Mathematics behind the Standard Model

Preview version

Zibo Wang

Mathematics Department
Rose-Hulman Institute of Technology
May 2019

Contents

I would like to thank my supervisor, David L. Finn, for his guidance through each stage of the process.

1 Introduction

The Standard Model of particle physics is one of the most successful theory describing three of the four known fundamental forces (the electromagnetic, weak, and strong interactions). The Mathematical formulation of the Standard Model contains the internal symmetries of the unitary product group SU(3)×SU(2)×U(1). This paper will start from group and symmetry, and go through some derivations of several important equations describing free particles and particle interactions.

2 Symmetries and groups

2.1 Groups

Group theory is a branch of mathematics that can be described as classifying symmetries. If an object is invariant under a set of transformations, this object has the symmetry under those transformation. A group can then be defined as a collection of transformations which leave the object invariant.

A group is a set G, together with a binary operation $\circ$ defined on G, if they satisfy the following axioms:

1. Closure: For all $g_1, g_2 \in G, g_1 \circ g_2 \in G$.
2. Identity: There exists an identity element $I \in G$ such that for all $g \in G, g = g \circ I = I \circ g$.
3. Inverse: For all $g \in G$, there is a g^{-1} such that $g^{-1} \circ g = g \circ g^{-1} = I$, with I the identity transformation.
4. Associativity: For all $g_1, g_2, g_3 \in G, g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$.

For example, consider rotations of a square: a square is invariant under rotation about its center for angle $n \cdot 90°$, where n is an integer. Denote the rotation by $R(\theta)$, and the set of rotation $G = \{R(n \cdot 90°)\}$.

1. For some integer n_1 and n_2, $R(n_1 \cdot 90°) \circ R(n_2 \cdot 90°) = R((n_1 + n_2) \cdot 90°)$. Since n_1 and n_2 are integers, $(n_1 + n_2)$ is also an integer. Thus $R((n_1 + n_2) \cdot 90°)$ belong to the group G, and the group is closed.
2. The identity transformation I is rotation for $0°$, $R(0°)$.
3. Each transformation has its inverse, i.e. $R(n \cdot 90°) \circ R(-n \cdot 90°) = I$ for any integer n.
4. For some integers n_1, n_2, and n_3:
   $[R(n_1 \cdot 90°) \circ R(n_2 \cdot 90°)] \circ R(n_3 \cdot 90°)$
   $=[R((n_1 + n_2) \cdot 90°)] \circ R(n_3 \cdot 90°)$
   $=R((n_1 + n_2 + n_3) \cdot 90°)$
   $=R(n_1 \cdot 90°) \circ [R((n_2 + n_3) \cdot 90°)]$
   $=R(n_1 \cdot 90°) \circ [R(n_2 \cdot 90°) \circ R(n_3 \cdot 90°)]$
   Therefore, the rotation is associative.

   In this example, there are really only four elements:

$$I = R(0°), R_1 = R(90°), R_1^2 = R(180°),$$
$$R_1^3 = R(270°) = R(-90°), \text{ as } R_1^4 = R(360°) = I.$$

2.2 Lie groups

A special type of Group theory that deals with continuous symmetries is Lie theory. In continuous symmetries, there are infinite amount of elements, which form a manifold and can be locally be described as $\mathbb{R}^n$. Since two elements can be arbitrary close to each other, the idea of "generator" needs to be introduced to describe a group. In terms of generator, an element close to the identity, or an infinitesimal transformation can be expressed as:

$$g(\epsilon) = I + \epsilon X,$$

where ϵ is a small number, I is the identity transformation, and X is a generator. Then a finite transformation, can be considered as doing the infinitesimal transformation many times.

Consider a finite transformation $h(\theta)$, and the infinitesimal transformation is now

$$g\left(\frac{\theta}{N}\right) = I + \frac{\theta}{N}X,$$

where N is some big number, and here I and X can be understand as linear maps, i.e. matrices. Then a finite transformation expressed in terms of the infinitesimal transformation is

$$h(\theta) = g(\epsilon)^N = (I + \frac{\theta}{N}X)^N.$$

Taking the limit as N goes to infinity, we have

$$\begin{aligned} h(\theta) &= \lim_{N\to\infty}(I + \frac{\theta}{N}X)^N \\ h(\theta) &= \mathrm{e}^{\theta X}. \end{aligned} \tag{2.1}$$

This is the matrix exponential given by the power series:

$$\mathrm{e}^{\theta X} = \sum_{k=0}^{\infty} \frac{1}{k!}\theta^k X^k.$$

In equation (2.1), a finite transformation is formed using the generator. This part of the reason how the name “generator” comes. The Taylor expansion of the above transformation is

$$h(\theta) = \sum_n \frac{1}{n!} \left.\frac{d^n h}{d\theta^n}\right|_{\theta=0} \theta^n.$$

Above form can be expressed in a more compact way using expansion of exponential function:

$$h(\theta) = \exp\left(\left.\frac{dh}{d\theta}\right|_{\theta=0} \theta \right).$$

Expressions above together with $h(\theta) = \mathrm{e}^{\theta X}$ implies that

$$X = \left.\frac{dh}{d\theta}\right|_{\theta=0}. \tag{2.2}$$

2.3 SO(3) and SU(2)

Most Lie groups can be represented as matrix groups. Listed below are some of these matrix groups.

Group name	Matrices in group
U(n)	unitary ($U^\dagger U = I$) *
SU(n)	unitary, determinant 1
O(n)	orthogonal ($O^T O = I$)
SO(n)	orthogonal, determinant 1

*: $\dagger$ denotes the conjugate transpose or Hermitian transpose.

2.3.1 SO(3) and rotation matrices

Rotations in three dimensions are usually represented using 3 by 3 matrices with bases:

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix},$$
$$R_y = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix},$$
$$R_z = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

Because $R_i^T R_i = 1$ and $\det(R_i) = 1$, these three bases belong to SO(3) group. In addition, we have

$$\begin{aligned} & R_i^T(\theta)(R_j^T(\phi)(R_k^T(\psi)R_k(\psi))R_j(\phi))R_i(\theta) \\ =& R_i^T(\theta)(R_j^T(\phi)(I)R_j(\phi))R_i(\theta) \\ =& R_i^T(\theta)(R_j^T(\phi)R_j(\phi))R_i(\theta) \\ =& R_i^T(\theta)(I)R_i(\theta) \\ =& R_i^T(\theta)R_i(\theta) \\ =& I, \end{aligned}$$

and

$$\det(R_i R_j R_k) = \det(R_i)\det(R_j)\det(R_k) = 1,$$

so rotation matrices in three dimensions in general belong to SO(3) group.

2.3.2 SU(2) and quaternion

Another way to represent three-dimensional rotations is to use quaternions. Quaternions are the four-dimensional complex numbers that can be construct in analogue to the regular two-dimensional complex numbers. Quaternions have three imaginary components, and their bases are named **i**, **j**, and **k**. These bases satisfy

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \tag{2.4}$$
$$\mathbf{ijk} = -1. \tag{2.5}$$

There are ways to represent quaternions as two by two matrices. One way is to replace the numbers **1**, **i**, **j**, and **k** by following matrices:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\mathbf{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$
$$\mathbf{j} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$
$$\mathbf{k} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{2.6}$$

These matrices still fulfill equation (2.4) and (2.5), using the identity matrix for the number 1. Then an arbitrary quaternion can be expressed as

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = \begin{pmatrix} a+di & -b-ci \\ b-ci & a-di \end{pmatrix},$$

The norm of the quaternion can also be expressed as the determinat of the above matrix

$$\det(q) = a^2 + b^2 + c^2 + d^2.$$

Then, for a unit quaternion,

$$\det(q) = a^2 + b^2 + c^2 + d^2 = 1.$$

In addition, we have

$$q^\dagger q = \begin{pmatrix} a^2 + b^2 + c^2 + d^2 & 0 \\ 0 & a^2 + b^2 + c^2 + d^2 \end{pmatrix} = I.$$

Therefore, any unit quaternion belongs to SU(2). One can also show every element in SU(2) can be represented as a quaternion.

Representing rotations using quaternions is bit complicated. Using a quaternion to represent a vector would be

$$v = (v_x, v_y, v_z) = 0\mathbf{1} + v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k}. \tag{2.7}$$

To specify a rotation, we need to know the rotation axis u and rotation angle θ. Suppose a point p is rotated about axis u for an angle θ, representing using quaternions would be

$$p' = \left(\cos\left(\frac{\theta}{2}\right) + u \sin\left(\frac{\theta}{2}\right)\right) p \left(\cos\left(\frac{\theta}{2}\right) + u \sin\left(\frac{\theta}{2}\right)\right)^{-1}. \tag{2.8}$$

Notice, the rotation axis u need to be a unit vector; otherwise, the rotation would not conserve length/norm. What inside the parentheses are always noted as q, so the above equation can be written in a more concise way:

$$p' = qpq^{-1}. \tag{2.9}$$

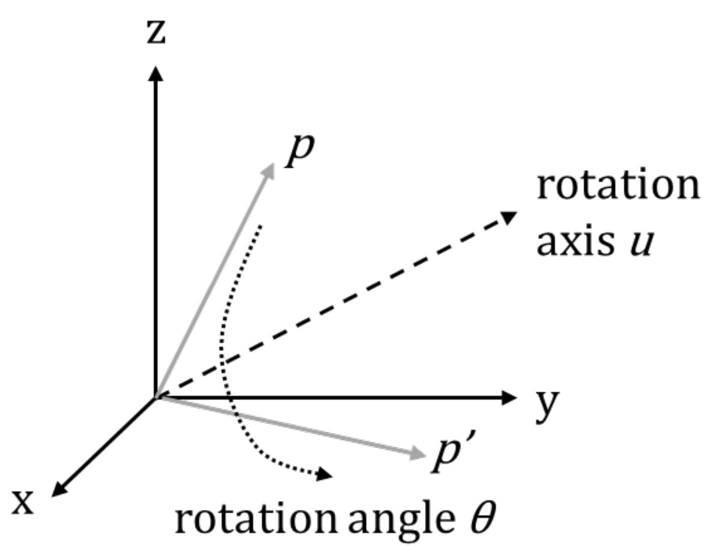

Figure 1: Representing rotation using quaternions

To show equation (2.9) gives the correct rotation, let us break the proof into three steps:

1. Is the length/norm conserved?

First, consider the norm of q:

$$\begin{aligned}|q| = &\cos^2\left(\frac{\theta}{2}\right) + u^2\cos^2\left(\frac{\theta}{2}\right)\\ &(u \text{ is demand to be a unit vector})\\ &= \cos^2\left(\frac{\theta}{2}\right) + \cos^2\left(\frac{\theta}{2}\right)\\ &= 1.\end{aligned}$$

Quaternion multiplication preseves norm, i.e. $|pq| = |p||q|$, and thus

$$\begin{aligned}|qpq^{-1}| &= |q||p||q^{-1}|\\ &= |p|.\end{aligned}$$

So the length/norm is conserved.

2. Is the rotation axis preserved?

   Applying the operation to a vector along the rotation axis u,

$$\begin{aligned}&qkuq^{-1}\\ =&\left[\left(\cos\left(\frac{\theta}{2}\right) + u\ \sin\left(\frac{\theta}{2}\right)\right) u \left(\cos\left(\frac{\theta}{2}\right) + u\ \sin\left(\frac{\theta}{2}\right)\right)^{-1}\right]\\ =&\left[\left(\cos\left(\frac{\theta}{2}\right) + u\ \sin\left(\frac{\theta}{2}\right)\right) u \left(\cos\left(\frac{\theta}{2}\right) + u^{-1}\sin\left(\frac{\theta}{2}\right)\right)\right]\\ =&\left[\left(\cos\left(\frac{\theta}{2}\right) + u\ \sin\left(\frac{\theta}{2}\right)\right)\left(u\ \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right)\right)\right]\\ =&u\left[(\cos^2\left(\frac{\theta}{2}\right) + (\sin^2\left(\frac{\theta}{2}\right)\right]\\ =&u.\end{aligned}$$

   So the rotation axis is preserved.

3. Is the rotation angle correct?

   We can rewrite q as $q = q_0 + \mathbf{q}$, where q_0 is the real part and $\mathbf{q}$ is the imaginary part. Consider a vector orthogonal to the rotation axis n,

$$\begin{aligned}quq^{-1} =&(q_0^2 - |\mathbf{q}|^2)n + 2(\mathbf{q}\cdot n)\mathbf{q} + 2q_0(\mathbf{q}\times n)\\ &(\text{dot product of orthogonal vectors is } 0)\\ =&(q_0^2 - |\mathbf{q}|^2)n + 2q_0(\mathbf{q}\times n).\end{aligned}$$

   Now let $u = \mathbf{q}/|\mathbf{q}|$ and $n_\perp = u \times n$,

$$\begin{aligned}quq^{-1} =&(q_0^2 - |\mathbf{q}|^2)n + 2q_0|\mathbf{q}|n_\perp\\ =&\left(\cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right)\right)n + \left(2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)\right)n_\perp\\ =&\cos(\theta)n + \sin(\theta)n_\perp,\end{aligned}$$

   which is rotation of n through an angle θ on the common plane of n and $n_\perp$.

Therefore, equation (2.9) is exactly what a three-dimensional rotation supposed to be.

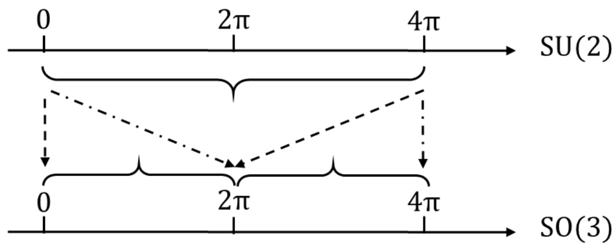

Figure 2: SU(2) double covers SO(3)

2.3.3 Relation between SO(3) and SU(2)

Because of half-angles in the formulas of rotation using quaternion, people sometimes say that quaternions have "720 degrees" of rotation internally. In other words, θ and $\theta' = \theta + 4\pi$ correspond the same three-dimensional rotation. Two unit-quaternions describing the same rotation is the reason why SU(2) is called a double cover of SO(3).

2.3.4 More about SU(2)

Matrices in SU(2) satisfy

$$\begin{aligned} U^\dagger U &= 1, \\ \det(U) &= 1. \end{aligned}$$

As mentioned before, the dagger symbol, †, denotes the conjugate transpose, i.e. $U^\dagger = (U^*)^T$. In terms of generators J, the above two conditions read

$$\begin{aligned} U^\dagger U &= (\mathrm{e}^{iJ})^\dagger \mathrm{e}^{iJ} = 1, \\ \det(U) &= \det(\mathrm{e}^{iJ}) = 1. \end{aligned}$$

An extra "i" is in the exponent which ensures that the final result stays real. Baker-Campbell-Hausdorff formula states that

$$\mathrm{e}^X \mathrm{e}^Y = \mathrm{e}^{X+Y+\frac{1}{2}[X,Y]+\ldots(\text{higher order terms})}, \tag{2.10}$$

Here e^X and e^Y do not commute, so $\mathrm{e}^X\mathrm{e}^Y$ is not $\mathrm{e}^Y\mathrm{e}^X$, as matrix multiplication is not commutative. The term $[X, Y]$ is the commutator of X and Y, $[X, Y] = XY - YX$. From Baker-Campbell-Hausdorff formula,

$$\begin{aligned} (\mathrm{e}^{iJ})^\dagger \mathrm{e}^{iJ} &= 1 \\ \mathrm{e}^{-iJ^\dagger} \mathrm{e}^{iJ} &= 1 \\ \mathrm{e}^{-iJ^\dagger + iJ + [-iJ^\dagger, iJ] + \ldots} &= 1 \\ \rightarrow -iJ^\dagger + iJ = 0 &\text{ and } [-iJ^\dagger, iJ] = 0 \\ \rightarrow J^\dagger &= J. \end{aligned}$$

A matrix fulfilling the condition $J_i^\dagger = J_i$ is called a Hermitian matrix. For the second condition,

$$\begin{aligned} \det(\mathrm{e}^{iJ_i}) = \mathrm{e}^{i\mathrm{tr}(J_i)} &= 1 \\ \mathrm{tr}(J_i) &= 0. \end{aligned}$$

Therefore, the generators of SU(2) must be traceless Hermitian matrices. A basis for traceless Hermitian matrices are Pauli spin matrices:

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.11}$$

These matrices have commutator relation

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l, \tag{2.12}$$

Where ϵ_{ijk} us the Levi-Civita symbol or sometimes refers to permutation tensor,

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } ijk \text{ is an even permutation;} \\ -1, & \text{if } ijk \text{ is an odd permutation;} \\ 0, & \text{if any of } ijk \text{ repeats.} \end{cases} \tag{2.13}$$

In order to get rid of the 2, it is convenient to define the generators as $J_k = \frac{1}{2}\sigma_k$, and the commutation relation becomes

$$[J_j, J_k] = i\epsilon_{jkl}J_l, \tag{2.14}$$

This is the Lie bracket relation for SU(2), which means generators have the above commutation relation forms a SU(2) group.

For Pauli spin matrices, usually we define an additional matrix:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.15}$$

2.4 The Lorentz group SO(1,3)

Before introducing the Lorentz group, we need a little background of four vector. In Euclidean geometry, we have length being invariant under transformations like translation and rotation. The length of a vector can be expressed as

$$x^2 = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

where the matrix in the middle can be treated as the metric for Euclidean space. In Special Relativity, time is considered as an additional dimension, usually referred to as x_0, thus we form a four-vector

$$x_\mu = \begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The Lorentz group is the set of all transformation that preserve the inner product (length/norm of a vector) of four vectors. Minkowski space, is the a combination of three-dimensional Euclidean space and time. Although its initially developed by mathematician Hermann Minkowski for Maxwell's equations of electromagnetism (to explain why the speed of light appears to be constant for all observer), the mathematical structure of Minkowski spacetime was shown to be an immediate consequence of the postulates of special relativity (which also means Maxwell's equations are indeed relativistic).

The metric of Minkowski space is

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.16}$$

With the metric defined above, the inner product of a vector can be expressed as:

$$x^\mu x_\mu = x^\mu \eta_{\mu\nu} x_\mu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2.$$

Suppose there is a Lorentz transformation, Λ, which preserve the inner product:

$$x^\mu \eta_{\mu\nu} x_\mu \to (x^\mu \Lambda^\sigma_\mu)\eta_{\sigma\rho}(\Lambda^\rho_\nu x_\nu) \equiv x^\mu \eta_{\mu\nu} x_\mu,$$

which indicates

$$\Lambda^{\sigma}_{\mu}\eta_{\sigma\rho}\Lambda^{\rho}_{\nu} \equiv \eta_{\mu\nu}, \tag{2.17}$$

or in matrix form

$$\Lambda^T \eta \Lambda \equiv \eta. \tag{2.18}$$

Use the fact of determinant $\det(AB) = \det(A)\det(B)$ to the above conclusion

$$\begin{aligned}
\det(\Lambda^T)\det(\eta)\det(\Lambda) &= \det(\eta) \\
\det(\Lambda^T)(-1)\det(\Lambda) &= -1 \\
\det(\Lambda)^2 &= 1 \\
\det(\Lambda) &= \pm 1.
\end{aligned} \tag{2.19}$$

To preserve the direction of time; in other words, the time after transformation, $x^{0\prime} = t'$, and time before transformation, $x^0 = t$, should have the same sign. This requires $\Lambda^0_0 > 0$, and thus equation (2.17) becomes

$$\det(\Lambda) = 1. \tag{2.20}$$

2.4.1 Generators of the Lorentz Group

If we only consider the spatial part of the Minkowski metric only, it reduce to Euclidean metric. Consider a rotation matrix R in three dimensions, from equation (2.17):

$$R^T R = I_{3\times 3},$$

where $I_{3\times 3}$ is a 3 by 3 identity matrix, which is the condition of $O(3)$. Also, from equation (2.20):

$$\begin{aligned}
\det(\Lambda) &= 1 \\
\det\begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} &= 1 \\
\det(R) &= 1.
\end{aligned}$$

Two conditions together make the conditions for $SO(3)$. Thus, rotation transformations can be expressed in terms of three dimensional rotations as:

$$\Lambda_{rotation\ i} = \begin{pmatrix} 1 & 0 \\ 0 & R_{i(3D)} \end{pmatrix}, \tag{2.21}$$

where three dimensional rotations are expressed as in equation (2.3). The generators of rotation transformations in Minkowski space are

$$\begin{aligned}
J_1 = J_x &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\
J_2 = J_y &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \\
J_3 = J_z &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{2.22}$$

For transformation involving both time and space, we can start with an infinitesimal transformation:

$$\Lambda^{\mu}_{\rho} = \delta^{\mu}_{\rho} + \epsilon K^{\mu}_{\rho},$$

where the Kronecker delta here is defined by $\delta^\mu_\rho = 1$ if $\mu = \rho$ and $\delta^\mu_\rho = 0$ otherwise. If in matrix form, the Kronecker delta is just the identity matrix. By equation (2.17):

$$\begin{aligned}
(\delta^\mu_\rho + \epsilon K^\mu_\rho)\eta_{\mu\nu}(\delta^\nu_\sigma + \epsilon K^\nu_\sigma) &= \eta_{\rho\sigma} \\
\eta_{\rho\sigma} + \epsilon K^\mu_\rho \eta_{\mu\sigma} + \epsilon K^\nu_\sigma \eta_{\rho\nu} + \cancel{\epsilon^2 K^\mu_\rho \eta_{\mu\nu} K^\nu_\sigma} &= \eta_{\rho\sigma} \\
\epsilon K^\mu_\rho \eta_{\mu\sigma} + \epsilon K^\nu_\sigma \eta_{\rho\nu} &= 0.
\end{aligned}$$

Term with ϵ^2 is ignored, since ϵ itself is infinitesimal. The above condition defines another set of generators. A transformation generated by these generators are called boost. A boost can be understand as transform of coordinate system into another coordinate system with constant velocity with respect to the original coordinate system, which is related to phenomena such as time dilation and length contraction. This condition in matrix form is

$$K^T\eta = -\eta K.$$

A set of basis solutions to the above form is:

$$\begin{aligned}
K_1 = K_x &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
K_2 = K_y &= \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
K_3 = K_z &= \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},
\end{aligned} \tag{2.23}$$

which correspond to a boost alone x-axis, y-axis, and z-axis respectively.

To combine the generators of a rotation and a boost together, we define

$$N^\pm_k = \frac{1}{2}(J_k \pm iK_k), \tag{2.24}$$

which has commutation relations:

$$[N^+_j, N^+_k] = i\epsilon_{jkl}N^+_l, \tag{2.25}$$

$$[N^-_j, N^-_k] = i\epsilon_{jkl}N^-_l, \tag{2.26}$$

$$[N^+_j, N^-_k] = 0. \tag{2.27}$$

First two relations are the commutation relations for the Lie algebra of SU(2), which indicates that the Lie algebra of the Lorentz group SO(1,3) consists of two copies of the Lie algebra of SU(2). The last relation shows those two copies of SU(2) commute with each other.

2.5 Representations of the Lorentz group

Representation theory is a branch abstract algebra. A representation makes an abstract algebraic object more concrete by describing its elements by matrices and matrix algebra.

2.5.1 The (0,0) Representation of the Lorentz group

(0,0) means both copies of SU(2) is represented in one dimension. The only 1 by 1 matrices fulfilling the commutation relations of the SU(2) are 0, which means in this case

$$N^+_i = N^-_i = 0,$$

and therefore a transformation would be

$$\mathrm{e}^{N_i^+} = \mathrm{e}^{N_i^-} = 1, \tag{2.28}$$

Since the transformation is an identity matrix, this means that the (0,0) representation of the Lorentz group acts on objects that do not change under Lorentz transformations. In addition, since the transformation is a 1 by 1 matrix, those objects have to be scalars. Therefore, the (0,0) representation is also know as the Lorentz scalar representation.

2.5.2 The (1/2,0) Representation of the Lorentz group

In the $(\frac{1}{2}, 0)$ representation, a two dimensional representation is used for one copy of SU(2), N_i^+; and the one dimensional representation is used for the other copy, N_i^-. From the (0,0) Representation, one dimensional representation for N_i^- is

$$\mathrm{e}^{N_i^-} = 1, \tag{2.29}$$

and by definition,

$$N_i^- = \frac{1}{2}(J_i - iK_i).$$

Therefore in the $(\frac{1}{2}, 0)$ representation,

$$J_i = iK_i.$$

Using two dimensional representation of SU(2), N_i^+ can be expressed as:

$$N_i^+ = \sigma_i/2,$$

where σ_i are Pauli matrices. Also, N_i^+ is defined as

$$N_i^+ = \frac{1}{2}(J_i + iK_i).$$

Since from expression of N_i^-, $J_i = iK_i$, the above form then becomes

$$N_i^+ = iK_i.$$

Comparing two expression for N_i^+ gives

$$K_i = \frac{-i}{2}\sigma_i, \tag{2.30}$$

and using the relation $J_i = iK_i$

$$J_i = \frac{1}{2}\sigma_i. \tag{2.31}$$

Then transformations in the $(\frac{1}{2}, 0)$ representation can be defined using these two generators (R for rotation transformation and B for boost transformation):

$$R_\theta = \mathrm{e}^{i\theta J} = \mathrm{e}^{i\theta\frac{\sigma}{2}}, \tag{2.32}$$

$$B_\phi = \mathrm{e}^{\phi K} = \mathrm{e}^{\phi\frac{\sigma}{2}}, \tag{2.33}$$

Since Pauli matrices are 2 by 2 matrices, transformations in the $(\frac{1}{2}, 0)$ representation are also represented by 2 by 2 matrices, and they should act on two-component objects. These objects are called left-chiral spinors, and can be expressed as:

$$\chi_L = \begin{pmatrix} (\chi_L)_1 \\ (\chi_L)_2 \end{pmatrix}. \tag{2.34}$$

Schwichtenberg (2015) in *Physics from Symmetry*, defined left-chiral spinors as "objects that transform under Lorentz transformation according to the $(\frac{1}{2}, 0)$ representation of the Lorentz group". One important fact is that spinors have properties that usual vectors do not have. For example, a spinor rotated by 2π (360°) is not the same before rotation, because of $\frac{1}{2}$ in the exponent. This property is due to spinors are related to particle with spin 1/2 as we will see in section 4.2, and spins are closely related to symmetries. In *The*

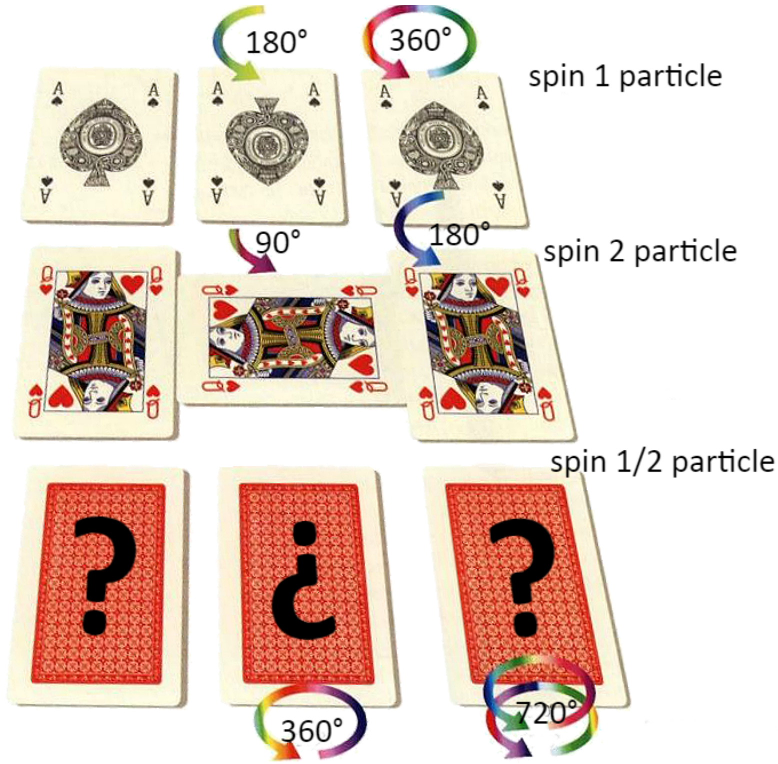

Figure 3: An analogy to spin

Universe in a Nutshell, spins are in an analogue to cards as illustrated in Fig. 1. Spin 1 particles, like the ace of spades, they are invariant under a 360° revolution. Spin 2 particles, like the queen, are invariant under a 180° revolution. And spin $(\frac{1}{2}, 0)$ particles has the bizarre property that require 720° revolution to be invariant.

The $(0, \frac{1}{2})$ representation is similar to the $(\frac{1}{2}, 0)$ representation, but with two dimensional representation for N_i^- and one dimensional representation for N_i^+. With derivation similar to above, the generators in the $(\frac{1}{2}, 0)$ representation can be found:

$$J_i = \frac{1}{2}\sigma_i, \tag{2.35}$$

$$K_i = \frac{-i}{2}\sigma_i, \tag{2.36}$$

and thus transformations are:

$$R_\theta = \mathrm{e}^{i\theta J} = \mathrm{e}^{i\theta\frac{\sigma}{2}}, \tag{2.37}$$

$$B_\phi = \mathrm{e}^{\phi K} = \mathrm{e}^{-\phi\frac{\sigma}{2}}, \tag{2.38}$$

Transformations in the $(0, \frac{1}{2})$ representation are similar to that in the $(\frac{1}{2}, 0)$ representation. It should also act on two-component objects. However, since the boost transformations in two representation are not exactly the same, the objects in two representations should not be exactly the same as well. Objects in the $(0, \frac{1}{2})$ representation are called right-chiral spinors:

$$\chi_R = \begin{pmatrix} (\chi_R)^1 \\ (\chi_R)^2 \end{pmatrix}. \tag{2.39}$$

The generic name for left- and right-chiral spinors is Weyl spinors.

2.5.3 Van der Waerden notation

Right- and left-chiral spinors are similar, so to keep track of which object transforms in what way, it is common to use Van der Waerden Notation. In Van der Waerden Notaion, left- and right- chiral spinors are defined as

$$\chi_L = \chi_a,\ \chi_R = \chi^{\dot{a}}. \tag{2.40}$$

Here the dot notates complex conjugation, i.e. $(x^a)^* = x^{\dot{a}}$. To turn a right-chiral spinor into a left-chiral spinor or vice versa, a "spinor metric" is required, which is defined as

$$\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.41}$$

The "metric" raises and lowers indices similar to the tensor notation

$$\epsilon\chi_L = \epsilon^{ab}\chi_b = \chi^a,$$

but the result is not a right-chiral spinor yet. In order to make the boost transformation be correct (since rotation transformation is same for left/right-chiral spinors, we do not need to worry about that), we need to take complex conjugate in addition to turn the left-chiral spinor to right-chiral, which means

$$\begin{aligned} \epsilon\chi_L^* &= \epsilon^{ab}\chi_b^* \\ &= \chi^{a*} \\ &= \chi^{\dot{a}} \\ &\equiv \chi_R. \end{aligned} \tag{2.42}$$

We can verify that this form transform correctly. Suppose there is a left-chiral spinor after a boost transformation, i.e. $\chi_L' = \mathrm{e}^{\frac{\phi}{2}\sigma}\chi_L$:

$$\begin{aligned} \epsilon(\chi_L')^* &= \epsilon\left(\mathrm{e}^{\frac{\phi}{2}\sigma}\chi_L\right)^* \\ &= \epsilon\left(\mathrm{e}^{\frac{\phi}{2}\sigma*}\chi_L^*\right) \\ &= \epsilon\left(\mathrm{e}^{\frac{\phi}{2}\sigma*}(-\epsilon)(\epsilon)\chi_L^*\right) \\ &(\ (-\epsilon)(\epsilon) = I\ ) \\ &= \left(\epsilon\mathrm{e}^{\frac{\phi}{2}\sigma*}\right)(\epsilon\chi_L^*) \\ &(\ (-\epsilon)\sigma^*(\epsilon) = -\sigma\ ) \\ &= \mathrm{e}^{-\frac{\phi}{2}\sigma*}(\epsilon\chi_L^*)\,, \end{aligned}$$

which is the boost transformation for right-chiral spinor acting on a right-chiral spinor.

From equation (2.42), we can work out how to turn a right-chirla spinor back into a left-chiral spinor by complex conjugate both side

$$\begin{aligned} \chi_R^* &= \epsilon\chi_L \\ \epsilon^{-1}\chi_R^* &= \epsilon^{-1}\epsilon\chi_L \\ \epsilon^{-1}\chi_R^* &= \chi_L. \end{aligned} \tag{2.43}$$

The inversed "metric" in Van der Waerden notation would be

$$\begin{aligned} \epsilon_{ab} = \epsilon^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= -\epsilon. \end{aligned} \tag{2.44}$$

2.5.4 The (1/2,1/2) Representation of the Lorentz group

The $(\frac{1}{2}, \frac{1}{2})$ representation means that both copies of SU(2) are two dimensional. Therefore an object in the $(\frac{1}{2}, \frac{1}{2})$ representation should have four components, and it can be represent by a 2 by 2 complex matrix. Any complex matrix can be written as the sum of a Hermitian matrix and an anti-Hermitian matrix. A Hermitian matrix A satisfies $A = A^\dagger$, and an anti-Hermitian matrix A satisfies $A = -A^\dagger$. The fact that generators of N^- and N^+ in the two dimensional representation are Hermitian restricts objects in the $(\frac{1}{2}, \frac{1}{2})$ representation to be Hermitian matrices. In Van der Waerden notation, suppose an object $v^{\dot{a}b}$ is

$$v^{\dot{a}b} = \begin{pmatrix} c+f & d-ie \\ d+ie & c-f \end{pmatrix},$$

where c, d, e, and f are all real. Notice this can be written as

$$\begin{aligned} v^{\dot{a}b} =& c\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + d\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + e\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + f\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ =& c\sigma_0 + d\sigma_1 + e\sigma_2 + f\sigma_3. \end{aligned}$$

Replace the c, d, e, and f with v^0, v^1, v^2, and v^3, then

$$v^{\dot{a}b} = v^\mu \sigma_m u^{\dot{a}b}.$$

Applying an arbitrary Lorentz transformation to $v^{\dot{a}b}$ gives

$$\begin{aligned} v^{\dot{c}d} &= \left(\mathrm{e}^{\frac{i}{2}\theta\sigma - \frac{1}{2}\phi\sigma}\right)^{\dot{c}}_{\ \dot{a}} v^{\dot{a}b} \left(\mathrm{e}^{\frac{i}{2}\theta\sigma + \frac{1}{2}\phi\sigma}\right)^{d}_{\ b} \\ &= \left(1 + \frac{i}{2}\theta\sigma - \frac{1}{2}\phi\sigma\right)^{\dot{c}}_{\ \dot{a}} v^{\dot{a}b} \left(1 + \frac{i}{2}\theta\sigma + \frac{1}{2}\phi\sigma\right)^{d}_{\ b} \\ &= \begin{pmatrix} 1 + \frac{1}{2}(i\theta_z - \phi_z) & \frac{1}{2}(i\theta_x + \theta_y - \phi_x + i\phi_y) \\ \frac{1}{2}(i\theta_x - \theta_y - \phi_x - i\phi_y) & 1 - \frac{1}{2}(i\theta_z - \phi_z) \end{pmatrix} \\ &\quad \begin{pmatrix} v^0 + v^3 & v^1 - iv^2 \\ v^1 + iv^2 & v^0 - v^3 \end{pmatrix} \\ &\quad \begin{pmatrix} 1 + \frac{1}{2}(i\theta_z + \phi_z) & \frac{1}{2}(i\theta_x + \theta_y + \phi_x - i\phi_y) \\ \frac{1}{2}(i\theta_x - \theta_y + \phi_x + i\phi_y) & 1 - \frac{1}{2}(i\theta_z + \phi_z) \end{pmatrix} \\ &= \begin{pmatrix} v^{0\prime} + v^{3\prime} & v^{1\prime} - iv^{2\prime} \\ v^{1\prime} + iv^{2\prime} & v^{0\prime} - v^{3\prime} \end{pmatrix}. \end{aligned}$$

Focus on each entry of the matrix, and we have

$$\begin{cases} v^{0\prime} + v^{3\prime} = v^0 + i\theta_z v^0 + i\phi_y v^1 + i\theta_x v^1 - i\phi_x v^2 + i\theta_y v^2 + v3 + i\theta_z v^3, \\ v^{1\prime} - iv^{2\prime} = i\theta_x v^0 + \theta_y v^0 + v1 - \phi_z v^1 - iv^2 + i\phi_z v^2 + \phi_x v^3 - i\phi_y v^3, \\ v^{1\prime} + iv^{2\prime} = i\theta_x v^0 - \theta_y v^0 + v1 + \phi_z v^1 + iv^2 + i\phi_z v^2 - \phi_x v^3 - i\phi_y v^3, \\ v^{0\prime} + v^{3\prime} = v^0 - i\theta_z v^0 - i\phi_y v^1 + i\theta_x v^1 + i\phi_x v^2 + i\theta_y v^2 - v3 + i\theta_z v^3. \end{cases}$$

This gives

$$\begin{cases} v^{0\prime} = v^0 + i\theta_x v^1 + i\theta_y v^2 + i\theta_z v^3, \\ v^{1\prime} = i\theta_x v^0 + v^1 + i\phi_z v^2 - i\phi_y v^3, \\ v^{2\prime} = i\theta_y v^0 - i\phi_z v^1 + v^2 + i\phi_x v^3, \\ v^{3\prime} = i\theta_z v^0 + i\phi_y v^1 - i\phi_x v^2 + v^3. \end{cases}$$

Express transformation of v^μ in matrix form

$$\begin{pmatrix} v^{0\prime} \\ v^{1\prime} \\ v^{2\prime} \\ v^{3\prime} \end{pmatrix} = \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} + \begin{pmatrix} 0 & i\theta_x & i\theta_y & i\theta_z \\ i\theta_x & 0 & i\phi_z & -i\phi_y \\ i\theta_y & -i\phi_z & 0 & i\phi_x \\ i\theta_z & i\phi_y & -i\phi_x & 0 \end{pmatrix} \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}, \tag{2.45}$$

which transform in a way same as a four-vector. Therefore, the $(\frac{1}{2}, \frac{1}{2})$ representation is the vector representation.

3 Lagrangian formalism

This section will show briefly how Euler-Lagrange equation can be obtained, which is an essential tool to describe motion of particles/fields. In following two sections (section 4 and 5), Euler-Lagrange equations will be used to help derive equations for particles with different spin. Note particles can be descirbed by fields.

3.1 Principle of least action

Principle of least action can be phrased as:

> "The path taken by the system between times t_1 and t_2 and configurations q_1 and q_2 is the one for which the action is stationary to first order." -R. Penrose (2007). *The Road to Reality.*

The action, S, is defined as the integral of the Lagrangian L, which is defined as the difference between the kinetic and potential energies: $L \equiv T - U$, between two instants of time t_1 and t_2:

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)dt, \tag{3.1}$$

where q is generalized coordinates, $\dot{q}$ is time derivative of coordinates, and t is time. Mathematically, the principle then is:

$$\delta S = 0. \tag{3.2}$$

where the δ notates a small change, i.e. $\delta S = \frac{\partial S}{\partial \epsilon} d\epsilon$ for some perturbation ϵ.

3.2 Euler-Lagrange Equations

In field theories, fields are in general not specified in certain location, so coordinates q are not good descriptions for field. Instead, field itself can be treat as a "coordinate". Therefore a Lagrangian describing a field is

$$\mathscr{L} = \mathscr{L}(\Phi, \partial_\mu \Phi, x). \tag{3.3}$$

Notice, x here is a four-vector including space and time. The action for field is thus

$$S = \int_{x_1}^{x_2} \mathscr{L}(\Phi, \partial_\mu \Phi, x)dx.$$

If the field varies a little bit, i.e. $\Phi \to \Phi + \epsilon$, and $\partial_\mu \Phi \to \partial_\mu \Phi + \partial_\mu \epsilon$. Also, the boundary condition should not change, which requires variation at boundary to be zero, i.e. $\epsilon(x_1) = \epsilon(x_2) = 0$.

As stated in Principle of least action, the action should be stationary to first order (otherwise, a different variation, ϵ, may be picked so that the action is even smaller, which makes the original action not the least). Therefore:

$$S = \int_{x_1}^{x_2} \left[\epsilon \frac{\partial \mathscr{L}}{\partial \Phi} + (\partial_\mu \epsilon) \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} + \text{...(higher order terms)} \right] dx$$

$$\int_{x_1}^{x_2} \left[\epsilon \frac{\partial \mathscr{L}}{\partial \Phi} + (\partial_\mu \epsilon) \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right] dx = 0. \tag{3.4}$$

Integrate the second term by part yields:

$$\int_{x_1}^{x_2} (\partial_\mu \epsilon) \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} dx = \epsilon \frac{\partial \mathscr{L}}{\partial \Phi} \bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} \epsilon \frac{d}{dx} \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right) dx.$$

Since variation at boundary is zero,

$$\epsilon \frac{\partial \mathscr{L}}{\partial \Phi} \bigg|_{x_1}^{x_2} = 0.$$

Now equation 3.4 can be expressed as:

$$\int_{x_1}^{x_2} \epsilon \left[\frac{\partial \mathscr{L}}{\partial \Phi} - \frac{d}{dx} \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)} \right) \right] dx = 0.$$

Since ϵ is arbitrary, terms inside square bracket has to be zero to make the integral always zero. This yields:

$$\frac{\partial \mathscr{L}}{\partial \Phi} - \frac{d}{dx}\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi)}\right) = 0, \tag{3.5}$$

which is the Euler-Lagrange equation.

4 Free theory

This section will show how equations of free particles, which means non-interaction particles/fields can be derived from both physical point of view, and from symmetry.

4.1 Klein-Gordon equation

The Klein-Gordon equation is related to the $(0,0)$ representation of Lorentz group, and it describes particles/fields with spin 0.

4.1.1 Derivation from physics

In non-relativistic quantum mechanics, particles are described by Schrodinger equation. Schrodinger equation can be understand as the energy-momentum relation:

$$\frac{p^2}{2m} + V = E. \tag{4.1}$$

The first term, $p^2/2m$, represents kinetic energy, where p is the momentum and m is the mass. Second term, V, is the potential energy. Sum of kinetic energy and potential energy gives the total energy, E. Replace momentum, p, and total energy, E, by their quantum operators,

$$p \to -i\hbar\nabla \ , \ E \to i\hbar\frac{\partial}{\partial t},$$

and let the resulting form acting on a wave function, Φ:

$$-\frac{\hbar^2}{2m}\nabla^2\Phi + V\Phi = i\hbar\frac{\partial\Phi}{\partial t.} \tag{4.2}$$

This is the Schrodinger equation. The Klein-Gordon equation can be thought as a more generalized version of the Schrodinger equation, which can be applied to relativistic particles/fields. It can be derived in a similar approach. Begin with relativistic energy-momentum relation, but ignoring the potential part, since we are dealing with free particles/fields:

$$p^\mu p_\mu - m^2c^2 = 0, \tag{4.3}$$

where p is a momentum four-vector, and energy is its zeroth component, which means:

$$p_0 = \frac{E}{c} \ , \ p_1 = p_x \ , \ p_2 = p_y \ , \ p_3 = p_z,$$

and the second term is the rest mass energy. As before, substitute in quantum mechanics operators for p and E:

$$p_\mu \to i\hbar\partial_\mu,$$

here

$$\partial_\mu = \frac{\partial}{\partial x^\mu,}$$

where x is also a four-vector:

$$x_0 = t, \ x_1 = -x, \ x_2 = -y, \ x_3 = -z,$$

and then apply the resulting form to a wave function, Φ, to get

$$-\frac{1}{c^2}\partial^\mu(\partial_\mu\Phi) - m^2c^2\Phi = 0, \tag{4.4}$$

which is the Klein-Gordon equation.

4.1.2 Derivation from symmetry

From the symmetry point of view, a spin 0 particle can be treat as a scalar field transform according to the $(0, 0)$ representation of the Lorentz group. To get the correct equation of motion, begin with a general Lagrangian with highest order 2. The idea here is to use lowest possible, non-trivial order term, since order 0 and order 1 terms only introduce constants as we will see, so lowest possible, non-trivial order term would be 2. In addition, all terms should be Lorentz invariant; otherwise, different observers would have different view of a same phenomenon:

$$\mathscr{L} = C_1\Phi^0 + C_2\Phi + C_3\Phi^2 + C_4\partial_\mu\Phi + C_5(\partial_\mu\Phi)(\partial^\mu\Phi) + C_6\Phi(\partial_\mu\Phi).$$

We have not included some term with order less than or equal to 2 in the form above. This is because, when calculate the action, $\mathscr{L}$ is integrated over all space and time, thus term like $\Phi(\partial_\mu(\partial^\mu\Phi))$ would be redundant, since it is equivalent to $(\partial_\mu\Phi)(\partial^\mu\Phi)$:

$$\int \Phi(\partial_\mu(\partial^\mu\Phi))dx = \cancel{\Phi\partial^\mu\Phi|_\infty} + \int (\partial_\mu\Phi)(\partial^\mu\Phi)dx,$$

or the other way round:

$$\int (\partial_\mu\Phi)(\partial^\mu\Phi)dx = \cancel{\Phi\partial^\mu\Phi|_\infty} + \int \Phi(\partial_\mu(\partial^\mu\Phi))dx.$$

A physical interpretation for the boundary term vanishing at infinity could be there is a upper speed limit in physics, and fields infinitely far away cannot have any influence at a finite distance instantaneously.

Since we are trying to find an object transform according to $(0, 0)$ representation of Lorentz group, which is a scalar, and therefore, terms with odd powers in ∂_μ, like ∂_μ which is a vector, are forbidden, so $C_4 = 0$, $C_6 = 0$. The constant term, Φ^0, can also be neglect, $C_1 = 0$. This is because when finding the equation of motion using Euler-Lagrange equation, adding or subtracting a constant to the Lagrangian does not have any influence:

$$\frac{\partial\mathscr{L}}{\partial\Phi} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Phi)}\right) = \frac{\partial(\mathscr{L}+C_1)}{\partial\Phi} - \partial_\mu\left(\frac{\partial(\mathscr{L}+C_1)}{\partial(\partial_\mu\Phi)}\right).$$

The term linear in Φ can be ignored as well, i.e. $C_2 = 0$, since it only adds a constant to the equation of motion:

$$\frac{\partial(\mathscr{L}+C_2\Phi)}{\partial\Phi} - \partial_\mu\left(\frac{\partial(\mathscr{L}+C_2)}{\partial(\partial_\mu\Phi)}\right) = \frac{\partial\mathscr{L}}{\partial\Phi} + C_2 - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Phi)}\right).$$

Therefore, the general Lagrangian now becomes:

$$\mathscr{L} = C_3\Phi^2 + C_5(\partial_\mu\Phi)(\partial^\mu\Phi).$$

A conventional way is to write this Lagrangian as:

$$\mathscr{L} = \frac{1}{2}((\partial_\mu\Phi)(\partial^\mu\Phi) - m^2\Phi^2). \tag{4.5}$$

Applying the Euler-Lagrange equation gives the equation of motion:

$$\frac{\partial\mathscr{L}}{\partial\Phi} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\Phi)}\right) = 0$$

$$\frac{\partial}{\partial\Phi}\left(\frac{1}{2}((\partial_\mu\Phi)(\partial^\mu\Phi) - m^2\Phi^2)\right) - \partial_\mu\left(\frac{\partial}{\partial(\partial_\mu\Phi)}\left(\frac{1}{2}((\partial_\mu\Phi)(\partial^\mu\Phi) - m^2\Phi^2)\right)\right) = 0$$

$$(\partial_\mu\partial^\mu + m^2)\Phi = 0, \tag{4.6}$$

which is the Klein-Gordon equation, and it is same as derived from the physical point of view (in natural units, i.e. $c = 1$).

4.2 Dirac equation

The Dirac equation is the equation of motion for spin $\frac{1}{2}$ particles/fields. It can be derived from energy-momentum relation, or from $(\frac{1}{2},0)\bigoplus(0,\frac{1}{2})$ representation of the Lorentz group.

4.2.1 Derivation from physics

Start with the relativistic energy-momentum relation (equation 4.3), and suppose it can be expressed as:

$$p^\mu p_\mu - m^2c^2 = (\beta^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc),$$

where β^κ and γ^λ are both tensors to be determined. Expand right-hand side of above equation gives:

$$\beta^\kappa\gamma^\lambda p_\kappa p_\lambda - mc(\beta^\kappa - \gamma^\kappa)p_\kappa - m^2c^2.$$

Since in the original energy-momentum relation, there is no linear term in p, this implies that $\beta^\kappa = \gamma^\kappa$, and thus

$$p^\mu p_\mu = \gamma^\kappa\gamma^\lambda p_\kappa p_\lambda.$$

Express the above equation in longhand:

$$\begin{aligned}(p^0)^2-(p^1)^2-(p^2)^2-(p^3)^2 =&(\gamma^0)^2(p^0)^2+(\gamma^1)^2(p^1)^2\\ &+(\gamma^2)^2(p^2)^2+(\gamma^3)^2(p^3)^2\\ &+(\gamma^0\gamma^1+\gamma^1\gamma^0)p_0p_1+(\gamma^0\gamma^2+\gamma^2\gamma^0)p_0p_2\\ &+(\gamma^0\gamma^3+\gamma^3\gamma^0)p_0p_3+(\gamma^1\gamma^2+\gamma^2\gamma^1)p_1p_2\\ &+(\gamma^1\gamma^3+\gamma^3\gamma^1)p_1p_3+(\gamma^2\gamma^3+\gamma^3\gamma^2)p_2p_3.\end{aligned}$$

It is impossible to pick any number for γ^i to make cross terms becomes zero; however, if γ^i are matrices, since matrices do not commute, eliminating cross terms becomes possible.

The standard "Bjorken and Drell convention" for gamma matrices are:

$$\gamma^0 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \gamma^i = \begin{pmatrix}0 & \sigma^i\\ -\sigma^i & 0\end{pmatrix}, \tag{4.7}$$

where σ^i are Pauli matrices, 1 is 2×2 identity matrix, and 0 is 2×2 zero matrix. Now, the energy-momentum relation becomes:

$$(\gamma^\kappa p_\kappa + mc)(\gamma^\kappa p_\kappa - mc) = 0.$$

The equation of motion can be either term, and it is conventional to use the latter, that is:

$$(\gamma^\kappa p_\kappa - mc) = 0.$$

At the end, replace momentum by the momentum operator $p_\mu = i\hbar\partial_\mu$, and apply the result to a wave function Ψ:

$$i\hbar\gamma^\mu\partial_\mu\Psi - mc\Psi = 0. \tag{4.8}$$

This is the Dirac equation. Here Ψ is a four-element column vector, called "bi-spinor" or Dirac spinor. Although it also has four components, it is not a four-vector, and it is not Lorentz invariant using Minkowski metric.

4.2.2 Derivation from symmetry

Using van der Waerden notation introduced in section 2.5.3, we can work out how spinors behave under Lorentz transformations.

$$\chi_L' = \chi_a' = \left(\mathrm{e}^{i\theta\frac{\sigma}{2}+\phi\frac{\sigma}{2}}\right)_a^{\ b}\chi_b, \tag{4.9}$$

$$\chi_L^{*}{}' = \chi_{\dot a}' = \left(\left(\mathrm{e}^{i\theta\frac{\sigma}{2}+\phi\frac{\sigma}{2}}\right)_a^{\ b}\right)^* \chi_b^*$$
$$= \left(\mathrm{e}^{-i\theta\frac{\sigma^*}{2}+\phi\frac{\sigma^*}{2}}\right)_{\dot a}^{\ \dot b}\chi_{\dot b}, \tag{4.10}$$

$$\chi_R' = \chi^{\dot a\prime} = \left(\mathrm{e}^{i\theta\frac{\sigma}{2}-\phi\frac{\sigma}{2}}\right)^{\dot a}_{\ \dot b}\chi^{\dot b}, \tag{4.11}$$

$$\chi_R^{*}{}' = \chi^{a\prime} = \left(\left(\mathrm{e}^{i\theta\frac{\sigma}{2}-\phi\frac{\sigma}{2}}\right)^{\dot a}_{\ \dot b}\right)^* (\chi^{\dot b})^*$$
$$= \left(\mathrm{e}^{-i\theta\frac{\sigma^*}{2}-\phi\frac{\sigma^*}{2}}\right)_a^{\ b}\chi^b. \tag{4.12}$$

Based on four equations above we are able to find terms that are invariant under Lorentz transformation. For example $(\chi^a)^T\chi_a$ is an invariant:

$$\begin{aligned}
(\chi^{a\prime})^T\chi_a' &= \left(\left(\mathrm{e}^{-i\theta\frac{\sigma^*}{2}-\phi\frac{\sigma^*}{2}}\right)_a^{\ b}\chi^b\right)^T \left(\mathrm{e}^{i\theta\frac{\sigma}{2}+\phi\frac{\sigma}{2}}\right)_a^{\ c}\chi_c \\
&= \left(\chi^b\right)^T \left(\mathrm{e}^{-i\theta\frac{(\sigma^*)^T}{2}-\phi\frac{(\sigma^*)^T}{2}}\right)^a_{\ b} \left(\mathrm{e}^{i\theta\frac{\sigma}{2}+\phi\frac{\sigma}{2}}\right)_a^{\ c}\chi_c \\
&= \left(\chi^b\right)^T \left(\mathrm{e}^{-i\theta\frac{\sigma}{2}-\phi\frac{\sigma}{2}}\right)^a_{\ b} \left(\mathrm{e}^{i\theta\frac{\sigma}{2}+\phi\frac{\sigma}{2}}\right)_a^{\ c}\chi_c \\
&= \left(\chi^b\right)^T \delta_b^c\chi_c \\
&= (\chi^c)^T\chi_c.
\end{aligned}$$

$\chi_{\dot a}$ and $\chi^{\dot a}$ can be combined in a similar way to form another invariant $(\chi_{\dot a})^T\chi^{\dot a}$. However, combinations of term with dotted index and term with undotted index are not invariant, since two transformation have different type of indices and there is no way to form a Kronecker delta. For example, $(\chi_{\dot a})^T\chi^a$:

$$\begin{aligned}
(\chi_{\dot a}')^T\chi^{a\prime} =& \chi^{\dot b}\left(\mathrm{e}^{i\theta\frac{\sigma^T}{2}-\phi\frac{\sigma^T}{2}}\right)^{\dot a}_{\ \dot b}\left(\mathrm{e}^{i\theta\frac{\sigma}{2}+\phi\frac{\sigma}{2}}\right)_a^{\ c}\chi_c \\
\neq& \chi^b\delta_b^c\chi_c.
\end{aligned}$$

Also, now it is clear why the "spinor metric" ϵ introduced in section 2.5.3 is called metric, since it makes spinor product invariant:

$$\chi_a^T\chi_a = \chi_a^T(\epsilon^{ab}\chi_b). \tag{4.13}$$

Left-chiral and right-chiral spinors are objects in $(\frac{1}{2},0)$, and $(0,\frac{1}{2})$ representation of Lorentz group respectively. Objects in $(\frac{1}{2},0)\bigoplus(0,\frac{1}{2})$ representation are just combination of Left-chiral and right-chiral spinors, which is the bi-spinor or Dirac spinor. Define the Dirac spinor to be

$$\begin{aligned}
\Psi &= \begin{pmatrix}\chi_L\\ \xi_R\end{pmatrix} && \text{(Different letters are used to avoid ambiguity)} \\
&= \begin{pmatrix}\chi_a\\ \xi^{\dot a}\end{pmatrix}. && \text{(In van der Waerden notation)}
\end{aligned} \tag{4.14}$$

We already found two invariants $\chi_{\dot a}^T\xi^{\dot a}$, and $(\xi^a)^T\chi_a$. Since they are Lorentz invariant, they are part of the Lagrangian. As we did in derivation of Klein-Gordon equation, we also need second order terms that are Lorentz invariant. A second order terms involves a first order derivative. Derivative operator for spinors is:

$$\partial_{a\dot b} = \partial_\nu\sigma^\nu_{a\dot b}, \tag{4.15}$$

where σ in written in matrix form would just be Pauli spin matrices. Two invariants involving first order derivatives are:

$$(\chi_a)^T\partial_\mu(\sigma^\mu)^{a\dot b}\chi_{\dot b}, \tag{4.16}$$
$$(\xi^a)^T\partial_\mu(\sigma^\mu)_{a\dot b}\xi^{\dot b}. \tag{4.17}$$

Here is how we can get $(\sigma^\mu)^{a\dot{b}}$ by raising indices of $(\sigma^\mu)_{a\dot{b}}$ twice using spinor metric:

$$
\begin{aligned}
(\sigma^\mu)^{a\dot{b}} =& ((\sigma^\mu)^T)^{\dot{b}a} \\
=& \epsilon^{\dot{b}\dot{c}}((\sigma^\mu)^T)_{\dot{c}d}\epsilon^{adT} \\
& \text{(in matrix form)} \\
=& \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} ((\sigma^\mu)^T) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^T \\
=& \bar{\sigma}_\mu .
\end{aligned}
$$

In last step of above, the bar of σ stands for adjoint matrix for σ, and for Pauli spin matrices we have:

$$\bar{\sigma}_0 = \sigma_0,\ \bar{\sigma}_i = \sigma_i \text{ for } i = 1, 2, 3. \tag{4.18}$$

Now, if we introduce the matrices

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix},\ \gamma_\mu = \begin{pmatrix} 0 & \bar{\sigma}_\mu \\ \sigma_\mu & 0 \end{pmatrix}, \tag{4.19}$$

which are same as matrices in "Bjorken and Drell convention" introduced in section 4.2.1, and γ_0 can be treat as the metric of bi-spinors or Dirac spinors. We can rewrite four invariants in terms of the Dirac spinor Ψ and the matrix γ as

$$
\begin{aligned}
& \Psi^\dagger\gamma_0\Psi \\
=& \begin{pmatrix} (\chi_L)^\dagger & (\xi_R)^\dagger \end{pmatrix} \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \\
=& (\chi_L)^\dagger\sigma_0\xi_R + (\xi_R)^\dagger\sigma_0\chi_L \\
=& \chi^T_{\dot{a}}\xi^{\dot{a}} + (\xi^a)^T\chi_a ,
\end{aligned} \tag{4.20}
$$

and

$$
\begin{aligned}
& \Psi^\dagger\gamma_0\gamma^\mu\partial_\mu\Psi \\
=& \begin{pmatrix} (\chi_L)^\dagger & (\xi_R)^\dagger \end{pmatrix} \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^m u\partial_\mu \\ \bar{\sigma}^\mu_{\partial\mu} & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \\
=& (\chi_L)^\dagger\bar{\sigma}_0\bar{\sigma}^\mu_{\partial\mu}\chi_L + (\xi_R)^\dagger\sigma_0\sigma^\mu_{\partial\mu}\xi_R \\
=& (\chi_a)^T\partial_\mu(\sigma^\mu)^{a\dot{b}}\chi_{\dot{b}} + (\xi^a)^T\partial_\mu(\sigma^\mu)_{a\dot{b}}\xi^{\dot{b}} .
\end{aligned} \tag{4.21}
$$

If we define the "adjoint spinor" as

$$\bar{\Psi} = (\Psi)^\dagger\gamma_0, \tag{4.22}$$

Invariants can be written in a more compact way

$$\bar{\Psi}\Psi \text{ and } \bar{\Psi}\gamma^\mu\partial_\mu\Psi, \tag{4.23}$$

These two terms are invariant under Lorentz transformation and with order lower than or equal to 2, thus we can write Lagrangian as linear combination of these two terms

$$\mathscr{L} = C_1\bar{\Psi}\Psi + C_2\bar{\Psi}\gamma^\mu\partial_\mu\Psi. \tag{4.24}$$

When $C_1 = -m$, and $C_2 = i$, the Lagrangian is the Dirac-Lagrangian

$$\mathscr{L}_{Dirac} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma^\mu\partial_\mu\Psi. \tag{4.25}$$

Applying the Euler-Lagrange equation gives the Dirac equation (in natural units, i.e. $c = 1$)

$$-m\Psi + i\partial_\mu\gamma^\mu\Psi = 0. \tag{4.26}$$

4.3 Proca equation

The Proca equation is used to describe particles/fields of spin 1, which transform according to $(\frac{1}{2}, \frac{1}{2})$ representation of Lorentz group.

4.3.1 Derivation from physics

Since a photon has spin 1, it would be reasonable to begin with Maxwell equation, which well described behavior of light:

$$\begin{cases} \nabla . E = 4\pi\rho, \\ \nabla \times E + \frac{1}{c}\frac{\partial B}{\partial t} = 0, \\ \nabla . B = 0, \\ \nabla \times B - \frac{1}{c}\frac{\partial E}{\partial t} = \frac{4\pi}{c} J. \end{cases} \tag{4.27}$$

In relativistic electrodynamics, the field strength tensor, $F^{\mu\nu}$, is defined as:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -E_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{4.28}$$

With $F^{\mu\nu}$, the Maxwell equation can be written in a more compact way:

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu. \tag{4.29}$$

Since B has zero divergence, it can be expressed as the curl of a vector potential, A:

$$B = \nabla \times A.$$

Thus,

$$\nabla \times E + \frac{1}{c}\frac{\partial(\nabla \times B)}{\partial t} = \nabla \times \left(E + \frac{1}{c}\frac{\partial A}{\partial t} \right) = 0.$$

Since $(E + (1/c)(\partial A/\partial t))$ has zero curl, it can be expressed as the gradient of a scalar potential, V:

$$E = -\nabla V - \frac{1}{c}\frac{\partial A}{\partial t}.$$

Define the four-vector potential, $A^\mu = (V, A)$. The field strength tensor, $F^{\mu\nu}$, in terms of A^μ then can be expressed as

$$F^\mu \nu = \partial^\mu A^\nu - \partial^\nu A^\mu,$$

and the Maxwell Equation becomes

$$\partial_\mu(\partial^\mu A^\nu) - \partial^\nu(\partial_\mu A^\mu) = \frac{4\pi}{c} J^\nu, \tag{4.30}$$

where J is the current four-vector:

$$J^\mu = (c\rho, J).$$

Again, since we are dealing with free theory, assume vacuum condition, so there is no external charge and current, the Maxwell Equation becomes

$$\partial_\mu(\partial^\mu A^\nu) - \partial^\nu(\partial_\mu A^\mu) = 0, \tag{4.31}$$

which is the Proca equation for a photon, and in fact for any massless particle of spin 1.

For massive particles, a mass term is required and the Proca equation becomes

$$m^2 A^\nu = \partial_\mu(\partial^\mu A^\nu) - \partial^\nu(\partial_\mu A^\mu). \tag{4.32}$$

4.3.2 Derivation from symmetry

A spin 1 particle can be treat as a vector field transform according to $(\frac{1}{2}, \frac{1}{2})$ representation of Lorentz group. Similar to section 4.1.2, start with a general Lagrangian:

$$\mathscr{L} = C_1(\partial^\mu A^\nu)(\partial_\mu A_\nu) + C_2(\partial^\mu A^\nu)(\partial_\nu A_\mu) + C_3 A^\mu A_\mu + C_4 \partial^\mu A_\mu.$$

Also, there are some redundant terms not included, for example, when doing integration over all space and time, $\partial_\mu(\partial^\mu(A^\nu A_\nu))$ is equivalent to $(\partial^\mu A^\nu)(\partial_\mu A_\nu)$.

When applying Euler-Lagrange equation to the Lagrangian above, term $C_4\partial^\mu A_\mu$ only introduce a constant C_4, so this term can be neglected. Apply Euler-Lagrange equation to the remaining terms gives:

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial A_\rho} &= \partial_\sigma \left(\frac{\partial \mathscr{L}}{\partial(\partial_\sigma A_\rho)} \right) \\
2C_3 A^\rho &= \partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma A_\rho)} (C_1(\partial^\mu A^\nu)(\partial_\mu A_\nu) + C_2(\partial^\mu A^\nu)(\partial_\nu A_\mu)) \right) \\
2C_3 A^\rho &= 2C_1 \partial_\sigma(\partial^\sigma A^\rho) + 2C_2 \partial^\rho(\partial_\sigma A^\sigma),
\end{aligned}$$

Change the to conventional constant:

$$m^2 A^\rho = \frac{1}{2}\partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma). \tag{4.33}$$

This is the Proca equation, which is same as derived from Maxwell equation.

5 Interaction theory

The previews section derived equations describing particles/fields with different spin. This section will show how particles/fields interact, using those equations.

5.1 U(1) interaction

In order to work out correct form that describes particle/field interaction, we need to focus on internal symmetries, often referred to as gauge symmetries.

For spin $\frac{1}{2}$ particles/fields, in section 4.2.2 we have worked out the Dirac Lagrangian

$$\mathscr{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi.$$

If a transformation is applied to a spin $\frac{1}{2}$ particle/field according to

$$\Psi \to \Psi' = \mathrm{e}^{ia}\Psi,$$

where a is a real number. The adjoint particle/field would also change

$$\begin{aligned}
\bar{\Psi} \to \bar{\Psi}' &= \Psi'^\dagger\gamma_0 \\
&= (\mathrm{e}^{ia}\Psi)^\dagger\gamma_0 \\
&= \bar{\Psi}\mathrm{e}^{-ia}.
\end{aligned}$$

Since $(\mathrm{e}^{-ia})^\dagger\mathrm{e}^{-ia} = 1$, transformation in this form belongs to the U(1) group. Also, since this transformation has no dependence on coordinate, which means it changes is same over the whole region, this transformation is called a global transformation. It is clear the Lagrangian would not be vary under a global transformation:

$$\begin{aligned}
\mathscr{L}'_{\text{Dirac}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\
&= -m(\bar{\Psi}\mathrm{e}^{-ia})(\mathrm{e}^{ia}\Psi) + i(\bar{\Psi}\mathrm{e}^{-ia})\gamma_\mu\partial^\mu(\mathrm{e}^{ia}\Psi) \\
&= -m\bar{\Psi}\Psi\mathrm{e}^{-ia}\mathrm{e}^{ia} + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi\mathrm{e}^{-ia}\mathrm{e}^{ia} \\
&= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi = \mathscr{L}_{\text{Dirac}}.
\end{aligned}$$

However, a global transformation is not physical, since it implies that a change at a point would immediately result same change over the whole region, and we cannot have signal travel faster than speed of light. Therefore, we need to consider a local transformation, which means the transformation would now have coordinate dependence:

$$\Psi \to \Psi' = \mathrm{e}^{ia(x)}\Psi,$$
$$\bar{\Psi} \to \bar{\Psi}' = \bar{\Psi}\mathrm{e}^{-ia(x)},$$

where the real number a would depend on coordinate. The Lagrangian after a local transformation is

$$\begin{aligned}\mathscr{L}'_{\text{Dirac}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\ &= -m(\bar{\Psi}\mathrm{e}^{-ia(x)})(\mathrm{e}^{ia(x)}\Psi) + i(\bar{\Psi}\mathrm{e}^{-ia(x)})\gamma_\mu\partial^\mu(\mathrm{e}^{ia(x)}\Psi) \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i(\bar{\Psi}\mathrm{e}^{-ia(x)})\gamma_\mu\Psi(\partial^\mu\mathrm{e}^{ia(x)}) \\ &= \mathscr{L}_{\text{Dirac}} - (\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi,\end{aligned}$$

which is no longer invariant. However the Lagrangian should be an invariant. To fix the problem, we can take a look if gauge symmetry of other particles/fields helps.

Based on section 4.3, a spin 1 particle/field has Lagrangian

$$\mathscr{L}_{\text{Proca}} = (\partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\mu A^\nu)(\partial_\nu A_\mu) + m^2 A^\mu A_\mu.$$

We can add constants a_μ to the Lagrangian, since it is a zeroth order term and it has no effect on equation of motion that can be obtained by applying Euler-Lagrange equation to the Lagrangian. Therefore, we can have a transformation defined as

$$A_\mu \to A'_\mu = A_\mu + a_\mu.$$

The Lagrangian after the transformation is

$$\begin{aligned}\mathscr{L}'_{\text{Proca}} =&(\partial^\mu A'^\nu)(\partial_\mu A'_\nu) - (\partial^\mu A'^\nu)(\partial_\nu A'_\mu) + m^2 A'^\mu A'_\mu \\ =&(\partial^\mu(A^\nu + a^\nu))(\partial_\mu(A_\nu + a_\nu)) \\ &- (\partial^\mu(A^\nu + a^\nu))(\partial_\nu(A_\mu + a_\mu)) + m^2(A^\mu + a^\mu)(A_\mu + a_\mu) \\ =&(\partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\mu A^\nu)(\partial_\nu A_\mu) + m^2(A^\mu + a^\mu)(A_\mu + a_\mu) \\ =&= \mathscr{L}_{\text{Proca}} + m^2(A^\mu + a^\mu)(A_\mu + a_\mu).\end{aligned}$$

If the particle/field is massless, its Lagrangian is invariant under such transformation. The Lagrangian for a spin 1 massless particle/field is denoted as Maxwell Lagrangian, $\mathscr{L}_{Maxwell}$. Because the transformation here is coordinate independent, it is a global transformation as well. Now consider a local transformation to a massless spin 1 particle/field

$$A_\mu \to A'_\mu = A_\mu + a_\mu(x),$$

and the transformed Lagrangian becomes

$$\begin{aligned}\mathscr{L}'_{\text{Maxwell}} =&(\partial^\mu A'^\nu)(\partial_\mu A'_\nu) - (\partial^\mu A'^\nu)(\partial_\nu A'_\mu) \\ =&(\partial^\mu(A^\nu + a^\nu(x)))(\partial_\mu(A_\nu + a_\nu(x))) \\ &- (\partial^\mu(A^\nu + a^\nu(x)))(\partial_\nu(A_\mu + a_\mu(x))) \\ =&(\partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\mu A^\nu)(\partial_\nu A_\mu) \\ &+ (\partial^\mu A^\nu)(\partial_\mu a_\nu(x)) + (\partial^\mu a^\nu(x))(\partial_\mu A_\nu) + (\partial^\mu a^\nu(x))(\partial_\mu a_\nu(x)) \\ &- (\partial^\mu A^\nu)(\partial_\nu a_\mu(x)) - (\partial^\mu a^\nu(x))(\partial_\nu A_\mu) - (\partial^\mu a^\nu(x))(\partial_\nu a_\mu(x)) \\ =&\mathscr{L}_{\text{Maxwell}} \\ &+ (\partial^\mu A^\nu)(\partial_\mu a_\nu(x)) + (\partial^\mu a^\nu(x))(\partial_\mu A_\nu) + (\partial^\mu a^\nu(x))(\partial_\mu a_\nu(x)) \\ &- (\partial^\mu A^\nu)(\partial_\nu a_\mu(x)) - (\partial^\mu a^\nu(x))(\partial_\nu A_\mu) - (\partial^\mu a^\nu(x))(\partial_\nu a_\mu(x)),\end{aligned}$$

which is clearly not conserved. However, if we let the local transformation be

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu a(x),$$

here $\partial_\mu a(x)$ is the derivative of some arbitrary function, the Lagrangian is

$$\begin{aligned}
\mathscr{L}'_{\text{Maxwell}} =& (\partial^\mu A'^\nu)(\partial_\mu A'_\nu) - (\partial^\mu A'^\nu)(\partial_\nu A'_\mu) \\
=& (\partial^\mu (A^\nu + \partial^\nu a(x)))(\partial_\mu (A_\nu + \partial_\nu a(x))) \\
& - (\partial^\mu (A^\nu + \partial^\nu a(x)))(\partial_\nu (A_\mu + \partial_\mu a(x))) \\
=& (\partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\mu A^\nu)(\partial_\nu A_\mu) \\
& + (\partial^\mu A^\nu)(\partial_\mu \partial_\nu a(x)) + (\partial^\mu \partial^\nu a(x))(\partial_\mu A_\nu) + (\partial^\mu \partial^\nu a(x))(\partial_\mu \partial_\nu a(x)) \\
& - (\partial^\mu A^\nu)(\partial_\nu \partial_\mu a(x)) - (\partial^\mu \partial^\nu a(x))(\partial_\nu A_\mu) - (\partial^\mu \partial^\nu a(x))(\partial_\nu \partial_\mu a(x)) \\
=& \mathscr{L}_{\text{Maxwell}},
\end{aligned}$$

which is invariant.

Notice, the term $\partial_\mu a(x)$ added to the spin 1 particle/field A_μ is same as the coefficient of the extra term in the Dirac Lagrangian when we apply a local transformation to Ψ and $\bar{\Psi}$. Therefore, if we add a term involving both spin 1 and spin $\frac{1}{2}$ particle/field to the Dirac Lagrangian, we should be able to cancel the extra term introduced in the local transformation. This term is $A_\mu \bar{\Psi}\gamma^\mu \Psi$. Term involving multiple particles/fields are representing the interaction, so denote this term by $\mathscr{L}_{int}$. Under the local transforms that preserve Lagrangian, the interaction Lagrangian becomes

$$\begin{aligned}
\mathscr{L}'_{\text{Int}} =& A'_\mu \bar{\Psi}' \gamma^\mu \Psi' \\
=& (A_\mu + \partial_\mu a(x))(\mathrm{e}^{ia(x)} \bar{\Psi}) \gamma^\mu (\mathrm{e}^{ia(x)} \Psi) \\
=& (A_\mu + \partial_\mu a(x)) \bar{\Psi} \gamma^\mu \Psi \\
=& A_\mu \bar{\Psi} \gamma^\mu \Psi + \partial_\mu a(x) \bar{\Psi} \gamma^\mu \Psi \\
=& \mathscr{L}_{\text{Int}} + \partial_\mu a(x) \bar{\Psi} \gamma^\mu \Psi.
\end{aligned}$$

The second term after transformation is exactly the same as the extra term of the Dirac Lagrangian after transformation but with different sign, so if we add this interaction Lagrangian to the Dirac Lagrangian, the total Lagrangian would be invariant under local transformation. Additionally, since a massless spin 1 particle/field is introduced, we should also include its Lagrangian in the total Lagragian as well. Therefore, the complete Lagrangian with conventional constant is

$$\begin{aligned}
&\mathscr{L}_{\text{Dirac+Int+Maxwell}} = \\
&- m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu \partial^\mu \Psi + gA_\mu \bar{\Psi}\gamma^\mu \Psi + \frac{1}{4}(\partial^\mu A^\nu)(\partial_\mu A_\nu) - \frac{1}{4}(\partial^\mu A^\nu)(\partial_\nu A_\mu).
\end{aligned} \tag{5.1}$$

A scalar g multiplied to the integration term is known as coupling constant, which measures how strong a given integration is. It is convenient to introduce a new operator called covariant derivative, which is defined as

$$D_\mu = i\partial_\mu + gA_\mu, \tag{5.2}$$

and the complete Lagrangian in terms of covariant derivative becomes

$$\begin{aligned}
&\mathscr{L}_{\text{Dirac+Int+Maxwell}} = \\
&- m\bar{\Psi}\Psi + \bar{\Psi}\gamma_\mu D^\mu \Psi + \frac{1}{4}(\partial^\mu A^\nu)(\partial_\mu A_\nu) - \frac{1}{4}(\partial^\mu A^\nu)(\partial_\nu A_\mu).
\end{aligned} \tag{5.3}$$

When spin $\frac{1}{2}$ particles/fields are used to represent electrons and spin 1 particle/field is used to represent a photon, the form is the Lagrangian for the quantum electrodynamics.

5.2 SU(2) interaction

Define a doublet as combination of two spin $\frac{1}{2}$ particles/fields

$$\mathbf{\Psi} = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}, \tag{5.4}$$

$$\bar{\mathbf{\Psi}} = \begin{pmatrix} \bar{\Psi}_1 \bar{\Psi}_2 \end{pmatrix}. \tag{5.5}$$

Based on Lagrangian for a single spin $\frac{1}{2}$ particle/field, the doublet has Lagrangian

$$\mathscr{L}_{2\times\text{Dirac}} = i\bar{\mathbf{\Psi}}\gamma_\mu\partial^\mu\mathbf{\Psi} - \bar{\mathbf{\Psi}}m\mathbf{\Psi}, \tag{5.6}$$

here m is the mass matrix

$$m = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}.$$

Since the object is now "two dimensional" (ignore the fact that a single Dirac spinor is actually has four components, consider a Dirac spinor as a variable), the transformation matrix should be two by two. Consider a global SU(2) transformation

$$\begin{aligned} \mathbf{\Psi} \to \mathbf{\Psi'} =& \mathrm{e}^{ia_i\frac{\sigma}{2}}\mathbf{\Psi}, \\ \bar{\mathbf{\Psi}} \to \bar{\mathbf{\Psi}}' =& \bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i\frac{\sigma}{2}}. \end{aligned}$$

here $\frac{\sigma}{2}$ is the generator of SU(2) as we found in section 2.3.4, with the Pauli spin matrices σ, a_i is arbitrary real constants. Here Einstein notation is used, and the Lagrangian is actually a summation over index i. The mass term under the transformation becomes

$$-\bar{\mathbf{\Psi}}'m\mathbf{\Psi'} = -\bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i\frac{\sigma_i}{2}}m\mathrm{e}^{ia_i\frac{\sigma_i}{2}}\mathbf{\Psi}.$$

The above form would be invariant only if $m_1 = m_2$, so that

$$\begin{aligned} &\mathrm{e}^{-ia_i\frac{\sigma_i}{2}}m\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\mathrm{e}^{ia_i\frac{\sigma_i}{2}} \\ =&m\mathrm{e}^{-ia_i\frac{\sigma_i}{2}}\mathrm{e}^{ia_i\frac{\sigma_i}{2}} \\ =&m. \end{aligned}$$

This restriction causes that the Yang-Mills theory, which deals with local SU(2) transformation, originally has little use, since in reality there is no such pair with exactly same mass. Apparently the mass term is not correct, so at this point let us ignore the mass term first by considering massless doublet only, and we will see later how mass can be restored through different process. For a massless doublet, the Lagrangian is

$$\mathscr{L}_{2\times\text{Dirac}} = i\bar{\mathbf{\Psi}}\gamma_\mu\partial^\mu\mathbf{\Psi}, \tag{5.7}$$

and it is invariant under a global transformation

$$\begin{aligned} \mathscr{L}'_{2\times\text{Dirac}} =& i\bar{\mathbf{\Psi}}'\gamma_\mu\partial^\mu\mathbf{\Psi'} \\ =& i\bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i\frac{\sigma}{2}}\gamma_\mu\partial^\mu\mathrm{e}^{ia_i\frac{\sigma}{2}}\mathbf{\Psi} \\ =& i\bar{\mathbf{\Psi}}\gamma_\mu\partial^\mu\mathbf{\Psi} \\ =& \mathscr{L}_{2\times\text{Dirac}}. \end{aligned}$$

Again, if the transformation is local, and a_i are coordinate dependent, the Lagrangian for the doublet is not invariant

$$\begin{aligned} \mathscr{L}'_{2\times\text{Dirac}} =& i\bar{\mathbf{\Psi}}'\gamma_\mu\partial^\mu\mathbf{\Psi'} \\ =& i\bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\partial^\mu\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\mathbf{\Psi} \\ =& i\bar{\mathbf{\Psi}}\gamma_\mu\partial^\mu\mathbf{\Psi} - \bar{\mathbf{\Psi}}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\mathbf{\Psi} \\ =& \mathscr{L}_{2\times\text{Dirac}} - \bar{\mathbf{\Psi}}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\mathbf{\Psi}. \end{aligned}$$

From U(1) interaction, we expect that spin 1 particles/fields would help to make the Lagrangian local SU(2) invariant. However, in this case, there are three generators, and each generator introduces an extra term in the transformed Lagrangian, thus we need three spin 1 particles/fields W_1^μ, W_2^μ, and W_3^μ to cancel out these extra terms. Define a interaction similar to that in U(1) interaction

$$\mathscr{L}_{\text{Int}} = \bar{\boldsymbol{\Psi}}\gamma_\mu \frac{\sigma_i}{2} W_i^\mu \boldsymbol{\Psi}.$$

Suppose the local transformation for $W_{\mu i}$ is same as that of A in U(1) interaction

$$(W_\mu)_i \to (W_\mu)_i' = (W_\mu)_i + \partial_\mu a_i(x),$$

The total Lagrangian of doublet with interaction under a local transformation is

$$\begin{aligned}
\mathscr{L}'_{2\times\text{Dirac+Int}} =& i\bar{\boldsymbol{\Psi}}'\gamma_\mu\partial^\mu\boldsymbol{\Psi}' + \bar{\boldsymbol{\Psi}}'\gamma_\mu\frac{\sigma_i}{2}W_i^{\mu\prime}\boldsymbol{\Psi}' \\
=& \mathscr{L}_{2\times\text{Dirac}} - \bar{\boldsymbol{\Psi}}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\boldsymbol{\Psi} \\
& + \bar{\boldsymbol{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\frac{\sigma_i}{2}W_j^\mu\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\boldsymbol{\Psi} \\
& + \bar{\boldsymbol{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\frac{\sigma_i}{2}(\partial_\mu a_j(x))\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\boldsymbol{\Psi} \\
& \text{(expend about } a_i = 0\text{, and ignore second or higher order terms)} \\
\approx& \mathscr{L}_{2\times\text{Dirac}} - \bar{\boldsymbol{\Psi}}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\boldsymbol{\Psi} \\
& + \bar{\boldsymbol{\Psi}}\gamma_\mu\left(\frac{\sigma_j}{2} - a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)W_j^\mu\boldsymbol{\Psi} \\
& + \bar{\boldsymbol{\Psi}}\gamma_\mu\left(\frac{\sigma_j}{2} - a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)(\partial_\mu a_j(x))\boldsymbol{\Psi} \\
=& \mathscr{L}_{2\times\text{Dirac}} - \cancel{\bar{\boldsymbol{\Psi}}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)}\boldsymbol{\Psi} \\
& + \bar{\boldsymbol{\Psi}}\gamma_\mu\frac{\sigma_j}{2}W_j^\mu\boldsymbol{\Psi} - \bar{\boldsymbol{\Psi}}\gamma_\mu\left(a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)W_j^\mu\boldsymbol{\Psi} \\
& + \cancel{\bar{\boldsymbol{\Psi}}\gamma_\mu\frac{\sigma_j}{2}(\partial_\mu a_j(x))\boldsymbol{\Psi}} - \bar{\boldsymbol{\Psi}}\gamma_\mu\left(a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)(\partial_\mu a_j(x))\boldsymbol{\Psi} \\
& \text{(last term is second order)} \\
=& \mathscr{L}_{2\times\text{Dirac+Int}} - \bar{\boldsymbol{\Psi}}\gamma_\mu\left(a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)W_j^\mu\boldsymbol{\Psi},
\end{aligned}$$

Here ϵ_{ijk} is Levi-Civita symbol again. Unfortunately, we introduced another extra term. In order to cancel that, the correct form of transformation for $W_{\mu i}$ should be

$$(W_\mu)_i \to (W_\mu)_i' = (W_\mu)_i + \partial^\mu a_i(x) + \epsilon_{ijk}a_j(x)(W_\mu)_k.$$

Now we can have total Lagrangian of doublet with interaction invariant under a local transformation

$$\begin{aligned}
\mathscr{L}'_{2\times\text{Dirac+Int}} =& i\bar{\mathbf{\Psi}}'\gamma_\mu\partial^\mu\mathbf{\Psi}' + \bar{\mathbf{\Psi}}'\gamma_\mu\frac{\sigma_i}{2}W_i^{\mu\prime}\mathbf{\Psi}' \\
=& \mathscr{L}_{2\times\text{Dirac}} - \bar{\mathbf{\Psi}}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\mathbf{\Psi} \\
&+ \bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\frac{\sigma_i}{2}W_j^\mu\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\mathbf{\Psi} \\
&+ \bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\frac{\sigma_i}{2}(\partial_\mu a_j(x))\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\mathbf{\Psi} \\
&+ \bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\frac{\sigma_i}{2}(\epsilon_{ijk}a_j(x)W_k^\mu)\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\mathbf{\Psi} \\
=& \mathscr{L}_{2\times\text{Dirac+Int}} - \bar{\mathbf{\Psi}}\gamma_\mu(a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2})W_j^\mu\mathbf{\Psi} \\
&+ \bar{\mathbf{\Psi}}\mathrm{e}^{-ia_i(x)\frac{\sigma}{2}}\gamma_\mu\frac{\sigma_i}{2}(\epsilon_{ijk}a_j(x)W_k^\mu)\mathrm{e}^{ia_i(x)\frac{\sigma}{2}}\mathbf{\Psi} \\
\approx& \mathscr{L}_{2\times\text{Dirac+Int}} - \cancel{\bar{\mathbf{\Psi}}\gamma_\mu(a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2})W_j^\mu\mathbf{\Psi}} \\
&+ \cancel{\bar{\mathbf{\Psi}}\gamma_\mu(a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2})W_j^\mu}\mathbf{\Psi}.
\end{aligned}$$

Again for complete Lagrangian, we need to include the Lagrangian of three spin 1 particles/fields

$$\begin{aligned}
\mathscr{L}_{3\times\text{Maxwell}} =& (\partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i - \epsilon_{ijk}(W_\mu)_j(W_\nu)_k) \\
&\times(\partial^\mu(W^\nu)_i - \partial^\nu(W^\mu)_i - \epsilon_{ijk}(W^\mu)_j(W^\nu)_k).
\end{aligned}$$

Terms with ϵ_{ijk} ensure Lagrangian for W_i^μ is invariant under the correct transformation. For convenience, define

$$(W_{\mu\nu})_i = (\partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i - \epsilon_{ijk}(W_\mu)_j(W_\nu)_k),$$

so

$$\mathscr{L}_{3\times\text{Maxwell}} = (W_{\mu\nu})_i(W^{\mu\nu})_i.$$

Therefore, the complete Lagrangian invariant under local SU(2) transformation with conventional constant is

$$\mathscr{L}_{2\times\text{Dirac+Int+3}\times\text{Maxwell}} = i\bar{\mathbf{\Psi}}\gamma_\mu\partial^\mu\mathbf{\Psi} + \bar{\mathbf{\Psi}}\gamma_\mu\frac{\sigma_i}{2}W_i^\mu\mathbf{\Psi} - \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i. \tag{5.8}$$

5.3 Mass terms

In SU(2) interaction, we cannot have mass term like $m\bar{\mathbf{\Psi}}\mathbf{\Psi}$ or $m(W_{\mu\nu})_i(W^{\mu\nu})_i$ in order to keep the local SU(2) symmetry. However, particles corresponding to $\mathbf{\Psi}$ or W_μ are usually not massless. For example, the doublet $\mathbf{\Psi}$ can be used to describe electron and electron neutrino pair, which are both massive particles; $W_\pm$ and Z_0 bosons, described by W_μ, are massive as well. Mass of these particles are conventionally interpreted as the spontaneous symmetry breaking of SU(2) and the Higgs mechanism.

5.3.1 Meson mass terms

From previous two sections, we discovered Lagragian invariant under local U(1) interaction and SU(2) interaction:

$$\begin{aligned}
\mathscr{L}_{\mathrm{U}(1)} =& -m\bar{\Psi}\Psi + \bar{\Psi}\gamma_\mu(i\partial^\mu + gB_\mu)\Psi - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\
&\text{(for a doublet, and ignore mass term)} \\
\rightarrow& \bar{\mathbf{\Psi}}\gamma_\mu(i\partial^\mu + gB_\mu)\mathbf{\Psi} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu},
\end{aligned}$$

with

$$B_{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu,$$

(here B_μ is used instead of A_μ. A_μ is commonly used for photon, but we will see in a moment, this spin 1 field here is not a photon), and

$$\begin{aligned}\mathscr{L}_{\mathrm{SU}(2)} =& i\bar{\mathbf{\Psi}}\gamma_\mu\partial^\mu\mathbf{\Psi} + \bar{\mathbf{\Psi}}\gamma_\mu\frac{\sigma_i}{2}W_i^\mu\mathbf{\Psi} - \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i \\ =& i\bar{\mathbf{\Psi}}\gamma_\mu(\partial^\mu + g'\sigma_i W_i^\mu)\mathbf{\Psi} - \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i,\end{aligned}$$

here the constant for interaction term is absorbed into another coupling constant g' that measure the coupling strength between $\mathbf{\Psi}$ and $(W_\mu)_i$, with

$$(W_{\mu\nu})_i = \partial^\mu(W^\nu)_i - \partial^\nu(W^\mu)_i + \epsilon_{ijk}(W_\mu)_j(W_\nu)_k,$$

The spin 1 field B^μ makes the Lagrangian U(1) invariant, so it is often called U(1) gauge field. For similar reason, $(W_\mu)_i$ are often called SU(2) gauge field.

Combining those two Lagrangian gives a Lagrangian that is both locally U(1) and SU(2) invariant

$$\begin{aligned}&\mathscr{L}_{\mathrm{U(1)\ and\ SU(2)}} \\ =&\bar{\mathbf{\Psi}}\gamma_\mu(i\partial^\mu + gB^\mu + g'\sigma_i W_i^\mu)\mathbf{\Psi} - \frac{1}{4}\left((W_{\mu\nu})_i(W^{\mu\nu})_i + B_{\mu\nu}B^{\mu\nu}\right).\end{aligned} \tag{5.9}$$

To preserve SU(2) symmetry, we did not include mass terms. However, we have not used the spin 0 field. Maybe there is a way to include mass terms without spoiling the SU(2) symmetry by include spin 0 field into the Lagrangian. From section 4.1.2, the Lagrangian for a spin 0 particle/field is

$$\begin{aligned}\mathscr{L}_{\mathrm{spin}\ 0} =& \frac{1}{2}((\partial_\mu\Phi)(\partial^\mu\Phi) - m^2\Phi^2) \\ & \text{(for complex field)} \\ =& \frac{1}{2}((\partial_\mu\Phi^*)(\partial^\mu\Phi) - m^2\Phi^*\Phi) \\ & \text{(rename constant)} \\ =& (\partial_\mu\Phi^*)(\partial^\mu\Phi) + \rho^2\Phi^*\Phi.\end{aligned}$$

In section 4.1.2, we restrict the highest order of Φ to be 2. However in fact we can add the next higher power term $(\Phi^\dagger\Phi)^2$, since the only strict constrain is that the Lagrangian of a spin 0 particle/field need to be a scalar so that it is an object in (0,0) representation of Lorentz group. After including the higher order term, the Lagrangian is now

$$\mathscr{L}_{\mathrm{spin}\ 0} = (\partial_\mu\Phi^*)(\partial^\mu\Phi) + \rho^2\Phi^*\Phi - \lambda(\Phi^*\Phi)^2.$$

Before coming up with a Lagrangian that is both locally U(1) and SU(2) invariant for a spin 1 particle/field, in order to match the dimension, we need to form a doublet for the spin 0 particle/field as well,

$$\mathbf{\Phi} = \begin{pmatrix}\Phi_1 \\ \Phi_2\end{pmatrix},$$

and the Lagrangian for a spin 0 doublet reads

$$\mathscr{L}_{\mathrm{spin}\ 0} = (\partial_\mu\mathbf{\Phi}^\dagger)(\partial^\mu\mathbf{\Phi}) + \rho^2\mathbf{\Phi}^\dagger\mathbf{\Phi} - \lambda(\mathbf{\Phi}^\dagger\mathbf{\Phi})^2.$$

In analogues to sections 5.1 and 5.2, the Lagrangian that is both locally U(1) and SU(2) invariant for a spin 1 doublet turned out to be

$$\begin{aligned}&\mathscr{L}_{\mathrm{U(1)\ and\ SU(2)}} \\ =&\left((\partial_\mu - ig'\sigma_i(W_\mu)_i - igB_\mu)\mathbf{\Phi}^\dagger\right)\left((\partial^\mu - ig'\sigma_i(W^\mu)_i - igB^\mu)\mathbf{\Phi}\right) \\ &+ \rho^2\mathbf{\Phi}^\dagger\mathbf{\Phi} - \lambda(\mathbf{\Phi}^\dagger\mathbf{\Phi})^2,\end{aligned} \tag{5.10}$$

with symmetries

$$\begin{cases} B_\mu \to B'_\mu = B_\mu + \partial_\mu a(x), \\ \mathbf{\Phi} \to \mathbf{\Phi}' = \mathrm{e}^{ia(x)\sigma_i}\mathbf{\Phi}; \end{cases}$$

$$\begin{cases} (W_\mu)_i \to (W_\mu)'_i = (W_\mu)_i + \partial^\mu b_i(x) + \epsilon_{ijk} b_j(x)(W_\mu)_k, \\ \mathbf{\Phi} \to \mathbf{\Phi}' = \mathrm{e}^{ib(x)_i\sigma_i}\mathbf{\Phi}; \end{cases}$$

Last two terms of the Lagrangian are often refered as the Higgs potential,

$$\begin{aligned} V(\mathbf{\Phi}) =& -\rho^2\mathbf{\Phi}^\dagger\mathbf{\Phi} + \lambda(\mathbf{\Phi}^\dagger\mathbf{\Phi})^2 \\ =& -\rho^2\Phi_1^*\Phi_1 + \lambda(\Phi_1^*\Phi_1)^2 - \rho^2\Phi_2^*\Phi_2 + \lambda(\Phi_2^*\Phi_2)^2 \\ =& V(\Phi_1) + V(\Phi_2). \end{aligned}$$

Here is the idea of spontaneous symmetry breaking: Spin 0 particles stays at state with lowest potential. In the early universe when the temperature is really high, the potential has minimum at $\Phi = 0$ (the left most plot in Fig. 4). As the temperature decreases, the parameters ρ and λ change. When ρ is large enough, there would be a circle where potential is minimized (two plots on the right side of Fig. 4), and a spin 0 particle would stay somewhere inside the region, and the system is no longer symmetric.

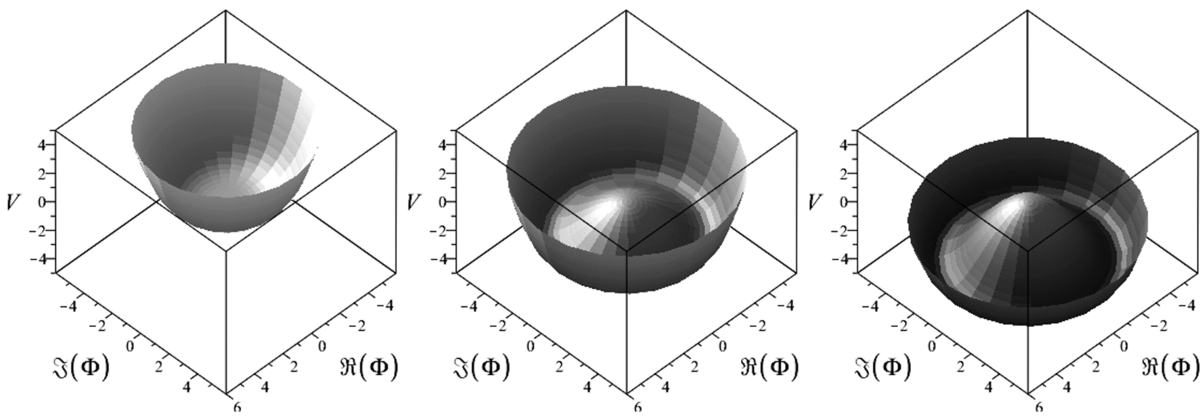

Figure 4: Higgs potential as ρ increases

Consider the Higgs potential for one spin 0 particle/field

$$V(\Phi) = -\rho^2\Phi_1^\dagger\Phi_1 + \lambda(\Phi_1^\dagger\Phi_1)^2.$$

We can find its minima by solving where the derivative of potential is 0,

$$\begin{gathered} \frac{\partial V(\Phi)}{\partial \Phi} = 0 \\ -2\rho^2|\Phi| + 4\lambda|\Phi|^3 = 0 \\ |\Phi|(-2\rho^2 + 4\lambda|\Phi|^2) = 0 \\ \text{(graphically we can see } \Phi = 0 \text{ is a local maximum)} \\ (-2\rho^2 + 4\lambda|\Phi|^2) = 0 \\ |\Phi| = \sqrt{\frac{\rho^2}{2\lambda}} \\ \to \Phi_{\min} = \mathrm{e}^{i\phi}\sqrt{\frac{\rho^2}{2\lambda}}. \end{gathered}$$

For a doublet, following the similar process, we can get

$$(-2\rho^2 + 4\lambda(|\Phi_1|^2 + |\Phi_2|^2)) = 0.$$

One possible minimum exists at

$$\mathbf{\Phi}_{\text{min}} = \begin{pmatrix} 0 \\ \sqrt{\frac{\rho^2}{2\lambda}} \end{pmatrix}.$$

Denote $v = \sqrt{\frac{\rho^2}{2\lambda}}$, so Φ_2 can be expressed as $\frac{v}{\sqrt{2}}$. With this choise of $\mathbf{\Phi}$ the Lagrangian is

$$\begin{aligned}
&\mathscr{L}_{\text{U(1) and SU(2)}} \\
=&\left((\partial_\mu - ig'\sigma_i(W_\mu)_i - igB_\mu)\mathbf{\Phi}^\dagger_{\text{min}}\right)((\partial^\mu - ig'\sigma_i(W^\mu)_i - igB^\mu)\mathbf{\Phi}_{\text{min}}) \\
&+ V(\mathbf{\Phi}_{\text{min}}) \\
=&\frac{v^2}{2}\left|\left((g'\sigma_i(W^\mu)_i + gB^\mu)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)\right|^2 + V(\mathbf{\Phi}_{\text{min}}).
\end{aligned}$$

Ignoring the potential part, in the explicit form of Pauli spin matrices, the Lagrangian is

$$\begin{aligned}
&\frac{v^2}{2}\left|\left((g'\sigma_i(W^\mu)_i + gB^\mu)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)\right|^2 \\
=&\frac{v^2}{2}\left|\begin{pmatrix} g'W_3^\mu + gB^\mu & g'W_1^\mu - ig'W_2^\mu \\ g'W_1^\mu + ig'W_2^\mu & -g'W_3^\mu + gB^\mu \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right|^2 \\
=&\frac{v^2}{2}\left((g')^2((W_1^\mu)^2 + (W_2^\mu)^2) + (g'W_3^\mu - gB^\mu)^2\right).
\end{aligned}$$

Define two new spin 1 fields

$$W_+^\mu = \frac{1}{\sqrt{2}}(W_1^\mu - iW_2^\mu), \tag{5.11}$$

$$W_-^\mu = \frac{1}{\sqrt{2}}(W_1^\mu + iW_2^\mu). \tag{5.12}$$

The first term of the Lagrangian becomes

$$(g'v)^2(W^+)_\mu(W^-)^\mu.$$

The second term in matrix form is

$$\begin{aligned}
&(g'W_3^\mu - gB^\mu)^2 \\
=&\begin{pmatrix} W_3^\mu & B^\mu \end{pmatrix}\begin{pmatrix} g^2 & -gg' \\ -gg' & g^2 \end{pmatrix}\begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix}.
\end{aligned}$$

Diagonalize the matrix by the eigen decomposition

$$\begin{aligned}
&\frac{v^2}{2}\begin{pmatrix} g^2 & -gg' \\ -gg' & g^2 \end{pmatrix} \\
=&\frac{v^2}{2}\frac{1}{\sqrt{g^2 + g'^2}}\begin{pmatrix} g & g' \\ g' & -g \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & (g^2 + g'^2) \end{pmatrix}\frac{1}{\sqrt{g^2 + g'^2}}\begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \\
=&\frac{v^2}{2}PDP^{-1}.
\end{aligned}$$

Define two more new field

$$P^{-1}\begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix} = \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix},$$

and the second term becomes

$$\begin{pmatrix} A^\mu & Z^\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & (g^2+g'^2) \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix}.$$

To sum up:

$$\begin{aligned} &\mathscr{L}_{\text{U(1) and SU(2)}} \\ =&\frac{v^2}{2}\left((g')^2((W_1^\mu)^2+(W_2^\mu)^2)+(g'W_3^\mu-gB^\mu)^2\right)+V(\boldsymbol{\Phi}_{\min}) \\ =&\frac{v^2}{2}\left[g'^2(W^+)_\mu(W^-)_\mu+(g^2+g'^2)Z_\mu^2+0A_\mu^2\right]+V(\boldsymbol{\Phi}_{\min}), \end{aligned}$$

with conventional constant

$$\begin{aligned} &\mathscr{L}_{\text{U(1) and SU(2)}} \\ =&\frac{v^2}{8}g'^2(W^+)_\mu(W^-)_\mu+\frac{v^2}{8}(g^2+g'^2)Z_\mu^2+\frac{v^2}{8}0A_\mu^2+V(\boldsymbol{\Phi}_{\min}), \end{aligned}$$

self-interaction terms is usually interpreted as mass term. For example,

$$\frac{v^2}{8}g'^2(W^+)_\mu(W^-)=\frac{1}{2}m_W^2(W^+)_\mu(W^-)_\mu.$$

Thus, mass of W^+, W^-, and Z bosons is well explained, and photon is massless as expected. Also, the common origin of Z boson and photon suggests that electromagnetism and weak interaction are somehow related, and in fact in early universe when temperature is high enough ($10^{15}K$) those two force merge into a single electroweak force.

5.3.2 Lepton and quark mass terms

Before the discussion about Lepton mass, we need little background about parity violation. In 1956, the Wu experiment discovered the violation of parity. The W^+, W^-, and Z bosons only couple with left-chiral particles/fields. Since left-chiral particles/fields are able to interact with those weak force mediator (W^+, W^-, and Z), one left-chiral particle/field can convert to another left-chiral particle/field throught the weak interaction. Therefore, left-chiral particles/fields are often expressed as doublets $\boldsymbol{\Phi}$. On the other hand, since right-chiral particles/fields are unable to do that, they are often expressed as singlets Φ.

Combination of a left-chiral doublet and a right-chiral singlet directly is not SU(2) invariant. In order to make this interaction term both U(1) and SU(2) invariant, a spin 0 doublet is required,

$$\bar{\boldsymbol{\Psi}}_L\boldsymbol{\Phi}\Psi_R,$$

and

$$\begin{aligned} \bar{\boldsymbol{\Psi}}_L\boldsymbol{\Phi}\Psi_R \to\ & \bar{\boldsymbol{\Psi}}'_L\boldsymbol{\Phi}\Psi'_R \\ &= \bar{\boldsymbol{\Psi}}_L \mathrm{e}^{-ia(x)}\boldsymbol{\Phi}\mathrm{e}^{ia(x)}\Psi_R \\ &= \bar{\boldsymbol{\Psi}}_L\boldsymbol{\Phi}\Psi_R; \end{aligned}$$

$$\begin{aligned} \bar{\boldsymbol{\Psi}}_L\boldsymbol{\Phi}\Psi_R \to\ & \bar{\boldsymbol{\Psi}}'_L\boldsymbol{\Phi}'\Psi_R \\ &= \bar{\boldsymbol{\Psi}}_L \mathrm{e}^{-ib(x)\sigma_i}\mathrm{e}^{ib(x)\sigma_i}\boldsymbol{\Phi}\Psi_R \\ &= \bar{\boldsymbol{\Psi}}_L\boldsymbol{\Phi}\Psi_R. \end{aligned}$$

This type of term is called Yukawa coupling term, and we can use that to form a Lagragian with coupling constant λ_2

$$\mathscr{L}=-\lambda_2(\bar{\boldsymbol{\Psi}}_L\boldsymbol{\Phi}\Psi_R+\bar{\Psi}_L\bar{\boldsymbol{\Phi}}\boldsymbol{\Psi}_R). \tag{5.13}$$

One minimum of the Higgs potential is at

$$
\begin{aligned}
\mathbf{\Phi} &= \begin{pmatrix} 0 \\ \mathrm{e}^{i\phi}\sqrt{\frac{\rho^2}{2\lambda}} \end{pmatrix} \\
&\quad \text{(let } v = \sqrt{\frac{\rho^2}{\lambda}} \text{ as we did before)} \\
&= \begin{pmatrix} 0 \\ \mathrm{e}^{i\phi}\frac{v}{\sqrt{2}} \end{pmatrix} \\
&\quad \text{(expand and keep the first order)} \\
&= \begin{pmatrix} 0 \\ \frac{v+vi\phi}{\sqrt{2}} \end{pmatrix} \\
&\quad \text{(denote } vi\phi\text{=h)} \\
&= \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}.
\end{aligned}
$$

With this choice of $\mathbf{\Phi}$, the Lagrangian becomes

$$
\begin{aligned}
\mathscr{L} =& -\frac{\lambda_2}{\sqrt{2}}\left(\begin{pmatrix} \bar{\mathbf{\Psi}}_{L1} & \bar{\mathbf{\Psi}}_{L2} \end{pmatrix}\begin{pmatrix} 0 \\ v+h \end{pmatrix}\Psi_R + \bar{\Psi}_R\begin{pmatrix} 0 & v+h \end{pmatrix}\begin{pmatrix} \bar{\mathbf{\Psi}}_{L1} \\ \bar{\mathbf{\Psi}}_{L2} \end{pmatrix}\right) \\
=& -\frac{\lambda_2(v+h)}{\sqrt{2}}(\bar{\mathbf{\Psi}}_{L2}\Psi_R + \bar{\Psi}_R\mathbf{\Psi}_{L2}) \\
=& -\frac{\lambda_2(v+h)}{\sqrt{2}}(\bar{\Psi}_L\Psi_R + \bar{\Psi}_R\Psi_L).
\end{aligned}
$$

The inner product of Ψ in tensor notation, with metric for Dirac spinor

$$
\gamma_0 = \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix},
$$

can be rewritten as

$$
\begin{aligned}
&\bar{\Psi}_L\Psi_R + \bar{\Psi}_R\Psi_L \\
\rightarrow &\begin{pmatrix} \bar{\Psi}_L & 0 \end{pmatrix}\begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ \Psi_R \end{pmatrix} + \begin{pmatrix} 0 & \bar{\Psi}_R \end{pmatrix}\begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix}\begin{pmatrix} \Psi_L \\ 0 \end{pmatrix} \\
= &\begin{pmatrix} \bar{\Psi}_L & \bar{\Psi}_R \end{pmatrix}\begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix}\begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} \\
= &\bar{\mathbf{\Psi}}\gamma_0\mathbf{\Psi} \\
\rightarrow &\bar{\mathbf{\Psi}}\mathbf{\Psi}.
\end{aligned}
$$

Therefore, the Lagrangian is equivalent to

$$
\begin{aligned}
\mathscr{L} =& -\frac{\lambda_2(v+h)}{\sqrt{2}}\bar{\mathbf{\Psi}}\mathbf{\Psi} \\
=& -\frac{\lambda_2 v}{\sqrt{2}}(\bar{\mathbf{\Psi}}\mathbf{\Psi}) - \frac{\lambda_2 h}{\sqrt{2}}(\bar{\mathbf{\Psi}}\mathbf{\Psi}).
\end{aligned}
$$

The first term is interpreted as mass term of lepton, and second term is the lepton-Higgs interaction term. In practice, one entry of the doublet can be an electron, muon, or tauon, and the other entry can be the corresponding neutrino of the first entry.

The up and down quarks can transfer into each other through weak interaction as well. Similar for the charm and strange or top and bottom quarks. The process is again violate parity. Thus we can form left-chiral doublets and right chiral singlets. And the rest process to get the quark mass term is just similar to how we get the lepton mass term.

6 Conclusion

Through this paper, although we skipped the SU(3) interaction and the chromodynamics, we can still get a sense of the self-consistency of the Standard Model. In free theory, equations derived using symmetry are essentially the same as if they were derived using conservation of energy and momentum; in the interaction theory, missing mass can also be restored through process like spontaneous symmetry breaking or the Higgs mechanism. Yet, the Standard Model leaves some unexplained phenomena and does not provide the whole picture of the physics. In particular, it is incompatible with general relativity, and it fails to quantize the gravitational integration, i.e. the graviton. Therefore, there are still work to be done to complete the framework.

Reference

[1] D. Griffiths
Introduction to Elementary Particles.
Wiley-VCH, second edition, 2011

[2] J. Schwichtenberg
Physics from Symmetry.
Springer, 2015

[3] M. Robinson
Symmetry and the Standard Model: Mathematics and Particle Physics.
Springer, 2011.

[4] S. Hawking
The Universe in a Nutshell.
Bantam Spectra, 2001

[5] R. Penrose
The Road to Reality: A Complete Guide to the Laws of the Universe.
Alfred A. Knopf, 2005