transformations

There is a large class of continuous transformations that depend on one or more continuous parameters, say α such that $|\psi'\rangle = U(\alpha)\,|\psi\rangle$. An example is the transformation induced by a rotation over an angle α of the coordinate system (passive rotation), or of the wave function (active rotation). Such transformations have the property that they can be written as a succession of infinitesimal deviations from the identity

$$U(\alpha) = \lim_{n\to\infty} \left(I + \frac{i\alpha}{n} F \right)^n = \exp(i\alpha F)$$

The factor 'i' is a matter of definition but important (see below). In the above, F is called the generator of U.[1] Now if U is unitary we have, to first order in α,

$$U^\dagger U = (I - i\alpha F^\dagger)\,(I + i\alpha F) = I + i\alpha(F - F^\dagger) = I$$

so that $F = F^\dagger$. In other words, **the generator of a unitary operator is Hermitian.** We now also understand the factor 'i' in the definition of a generator: without it the generator $G \equiv iF$ of a unitary operator would not be Hermitian but **anti-Hermitian**: $G = -G^\dagger$.

We have seen that a symmetry operator U commutes with the Hamiltonian so it remains to show that its generator will then also commute with H. The proof is very simple and starts with $U(\alpha)$, a symmetry operator. The infinitesimal transformation $U(\varepsilon)$ will also be a symmetry operator. Expanding to the first order in ε obtains:

[1] Exponentiation of an operator F should be interpreted as $\exp(i\alpha F) = I + i\alpha F + \frac{1}{2!}(i\alpha F)^2 + \cdots$. However, the familiar relation $e^A e^B = e^{A+B}$ is *only* true when A and B commute.

$$[H,U] \doteq [H,I+i\varepsilon F] = \underbrace{[H,I]}_{0} + i\varepsilon\,[H,F] = 0 \quad \rightarrow \quad [H,F] = 0$$

It is seen that if U is a unitary operator that commutes with the Hamiltonian, then its generator F is a Hermitian operator that also commutes with the Hamiltonian. A multiplication of continuous symmetry operators corresponds to the addition of their generators in the exponent. The conserved quantum numbers, which are related to F and not to U, are therefore additive.

1.5.1 Invariance under translations

Let us assume that we impose an infinitesimally small translation. For simplicity we also assume that this translation takes place along one axis, let's say along the x–axis (you can easily generalise this example into translations in whichever axis). This action is then described by

$$x \rightarrow x^{'} = x + \varepsilon \Rightarrow dx^{'} = x^{'} - x = \varepsilon,$$

where the parameter ε describes the infinitesimally small translation.

As we have seen, the generators of a transformation U are given by Eq. 1.4.2. In this particular case, the group f_i that describe the transformations is simply $f_x = dx^{'}$ while the variable that describes this transformation is $\alpha_v = \varepsilon$.

The generator is then given by

$$\hat{X}_x = i\frac{\partial f_x}{\partial \varepsilon}\frac{\partial}{\partial x} = i\frac{\partial (dx^{'})}{\partial \varepsilon}\frac{\partial}{\partial x} = i\frac{\partial}{\partial x}$$

But what does this generator represent? The quantum mechanical operator of momentum is $\hat{P} = -i\hbar\nabla$, which for the component along the x–axis reads $\hat{P}_x - i\hbar\frac{\partial}{\partial x}$. That means that the generator $\hat{X}_x$ is identified as

$$\hat{X}_x = i\frac{\partial}{\partial x} = -\frac{1}{\hbar}\hat{P}_x$$

That means that if a system is invariant under translations along the x–axis, then the x–coordinate of the momentum vector is conserved.

Another way to extract the same as above is to start from the quantum state $\Psi(x)$ and apply the same transformation that will change the wave function accordingly: $\Psi(x) \rightarrow \Psi^{'}(x') = \Psi(x+\varepsilon)$. Let us now expand $\Psi(x)$ according to

$$\Psi(x) \approx \Psi^{'}(x) + \frac{d\Psi^{'}(x)}{dx}\varepsilon = \left(1 + \varepsilon\frac{d}{dx}\right)\Psi^{'}(x) \Rightarrow$$

$$\left(1-\varepsilon\frac{d}{dx}\right)\Psi(x) = \left(1-\varepsilon\frac{d}{dx}\right)\left(1+\varepsilon\frac{d}{dx}\right)\Psi^{'}(x) \Rightarrow \left(1-\varepsilon\frac{d}{dx}\right)\Psi(x) = \Psi^{'}(x) + O(\varepsilon^2)\Psi^{'}(x) \Rightarrow \Psi^{'}(x) \approx \left(1-\varepsilon\frac{d}{dx}\right)\Psi(x)$$

The operator of a transformation can be written as $U = e^{i\varepsilon\hat{X}} \approx 1 + i\varepsilon\hat{X}$. Comparing this equation with the previous one, one can identify that

$$\hat{X}_x = i\frac{d}{dx} = -\frac{1}{\hbar}\hat{P}_x$$

1.5.2 Invariance under rotations

Rotations in three dimensions are usually represented by a matrix R which is 3×3:

$$\mathbf{x}^{'} = R\mathbf{x}$$

It follows that the length is invariant under such transformations, such that $\mathbf{x}^{'\dagger} \cdot \mathbf{x}^{'} = \mathbf{x}^{\dagger} \cdot \mathbf{x}$. This in turns implies that $R^{\dagger}R = I$ so that the matrix R is orthogonal. Based on the previous:

$$1 = det(R^{\dagger}R) = det(R^{\dagger}) \cdot det(R) = [det(R)]^2,$$

which needs $det(R) = \pm 1$. The matrices with a negative determinant include a parity transformation which is not continuous and thus not connected with the identity of the transformation. This leaves only the matrices with $det(R) = 1$ to form the elements of the SO(3) group i.e. the *Special Orthogonal group in three dimensions*.

An expansion of R close to the identity matrix I can be written as $R = I + \delta R$, where

$$(I + \delta R)^{\dagger}(I + \delta R) = I$$

Expanding this expression and keeping terms out to first order one gets:

$$\delta R^{\dagger} = -\delta R$$

so that δR is an anti-symmetric 3×3 matrix that can be parameterised as

$$\delta R = \begin{pmatrix} 0 & \varepsilon_3 & -\varepsilon_2 \\ -\varepsilon_3 & 0 & \varepsilon_1 \\ \varepsilon_2 & -\varepsilon_1 & 0 \end{pmatrix}$$

An infinitesimal rotation, that transforms x to $x^{'}$ can be written as

$$x^{'} = x - \delta R \times x$$

from where one gets $dx_1 = -\varepsilon_2 x_3 + \varepsilon_3 x_2$, $dx_2 = -\varepsilon_3 x_1 + \varepsilon_1 x_3$ and $dx_3 = -epsilon_1 x_2 + \varepsilon_2 x_1$.

Since the transformation in the neighbourhood of the identity is then given by

$$dx_i = \sum_{\nu=1}^{r} \frac{\partial f_i}{\partial \alpha_\nu} d\alpha_\nu,$$

it is easy to see that $d\alpha_1 \equiv \varepsilon_1$, $d\alpha_2 \equiv \varepsilon_2$ and $d\alpha_3 \equiv \varepsilon_3$. It follows that

$$\frac{\partial f_1}{\partial \alpha_1} = 0$$

$$\frac{\partial f_1}{\partial \alpha_2} = -x_3$$

$$\frac{\partial f_1}{\partial \alpha_3} = x_2$$

$$\frac{\partial f_2}{\partial \alpha_1} = -x_3$$

$$\frac{\partial f_2}{\partial \alpha_2} = 0$$

$$\frac{\partial f_2}{\partial \alpha_3} = -x_1$$

$$\frac{\partial f_3}{\partial \alpha_1} = -x_2$$

$$\frac{\partial f_3}{\partial \alpha_2} = x_1$$

$$\frac{\partial f_3}{\partial \alpha_3} = 0$$

The generators, as we have seen before, are generally given by $\hat{X}_\nu \equiv i\sum_{i=1}^{N} \frac{\partial f_i}{\partial \alpha_\nu} \frac{\partial}{\partial x_i}$ (see Eq. 1.4.2). This gives the generators of the SO(3) group:

$$\hat{X}_1 = ix_3 \frac{\partial}{\partial x_2} - ix_2 \frac{\partial}{\partial x_3}$$

$$\hat{X}_2 = ix_1 \frac{\partial}{\partial x_3} - ix_3 \frac{\partial}{\partial x_1}$$

$$\hat{X}_3 = ix_2 \frac{\partial}{\partial x_1} - ix_1 \frac{\partial}{\partial x_2}$$

These generators are easily recognised as the quantum mechanical angular momentum operators:

$$\hat{X} = x \times (-i)\nabla$$

which satisfies the SO(3) algebra $[\hat{X}_i, \hat{X}_j] = i\varepsilon_{ijk}\hat{X}_k$.

1.6 Conservation of charge

All particles have electric charge in multiples of the charge of quarks i.e. $N(\frac{1}{3})$. In an interaction of the form $A + B \to C + D + E$ the incoming charge from A and B is compensated by the outgoing charge of C, D and E. This is one of the most basic conservation laws in nature, respected by all interactions in the Standard Model and is respected for every interaction: the electromagnetic, the strong and the weak. As we have seen, each conservation law is connected to a symmetry principle. Let's try to review the transformation and the corresponding symmetry that emerges for the case of the electric charge. In what comes we are going to plant a seed. We are going to introduce the notion of gauge transformations[2] that you were all familiar with from classical mechanics, but you were not aware of! It turns out that all interactions described in the Standard Model are produced via the exploitation of such gauge transformations. It is not that important if you do not catch the full essence behind this notion at this stage. The seeds we are going to plant here are going to start getting their meaning once we reach the chapters where the basic interactions of the Standard Model will be discussed in detail.

[2] Gauge transformations are any systematic transformations of the potential of a given field. As an example, the electrostatic potential can be redefined by introducing any additive constant: physics does not change by introducing this constant. In other words no physics law depends on the value of this constant. Results depend only on differences between potentials where the constants cancel out.

1.6.1 Global gauge transformation

Suppose that we have a particle of charge q that is represented by a wave function ψ that satisfies the Schrodinger equation

$$i\hbar\frac{\partial\psi}{\partial t}=\mathscr{H}\psi$$

If $\hat{Q}$ is the quantum mechanical charge operator, then the fact that charge is conserved can be written as $\partial\langle Q\rangle/\partial t=0$ which in turns means that $[\hat{Q},\mathscr{H}]=0$ and ψ can be the eigenfunction of $\hat{Q}$ as well: $\hat{Q}\psi=q\psi$.

Let us now consider the transformation of the form $U=e^{i\varepsilon Q}$ such that $\psi\rightarrow\psi^{'}=U\psi=e^{i\varepsilon Q}\psi$. In this transformation, ε represents a small change in charge and is represented by a real number which is obviously independent of space and time. This type of transformation is called global gauge transformation. Gauge transformation means that the new wave function that emerges should satisfy the same Schrodinger equation as the initial one:

$$i\hbar\frac{\partial\psi^{'}}{\partial t}=\mathscr{H}\psi^{'}\Rightarrow i\hbar\frac{\partial(e^{i\varepsilon Q}\psi)}{\partial t}=\mathscr{H}e^{i\varepsilon Q}\psi$$

But both ε and Q are space and time independent so that the previous can be written as

$$i\hbar e^{i\varepsilon Q}\frac{\partial\psi}{\partial t}=\mathscr{H}e^{i\varepsilon Q}\psi\Rightarrow i\hbar e^{-i\varepsilon Q}e^{i\varepsilon Q}\frac{\partial\psi}{\partial t}=e^{-i\varepsilon Q}\mathscr{H}e^{i\varepsilon Q}\psi\Rightarrow$$

$$i\hbar\frac{\partial\psi}{\partial t}=e^{-i\varepsilon Q}\mathscr{H}e^{i\varepsilon Q}\psi\Rightarrow\mathscr{H}\psi=e^{-i\varepsilon Q}\mathscr{H}e^{i\varepsilon Q}\psi\Rightarrow\mathscr{H}=e^{-i\varepsilon Q}\mathscr{H}e^{i\varepsilon Q}$$

Considering infinitesimally small transformations, such that $\varepsilon Q\leq\leq 1$ one gets:

$$e^{\pm i\varepsilon Q}\approx 1\pm i\varepsilon Q$$

and the previous equation becomes

$$(1-i\varepsilon Q)\mathscr{H}(1+i\varepsilon Q)=\mathscr{H}\Rightarrow\mathscr{H}+i\varepsilon Q\mathscr{H}-i\varepsilon Q\mathscr{H}+\varepsilon^2 Q\mathscr{H}Q=\mathscr{H}\Rightarrow[\mathscr{H},Q]=0$$

That means that invariance under global gauge transformation of the type $U=e^{i\varepsilon Q}$ leads to conservation of charge.

1.6.2 Local gauge transformation

Let us now promote the previous transformation from global to local. That is, this time ε has a dependence on space and time: $\varepsilon\rightarrow\varepsilon(\mathbf{x},t)$. We assume that q is the electric charge and we place out test particle in a static electric field $\mathbf{E}$ defined by a scalar potential A_0 according to

$$\mathbf{E}=-\nabla A_0$$

The Hamiltonian of our new system is written as

$$\mathscr{H}=\mathscr{H}+qA_0,$$

where $H_0 = \frac{p^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2$ is the Hamiltonian of the free particle without the presence of the external field. In general $\mathbf{E}$ and $\mathbf{B}$, the latter being the magnetic field, remain unchanged under a gauge transformation of the form $A_0 \rightarrow A_0^{'}$ and $\mathbf{E} \rightarrow \mathbf{E}^{'}$ according to

$$A_0^{'} = A_0 - \frac{1}{c}\frac{\partial \Lambda(\mathbf{x},t)}{\partial t}$$

and thus

$$\mathbf{E}^{'} = \mathbf{E} + \nabla\Lambda(\mathbf{x},t),$$

with $\Lambda(\mathbf{x},t)$ any arbitrary function of $\mathbf{x}$ and t. Let us now try to identify this arbitrary function. For simplicity and without any loss of generality we assume that ε and Λ depend only on time and not on both space and time. You can work out how things would change (if at all) on your own if the dependence on both space and time is preserved.

Invariance under local gauge transformation[3] requires that

$$i\hbar\frac{\partial \psi}{\partial t} = \mathscr{H}\psi = (\mathscr{H} + qA_0)\psi$$

Similarly for the new wave function that emerges after the transformation:

$$i\hbar\frac{\partial \psi^{'}}{\partial t} = \mathscr{H}\psi^{'} = (\mathscr{H} + qA_0^{'})\psi \Rightarrow i\hbar\frac{\partial}{\partial t}\Big(e^{i\varepsilon(t)Q}\psi\Big) = \Big(\mathscr{H} + qA_0 - \frac{q}{c}\frac{\partial\Lambda(\mathbf{x},t)}{\partial t}\Big)e^{i\varepsilon(t)Q}\psi \Rightarrow$$

$$i\hbar\Big[i\frac{\partial\varepsilon(t)}{\partial t}Q\psi + \frac{\partial\psi}{\partial t}\Big]e^{i\varepsilon(t)Q} = e^{i\varepsilon(t)Q}\Big[-\frac{\hbar^2\nabla^2}{2m} + qA_0 - \frac{q}{c}\frac{\partial\Lambda(\mathbf{x},t)}{\partial t}\Big]\psi \Rightarrow$$

$$e^{i\varepsilon(t)Q}\Big[i\hbar\frac{\partial\psi}{\partial t} - \hbar Q\psi\frac{\partial\varepsilon(t)}{\partial t}\Big] = e^{i\varepsilon(t)Q}\Big[-\frac{\hbar^2\nabla^2}{2m} + qA_0 - \frac{q}{c}\frac{\partial\Lambda(\mathbf{x},t)}{\partial t}\Big]\psi \Rightarrow$$

$$e^{i\varepsilon(t)Q}\Big[-\frac{\hbar^2\nabla^2}{2m} + qA_0 - \hbar Q\frac{\partial\varepsilon(t)}{\partial t}\Big]\psi = e^{i\varepsilon(t)Q}\Big[-\frac{\hbar^2\nabla^2}{2m} + qA_0 - \frac{q}{c}\frac{\partial\Lambda(\mathbf{x},t)}{\partial t}\Big]\psi \Rightarrow$$

$$\hbar Q\frac{\partial\varepsilon(t)}{\partial t} = \frac{q}{c}\frac{\partial\Lambda(\mathbf{x},t)}{\partial t}$$

Since $\varepsilon(t)$ and $\Lambda(t)$ are arbitrary functions, they can be connected via

$$\Lambda(t) = \hbar c\varepsilon(t)$$

with $Q \equiv q$.

1.7 Discrete transformations

1.7.1 Parity transformations

Parity transformations imply a mirroring of a quantum state around the origin. Invariance under parity operation, in turns, implies an invariance under left–right or a symmetry of mirror image and is translated into $[\mathscr{H}, \hat{P}] = 0$.

[3] At this stage it is important to point out that local gauge invariance does not come naturally! It is something we impose by hand.

Parity operation acts as a space inversion mechanism, changing the sign of vectors e.g. $\hat{P}\mathbf{x} \to -\mathbf{x}$, $\hat{P}\mathbf{p} \to -\mathbf{p}$. Axial vectors, however, remain unchanged:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \to (-\mathbf{r}) \times (-\mathbf{p}) = \mathbf{r} \times \mathbf{p} = \mathbf{L}$$

Applying the parity operator on a quantum state $|\psi(\mathbf{x})\rangle$ leads to

$$\hat{P}|\psi(\mathbf{x})\rangle = |\psi(-\mathbf{x})\rangle$$

If we apply the parity operator a second time, we return to the original state:

$$\hat{P}^2|\psi(\mathbf{x})\rangle = \hat{P}\Big(\hat{P}|\psi(\mathbf{x})\rangle\Big) = \hat{P}|\psi(-\mathbf{x})\rangle = \psi(\mathbf{x})$$

That means that the eigenvalues are $P^2 = 1 \Rightarrow P = \pm 1$. We say then that the wavefunctions have either even i.e. $P = +1$ or odd i.e. $P = -1$ parity. Parity is a multiplicative quantum number.

1.7.2 Charge conjugation

Charge conjugation changes a particle to its antiparticle. The quantum mechanical operator is represented by $\hat{C}$. Its application on a quantum state changes the sign of all additive quantum numbers.

For a particle of charge q we have

$$\hat{Q}|q\rangle = q|q\rangle$$

$$\hat{C}|q\rangle = |-q\rangle$$

If we apply the charge conjugation operator to the first equation, we get

$$\hat{C}\hat{Q}|q\rangle = q\hat{C}|q\rangle = q|-q\rangle$$

Applying the charge operator to the second equation, we get

$$\hat{Q}\hat{C}|q\rangle = \hat{Q}|-q\rangle = -q|-q\rangle$$

Subtracting these new two relations, we get

$$(\hat{C}\hat{Q} - \hat{Q}\hat{C})|q\rangle = 2q|-q\rangle = 2\hat{C}\hat{Q}|q\rangle$$

It follows that the operators $\hat{C}$ and $\hat{Q}$ do not commute.

1.7.3 Baryon number

Another quantity that can be transformed in a discrete way is the quantum number that a certain category of composite particles carry. These particles, as we will see in the following chapters, consist of three quarks or antiquarks and are called baryons or antibaryons, respectively. We are very much familiar with baryons: protons (p) and neutrons (n) are the most known members of this family. The quantum number associated with such particles is called the baryon number.

Every baryon (e.g. p, n, Λ, Σ^-, Ξ^-, Ω^-) carry $B = +1$ i.e. posotive baryon number. On the other hand, every antibaryon (e.g. $\overline{\text{p}}$, $\overline{\text{n}}$, $\overline{\Lambda}$, Σ^+, $\overline{\Xi}^+$, $\overline{\Omega}^+$) carry a negative baryon number i.e. $B = -1$.

It is important to note that the conservation of the baryon number of not connected to an exact symmetry i.e. it can not be connected to any symmetry from first principles. Therefore, there are theories which extend the Standard Model (also known as beyond the Standard Model or BSM) that we will see towards the end of these lectures where the baryon number is not conserved e.g. a proton decays into a positron and a photon.

1.7.4 Isospin

When the neutron was found by Chadwick, its mass indicated that it was quite similar to the proton. However, the main difference between those two was that the neutron is neutral while the proton has a positive electric charge. Heisenberg proposed that both particles belong to the same particle, the nucleon, but they are simply different states of it. To describe the two states, a new quantum number, called isospin, was introduced. Based on this classification protons and neutrons were represented according to:

$$|\text{p}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv |\text{up}\rangle$$

and

$$|\text{n}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv |\text{down}\rangle$$

Thus the proton and the neutron form a doublet of isospin $I = 1/2$ state and third component of isospin $I_3 = +1/2$ for the proton and $I_3 = -1/2$ for the neutron:

$$|\text{p}\rangle = |\frac{1}{2}, +\frac{1}{2}\rangle$$

$$|\text{n}\rangle = |\frac{1}{2}, -\frac{1}{2}\rangle$$

The electric charge of a particle is connected to its isospin value via:

$$q = e(I_3 + \frac{1}{2})$$

The strong interactions of the Standard Model are not depending on this “isospin charge“ of the nucleons. That can also be phrased as follows: the strong interactions are invariant under rotations in the isospin space. This symmetry implies that

$$[\mathscr{H}_{\text{strong}}, \mathbf{I}] = 0$$

However, in the presence of an external magnetic field or in electromagnetic interactions $\mathbf{I}$ is not conserved:

$$[\mathscr{H}_{\text{strong}} + \mathscr{H}_{\text{EM}}, \mathbf{I}] \neq 0$$

But in both the electromagnetic and the strong interactions, charge is conserved

$$[\mathscr{H}_{\text{strong}} + \mathscr{H}_{\text{EM}}, Q] = 0 \Rightarrow [\mathscr{H}_{\text{strong}} + \mathscr{H}_{\text{EM}}, e(I_3 + \frac{1}{2})] = 0 \Rightarrow$$

$$[\mathscr{H}_{\rm EM}, I_3] = 0$$

the third component of isospin is conserved even in electromagnetic interactions.

1.7.5 Flavour

The idea of isospin can be extended to quarks. For a system of quarks the Hamiltonian can be written as

$$\mathscr{H} = \mathscr{H} + \mathscr{H}_{Strong} + \mathscr{H}_{E/M}$$

where the term $\mathscr{H}$ represents the kinetic energy of the system. Since the strong interactions are stronger than the electromagnetic ones as we will see in detail in the following chapters and assuming that the masses of the u and d quarks are the same it can be seen that the Hamiltonian written above possesses an up–down symmetry. We can then represent the two quarks by their quantum mechanical states denoted as one column matrices:

$$|u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$|d\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

If the up and down quarks are identical, then the strong interaction would have a flavour independence property which could be expressed by a flavour transformation in this abstract, flavour space according to

$$\begin{pmatrix} u' \\ d' \end{pmatrix} = \hat{U} \begin{pmatrix} u \\ d \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} u \\ d \end{pmatrix}$$

A 2×2 matrix depends on four complex numbers i.e. eight real ones. The matrix needs to be unitary i.e. $UU^{\dagger} = I$, which by itself imposes a constrain in four out of the eight free parameters. Thus, a transformation of this kind can be described by four real numbers or four linearly independent 2×2 matrices representing the generators of this transformation, according to

$$\hat{U} = e^{i\alpha_i \hat{G}_i}$$

One such generator can be formed by considering

$$\hat{U} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} e^{i\phi}$$

But this is a typical U(1) transformation as we will see soon, which is not particularly relevant for transformations in flavour space. The three remaining unitary matrices for a special group which is the SU(2) one with $det(U) = 1$:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where σ_1, σ_2 and σ_3 are the Pauli matrices.

Quark flavour is conserved in the strong interactions but it can be violated in the weak sector of the Standard Model.