

Notes on The Theoretical Minimum
— Elementary and Advanced Quantum Mechanics,
Quantum Field Theory and Particle Physics

Gui-Rong Liang

September 9, 2019

Contents

1 Mathematical Foundations and Spin systems	3
1.1 Basic concepts and fundamental principles	3
1.2 Spin states, spin operators, and Pauli Matrices	6
1.3 States and Operators in Continuous basis	8
2 Symmetry Operations of Quantum States and Conserved Quantities	14
2.1 Time evolution of isolated states and the Hamiltonian	14
2.2 General symmetry transformations of states and wave functions	18
2.3 Symmetry operations and Energy degeneracy	20
3 Non-relativistic Quantum Fields	25
3.1 Quantum mechanical harmonic oscillators and ladder operators	25
3.2 Many-harmonic-oscillator system and quantum field operators	27
3.3 Exchange properties for bosons and fermions	31
4 Particle-wave Duality	34
4.1 Property of waves and waves packets	34
4.2 Wave properties of particles and different dispersion relations	36
4.3 The wave equations	38
5 Relativistic Quantum Mechanics and Quantum Field Theory	39
5.1 Dirac equation in one-dimension and the negative energy problem	39
5.2 Dirac equation in three-dimension and chirality	41
5.3 The Dirac Lagrangian and the Covariant Dirac equation	44
6 The Standard Model of Particle Physics	47
6.1 Basics concepts of the Standard Model	47
6.2 U(1) Gauge symmetry in QED	47
6.2.1 Global and local phase transformation symmetry and covariant derivatives	47
6.2.2 Spontaneous Symmetry Breaking and the mass of a field	48
6.2.3 The mass of a fermion field and chirality transition	55
6.3 SU(2) symmetry in Spin multiplets, Isospin, and in weak interaction	58
6.3.1 Group theory of orthogonal and unitary rotations	58
6.3.2 From Spin Multiplets to Isospin	59
6.3.3 The weak interaction mechanism	62
6.4 SU(3) in QCD	65
6.5 The Grand Unification	67

6.5.1 The screening effect and the running coupling constant	67
--	----

1 Mathematical Foundations and Spin systems

1.1 Basic concepts and fundamental principles

1. Quantum states, operators, Hermitian properties

A quantum state is represented by a ket vector that can be decomposed using the orthogonal basis vectors, in this subsection we mainly focus on discrete basis, thus

$$|A\rangle = \sum_j \alpha_j |j\rangle, \quad \text{with } \alpha_j = \langle j|A\rangle, \quad (1)$$

where the coefficient α_j is a complex number, and is also called a wave function. Thus

$$|A\rangle = \sum_j |j\rangle \langle j|A\rangle \implies \sum_j |j\rangle \langle j| = 1. \quad (2)$$

The probability of measuring a state corresponding to a certain basis is given by

$$P_j = \alpha_j^* \alpha_j = \langle A|j\rangle \langle j|A\rangle, \quad (3)$$

then the normalization of probability gives to

$$1 = \sum_j P_j = \sum_j \langle A|j\rangle \langle j|A\rangle = \langle A|A\rangle. \quad (4)$$

So the general principle is the state of a system is represented by a unit vector.

Now we introduce the matrix representation of an operator, which turn a vector to another vector,

$$M|A\rangle = |B\rangle. \quad (5)$$

If we decompose A and B into components on basis vectors, we can represent the above equation as,

$$\sum_j M|j\rangle \alpha_j = \sum_j \beta_j |j\rangle, \quad (6)$$

multiplying the above by a basis bra, we have

$$\sum_j \langle k|M|j\rangle \alpha_j = \sum_j \beta_j \langle k|j\rangle = \beta_k, \quad (7)$$

which is denoted by

$$\sum_j M_{kj} \alpha_j = \beta_k, \quad (8)$$

where $M_{ij} \equiv \langle k|M|j\rangle$ is matrix element of the operator.

We write the above as explicit matrix form (taking 2 dimension as an example) as

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}. \quad (9)$$

Now we take its complex conjugate and make it a equation of row vectors,

$$\begin{pmatrix} \alpha_1^* & \alpha_2^* \end{pmatrix} \begin{pmatrix} M_{11}^* & M_{21}^* \\ M_{12}^* & M_{22}^* \end{pmatrix} = \begin{pmatrix} \beta_1^* & \beta_2^* \end{pmatrix}, \quad (10)$$

which can be represented by the bras notation,

$$\langle A|M^\dagger = \langle B|, \quad (11)$$

where M^\dagger is called the *Hermitian conjugate* of M , with the elements of M^\dagger given by

$$(M^\dagger)_{jk} = M_{kj}^*. \quad (12)$$

An Hermitian operator is an operator which equals to its own Hermitian conjugate,

$$M = M^\dagger. \quad (13)$$

Physical observable quantities are represented by Hermitian operators, due to the fact that the eigenvalues of Hermitian operators are real. And further, eigenvectors of an Hermitian operator corresponding to different eigenvalues are orthogonal. Next we will prove this.

If

$$L|\lambda\rangle = \lambda|\lambda\rangle, \quad (14)$$

where L is Hermitian, we have

$$\lambda^* \langle \lambda | \lambda \rangle = \langle \lambda | L^\dagger | \lambda \rangle = \langle \lambda | L | \lambda \rangle = \lambda \langle \lambda | \lambda \rangle \implies \lambda^* = \lambda, \quad (15)$$

Thus λ is real. Then if

$$\begin{cases} L|\lambda_1\rangle = \lambda_1|\lambda_1\rangle \\ L|\lambda_2\rangle = \lambda_2|\lambda_2\rangle, \end{cases} \quad (16)$$

we have

$$\lambda_1 \langle \lambda_1 | \lambda_2 \rangle = \langle \lambda_1 | L | \lambda_2 \rangle = \lambda_2 \langle \lambda_1 | \lambda_2 \rangle \implies (\lambda_1 - \lambda_2) \langle \lambda_1 | \lambda_2 \rangle = 0 \implies \langle \lambda_1 | \lambda_2 \rangle = 0. \quad (17)$$

Even if eigenvectors of the same eigenvalue are not orthogonal, we can make it orthogonal by Gram-Schmidt procedure. Finally we can make all eigenvectors of an Hermitian operator orthogonal, thus they can be used as a basis for the vector space.

Another inference is that if two operators L, M has the same set of eigenvectors, they commute. We'll use the labels λ_i and μ_α to denote the eigenvalues of L and M , and assume that there is a basis of state-vectors $|\lambda_i, \mu_\alpha\rangle$ that are simultaneous eigenvectors of both observables.

$$\begin{aligned} L|\lambda_i, \mu_\alpha\rangle &= \lambda_i|\lambda_i, \mu_\alpha\rangle \\ M|\lambda_i, \mu_\alpha\rangle &= \mu_\alpha|\lambda_i, \mu_\alpha\rangle, \end{aligned} \quad (18)$$

the by acting on any of the basis vector, which we omit the subscript later, by the two operators,

$$LM|\lambda, \mu\rangle = \lambda\mu|\lambda, \mu\rangle = ML|\lambda, \mu\rangle \implies [L, M]|\lambda, \mu\rangle = 0, \quad (19)$$

since any vector is composed of the basis vectors, $[L, M]$ acting on any vector gives zero,

$$[L, M] = 0, \quad (20)$$

thus they commute. It turns out that the converse of this theorem is also true: if two observables commute, then there is a complete basis of simultaneous eigenvectors of the two observables. Moreover, one may need to specify a larger number of observables to completely label a basis, regardless of the number of observables that are needed, they must all commute, we call this collection *a complete set of commuting observables*.

2. The Uncertainty Principle

Suppose the eigenvalues of the observable A is called a , then the expectation value of A is

$$\langle A \rangle = \sum_a aP(a). \quad (21)$$

Roughly speaking this means that $P(a)$ is centered around the expectation value. What we will mean by "the uncertainty in A is the so-called *standard deviation* \bar{A} , which is computed by subtracting from A its expectation value,

$$\bar{A} = A - \langle A \rangle = A - \langle A \rangle I. \quad (22)$$

The eigenvectors of \bar{A} are the same as those of A and the eigenvalues of \bar{A} is shifted from that of A by the average value $\langle A \rangle$,

$$\bar{A} = a - \langle A \rangle. \quad (23)$$

The square of the uncertainty of A is defined by

$$(\Delta A)^2 := \sum_a \bar{A}^2 P(a) = \sum_a (a - \langle A \rangle)^2 P(a) = \langle \Psi | \bar{A}^2 | \Psi \rangle, \quad (24)$$

if the expectation value of A is zero, then the uncertainty takes the simpler form

$$(\Delta A)^2 = \langle \Psi | A^2 | \Psi \rangle. \quad (25)$$

If \vec{X} and \vec{Y} are two vectors, then the *Triangle Inequality* gives

$$|\vec{X}| + |\vec{Y}| \geq |\vec{X} + \vec{Y}|, \quad (26)$$

by squaring and sorting it we have

$$|\vec{X}| |\vec{Y}| \geq \vec{X} \cdot \vec{Y}, \quad (27)$$

and by squaring it again we have

$$|\vec{X}|^2 |\vec{Y}|^2 \geq |\vec{X} \cdot \vec{Y}|^2, \quad (28)$$

which is called the *Cauchy-Schwarz inequality*.

For complex vector spaces, the magnitude of the vectors are given by

$$\begin{aligned} |X| &= \sqrt{\langle X | X \rangle} \\ |X + Y| &= \sqrt{(\langle X | + \langle Y |)(|X\rangle + |Y\rangle)}, \end{aligned} \quad (29)$$

squaring them and using the Triangle Inequality we have

$$2|X||Y| \geq |\langle X | Y \rangle + \langle Y | X \rangle|. \quad (30)$$

Now we take $|X\rangle$ and $|Y\rangle$ as follows,

$$\begin{aligned} |X\rangle &= A|\Psi\rangle \\ |Y\rangle &= iB|\Psi\rangle, \end{aligned} \quad (31)$$

where A and B are any two observables and $|\Psi\rangle$ is any ket. Now substitute it into the Triangle Inequality for complex vectors, to get

$$2\sqrt{\langle A^2 \rangle \langle B^2 \rangle} \geq |\langle \Psi | AB | \Psi \rangle - \langle \Psi | BA | \Psi \rangle| = |\langle \Psi | [A, B] | \Psi \rangle|. \quad (32)$$

Let's suppose for the moment that A and B have expectation values of zero, so $A = \bar{A}$, and $\langle A^2 \rangle = (\Delta A)^2$, thus we rewrite the above inequality as

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \Psi | [A, B] | \Psi \rangle|. \quad (33)$$

This is the general version of the uncertainty inequality. It says that the product of the uncertainties cannot be smaller than half the magnitude of the expectation value of the commutator. Or to be less quantitative, if the commutator of A and B is not zero, then both observables cannot simultaneously be certain. By the way, if the expectation values of A and B are not zero, we just use a trick to redefine two new operators in which the expectation values have been subtracted off, then the same relation holds as we expected.

Later we will come to a specific version of the inequality — *Heisenberg's Uncertainty Principle*: The product of the uncertainties of the position and momentum of a particle cannot be less than half of Plank's constant.

1.2 Spin states, spin operators, and Pauli Matrices

All possible spin states can be represented in a two dimensional vector space.

$$|A\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle, \quad (34)$$

where the u and d represents for pointing up and down when we measure spin along z-axis, and the orthogonality means if we get up we never get down, vice versa.

Now we choose A to be pointing left and right, considering that they're orthonormal, and it is equally likely to be up and down when measured along z, we write them as

$$\begin{cases} |r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle \\ |l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle. \end{cases} \quad (35)$$

And when we represent forward and backward, we must further take into account the equal likelihood of left and right when measured along x-axis, so we take the following form,

$$\begin{cases} |f\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle \\ |b\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle. \end{cases} \quad (36)$$

Notice that all the above choices have the same freedom by a phase factor $e^{i\theta}$.

We will use σ to denote spin operator. Measuring along z-axis gives us

$$\begin{cases} \sigma_z|u\rangle = |u\rangle \\ \sigma_z|d\rangle = -|d\rangle, \end{cases} \quad (37)$$

if we denote $|u\rangle$ and $|d\rangle$ as unit vectors as $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we may figure out that

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (38)$$

Measuring along x-axis gives us

$$\begin{cases} \sigma_x|r\rangle = |r\rangle \\ \sigma_x|l\rangle = -|l\rangle, \end{cases} \quad (39)$$

if we denote $|r\rangle$ and $|l\rangle$ as unit vectors as $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$, we may figure out that

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (40)$$

Measuring along y-axis gives us

$$\begin{cases} \sigma_y|f\rangle = |f\rangle \\ \sigma_y|b\rangle = -|b\rangle, \end{cases} \quad (41)$$

if we denote $|f\rangle$ and $|b\rangle$ as unit vectors as $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix}$, we may figure out that

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (42)$$

The matrix representations of σ_x , σ_y and σ_z are called the *Pauli Matrices*. It is easy and useful to verify the commutation relations of components of spins,

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l. \quad (43)$$

Now we may treat σ_x , σ_y and σ_z as three components of σ , thus σ can be viewed as a general vector, and it can be projected onto a unit vector \hat{n} point any direction in space, so as to measure spin along that direction.

$$\sigma_n = \vec{\sigma} \cdot \hat{n} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix}, \quad (44)$$

from the fact that

$$\begin{aligned} \text{tr}(\sigma_n) &= 0 \\ \det(\sigma_n) &= -1, \end{aligned} \quad (45)$$

we could immediately figure out that the two eigenvalues of σ_n are give by ± 1 . And we could work out its eigenvectors in a certain configuration, if \hat{n} lies x-z plane, we may write

$$\begin{cases} n_z = \cos \theta \\ n_x = \sin \theta \\ n_y = 0, \end{cases} \quad (46)$$

thus

$$\sigma_n = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}, \quad (47)$$

so the eigenvectors correspond to each eigenvalues are given by

$$\begin{cases} \lambda_1 = 1, \quad |\lambda_1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} \\ \lambda_2 = -1, \quad |\lambda_2\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix}. \end{cases} \quad (48)$$

Now suppose we prepare a $|u\rangle$ state, and we measure the probabilities of $+1$ and -1 along the \hat{n} direction, we get

$$\begin{cases} P(+1) = \langle u | \lambda_1 \rangle \langle \lambda_1 | u \rangle = \cos^2 \frac{\theta}{2} \\ P(-1) = \langle u | \lambda_2 \rangle \langle \lambda_2 | u \rangle = \sin^2 \frac{\theta}{2} \end{cases} \quad (49)$$

Now we're ready to compute the average value of spin along the n direction,

$$\langle \sigma_n \rangle = \sum_j \lambda_j P(\lambda_j) = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} = \cos \theta, \quad (50)$$

which agrees perfectly with our expectation.

To proceed, we come to an important theorem,

The Spin-Polarization Principle: Any state of a single spin is an eigenvector of some component of the spin.

In other words, given any state

$$|A\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle, \quad (51)$$

there exists some direction \hat{n} , such that

$$\vec{\sigma} \cdot \vec{n}|A\rangle = |A\rangle. \quad (52)$$

In physics language, we say that the states of a spin are characterized by a *polarization vector*, and along that polarization vector the component of spin is predictably +1, assuming of course that you know the state-vector. It follows that the expectation value of the spin along the direction \hat{n} can be expressed as

$$\langle \sigma_n \rangle = \langle \vec{\sigma} \cdot \vec{n} \rangle = 1. \quad (53)$$

In fact, any observable of a spin is represented by a 2×2 Hermitian matrix, and has the form

$$\begin{pmatrix} r & w \\ w^* & r' \end{pmatrix}, \quad (54)$$

the implication is that it takes exactly four real parameters to specify this observable. And, there is a neat way to write any spin observable in terms of the Pauli matrices, σ_x , σ_y , σ_z , and the unit matrix I ,

$$S = aI + b_1\sigma_x + b_2\sigma_y + b_3\sigma_z = aI + \vec{b} \cdot \vec{\sigma}, \quad (55)$$

with a and b_i real numbers.

1.3 States and Operators in Continuous basis

1. States and the inner product

A particle at a fixed position x_0 (in one dimension) can be represented by a state vector $|x_0\rangle$. In general, a state is a superposition of various position state, since position is continuous variable, we must turn the sum into integral,

$$|\psi\rangle = \int dx \psi(x)|x\rangle, \quad (56)$$

where the coefficient, or wave function $\psi(x)$ is computed as

$$\psi(x) = \langle x|\psi\rangle, \quad (57)$$

substituting the coefficient into the decomposition we have

$$|\psi\rangle = \int dx |x\rangle \langle x|\psi\rangle, \quad (58)$$

from which we find an identity as

$$I = \int dx |x\rangle \langle x|, \quad (59)$$

which implies summing over all projections $|x\rangle \langle x|$, we can recover the original state.

Left multiplying the decomposition of state by a bra vector $|x'\rangle$, we have

$$\psi(x') = \langle x'|\psi\rangle = \int dx \psi(x) \langle x'|x\rangle, \quad (60)$$

from which we can see the inner product between basis is nothing but the Dirac-delta function,

$$\langle x'|x\rangle = \delta(x - x'), \quad (61)$$

this is the orthonormalization relation of the basis vectors.

Using the above relations we may naturally find the expression of inner product of states in the continuous position basis, just by inserting two different sets of the identity relations of x and x' ,

$$\langle \phi|\psi \rangle = \int dx \int dx' \langle \phi|x'\rangle \langle x'|x\rangle \langle x|\psi \rangle = \int dx \int dx' \phi^*(x') \delta(x' - x) \psi(x) = \int dx \phi^*(x) \psi(x), \quad (62)$$

thus the inner product of a state with itself is given by

$$\langle \psi|\psi \rangle = \int dx \psi^*(x) \psi(x), \quad (63)$$

this will remind us of the expression of the probability, and due to the continuous property we will denote our corresponding concept as the "probability density", which is given by

$$P(x) := \psi^*(x) \psi(x), \quad (64)$$

and the probability of state in an interval $[a, b]$ is given by

$$P([a, b]) = \int_a^b dx P(x) = \int_a^b dx \psi^*(x) \psi(x), \quad (65)$$

and because the total probability equal to 1, our state here is normalized,

$$\langle \psi|\psi \rangle = \int dx \psi^*(x) \psi(x) = \int dx P(x) = 1. \quad (66)$$

2. Linear Operators X and P , their eigenvectors, eigenvalues and relations

An operator is a machine which can turn a given state into a new state, in terms of the continuous basis, if the coefficients (or the wave function) of the new state is specified, the new state would be determined, thus the effect of the operator is made clear.

Let's set an example. We want to define an operator X which can turn a state $|\psi\rangle$ in to a new state $|\bar{\psi}\rangle$, now the state $|\psi\rangle$ is given, so the definition of X is clear only if we know what the new state $|\bar{\psi}\rangle$ is. By "the state $|\psi\rangle$ is given", we mean in terms of any basis, we know exactly what the coefficients of the state $|\psi\rangle$ are, here we choose the continuous position basis, where we know exactly the wave function $\psi(x)$. So the task now shifts from defining X to giving the expression of $\bar{\psi}(x)$, and we define it to be $\bar{\psi}(x) = x\psi(x)$, thus we say we complete the definition of the operator X .

By mathematical expressions, we write the definition of X as

$$X|\psi\rangle = |\bar{\psi}\rangle \quad \text{s.t.} \quad \bar{\psi}(x) = x\psi(x), \quad (67)$$

or in a more neat form,

$$\langle x|X|\psi\rangle := x\langle x|\psi\rangle. \quad (68)$$

Here we say that x is a *basis-dependent operator corresponding to X* . Note that an operator only acts on a state in the abstract vector space, while a basis-dependent operator acts on the wave function. Due to the fact that both the abstract state vector and the wave function can be multiplied by a number, sometimes we blur the distinction of the two concept and just write $X = x$, but this is not generally true. The clear relation between X and x can be seen by inserting an identity relation to the definition, as

$$\int dx' X_{xx'} \psi(x') = \int dx' \langle x|X|x'\rangle \langle x'|\psi\rangle = \langle x|X|\psi\rangle = x\langle x|\psi\rangle = x\psi(x), \quad (69)$$

and by observing we have

$$\begin{cases} X_{xx'} = \langle x|X|x'\rangle = x\delta(x - x') \\ x = \int dx' X_{xx'} = \int dx' \langle x|X|x'\rangle, \end{cases} \quad (70)$$

which clearly shows in position basis X is a matrix with only non-zero diagonalized elements $x\delta(x - x')$, and x is the sum of the elements of the x' 'th column.

Then by bearing the above in mind, we proceed to define a new operator D as

$$\langle x|D|\psi\rangle := \frac{\partial}{\partial x}\langle x|\psi\rangle, \quad (71)$$

where the partial derivative $\frac{\partial}{\partial x}$ is the basis-dependent operator corresponding to D . Notice that now we CANNOT write $D = \frac{\partial}{\partial x}$, because $\frac{\partial}{\partial x}|\psi\rangle$ is meaningless due to the basis-independence of $|\psi\rangle$. The relation between D and $\frac{\partial}{\partial x}$ can be seen by the similar process as in the case of X and x , and here we directly list the results:

$$\begin{cases} D_{xx'} = \langle x|D|x'\rangle = \delta(x - x')\frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} = \int dx' D_{xx'} = \int dx' \langle x|D|x'\rangle. \end{cases} \quad (72)$$

Operators with physical interests are Hermitian operators. A convenient way to judge the hermitian property is by using the inner product. The inner product is number, assuming O is hermitian, we take the complex conjugate of the inner product,

$$\langle \phi|O|\psi\rangle^* = \langle \psi|O^\dagger|\phi\rangle = \langle \psi|O|\phi\rangle, \quad (73)$$

finding that switching the left and right vector gives the complex conjugate of the original, and this will serve as the criteria for judging the hermitian property in the following.

So are X and D hermitian? We first deal with X ,

$$\begin{aligned} \langle \phi|X|\psi\rangle^* &= \left[\int dx \langle \phi|x\rangle \langle x|X|\psi\rangle \right]^* = \left[\int dx \phi^*(x) x\psi(x) \right]^* \\ &= \int dx \psi^*(x) x^* \phi(x) = \int dx \psi^*(x) x\phi(x) \\ &= \langle \psi|X|\phi\rangle, \end{aligned} \quad (74)$$

where we used the fact that the position variable is real. Thus the position operator is indeed Hermitian.

And next the D operator,

$$\begin{aligned} \langle \phi|D|\psi\rangle^* &= \left[\int dx \langle \phi|x\rangle \langle x|D|\psi\rangle \right]^* = \left[\int dx \phi^*(x) \frac{\partial}{\partial x} \psi(x) \right]^* \\ &= \left[\int_{-\infty}^{\infty} dx \psi^*(x) \overbrace{\phi(x)}^{} - \int dx \frac{\partial}{\partial x} \psi^*(x) \phi(x) \right]^* = - \int dx \phi^*(x) \frac{\partial}{\partial x} \psi(x) \\ &= -\langle \psi|D|\phi\rangle \implies D^\dagger = -D, \end{aligned} \quad (75)$$

where we have used the boundary condition that the wave functions vanish at infinity. Thus the D operator is not Hermitian, but anti-Hermitian. But actually we still want to construct an Hermitian operator out of D , recall that the easiest way to construct an Hermitian operator out of an anti-Hermitian operator is multiplying it by i or $-i$. So we define a new operator as

$$P := -i\hbar D, \quad (76)$$

and it's easy to check that P is Hermitian,

$$\langle \phi|P|\psi\rangle^* = [-i\hbar\langle \phi|D|\psi\rangle]^* = i\hbar[-\langle \psi|D|\phi\rangle] = \langle \psi|P|\phi\rangle, \quad (77)$$

and its basis-dependent counterpart is naturally given by $\hat{p} = -i\hbar\frac{\partial}{\partial x}$, which can act directly on the wave function,

$$\hat{p}\psi(x) = -i\hbar\frac{\partial}{\partial x}\psi(x). \quad (78)$$

Next we explore the eigenvalues and eigenvectors of X and P . Eigenvectors of an operator are, by definitions, vectors that acted by the operator gives the same vector multiplied by a number which is called the eigenvalue. For the X operator, we are looking for a solution that satisfies,

$$X|\psi\rangle = x_0|\psi\rangle, \quad (79)$$

where x_0 is a specific number. By definition, the LHS in x -basis is given by

$$\langle x|X|\psi\rangle = x\langle x|\psi\rangle = x\psi(x), \quad (80)$$

and the RHS in the x -basis would be

$$\langle x|x_0|\psi\rangle = x_0\langle x|\psi\rangle = x_0\psi(x). \quad (81)$$

So the equation we're going to solve is nothing but

$$x\psi(x) = x_0\psi(x), \quad (82)$$

and when obtaining the expression for $\psi(x)$, we can easily recover the state vector by the superposition principle, $|\psi\rangle = \int dx \psi(x)|x\rangle$.

By rearranging the equation to solve, we get

$$(x - x_0)\psi(x) = 0, \quad (83)$$

which says that $\psi(x)$ could be non-zero only when $x = x_0$, or equivalently, when $x \neq x_0$, $\psi(x)$ must be zero. We summarize this as follows,

$$\psi(x) = \begin{cases} (\text{non-})\text{zero}, & x = x_0 \\ 0, & x \neq x_0, \end{cases} \quad (84)$$

this reminds us of the Dirac-delta function $\delta(x - x_0)$, but we didn't strictly prove that $\psi(x)$ is indeed $\delta(x - x_0)$. Since a straightforward proof is a little troublesome and will lead to some ambiguities (as I tried), we go around that and try to find another clever way. As all we want is a function that satisfies $x\psi(x) = x_0\psi(x)$, if we somehow happen to find such a function, we can say that it is the eigenfunction of the operator. The trick here is to use the relation between X and x , as follows,

$$x_0\delta(x - x_0) = [x_0\delta(x_0 - x)]^* = \langle x_0|X|x\rangle^* = \langle x|X|x_0\rangle = x\delta(x - x_0), \quad (85)$$

where we have used the Hermitian property of X , the realness of x_0 and $\delta(x_0 - x)$, and the evenness of $\delta(x_0 - x)$. Now we see that the Dirac-delta function $\delta(x - x_0)$ indeed is a solution to the equation, so it is the eigenfunction, $\psi(x) = \delta(x - x_0)$, now we recover our eigenvector as

$$|\psi\rangle = \int dx \psi(x)|x\rangle = \int dx \delta(x - x_0)|x\rangle = |x_0\rangle, \quad (86)$$

and finally our eigen-equation will be

$$X|x_0\rangle = x_0|x_0\rangle. \quad (87)$$

Since x_0 can be any value along the x -axis, this equation can be generalized to

$$X|x\rangle = x|x\rangle, \quad (88)$$

which tells us the simple fact that any state representing a definite position is an eigenvector of the operator X , and the eigenvalue is the position coordinate; thus the operator X is called the position operator, and the wave function in the position representation $\psi(x)$ is the projection of the state-vector onto the eigenvectors of position.

Then we play around with P . Similarly the eigen-equation is given by

$$P|\psi\rangle = p_0|\psi\rangle, \quad (89)$$

and in x -basis,

$$p_0\psi(x) = p_0\langle x|\psi\rangle = \langle x|p_0|\psi\rangle = \langle x|P|\psi\rangle = -i\hbar\frac{d}{dx}\langle x|\psi\rangle = -i\hbar\frac{d}{dx}\psi(x), \quad (90)$$

which is nothing but a simple first-order equation. The solution is given by

$$\psi(x) = Ae^{\frac{i}{\hbar}p_0x}, \quad (91)$$

where A will be normalized as $\frac{1}{\sqrt{2\pi}}$. The eigenvector is constructed as

$$|\psi\rangle = \int dx \psi(x)|x\rangle = \int dx e^{\frac{i}{\hbar}p_0x}|x\rangle \equiv |p_0\rangle, \quad (92)$$

which is just the Fourier-transformation of the state $|x\rangle$. So the eigen-equation will be recovered as

$$P|p_0\rangle = p_0|p_0\rangle \quad \text{or} \quad P|p\rangle = p|p\rangle, \quad (93)$$

where the physical meaning of p and the operator P will be clear later when we take the classical limit, but actually you may already guess that they are momentum operators. As an aside, here we get a commonly used relation by re-express the solution,

$$\langle x|p\rangle = \langle x|\psi\rangle = \psi(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{i}{\hbar}px}, \quad \text{and} \quad \langle p|x\rangle = \frac{1}{\sqrt{2\pi}}e^{-\frac{i}{\hbar}px}, \quad (94)$$

where the left equation represents the a momentum eigenfunction in the position basis, with p fixed and x changeable; the right equation represents the converse case.

To this end, we really see that the eigenfunction of the P operator in the position basis has the form of a wave, and that's why we call it a "wave function". Moreover, the wave length is given by

$$\lambda = \frac{2\pi\hbar}{p}, \quad (95)$$

which is the famous de Broglie's relation between the momentum and the wavelength. An implication of this formula is the fact that to probe objects of ever smaller size one needs particles of ever larger momentum, that's why we need increasingly powerful particle accelerators.

Now that we have an Hermitian operator P and its eigenvectors $|p\rangle$, we can decompose the state vector $|\psi\rangle$ in this basis, just parallel to do it in the position basis,

$$|\psi\rangle = \int dx \tilde{\psi}(p)|p\rangle, \quad \text{where} \quad \tilde{\psi}(p) = \langle p|\psi\rangle, \quad (96)$$

and the identity is subtracted as

$$|\psi\rangle = \int dp |p\rangle\langle p|\psi\rangle \implies I = \int dp |p\rangle\langle p|. \quad (97)$$

Next we want to find the relation between $\tilde{\psi}(p)$ and $\psi(x)$, that is to say, if we knew one of them, we could calculate the other, so as to know the probability of the given interval of the other, e.g, in p -basis, $P(p) = \tilde{\psi}^*(p)\tilde{\psi}(p)$. We simultaneously solve the relation using the Dirac's elegant bra-and-ket's notation,

$$\begin{cases} \tilde{\psi}(p) = \langle p|\psi\rangle = \int dx \langle p|x\rangle\langle x|\psi\rangle = \frac{1}{\sqrt{2\pi}} \int dx e^{-\frac{i}{\hbar}px}\psi(x) \\ \psi(x) = \langle x|\psi\rangle = \int dp \langle x|p\rangle\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi}} \int dp e^{\frac{i}{\hbar}px}\psi(p), \end{cases} \quad (98)$$

which are reciprocal *Fourier transformations* of one another.

Now that we know the exact form of X and P , we may find their commutator and hence we know the uncertainty principle of them. Their commutator is computed as,

$$\begin{aligned}\langle x|[X, P]|\psi\rangle &= \langle x|XP|\psi\rangle - \langle x|PX|\psi\rangle = \int dx' \langle x|X|x'\rangle\langle x'|P|\psi\rangle - \int dx' \langle x|P|x'\rangle\langle x'|X|\psi\rangle \\ &= -i\hbar \int dx' x' \delta(x' - x) \frac{d}{dx'} \psi(x') + i\hbar \int dx' \delta(x' - x) \frac{d}{dx} (x' \psi(x')) \\ &= i\hbar \left(\frac{d}{dx} \left(x\psi(x) \right) - x \frac{d}{dx} \psi(x) \right) \\ &= i\hbar\psi(x) = \langle x|i\hbar|\psi\rangle,\end{aligned}\tag{99}$$

thus we find the commutator is nothing but a complex number,

$$[X, P] = i\hbar.\tag{100}$$

And according to the uncertainty principle,

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \Psi | [A, B] | \Psi \rangle|,\tag{101}$$

we have the uncertainty between X and P ,

$$\Delta X \Delta P \geq \frac{1}{2} |\langle \Psi | [X, P] | \Psi \rangle| = \frac{1}{2} |i\hbar \langle \psi | \psi \rangle| = \frac{\hbar}{2},\tag{102}$$

this is the famous *Heisenberg's uncertainty principle*, which means you cannot measure the position and P in the same time, with $\frac{\hbar}{2}$ the limitation.

As a qualitative illustration of the Heisenberg's uncertainty principle, we knew previously the wave function of an eigenstate of X is highly concentrated about some point x_0 , the probability is perfectly localized, and so is true for the eigenstate of P . But what about the probability of P eigenstate in the X -basis? It's just

$$\psi_p^*(x)\psi_p(x) = \langle p|x\rangle\langle x|p\rangle = \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{i}{\hbar}px}\right)\left(\frac{1}{\sqrt{2\pi}}e^{\frac{i}{\hbar}px}\right) = \frac{1}{2\pi}.\tag{103}$$

The result is completely uniform, with no peaks anywhere on the x axis. Evidently, a state with definite p_0 is completely uncertain in its position, vice versa.

2 Symmetry Operations of Quantum States and Conserved Quantities

2.1 Time evolution of isolated states and the Hamiltonian

In classical mechanics, if two identical isolated systems start out in different states, they stay in different states. It underlies the fact that information is never lost. The same is true in quantum mechanics, and even much stronger. The direct manifestation is that if two states are initially orthogonal, they remain orthogonal throughout the evolution, or *conservation of orthogonality*.

We denote a state at time t as $|\Psi(t)\rangle$, which originates from the state at time $t = 0$, $|\Psi(0)\rangle$, by acting an *time-evolution operator* $\hat{U}(t)$ on it,

$$|\Psi(t)\rangle = \hat{U}(t)|\Psi(0)\rangle. \quad (104)$$

This equation reflects the fact that quantum evolution of states allows us to compute the probabilities of the outcomes of later experiments, in contrast to the case in classical mechanics, classical determinism allows us to predict the results of experiments.

Mathematically, the conservation of orthogonality tells us

$$\langle\Psi(0)|\Phi(0)\rangle = 0 \implies \langle\Psi(t)|\Phi(t)\rangle = \langle\Psi(0)|\hat{U}^\dagger(t)\hat{U}(t)|\Phi(0)\rangle = 0, \quad (105)$$

if we choose $|\Psi(0)\rangle$ and $|\Phi(0)\rangle$ to be any two orthonormal basis of vectors $|i\rangle$ and $|j\rangle$, we could obtain a property of the evolution operator,

$$\langle i|j\rangle = \delta_{ij} = \langle i|\hat{U}^\dagger(t)\hat{U}(t)|j\rangle \implies \hat{U}^\dagger(t)\hat{U}(t) = I, \quad (106)$$

with which the evolution operator satisfies is called *unitary*.

Principle: The evolution of state-vectors with time is unitary.

A direct consequence of the unitarity leads to a much stronger version of conservation of orthogonality — *conservation of overlaps*,

$$\langle A(t)|B(t)\rangle = \langle A(0)|\hat{U}^\dagger(t)\hat{U}(t)|B(0)\rangle = \langle A(0)|B(0)\rangle, \quad (107)$$

which says the inner product of $|A\rangle$ and $|B\rangle$ does not change with time, or the inner product is conserved. Thus superficially conservation of orthogonality is a special case of conservation of overlap, we have seen that conservation of orthogonality can lead to conservation of overlap, thus they are equivalent.

Time evolution is led by incremental changes, so it would be inspiring to study the evolution operator of an infinitesimal time interval ϵ . Another property besides unitarity that an evolution operator should have is *continuity*, which means that the state-vector changes smoothly.

When ϵ is very small, $\hat{U}(\epsilon)$ is close to the unit operator, only differing from it by something of order ϵ , and when ϵ is taken to be zero, the state does not change.

Or to be clearer, the most general form of a state obtained by acting the infinitesimal evolution operator on the initial state, considering continuity, takes the form as

$$U(\epsilon)|\Psi\rangle = |\Psi\rangle + \epsilon|\Theta\rangle, \quad (108)$$

where $|\Theta\rangle$ is some unknown state though, it can be obtained by acting some operator to the original state $|\Theta\rangle \equiv A|\Psi\rangle$, so we have

$$U(\epsilon)|\Psi\rangle = |\Psi\rangle + \epsilon A|\Psi\rangle = (I + \epsilon A)|\Psi\rangle, \quad (109)$$

then we have a form of $U(\epsilon)$ as

$$U(\epsilon) = I + \epsilon A, \quad (110)$$

where the property and meaning of A is to be explored. Using the unitarity, we have

$$I = U^\dagger(\epsilon)U(\epsilon) = (I + \epsilon A^\dagger)(I + \epsilon A) = I + \epsilon(A^\dagger + A), \quad (111)$$

or

$$A^\dagger = -A, \quad (112)$$

which says A is anti-Hermitian.

But it is easy to construct an Hermitian operator, which is more desired by us, out of the anti-Hermitian operator just by multiplying the imaginary unit on it, ie,

$$\mathcal{H} \equiv iA = (-i)(-A) = (iA)^\dagger = \mathcal{H}^\dagger, \quad (113)$$

thus we have the form of the infinitesimal evolution operator as

$$U(\epsilon) = I + \epsilon A = I + \epsilon(-i)(iA) = 1 - i\epsilon\mathcal{H}, \quad (114)$$

and due to the Hermitian property of H , it has the great significance that H is an observable and has a complete set of orthonormal eigenvalues and eigenvectors.

Now we're going to explore more on the evolution of states at an infinitesimal time interval ϵ , we write

$$|\Psi(\epsilon)\rangle = \hat{U}(\epsilon)|\Psi(0)\rangle = |\Psi(0)\rangle - i\epsilon\mathcal{H}|\Psi(0)\rangle, \quad (115)$$

and rearranging this we have

$$\frac{|\Psi(\epsilon)\rangle - |\Psi(0)\rangle}{\epsilon} = -i\mathcal{H}|\Psi(0)\rangle, \quad (116)$$

where the LHS is nothing but the time derivative of the state $|\Psi\rangle$ at time $t = 0$, but since the evolution is assumed to be linear, this should hold at any time t , so we may write

$$\frac{d|\Psi\rangle}{dt} = -i\mathcal{H}|\Psi\rangle, \quad (117)$$

where we wrote partial or total derivative make no difference. This equation tells us how an isolated (not contact with apparatus) state evolves with time, and it is called the *generalized Schrodinger equation*. This is a simple first-order differential equation, and its solution is given by

$$|\Psi(t)\rangle = e^{-i\mathcal{H}t}|\Psi(0)\rangle, \quad (118)$$

thus we may extract the explicit form of the evolution operator as $U(t) = e^{-iHt}$, and we check it by taking the time to be infinitesimal time interval ϵ ,

$$U(\epsilon) = e^{-i\mathcal{H}\epsilon} \approx 1 - i\mathcal{H}\epsilon, \quad (119)$$

which is consistent with our previous reasoning. From this exponential form we now know why we call it unitary, and a unitary operator has another property that its eigenvalues are eigenvalues of the phase operator, in this case, \mathcal{H} .

Next we want to compare it with classical mechanics to see the meaning of \mathcal{H} clearly. The average, or expectation value, of an observable is the closest thing in quantum mechanics to a classical value. From a mathematical point of view, an average is defined by the equation

$$\langle \lambda \rangle = \sum_i \lambda_i P(\lambda_i). \quad (120)$$

If we expand a quantum state $|\Psi\rangle$ in the orthonormal basis of eigenvectors of L ,

$$|\Psi\rangle = \sum_i \alpha_i |\lambda_i\rangle, \quad (121)$$

we try to compute the following quantity,

$$\langle L \rangle \equiv \langle \Psi | L | \Psi \rangle = \sum_i \sum_j \langle \lambda_j | \alpha_j^* \lambda_i \alpha_i | \lambda_i \rangle = \sum_i \lambda_i (\alpha_i^* \alpha_i) = \sum_i \lambda_i P(\lambda_i) = \langle \lambda \rangle, \quad (122)$$

that's why we use the notation $\langle L \rangle$ to denote the average. Let's see how this changes with time,

$$\begin{aligned}\frac{d}{dt} \langle L \rangle &= \frac{d}{dt} \langle \Psi(t) | L | \Psi(t) \rangle = \langle \dot{\Psi}(t) | L | \Psi(t) \rangle + \langle \Psi(t) | L | \dot{\Psi}(t) \rangle \\ &= i \langle \Psi(t) | (\mathcal{H}L - L\mathcal{H}) | \Psi(t) \rangle = i \langle [\mathcal{H}, L] \rangle,\end{aligned}\tag{123}$$

where in the second row we used the generalized Schrodinger equation, and the notation for a commutator. Note that the i before the commutator is important, since the commutator itself is always imaginary.

Equivalently we rewrite it in a shorthand form:

$$\frac{dL}{dt} = -i[L, \mathcal{H}],\tag{124}$$

where letters denote their average value. This is a very interesting and important equation. It relates the time derivative of the expectation value of an observable L to the expectation value of another observable $[L, \mathcal{H}]$. If we assume that the probabilities are nice, narrow, bell-shaped curves, this equation tells us how the peaks of the curves move with time. Equations like this are the closest thing in quantum mechanics to the equations of classical physics.

The equation tells us the expectation value of an observable $\langle Q \rangle$ does not change with time if it commutes with the \mathcal{H} ,

$$\frac{d\bar{Q}}{dt} = -\frac{i}{\hbar} \overline{[Q, \mathcal{H}]} = 0.\tag{125}$$

Moreover if Q commutes with the Hamiltonian, the expectation values of all powers of Q , and even all functions of Q are conserved.

As a special case, the most obvious conserved quantity is \mathcal{H} itself, since it commutes with itself,

$$\frac{d\bar{\mathcal{H}}}{dt} = -i \overline{[\mathcal{H}, \mathcal{H}]} = 0,\tag{126}$$

which says it is a conserved quantity, and we define it to be the energy, since it is conserved in the most general state, considering only the unitarity and continuity of the evolution operator, regardless of any vector basis or specific physical situations.

The only thing to be noticed is the fact that energy has the dimension of energy, by checking the dimensionality of the time-dependent Schrodinger equation, we may fix \mathcal{H} by a factor of \hbar , the Planck constant, i.e, to define the true Hamiltonian operator as $H := \hbar\mathcal{H}$, and we have

$$i\hbar \frac{d|\Psi\rangle}{dt} = H|\Psi\rangle.\tag{127}$$

This is the standard form of the *time-dependent Schrodinger equation*.

Recall that in classical mechanics, the time derivative of a physical quantity is given by its Poisson Bracket with the Hamiltonian,

$$\frac{dL}{dt} = \{L, H\}.\tag{128}$$

Due to their similarity, we may identify a correspondence between the commutators and the Poisson Brackets,

$$[L, H] \iff i\hbar \{L, H\}.\tag{129}$$

Actually they have the same mathematical structures (like anti-symmetric, Bianchi identity) called the Lie Algebra, which would be explored later if we have time.

In classical physics, commutators between ordinary observables are zero. But if we want to apply (124) to classical physics, we must assume the commutators are not zero but a very small value of the order of \hbar . Vice versa, the classical limit is the limit at which \hbar is negligibly small, thus the $i\hbar \{L, H\}$ goes back to the vanishing commutator.

Now it's more concrete to set the spin as an example. The energy of a spin in a magnetic field is given by

$$H \sim \vec{\sigma} \cdot \vec{B} = \sigma_x B_x + \sigma_y B_y + \sigma_z B_z. \quad (130)$$

If the magnetic field lies along the z-axis, the Hamiltonian is proportional to σ_z , in the following expression we'll absorb all the irrelevant numerical constants into a single constant ω , so

$$H = \frac{\hbar\omega}{2} \sigma_z, \quad (131)$$

where we kept $\frac{\hbar}{2}$ for later convenience.

Now we want to find out how the expectation value of the spin varies with time, we can get

$$\begin{aligned} \langle \dot{\sigma}_x \rangle &= -\frac{i}{\hbar} \langle [\sigma_x, H] \rangle = -\frac{i\omega}{2} \langle [\sigma_x, \sigma_z] \rangle = -\omega \langle \sigma_y \rangle \\ \langle \dot{\sigma}_y \rangle &= -\frac{i}{\hbar} \langle [\sigma_y, H] \rangle = -\frac{i\omega}{2} \langle [\sigma_y, \sigma_z] \rangle = \omega \langle \sigma_x \rangle \\ \langle \dot{\sigma}_z \rangle &= -\frac{i}{\hbar} \langle [\sigma_z, H] \rangle = -\frac{i\omega}{2} \langle [\sigma_z, \sigma_z] \rangle = 0 \end{aligned} \quad (132)$$

This result implies that the 3-vector-operator $\vec{\sigma}$ precesses like a gyroscope around the direction of the magnetic field. The precession is uniform, with angular velocity ω . To be specific, the expectation value for a σ_z measurement does not change with time, but the other two expectation values do change.

Now that we have defined that H is a Hermitian operator, we can expand our states in the basis of H 's eigenvectors, to fully solve the Schrodinger equation,

$$|\Psi\rangle = \sum_j \alpha_j |E_j\rangle, \quad (133)$$

where $|E_i\rangle$ satisfies

$$H|E_i\rangle = E_i|E_i\rangle, \quad (134)$$

and E_i s are the eigenvalues of H , or energy.

Using the time-dependent Schrodinger equation and due the fact that basis vector do not change with time, we have

$$\sum_j \dot{\alpha}_j(t) |E_j\rangle = -\frac{i}{\hbar} H \sum_j \alpha_j(t) |E_j\rangle = -\frac{i}{\hbar} \sum_j E_j \alpha_j(t) |E_j\rangle, \quad (135)$$

or regrouping

$$\sum_j \left\{ \dot{\alpha}_j(t) + \frac{i}{\hbar} E_j \alpha_j \right\} |E_j\rangle = 0, \quad (136)$$

thus every coefficient must be zero, and we have again the simplest differential equation,

$$\frac{d\alpha_j(t)}{dt} = -\frac{i}{\hbar} E_j \alpha_j(t) \quad (137)$$

, and the solution is given by

$$\alpha_j(t) = \alpha_j(0) e^{-\frac{i}{\hbar} E_j t} \quad (138)$$

. This equation contains the underlying assumption that Hamiltonian does not depend explicitly on time. And it's the first time for us to see the deep connection between energy and time. We may extract out coefficients at time zero as

$$\alpha_j(0) = \langle E_j | \Psi(0) \rangle \quad (139)$$

, thus we can write the full solution fo the time-dependent Schrodinger equation as

$$|\Psi(t)\rangle = \sum_j \alpha_j(0) e^{-\frac{i}{\hbar} E_j t} |E_j\rangle = \sum_j |E_j\rangle \langle E_j | \Psi(0) \rangle e^{-\frac{i}{\hbar} E_j t}, \quad (140)$$

which emphasizes that we're summing over the basis vectors, and implies that as long as we have the initial state $|\Psi(0)\rangle$ and the Hamiltonian H which governs the physical law of evolution, we can have state $|\Psi(t)\rangle$ at any time as we want.

All the above formalism in this subsection describes how state evolves with time during two measurements. But something different happens when an observation is made, during an experiment the state of a system jumps unpredictably to an eigenstate of the observable that was measured. This phenomenon is called *the collapse of the wave function*. This implies that to examine the measurement process itself as a quantum mechanical evolution, we must consider the entire experimental setup, including the apparatus, as part of a single quantum system.

2.2 General symmetry transformations of states and wave functions

Time evolution is a special case of symmetry transformations. Symmetry transformations are operations that you can do to the system, which don't change the phenomenon or the description of it, e.g, EOM. In this subsection, we are going to discuss more general properties of symmetry transformations and some specific cases like space translations and rotations.

A general symmetry transformation operator V has two properties. The first is that it preserves orthogonality, which gives it the property of unitarity,

$$\langle \phi | V^\dagger V | \psi \rangle = \langle \phi | \psi \rangle \implies V^\dagger V = I. \quad (141)$$

The second is that it preserves time-evolution, which means that a state $|\psi(0)\rangle$ and its transformed state $|\psi'(0)\rangle = V|\psi(0)\rangle$ together evolves with time, and the resulting new state $|\psi'(t)\rangle$ is just the transformed version of $|\psi(t)\rangle$,

$$\begin{array}{ccc} |\psi(0)\rangle & \xrightarrow{U(t)} & |\psi(t)\rangle \\ V \downarrow & & \downarrow V \\ |\psi'(0)\rangle & \xrightarrow{U(t)} & |\psi'(t)\rangle, \end{array} \quad (142)$$

in mathematical language,

$$U(t)V|\psi(0)\rangle = VU(t)|\psi(0)\rangle \implies [V, U(t)] = 0 \quad (143)$$

thus it indicates that it commutes with the time-evolution operator U , where we use the fact that $|\psi(0)\rangle$ is arbitrary. Now that we know that $U(t) = e^{-\frac{i}{\hbar}Ht}$, expanding it gives us

$$[V, H] = 0. \quad (144)$$

So we conclude here that *symmetry transformation is a unitary operation that commutes with the Hamiltonian, thus it is conserved*.

Moreover, if the symmetry transformation is continuous, we can expand it as before, $V = I - i\epsilon G$, where G is a Hermitian operator called the *generator*, which means the operation can be built up by incremental changes with it. Having the generator in hand we get a stronger result as

$$[G, H] = 0. \quad (145)$$

Thus the generator itself commutes with the Hamiltonian and is conserved. By saying that finding all the symmetries, we mean finding all the conserved quantities, as the above equation indicates.

As a special case, we take *the space translation of the wave function* as the first example. By saying “wave function” rather than “states”, we meant we’ve already choose a set of basis, so in below, an operator means it’s in the basis-dependent form.

An infinitesimal space translation operator $V_x(\epsilon)$ takes the original wave function $\psi(x)$ to its nearest point,

$$V_x(\epsilon)\psi(x) = \psi(x - \epsilon) = \psi - \epsilon \frac{\partial \psi}{\partial x}, \quad (146)$$

so we may extract its form as

$$V_x(\epsilon) = I - \epsilon \frac{\partial}{\partial x} = I - i\epsilon \frac{\hat{p}_x}{\hbar}, \quad (147)$$

where $\frac{\hat{p}_x}{\hbar}$ serves as its generator. To judge whether V_x is a symmetry transformation, we see that whether the generator \hat{p}_x commutes with the Hamiltonian. For a free particle, the Hamiltonian is given as

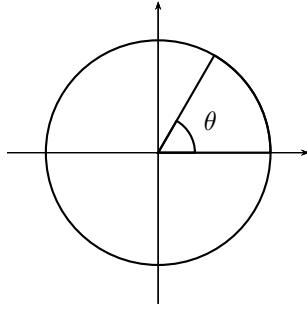
$$H = \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m}, \quad (148)$$

we see that it complies with

$$[\hat{p}_x, H] = 0, \quad (149)$$

since \hat{p}_x commutes with itself and it commutes with \hat{p}_y and \hat{p}_z . So V_x is indeed a translation symmetry under the condition of a free particle.

As another case, we discuss the *space rotation of the wave function*. The wave function has a particular form on a circular wire, as $\psi(\theta)$.



An infinitesimal rotation operator $R_\theta(\epsilon)$ makes the wave function moves around the circle at an infinitesimal angle θ anti-clockwise.

$$R_\theta(\epsilon)\psi(\theta) = \psi(\theta - \epsilon) = \psi - \epsilon \frac{\partial \psi}{\partial \theta}, \quad (150)$$

so the infinitesimal change is

$$\delta\psi = -\epsilon \frac{\partial \psi}{\partial \theta} = -i\frac{\epsilon}{\hbar} \left(-i\hbar \frac{\partial \psi}{\partial \theta} \right) \equiv -i\frac{\epsilon}{\hbar} L\psi, \quad (151)$$

where $L \equiv -i\hbar \frac{\partial}{\partial \theta}$ serves as the generator of the rotation, which we will see is the angular momentum operator. Since the Hamiltonian is expressed in terms of the angular momentum as:

$$H = \frac{L^2}{2I}, \quad (152)$$

where I is the rotation inertia, so the angular momentum commutes with the Hamiltonian $[L, H] = 0$, thus it is a conserved quantity and R_θ is a rotation symmetry.

Next we want to explore the eigenvalues and eigenvectors of the angular momentum operator. The eigen-equation is

$$-i\hbar \frac{\partial}{\partial \theta} \psi(\theta) \equiv L\psi(\theta) = l\psi(\theta), \quad (153)$$

where l is its eigenvalue (of z -component of angular momentum), and the eigenfunction is solved as

$$\psi(\theta) = e^{i\frac{l}{\hbar}\theta}\psi(0) \equiv e^{im\theta}\psi(0), \quad (154)$$

where m is multiples of \hbar by l . Note that in this form we have an additional constraint, which says that rotating the function by the angle of 2π , physical observables should stay the same, for bosons the wave function also stays the same, for fermions the wave function may differ by a phase factor of -1 , thus corresponding to the fact that m must be an integer or half-integer, which we will discuss later.

2.3 Symmetry operations and Energy degeneracy

In the rotation symmetry case we discussed above, we see that the hamiltonian, or the energy, depends only on the magnitude of the angular momentum, but not on the direction of it, so in this case we have

$$E(m) = E(-m), \quad (155)$$

which says that the energies are the same in both anti-clockwise and clockwise cases of angular momentum m , or the energy level m is *degenerate*. *Degeneracy* means there are more than one state at the same energy level. Symmetry sometimes implies degeneracy, but sometimes only one symmetry is not enough to tell.

In the above case, it seems that rotation symmetry leads to the degeneracy of energy level m , but the degeneracy will be broken by a magnetic field pointing outwards or inwards the plane. It'll raise the energy a little bit of the angular momentum in one direction, and lower the energy a little bit of the other direction. Note that the magnetic field however does not violate the rotation symmetry, yet with its existence the degeneracy is gone. To conclude, the wave function has the rotation symmetry, and the magnetic field also has the rotation symmetry, but the energy level is not degenerate. So the rotation symmetry itself is not enough to tell the degeneracy of energy levels, we need one more symmetry which is discrete—*reflection symmetry*.

In this case the reflection is about x -axis, we will denote the operator as M which represents the “mirror reflection”, to avoid the confusion with the rotation operator R . The mirror reflection typically acts on the wave function as

$$M\psi(\theta) = \psi(-\theta). \quad (156)$$

Here we explore the commutator of M with L and H . Note that each wave function can be decomposed in terms of the eigenfunctions of L , so for simplicity, we make the operator act on the eigenfunctions of L .

$$\begin{cases} [M, L]e^{im\theta} = MLe^{im\theta} - LM e^{im\theta} = me^{-im\theta} - (-me^{-im\theta}) = 2me^{-im\theta} \\ [M, L^2]e^{im\theta} = ML^2e^{im\theta} - L^2Me^{im\theta} = m^2e^{-im\theta} - (-m)^2e^{-im\theta} = 0, \end{cases} \quad (157)$$

so M commutes with the Hamiltonian, thus it's a symmetry operation; but it doesn't commute with L .

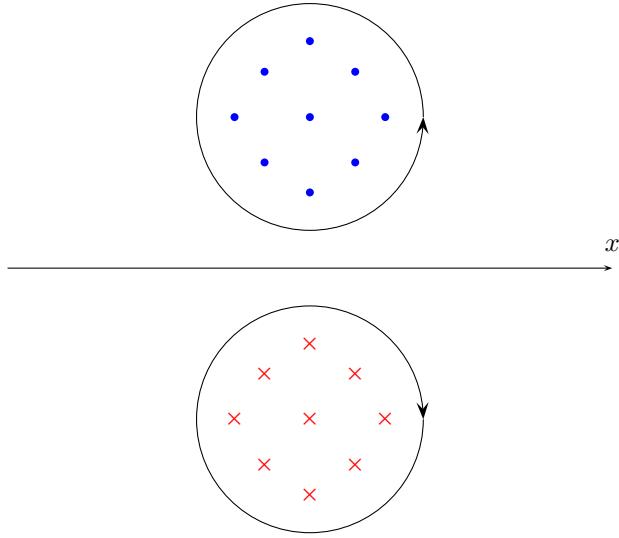
Consequently, if we consider the existence of the magnetic field, we have to add to the Hamiltonian an extra term which is proportional to the product of B and L ,

$$H = \frac{L^2}{2I} + \gamma BL, \quad (158)$$

in this case the reflection operation will not commute with the Hamiltonian, thus it's not a symmetry.

To conclude, in the case without the magnetic field, both the rotation symmetry and the mirror symmetry is guaranteed, thus we have the degeneracy; but in the case with the existence of the magnetic field, the mirror symmetry is broken, thus degeneracy is also gone.

The reflection operation leads to the inversion of the magnetic field, turning pointing outwards into inwards, since the magnetic field is generated by the electric current, which is the actual ingredient we want reflection.



Reflection of the magnetic field about the x-axis

Actually we have a more accurate statement which says,

Symmetry operations that don't commute with each other imply degeneracy.

In mathematical language, “symmetry operations that don't commute with each other” should be written as

$$\begin{cases} [A, H] = 0 \\ [B, H] = 0 \\ [A, B] = iC, \end{cases} \quad (159)$$

where A and B are generators of the corresponding symmetry. C is a new operator, with its Hermitian property guaranteed by the factor of i in front of it. It's easy to prove that C also commutes with the Hamiltonian,

$$[C, H] = 0, \quad (160)$$

thus it's also a symmetry. By the same procedure we create C , we could create more symmetry generators as,

$$[A, C] = D, \quad [B, D] = \dots \quad (161)$$

We keep doing this until we don't get something new, but rarely we might never end. Regardless of endless case, we have a collection of symmetry generators, $\{A, B, C, D, \dots\}$, which form a closed system called a group, with the group multiplication defined as the commutation. With this definition of multiplication, the generator group is called a *Lie algebra*.

Next we're going to explore an important case of energy degeneracy led by non-commuting symmetry operations — three components of the 3-dimensional rotation.

If a system has rotation invariance, all the three components L_x, L_y, L_z should commute with the Hamiltonian H ,

$$[L_i, H] = 0, \quad i = x, y, z. \quad (162)$$

But do the three components commute with each other?

In order to figure out this more easily, we firstly express one components L_z in terms of $x - y$ coordinates. L_z is the generator of a rotation of a particle in 2-dimension. The rotation is led by infinitesimal changes in $x - y$ coordinates as

$$\begin{cases} x' = x \cos \epsilon + y \sin \epsilon = x + \epsilon y \\ y' = -x \sin \epsilon + y \cos \epsilon = y - \epsilon x, \end{cases} \quad (163)$$

thus the variation of the coordinates are

$$\begin{cases} \delta x = \epsilon y \\ \delta y = -\epsilon x, \end{cases} \quad (164)$$

and the wave function is changed by

$$\begin{aligned} \delta\psi &= \frac{\partial\psi}{\partial x}\delta x + \frac{\partial\psi}{\partial y}\delta y \\ &= \epsilon y \frac{\partial\psi}{\partial x} - \epsilon x \frac{\partial\psi}{\partial y} \\ &= -i\frac{\epsilon}{\hbar}(x\hat{p}_y - y\hat{p}_x)\psi, \end{aligned} \quad (165)$$

and in the last subsection we denote $\delta\psi$ as $-i\frac{\epsilon}{\hbar}L_z\psi$, so

$$L_z = xp_y - yp_x, \quad (166)$$

which is consistent with the classical definition

$$(L)_z = (x \times p)_z. \quad (167)$$

By the same reasoning, we give the expression of the two other components:

$$\begin{cases} L_x = yp_z - zp_y \\ L_y = zp_x - xp_z. \end{cases} \quad (168)$$

Then we can calculate the commutator of two of the components, say, L_x and L_y ,

$$[L_x, L_y] = [yp_z - zp_y, zp_x - xp_z] = -i\hbar yp_x + i\hbar xp_y = i\hbar L_z, \quad (169)$$

where we have used the fact that $[x_i, p_j] = i\hbar\delta_{ij}$. Thus we find the commutator of two symmetry generators will give a new symmetry generator, as we stated. Actually the commutators of the three components of angular momentum can be summarized as

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k, \quad (170)$$

where ϵ_{ijk} is the Levi-Civita symbol. For convenience, we sometimes just drop the \hbar in below.

Next we arbitrarily pick one of the components, say L_z , and work with eigenvalues and eigenvectors of it. The eigen-function of L_z is given by

$$L_z|m\rangle = m|m\rangle. \quad (171)$$

But now let's temporarily forget the fact that "m"s are integers, and explore what possible values "m"s can have. As a trick, we may convert the question in another way: suppose we already have an "m", how to find other values of "m"? Here Dirac suggest the clever tools of raising and lowering operators of L_z :

$$\begin{cases} L_+ = L_x + iL_y \\ L_- = L_x - iL_y. \end{cases} \quad (172)$$

It's straightforward to compute their commutators with L_z ,

$$\begin{cases} [L_+, L_z] = -L_+ \\ [L_-, L_z] = L_-. \end{cases} \quad (173)$$

Now let's take one of the commutators to act on an eigenstate of L_z ,

$$-L_+|m\rangle = [L_+, L_z]|m\rangle = mL_+|m\rangle - L_zL_+|m\rangle, \quad (174)$$

by sorting the equation we have

$$L_z L_+|m\rangle = (m + 1) L_+|m\rangle, \quad (175)$$

this equation implies that $L_+|m\rangle$ is also an eigenstate of L_z , and its corresponding eigenvalue is $m + 1$, so we denote this eigenstate as

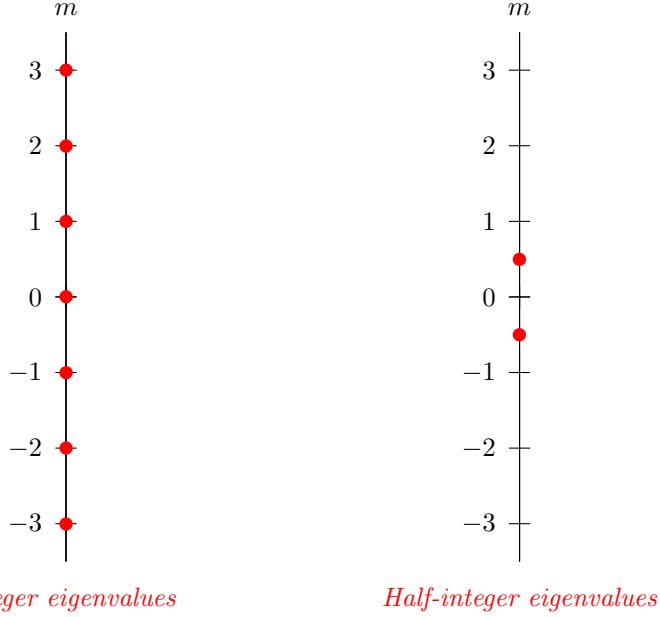
$$L_+|m\rangle \equiv |m + 1\rangle, \quad (176)$$

and the similar procedure goes with L_- ,

$$L_-|m\rangle \equiv |m - 1\rangle, \quad (177)$$

where we actually omit the normalization factor in front of the state by setting it to be 1. From these we see that the effect of L_{\pm} is to raise or lower the energy eigenstate by one unit, thus generating spectrum of energy from a single energy eigenstate.

Since L_z is symmetric when converted upside-down, thus the spectrum has no difference with that of $-L_z$, hence the spectrum is symmetric about the zero point in m -axis. There are two possibilities satisfying this symmetry: “ m ”s are all integers or half integers, which represent orbital or spin angular momentum when truncated at certain maximum and minimum values.



Next we rotate the wave function by an angle of 2π , when “ m ”s are integers, we have

$$R(2\pi)\psi(\theta) = \psi(\theta + 2\pi) = e^{i2\pi m}\psi(\theta) = \psi(\theta), \quad (178)$$

the wave function does not change, and neither are the physical quantities, as we expected, thus we call the particles that has the kind of wave functions, the bosons.

when “ m ”s are half-integers as $m = n + 1/2$, where n is an integer, we have

$$R(2\pi)\psi(\theta) = \psi(\theta + 2\pi) = e^{i2\pi m}\psi(\theta) = e^{i2\pi(n+\frac{1}{2})}\psi(\theta) = -\psi(\theta), \quad (179)$$

the wave function changes by a minus sign, but the physical quantities still do not change since they contain product of conjugates of wave functions, also as we expected. But in these cases we call the particles that has this kind of wave functions, the fermions. Note that for fermions, if we want the wave function itself to be unchanged under rotation, the minimal angle would be 4π .

Experimentally if we want to test this phase factor of -1 of fermion wave function caused by rotation by 2π , we choose the spin direction to denote the direction of the wave function $\psi(x, \sigma)$, and align it to the magnetic field. Then we rotate the magnetic field slowly by the angle of 2π , and interfere it with the original wave function, we can see the interference pattern differs from that of two wave functions not rotated at all.

Each state point in the spectrum axis has the same energy, thus we call these states “*multiplets*”. There are $2l + 1$ states of the angular momentum characterized by l , the maximum value of and the minus minimum value of L_z , and there’re 2 states of the spin characterized by $\frac{1}{2}$.

As an aside, the eigenstates of L_x or L_y are just linear combination of the “ $|m\rangle$ ”s, which is an obvious fact.

We will firstly prove that different states in multiplets, say $|m\rangle$ and $|m - 1\rangle$, have the same energy, without showing explicitly what value the energy is, and after that we calculate the energy value of all the multiplets. Since the Hamiltonian is proportional to the squared angular momentum, the eigenstate of the angular momentum is also the eigenstate of the Hamiltonian, thus it has a definite energy, we set it to be E ,

$$H|m\rangle = E|m\rangle, \quad (180)$$

and we calculate the energy of the lower level by one unit, $|m - 1\rangle$,

$$H|m - 1\rangle = HL_-|m\rangle = L_-H|m\rangle = L_-E|m\rangle = E|m - 1\rangle, \quad (181)$$

where we used the fact that $[H, L_-] = 0$. Hence we find that $|m - 1\rangle$ has the same energy with $|m\rangle$. And by the same reasoning, every state in the multiplets has the same energy E . This degeneracy follows from that the L_x and L_y symmetries don’t commute.

Next we are going to find what exact value does E have. We expand the squared angular momentum as

$$\begin{aligned} L^2 &= L_z^2 + L_x^2 + L_y^2 \\ &= \underbrace{L_z^2 + (L_x - iL_y)(L_x + iL_y)}_{\text{classical terms}} - \underbrace{i[L_x, L_y]}_{\text{quantum correction}} \\ &= L_z^2 + L_z + L_-L_+, \end{aligned} \quad (182)$$

and we apply it to the state $|l\rangle$, where l is the maximum of m ,

$$\begin{aligned} L^2|l\rangle &= L_z^2|l\rangle + L_z|l\rangle + L_- \underbrace{L_+|l\rangle}_0 \\ &= (\underbrace{l^2}_{\text{classical}} + \underbrace{l}_{\text{quantum}})|l\rangle = l(l+1)|l\rangle, \end{aligned} \quad (183)$$

where we used the fact that $L_+|l\rangle = 0$, which means the spectrum is truncated at the highest m , and we see that the eigenvalue of L^2 is $l(l+1)$, differing from that of a classical operator by a factor of l .

Next we apply the same operator on a lowest level $|l-1\rangle$,

$$L^2|l-1\rangle = L^2L_-|l\rangle = L_-L^2|l\rangle = l(l+1)L_-|l\rangle = l(l+1)|l-1\rangle, \quad (184)$$

where we used that $[L^2, L_\pm] = 0$ following from $[L^2, L_i] = 0$. Here we didn’t use directly the form of L^2 as $L_z^2 + L_z + L_-L_+$ because we didn’t get the normalization factor of L_+ acting on $|l-1\rangle$, or more generally $|m\rangle$. From here we see that when you find a multiplet like this, going from l to minus l , such that there’s no more above it or no more below it, you find not only eigenvectors of L_z with eigenvalues m , but you also find the eigenvectors of the squared angular momentum L^2 with eigenvalue $l(l+1)$ where l is just the maximum value of m ; so for each one of these little l s, there are $2l + 1$ states and they all have the same squared angular momentum, and hence the energy, except a factor considering units.

3 Non-relativistic Quantum Fields

3.1 Quantum mechanical harmonic oscillators and ladder operators

The Hamiltonian of a harmonic oscillator is given by

$$\begin{aligned}
H &= \frac{p^2}{2} + \frac{\omega^2 x^2}{2} \\
&= \frac{1}{2} \left\{ (p + i\omega x)(p - i\omega x) - i\omega [x, p] \right\} \\
&= \hbar\omega \left(\frac{p + i\omega x}{\sqrt{2\omega\hbar}} \right) \left(\frac{p - i\omega x}{\sqrt{2\omega\hbar}} \right) + \frac{\hbar\omega}{2} \\
&\equiv \hbar\omega a^+ a^- + \frac{\hbar\omega}{2} \\
&\equiv \hbar\omega \left(N + \frac{1}{2} \right),
\end{aligned} \tag{185}$$

where we set the mass to be 1 and the “spring constant” to be ω^2 , and we used the commutation relation $[x, p] = i\hbar$, and induced two new symbols a^\pm .

We see that there’s constant $\frac{\hbar\omega}{2}$ appears in the expression, we call it the “zero point energy”, or the ground state energy, it’s there even when the harmonic oscillator is in its ground state. Classically, the ground state energy of a harmonic oscillator is zero; but quantum mechanically, the position and the momentum cannot simultaneously be zero, thus the energy being square of them got to be a little energy. It’s the Heisenberg uncertainty principle which tells you that you can’t have zero energy in the ground state.

Since $\frac{\hbar\omega}{2}$ is nothing but a constant we may add anytime and it doesn’t affect the eigenvectors of the Hamiltonian, we can drop it for a while until the last and then we restore it. So we list the Hamiltonian and the ladder operator, as well as a useful tool which is the commutator of the ladder operators, as below,

$$\begin{cases} H = \hbar\omega a^+ a^- \equiv \hbar\omega N \\ a^\pm = \frac{p \pm i\omega x}{\sqrt{2\omega\hbar}} \\ [a^-, a^+] = 1, \end{cases} \tag{186}$$

where the a^+ and a^- are Hermitian conjugates of each other, since p and x are Hermitian operators.

Next we explore the eigenvectors and eigenvalues of the Hamiltonian or the number operator N . The eigen-equation is given by

$$a^+ a^- |n\rangle \equiv N |n\rangle = n |n\rangle, \tag{187}$$

where $|n\rangle$ is its eigenvector and n its eigenvalue. According to the commutation relation of the ladder operators, we have

$$a^+ |n\rangle = a^+ [a^-, a^+] |n\rangle = a^+ a^- a^+ |n\rangle - a^+ a^+ a^- |n\rangle = N a^+ |n\rangle - a^+ N |n\rangle = N a^+ |n\rangle - n a^+ |n\rangle, \tag{188}$$

by moving the last term to the left, we get

$$N a^+ |n\rangle = (n + 1) a^+ |n\rangle, \tag{189}$$

which tells us that the raising operator a^+ acting on the state $|n\rangle$ gives us a new eigenvector of N with an eigenvalue one unit up, $n + 1$. And the same reasoning goes with the lowering operator a^- , it takes the level one unit down, except that it cannot go infinitely down, since the Hamiltonian is always positive, there’s a ground state $|0\rangle$ exerted by a^- gives 0,

$$a^- |0\rangle = 0. \tag{190}$$

Thus, we write the effect that a^\pm act on $|n\rangle$ as

$$\begin{cases} a^+ |n\rangle = c_n |n + 1\rangle \\ a^- |n\rangle = d_n |n - 1\rangle, \end{cases} \tag{191}$$

and by using the orthonormal relation

$$\langle n|m\rangle = \delta_{nm}, \quad (192)$$

the normalization factor can be computed as

$$\begin{aligned} c_n^2 &= \langle n+1|c_n^2|n+1\rangle = \langle n|a^-a^+|n\rangle = \langle n|a^+a^-|n\rangle + \langle n|[a^-, a^+]|n\rangle = n+1 \implies c_n = \sqrt{n+1} \\ d_n^2 &= \langle n-1|d_n^2|n-1\rangle = \langle n|a^+a^-|n\rangle = \langle n|n|n\rangle = n \implies d_n = \sqrt{n}, \end{aligned} \quad (193)$$

so we can write

$$\begin{cases} a^+|n\rangle = \sqrt{n+1}|n+1\rangle \\ a^-|n\rangle = \sqrt{n}|n-1\rangle, \end{cases} \quad (194)$$

which are consistent with the fact that $a^+a^-|n\rangle = n|n\rangle$ and $a^-|0\rangle = 0$. Every other state can be built up by acting a^+ s on the ground state $|0\rangle$.

Next we're going to demonstrate what explicitly the ground state is, by solving the differential equation it satisfies. We work in x -basis, where the ground state wave function is denoted by $\langle x|0\rangle \equiv \psi_0(x)$, and x and p operators are in their basis-dependent form. Instead of solving the Schrodinger's equation directly, we may find a clever way to solve another equation which nothing but $\langle x|a^-|0\rangle = 0$, in x -basis, the equation is given by

$$\left(-i\hbar\frac{d}{dx} - i\omega x\right)\psi_0(x) = 0 \implies \left(\frac{d}{dx} + \frac{\omega}{\hbar}x\right)\psi_0(x) = 0, \quad (195)$$

we propose an ansatz in the form of

$$\psi_0(x) \sim e^{f(x)}, \quad (196)$$

thus the differential equation becomes

$$f'(x) + \frac{\omega}{\hbar}x = 0, \quad (197)$$

so the unknown function is given by

$$f = -\frac{\omega}{2\hbar}x^2 + C, \quad (198)$$

and the ground state function is proportional to

$$\psi_0(x) \sim e^{-\frac{\omega}{2\hbar}x^2}. \quad (199)$$

This is a bell-shaped even function with no nodes, which goes to zero when x approached infinity, as we aspired. picture to be added

We check this solution by inserting it into the time-independent Schrodinger's equation,

$$-\frac{\hbar^2}{2}\frac{d^2}{dx^2}\psi(x) + \frac{1}{2}\omega^2x^2\psi(x) = E\psi(x), \quad (200)$$

for convenience, we list the first and second order derivative of $\psi_0(x)$ as follows

$$\begin{cases} \psi'_0(x) = -\frac{\omega}{\hbar}xe^{-\frac{\omega}{2\hbar}x^2} = -\frac{\omega}{\hbar}x\psi_0(x) \\ \psi''_0(x) = -\frac{\omega}{\hbar}e^{-\frac{\omega}{2\hbar}x^2} + \frac{\omega^2}{\hbar^2}x^2e^{-\frac{\omega}{2\hbar}x^2} = \left(\frac{\omega^2}{\hbar^2}x^2 - \frac{\omega}{\hbar}\right)\psi_0(x), \end{cases} \quad (201)$$

thus the LHS of the SEQ is given by

$$\left(-\frac{1}{2}\omega^2x^2 + \frac{\hbar\omega}{2} + \frac{1}{2}\omega^2x^2\right)\psi_0(x) = \frac{\hbar\omega}{2}\psi_0(x), \quad (202)$$

joining it with RHS, we see that

$$E_0 = \frac{\hbar\omega}{2}, \quad (203)$$

which is the zero point energy, as we desired.

Excitation states' wave functions can be obtained by acting the raising operator onto the ground state function, with each result the same order the coefficient in front of the $\psi_0(0)$ has as the excitation level. For example, the first excitation state can be gained by

$$\psi_1(x) = a^+ \psi_0(x) = \left(-i\hbar \frac{d}{dx} + i\omega x \right) \psi_0(x), \quad (204)$$

but we already knew

$$\left(-i\hbar \frac{d}{dx} - i\omega x \right) \psi_0(x) = 0, \quad (205)$$

thus the first excitation wave function is

$$\psi_1(x) = 2i\omega x \psi_0(x) = 2i\omega x e^{-\frac{\omega}{2\hbar}x^2}, \quad (206)$$

without considering the proper factor in front of it. This is an odd function with one nodes, which also shrinks to zero at infinity. picture to be added

The following second excitation wave function is straightforward,

$$\psi_2(0) = (a^+)^2 \psi_0(x) = a^+ \psi_1(x) = \left(-i\hbar \frac{d}{dx} + i\omega x \right) (2i\omega x \psi_0(x)) = (2\omega x^2 - \hbar) \psi_0(x), \quad (207)$$

which becomes even again and has two nodes. picture to be added

As more and more raising operator acting on the ground state function, the resulting function becomes more and more wiggled. With the most wiggled part concentrated at the origin which represents the high momentum, and the larger amplitude away from the origin represents the big potential energy, and finally damps down to zero at infinity. This picture resembles much as the classical picture of a harmonic oscillator vibrating back and forth in a given potential. picture to be added

3.2 Many-harmonic-oscillator system and quantum field operators

Consider a system containing many harmonic oscillators, each at its energy eigenstate, this state can be represented as

$$|n_1, n_2, \dots, n_i, \dots\rangle, \quad (208)$$

if we want to raise or lower the energy of the i th harmonic oscillator, we may make use of the ladder operator as

$$\begin{aligned} a_i^+ |n_1, n_2, \dots, n_i, \dots\rangle &= \sqrt{n_i + 1} |n_1, n_2, \dots, n_i + 1, \dots\rangle \\ a_i^- |n_1, n_2, \dots, n_i, \dots\rangle &= \sqrt{n_i} |n_1, n_2, \dots, n_i - 1, \dots\rangle. \end{aligned} \quad (209)$$

Next we're going to generalize this set of notations to the many-particle system in an arbitrary potential $V(x)$. The energy eigenstate wave functions by solving the Schrodinger's equation are given by $\psi_i(x)$ (schematic picture to be added), their orthogonality relations are given by

$$\begin{cases} \int dx \psi_i^*(x) \psi_j(x) = \int dx \langle i|x\rangle \langle x|j\rangle = \langle i|j\rangle = \delta_{ij} \\ \sum_i \psi_i^*(x) \psi_i(y) = \sum_i \langle y|i\rangle \langle i|x\rangle = \langle y|x\rangle = \delta(x - y). \end{cases} \quad (210)$$

Suppose we have n_i bosonic particles in the i th state, the state of the system is denoted by the same notation as the above many-SHO system as,

$$|n_1, n_2, \dots, n_i, \dots\rangle, \quad (211)$$

and we replace the raising/lowering operator to a set of new operators called the creation/annihilation operator,

$$\begin{cases} a_i^+ &\rightarrow a_i^\dagger \\ a_i^- &\rightarrow a_i, \end{cases} \quad (212)$$

meaning that when it acts on a state, it creates/annihilates a particle at the corresponding energy eigenstate,

$$\begin{aligned} a_i^\dagger |n_1, n_2, \dots, n_i, \dots\rangle &:= \sqrt{n_i + 1} |n_1, n_2, \dots, n_i + 1, \dots\rangle \\ a_i |n_1, n_2, \dots, n_i, \dots\rangle &:= \sqrt{n_i} |n_1, n_2, \dots, n_i - 1, \dots\rangle, \end{aligned} \quad (213)$$

where the normalization factor is a convention satisfying

$$\begin{aligned} a_i^\dagger a_i |n_1, n_2, \dots, n_i, \dots\rangle &= n_i |n_1, n_2, \dots, n_i + 1, \dots\rangle \\ a_i |n_1, n_2, \dots, 0, \dots\rangle &= 0, \end{aligned} \quad (214)$$

where we denote $a_i^\dagger a_i \equiv N_i$ as the occupation operator, and the commutator is easily computed as

$$[a_i, a_j^\dagger] = \delta_{ij}, \quad (215)$$

the similar result as before.

Specially, when the creation/annihilation operator acts on a single particle state, it gives

$$\begin{aligned} a_i^\dagger |0\rangle &= |i\rangle \\ a_i |j\rangle &= \delta_{ij} |0\rangle, \end{aligned} \quad (216)$$

where $|i\rangle$ is the i th energy eigenstate, and in x -basis,

$$\langle x | i \rangle = \psi_i(x). \quad (217)$$

Next, importantly, we introduce two field operators defined as

$$\begin{cases} \hat{\Psi}^\dagger(x) := \sum_i a_i^\dagger \psi_i^*(x) \\ \hat{\Psi}(x) := \sum_i a_i \psi_i(x). \end{cases} \quad (218)$$

Note that these “field operators” are actually “operator fields”, meaning that at each point x in the space there’s an operator $\hat{\Psi}^\dagger$ and $\hat{\Psi}$. Then we are going to see what effects these operators have, firstly by acting the $\hat{\Psi}^\dagger(x)$ on a vacuum state,

$$\hat{\Psi}^\dagger(x) |0\rangle = \sum_i \psi_i^*(x) a_i^\dagger |0\rangle = \sum_i |i\rangle \langle i | x \rangle = |x\rangle, \quad (219)$$

it means that we created a particle at point x ! By the similar reasoning, $\hat{\Psi}(x)$ will annihilate a particle at point x ,

$$\hat{\Psi}(x) |y\rangle = \sum_i \psi_i(x) a_i \sum_j |j\rangle \langle j | y \rangle = \sum_{ij} \psi_i(x) \delta_{ij} |0\rangle \psi_j^*(y) = \sum_i \psi_i(x) \psi_i^*(y) |0\rangle = \delta(x - y) |0\rangle, \quad (220)$$

which means if this particle is located at x , the operator annihilates it and gives a vacuum state (though with a divergent factor), and if this particle is at another position y , the operator just kills it to give 0. Instead of creating/annihilating a particle at definite energy as a_i^\dagger and a_i , $\hat{\Psi}^\dagger(x)$ and $\hat{\Psi}(x)$ creates/annihilates a particle at definite position.

The only non-vanishing commutation relation is computed as

$$\begin{aligned} [\Psi(x), \Psi^\dagger(y)] &= \left[\sum_i a_i \psi_i(x), \sum_j a_j^\dagger \psi_j^*(y) \right] \\ &= \sum_{ij} \psi_i(x) \psi_j^*(y) [a_i, a_j^\dagger] \\ &= \sum_{ij} \psi_i(x) \psi_j^*(y) \delta_{ij} \\ &= \sum_i \psi_i(x) \psi_i^*(y) = \delta(y - x). \end{aligned} \quad (221)$$

Next we are going to look at an interesting quantity composed of the creation and annihilation operators,

$$\hat{N} \equiv \int dx \Psi^\dagger(x)\Psi(x) = \int dx \sum_i \psi_i^*(x)a_i^\dagger \sum_j \psi_j(x)a_j = \sum_{ij} \delta_{ij}a_i^\dagger a_j = \sum_i a_i^\dagger a_i = \sum_i N_i, \quad (222)$$

where we get nothing but the total number of particle at all energy levels, or in all field; thus $\Psi(x)^\dagger\Psi(x)$ represents the particle number density at the given point x . This is to be compared with the equation that

$$1 = \int dx \psi^*(x)\psi(x), \quad (223)$$

where the $\psi(x)$ is the normal quantum wave function; the 1 here represents the total probability, or incidentally also represents the particle number. The similar form by substituting the wave function to the field operator is inspiring in the next few cases.

In quantum mechanics, the average value of a observable is given by the integral of the corresponding Hermitian operator sandwiched by the wave function and its complex conjugate, for example, the average energy is

$$\langle H \rangle = \int dx \psi^*(x)H\psi(x); \quad (224)$$

in quantum field theory, the corresponding operator has the similar structure except that replacing the wave function by the field operator, e.g,

$$\begin{aligned} \hat{H}_{field} &= \int dx \Psi^\dagger(x)\hat{H}_{par}\Psi(x) \\ &= \int dx \sum_i \psi_i^*(x)a_i^\dagger \hat{H}_{par} \sum_j \psi_j(x)a_j \\ &= \int dx \sum_{ij} \psi_i^*(x)a_i^\dagger \epsilon_j \psi_j(x)a_j \\ &= \sum_{ij} \delta_{ij}a_i^\dagger a_j \epsilon_j = \sum_i a_i^\dagger a_i \epsilon_i = \sum_i \hat{N}_i \epsilon_i \end{aligned} \quad (225)$$

where we have used the time-independent Schrodinger's equation, i.e, the eigen-equation of energy $\hat{H}_{par}\psi_i(x) = \epsilon_i \psi_i(x)$. The result is as we expected. And the same goes with the field momentum operator

$$\begin{aligned} \hat{P}_{field} &= \int dx \Psi^\dagger(x)\hat{P}_{par}\Psi(x) = -i\hbar \int dx \Psi^\dagger(x) \frac{\partial}{\partial x} \Psi(x) \\ &= \int dx \sum_i \psi_i^*(x)a_i^\dagger \hat{P}_{par} \sum_j \psi_j(x)a_j \\ &= \int dx \sum_i \psi_i^*(x)a_i^\dagger p_j \sum_j \psi_j(x)a_j \\ &= \sum_i a_i^\dagger a_i p_i = \sum_i N_i p_i \end{aligned} \quad (226)$$

Next we're going to develop the field operator of free particles, i.e, without any potential energy $V(x)$. In this case, the Hamiltonian is written as $H = p^2/2m$, the momentum operator commutes with the Hamiltonian, thus the momentum eigenfunctions are the same with the energy eigenfunctions. The momentum eigenfunction in x -basis is nothing but the plane wave $e^{ipx}/\sqrt{2\pi}$, thus the sum in the definition of the creation/annihilation operator becomes an integral,

$$\begin{cases} \Psi^\dagger(x) = \frac{1}{\sqrt{2\pi}} \int dp a^\dagger(p) e^{-ipx} \\ \Psi(x) = \frac{1}{\sqrt{2\pi}} \int dp a(p) e^{ipx}, \end{cases} \quad (227)$$

with a_p^\dagger and a_p meaning creating/annihilating a particle at the corresponding momentum. The creation/annihilation operator in position and momentum basis are just Fourier transformations of each other, so the inverse relation is given by

$$\begin{cases} a^\dagger(p) = \frac{1}{\sqrt{2\pi}} \int dx \Psi^\dagger(p) e^{ipx} \\ a(p) = \frac{1}{\sqrt{2\pi}} \int dx \Psi(p) e^{-ipx}, \end{cases} \quad (228)$$

To proceed, let's study the Hamiltonian of a free particle field, which is given by

$$H = mc^2 \int dx \Psi^\dagger(x) \Psi(x) + \int dx \Psi^\dagger(x) \left(-\frac{\nabla^2}{2m} \right) \Psi(x), \quad (229)$$

where the first term is the rest mass energy of the field, and the second term is the kinetic energy of the field.

Recall that the Hamiltonian governs the time-evolution of the state, or to say, to update the state; so if in this updating process a certain quantity is unchanged, we say that this quantity is conserved. Let's study the first term, which is proportional to the total particle number in the field,

$$\begin{aligned} \int dx \Psi^\dagger(x) \Psi(x) &= \frac{1}{2\pi} \int dx \int dp a^\dagger(p) e^{-ipx} \int dq a(q) e^{iqx} \\ &= \int dp a^\dagger(p) \int dq a(q) \left(\frac{1}{2\pi} \int dx e^{i(q-p)x} \right) \\ &= \int dp a^\dagger(p) \int dq a(q) \delta(q-p) \\ &= \int dp a^\dagger(p) a(p). \end{aligned} \quad (230)$$

When this term acts on a state, it finds a particle at momentum p and annihilates it and then creates a particle at the same momentum, in this way, the momentum is conserved. The working ingredient here is the integral of the exponential gives the Dirac-delta function

$$\frac{1}{2\pi} \int dx e^{i(q-p)x} = \delta(q-p), \quad (231)$$

so we can say that the delta function is the delta function of momentum conservation.

Then we can generate this type of terms to any term has the following form:

$$\int dx \Psi_A^\dagger(x) \Psi_B^\dagger(x) \Psi_A(x) \Psi_B(x), \quad (232)$$

where different letters denote different type of particles. And this term means that two particles A and B disappear and they reappear, i.e, they scatter off each other. The integral over space guarantees that wherever the scattering happens, this term always takes it into account. No matter how many particle type creation/annihilation operators there in a term, the integral will always contain the working ingredient term

$$\frac{1}{2\pi} \int dx e^{i(q_A+q_B-p_A-p_B)x} = \delta(q_A+q_B-p_A-p_B), \quad (233)$$

the total momentum will always be conserved.

As another example, let's consider

$$\int dx \Psi_A^\dagger(x) \Psi_B^\dagger(x) \Psi_C(x), \quad (234)$$

which describes a decaying process — a particle C becomes particles A and B , and the same momentum conservation ingredient is still the same. But note that this term is not Hermitian so if it's in the Hamiltonian we must

add its Hermitian conjugate which represents the inverse process — particles B and C becomes A . And usually a coupling constant denoted by g reflecting the strength of the process, mostly measured in the experiments, is given in front of the term. So we may rewrite the term as

$$g \int dx \left(\Psi_A^\dagger(x) \Psi_B^\dagger(x) \Psi_C(x) + \Psi_C^\dagger(x) \Psi_B(x) \Psi_A(x) \right). \quad (235)$$

If we further expand the evolution operator to the second order, we will have the H^2 term. Considering the decay process in the above, we will find terms as

$$\begin{cases} g \int dx \Psi_A^\dagger(x) \Psi_B^\dagger(x) \Psi_C(x) \Psi_C^\dagger(x) \Psi_B(x) \Psi_A(x) \\ g \int dx \Psi_C^\dagger(x) \Psi_B(x) \Psi_A(x) \Psi_A^\dagger(x) \Psi_B^\dagger(x) \Psi_C(x), \end{cases} \quad (236)$$

where the first term represents A and B combined to form C then C decays into A and B again, just like a scattering process of A and B , except for that it doesn't happen at a point, but with C as the intermediate particle; the second term represents C becomes C again, with A and B being the temporary particles.

Then we study the second term of the free particle field Hamiltonian,

$$\begin{aligned} H &= \int dx \Psi^\dagger(x) \left(-\frac{\nabla^2}{2m} \right) \Psi(x) \\ &= \frac{1}{2\pi} \int dx \int dp a^\dagger(p) e^{-ipx} \left(-\frac{1}{2m} \frac{\partial^2}{\partial x^2} \right) \int dq a(q) e^{iqx} \\ &= \frac{1}{2\pi} \int dx \int dp a^\dagger(p) e^{-ipx} \left(\frac{q^2}{2m} \right) \int dq a(q) e^{iqx} \\ &= \frac{1}{2m} \int dp p^2 a^\dagger(p) a(p), \end{aligned} \quad (237)$$

which also preserves the conservation of momentum.

3.3 Exchange properties for bosons and fermions

Previously we talked a little bit about bosons and fermions in terms of the phase factor caused by rotating their single particle wave function. Now we come to the second property of these two species of particles, or to say, another equivalent definition of them.

For identical particles, if we exchange the positions, or to be more precise the states of them, the physical properties of the system should remain the same, which is intuitive since they are “identical”. But the wave function describing the system may differ by a phase factor since this does not affect the physical observables. According to the different phase factors, particles can be divided into two kinds — fermions and bosons. In the following, we will discuss the classification both in quantum mechanical sense and quantum field sense.

Quantum mechanically, a state of a system consists of two particle at position x and y is denoted as $|x, y\rangle$, naturally the corresponding wave function is given by $\psi(x, y) = \langle x, y|\psi\rangle$. We denote the “swapping operator” as S , when it acts on the state of the system, it just switches the positions of the two particles,

$$S|x, y\rangle = |y, x\rangle, \quad (238)$$

and when acting twice the operator on the state, it must turn back into the original state, thus the square of the swapping operator must be the unit operator,

$$S^2|x, y\rangle = S|y, x\rangle = |x, y\rangle \implies S^2 = I, \quad (239)$$

which mathematically says that the eigenvalues of S would be ± 1 . The identical particles' state described by the eigenvector corresponding to the eigenvalue 1 are called bosons; and particles are called fermions if their state corresponds to the eigenvalue of -1 ,

$$\begin{aligned} S|x, y\rangle &= |y, x\rangle = |x, y\rangle, && \text{Bosons;} \\ S|x, y\rangle &= |y, x\rangle = -|x, y\rangle, && \text{Fermions.} \end{aligned} \quad (240)$$

Pauli exclusion principle goes naturally out of this definition for fermions, if two particles are put in the same state (here the same positon), the equation goes like

$$|x, x\rangle = -|x, x\rangle \implies |x, x\rangle = 0, \quad (241)$$

which says there is no such a state existed, or, you can't put two particles in the same position. Naturally these can also be represented by wave functions, which we won't bother writing and will directly use it below.

If we refer the state of a particle as the energy level, without considering the spin degree of freedom, let's see how we construct boson and fermion systems and how Pauli exclusion principle works. For convenience, we will work in wave function representation here. If two particles are independent of each other, naming they don't have interactions, the wave function of the two-particle system can be written as the product of the single particle wave function of the two particles,

$$\psi(x, y) = \psi_1(x)\psi_2(y), \quad (242)$$

with 1 and 2 representing two energy levels, same or different. If the two particles are in the same energy level, say 0, the wave function is then symmetric under exchanging the two particles,

$$S\psi(x, y) = \psi(y, x) = \psi_0(y)\psi_0(x) = \psi_0(x)\psi_0(y) = \psi(x, y), \quad (243)$$

in this sense the particles must be bosons. If the two particles are in different energy levels, swopping the two particles doesn't give the minus or the original wave function at all,

$$S\psi(x, y) = \psi(y, x) = \psi_1(y)\psi_2(x) \neq \pm\psi_1(x)\psi_2(y) = \pm\psi(x, y). \quad (244)$$

But we may construct new wave functions by superposition of the original and the swopped ones, with plus or minus operation giving symmetric or anti-symmetric wave function,

$$\begin{cases} \psi_S(x, y) := \psi(x, y) + S\psi(x, y) = \psi_1(x)\psi_2(y) + \psi_1(y)\psi_2(x) \\ \psi_A(x, y) := \psi(x, y) - S\psi(x, y) = \psi_1(x)\psi_2(y) - \psi_1(y)\psi_2(x), \end{cases} \quad (245)$$

and it is easy to verify their symmetric/anti-symmetric property by acting the swopping operator on,

$$\begin{cases} S\psi_S(x, y) = S\psi(x, y) + S^2\psi(x, y) = S\psi(x, y) + \psi(x, y) = \psi_S(x, y) \\ S\psi_A(x, y) = S\psi(x, y) - S^2\psi(x, y) = S\psi(x, y) - \psi(x, y) = -\psi_A(x, y). \end{cases} \quad (246)$$

Thus the symmetric wave function describes a system of two bosons, and the anti-symmetric wave function describes a system of two fermions. For the boson system, if we put two particles in the same position or energy level, the wave function just doubles; but for the fermions system, if we do the same thing, the wave function just goes to zero,

$$\begin{cases} \psi'_S(x, y) = \psi_1(x)\psi_1(y) + \psi_1(y)\psi_1(x) = 2\psi_1(x)\psi_1(y) \\ \psi_S(x, x) = \psi_1(x)\psi_2(x) + \psi_1(x)\psi_2(x) = 2\psi_1(x)\psi_2(x) \\ \psi'_A(x, y) = \psi_1(x)\psi_1(y) - \psi_1(y)\psi_1(x) = 0 \\ \psi_A(x, x) = \psi_1(x)\psi_2(x) - \psi_1(x)\psi_2(x) = 0, \end{cases} \quad (247)$$

with the primed wave function denoting putting particles in the same energy level and $\psi(x, x)$ denoting putting particle in the same position. The zero wave function for fermions says that the probability for finding such a state is 0, or to say, this kind of situation where two particles are in the same state cannot happen at all. This

is also a natural manifestation of the Pauli exclusion principle, and this kind of symmetric and anti-symmetric wave functions occurs frequently in particle physics multiplets, thus one should get familiar with them.

The above logic goes the same with the system having more than two particles, denoted as the multiple-particle state $|x_1, x_2, x_3, \dots\rangle$ and system with states denoted by both position and spin $|x_1, \sigma_1; x_2, \sigma_2; x_3, \sigma_3; \dots\rangle$.

In quantum field theory, the state of a system of two particles is generated by acting two creation operators on the vacuum state, $|x, y\rangle = \Psi^\dagger(x)\Psi^\dagger(y)|0\rangle$. From the swapping properties, we can obtain the commutation relations for boson fields and the anti-commutation relations for fermion fields,

$$\begin{aligned}\Psi_B^\dagger(x)\Psi_B^\dagger(y)|0\rangle &= \Psi_B^\dagger(y)\Psi_B^\dagger(x)|0\rangle \implies [\Psi_B^\dagger(x), \Psi_B^\dagger(y)] = 0 && \text{Bosons;} \\ \Psi_F^\dagger(x)\Psi_F^\dagger(y)|0\rangle &= -\Psi_F^\dagger(y)\Psi_F^\dagger(x)|0\rangle \implies \{\Psi_F^\dagger(x), \Psi_F^\dagger(y)\} = 0 && \text{Fermions.}\end{aligned}\quad (248)$$

And still, mathematical consistency of the Pauli exclusion principle goes well with creation operators' representation,

$$\{\Psi_F^\dagger(x), \Psi_F^\dagger(x)\} = 2\Psi_F^\dagger(x)\Psi_F^\dagger(x) = 0 \implies \Psi_F^\dagger(x)\Psi_F^\dagger(x)|0\rangle = 0, \quad (249)$$

meaning that you cannot create two particles at the same point from the vacuum.

Now that the relation of fermion field operators are given by anti-commutators, we go further to derive the anti-commutator for annihilation and creation operators, by acting it on the vacuum state

$$\{\Psi_F(x), \Psi_F^\dagger(y)\}|0\rangle = \Psi_F(x)\Psi_F^\dagger(y)|0\rangle + \underline{\Psi_F^\dagger(y)\Psi_F(x)}|0\rangle = \Psi_F(x)|y\rangle = \delta(x-y)|0\rangle, \quad (250)$$

and on single particle state $|y\rangle$ as

$$\{\Psi_F(x), \Psi_F^\dagger(y)\}|y\rangle = \underline{\Psi_F(x)\Psi_F^\dagger(y)}|y\rangle + \Psi_F^\dagger(y)\Psi_F(x)|y\rangle = \delta(x-y)\Psi_F^\dagger(y)|0\rangle = \delta(x-y)|y\rangle, \quad (251)$$

thus we may conclude

$$\{\Psi_F(x), \Psi_F^\dagger(y)\} = \delta(x-y). \quad (252)$$

Yet this anti-commutator acting on $|x\rangle$ will result in some subtleties which are not important, so we just ignore it.

The creation/annihilation operators of fermions of a certain energy level also takes the anti-commutation relations, which are given as

$$\{a_F, a_F^\dagger\} = 1 \iff \{a_F^\dagger, a_F\} = 1, \quad (253)$$

this is easy to be proved by acting this to the no-particle state or 1-particle state on that given energy level,

$$\begin{aligned}\{a_F, a_F^\dagger\}|0\rangle &= a_F a_F^\dagger|0\rangle + \underline{a_F^\dagger a_F}|0\rangle = a_F|1\rangle = |0\rangle \\ \{a_F, a_F^\dagger\}|1\rangle &= \underline{a_F a_F^\dagger}|1\rangle + a_F^\dagger a_F|1\rangle = a_F^\dagger|0\rangle = |1\rangle.\end{aligned}\quad (254)$$

Observing the anti-commutation relations, we see that from the algebraic structure, there's no clear distinction between the creation and annihilation operator; they are totally symmetric in mathematical sense. On contrast, the commutation relations of boson creation and annihilation operators show a clear unequal footing of them,

$$[a_B, a_B^\dagger] = 1 \iff [a_B^\dagger, a_B] = -1. \quad (255)$$

This is basically due to the physical fact of the following. For fermions the 0-particle state and the 1-particle state are the only two possible state on a certain energy level, thus they are treated as equal — acting the creation operator twice on the 0-particle state will give zero, as well as acting the annihilation operator twice on the 1-particle state, the particle number axis can turn upside-down without any difference to be seen. But for bosons the 0-particle state is the ground state, above it, there are 1-particle, 2-particle, ... states etc, hence there is no reason we can treat them equal, to illustrate more explicitly, for a state which number of particles is given, one can always act the creation operators on it and never get zero, but acting the annihilation operators on for finite number of times, one can finally get zero; you turn upside-down the particle number axis, it's not

as same as before, i.e, it's asymmetric. (Figures of boson and fermion particle number axis and the effect of creation/annihilation operators are to be added.)

One last thing, we recover the anti-commutation relation of the field operators from the generalized anti-commutation relation of the annihilation/creation operators of two energy levels i and j ,

$$\{a_i, a_j^\dagger\} = \delta_{ij}. \quad (256)$$

The approach is by simply using the definition of the field operators,

$$\begin{aligned} \{\Psi_F(x), \Psi_F^\dagger(y)\} &= \left\{ \sum_i a_i \psi_i(x), \sum_j a_j^\dagger \psi_j^*(y) \right\} \\ &= \sum_{ij} \psi_i(x) \psi_j^*(y) \{a_i, a_j^\dagger\} = \sum_{ij} \psi_i(x) \psi_j^*(y) \delta_{ij} = \sum_i \psi_i(x) \psi_i^*(y) = \delta(x - y), \end{aligned} \quad (257)$$

which is consistent with the result when we act the field anti-commutator on single-particle states $|0\rangle$ and $|y\rangle$.

4 Particle-wave Duality

4.1 Property of waves and waves packets

A wave has certain characteristic quantities attached to it, they can be divided into spatial quantities and temporal quantities, and these kinds are related by the speed of the wave. The simplest relation would be

$$v = \frac{\lambda}{T}, \quad (258)$$

with λ its wavelength as the spatial quantity, and the T as its period as the temporal quantity.

Another pair of temporal quantities are the frequency and the phase frequency, their definition and relation with the above equation are given as

$$\begin{cases} f := \frac{1}{T} = \frac{v}{\lambda} \\ \omega := 2\pi f = \frac{2\pi v}{\lambda}. \end{cases} \quad (259)$$

The frequency means how many wavelength pass through a fixed point in space per unit time, so the definition is straightforward: the numerator 1 denotes 1 wavelength, it's dimensionless, and the denominator T is the time that a wavelength pass through a point; when you wait time T , the number of wavelength goes through is 1, their ratio gives the frequency. And v/λ can also be understood: v is distance of the wave passes the point per unit time, when it's divided by the wavelength, it just counts the number of wavelengths passes through that point per unit time.

The phase frequency means how much phase of a space point varies when a unit time passes. It's just 2π times the frequency, since when a wavelength passes through, a full cycle in phase space is completed. The phase frequency is used more often than the frequency, thus we always call ω the frequency.

A pair of spatial quantities are the wave number and the phase wave number, they're given as

$$\begin{cases} \tilde{k} := \frac{1}{\lambda} \\ k := 2\pi \tilde{k} = \frac{2\pi}{\lambda}. \end{cases} \quad (260)$$

The wave number counts how many wavelength passes through a point per unit space, so the definition is natural as a wavelength passes through, the number of wavelength is "1", thus the numerator is dimensionless. The phase wave number shows how much phase of a certain space point it goes through when a unit length on

the wave passes, thus the definition is also natural: when a wavelength passes through, the phase varies by completing a circle which is 2π . The wave number is already somehow abandoned and we just call k the wave number.

By comparison, the temporal quantity ω is connected with the spatial quantity k as

$$\omega = vk, \quad (261)$$

which is also easy to understand intuitively: the phase change per unit time is equal to the distance it travels per unit time multiplying the phase change in unit spatial distance. This equation is usually called the dispersion relation, since generally to each wave number k (or wavelength λ), there is a specific velocity attached. Thus the velocity is a function of k , and so is the frequency ω ,

$$\omega(k) = v(k)k. \quad (262)$$

Since $\omega(k)$ and $v(k)$ are not independent, if we happen to know one of them, we can calculate the other by this equation. In most cases by certain physical laws, we somehow know the $\omega(k)$ first, and because the velocity here is attached to a specific wavenumber, it's called the phase velocity,

$$v_p(k) = \frac{\omega(k)}{k}. \quad (263)$$

A monochromatic wave has a single k hence the single $\omega(k)$, and due to its periodic property, it's always written as $\sin(kx - \omega t)$, $\cos(kx - \omega t)$, or $e^{i(kx - \omega t)}$. Note that in these forms, the argument $kx - \omega t$ has the dimension of the phase: k is the phase change per unit space and x is just the space, ω is the phase change per unit time and t is just the time. Since the wave as a whole travels at a constant velocity, you can just pick a point on it and see its velocity, the point you are tracing has a constant phase so that the wave function value doesn't change when it's moving. So for this specific point, we have

$$kx - \omega t = const. \implies kx = \omega t + const, \quad (264)$$

then we have its velocity just as

$$v = \frac{dx}{dt} = \frac{\omega}{k} = v_p, \quad (265)$$

which is nothing but the phase velocity, so in this sense the name of the phase velocity is clear: it's the "constant phase velocity".

Usually a general wave packet $p(kx - \omega t)$ is composed of waves of a spectrum of frequencies, and it can be Fourier decomposed to

$$p(x, t) = \int_{-\infty}^{\infty} dk A(k) e^{i(kx - \omega(k)t)}, \quad (266)$$

where $A(k)$ is the amplitude of each wavenumber, and $\omega(k)$ is a function of k .

When it forms a stable wave packet, it has a velocity as a whole, which we call it a group velocity. To assume the wave packet is almost monochromatic, so that the amplitude $A(k)$ is sharply peaked around a central wavenumber k_0 . In this region, the frequency can be expanded as

$$\omega(k) \approx \omega_0 + (k - k_0)\omega'_0, \quad (267)$$

where

$$\omega_0 = \omega(k_0) \quad \text{and} \quad \omega'_0 = \left. \frac{\partial \omega(k)}{\partial k} \right|_{k=k_0}. \quad (268)$$

Then, after some algebra,

$$p(x, t) = e^{i(k_0 x - \omega_0 t)} \int_{-\infty}^{\infty} dk A(k) e^{i(k - k_0)(x - \omega'_0 t)}. \quad (269)$$

There are two factors in this expression. The first factor, $e^{i(k_0 x - \omega_0 t)}$, describes a perfect monochromatic wave with wavenumber k_0 , with peaks and troughs moving at the phase velocity ω_0/k_0 within the envelope of the wave packet. The other factor,

$$\int_{-\infty}^{\infty} dk A(k) e^{i(k - k_0)(x - \omega'_0 t)}, \quad (270)$$

gives the envelope of the wave packet. This envelope function depends on position and time only through the combination $(x - \omega'_0 t)$. This combination is kept constant at a point when the packet travels, $(x - \omega'_0 t) = const$. Therefore, the envelope of the wave packet travels at velocity

$$v_g = \frac{dx}{dt} = \omega'_0 = (d\omega/dk)_{k=k_0}, \quad (271)$$

which explains the group velocity formula. And yet there's still a simplified version of analyzing this, if we just take two waves with slightly different wave numbers k and k' , as $\sin(kx - \omega(k)t)$, and $\sin(k'x - \omega'(k')t)$. When they form an envelope, any point on it must have a constant phase, and this phase is the phase difference of the two component wave, as $(k'x - \omega'(k')t) - (kx - \omega k t) = constant phase$. The place where this constant phase is 0, is the place where the two waves reinforce each other, and where this phase is π , there the two waves cancel each other. Thus the group velocity is easily calculated as

$$v_g = \frac{dx}{dt} = \frac{\omega' - \omega}{k' - k} = \frac{d\omega}{dk}. \quad (272)$$

Combining the expressions of the group and phase velocity, we can have the relation between them, which is derived as

$$\begin{aligned} v_g &= \frac{d\omega(k)}{dk} = \frac{d(v_p(k)k)}{dk} = \boxed{v_p(k) + k \frac{dv_p(k)}{dk}} = v_p + \frac{\omega}{v_p} \frac{dv_p}{d\omega} v_g \\ \implies v_g &= \frac{v_p}{1 - \frac{\omega}{v_p} \frac{dv_p}{d\omega}}. \end{aligned} \quad (273)$$

So if $\frac{dv_p}{d\omega} = 0$, i.e, the phase velocity is irrelevant to the frequency, the group velocity is equal to the phase velocity, then there's no dispersion; if $\frac{dv_p}{d\omega} < 0$, i.e, the phase velocity decreases as the frequency increases, the group velocity is smaller than the phase velocity, and this case is called the normal dispersion; if $\frac{dv_p}{d\omega} > 0$, i.e, the phase velocity increases as the frequency increases, the group velocity is greater than the phase velocity, and this case is called the a-normal dispersion.

4.2 Wave properties of particles and different dispersion relations

Previously we said that generally we know the dispersion relation $\omega(k)$ first and then we can calculate the phase and group velocity. This is possibly due to the fact that in quantum mechanics, the energy and momentum is associated with the frequency and the wave number by the particle-wave duality relations as

$$\begin{cases} E = \hbar\omega = 2\pi\hbar f = hf \\ p = \hbar k = \frac{2\pi\hbar}{\lambda} = \frac{h}{\lambda}, \end{cases} \quad (274)$$

and we have the relations between the energy and the momentum in different contexts.

In non-relativistic case, for a free particle dual to a Schrodinger wave, the energy relation and the resulting dispersion relation is given by

$$E = \frac{p^2}{2m} \implies \omega = \frac{\hbar k^2}{2m}, \quad (275)$$

thus we have the phase and the group velocity as

$$\begin{cases} v_p = \frac{\omega}{k} = \frac{\hbar k}{2m} = \frac{p}{2m} \\ v_g = \frac{d\omega}{dk} = \frac{\hbar k}{m} = \frac{p}{m}. \end{cases} \quad (276)$$

We see that the group velocity is consistent with the particle's velocity p/m while the phase velocity isn't. The above relation can also be written in terms of f and λ as

$$hf = \frac{h^2}{2m\lambda^2} \implies f = \frac{h}{2m}\lambda^{-2} \implies v_p = \frac{h}{2m}\lambda^{-1}. \quad (277)$$

If the energy is somehow shifted by a constant C , then the relations would be

$$E = \frac{p^2}{2m} + C \implies \omega = \frac{\hbar k^2}{2m} + \frac{C}{\hbar}, \quad (278)$$

and the resulting two velocities are

$$\begin{cases} v_p = \frac{\omega}{k} = \frac{\hbar k}{2m} + \frac{C}{\hbar k} = \frac{p}{2m} + \frac{C}{p} \\ v_g = \frac{d\omega}{dk} = \frac{\hbar k}{m} = \frac{p}{m}. \end{cases} \quad (279)$$

We see that the phase velocity depends on C while the group velocity doesn't, and the group velocity still coincides with the particle's velocity. In this sense we may guess that the phase velocity is somehow unphysical and doesn't influence the real particle's motion. To explicitly see this, we recall that a time-dependent field operator is given as

$$\psi = \int dk a_k e^{i(kx - \omega t)}, \quad (280)$$

when the energy is shifted by a constant, the new field operator would be

$$\psi \rightarrow \psi' = \int dk a_k e^{i(kx - (\omega + C)t)} = e^{-iCt} \psi, \quad (281)$$

which is nothing but a phase change of the original field thus would not change the physics. So we concluded that the change of the phase velocity does not influence the physics and it's unmeasurable.

For a relativistic massive particle, the energy-momentum relation and the dispersion relation are given as

$$E = \sqrt{p^2 + m^2} \implies \omega = \sqrt{k^2 + m^2/\hbar^2}, \quad (282)$$

where we choose the speed of light $c = 1$. So the resulting velocities are

$$\begin{cases} v_p = \frac{\omega}{k} = \sqrt{1 + m^2/(\hbar k)^2} = \sqrt{1 + m^2/p^2} > 1 \\ v_g = \frac{d\omega}{dk} = \frac{1}{\sqrt{1 + m^2/(\hbar k)^2}} = \frac{k}{\omega} < 1. \end{cases} \quad (283)$$

We see that the unphysical phase velocity is greater than the speed of light, while the group velocity still keeps slower than the light.

For a relativistic massless particle like a photon, the energy relation and the dispersion is given as

$$E = p \implies \omega = k, \quad (284)$$

so we have

$$v_g = v_p = 1, \quad (285)$$

where the group and the phase velocity is the same, all equal to the speed of light.

4.3 The wave equations

The simplest wave equations are

$$\begin{cases} \frac{\partial\psi(x,t)}{\partial x} = -\frac{\partial\psi(x,t)}{\partial t} & \iff \frac{\partial\psi(x,t)}{\partial x} + \frac{\partial\psi(x,t)}{\partial t} = 0 \iff \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right)\psi(x,t) = 0 \\ \frac{\partial\psi(x,t)}{\partial x} = \frac{\partial\psi(x,t)}{\partial t} & \iff \frac{\partial\psi(x,t)}{\partial x} - \frac{\partial\psi(x,t)}{\partial t} = 0 \iff \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right)\psi(x,t) = 0. \end{cases} \quad (286)$$

In order to explicitly solve them, we induce a pair of new variables (ξ, η) instead of the old pair (x, t) , their transformation relations are given by

$$\begin{cases} \xi = x + t \\ \eta = x - t \end{cases} \implies \begin{cases} x = \frac{\xi + \eta}{2} \\ t = \frac{\xi - \eta}{2} \end{cases} \implies \begin{cases} \frac{\partial}{\partial\xi} = \frac{\partial x}{\partial\xi} \frac{\partial}{\partial x} + \frac{\partial t}{\partial\xi} \frac{\partial}{\partial t} = \frac{1}{2} \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial t} \right) \\ \frac{\partial}{\partial\eta} = \frac{\partial x}{\partial\eta} \frac{\partial}{\partial x} + \frac{\partial t}{\partial\eta} \frac{\partial}{\partial t} = \frac{1}{2} \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial t} \right), \end{cases} \quad (287)$$

thus the equations we want to solve become

$$\begin{cases} \frac{\partial}{\partial\xi} \psi(\xi, \eta) = 0 \\ \frac{\partial}{\partial\eta} \psi(\xi, \eta) = 0. \end{cases} \quad (288)$$

The first equation doesn't contain η and the second doesn't contain ξ , so the solutions have the form of

$$\begin{cases} \psi(\xi, \eta) = \psi(\eta) = \psi(x - t) \\ \psi(\xi, \eta) = \psi(\xi) = \psi(x + t). \end{cases} \quad (289)$$

We can check them by differentiate them with respect to x and t ,

$$\begin{cases} \frac{\partial\psi(x-t)}{\partial x} + \frac{\partial\psi(x-t)}{\partial x} = \frac{\partial\psi(x-t)}{\partial(x-t)} \frac{\partial(x-t)}{\partial x} + \frac{\partial\psi(x-t)}{\partial(x-t)} \frac{\partial(x-t)}{\partial t} = \frac{\partial\psi(x-t)}{\partial(x-t)} - \frac{\partial\psi(x-t)}{\partial(x-t)} = 0 \\ \frac{\partial\psi(x+t)}{\partial x} - \frac{\partial\psi(x+t)}{\partial x} = \frac{\partial\psi(x+t)}{\partial(x+t)} \frac{\partial(x+t)}{\partial x} - \frac{\partial\psi(x+t)}{\partial(x+t)} \frac{\partial(x+t)}{\partial t} = \frac{\partial\psi(x+t)}{\partial(x+t)} - \frac{\partial\psi(x+t)}{\partial(x+t)} = 0. \end{cases} \quad (290)$$

The first equation will describe a right-moving function, there are three ways to prove it. The first approach is that if we choose a point on the curve and observe its motion, we see that its function values doesn't change so that its argument would keep a constant as $x - t = const$, in this way we may find its velocity as $v = \frac{dx}{dt} = 1 > 0$, so its traveling to the right. The same reasoning goes with the second equation which describes a left-moving function.

The second way is to intuitively understand the equation $\frac{\partial\psi(x,t)}{\partial x} = -\frac{\partial\psi(x,t)}{\partial t}$, to see why the partial derivatives differ by a minus sign describes a right-moving wave. We draw the picture as follows.

The third way is to treat the partial derivative as the quantum operators, with $i\frac{\partial}{\partial t}$ the frequency operator and $-i\frac{\partial}{\partial x}$ the wave number operator. And if the solution of the first equation is one of its eigenfunctions, it'll give $\omega = k$, which means $v_g = \frac{d\omega}{dk} = 1 > 0$, it's right-moving. And the second equation will give $\omega = -k$, then it's left-moving. The picture is drawn as follows.

The second-order differential equation which can describe both the left- and right- moving particles is given by

$$\frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2\psi}{\partial t^2} = 0 \iff \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right) \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right) \psi(x, t) = 0 \iff \frac{\partial^2}{\partial\xi\partial\eta} \psi(\xi, \eta) = 0. \quad (291)$$

It's easy to see the solution is given by linear combination of function of only ξ and function of only η as

$$\psi(\xi, \eta) = \psi_1(\xi) + \psi_2(\eta) = \psi_1(x + t) + \psi_2(x - t). \quad (292)$$

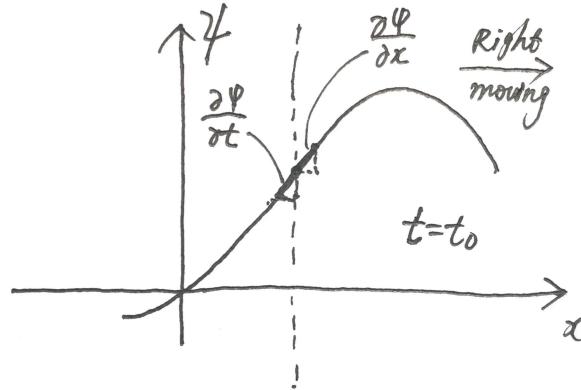


Figure 1: Intuitively understand the first-order partial differential equations of the wave function

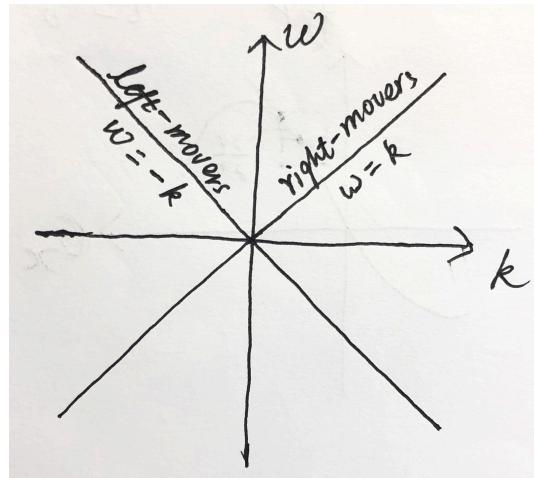


Figure 2: ω - k diagram for left- and right-movers

5 Relativistic Quantum Mechanics and Quantum Field Theory

5.1 Dirac equation in one-dimension and the negative energy problem

When dealing the case of particle moving very fast, comparable with the speed of light, we need a relativistic version of the quantum mechanics. In special relativity, the relation between the energy, momentum and the mass is given by

$$E^2 = p^2 + m^2 \quad (293)$$

When we replace the quantity to the corresponding operators as

$$\begin{cases} E & \rightarrow i \frac{\partial}{\partial t} \\ p & \rightarrow -i \frac{\partial}{\partial x}, \end{cases} \quad (294)$$

we have

$$\frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial x^2} = m^2 \psi, \quad (295)$$

which is the Klein-Gorden equation. But instead of a second order differential equation, we want a first order one, which is relevantly more difficult.

To begin with, we consider a massless particle moving with the speed of light in one dimension, the energy relations and the dispersion relation read as

$$E = \pm p \iff \omega = \pm k \quad (296)$$

which can represent right moving particles with the plus sign, or left moving particles with the minus sign; and no matter which direction the particle is moving, the momentum and the wave number itself can be positive or negative. We stick with right-moving particles, and with the substitution we have

$$i \frac{\partial \psi}{\partial t} = \mp i \frac{\partial \psi}{\partial x} \quad \text{or} \quad \frac{\partial \psi}{\partial t} \pm \frac{\partial \psi}{\partial x} = 0, \quad (297)$$

the solution has the form of $\psi(x \mp t)$ describes a right/left moving wave function with the speed of light 1. Dirac treat the right/left moving property as a degree of freedom, thus any wave function has two components representing right/left moving just as the spin. In this treatment, the above equations can be composed into a single one in matrix version

$$i \frac{\partial}{\partial t} \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} = -i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial}{\partial x} \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}, \quad (298)$$

in this way, the Hamiltonian is given in terms of p as

$$H = \alpha p \quad \text{with} \quad \alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (299)$$

where α is the third Pauli matrix. When it acts on the right-moving state, it gives the eigenvalue of 1, which can also mean the spin is pointing right if we treat α as σ_x ; when it acts on the left-moving state, it gives the eigenvalue of -1 , which can also mean the spin is pointing left,

$$\begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \psi_R \\ 0 \end{pmatrix} = \begin{pmatrix} \psi_R \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ \psi_L \end{pmatrix} = -\begin{pmatrix} 0 \\ \psi_L \end{pmatrix}, \end{cases} \quad (300)$$

In this sense, the spin naturally arises by mixing the right- and left- moving particles, and what's more, the spin is locked with the moving direction. Thus if we treat the αp as a quantity known as the handedness, it is the same with the left- and right- moving particles! We call this property of states satisfying

$$H\psi = \alpha p \psi \iff i \frac{\partial}{\partial t} \psi = -i\alpha \frac{\partial}{\partial x} \psi \iff \frac{\partial}{\partial t} \psi + \alpha \frac{\partial}{\partial x} \psi = 0, \quad (301)$$

the right-handedness. And on the contrary the states satisfying

$$H\psi = -\alpha p \psi \iff i \frac{\partial}{\partial t} \psi = i\alpha \frac{\partial}{\partial x} \psi \iff \frac{\partial}{\partial t} \psi - \alpha \frac{\partial}{\partial x} \psi = 0 \quad (302)$$

are left-handed.

Then we proceed to construct the Hamiltonian, by adding a mass term as,

$$H = \alpha p + \beta m, \quad (303)$$

the new symbol β denoting also a matrix due to the consistency of the matrix form. If we square it, we should recover the energy relation of the massive particle eq (293), which requires

$$\begin{cases} \alpha^2 = 1 \\ \beta^2 = 1 \\ \alpha\beta + \beta\alpha = 0, \end{cases} \quad (304)$$

and we know the β matrix which satisfies this can be any of others Pauli matrices. We set it to be

$$\beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (305)$$

Now we apply this new first order Hamiltonian to a two-component state to get

$$i\frac{\partial}{\partial t}\psi = -i\alpha\frac{\partial}{\partial x}\psi + \beta m\psi, \quad (306)$$

or explicitly in its matrix form

$$i\frac{\partial}{\partial t}\begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} = -i\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} + m\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}, \quad (307)$$

which can be separately written as

$$\begin{cases} i\frac{\partial\psi_R}{\partial t} + i\frac{\partial\psi_R}{\partial x} = m\psi_L \\ i\frac{\partial\psi_L}{\partial t} - i\frac{\partial\psi_L}{\partial x} = m\psi_R. \end{cases} \quad (308)$$

We see that when we introduce the mass term, the two components of the state couple with each other, they're not independent any more. Or to say inversely, the coupling of the two components of the state gives the mass of the particle, which makes the particle move slower than the light. This is consistent with physical significance of a massive particle — the right moving particle can be transformed into the left moving particle in a reference frame which moves faster than the particle, while the massless particle moving in the speed of light doesn't have this property since no reference frame can move faster than the light.

Next we come to the problem of the negative energy.

The vacuum is the state that all the negative energy states are filled with fermions like electrons, which are called the Dirac sea. When a photon hits the vacuum, the energy excites an electron at a negative energy level to a positive energy level, thus leaving a hole in the negative energy sea. The effect of the hole is in general to have a positive charge in the sea, which we call a positron. This is the mechanism of how a photon creates a electron/positron pair. (figure to be added)

The Dirac field operator is constructed as

$$\begin{aligned} \Psi(x) &= \int_{-\infty}^{\infty} dp a(p)e^{ipx} \\ &= \int_0^{\infty} dp a(p)e^{ipx} + \int_{-\infty}^0 dp a(p)e^{ipx} \\ &= \int_0^{\infty} dp a(p)e^{ipx} + \int_0^{\infty} dp' b^\dagger(p')e^{-ip'x}, \end{aligned} \quad (309)$$

where annihilation operators for electrons can be seen as creation operators for positions. Since the annihilation and creation operators for fermions are on the same footing, the creation operators would take the same form.

In this sense, a term like $\Psi^\dagger\Psi A$, where A represents the photon field operator, contains four separate terms. Thus the original single term could contain multiple processes as: an electron emits a photon, an electron and a positron annihilate and create a photon, the vacuum spontaneously create an electron and a positron and a photon, etc. (Figures of Feynman diagrams of the three processes to be added.)

5.2 Dirac equation in three-dimension and chirality

Now we generalize our relativistic equation into 3-dimensions. Still, we study massless particles and then add the mass term. In 3D case, the momentum must have three components, thus it's a vector; but the energy remains a scalar which has rotation invariance. This requires the α matrix also has three component, so as to get a

“dot product” with the momentum to get an invariant scalar, in this sense α can also be treated as a “vector”. Therefore the Hamiltonian can be generalized to

$$H = \vec{\alpha} \cdot \vec{p} = \alpha_x p_x + \alpha_y p_y + \alpha_z p_z, \quad (310)$$

and we want this generalization to satisfy the good old energy-momentum relation,

$$H^2 = \vec{p}^2 = p_x^2 + p_y^2 + p_z^2, \quad (311)$$

which requires the α matrices to have the property as

$$\begin{cases} \alpha_i^2 = 1 \\ \{\alpha_i, \alpha_j\} = 0 \quad \text{for } i \neq j, \end{cases} \quad (312)$$

or to be summarized as

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}. \quad (313)$$

There are three matrices which can satisfy all the above conditions, Pauli matrices; thus we naturally attribute α matrices as $\alpha_i = \sigma_i$, hence the Hamiltonian is given by

$$H = \vec{\sigma} \cdot \vec{p}. \quad (314)$$

Since the Pauli matrices are attached to the spin, this equation would serve as the origin of the spin. we don’t see spin in 1-dimension, we only see spin-like things as the left/right moving property; spin is truly seen in 3-dimension. Additionally, in this form the energy is proportional to the component of the spin along the direction of motion. But we know the component of the spin is always ± 1 , so in one case the spin is pointing along the momentum giving positive energy, and in the other case the spin is opposite to the momentum giving the negative energy. Again, we restate that, for a massless particle with positive energy, the spin is always locked to the direction of the momentum.

And then we follow the old steps to add the mass term multiplied by a matrix which has the same dimension with that of σ_i ,

$$H = \vec{\sigma} \cdot \vec{p} + \beta m, \quad (315)$$

and make it to satisfy

$$H^2 = p^2 + m^2, \quad (316)$$

which additionally requires the β matrix to satisfy

$$\begin{cases} \beta^2 = 1 \\ \{\beta, \sigma_i\} = 0. \end{cases} \quad (317)$$

But unfortunately, there’s not a forth 2×2 matrix beyond the three Pauli matrices which could both squares to be 1 and anti-commutes with the σ_i . Nor can we find four 3×3 matrices which can simultaneously meet all the requirements. The minimal dimension we can find such four matrices will be 4×4 . These matrices are called **Dirac matrices**. Incidentally, there are many representations of Dirac matrices, we just have to pick one and stick with it. Thus the wave function would have two more degrees of freedom, in total four components, together known as a **spinor**. But note that these four components have nothing to do with the $3+1$ dimensional spacetime, their existence is just a mathematical result. In 4D spinor space, we rewrite the Hamiltonian once more as

$$H = \vec{\alpha} \cdot \vec{p} + \beta m, \quad (318)$$

and one set of 4×4 matrices called the **Weyl representation** of the Dirac matrices are given as

$$\vec{\alpha} = \begin{pmatrix} \vec{\sigma} & O \\ O & -\vec{\sigma} \end{pmatrix} \quad \beta = \begin{pmatrix} O & I \\ I & O \end{pmatrix}, \quad (319)$$

where O , I and $\vec{\sigma}$ are themselves 2×2 matrices. One can check them by squaring the Hamiltonian to give the energy-momentum relation.

Now we apply this $3D$ Hamiltonian to the four-component state to get

$$i \frac{\partial}{\partial t} \psi = -i\alpha_j \frac{\partial}{\partial x^j} \psi + \beta m \psi, \quad (320)$$

or explicitly in its block-matrix form

$$i \frac{\partial}{\partial t} \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} = -i \begin{pmatrix} \sigma_j & O \\ O & -\sigma_j \end{pmatrix} \frac{\partial}{\partial x^j} \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} + m \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}, \quad (321)$$

where ψ_R and ψ_L are two-component vector states. Then this equation can be separately written as

$$\begin{cases} i \left(\frac{\partial \psi_R}{\partial t} + \sigma_j \frac{\partial \psi_R}{\partial x^j} \right) = m \psi_L \\ i \left(\frac{\partial \psi_L}{\partial t} - \sigma_j \frac{\partial \psi_L}{\partial x^j} \right) = m \psi_R. \end{cases} \quad (322)$$

A property describes the relations between the direction of the spin and the momentum is called the **chirality**. A particle with the spin and the momentum in the same direction is *right-handed*, a right-handed particle has both positive energy and negative energy components; a particle with the spin and the momentum in the opposite direction is *left-handed*, a left-handed particle has also both positive energy and negative energy components. To leave the negative energy away for the moment, we restate that for positive energy particles, they can be classified into right-handed and left-handed ones. When the two chiral components are coupled together by the β matrix, the mass term of the particle comes out; or to say, the effect of the mass of a Dirac particle is to make transition of the two chiral particles; this is consistent with the fact that one can choose a reference frame which moves faster than the massive particle and sees the momentum pointing in the opposite direction (with spin direction unchanged), thus from his point of view the right-handedness becomes left-handedness.

In this sense, we conclude that for massive particles, chirality depends on the reference frame, and a particle at rest is a mixture of the LH and RH. To see this clearly, we choose a reference frame where the momentum $\vec{p} = 0$, in this way we have $H = \beta m$, then the spatial derivative terms vanish, from equation (325) we have

$$\begin{cases} i \frac{\partial \psi_R}{\partial t} = m \psi_L \\ i \frac{\partial \psi_L}{\partial t} = m \psi_R. \end{cases} \quad (323)$$

If we add or subtract these two equations and define two new states as $\psi_{\pm} = \psi_R \pm \psi_L$, we have

$$\begin{cases} i \frac{\partial \psi_+}{\partial t} = i \frac{\partial}{\partial t} (\psi_R + \psi_L) = m(\psi_R + \psi_L) = m \psi_+ \\ i \frac{\partial \psi_-}{\partial t} = i \frac{\partial}{\partial t} (\psi_R - \psi_L) = m(\psi_R - \psi_L) = -m \psi_-. \end{cases} \quad (324)$$

Since the $i\partial/\partial t$ is the Hamiltonian operator, when it acts on a state it gives the energy as the eigenvalue, so for ψ_+ we have $E(\psi_+) = \omega(\psi_+) = m > 0$ which is a positive energy, and for ψ_- we have $E(\psi_-) = \omega(\psi_-) = -m < 0$ which is a negative energy. Therefore for both the positive and negative energy states of a rest particle, they are mixtures of the LH and RH.

On contrast, for a massless particle the chirality is conserved, it stays right-handed or left-handed forever, which is also consistent with the fact that nothing can move faster than light and change the direction of the momentum. To see this conservation, we write

$$\begin{cases} i \left(\frac{\partial \psi_R}{\partial t} + \sigma_j \frac{\partial \psi_R}{\partial x^j} \right) = 0 \\ i \left(\frac{\partial \psi_L}{\partial t} - \sigma_j \frac{\partial \psi_L}{\partial x^j} \right) = 0 \end{cases} \implies \begin{cases} (H - \vec{\sigma} \cdot \vec{p}) \psi_R = 0 \\ (H + \vec{\sigma} \cdot \vec{p}) \psi_L = 0, \end{cases} \quad (325)$$

which satisfies with the handedness equations we discussed in one dimension.

Besides, chirality are also well-defined and useful in context of high-energy particles where the momentum is large enough.

Next we're going to leave the dot product away and study the $\vec{\alpha}$ matrices alone. It's physical meaning, or to say, its expectation value in any state, is the velocity of the particle! This seems very bizarre, but we will prove it conversely, by using the commutator of the position with the Hamiltonian,

$$\frac{d}{dt}\langle x \rangle = \frac{i}{\hbar} \langle [H, x] \rangle = \frac{i}{\hbar} \langle [\vec{\alpha} \cdot \vec{p} + \beta m, x] \rangle = \frac{i}{\hbar} \langle \alpha_x [p_x, x] \rangle = \frac{i}{\hbar} \langle \alpha_x \rangle (-i\hbar) = \langle \alpha_x \rangle, \quad (326)$$

which holds not only for x -component, but also for y and z . Thus we have

$$\vec{v} = \frac{d}{dt}\langle \vec{r} \rangle = \langle \vec{\alpha} \rangle. \quad (327)$$

The velocity is the average of the α matrix, not the momentum like in classical mechanics! So if you want to know what the velocity is, you measure α .

On the other hand, the α itself is not conserved. This is easy to be seen in the fact that the time derivative of α is proportional to the commutator of the Hamiltonian and the α as $[H, \alpha_x]$, since H contains α_y and α_z , this commutator would be non-zero. This means that the particle is accelerating. This does not violate the Newton's law, since firstly Newton's law says the time derivative of the momentum, not the derivative of the velocity, is zero; secondly the time average of the velocity over a certain period is indeed zero, which says the particle moves in a crazy manner in every instant, but in general it moves along a line.

Besides, since α is the velocity, it's vector quantity; it's not spin, which is an axial vector, or pseudo-vector quantity. We list the similar quantities as below

$$\begin{aligned} \text{Velocity: } \vec{\alpha} &= \begin{pmatrix} \vec{\sigma} & O \\ O & -\vec{\sigma} \end{pmatrix} \\ \text{Spin: } \vec{\Sigma} &= \begin{pmatrix} \vec{\sigma} & O \\ O & \vec{\sigma} \end{pmatrix} \\ \text{Chirality: } \vec{\sigma} \cdot \vec{p}. \end{aligned} \quad (328)$$

A photon injects into a vacuum and creates a hole in the Dirac negative sea and an electron, both of which have positive energy ϵ , or ω and momentum p , or k .

5.3 The Dirac Lagrangian and the Covariant Dirac equation

Next we are going to write the Lagrangian of the Dirac equation. To start from the 3D Dirac equation,

$$i\partial_0\psi = -i\alpha_j\partial_j\psi + \beta m\psi, \quad (329)$$

where we used the spacetime indices to denote the derivatives. By moving terms we have

$$(i\partial_0 + i\alpha_j\partial_j - \beta m)\psi = 0, \quad (330)$$

The Dirac Lagrangian is given by multiplying the LHS of the above by the Hermitian conjugate of the the Dirac field ψ^\dagger as

$$\mathcal{L} = \psi^\dagger(i\partial_0 + i\alpha_j\partial_j - \beta m)\psi. \quad (331)$$

Then the Dirac equation is recovered by EL equation with respect to ψ^\dagger as

$$\frac{\partial \mathcal{L}}{\partial \psi^\dagger} = \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi^\dagger)}, \quad (332)$$

where the RHS vanishes since there's no derivative terms with respect to ψ^\dagger , and we have

$$\frac{\partial \mathcal{L}}{\partial \psi^\dagger} = (i\partial_0 + i\alpha_j\partial_j - \beta m)\psi = 0. \quad (333)$$

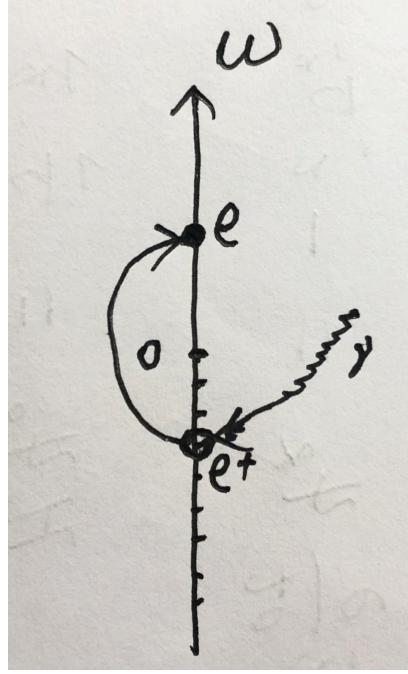


Figure 3: A photon injects into a vacuum and creates an electron/positron pair

Next we try to make the Dirac equation and its Lagrangian a more neat and modern way. We multiply the Dirac equation by the matrix β on the left and get

$$(i\beta\partial_0 + i\beta\alpha_j\partial_j - m)\psi = 0, \quad (334)$$

where we used the fact that $\beta^2 = I$ on the mass term. Then we rename the matrices as

$$\begin{cases} \gamma^0 \equiv \beta \\ \gamma^j \equiv \beta\alpha_j, \end{cases} \quad (335)$$

and using these Gamma matrices we write the equation as

$$(i\gamma^0\partial_0 + i\gamma^j\partial_j - m)\psi = 0 \implies (i\gamma^\mu\partial_\mu - m)\psi = 0, \quad (336)$$

where summation over repeated indices is implied. This is the standard modern form of the Dirac equation.

The Dirac Lagrangian with this notation is given by multiplying the LHS of the above by a newly defined vector $\bar{\psi} \equiv \psi^\dagger\beta$ as

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi. \quad (337)$$

The Dirac equation can be gain through the Euler-Lagrangian equation with respect to $\bar{\psi}$ as

$$\frac{\partial \mathcal{L}}{\partial \bar{\psi}} = \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \bar{\psi})}, \quad (338)$$

where the RHS is again zero, so we have

$$\frac{\partial \mathcal{L}}{\partial \bar{\psi}} = (i\gamma^\mu\partial_\mu - m)\psi = 0. \quad (339)$$

The mathematical mechanism of $\bar{\psi}$ is that it can be converted back to ψ^\dagger by right multiplying β as $\bar{\psi}\beta = \psi^\dagger$, thus it compensate the β we already multiplied in.

We know from the previous lectures that a derivative term in the Lagrangian would have the effect of absorbing a particle at one spacetime point and emitting it at another point, thus making the particle move; while the non-derivative term means absorbing the particle at a point and emitting it at the same point, thus giving it some properties like the mass and charge.

Then we analysis each term of the Dirac Lagrangian. By comparing the two forms of it, we have the first term as

$$\begin{aligned}
i\bar{\psi}\gamma^\mu\partial_\mu\psi &= i\psi^\dagger\beta(\beta\partial_0 + \beta\alpha_j\partial_j)\psi = i\psi^\dagger(\partial_0 + \alpha_j\partial_j)\psi \\
&= i(\psi_L^\dagger, \psi_R^\dagger) \begin{pmatrix} \partial_0\psi_L \\ \partial_0\psi_R \end{pmatrix} + i(\psi_L^\dagger, \psi_R^\dagger) \begin{pmatrix} \sigma_j & O \\ O & -\sigma_j \end{pmatrix} \begin{pmatrix} \partial_j\psi_L \\ \partial_j\psi_R \end{pmatrix} \\
&= i(\psi_L^\dagger\partial_0\psi_L + \psi_R^\dagger\partial_0\psi_R) + i(\psi_L^\dagger\sigma_j\partial_j\psi_L - \psi_R^\dagger\sigma_j\partial_j\psi_R) \\
&\quad + i(\psi_L^\dagger\partial_0\psi_L + \psi_L^\dagger\sigma_j\partial_j\psi_L) + i(\psi_R^\dagger\partial_0\psi_R - \psi_R^\dagger\sigma_j\partial_j\psi_R),
\end{aligned} \tag{340}$$

where ψ_L and ψ_R themselves are two-component vectors. We see from this form that the LH and RH component of the field doesn't mix with each other. The technical reason is the β in $\bar{\psi}$ and the β in γ^μ cancel each other, thus the net effect is without the β matrix which could flip the handedness.

The second term is expanded as

$$m\bar{\psi}\psi = m\psi^\dagger\beta\psi = m(\psi_L^\dagger, \psi_R^\dagger) \begin{pmatrix} O & I \\ I & O \end{pmatrix} \begin{pmatrix} \psi^\dagger \\ \psi_R \end{pmatrix} = m(\psi_L^\dagger\psi_R + \psi_R^\dagger\psi_L), \tag{341}$$

it mixes the LH and RH components. The matrix that characterize chirality is defined as $\gamma^5 \equiv \gamma^0\gamma^1\gamma^2\gamma^3$, it squares to unity $(\gamma^5)^2 = I$, thus it has a positive eigenvalue $+1$ denoting the RH state and -1 corresponding to the LH state. By the way, the quantity $\bar{\psi}\gamma^\mu\psi$ is a 4-vector, and the $\bar{\psi}\psi$ is a scalar.

But in the weak interaction context, only the LH part is charged, thus we have to modify the mass term as

$$g\phi\bar{\psi}\psi = g\psi_L^\dagger\phi\psi_R + g\psi_R^\dagger\psi_L\phi, \tag{342}$$

where the ϕ transforms as ψ_L , and ϕ^* the opposite. When SSB occurs, the field ϕ sits in $f+H$, the term becomes

$$g\phi\bar{\psi}\psi = gf\phi\bar{\psi}\psi + gH\phi\bar{\psi}\psi, \tag{343}$$

where the first term gives the fermion mass, and the second term describes fermion emitting a Higgs or Higgs decaying into particle/anti-particle pairs, as following figures depicts. (Move this to SSB chapter.)

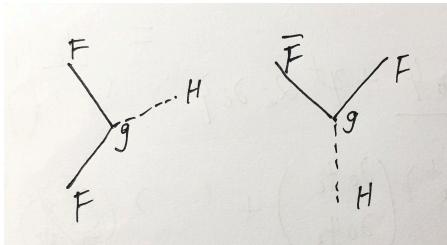


Figure 4: Interaction between Fermions and Higgs in context of weak interactions

6 The Standard Model of Particle Physics



6.1 Basics concepts of the Standard Model

From the group symmetry point of view, the standard model of particle physics has the symmetry of

$$SU(3) \times SU(2) \times U(1). \quad (345)$$

The $U(1)$ symmetry accounts for the electro-magnetic interaction, and the group has 1 generator corresponding to 1 gauge field (photon); the $SU(2)$ symmetry accounts for the weak interaction and has 3 generator corresponding to 3 gauge fields (W^\pm and Z); the $SU(3)$ symmetry accounts for the strong interaction and has 8 generator corresponding to 8 gauge fields (gluon). Every gauge field will have three features: 1, it has a source; 2, it has a conservation law corresponding to a symmetry; 3, it has a coupling constant which is the quantum amplitude.

As we know the electric charge is the coupling constant of e-m interaction amplitude, its square is the quantum probability for that process to happen. The fine-structure constant is $\alpha = \frac{e^2}{4\pi} \sim \frac{1}{137}$, it means you will need about 137 electrons to have one of them emit a photon. The fine structure constant governs everything about the magnitude of the an atom (in natural units $1eV^{-1} = 10^{-7}m$), like the atom diameter ($10^{-10}m \sim 1KeV^{-1}$), the time for light to pass an atom called the transit time ($10^{-18}s$), the electron orbital time ($10^{-16}s$), and the atom decay time ($10^{-9}s$).

Analog to e-m interaction amplitude, the strong interaction amplitude is approximately the same order $g_s \sim \frac{1}{20}$, and it controls the magnitude of every size of a hadron. The nuclei diameter $\sim 10^{-15}m$, the transit time $\sim 10^{-23}m$, the quark orbital time $\sim 10^{-23}s$, the hadron decay time or the oscillation period between hadrons and mesons $\sim 10^{-23}s$. We see from these numbers that there're no hierarchy scales among them.

6.2 $U(1)$ Gauge symmetry in QED

6.2.1 Global and local phase transformation symmetry and covariant derivatives

If a complex charged scalar field ϕ , and its complex conjugate transforms as

$$\begin{cases} \phi \rightarrow \Phi = e^{i\theta} \phi \\ \phi^* \rightarrow \Phi = e^{-i\theta} \phi^*, \end{cases} \quad (346)$$

where the phase factor θ is a constant throughout the whole spacetime, and the gauge field just transforms to it self,

$$A^\mu \rightarrow A^\mu, \quad (347)$$

then the Lagrangian which has the form

$$\mathcal{L} = \partial_\mu \phi^* \partial_\mu \phi - V(\phi^* \phi) + F_{\mu\nu} F^{\mu\nu} \quad (348)$$

will have the symmetry under the transformation. This is called the global $U(1)$ symmetry. The electron and positron are called the complex conjugate representations of $U(1)$ group.

If the above scalar field and its complex conjugate transform with a phase factor dependent on the spacetime coordinate x ,

$$\begin{cases} \phi \rightarrow \Phi = e^{i\theta(x)} \phi \\ \phi^* \rightarrow \Phi = e^{-i\theta(x)} \phi^*, \end{cases} \quad (349)$$

meaning that it transforms locally. In this case we want to know whether the symmetry invariance is still kept. To go on with it, the ordinary derivative of the new fields are given by

$$\begin{cases} \partial_\mu \Phi = (\partial_\mu \phi + i\phi \partial_\mu \theta) e^{i\theta(x)} \\ \partial_\mu \Phi^* = (\partial_\mu \phi^* - i\phi^* \partial_\mu \theta) e^{-i\theta(x)}, \end{cases} \quad (350)$$

thus the new inner product will be the old inner product plus some extra term, which is of course not invariant. Here we could introduce a gauge field (in this case the e-m field) and the covariant derivative to eliminate the extra terms. The transformations of the gauge field and the covariant derivative are given as follows

$$\begin{cases} A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu \theta \\ \partial_\mu \rightarrow D_\mu = \partial_\mu + iA_\mu, \end{cases} \quad (351)$$

and hence naturally the covariant derivative of the complex conjugate field is given by $D_\mu \phi^* = \partial_\mu - iA_\nu \theta$. With this definition, we could calculate the covariant derivatives of the new field and its complex field as,

$$\begin{cases} D_\mu \Phi = \partial_\mu \Phi + iA' \Phi = (\partial_\mu \phi + i\phi \partial_\mu \theta) e^{i\theta(x)} + i(A_\mu - \partial_\mu \theta) \phi e^{i\theta(x)} = (\partial_\mu \phi + iA_\mu \phi) e^{i\theta(x)} = (D_\mu \phi) e^{i\theta(x)} \\ D_\mu \Phi^* = \partial_\mu \Phi^* - iA' \Phi^* = \dots = (D_\mu \phi^*) e^{-i\theta(x)}, \end{cases} \quad (352)$$

only differing from the original derivative by a phase, thus can be canceled in the inner product, explicitly as

$$D_\mu \Phi^* D_\mu \Phi = D_\mu \phi^* D_\mu \phi. \quad (353)$$

Thus we reserve the gauge invariance of the kinetic term. And next the potential term is easy to be kept invariant only if it's a function of $\phi^* \phi$, so if the Lagrangian has the following form as

$$\mathcal{L} = D_\mu \phi^* D_\mu \phi - V(\phi^* \phi), \quad (354)$$

it'll be invariant under the gauge transformation.

But we still have to add the gauge field term involving $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, note that it's itself gauge invariant as

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = \partial_\mu A_\nu - \partial_\mu \partial_\nu \theta - \partial_\nu A_\mu + \partial_\nu \partial_\mu \theta = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}, \quad (355)$$

so the only thing to do is to make a scalar out of it as $F_{\mu\nu} F^{\mu\nu}$. Therefore we have a Lagrangian which has the local gauge symmetry, as

$$\mathcal{L} = D_\mu \phi^* D_\mu \phi - V(\phi^* \phi) + \frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (356)$$

where the factor of $\frac{1}{4}$ is due to convention (referring to classical field theory). Note that it's usually common to assign a mass to a field by a quadratic term of it, we may try to make a term as $\frac{m^2}{2} A^\mu A_\mu$ to make the gauge particle have the mass, but this term is itself not invariant under gauge transformation (also see the classical field theory), so it's not possible to add such a term; this is consistent with fact that photons don't have mass. In another way,

Gauge invariance is the symmetry which prohibits the photon from having a mass.

But mass of a gauge field can be given by another mechanism known as the *Spontaneous Symmetry Breaking*, as we will talk later.

6.2.2 Spontaneous Symmetry Breaking and the mass of a field

The concept of discrete Spontaneous Symmetry Breaking.

We begin with the case of a real scalar field which has only the simplest form, the kinetic term minus the potential term.

$$\mathcal{L} = \partial_\mu \phi \partial_\mu \phi - V(\phi), \quad (357)$$

If there's no potential term, the Lagrangian with only the gradient terms would indicate the field doesn't have mass, just like photons.

We know that a field state will stay at the stationary point of the potential, which has the lowest value of the potential term. Consider a potential term takes the form of a "Mexican hat" as

$$V(\phi) = -\frac{\mu^2}{2}\phi^2 + \frac{\lambda}{4}\phi^4, \quad (358)$$

where the parameter μ has the dimension of mass, and the λ which is dimensionless called a quartic coupling constant. The potential has a symmetric form where two lowest points exists, they're calculated as

$$V'(\phi) = -\mu^2\phi + \lambda\phi^3 = 0 \implies \phi^2 = \frac{\mu^2}{\lambda} \equiv f^2 \implies f = \pm \frac{\mu}{\sqrt{\lambda}}, \quad (359)$$

where f is dominated by μ , while λ is less important. By the way, for symmetry we only pick even powers of the field, and powers beyond the fourth would make the theory non re-normalizable, so we just neglect higher-power terms.

The figure of this potential is shown below (see Figure 5), the stable field would not be preferable to any of these two lowest point, but it'll randomly pick one of them to stay in. This is the phenomenon of *spontaneous symmetry breaking* (SSB). The contrary concept is, we name it, the explicit symmetry breaking, corresponding to the case where there are also two local lowest point in potential term but one is lower than the other. And since the two points are not connected or joint somehow, we say that this spontaneous symmetry breaking is discrete.

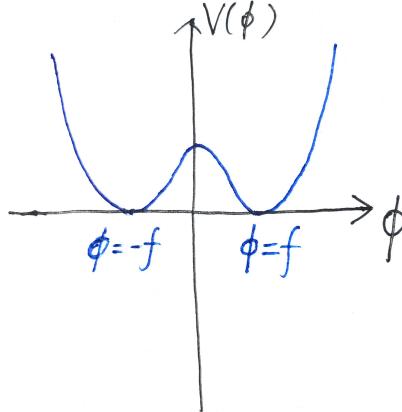


Figure 5: Symmetric potential term about real ϕ -axis

Consider a field which at a far left place the field locates at $\phi = -f$ while at a far right place locates at $\phi = f$, these field will influence their neighborhood and thus spread away out. Then what happens at the place where these two different values of field meet? The points at this place doesn't know whether they should have the value $\phi = -f$ or $\phi = f$ due to the difference of their neighbors, thus creating a phenomenon called the *domain wall* (see Figure), indicating that it's very hard to transit from the state in one side of this wall to the state in another side. The domain wall is a signature of the discrete spontaneous symmetry breaking.

Then we move to the case of continuous symmetry breaking, in which the Goldstone bosons replace the domain walls, and Higgs mechanism gives rise to the mass to photon-like particles (Z bosons).

In this case we consider a complex scalar field which has only the global U(1) phase symmetry. The Lagrangian is written as

$$\mathcal{L} = \partial_\mu \phi^* \partial_\mu \phi - V(\phi^* \phi), \quad (360)$$

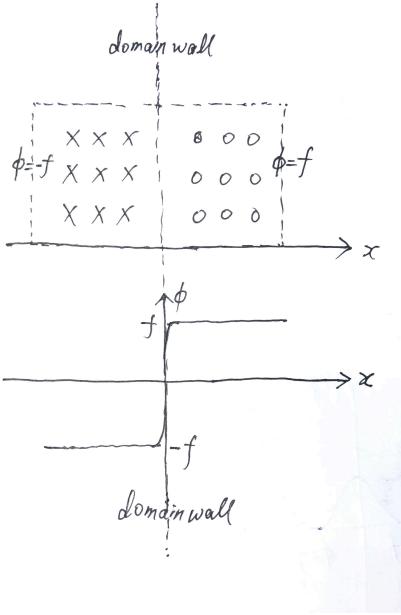


Figure 6: Schematic picture of the Domain Wall

where the kinetic term and the potential term have the same phase symmetry. The potential term $V(\phi^*\phi)$ is function in the complex ϕ -plane where the axis are ϕ_R and ϕ_I .

If the potential has the simple form of $V(\phi^*\phi) = \phi^*\phi$, the Lagrangian is re-expressed in terms of the real and imaginary part as

$$\mathcal{L} = (\partial\phi_R)^2 + (\partial\phi_I)^2 - (\phi_R^2 + \phi_I^2) = \mathcal{L}_R + \mathcal{L}_I, \quad (361)$$

where the motion of the real and imaginary field look like indepedent. The potential $V(\phi^*\phi) = \phi^*\phi$ has the shape as in Figure 7, which is the most trivial case.

If the potential has the form of $V(\phi^*\phi) = -a(\phi^*\phi) + b(\phi^*\phi)^2$, it'll has the shape as in Figure 8, which is symmetric about the vertical axis and has continuous circle of minimums, thus it has the continuous symmetry. If the field resits itself at some point of that continuous minimal circle, we say that this is the continuous symmetry breaking.

If the field at some place far left is at a in the ground state circle, and at some other place far right is at b in the ground state circle, then the field between them will gradually change to each other, varying continuously by a phase factor. In this case, the domain wall no longer exists, see Figure 9.

Now that the potential has the rotation symmetry in the complex ϕ -plane, it is more convenient to express the Lagrangian in terms of polar coordinates

$$\begin{cases} \phi = \rho e^{i\alpha} \\ \phi^* = \rho e^{-i\theta}, \end{cases} \quad (362)$$

thus the derivatives of them will be

$$\begin{cases} \partial_\mu \phi = (\partial_\mu \rho + i\rho \partial_\mu \alpha) e^{i\alpha} \\ \partial_\mu \phi^* = (\partial_\mu \rho - i\rho \partial_\mu \alpha) e^{-i\alpha}, \end{cases} \quad (363)$$

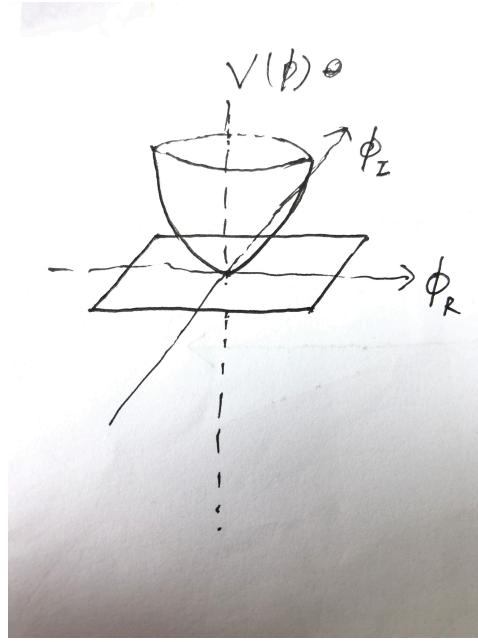


Figure 7: $V(\phi^*\phi) = \phi^*\phi$ diagram

and the Lagrangian is written as

$$\mathcal{L} = \partial_\mu \rho \partial^\mu \rho + \rho^2 \partial_\mu \alpha \partial^\mu \alpha - V(\rho^2). \quad (364)$$

In the ground state circle, the field has the value $\rho = f$, thus $\partial\rho = 0$ and $V(\rho) = \text{const}$ which can be dropped out; the only left term $f^2 \partial_\mu \alpha \partial^\mu \alpha$ can be rewritten in terms of a newly-defined field $\beta \equiv f\alpha$, hence the Lagrangian can be written as

$$\mathcal{L} = \partial_\mu \beta \partial^\mu \beta, \quad (365)$$

which contains only the gradient term and behaves like a field with no mass.

By varying α (or β), the field ϕ varies only within the ground state circle which costs very little energy, and this will give a Goldstone boson. Since in this way, the $\phi = \rho e^{i\alpha}$ itself transforms like a gauge transformation, thus the corresponding particle is called a Goldstone boson for a U(1) symmetry. If ϕ vibrates with respect to ρ , which costs bigger energy, this will give rise to a massive particle known as the Higgs. The above two mechanism give significance to both the ρ and α parts of the field.

At this time we want to check what will happen if we induce a gauge field A_μ . In this situation, the ordinary derivatives should naturally replaced by the covariant derivative, which are computed in the ground state case as

$$\phi = fe^{i\alpha} \implies \partial\phi = if\partial\alpha e^{i\alpha} \implies D\phi = ife^{i\alpha}(\partial\alpha + A) \implies D\phi^* = -ife^{-i\alpha}(\partial\alpha + A), \quad (366)$$

where we neglect the summation indices and only abbreviate them as squares. Thus we have the squared covariant term as

$$D\phi^* D\phi = f^2(\partial\alpha + A)^2. \quad (367)$$

Thus we have the Lagrangian as

$$\mathcal{L} = f^2(\partial\alpha + A)^2 + F^2, \quad (368)$$

where we have neglected the potential term which is nothing but a constant.

But at present we didn't tell how the gauge field A_μ transform, or to say, in the general transformation $A \rightarrow A' = A - \partial\theta$, we didn't assign a specific function to $\theta(x)$. Here we may define it to be $\theta = \alpha$, then the transformed

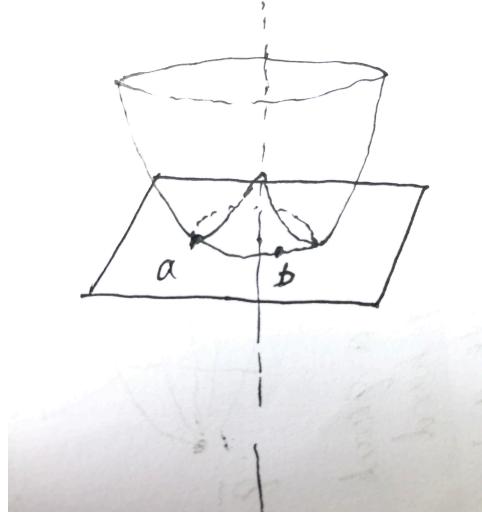


Figure 8: $V(\phi^*\phi) = -a(\phi^*\phi) + b(\phi^*\phi)^2$ diagram

gauge field would be $A' = A - \partial\alpha$. By inserting it into eq (368), we may find an interesting fact that it cancels the derivative term of α , thus the Lagrangian is given by

$$\mathcal{L}' = f^2(\partial\alpha + A')^2 + F^2 = f^2A^2 + F^2. \quad (369)$$

By inspecting the form of the new Lagrangian, we find that the first term f^2A^2 just behaves like a mass term, with f^2 playing the role of the mass of the gauge field A . In this sense, we say that the SSB gives the gauge field the mass. Furthermore, since the gauge transformation $\alpha \rightarrow \alpha + \theta$ doesn't change energy, the system stays in the ground state, but it kills the angular degree of freedom — the field sits in a fixed point and the slowly varying α no longer exists, so the Goldstone boson disappears, we say that it's eaten by the gauge field. To conclude,

The SSB with the gauge transformation has simultaneously two effects: the first is that it gives the mass to the gauge field, the second is that it kills the Goldstone boson.

Here we're going to talk a little about the lost degree of freedom(DOF) of the phase angular α . Why a degree of freedom can be lost, and what happens on earth then? The answer is that it doesn't lose at all but become another DOF. We all know that a photon has two polarizations which are perpendicular to its direction of travel, but if it got a mass, it would not travel at the speed of light, thus in a certain frame it can be brought to rest, and again it can be brought to travel in one of the polarization direction, thus that polarization becomes longitudinal; but there still should be two polarizations perpendicular to the travel direction, thus a massive photon can have three polarization directions, and there are indistinguishable. In this way, the above two effects of SSB are connected to each other, it kills the Goldstone boson thus it kills the phase angular degree of freedom, meanwhile it gives the mass to the gauge field thus it creates an extra polarization direction. The overall effect will be, originally we have four DOF: two components of the charged complex scalar field and two polarizations of photon; but after the SSB, we have 3 polarizations of the massive “photon” and 1 Higgs bosons, which accounts for the 4 DOF. No DOF is lost, but one type becomes another.

Next we are going to elaborate what a mass is in sense of a field. Firstly let's take a closer look at the relativistic relations of photons and massive particles. The relativistic relations are also the dispersion relations, since they're

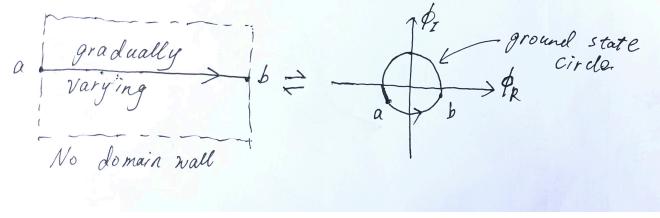


Figure 9: Continuous phase varying in ground state circle

relations between the frequency and the wave number. The relation of photons is given by

$$E = cp \quad \text{or} \quad \omega = ck. \quad (370)$$

We see that it's the frequency is linearly proportional to the wave number, so there's no dispersion — the wave of different wave number/momenta will travel independently as if they never meet. And an obvious mathematical fact we want to mention is that when $k \rightarrow 0 (p \rightarrow 0, \lambda \rightarrow \infty)$, $\omega \rightarrow 0$.

The relation of massive particles is

$$E = \sqrt{p^2 + m^2} \quad \text{or} \quad \omega = \sqrt{k^2 + m^2}. \quad (371)$$

We see that it's a quadratic relation and when $p \rightarrow 0$, $E \rightarrow m$. We plot the above two relations in the diagrams below.

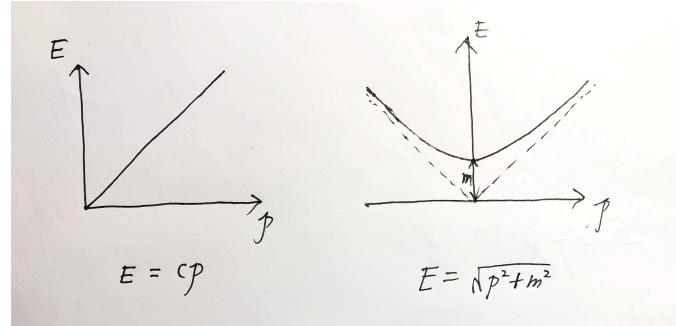


Figure 10: Linear and quadratic relations between E and p

From the above mathematical facts and diagrams we can say that the rest mass is the value of energy when the momentum goes to zero. In the linear case, when momentum goes to zero the energy also goes to zero, so the rest mass for a photon is zero. (That's why we say the "rest" mass of a photon although the photon is never at rest, since we treat this as the definition.) In the quadratic case, when momentum goes to zero the energy goes

to the value m which we call it a mass.

Next we explain what it mean by “the momentum goes to zero” for a field. The momentum is nothing but the wave number in quantum regime, the wave number is inverse proportional to the wave length. The wave length for a field reflects to what extent the field value varies with respect to space distance. For example, a very small wave length depicts intuitively, when you shift the space coordinates by a small distance, the field value changes dramatically, so that it goes back and forth and already went through several wavelength; a very large wave length reflects that when you displace some distance, you will almost see no change of the field value. In this sense, “the momentum goes to zero” corresponds to “the wave length goes to infinity”, which depicts the fact that the field throughout all space changes simultaneously. Thus it would not correspond to the derivative terms in Lagrangian, since the derivative term means a difference between neighboring points, which are not changed simultaneously; when you change the field simultaneously, the derivative term would be zero. By “changing” simultaneously, the field can oscillate simultaneously, or shift to another value simultaneously, or even sits there with no change at all. In the oscillation situation the varying field will have a non-zero frequency, which depends on the curvature of the potential energy of the field at the equilibrium point, and this corresponds to the non-zero energy with zero momentum, i.e, the rest mass. In the shifting and unchanging situation the field will not have a frequency, thus it corresponds to the massless case. To summarize it as follows.

The mass is the energy you can get when you disturb the homogeneous field everywhere and change it simultaneously.

We illustrate these two cases in the following pictures.

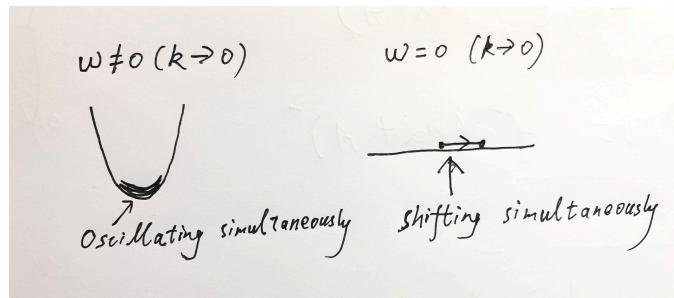


Figure 11: Non-zero and zero frequency when the field varies simultaneously

The oscillation field will correspond to the Higgs boson, and the shifting field (by the phase factor) will correspond to the Goldstone boson, as depicted below.

With this mechanism, the Goldstone boson eaten by the gauge boson would fix the phase of the field so that it could only oscillate in the radial ρ direction, thus it would result in the mass of the Higgs boson. In the standard model, a bunch of particles' mass can only be given by SSB with gauge fixing; other particles beyond the standard model like the dark matter particles, do not gain the mass from SSB but other mechanism which we may not know, but these particles are still not seen in the laboratories.

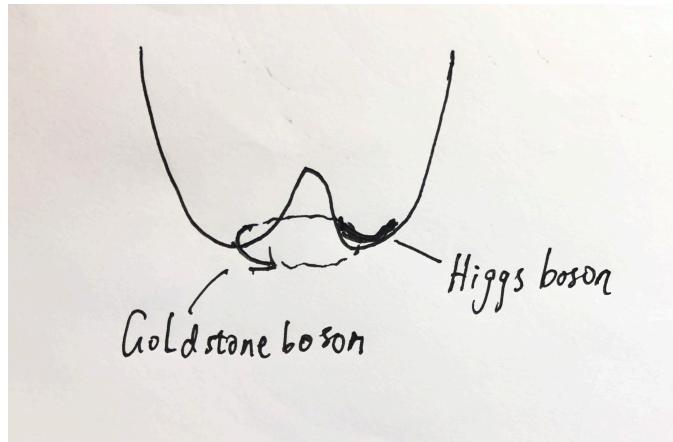


Figure 12: Higgs boson and Goldstone boson mechanism in terms of the field varying

6.2.3 The mass of a fermion field and chirality transition

In this subsection, we are going to explain how the fermion field gets the mass. But before that we firstly introduce some facts. In the last section we talk about Dirac particles which can be decomposed in terms of chirality into left-handedness and right-handedness, they are both carrying charge and in some sense mirror reflections of each other. However, mirror symmetry does not always hold. It holds for e-m force, strong force, but not for the weak force. The weak interaction is always left-handed. If we use Dirac equation for electrons as a model, it is as if the LH particles carry charge while the RH particles don't; of course the charge here should actually refer to that associated with weak interactions.

We list the right- and left-handed component equations for the Dirac field,

$$\begin{cases} i \left(\frac{\partial \psi_R}{\partial t} + \sigma_j \frac{\partial \psi_R}{\partial x^j} \right) = m\psi_L \\ i \left(\frac{\partial \psi_L}{\partial t} - \sigma_j \frac{\partial \psi_L}{\partial x^j} \right) = m\psi_R, \end{cases} \quad (372)$$

Suppose here only the LH particle is charged while the RH ones don't, then these equations do not hold at all, since the charge conservation is violated on two sides of the equators. In the mathematical way, the charged field gets multiplied by a phase factor $e^{i\theta}$ under the U(1) gauge transformation, while the uncharged field doesn't change, as follows

$$\begin{cases} \psi_L \rightarrow \psi'_L = e^{i\theta} \psi_L \\ \psi_R \rightarrow \psi'_R = \psi_R, \end{cases} \quad (373)$$

then the equations obviously does not have the same form as the unchanged ones. But notice a fact that if the mass is zero, the right side of the equations vanish, then everything is fine with the case that only one handedness is charged. We conclude it as follows.

In the sense of only one handedness is charged, charge conservation law would not permit the mixing of left- and right-handedness, or would forbid the mass of the particle.

Then how to coordinate the conflict between the single handed charge and the mass of the particle? Or, how in this single handed charged world, a particle can have a mass? We resolve it by inducing a complex scalar field, i.e, charged boson field ϕ and play a trick to the original equations. First we drop the mass m of the particle and couple the two handedness of the fields with the scalar field and its complex conjugate, with a coupling constant

g , as follows,

$$\begin{cases} i \left(\frac{\partial \psi_R}{\partial t} + \sigma_j \frac{\partial \psi_R}{\partial x^j} \right) = g\phi^* \psi_L \\ i \left(\frac{\partial \psi_L}{\partial t} - \sigma_j \frac{\partial \psi_L}{\partial x^j} \right) = g\phi \psi_R. \end{cases} \quad (374)$$

Then it is easy to check that under the gauge transformation of

$$\begin{cases} \phi \rightarrow \phi' = e^{i\theta} \phi \\ \phi^* \rightarrow \phi'^* = e^{-i\theta} \phi \\ \psi_L \rightarrow \psi'_L = e^{i\theta} \psi_L \\ \psi_R \rightarrow \psi'_R = \psi_R, \end{cases} \quad (375)$$

the equations are kept the same form as the original, thus it is consistent with the U(1) symmetry. The charge conservation in transition between handedness is guaranteed by emitting a charge particle. Note that the real weak interactions are transformed in SU(2) group, so the above analysis is just a simplified version. We schematically draw the pictures for the two equations.

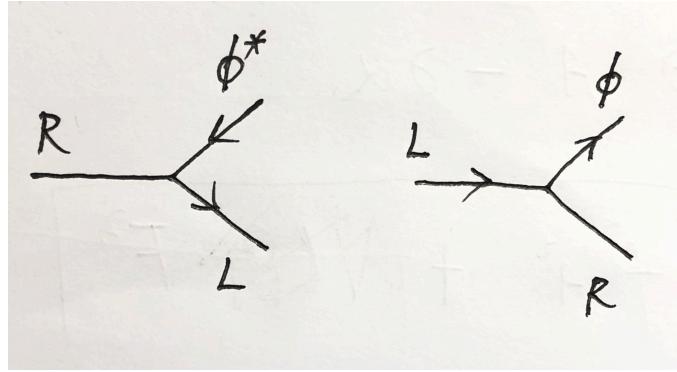


Figure 13: The left- and right- handed particle conversion through emitting a charged particle

The strength of emitting the scalar field ϕ denoted by g is called the Yukawa coupling constant. It has a great significance in explaining the mass of particles.

The SSB for the Higgs field $\phi = \rho e^{i\alpha}$ fixes its radial value to $f + H$, where H is a small quantity reflects oscillation that we neglected here (*collisions of two fermion fields will stop the oscillation of Higgs field*), thus gives the gauge particle W/Z boson the mass f , and gauge transformation locks the phase α to 0. In this way, the factor in front of the fermion field becomes $g\phi^* = g\phi = gf$, which serves as the mass of the particle. Then we have a spectrum of values of g , which correspond to different particles. The mass of a W/Z boson is about 100GeV, thus if we measure the mass of a certain particle, we have that corresponding coupling constant g . The electron mass is about 0.5MeV, which makes the g_e very small; the top quark mass is about 1.5 – 2 times the mass of W/Z bosons, thus g_t would be order 1.

Meanwhile, the Yukawa coupling constant is also the emission amplitude for chirality transition, which would have the chance to serve as a good experimental confirmation about the theory. If the probability for the emission to occur is measured to be m_F^2/f^2 , where m_F is the mass of the corresponding fermion, then it would be a consistency with the theory of Higgs mechanism and chirality transition.

In standard model, all of the neutrinos are left-handed, and all the anti-neutrinos are right-handed. They were thought to be massless, but they aren't because of the neutrinos oscillations, yet their masses are extremely

small. (Beyond the SM, there're right-handed neutrinos, but they have very big masses, which we're not going to talk about.) Neutrinos get their masses by transitions between LH neutrinos and RH anti-neutrinos: $\nu_L \leftrightarrow \bar{\nu}_R$.

The mass of a fermion is the thing which switches LH and RH: $p_L \leftrightarrow p_R$; but particles that get their masses by mixing LH particles and RH anti-particles $p_L \leftrightarrow \bar{p}_R$ are called the **Majorana Particles**. In this sense, neutrinos are Majorana particles, but electrons are not, since an electron carries the charge and if it transits to a positron it'll violate the charge conservation law. So Majorana particles can only be part of those particles who have no electric charges.

Actually there're different saying about the type of neutrinos. Some says there's only one type of neutrinos ν and it's both LH and RH, while others say there're only LH neutrinos ν and RH anti-neutrinos $\bar{\nu}$. Both sayings are right and equivalent, the point is that different types of neutrinos can mix and form massive Majorana particles which are their own anti-particles, because they're superpositions of ν and $\bar{\nu}$. By the way, ν and $\bar{\nu}$ interact differently with Z bosons, they have the opposite weak charges.

To conclude the mechanism for particles' mass, all mass of particles in standard model can be traced back to f , the shift of ground state of Higgs field and SSB. They couldn't have mass if the Higgs field is not shifted, i.e. if $f \rightarrow 0$, all particles would become massless, except for the Higgs particle itself. The gauge bosons normally don't have mass, the only known way to give them mass is the phenomenon of SSB associated with Higgs field, their mass is typically proportional to the W/Z bosons' mass f . All fermions, because of the peculiar property that weak interaction is purely left-handed, can only get their mass by the shift of the compensate charged scalar Higgs fields.

All particles that have energies low enough to be detected, get their mass from the same origin — a small amount of SSB. Yet, there may higher-energy particles ($m > f$) that we haven't discovered which mass are not gain with this mechanism, for example probably the dark matter particles. An imaginable case that can allow mass to be bigger than f is that in weak interactions, both LH and RH particles are charged, thus we don't need an extra scalar field to compensate the charge conservation, the original Dirac equation with a mass term exists, the mass m would be nothing to do with f , thus it could be arbitrarily big.

6.3 SU(2) symmetry in Spin multiplets, Isospin, and in weak interaction

6.3.1 Group theory of orthogonal and unitary rotations

We know from elementary physics that a vector in 3D is represented by a 3-component vector, thus the rotation is represented by a 3×3 matrix, then the resulting vector is given by

$$V'_i = R_{ij}(\theta, \hat{n})V_j, \quad (376)$$

where the \hat{n} is the axis along which the vector rotates and the θ is the rotation angle. Or this can be simply denoted by

$$V' = RV. \quad (377)$$

From physical experience we know that rotation reserves the length of the vector, which can give an important property of the rotation matrix which we will derive as follows:

$$\begin{aligned} V_j V_j &= V'_i V'_i = R_{ij} V_j R_{ik} V_k = R_{ij} R_{ik} V_j V_k \implies R_{ij} R_{ik} = \delta_{jk} \\ \implies R_{ji}^T R_{ik} &= \delta_{jk} \implies R^T R = I \implies R^T = R^{-1}, \end{aligned} \quad (378)$$

where the implication of summation over repeated index is used, and this result says the R matrix is an orthogonal matrix.

A particle which has $S = 1$ has three spin S_z components denoted the probability amplitude of $S_z = 1, 0, -1$, as $(\alpha_1, \alpha_2, \alpha_3)^T$, which just behaves like a ordinary vector. So if we want to rotate its direction of spin, we just rotate it by a 3×3 rotation matrix, in this sense we called a particle with $S = 1$ a vector quantity.

Another kind of quantity is called the spinor quantity. We know that a spin state of a half-spin particle can be in terms of its S_z component denoted as

$$|S_z\rangle = \alpha_\uparrow |\uparrow\rangle + \alpha_\downarrow |\downarrow\rangle. \quad (379)$$

If we denote the up and down state as unit vectors as $|\uparrow\rangle = (1, 0)^T$, $|\downarrow\rangle = (0, 1)^T$, then a general state can be represented as a complex column vector as $|S_z\rangle = (\alpha_\uparrow, \alpha_\downarrow)^T$. If we want to rotate the direction of the spin, we must act a operator on it as $|S'_z\rangle = U(\theta)|S_z\rangle$, in matrix form it's denoted as

$$\begin{pmatrix} \alpha'_\uparrow \\ \alpha'_\downarrow \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \alpha_\uparrow \\ \alpha_\downarrow \end{pmatrix}. \quad (380)$$

Now that the matrix is complex it has four free parameters, but not all of them are independent. One of the constraints is the fact that the total probability should be equal to one: $\alpha_1^* \alpha_1 + \alpha_2^* \alpha_2 = 1$, while the second extra parameter is the phase factor which doesn't account for any physical difference. With these two conditions, the spinor quantity has only two independent parameters, which are as we expected consistent with the number of parameters needed to indicate a direction in space, say the longitude and the latitude.

If we want rotate the spinor, the matrix should preserve the length, or the probability of the spinor, as $\alpha_1'^* \alpha'_1 + \alpha_2'^* \alpha'_2 = \alpha_1^* \alpha_1 + \alpha_2^* \alpha_2 = 1$. This gives the similar property as that of the rotation matrix, which is not called the orthogonality, but unitarity, denoted as $U^\dagger U = I$. Note that the unitary equation is a matrix equation thus it contains four constraints, each one for each matrix elements, thus we reduce the independent parameters of the rotation matrix to $8 - 4 = 4$. And yet there's an additional constraint which states that the determinant of the unitary matrix is 1, $\det(U) = 1$, which implies that this is a "genuine" rotation, not the inversion of it ($\det(U)=-1$). In this sense, there're $4 - 1 = 3$ independent parameters for the rotation, which again coincides with our intuition that we need 2 parameters to indicated the axis and 1 parameter to indicate the angle that rotated around the axis. This kind of unitary rotation is called the $SU(2)$, where S means "Special", i.e, the determinant is one, and U means unitarity, and 2 means what we rotate is a 2-component complex vector. The $SU(2)$ symmetry group is a 2×2 complex representation of the rotation group, i.e, it's isomorphic to rotation in 3D space, while not every $SU(N)$ is like so. By the way, the interchange group is a subgroup of the rotation

group, eg, $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ interchange the two components of a spinor, its determinant is one and it belongs to the $SU(2)$ group.

Next we build the unitary rotation by incremental changes and induce the generator, the process is much similar to those of continuous time and space symmetry transformations. An infinitesimal unitary rotation is built up by

$$U = I - i\epsilon T \quad (381)$$

, where $-i$ is due to convention and ϵ is an infinitesimal quantity, and T is called the generator. From the unitarity we may easily derive that the generator T is Hermitian,

$$U^\dagger U = I \implies T^\dagger = T, \quad (382)$$

which means that it's an observable, and actually it's the angular momentum in 2×2 complex representation. And further since the determinant of the unitary matrix is one, we may find another fact that T is traceless,

$$U = I - i\epsilon T = \begin{pmatrix} 1 - i\epsilon T_{11} & -iT_{12} \\ -iT_{21} & i - iT_{22} \end{pmatrix} \implies 1 = \det(U) = 1 - i\epsilon(T_{11} + T_{22}) = 1 - i\epsilon \text{tr}(T) \implies \text{tr}(T) = 0, \quad (383)$$

where we neglect the higher-order terms of ϵ . In 2×2 complex matrix space, there're 3 independent traceless hermitian matrices, which are nothing but the Pauli matrices. Thus the generator can be built up by linear combinations of the Pauli matrices as

$$T = \sum_i a_i \sigma_i. \quad (384)$$

Then we come to the quark color symmetries. Quarks have three colors denoted by R, G and B, just like the ordinary spin, the color of a quark can be any linear combination of these three, thus it'll appear as a complex 3-component vector, as $(\alpha_1, \alpha_2, \alpha_3)^T$. A rotation group in color space is called $SU(3)$ symmetry group, it's not an ordinary rotation isomorphic to rotation in $3D$ as $SU(2)$, but a new rotation that is hard to visualize. Just like $SU(2)$, there're two constraints for free parameters of an element of $SU(3)$, the unitarity and the unit determinant, thus they will reduce the number of independent parameters to $9 \times 2 - 9 - 9 = 8$, corresponding to the mathematical structure of 8 independent types of gluons. $SU(3)$ symmetry is implied in the Lagrangian of the color fields, i.e, every possible term in QCD Lagrangian should be invariant under $SU(3)$ transformation; this is parallel to every term in QED Lagrangian is invariant under $SU(1)$, which is there simply a phase factor and leads to the consequence that the number of fields and their complex conjugate fields must be the same.

6.3.2 From Spin Multiplets to Isospin

What if we put two half-spin particles together to make a composite particle, then what's the possible total spin and z -component of spin could it have? And what is its rotation properties?

Intuitively, there're four ways to align the spin of the two particles, with each particle spinning up or down along the z -axis, we pictorially denote them as

$$|\uparrow\uparrow\rangle, \quad |\downarrow\downarrow\rangle, \quad |\uparrow\downarrow\rangle, \quad |\downarrow\uparrow\rangle. \quad (385)$$

We see that the maximum spin along the z -axis is 1 when the two particle spins are parallel and pointing up, so there're some of the states which correspond to the total spin of $S = 1$. But for $S = 1$, there're only three possible values for S_z , namely $S_z = 1$ corresponds to $|\uparrow\uparrow\rangle$, $S_z = -1$ corresponds to $|\downarrow\downarrow\rangle$, and $S_z = 0$ corresponds to one of the left two or some combination of them, so there's one independent state left there which does not belong to this multiplets, what total spin does it belong to? The natural answer is $S = 0$.

Actually these four states are the direct product of the single two-component states, denoted as $2 \otimes 2$, and can be treated as four components of a bigger column vector $(\alpha_{\uparrow\uparrow}, \alpha_{\downarrow\downarrow}, \alpha_{\uparrow\downarrow}, \alpha_{\downarrow\uparrow})^T$. This vector can also be rotated by a 4×4 matrix called the 4-dimensional representation of $SU(2)$. A meaningful fact about this 4-component rotation is that with a certain combination of the states, the rotation doesn't change one of the component while

varying the other three; thus the matrix can be separate into blocks in which one of them is unity. We denote this decomposition as $2 \otimes 2 = 1 + 3$, and schematically write it as follows:

$$\begin{pmatrix} \alpha_{\frac{1}{\sqrt{2}}(\uparrow\downarrow-\downarrow\uparrow)} \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & U_{22} & U_{23} & U_{24} \\ 0 & U_{32} & U_{33} & U_{34} \\ 0 & U_{42} & U_{43} & U_{44} \end{pmatrix} \begin{pmatrix} \alpha_{\frac{1}{\sqrt{2}}(\uparrow\downarrow-\downarrow\uparrow)} \\ \alpha_{\frac{1}{\sqrt{2}}(\uparrow\downarrow+\downarrow\uparrow)} \\ \alpha_{\uparrow\uparrow} \\ \alpha_{\downarrow\downarrow} \end{pmatrix}. \quad (386)$$

We call the three states which can be converted by SU(2) rotation to each other as the $S = 1$ multiplets, they are

$$|\uparrow\uparrow\rangle, \quad |\downarrow\downarrow\rangle, \quad \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \quad (387)$$

which are all symmetric under exchange of particles, and this set of multiplets is simply denoted by the number “3”. An intuitive picture of this rotation is that SU(2) rotates the direction of the total spin S , thus changing its z -component S_z , but remain the length of S .

The only one state that is invariant under the SU(2) rotation correspond to the total spin $S = 0$ and is called a singlet,

$$\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (388)$$

it is anti-symmetric under exchange, and is denoted by the number “1”. By the way, if we only consider the conservation of spin angular momentum, this 0 spin state could have the possibility to decay or vanish.

The singlets are of great significance in standard model, they are the states which are “neutral”; the U(1) singlet is electric-neutral, and the SU(3) singlet is color-neutral.

A concept in quark physics that is parallel to the ordinary spin is called the isospin, it’s also a two-component quantity with each one called up — “u”, and down — “d”. Here we state that the isospin is an approximate concept since we ignore the mass and the electric charge difference of the up and down quarks. The isospin of up and down quarks is given by $I = \frac{1}{2}$, and we assign each one a “third component of the isospin” with $I_3 = \frac{1}{2}$ to up quark and $I_3 = -\frac{1}{2}$ to down quark, with their anti-quarks the opposite I_3 . We denote the analog of isospin and ordinary spin as below

$$\begin{array}{ccc} (\uparrow, \downarrow) & \longrightarrow & \text{spin} \\ (u, d) & \longrightarrow & \text{isospin} \end{array} \quad (389)$$

In this sense, we can compose quark-antiquark pairs to form a meson only of the up and down and their antiquarks to give isospin multiplets. The $I = 1$ multiplets are called the pions, the states of which are classified by their I_3 and charge as,

$$\pi^+(I_3 = 1) : |u\bar{d}\rangle, \quad \pi^-(I_3 = -1) : |\bar{u}d\rangle, \quad \pi^0(I_3 = 0) : \frac{1}{\sqrt{2}}(|u\bar{u}\rangle - |\bar{d}d\rangle), \quad (390)$$

with their mass close to each other, about 140MeV. And the $I = 0$ state is named as the η particle,

$$\eta(I_3 = 0) : \frac{1}{\sqrt{2}}(|u\bar{u}\rangle + |\bar{d}d\rangle), \quad (391)$$

with its mass a little heavier than the pions, about 500MeV. By the way, the η and the π^0 are entangled through the isospin.

Next we state the possibility of putting three half-spin particles together to make a composite, then this composite particle would have $2^3 = 8$ independent states, we may easily know that 4 of them correspond to the total spin $S = \frac{3}{2}$, and what about the left 4 states? They correspond to 2 distinct ways of making $S = \frac{1}{2}$ particles. We conclude this as

$$8 = \begin{array}{c} 4 \\ \uparrow \\ S = \frac{3}{2} \end{array} + \begin{array}{c} 2 \\ \uparrow \\ S = \frac{1}{2} \end{array} \times \begin{array}{c} 2 \\ \uparrow \\ \text{2 ways of } "S = \frac{1}{2}" \end{array} \quad (392)$$

The parallel is that we come to particles composed with three quarks which are called the hadrons. The most commonly seen hadrons are nucleons, i.e, protons and neutrons, they are made of up and down quarks and have the same ordinary spin $S = \frac{1}{2}$ and isospin $I = \frac{1}{2}$, we list them below

$$p(I_3 = \frac{1}{2}) : uud, \quad n(I_3 = -\frac{1}{2}) = udd, \quad (393)$$

they have the similar mass about $m_{p,n} \sim 940$ MeV.

Then the less commonly seen hadrons made up of three up or down quarks are isospin $I = \frac{3}{2}$ particles, called $\Delta_{\frac{3}{2}}$, they are listed by their electric charge Q and the third component of isospin I_3 below,

$$\begin{aligned} uuu : \quad & Q = 2, \quad I_3 = \frac{3}{2} \\ uud : \quad & Q = 1, \quad I_3 = \frac{1}{2} \\ udd : \quad & Q = 0, \quad I_3 = -\frac{1}{2} \\ ddd : \quad & Q = -1, \quad I_3 = -\frac{3}{2}, \end{aligned} \quad (394)$$

the mass the Δ particles are about $m_\Delta \sim 1200$ MeV. By the way, the parameters like the mass that we describe particles with, are dependent on frequencies and wavelength of the interactions that engage in, thus they are given the name “the running constant”.

6.3.3 The weak interaction mechanism

There are three generations of quarks and leptons as the table below.

<i>generations species</i>	1st	2nd	3rd
quarks	u d	c s	t b
leptons	ν_e e	ν_μ μ	ν_τ τ

Figure 14: Three generations of elementary particles

In each generation, the left particle is one unit more electric charge than the right one, e.g, the up quark has the charge of $\frac{2}{3}$ while the down quark has the charge of $-\frac{1}{3}$, and the ν_e neutrino has charge 0 while the electron has charge -1.

The different types in a single generation can be converted to each other via intermediate particles, which behave like particle/anti-particle pairs in that generation. Quarks can be converted by strong interaction particles like mesons (e.g, $\pi^+ \leftrightarrow u\bar{d}$), or by weak interaction particles like W bosons (e.g, $W^+ \leftrightarrow u\bar{d}$); leptons can only be converted by weak interaction particles as W bosons (e.g, $W^+ \leftrightarrow e^+\bar{\nu}_e$). Here we only concern weak conversions, and considering the charge conservation, we draw the conversion picture in the first generation as an example.

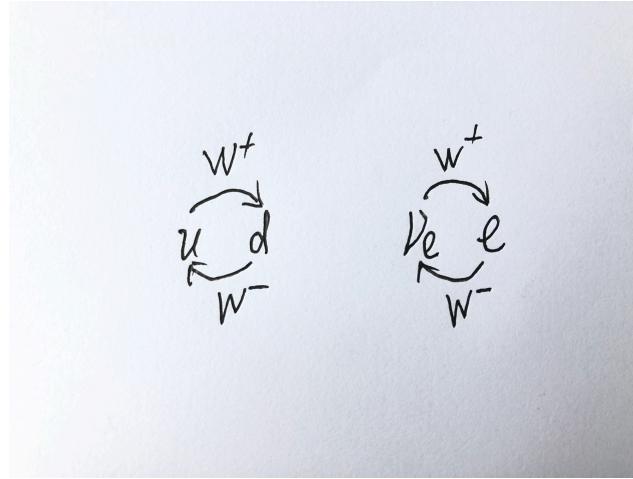


Figure 15: Conversion within a generation by weak interactions

We focus on conversion between the electron and the neutrino, and list all the possible processes as below. Then we use only the first one as an example to illustrate that W bosons behave like a particle/anti-particle pair.

Additionally, conversions from leptons in one generation to another generation through W bosons are possible. We list the only decay channel of muon particles as below.

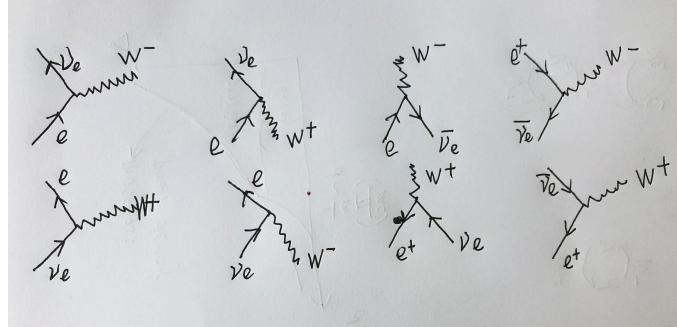


Figure 16: Possible conversions within the first generation

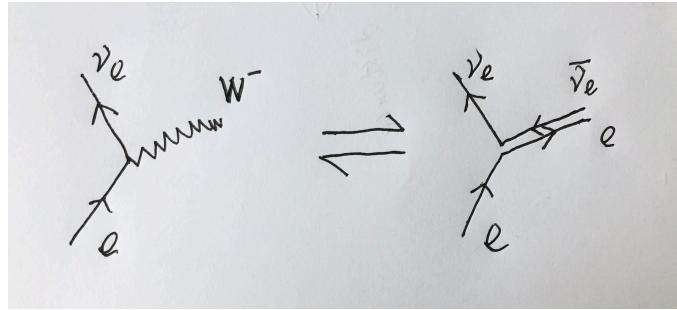


Figure 17: Illustration of the W boson as particle/anti-particle pair

Moreover, since the W bosons are involved within both the quarks and leptons, the conversions between them are also possible. We list the primary decay channel of the pions as below.

In the above process, the pion mass is about 140MeV and the mass of the muon is about 100MeV , which is consistent with the energy difference between initial and final states, where the rest of the energy goes to the neutrino and the kinetic energy of the particles.

Finally, we have the beta-decay of the proton process which we will leave in the below question.

The weak coupling constant is of the same order with that of electric interaction, then why is the weak force so weak? That has something to do with the propagator and the mass of it. Consider the following proton decay as an example.

There are two vertices, so the decay amplitude will be proportional to the squared coupling constant g_W^2 which is of the same order with the fine structure constant $\alpha \sim e^2$. What additionally matters here is the propagator Δ corresponding to the intermediate W^- , it is a function of the spacetime distance (the proper time) between the emission and the absorption points, and it depends crucially on the mass of the propagator. An then we give the propagator versus distance diagram regarding to massless and massive case as follows.

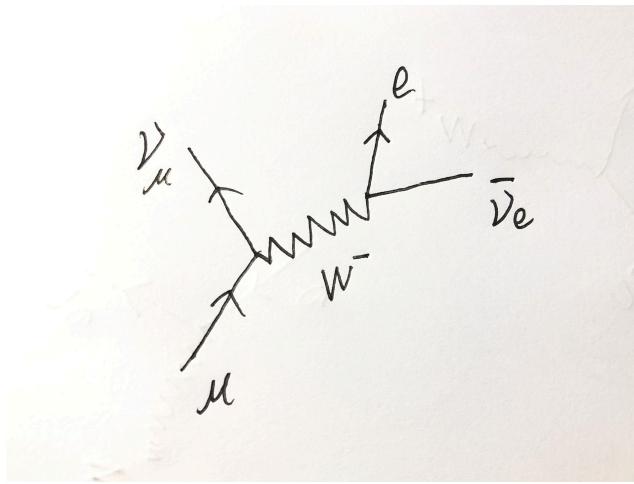


Figure 18: The only decay channel of a muon

The blue line indicates the relation between the propagator and the distance, which is dominated by the power decay thought the entire distance; while the red line denotes the massive case, which at some point according to the dimensional analysis to be the order of $\frac{1}{m}$ (or $\frac{\hbar}{mc}$ in normal units) begins to decay at an exponential rate. In experiment, we can't measure the distance that propagator runs directly, but we can detect the momentum of the W^- particle, which is given by $p_{W^-} = p_n - p_p$ and is related to the distance through the Fourier transformation, and then test the amplitude for that given momentum, to judge whether we give the right form of the postulate of the propagator.

The Fourier transformation of the red line is calculated to be as $\tilde{\Delta}(k) = \frac{1}{k^2 + m_W^2} \leq \frac{1}{m_W^2}$, where k is the momentum, and when it is low, it gives the large amplitude which is easy to detect. Thus we have the final amplitude combining the effect from both the coupling constant and the propagator, largest to be $\frac{g_W^2}{m_W^2} \equiv G_F$, which is usually pretty small, let alone the probability per unit time to be $\frac{g_W^4}{m_W^4}$. The crucial ingredient here is the mass of the W boson is so large, as about 100 times of that of a proton. That accounts for the reason why the weak force is so weak, and also accounts for the lifetime of the W is so long since it's inverse proportional to the decay rate, and generally is k -dependent as $\frac{(k^2 + m_W^2)}{g_W^2}$.

By the way, to see from the form of the propagator, the massless photon propagator will have the form of $\tilde{\Delta}(k) = \frac{1}{k^2}$, where k is usually experimentally very small, so the propagator will be relatively big, thus the electric force will big compared to that of the weak force.

Finally we list three ingredient that control the interaction ratios, by adding another one, as: 1, the coupling constant; 2, the mass of the intermediate particles; 3, the available energy, i.e, the energy difference between the initial and the final states $E_i - E_f$. Those close the question that why the weak force is so weak.

Why can't we knock out a virtual particle?

The proton-neutron oscillation has a frequency. If we want to detect the intermediate particle W^+ , we have to use the wave of the same frequency to knock it out, but in this sense the knocked-out particle can't be called an "intermediate" particle but a "final" particle, thus the detection will make it another process, not the original one. So in this sense, the intermediate or the virtual particle is undetectable.

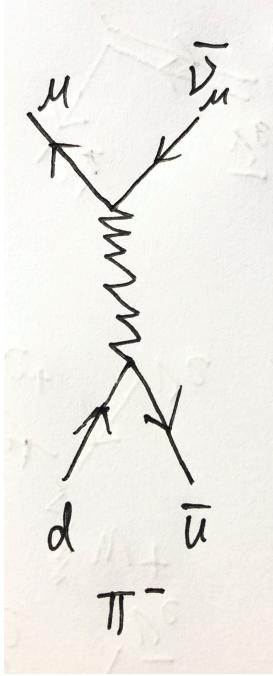


Figure 19: Primary decay channel of the pions

6.4 SU(3) in QCD

A single flavor of quark can have three colors, assigned as R, G, and B, or linear superpositions of them. Just like the spin, we can denote them as three-component vector $(q_1, q_2, q_3)^T$, and their anti-quark components as $(q_1^*, q_2^*, q_3^*)^T$. If we want to rotate them in color space we must act a 3×3 matrix on them, so they'll transform as

$$\begin{cases} q_j \rightarrow q'_j = U_{ij} q_j \\ q_j^* \rightarrow q'^*_j = U_{ij}^* q_j^* \end{cases} \quad (395)$$

where the U_{ij} is called an element of SU(3) group, and the q_i and q_i^* are called the complex conjugate representations of SU(3), denoted as “3” and “ $\bar{3}$ ”. Their generators differ from each other a minus sign, as $T = -\bar{T}$.

The SU(3) gauge field (gluon) is transformed as if it was a direct product of the anti-quark vector and the quark vector,

$$A_{ij} \longleftrightarrow q_i^* q_j. \quad (396)$$

Below we show some schematic pictures involving the gluons.

To proceed, we talk about composite particles. Mesons are composed of quark/anti-quark pairs $q\bar{q}$, they are direct products of the 3 and $\bar{3}$, and can be decomposed to a singlet and an octet as $3 \otimes \bar{3} = 1 + 8$. The singlet which has a great significance, is denoted as $R\bar{R} + G\bar{G} + B\bar{B}$, it's colorless (or color-neutral), meaning that it's invariant under the SU(3) group, just like the $S = 0$ state for a two half-spin composite particle; the other 8 states are called the $8D$ representations of SU(3) and can be transformed to each other, meanwhile 8 is the number of generators of SU(3), each of which corresponds to a member of $8D$ representations. By the way, the color property of a gluon is just like a meson, it has an octet denoted as the “8”.

A baryon is composed of three quarks. But firstly the two quark direct product can be decomposed as $3 \otimes 3 = \bar{3} + 6$; and then we make the direct product with the third quark, with the first part being the familiar $\bar{3} \otimes 3 = 1 + 8$, the same as mesons, and the second part being $6 \otimes 3 = 8 + 10$. Again the most important ingredient is the singlet, which is colorless, denoted as RGB .

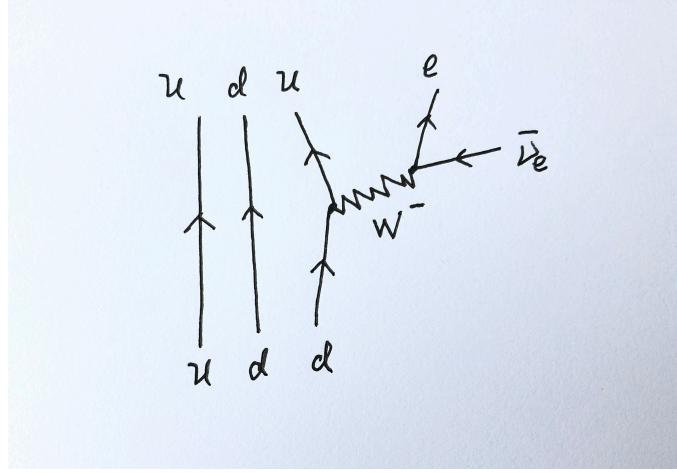


Figure 20: Beta Decay of proton

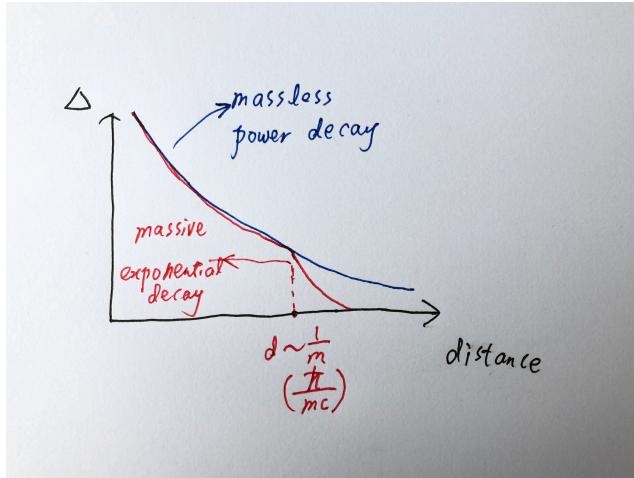


Figure 21: Propagator versus distance diagram

A final color-composite particle is the gluon ball which is composed of two gluons as gg , it's decomposed in the way as $8 \otimes 8 = 1 + 63$. The singlet is again colorless.

Baryons, mesons, and glue balls are three ingredients of hadrons, and all free/unconfined particles transform as color singlets “1”. Singlet states are always anti-symmetric like $\epsilon_{ij}\phi_i\psi_j$ and $\epsilon_{ijk}q_iq_jq_k$. If you make color singlets, then you will not get fraction charged particles.

What if we hit a meson very hard and try to kick out a single quark?

Firstly the quark and the anti-quark are separated, and they still interact with each other via the gluons. Gluons, unlike the charge-neutral photons, are not color-neutral, thus they attract each other and form a flux tube, where the energy is stored in. As you draw the quark and anti-quark farther and farther, the flux tube will become longer and longer but not thinner, the uniform energy per unit length is kept fixed, thus the total energy stored in is proportional to the distance between them $E \sim l$. At some time, the pair will stop flying away and begin to pull together. If you hit a quark very very hard inside a meson to keep it from flying back, then suddenly at a point between them a new quark/anti-quark pair is created out of vacuum, thus a meson becomes two mesons, each of which is still color-neutral; similarly to proceed, if the quark is flying even faster, then many pairs form in between, and we have a whole bunch of mesons, called a jet.

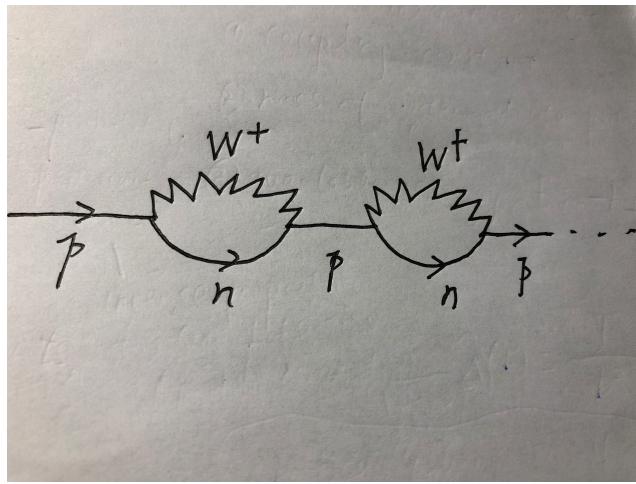


Figure 22: Proton-Neutron Decay Equilibrium

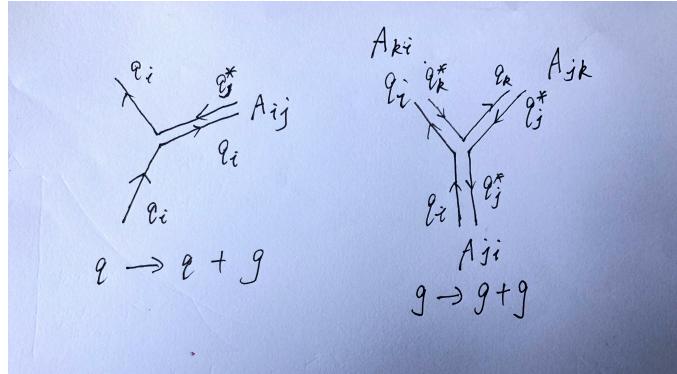


Figure 23: Gluons behave like quark/anti-quark pairs

6.5 The Grand Unification

6.5.1 The screening effect and the running coupling constant

We know the electric charge serves as the scattering amplitude in QED, but actually they're not constant, but a function of the distance between two of them.

The Coulomb's law tells us the force between two charges are given by

$$F = \frac{e^2}{R^2}, \quad (397)$$

but actually this is only an asymptotic result in case of large space separation. The charge you feel at a large distance is not the same as the numerical value of charge you put into the Lagrangian, which is called the bare electric charge. In short distance, they increase due to the decrease of the *vacuum polarization*.

In vacuum, the virtual electron/positron pairs pop up and annihilate everywhere. When you put two opposite real charges into the vacuum, the virtual particle pairs in between would get a little separation from each other and form a dipole. But unlike the polarization in dielectric, the e/e^+ pair's separation depends on their energy — when the energy of them is large, their separation gets smaller. When you fix the one of the real charges, say the positive charge, and move the negative charge towards it, as the negative charge gets closer and closer, the low-energy virtual far apart just get outside the real pairs and would have no influence on the real charge, but high-energy virtual particle pairs always pop out, so the net charge you measure will get increase very slowly like a logarithm, and this increase would last until some new physics comes into play. We depict the charge versus inverse distance diagram below. Yet this increase is very hard to see in experiments.

The same goes with scattering, the higher energy you collide them, the bigger charge you would see that should be put into the calculations.

In strong interaction, the situation is opposite, as $q(R) \sim R$, $E = \frac{q(R)^2}{R} \sim R$. At small R , it'll be back to $\frac{q^2}{R}$.

The similar situation goes with gravity, Newton's formula

$$F = G \frac{mm}{R^2}, \quad (398)$$

only holds for particles at rest. What really matters is the energy, so more accurately we have

$$F = G \frac{EE}{c^4 R^2}. \quad (399)$$

Putting $p \sim \frac{\hbar}{R}$ and $E = pc$, we have $E = \frac{\hbar c}{R}$, so

$$F = \frac{G \hbar^2}{c^4 R^4} \quad \Rightarrow \quad E \sim R^{-3}. \quad (400)$$

And then the distance that make gravity and e-m force meet is to set

$$G \frac{G \hbar^2}{c^4 R^4} = \frac{\hbar c}{R^2} \quad \Rightarrow \quad R = \sqrt{\frac{\hbar G}{c^3}}, \quad (401)$$

which is called the Plank length.

At some extremely high energy, the four coupling constant will more or less come to a point, the upshot is that four forces become comparable somewhere near the plank scale — they unify.

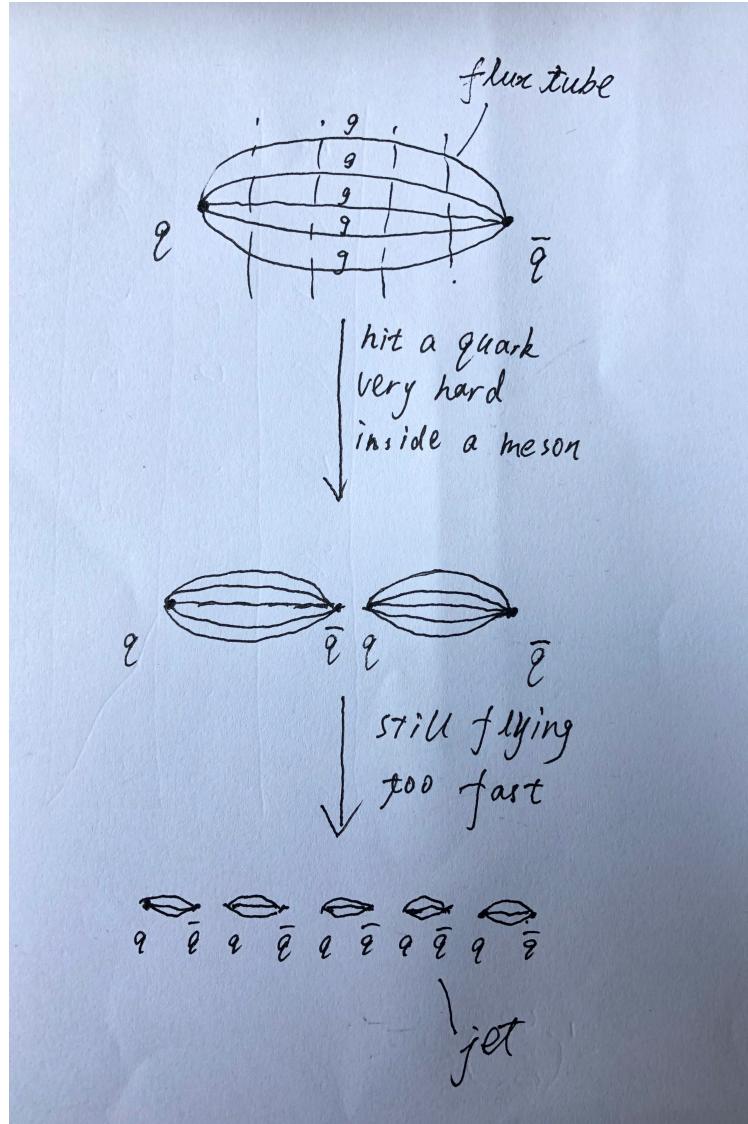
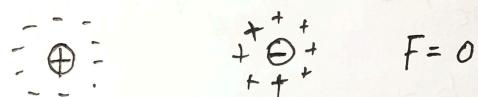


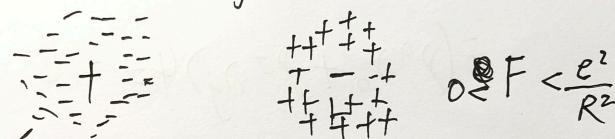
Figure 24: Breaking of the gluon flux tube and formation of a jet

conductor: large R , $\epsilon(R) = \infty$

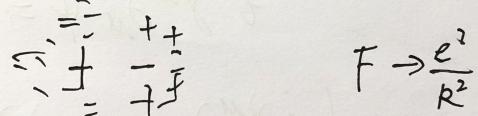


$$F = \frac{e^2}{R^2}$$

dielectric: dressed charge $\equiv \epsilon(R) < \epsilon(\infty)$ bare charge



Screening
cloud



$$0 < F < \frac{e^2}{R^2}$$

$$F \rightarrow \frac{e^2}{R^2}$$

$\nearrow R, \frac{\epsilon(R)^2}{R^2}$ approaches some value
less than $\frac{\epsilon(\infty)^2}{R^2}$

Figure 25: Two analog examples of polarizations

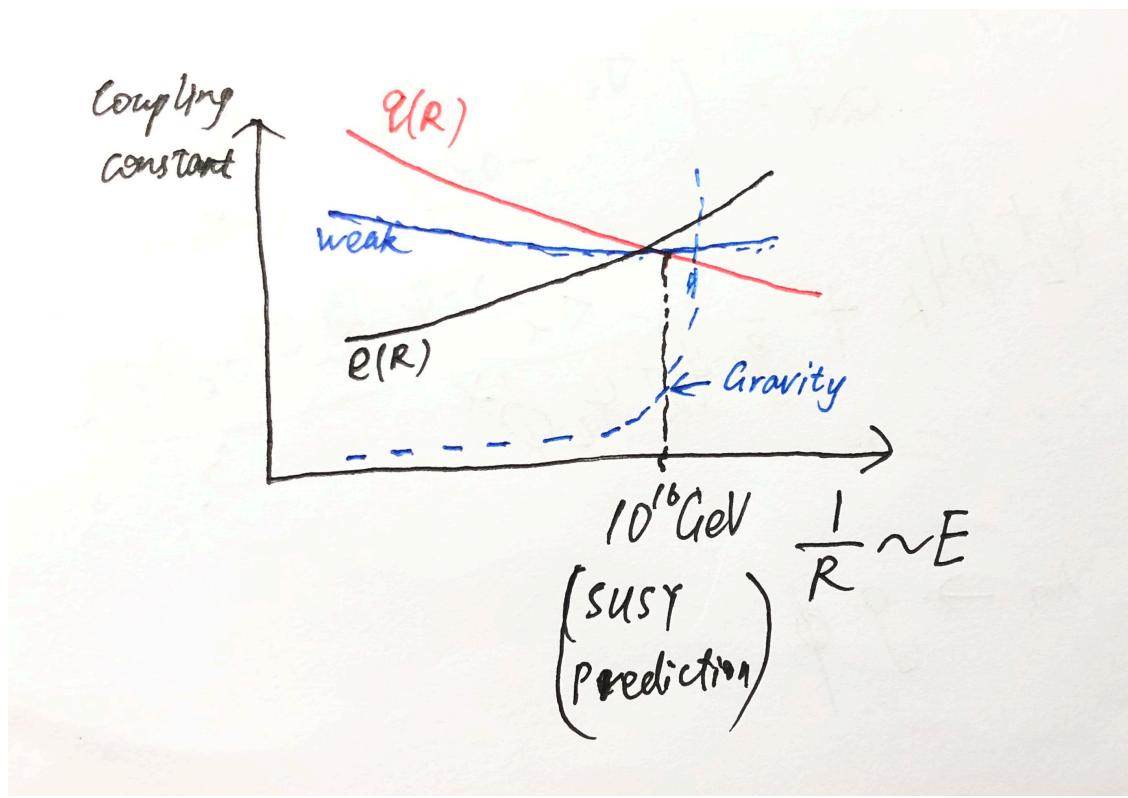


Figure 26: The curve trends for different coupling constants