Bits and Qubits

1. Bit

A bit is a two-level system, which at any time can have the value of logical zero OR logical one. (If the capacitor is charged below the threshold value, then this is a logical zero, if above, then this is a logical one.)

1.1. 1 bit

The number of states of one bit is 2: 0 or 1 From mathematical point of view, number of states of single bit is: $2^1 = 2$

0	0
1	1

The amount of memory required to store all the states of one bit is:

$1 * 2^1 = 2_{bits}$

1.2. 2 bits

The number of states of a system consisting of 2 bits is: $2^2 = 4$

0	0	0
1	0	1
2	1	0
3	1	1

At any given time, the system can only be in one state.

The amount of memory required to store all the states of 2 bit system is:

$2 * 2^2 = 8_{bits}$

1.3. 3 bits

The number of states of a system consisting of 3 bits is: $2^3 = 8$

0	0	0	0
1	0	0	1
2	0	1	0
3	0	1	1
4	1	0	0
5	1	0	1
6	1	1	0
7	1	1	1

The amount of memory required to store all the states of 3 bit system is:

$3 * 2^3 = 24_{bits}$

1.4. 300 bits

The number of states of a system consisting of 300 bits is:

2^{300} = more than number of atoms in visible universe

The amount of memory required to store all the states of 300 bit system is:

$300 * 2^{300}$ = more than number of atoms in visible universe

2. Qbit

A qubit is a two-level system that at one time can take logical 0, logical 1, **and can also be in a superposition** of the **classical** states of logical 0 and logical 1.

It is denoted as follows:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{1}$$

Also according to quantum theory, the factors α and β is a complex numbers.

And $|\alpha|^2$ it is the probability of finding a qubit in state 0,

and $|\beta|^2$ it is the probability of finding a qubit in state 1.

The (1) is also called a **pure quantum state** (usually just called **state**).

A pure quantum state can be written as a superposition of basis states (1)

2.1. Normalization condition

Since, in the end, the system is with 100 percent probability in state 0 or 1, then acording to the probablity theory:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{2}$$

This (2) is the normalization constraint.

2.2. Single Qubit is a linear space. Superposition.

If we look at the formula (1), we could recognize that a single qubit represents whole 2 dimensional linear space. Where the basis consists of $|0\rangle$ and $|1\rangle$ vectors

In other words, we say that the qubit $|\psi\rangle$ is in a **superposition** of states $|0\rangle$ and $|1\rangle$

The superposition is when the system is simultaneously in different states with some probability.

3. Bra-ket notation

3.1. Ket vector

As we can see, the $|\psi\rangle$, in formula (1) is a vector when the α and β have some fixed values.

And cccording to the quantum theory the $|0\rangle$ and $|1\rangle$ is a basis vectors (i.e they are orthogonal).

The numbers in notations like $|n\rangle$ are the analogues of indices in matrix notation.

For qubits in quantum computers, the dimension is 2, so we only have

$|0\rangle = e_0 = \begin{bmatrix}1\\0\end{bmatrix}$, $|1\rangle = e_1 = \begin{bmatrix}0\\1\end{bmatrix}$

etc., where e_n is the vector which has a 1 in the **n**th position and 0 in the other entries.

Thus, we can write the (1) as follows:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \alpha \begin{bmatrix}1\\0\end{bmatrix} + \beta \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}\alpha\\\beta\end{bmatrix} \tag{3}$$

The (3) is a **unit vector**. I.e the length of the $|\psi\rangle$ vector is:

$$|(|\psi\rangle\,| = 1$$
$$\langle\psi|\psi\rangle = 1$$

Actually each $|\rangle$ vector, is called **ket** vector. And it is represented as a column vector $|\psi\rangle = \begin{bmatrix}\alpha\\\beta\end{bmatrix}$ (in 2 dimensional state space).

3.2. Combined states. Tensor product of linear spaces.

Another important property of qubits is that they can be "prepared" in such a way that their total state is stored in the product of the linear spaces formed by each of the qubits.

Such a state is called a **Combined state** of qubits.

The operation of products of linear spaces is called **tensor product**. And is denoted as $\otimes$.

If two qubits, the $|\psi\rangle$ and the $|\varphi\rangle$ are in the **combined** state, it is denoted as:

$|\psi\rangle \otimes |\varphi\rangle$

The number of states of a system consisting of 2 qubits is equal to the dimension of the tensor product of linear spaces formed by these qubits.

The dimension of the tensor product of linear spaces of a two-level system is 2^n, where the n is the number of qubits.

And this coincides with the number of states for a system of classical bits, with the difference that to store all the states of a system of 3 bits, it is required

$3 * 2^3 = 24_{bits}$

but to store the same amount in qubits, only 3 qubits are enough.

To store all the states of 300 classical bits, $300 * 2^{300}$ bits required (more than numbers of atoms in visible universe), but only 300 combined qubits is required to store the same amount of states.

3.3. Examples of combined(joint) state.

Example 1

Start with a system having 4 energy levels. Let it interact with a 2 level system. The state space of the combined system has a dimension of

$4 * 2 = 8$

Example 2.

Two qubits (each qubit is a two level system). Combined system has a dimension of

$2 * 2 = 2^2 = 4$

Example 3.

3 qubit system:

$2 * 2 * 2 = 2^3 = 8$

Example 4.

10 qubit system:

$2 * 2 * 2 * 2 * 2 * 2 * 2 * 2 * 2 * 2 = 2^{10} = 1024$

Example 5. Let's consider two vectors

$$V = \begin{bmatrix} 1 \\ 0 \\ 7 \end{bmatrix} \text{ and } W = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

- then

$$V \otimes W = \begin{bmatrix} 1 \cdot W \\ 0 \cdot W \\ 7 \cdot W \end{bmatrix} = \begin{bmatrix} 1 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ 0 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ 7 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 7 \\ 14 \end{bmatrix}$$

Example 6. In quantum mechanis the tensor product is a product of **ket** vectors

$$|\phi\rangle = \alpha_0 |0\rangle + \beta_0 |1\rangle ,$$
$$\alpha_0 = -\frac{4}{5}, \beta_0 = \frac{3}{5}$$

$$|\psi\rangle = \alpha_1 |0\rangle + \beta_1 |1\rangle ,$$
$$\alpha_1 = -\frac{1}{2}, \beta_1 = \frac{\sqrt{3}}{2}$$

How can we describe their **combined(joint)** state?

The first guess might be using multiplication of some sort.

Let's use our **tensor product notation** - $\otimes$

$$|\phi\rangle \otimes |\psi\rangle = (\alpha_0 |0\rangle + \beta_0 |1\rangle) \otimes (\alpha_1 |0\rangle + \beta_1 |1\rangle) =$$
$$\alpha_0\alpha_1 \underbrace{|0\rangle \otimes |0\rangle}_{|00\rangle} + \alpha_0\beta_1 \underbrace{|0\rangle \otimes |1\rangle}_{|01\rangle} + \beta_0\alpha_1 \underbrace{|1\rangle \otimes |0\rangle}_{|10\rangle} + \beta_0\beta_1 \underbrace{|1\rangle \otimes |1\rangle}_{|11\rangle} \tag{4}$$

The $|\psi\rangle \otimes |\phi\rangle$ looks exactly the same as a linear combination of the four basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$.

However, we do need to check whether the result of $|x\rangle \otimes |y\rangle$ is actually a quatnum state. I.e. we have to check the **normalization condition** (2):

$$|\alpha_0\alpha_1|^2 + |\alpha_0\beta_1|^2 + |\beta_0\alpha_1|^2 + |\beta_0\beta_1|^2 = (|\alpha_0|^2 + |\beta_0|^2) \cdot (|\alpha_1|^2 + |\beta_1|^2) = 1 \tag{5}$$

Also, let's consider $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. According to (3) an (4):

$$\begin{aligned}
|00\rangle = |0\rangle \otimes |0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \\
|01\rangle = |0\rangle \otimes |1\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \\
|10\rangle = |1\rangle \otimes |0\rangle = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \\
|11\rangle = |1\rangle \otimes |1\rangle = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}
\end{aligned} \tag{6}$$

In the matrix notation:

$$|\phi\rangle|\psi\rangle = |\phi\psi\rangle = |\phi\rangle \otimes |\psi\rangle = \begin{bmatrix}-\frac{4}{5}\\ \\ \frac{3}{5}\end{bmatrix} \otimes \begin{bmatrix}-\frac{1}{2}\\ \\ \frac{\sqrt{3}}{2}\end{bmatrix} = \begin{bmatrix}-\frac{4}{5}\begin{bmatrix}\frac{1}{2}\\ \\ \frac{\sqrt{3}}{2}\end{bmatrix}\\ \\ \frac{3}{5}\begin{bmatrix}\frac{1}{2}\\ \\ \frac{\sqrt{3}}{2}\end{bmatrix}\end{bmatrix} = \begin{bmatrix}-\frac{4}{10}\\ \\ -\frac{4\sqrt{3}}{10}\\ \\ \frac{3}{10}\\ \\ \frac{3\sqrt{3}}{10}\end{bmatrix}$$

Now let's rewrite it according to the (4).

$$|\phi\psi\rangle = -\frac{4}{10}|00\rangle - \frac{4\sqrt{3}}{10}|01\rangle + \frac{3}{10}|10\rangle + \frac{3\sqrt{3}}{10}|11\rangle$$

Now let's check if it a correct quantum state, i.e. if a normalization conditian (5) is met:

$$\left|-\frac{4}{10}\right|^2 - \left|\frac{4\sqrt{3}}{10}\right|^2 + \left|\frac{3}{10}\right|^2 + \left|\frac{3\sqrt{3}}{10}\right|^2 = \frac{16}{100} + \frac{48}{100} + \frac{9}{100} + \frac{27}{100} = \frac{100}{100} = 1$$

We see that the normalization condition is met.

Therefore, in accordance with the meaning of the normalization condition (2), we get:

Probability of measuring $|00\rangle = \left|-\frac{4}{10}\right|^2 = \frac{16}{100} = 16\%$.

Probability of measuring $|01\rangle = \left|-\frac{4\sqrt{3}}{10}\right|^2 = \frac{48}{100} = 48\%$.

Probability of measuring $|10\rangle = \left|-\frac{3}{10}\right|^2 = \frac{9}{100} = 9\%$.

Probability of measuring $|11\rangle = \left|-\frac{3\sqrt{3}}{10}\right|^2 = \frac{27}{100} = 27\%$.

Example 7. Tensor product of two matrices

Given a matrix.

$$H = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

In quantum mechanics it is called Hadamard matrix.
Let's find $H^{\otimes^2} = H \otimes H$

$$H \otimes H = \begin{bmatrix} \frac{1}{\sqrt{2}}H & \frac{1}{\sqrt{2}}H \\ \frac{1}{\sqrt{2}}H & -\frac{1}{\sqrt{2}}H \end{bmatrix} = H^{\otimes^2} =$$

$$= \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

3.4. Properties of the tensor product

1) The tensor product is not commutative: $|0\rangle \otimes |1\rangle = |01\rangle \neq |10\rangle = |1\rangle \otimes |0\rangle = |00\rangle$
2) $A \otimes (B \otimes C) = (A \otimes B) \otimes C = A \otimes B \otimes C$ (associativity)
3) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
4) $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ (distributive law)
$(A + B) \otimes C = (A \otimes C) + (B \otimes C)$
5) α - is a scalar
$\alpha A \otimes B = A \otimes \alpha B = \alpha(A \otimes B)$ (freely floating scalar)

3.5. Computational Basis.

In the (3)

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{7}$$

- the

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- are called 2 dimensional computational basis. Let's denote it as follows:

$$\text{2D, } 2^1$$
$$B_2 = \{|0\rangle, |1\rangle\}$$

-and we can expand any vector in two-dimensional space in this basis.
But we can extend this basis for 4 dimension space. According to the (6)

$$\text{4D, } 2^2$$
$$B_4 = B_2 \otimes B_2 = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$$

Similarly we can make any bases a power of 2. For example 8D 2^3

$$\text{8D, } 2^3$$
$$B_2 \otimes B_4 = \{|000\rangle, |001\rangle, |010\rangle, ..., |111\rangle\}$$

Also, let's consider next ket vector: $|101\rangle$. 101 is quals to 5. Therefore 5th element of a row vector will be equal to 1, i.e.:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Example. Write the $|\psi\rangle = 7\,|101\rangle + 3i\,|111\rangle$. Write it in a vector form:

$$|\psi\rangle = 7 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 3i \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 7 \\ 0 \\ 3i \end{bmatrix}$$

3.6. Entangled states.

There is states, that cannot be written as a tensor product of other states. A superposition which cannot be written as a tensor product is called **entangled state**.

Suppose the first qubit is in the state

$$|\psi_0\rangle = \frac{3}{5}\,|0\rangle + \frac{4}{5}\,|1\rangle$$

and the second qubit is in the

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}\,|0\rangle - \frac{1}{\sqrt{2}}\,|1\rangle$$

then the joint state of the two qubits is

$$|\Psi\rangle = \left(\frac{3}{5}\,|0\rangle + \frac{4}{5}\,|1\rangle\right) \otimes \left(\frac{1}{\sqrt{2}}\,|0\rangle - \frac{1}{\sqrt{2}}\,|1\rangle\right) =$$
$$\frac{3}{5\sqrt{2}}\,|00\rangle - \frac{3}{5\sqrt{2}}\,|01\rangle + \frac{4}{5\sqrt{2}}\,|10\rangle - \frac{4}{5\sqrt{2}}\,|11\rangle =$$
$$\alpha_0\alpha_1|00\rangle + \alpha_0\beta_1|01\rangle + \beta_0\alpha_1|10\rangle + \beta_0\beta_1|11\rangle$$

As we can see joint state $|\Psi\rangle$ above, can be represented as a tensor product of two states $|\psi_1\rangle$ and $|\psi_2\rangle$

But let's consider a state, that cannot be represented in such a way:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle)$$

Let's put:

$$|\psi_0\rangle = \alpha_0\,|0\rangle + \beta_0\,|1\rangle$$

$$|\psi_1\rangle = \alpha_1 |0\rangle + \beta_1 |1\rangle$$

Hence, as we already know:

$$|\psi_0\psi_1\rangle = \alpha_0\alpha_1|00\rangle + \alpha_0\beta_1|01\rangle + \beta_0\alpha_1|10\rangle + \beta_0\beta_1|11\rangle$$

Comparing this with the $|\Psi\rangle$ state we see that multipliers $\alpha_0\beta_1$ and $\beta_0\alpha_1$ have to be zero. But on the other hand, we see that $\alpha_0\alpha_1 = \frac{1}{\sqrt{2}}$ and $\beta_0\beta_1 = \frac{1}{\sqrt{2}}$, which implies that, $\alpha_0\beta_1$ and $\beta_0\alpha_1$ cannot be zero. **That is, we have a contradiction.**

Therefore, the state Ψ cannot be represented as a product state of two other states.

Such states called **entangled states**.

3.7. Dot product. Inner product. Outer product

Consider the projection of one state vector $|\psi\rangle$ onto another $|\phi\rangle$ I.e how much of a "shadow" $|\psi\rangle$ casts on $|\phi\rangle$. The operation of finding the length of this shadow is called **inner product.** It is called inner product since we use complex space. Our space is complex because in the (1) the α and β are complex numbers. In regular euclidean space is is called **dot product** or **scalar product**. So we will call it an inner product, since we work in complex space.

If we project the $|\psi\rangle$ vector onto itself, we will get its **length**. The length, by definition it is a non-negative real number (not complex).

From algebraic point of view the inner product takes two vectors and maps them into a real, non-negative number.

3.8. Euclidean space. Dot product.

Let's consider Euclidean space first.

Lt's recollect well known formula from the college:

$$(v, w) = v_1w_1 + ... + v_nw_n \tag{8}$$

It is a or dot product (or scalar product).

Define a **length** or **norm** ov v by the formula:

$$||v|| = \sqrt{(v, v)} = \sqrt{v_1^2 + ... + v_n^2} \tag{9}$$

In the **matrix notation**, the scalar product of two vectors:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_n \end{bmatrix}, \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ . \\ . \\ . \\ w_n \end{bmatrix} \tag{10}$$

defined as product of **row wector** on **columnt vector**, i,e (v^T, w):

$$\begin{bmatrix} v_1, v_2, ..., v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ . \\ . \\ . \\ w_n \end{bmatrix} = v_1w_1 + ... + v_nw_n \tag{11}$$

Main property of the scalar product (8), is that it always a **real** and **positive** number, i.e:

for all $v \in R^n, (v,v) = ||v||^2 \geq 0$, and $(v,v) = 0$ if and only if $v = 0$.

There is also other properties, but we, maybe, consider them later, when needed.

Important. When in the euclidean space we write a (v,w), we mean (v^T, w)

$$(v,w) \sim (v^T, w) \tag{12}$$

3.9. Complex space. Inner product.

The thing, is that in the complex space the formula (11) in not always a real and positive number:

Let's consider complex vecrots:

$$w_1 = \begin{bmatrix} 1+i \\ 1-i \\ 0 \end{bmatrix}, w_2 = \begin{bmatrix} -i \\ 0 \\ 2-i \end{bmatrix}, w_3 = \begin{bmatrix} 2+i \\ 1-3i \\ 2i \end{bmatrix} \tag{13}$$

$$(w1, w2) = \begin{bmatrix} 1+i, 1-i, 0 \end{bmatrix} \begin{bmatrix} -i \\ 0 \\ 2-i \end{bmatrix} = (1+i)(-i) + 0 + 0 = 1 - i \tag{14}$$

$$\begin{gathered} (w1, w3) = \begin{bmatrix} 1+i, 1-i, 0 \end{bmatrix} \begin{bmatrix} 2+i \\ 1-3i \\ 2i \end{bmatrix} = \\ (1+i)(2+i) + (1-i)(1-3i) + 0 = \\ (1+3i) + (-2-4i) = -1 - i = -(1+i) \end{gathered} \tag{15}$$

Therefore, the dot product of two vectors over the field of complex numbers is, in general, a complex number.

But how about a **length** of a vector (9)?

Let's try to cacluate the **length** of the vector w_1 with the (9) formula:

$$\sqrt{(w1, w1)} = \sqrt{\begin{bmatrix} 1+i, 1-i, 0 \end{bmatrix} \begin{bmatrix} 1+i \\ 1-i \\ 0 \end{bmatrix}} = \sqrt{(1+i)^2 + (1-i)^2} = \sqrt{2i - 2i} = \sqrt{0} = 0 \tag{16}$$

Doesn't this look like a BUG?, Is the length of a vector with non-zero coordinates zero?

That is, we see that our algebra for real numbers breaks down in the case of complex numbers.

Consequently, mathematicians figured out how to get around this, they introduced a new definition of the dot product in complex space, and called it **inner product**:

$$||v|| = \sqrt{(v^\dagger, v)} \tag{17}$$

where the $v^\dagger$ means **hermitian conjugate**. Hermitian conjugate vector is a transpose vector v^T with additional **complex conjugation** v^* on it's components.

Let's find a hermitian conjugate of $w1 = \begin{bmatrix} 1+i \\ 1-i \\ 0 \end{bmatrix}$

$$w1^\dagger = \begin{bmatrix} 1-i, 1+i, 0 \end{bmatrix} \tag{18}$$

The important thing is that whenever we multiply a complex number by its complex conjugate we get a real number.

Example:

$$(4+7i)(4-7i) = 16 - 28i + 28i - 49i^2 = \\ 16 + 49 = 65 \tag{19}$$

That is, the answer is a purely real number - it has no imaginary part.

Let's use (17) to calculate (16):

$$\sqrt{(w1^\dagger, w1)} = \sqrt{\begin{bmatrix}1-i, 1+i, 0\end{bmatrix}\begin{bmatrix}1+i\\1-i\\0\end{bmatrix}} = \\ \sqrt{(1-i)(1+i)+(1+i)(1-i)+0} = \sqrt{2+2} = \sqrt{4} = 2 \tag{20}$$

And we see, that with our fixed dot product for complex spaces (**inner product**), we always get a length as a non negative, real number.

Important. There is no reason for the inner product of two states to be real, unless they are the same state.

Important. The inner product satisfies $(\phi, \psi) = (\psi, \phi)^*$

Important. Actually, for complex spaces we have to write $(\phi^\dagger, \psi)$, which is just (ϕ, ψ) for Euclidean spaces.

Important. Many textbooks define $\phi^\dagger \cdot \psi$ as a Hermitian product

Hence for the formula (2) we have

- probability of finding state $|0\rangle$

$$|\alpha|^2 = \alpha^\dagger \cdot \alpha \tag{21}$$

- probability of finding state $|1\rangle$

$$|\beta|^2 = \beta^\dagger \cdot \beta \tag{22}$$

-and

$$|\alpha|^2 + |\beta|^2 = 1 \\ |\psi\rangle = \alpha\begin{bmatrix}1\\0\end{bmatrix} + \beta\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}\alpha\\\beta\end{bmatrix} \longleftarrow \text{unit vector} \\ |\,|\psi\rangle\,| = 1 \text{ or} \\ \langle\psi|\psi\rangle = 1 \tag{23}$$

Example.

A qubit is given:

$$|\psi\rangle = -\frac{4}{5}i\,|0\rangle + \frac{3}{5}\,|1\rangle$$

1) If a $|\psi\rangle$ is a valid qubit.

2) What is the probability of finding $|0\rangle$ state and the $|1\rangle$ state.

Solution for part 1.

The $|\,|\psi\rangle\,|$ (and the $\langle\psi|\psi\rangle$) must be equals to 1. Hence:

$$|-\frac{4}{5}|^2 + \left|\frac{3}{5}\right| = 1$$

- according to the (21) and (22)

$$\left(\frac{4}{5}i \cdot \frac{-4}{5}i\right) + \left(\frac{3}{5} \cdot \frac{3}{5}\right) = \frac{16}{25} + \frac{9}{25} = 1$$

- i.e. it is a valid qubit.

Solution for part 2.

Prob. for finding $|0\rangle$ state: $\left|\frac{-4}{5}\right|^2 = \frac{16}{25} = 0.64$

Prob. for finding $|1\rangle$ state: $\left|\frac{3}{5}\right|^2 = \frac{9}{25} = 0.36$

Example

Two qubits is given:

$$|\psi\rangle = -\frac{4}{5}i\,|0\rangle + \frac{3}{5}\,|1\rangle$$
$$|\phi\rangle = \frac{|0\rangle}{\sqrt{2}} + \frac{|1\rangle}{\sqrt{2}}$$

Now we want to combine this two qubits. To combine qubits we have to compute tensor products of them $|\psi\rangle \otimes |\phi\rangle \sim |\psi\phi\rangle$ (and we will get a pure state again!).

$$|\psi\phi\rangle =$$
$$\left(\frac{-4}{5}i\,|0\rangle + \frac{3}{5}\,|1\rangle\right) \otimes \left(\frac{|0\rangle}{\sqrt{2}} + \frac{|1\rangle}{\sqrt{2}}\right) =$$
$$\frac{-4i}{5\sqrt{2}}\,|00\rangle + \frac{-4i}{5\sqrt{2}}\,|01\rangle + \frac{3}{5\sqrt{2}}\,|10\rangle + \frac{3}{5\sqrt{2}}\,|11\rangle =$$
$$\alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \theta\,|11\rangle$$

- and the next constraint must be met

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\theta|^2 = 1$$

We can also find probabilities of endividual outcome:

$$|\alpha|^2 = 0.32$$
$$\text{etc.}$$

3.10. Basis.

Definition. We say that 2 vectors are **orthogonal** if they are perpendicular to each other. i.e. the dot product of the two vectors is zero.

Definition. We say that a set of vectors $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ are mutually orthogonal if every pair of vectors is orthogonal. i.e.

$$\vec{v}_i \cdot \vec{v}_j = 0, \text{for all } i \neq j \tag{24}$$

Example The set of vectors $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$ is mutually orthogonal.

$$\begin{aligned} (1,0,-1)\cdot(1,\sqrt{2},1) &= 0 \\ (1,0,-1)\cdot(1,-\sqrt{2},1) &= 0 \\ (1,\sqrt{2},1)\cdot(1,-\sqrt{2},1) &= 0 \end{aligned} \tag{25}$$

Definition. A set of vectors S is **orthonormal** if every vector in S has magnitute 1 and the set of vectors are mutually orthogonal

$$\begin{aligned} \vec{v}_i \cdot \vec{v}_j &= 0, \text{for all } i \neq j \\ \vec{v}_i \cdot \vec{v}_j &= 1, \text{for all } i == j \end{aligned} \tag{26}$$

The vectors from the example above are mutually orhtogonal, but they are not notrmilized (this term is sometimes used to say that the vectors are not magnitude 1). Let

$$\begin{aligned} \vec{u}_1 = \frac{\vec{v}_1}{|\vec{v}_1|} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \\ \vec{u}_2 = \frac{\vec{v}_2}{|\vec{v}_2|} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{2}}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \\ \vec{u}_3 = \frac{\vec{v}_3}{|\vec{v}_3|} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{\sqrt{2}}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned} \tag{27}$$

The set of vectors $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ is **orthonormal**.

Proopsition An orthogonal set of non-zero vectors is linearly independent.

In the linear algebra, a set of linearly independent vectros usually are called **basis** set.

In the linear algebra, it is easier to choise some orthogonal basis for the computation. Because of the formula for computing the scalr product (69) and (28) is simplified. And even better to choice the orthonormal basis because of the (75), it is because the unit vector just show us a direction.

Example.

The three vectors

$$\vec{v}_1 = [1,2,1]^T, \vec{v}_2 = [2,1,-4]^T, \vec{v}_3 = [3,-2,1]^T$$

are mutually orthogonal. Express the vector $\vec{v} = [7,1,9]^T$ as a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3$

Set

$$x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{v}$$

There are two ways to find x_1, x_2, x_3

Method 1: Solving the linear system by performing row operations to its augmented matrix

$$[\vec{v}_1, \vec{v}_2, \vec{v}_3 | \vec{v}]$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 2 & 1 & -2 & 1 \\ 1 & -4 & 1 & 9 \end{array}\right]$$

we obtain $x_1 = 3, x_2 = -1, x_3 = 2$.

So $\vec{v} = 3\vec{v}_1 - \vec{v}_2 + 2\vec{v}_3$

Method 2: Since $\vec{v}_i \perp \vec{v}_j$ for $i \neq j$, we have

$$\langle \vec{v}_i | \vec{v} \rangle = \langle \vec{v}_i | x_1 \vec{v}_1 + x_2 \vec{v}_2 + x_3 \vec{v}_3 \rangle = x_i (\vec{v}_i, \vec{v}_i) \tag{28}$$

-hint, since $\vec{v}_i \perp \vec{v}_j$ for $i \neq j$ are vanish.

where $i = 1, 2, 3$. Then

$$x_i = \frac{\langle \vec{v}_i | \vec{v} \rangle}{\langle \vec{v}_i | \vec{v}_i \rangle}, \ i = 1, 2, 3$$

We then have

$$\begin{aligned} x_1 &= \frac{7 + 2 + 9}{1 + 4 + 1} = \frac{18}{6} = 3 \\ x_2 &= \frac{14 + 1 - 36}{4 + 1 + 16} = \frac{-21}{21} = -1 \\ x_3 &= \frac{21 - 2 + 9}{9 + 4 + 1} = \frac{28}{14} = 2 \end{aligned}$$

Theorem. Let $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k$ be an orthogonal basis of a subspace W. Then for any $\vec{v} \in W$,

$$\vec{w} = \frac{(\vec{v}_1, w)}{(\vec{v}_1, \vec{v}_1)} \vec{v}_1 + \frac{(\vec{v}_2, w)}{(\vec{v}_2, \vec{v}_2)} \vec{v}_2 + \cdots + \frac{(\vec{v}_k, w)}{(\vec{v}_k, \vec{v}_k)} \vec{v}_k \tag{29}$$

Now let's return to the (3):

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Let's show that basis vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are orthonormal.

$$\langle 0 | 1 \rangle = (1; 0) \cdot (0; 1) = 1 \cdot 0 + 0 \cdot 1 = 0 \quad \text{(othrogonality)}$$

$$\begin{aligned} \langle 0 | 0 \rangle &= (1; 0) \cdot (1; 0) = 1 \cdot 1 + 0 \cdot 1 = 1 \quad \text{(unit vector)} \\ \langle 1 | 1 \rangle &= (0; 1) \cdot (0; 1) = 0 \cdot 0 + 1 \cdot 1 = 1 \quad \text{(unit vector)} \end{aligned}$$

3.11. Dual space

The bra $\langle \phi |$ associated with a ket $|\psi\rangle$ is the linear transformation that maps any ket $|\psi\rangle$ to the complex number $\langle \phi | \psi \rangle$. In the linear algebra if you take a vector and generate the inner product of that vector with all other vectors then that is a **linear transformation** (linear functional, linear form, one-form, covector) from the vector space to the complex numbers.

Therefore, the (8) can be written as:

$$\varphi = f_\phi(\psi) = a_1 b_1 + \cdots + a_n b_n \tag{30}$$

- where the φ is some number (complex in general).

This can be interpreted as either the matrix product or the dot product of the row vector ϕ and the column vector ψ

$$\varphi = f_\phi(\psi) = \langle\phi|\psi\rangle = \begin{bmatrix} a_1, \cdots, a_n \end{bmatrix} \begin{bmatrix} b_1, \\ \cdots, \\ b_n \end{bmatrix} \tag{31}$$

From this notation we can see that indeed, the bra vector $f_\phi()$ is some linear functional. Later we will consider it's properties in more detail.

So, to get a bra vector from a ket vector, we need to perform two operations:

1) Changing the sign of the imaginary part of the vector components, This corresponds to the mirroring of a vector around the real axis.

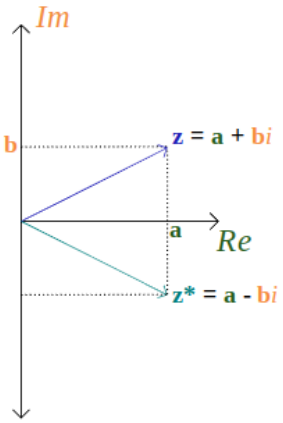


Figure 1. Complex Conjugate

And it solves the problem with the length of a vector (see above).

Also se can see that we can construct complex conjugate vector for any given vector.

2) Tanspose the vector column to get the vector row.

But why we have to do this.

To answer this question, let's first look at how a transposed matrix arises in linear algebra.

Let's consider two bases:

$O\boldsymbol{e_1}\boldsymbol{e_2}\boldsymbol{e_3}$ - the "old one",

$O^{'}\boldsymbol{e_1}^{'}\boldsymbol{e_2}^{'}\boldsymbol{e_3}^{'}$ - and the "new one".

And let's take that O and $O^{'}$ it is the same point in space - the origin.

Then the new basis is well defined if vecotrs $\boldsymbol{e}^{'}, \boldsymbol{e}^{'}, \boldsymbol{e}^{'}$ is defined by coordinates in "old one" basis, i.e. if a_{ik} $i, k = 1, 2, 3$ is defined in equations

$$\left.\begin{array}{l} \boldsymbol{e_1}^{'} = a_{11}\boldsymbol{e_1} + a_{21}\boldsymbol{e_2} + a_{31}\boldsymbol{e_3} \\ \boldsymbol{e_2}^{'} = a_{12}\boldsymbol{e_1} + a_{22}\boldsymbol{e_2} + a_{32}\boldsymbol{e_3} \\ \boldsymbol{e_3}^{'} = a_{13}\boldsymbol{e_1} + a_{23}\boldsymbol{e_2} + a_{33}\boldsymbol{e_3} \end{array}\right\} \tag{32}$$

The matrix

$$A^{*} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \tag{33}$$

is called the **transition matrix** from the $\boldsymbol{e}, \boldsymbol{e}, \boldsymbol{e}$ to the $\boldsymbol{e}^{'}, \boldsymbol{e}^{'}, \boldsymbol{e}^{'}$ basis.

Since the basis vectors $\boldsymbol{e}^{'}, \boldsymbol{e}^{'}, \boldsymbol{e}^{'}$ are linearly independent, then the determinant of the matrix A^{*} is non-zero - the transition matrix from one basis to another is always a non-singular matrix.

Since the vectors $\boldsymbol{e}^{'}, \boldsymbol{e}^{'}, \boldsymbol{e}^{'}$ are the basis vectors therefore each of $\boldsymbol{e}, \boldsymbol{e}, \boldsymbol{e}$ vectors in turn can be uniquely represented as linear combination of the vectors $\boldsymbol{e}^{'}, \boldsymbol{e}^{'}, \boldsymbol{e}^{'}$:

$$\left.\begin{array}{l} \boldsymbol{e_1} = a_{11}^{'}\boldsymbol{e_1}^{'} + a_{21}^{'}\boldsymbol{e_2}^{'} + a_{31}^{'}\boldsymbol{e_3}^{'} \\ \boldsymbol{e_2} = a_{12}^{'}\boldsymbol{e_1}^{'} + a_{22}^{'}\boldsymbol{e_2}^{'} + a_{32}^{'}\boldsymbol{e_3}^{'} \\ \boldsymbol{e_3} = a_{13}^{'}\boldsymbol{e_1}^{'} + a_{23}^{'}\boldsymbol{e_2}^{'} + a_{33}^{'}\boldsymbol{e_3}^{'} \end{array}\right\} \tag{34}$$

Equations above are uniquely solvable with respect to the old unit vectors $\boldsymbol{e}, \boldsymbol{e}, \boldsymbol{e}$

Now, let's consider how the coordinates x, y, z and $x^{'}, y^{'}, z^{'}$ of some point M (vector $\boldsymbol{u} = \overrightarrow{OM}$) are related to each other in the "old one" and "new one" coordinate systems.

Because the $\boldsymbol{u}$ it is the same vector in different coordinate systems:

$$\boldsymbol{u} = x\boldsymbol{e_1} + y\boldsymbol{e_2} + z\boldsymbol{e_3} = x^{'}\boldsymbol{e_1}^{'} + y^{'}\boldsymbol{e_2}^{'} + z^{'}\boldsymbol{e_3}^{'} \tag{35}$$

Let's introduce into this identity the expressions from (32),

so we get:

$$\begin{gathered} \boldsymbol{u} = \\ x\boldsymbol{e_1} + y\boldsymbol{e_2} + z\boldsymbol{e_3} = \\ x^{'}(a_{11}\boldsymbol{e_1} + a_{21}\boldsymbol{e_2} + a_{31}\boldsymbol{e_3}) + y^{'}(a_{12}\boldsymbol{e_1} + a_{22}\boldsymbol{e_2} + a_{32}\boldsymbol{e_3}) + z^{'}(a_{13}\boldsymbol{e_1} + a_{23}\boldsymbol{e_2} + a_{33}\boldsymbol{e_3}) \end{gathered} \tag{36}$$

- or, after grouping the terms

$$\begin{gathered}\boldsymbol{u} = \\ x\boldsymbol{e_1} + y\boldsymbol{e_2} + z\boldsymbol{e_3} = \\ (a_{11}\boldsymbol{x}^{'} + a_{12}\boldsymbol{y}^{'} + a_{13}z^{'})\boldsymbol{e_1} + (a_{21}\boldsymbol{x}^{'} + a_{22}\boldsymbol{y}^{'} + a_{23}z^{'})\boldsymbol{e_2} + (a_{31}\boldsymbol{x}^{'} + a_{32}\boldsymbol{y}^{'} + a_{33}z^{'})\boldsymbol{e_3}\end{gathered} \tag{37}$$

Looking at the equation above we can see that:

$$\left.\begin{aligned} x &= a_{11}x^{'} + a_{12}y^{'} + a_{13}z^{'} \\ y &= a_{21}x^{'} + a_{22}y^{'} + a_{23}z^{'} \\ z &= a_{31}x^{'} + a_{32}y^{'} + a_{33}z^{'} \end{aligned}\right\} \tag{38}$$

These formulas express the "old one" coordinates x, y, z of the point M (vector $\mathbf{u}$) in terms of the "new ones". The matrix, gives this expression:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{33} & a_{33} \end{bmatrix} \tag{39}$$

- is called coortinate transform matrix, and it is **transposed** with respect to the matrix A^* transition from basis $\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}$ to the basis $\boldsymbol{e_1^{'}}, \boldsymbol{e_2^{'}}, \boldsymbol{e_3^{'}}$

The fact that basis elements $\boldsymbol{e}, \boldsymbol{e}, \boldsymbol{e}$ change in one way A^* while coordinates (x, y, z) change in the inverse way A is why we sometimes call the basis elements **covariant** and the vector coordinates **contravarian**.

Numbers (x, y, z) and $(x^{'}, y^{'}, z^{'})$ from the (35) are called **contravariant** components of the vector

Now let's consider the (30) again. As we already mention the $f_\phi()$ function assigns a number φ to each vector ψ. The dot product (Euclidean space) is linear function of a vector. It means that

$$f_\phi(\boldsymbol{\psi_1} + \boldsymbol{\psi_2}) = f_\phi(\boldsymbol{\psi_1}) + f_\phi(\boldsymbol{\psi_2}) \tag{40}$$

and

$$f_\phi(\alpha\boldsymbol{\psi}) = \alpha f_\phi(\boldsymbol{\psi}) \tag{41}$$

Let's consider the $f_\phi(\psi)$ in some coordinate system, i.e let's define argument ψ by it's coordinates

$$\boldsymbol{\psi} = x_1\boldsymbol{e_1} + \cdots x_n\boldsymbol{e_n} \tag{42}$$

Then using (40) and (41) we get

$$f_\phi(\boldsymbol{\psi_1}) = f_\phi(x_1\boldsymbol{e_1} + \cdots x_n\boldsymbol{e_n}) = x_1 f_\phi(\boldsymbol{e_1}) + \cdots + x_n f_\phi(\boldsymbol{e_n}) \tag{43}$$

Let us introduce the notation for brevity:

$$\varphi_i = f_\phi(\boldsymbol{e_i}) \tag{44}$$

The we can write the (43) in the following way:

$$f_\phi(\boldsymbol{\psi}) = \varphi_i x_i \tag{45}$$

- here we mean summation over index i

The (45) is also called **linear form**, it is not a vector. I.e. linear function of a vector $f_\phi(\boldsymbol{\psi})$ is expressed in terms of its coordinates by linear form (45)

Let's write dependency (45) in some other coordinate system (like we did for a vector):

$$f_\phi(\boldsymbol{\psi}) = \varphi_i^{'} x_i{}' \tag{46}$$

The function $f_\phi(\boldsymbol{\psi})$ is remains the same, but since x_i - vector agrument coordinates changes to $x_i{}'$ then the coefficients φ_i of linear form (45) have to change also.

Let's find out how they change. Let's apply the (44) in a new coordinate system:

$$\varphi_i^{'} = f_\phi(\boldsymbol{e_i^{'}}) \tag{47}$$

And, according to the (32):

$$\boldsymbol{e_i^{'}} = \sum_j a_{ji} \boldsymbol{e_j} \tag{48}$$

Then, using (40) and (41) we can rewrite (47) in the form of:

$$\varphi_i^{'} = \varphi(a_{1i}\boldsymbol{e_1} + \cdots + a_{ni}\boldsymbol{e_n}) = a_{1i} f_\phi(\boldsymbol{e_1}) + \cdots + a_{ni} f_\phi(\boldsymbol{e_n}) \tag{49}$$

- comparing with the (44) we are getting:

$$\varphi_i^{'} = \sum_j a_{ji} \varphi_j \tag{50}$$

The $\varphi_i^{'}$ - is the **covariant** components of the vector ψ (the smae as $\boldsymbol{u}$ (35))

Comparing (50) and (48), we can see that components $\varphi_i^{'}$ have the same transformation law as the "basis vector" not a "regular vector".

And recollectiong the (31) now we can conclude that indeed the "bra" vector is transposed with respect the "ket" vector - because the bra vector is a linear form - not a vector, and have different components transformation law.

Also, we can rewrite the (38) as follows

$$\psi_j^{'} = \sum_i a_{ji} x_i \tag{51}$$

We see that the index of φ from the (50) corresponds to the second index of matrix a, and the index of ψ from the (51) corresponds to the first argument of matrix a.

3.12. Bra vector

The famous physicist and matematician Paul Dirac suggested using the following notation for inner product:

$$\langle \phi | \psi \rangle \tag{52}$$

where $\langle \phi |$ means $(|\phi\rangle)^\dagger$ and it is called **bra** vector.

The (52) is an inner product.

Quantum mechanics is a set of principles. Here is one of the most important:

The probability of an event occurring is equal to the square of the absolute value of the inner product of the final and initial state vectors.

$$P_(\psi \Rightarrow \phi) = \| \langle \phi | \psi \rangle \|^2 \tag{53}$$

The (53) gives the probability of finding state $|\psi\rangle$ in the state $|\phi\rangle$

The (53) is also known as a **Born's Rule** (simplified):

The probability of a state $|\psi\rangle$ **"collaplsing"** onto state $|\phi\rangle$ is equal to (53)

3.13. Stern and Gerlach experiment

Each electron behaves as a magnet but with only two possible directions. Spin up and spin down (related to the external magnetic field applied). This behavior is described by the **Stern and Gerlach experiment**.

Suppose that we know that initially electron spin is up. Then the probability that after Stern and Gerlach experiment we will find an electron with the spin up is

$$P_\uparrow = \left| \left\langle [1,0] \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle \right|^2 = |1 \cdot 1 + 0 \cdot 0|^2 = 1 \tag{54}$$

i.e with the 100% it will be a spin up $\uparrow$ state.

Let's consider **another experiment**. Suppose that we know that initially electron spin is down. The probability that after Stern and Gerlach experimant we will find and electron with the spin up is

$$P_\downarrow = \left| \left\langle [1,0] \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle \right|^2 = |1 \cdot 0 + 0 \cdot 1|^2 = 0 \tag{55}$$

If the spin is prepared *up*, then the probability to detect it *down* is zero, and vice versa. Two orthogonal states are physically distinct and mutually exclusive. If the spin is one of these states, it *cannot* be (has zoro probability to be) in the other one. This idea applies to all quantum systems, not just spin. But don't mistake the orthogonality of the state-vectors for orthogonal directions in space. In fact, the directions *up* and *down* are not orthogonal directions in space, even though their associated state-vectors are orthogonal in state space.

Let's consider **another experiment**. Suppose that we know that initially electron spin is down. What is the probability that it ends up in the $\begin{bmatrix} \frac{-i}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$ state? (Actually it is a "left" state).

$$P_\leftarrow = \left| \left\langle \frac{1}{\sqrt{2}} [i,-1] \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle \right|^2 = \left| \frac{-1}{\sqrt{2}} \right|^2 = \frac{1}{2} \tag{56}$$

We get 0.5, that is, the probability of 50%.

We also could introduce other notation:

$$\begin{aligned} |\langle \uparrow | \uparrow \rangle|^2 &= 1 \\ |\langle \uparrow | \downarrow \rangle|^2 &= 0 \\ |\langle \leftarrow | \downarrow \rangle|^2 &= \frac{1}{2} \end{aligned} \tag{57}$$

4. Projections.

4.1. Orthogonal Projection. Outer product.

Let's first recollect how to **add and subtract vectors.**

4.2. Add and subtract vectors

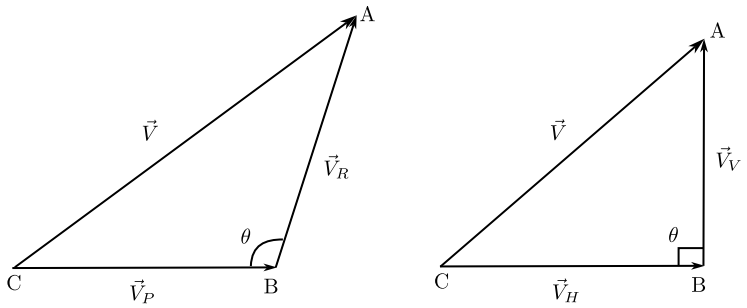


Figure 2. Triangle Law of Vectors

a. The 2 vectors are in Continuation with each other i.e Head Point of one vector is the Tail Point of the Other Vector. The vector containing the Common point as Tail is called the Leading Vector. The Vector containing the Common Point as Head is called the Trailing Vector.

b. The 2 ectors share a common Tail Point.

c. The 2 vectors share a common Head Point.

Triangular Law of Vector Addition states that If 2 Vectors that are in Continuation (i.e. Head Point of one Vector is the Tail Point of the Other Vector) are added then the Resultant Vector has the Head Point of the Lead Vector and Tail Point of the Trailing Vector. Since the 2 Vectors and the Resultant Vector form a Triangle, it is know as the Triangular Law of Vector Addition.

$$\vec{AC} = \vec{AB} + \vec{BC} = \vec{BC} + \vec{AB} \tag{58}$$

And as per Cosine Law of Triangles

$$|\vec{AC}| = |\vec{CA}| = \sqrt{|\vec{AB}|^2 + |\vec{BC}|^2 - 2|\vec{AB}||\vec{BC}|\cos\theta} \tag{59}$$

Triangle Law can be used to Find the Resultant Vector under following conditions

a. Addition of 2 Vectors that are in Continuation with each other: The resultant Vector of addition of 2 given vectors $\vec{AB}$ and $\vec{BC}$ is given as

$$\vec{AB} + \vec{BC} = \vec{BC} + \vec{AB} = \vec{AC} \tag{60}$$

b. Subtraction of 2 Vectors that Have a Common Tail Point: The resultant Vector of Subtration of 2 Vectors $\vec{AB}$ and $\vec{CB}$ is given as

$$\begin{aligned} \vec{AB} - \vec{CB} &= \vec{AB} + \vec{BC} = \vec{AC} \\ \vec{CB} - \vec{AB} &= \vec{CB} + \vec{BA} = \vec{CA} \end{aligned} \tag{61}$$

c. Subtraction of 2 Vectors that Have a Common Head Point: The resultant Vector of Subtration of 2 Vectors AB and AC is given as

$$\begin{aligned} \vec{AB} - \vec{AC} &= \vec{AB} + \vec{CA} = \vec{CA} + \vec{AB} = \vec{CB} \\ \vec{AC} - \vec{AB} &= \vec{AC} + \vec{BA} = \vec{BA} + \vec{AC} = \vec{BC} \end{aligned} \tag{62}$$

Any vector $\vec{V}$ can be represented as a sum of its Projection vector ($\vec{V_P}$) and Rejection vector ($\vec{V_R}$) on any other vector. That is

$$\vec{V} = \vec{V}_P + \vec{V}_V \tag{63}$$

When the vector $\vec{V}$'s Projection and Rejection vectors are perpendicular to each other then Projection vector ($\vec{V}_P$) is also called its Horizontal component ($\vec{V}_H$) and Rejection vector ($\vec{V}_R$) is also called its Vertical component ($\vec{V}_V$). Hence, any vector $\vec{V}$ can be presented as sum of its Horizontal ($\vec{V}_H$) and Vertical ($\vec{V}_V$) components based on any other vector. That is

$$\vec{V} = \vec{V}_H + \vec{V}_V \tag{64}$$

Hence:

$$\vec{V}_V = \vec{V} - \vec{V}_H \tag{65}$$

4.3. Orthogonal Scalar Projection

Definition of the $\cos(\alpha)$ function.

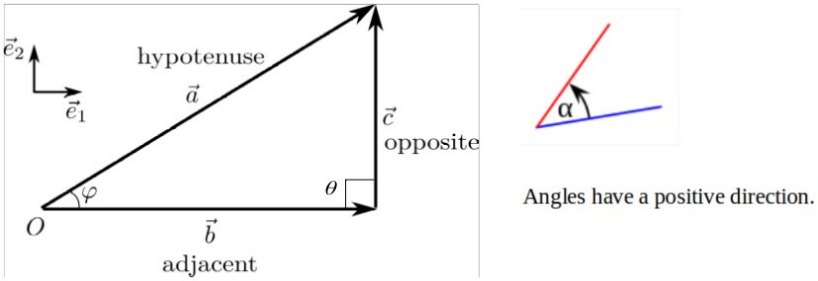


Figure 3. Definition of the $\cos(\alpha)$ function.

$$\cos(\varphi) = \frac{\text{adjacent}}{\text{hypotenuse}} = \frac{|\vec{b}|}{|\vec{a}|} \tag{66}$$

$$|\vec{b}| = \text{adjacent} = \text{hypotenuse} \cdot \cos(\varphi) = |\vec{a}| \cdot \cos(\varphi) \tag{67}$$

$$|\vec{c}| = |\vec{a}| \sin\varphi \tag{68}$$

$$|\vec{a}|^2 = \sqrt{|\vec{b}|^2 + |\vec{c}|^2 - 2|\vec{b}||\vec{c}|\cos\theta} = \sqrt{|\vec{b}|^2 + |\vec{c}|^2} \tag{69}$$

That is why it is easier to choice orthogonal basis for the computations - because the $\cos\theta = 0$ and the formula (69) is simplified.

The *adjacent* is a scalar projection of *hypotenuse* onto the direction parallel to the *adjacent*.

Let's introduce new notation for a scalar orthogonal projection:

$$\text{Proj}_{\vec{b}}\vec{a} = |\vec{a}| \cos\left(\widehat{\vec{b}, \vec{a}}\right) \tag{70}$$

In the case of figure (3)

$$|\vec{b}| = \text{Proj}_{\vec{b}}\vec{a} \tag{71}$$

- in this case:

$$|\vec{b}||\vec{b}| = |\vec{b}|^2 = |\vec{b}| \cdot \text{Proj}_{\vec{b}}\vec{a} = |\vec{b}||\vec{a}| \cos\left(\widehat{\vec{b}, \vec{a}}\right) = (\vec{b}, \vec{a}) \tag{72}$$

But this is a special case.

The $\vec{b}$ vector can have arbitrary length but the same direction. In this case:

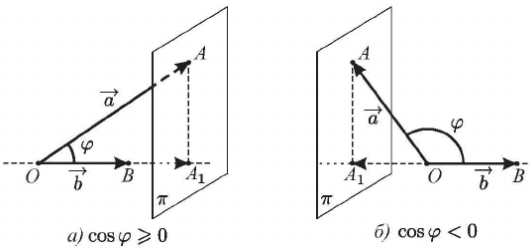


Figure 4. Orthogonal Scalar Projection

$$(\vec{b},\vec{a}) = |\vec{b}| \cdot \text{Proj}_{\vec{b}}\vec{a} = |\vec{b}||\vec{a}| \cos\left(\widehat{\vec{b},\vec{a}}\right) \neq |\vec{b}|^2 \tag{73}$$

$$\cos\left(\widehat{\vec{b},\vec{a}}\right) = \frac{(\vec{b},\vec{a})}{|\vec{b}||\vec{a}|} \tag{74}$$

So, in general, in the (70) we are taking the projection in the direction of $\vec{b}$. And the $\vec{b}$ can have an arbitrary length. If we take the length of the $\vec{b}$ equals to 1, and denote it as $\vec{e}_1$ - **unit vector** collinear with $\vec{b}$ the result will be:

$$|\vec{b}| = \text{Proj}_{\vec{e}_1}\vec{a} = |\vec{a}| \cos\left(\widehat{\vec{e}_1,\vec{a}}\right) = |\vec{a}||\vec{e}_1| \cos\left(\widehat{\vec{e}_1,\vec{a}}\right) = (\vec{e}_1,\vec{a}) = a_1 \tag{75}$$

I.e, **the scalar projection of some vector onto it's first basis vector coincides with it's first coordinate -** $|\vec{b}| = a_1$.

Therefore:

$$(\vec{b},\vec{a}) = a_1|\vec{b}| = (\vec{e}_1,\vec{a})|\vec{b}| \tag{76}$$

The (73) and (76) is a geometrical definition of a dot product.

In multidimensional euclidean space:

$$\begin{aligned} \vec{a} &= [a_1, \cdots, a_n] = \sum_i a_i\vec{e}_i \\ \vec{b} &= [b_1, \cdots, b_n] = \sum_i b_i\vec{e}_i \end{aligned} \tag{77}$$

$$(\vec{e}_i,\vec{a}) = |\vec{a}| \cos\left(\widehat{\vec{e}_i,\vec{a}}\right) = |\vec{e}_i||\vec{a}| \cos(\varphi_i) = a_i \tag{78}$$

The algebraic definition of a dot product is:

$$(\vec{b},\vec{a}) = \sum_i b_i a_i = \sum_i a_i b_i \tag{79}$$

Hence:

$$(\vec{b},\vec{a}) = \sum_i b_i a_i = \sum_i a_i b_i = \sum_i \vec{b}_i(\vec{e}_i,\vec{a}) = \sum_i \vec{a}_i(\vec{e}_i,\vec{b}) \tag{80}$$

4.4. Linearity of orthogonal scalar projection

The scalar projection is a linear operation. I.e the projection of the sum of vectors is the same as the sum of the projections, and the scalar factor can be taken out of the brackets:

$$\text{Proj}_{\vec{b}}(\vec{a}_1 + \vec{a}_2) = \text{Proj}_{\vec{b}}(\vec{a}_1) + \text{Proj}_{\vec{b}}(\vec{a}_2) \tag{81}$$

$$\text{Proj}_{\vec{b}}(\lambda\vec{a}) = \lambda\text{Proj}_{\vec{b}}\vec{a} \tag{82}$$

4.5. Orthogonal Vector Projection

The (70) is a scalar and as we already show, it does not dependent on the length of a $\vec{b}$. To make this quantity directional, we can multiply it by a unit vector $\vec{e}_1$

$$\overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = |\vec{a}|\cos\left(\widehat{\vec{b},\vec{a}}\right)\vec{e}_1 = a_1\vec{e}_1 \tag{83}$$

Multiplying a vector by a scalar simply scales the vector without changing its direction.

Example:

Let $\vec{u} = [-1, 3]$. Find $7\vec{u}$:

$7\vec{u} = [-7, 21]$

Since:

$$\begin{gathered}\vec{e}_1 = \frac{\vec{b}}{|\vec{b}|} \\ \vec{b} = |\vec{b}|\vec{e}_1 = \text{Proj}_{\vec{e}_1}\vec{a}\cdot\vec{e}_1 = a_1\vec{e}_1\end{gathered} \tag{84}$$

- then:

$$\overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = \frac{|\vec{a}|\cos\left(\widehat{\vec{b},\vec{a}}\right)}{|\vec{b}|}\vec{b} \tag{85}$$

- substitute the (74):

$$\overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = \frac{(\vec{b},\vec{a})}{|\vec{b}|^2}\vec{b} \tag{86}$$

Let's donote it as:

$$\overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = \lambda\vec{b} = \lambda a_1\vec{e}_1 = \lambda|\vec{b}|\vec{e}_1 = \lambda\vec{b} \tag{87}$$

-where:

$$\begin{gathered}\lambda = \frac{|\vec{a}|\cos(\varphi)}{|\vec{b}|} = \frac{(\vec{b},\vec{a})}{|\vec{b}|^2} \\ \cos(\varphi) = \cos\left(\widehat{\vec{b},\vec{a}}\right)\end{gathered} \tag{88}$$

The λ coefficient is called **Fourier coefficient.**

4.6. Properties of a dot product

In the (3.8) we already considered some of the properties of the dot product. Let's consider the properties of the scalar product in more detail.

Let V be a vector space over $\mathbb{R}$.

A dot product $(\cdot, \cdot)$ is a function $V \times V \to \mathbb{R}$ with the following properties

Property 1: Dot product with itself is non-negative

$$\begin{gathered} \forall \vec{a} \in V \\ (\vec{a}, \vec{a}) \geqslant 0 \end{gathered} \tag{89}$$

We know that the $(\vec{a}, \vec{a}) = |\vec{a}|^2$ - the lenght of a vector. The length cannot be negative.

Property 2: Dot product with self is zero iff zero vector

$$(\vec{a}, \vec{a}) = 0 \Leftrightarrow \vec{a} = 0 \tag{90}$$

Property 3: Orthogonality.

If $(\vec{b}, \vec{a}) = 0$ then it can be clearly seen that either $\vec{a}$ or $\vec{b}$ is zero or $\cos(\Theta) = 0 \Rightarrow \Theta = \frac{\pi}{2}$

It suggests that either of the vectors is zero or they are perpendicular to each other.

Property 4: Norm of a vector.

The dot product of a vector to itself is the magnitude squared of the vector i.e.

$(\vec{a}, \vec{a}) = |\vec{a}||\vec{a}| \cos(0) = |\vec{a}|^2$ which is a square of the length of a vector.

Property 5: Scalar projection.

$a_b = |\vec{a}| \cos(\theta)$ where the θ is the angle between $\vec{a}$ and $\vec{b}$

In terms of the geometric definition of the dot product, this can be rewritten as

$a_b = (\hat{b}, \vec{a})$

where $\hat{b} = \frac{\vec{b}}{|\vec{b}|}$

The dot product is thus characterized geometrically by

$(\vec{b}, \vec{a}) = a_b|\vec{b}| = b_a|\vec{a}|$

Property 6: Disctibutive over vector addition.

$$(\vec{a}, \vec{b} + \vec{c}) = (\vec{a}, \vec{b}) + (\vec{a}, \vec{c}) \tag{91}$$

Property 7: Dot product is **symmetric**, i.e - commutative.

$$(\vec{a}, \vec{b}) = (\vec{b}, \vec{a}) \tag{92}$$

This property follows from the geometric (73) and algebraic (79) definition of a dot product.

But the **projections are not symmetric**.

Property 8: Bilinearity. Bilinear means linear in every argument.

$$\forall \vec{a}, \vec{a}, \vec{c} \in V \text{ and } \forall \lambda, \mu \in \mathbb{R}$$

$$(\lambda \vec{a} + \mu \vec{b}, \vec{c}) = \lambda(\vec{a}, \vec{c}) + \mu(\vec{b}, \vec{c}) \tag{93}$$

$$(\vec{a}, \lambda\vec{b} + \mu\vec{c}) = \lambda(\vec{a}, \vec{b}) + \mu(\vec{a}, \vec{c}) \tag{94}$$

The (93) and the (94) follows from the linearity of a scalar orthogonal projection - (81) and (82)

Property 9: Not associative.

Formally, a binary operation $*$ on a set S is called **associative** if it satisfies the associative law:

$$\begin{gathered} \forall x, z, z \text{ in } S \\ (x * y) * z = x * (y * x) \end{gathered} \tag{95}$$

But because the dot product between a scalar $(\vec{a}, \vec{b})$ and a vector $\vec{c}$ is not defined the associative law is **ill-defined**.

But the dot product is associative with respect to scalar multiplication $c(\vec{a}, \vec{b}) = (c\vec{a}, \vec{b}) = (\vec{a}, c\vec{b})$

Property 10: No cancellation.

Unlike multiplication of ordinary numbers, where if $ab = ac$, then b always equals c unless a is zero, the dot product does not obey the **cancelation law**.

If $(\vec{a}, \vec{b}) = (\vec{a}, \vec{c})$, then we can write: $(\vec{a}, \vec{b} - \vec{c}) = 0$ by the distribution law. The result above says this just means that $\vec{a}$ is perpendicular to $(\vec{b} - \vec{c})$, which still allows $\vec{b} - \vec{c} \neq 0$, and therefore allows $\vec{b} \neq \vec{c}$.

4.7. Properties of an inner product

As we already show in the (3.9) for a complex spaces we have to modify the dot product. In this case we also call it **hermitian product**.

Let's consider this modifications in more details.

Property 1: Symmetric property becomes conjugate-symetric.

The symmetry property (92) is no longer met. Indeed for any $\vec{a}$ and $\vec{b}$ in $\mathbb{C}$ so that:

$$\vec{a} = \begin{bmatrix} a_1 \\ \cdots \\ a_n \end{bmatrix}, \vec{b} = \begin{bmatrix} b_1 \\ \cdots \\ b_n \end{bmatrix} \tag{96}$$

$$\left\langle \vec{b} \middle| \vec{a} \right\rangle \sim (\vec{b}^{\dagger}, \vec{a}) \tag{97}$$

where the $\vec{b}^{\dagger}$ is a transposed vector with complex conjugate components.

The $\dagger$ operation - is called **Conjugate Transpose of a Vector**.

To swap $\vec{a}$ and $\vec{b}$ we have to apply conjugate transpose to both of them.

$$(\vec{b}^{\dagger}, \vec{a}) = (\vec{a}^{\dagger}, \vec{b})^{\dagger} \tag{98}$$

Therefore the hermitian product is conjugate-symmetric, but the dot product is symmetric.

Property 2: Linear as a function of its second argument.

$$\left\langle \vec{c} \middle| \alpha\vec{a} + \beta\vec{b} \right\rangle = \alpha \langle \vec{c} | \vec{a} \rangle + \beta \left\langle \vec{c} \middle| \vec{b} \right\rangle \tag{99}$$

Property 3: Antilinear function of tis first argument

$$\left\langle \alpha\vec{a} + \beta\vec{b} \middle| \vec{c} \right\rangle = \bar{\alpha} \langle \vec{a} | \vec{c} \rangle + \bar{\beta} \left\langle \vec{b} \middle| \vec{c} \right\rangle \tag{100}$$

4.8. Outer product

It is a standard matrix of the vector orthogonal projection.

Let's consider the (86) again.

As we remember, in the case of complex space, we have to write it in the form of:

$$\overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = \frac{\left\langle \vec{b} \middle| \vec{a} \right\rangle}{|\vec{b}|^2} \left| \vec{b} \right\rangle \tag{101}$$

- and since the $\left\langle \vec{b} \middle| \vec{a} \right\rangle$ it is a scalar we can rewrite it as follows:

$$\overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = \frac{1}{|\vec{b}|^2} \left| \vec{b} \right\rangle \left\langle \vec{b} \middle| \vec{a} \right\rangle \tag{102}$$

Hence, we have a formula:

$$\left\langle \vec{b} \middle| \vec{a} \right\rangle \left| \vec{b} \right\rangle = \left| \vec{b} \right\rangle \left\langle \vec{b} \middle| \vec{a} \right\rangle \tag{103}$$

And according to the vector(matrix) multiplication rules the $\left| \vec{b} \right\rangle \left\langle \vec{b} \right|$ produces a matrix.

Sometimes, the cross product is denoted as follows $\vec{c} \otimes \vec{b}$, then there is some more general relation:

$$\left(\vec{c} \otimes \vec{b} \right) \vec{a} = \vec{c} \left(\vec{b} \cdot \vec{a} \right) \tag{104}$$

This therefore takes a vector in the $\vec{a}$ direction and gives back one in a $\vec{c}$ direction.

To proof it just rewrite it in a bra-ket notation.

$$|\vec{c}\rangle \left\langle \vec{b} \middle| \vec{a} \right\rangle = |\vec{c}\rangle \left\langle \vec{b} \middle| \vec{a} \right\rangle \tag{105}$$

If we have an orthonormal basis, $\vec{e}_i, i = 1...N$, then the outer product $\vec{e}_i \otimes \vec{e}_j$ acting on a vector $\vec{a}$ picks out the j-th component of the vector and gives it back in the i-th direction; that is

$$\left(\vec{e}_i \otimes \vec{e}_j \right) \vec{a} = \vec{e}_i \left(\vec{e}_j \cdot \vec{a} \right) = a_j \vec{e}_i \tag{106}$$

The matrix that does this job is one with 1 in the j-th row and i-th columnt and 0 everywhere else.

4.9. Completeness Relation

Acctording to the (7) and the (77) arbitrary state ψ can be written as linear superposition of basis vectors (in computational basis, as linear combination of eigenkets)

$$|\psi\rangle = \sum_i a_i |e_i\rangle \tag{107}$$

Taking inner product with $\langle e_j|$, and using the ortho normality of $|e_i\rangle$, we have

$$\langle e_j|\psi\rangle = \sum_i \langle e_j| a_i |e_i\rangle = \sum_i a_i \langle e_j|e_i\rangle = \sum_i a_i \delta_{ij} = a_j \tag{108}$$

For 2D spaces this coincides with (75) and for multidimensional spaces it coincides with (78)

Substituting back in Eq. (107) we have

$$|\psi\rangle = \sum_i \langle e_i|\psi\rangle |e_i\rangle = \left(\sum_i |e_i\rangle \langle e_i|\right) |\psi\rangle \tag{109}$$

Similarly we can write any matrix as an outer product operator

$$A = \sum_{i,j=1}^{N} A_{ij} e_i \otimes e_j = \sum_{i,j=1}^{N} A_{ij} |e_i\rangle \langle e_j| = \sum_{i,j=1}^{N} |e_i\rangle A_{ij} \langle e_j| \tag{110}$$

$$\mathbb{1} = \sum_{i,j=1}^{N} |e_i\rangle \delta_{ij} \langle e_j| = \sum_{i=1}^{N} |e_i\rangle \langle e_i| \tag{111}$$

- the (111) is also called **resolution of identity.**

- this can be inserted in any expression without affecting its value, for example:

$$\langle\phi|\psi\rangle = \langle\phi| \sum_{i\in N} |e_i\rangle \langle e_i|\psi\rangle = \langle\phi| \sum_{i\in N} |e_i\rangle \langle e_i| \sum_{j\in N} |e_j\rangle \langle e_j|\psi\rangle = \langle\phi|e_i\rangle \langle e_i|e_j\rangle \langle e_j|\psi\rangle \tag{112}$$

Hence the (110) can be rewritten in the form of:

$$A = \mathbb{1}A\mathbb{1} = \left(\sum_i |e_i\rangle \langle e_i|\right) A \left(\sum_j |e_j\rangle \langle e_j|\right) = \sum_{ij} |e_i\rangle \langle e_i| A |e_j\rangle \langle e_j| = \sum_{ij}^{N} |e_i\rangle A_{ij} \langle e_j| \tag{113}$$

$A_{ij} = \langle e_i| A |e_j\rangle$ – just a number

In the case of quantum mechanics, state vectors are normilized i.e $|\phi|^2 = 1$

Hence, according to the (102):

$$\overrightarrow{\text{Proj}}_\phi \psi = |\phi\rangle \langle\phi|\psi\rangle \tag{114}$$

In general, product of ket $|\phi\rangle$ and bra $\langle\psi|$, is called an **outer product**.

An outer product, producing a matrix. Example:

$$|\phi\rangle \langle\psi| = |\phi\psi| = \begin{bmatrix} -\frac{-4}{5} \\ \\ \frac{3}{5} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} -\frac{4}{10} & -\frac{4\sqrt{3}}{10} \\ \\ \frac{3}{10} & \frac{3\sqrt{3}}{10} \end{bmatrix} \tag{115}$$

An outer product $|\phi\rangle \langle\phi|$ is also known as a **projection operator**. Projection operator it is a standard matrix of orthogonal projection.

According to the (65), (102) and Fig(3):

$$\overrightarrow{\text{Proj}}_{\vec{c}}\vec{a} = \vec{a} - \overrightarrow{\text{Proj}}_{\vec{b}}\vec{a} = \left(I - \frac{1}{|\vec{b}|^2} \left|\vec{b}\right\rangle \left\langle\vec{b}\right|\right) |\vec{a}\rangle \tag{116}$$

The (116) is called **vector rejection**.

In two dimensions, the scalar rejection is equivalent to the projection of $\vec{a}$ onto $\vec{b}^{\perp} = (-\vec{b}_y, \vec{b}_x)$, which is $\vec{b} = (\vec{b}_x, \vec{b}_y)$ rotated 90° to the left,

Hence

$$|\vec{c}| = |\vec{a}| \sin\theta = \frac{(\vec{a}, \vec{b}^{\perp})}{|\vec{b}|} = \frac{\vec{a}_y \vec{b}_x - \vec{a}_x \vec{b}_y}{|\vec{b}|} \tag{117}$$

Let's return to the orthogonal projection in multidimensional space:

Theorem (Parseval's Identity.) Let $S = \{\vec{v}_1, \cdots, \vec{v}_p\} \subseteq F^n$ be an orthogonal subset. If $\vec{y}$ is a linear combination of the vectors in S, that is,

$$\vec{y} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p \tag{118}$$

then

$$c_i = \frac{\vec{v}_i \cdot \vec{y}}{\vec{v}_i \cdot \vec{v}_i} \tag{119}$$

-is called **Fourier coefficients.**

Proof. Taking the dot product, we get

$$\vec{v}_i \cdot \vec{y} = \vec{v}_i(c_1\vec{v}_1 + c_2\vec{v}_{2+\cdots+c_p\vec{v}_p}) = c_i(\vec{v}_i, \vec{v}_i) \tag{120}$$

Since $\vec{v}_i \cdot \vec{v}_i \neq 0$, dividing both sides by $\vec{v}_i \cdot \vec{v}_i$ gives the formula.

Example of projection operator.

$$\hat{\mathrm{P}}\vec{v} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \tag{121}$$

Thus, the action of this matrix on an arbitrary vector is

$$P\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

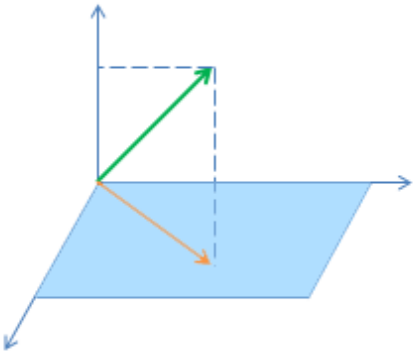


Figure 5. Orthogonal projection of a vector onto the plane

Example. Find the linear mapping from $\mathbb{R}^3$ to $\mathbb{R}^3$ that is a the orthogonal projection of $\mathbb{R}^3$ onto the plane $x_1 + x_2 + x_3 = 0$

To find the orthogonal projection of $\mathbb{R}^3$ onto the subspace $\vec{v}^{\perp}$, where $\vec{v} = [1, 1, 1]^T$, we find the following orthogonal projection

$$\text{Proj}_{\vec{v}}(\vec{y}) = \left(\frac{\vec{v}\cdot\vec{y}}{\vec{v}\cdot\vec{v}}\right)\vec{v} = \frac{y_1+y_2+y_3}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Then the orthogonal projection of $\vec{y}$ onto $\vec{v}^{\perp}$, according to (65) is given by

$$\text{Proj}_{\vec{v}}\vec{y} = \vec{y} - \text{Proj}_{v}(\vec{y}) = \left(I - \frac{1}{\vec{v}\cdot\vec{v}}\vec{v}\vec{v}^T\right)\vec{y} = \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

I.e. the projection of $\vec{y}$ **onto a plane** can be calculated by substracting the component $\text{Proj}_{v}(\vec{y})$ that is orthogonal to the plane from $\vec{y}$.

Let W be a subspace of V, and let $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k$ be an orthogonal basis of W. We want to decompose an arbitrary vector $\vec{v} \in V$ into the form

$$\vec{y} = \vec{w} + \vec{z}$$

with $\vec{w} \in W$ and $\vec{z} \in W^{\perp}$. Then there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_k$ such that

$$\hat{\vec{y}} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k$$

Since $\vec{z} \perp \vec{v}_1, \vec{z} \perp \vec{v}_2, \cdots, \vec{z} \perp \vec{v}_k$, we have

$$(\vec{v}_i, \vec{y}) = (\vec{v}_i, \alpha_1\vec{v}_1 + \cdots + \alpha_k\vec{v}_k + \vec{z}) = \alpha_i(\vec{v}_i, \vec{v}_i)$$

Then

$$\alpha_i = \frac{\vec{v}_i \cdot \vec{y}}{\vec{v}_i \cdot \vec{v}_i},\ 1 \leq i \leq k.$$

We thus define

$$\text{Proj}_W(\vec{y}) = \frac{\vec{v}_1 \cdot \vec{y}}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{v}_2 \cdot \vec{y}}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 + \cdots + \frac{\vec{v}_k \cdot \vec{y}}{\vec{v}_k \cdot \vec{v}_k}\vec{v}_k,$$

called the **orthogonal projection of $\vec{v}$ along W**. The linear transformation is called the **orthogonal projection of V onto W**.

Theorem. Let V be an n-dimensional inner product space. Let W be a subspace with an orthogonal basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$. Then for any $\vec{y} \in V$

$$\text{Proj}_W(\vec{y}) = \frac{\vec{v}_1 \cdot \vec{y}}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{v}_2 \cdot \vec{y}}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 + \cdots + \frac{\vec{v}_k \cdot \vec{y}}{\vec{v}_k \cdot \vec{v}_1}\vec{v}_k$$

$$\text{Proj}_{W^\perp}(\vec{y}) = \vec{y} - \text{Proj}_W(\vec{y})$$

In particular, if $\mathcal{B}$, is an orthonormal basis of W, then

$$\text{Proj}_W(\vec{y}) = (\vec{v}_1 \cdot \vec{y})\vec{v}_1 + (\vec{v}_2 \cdot \vec{y})\vec{v}_2 + \cdots + (\vec{v}_k \cdot \vec{y})\vec{v}_k$$

Proposition Let W be a subspace of $\mathbb{R}^n$. Len $U = [\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_k]$ be an $n \times k$ matrix, whose columns form an orthonormal basis of W. Then the orhtogonal projection $\text{Proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\text{Proj}_W(\vec{y}) = UU^T\vec{y}$$

Proof. For any $\vec{y} \in \mathbb{R}^n$, we have

$$\text{Proj}_W(\vec{y}) = (\vec{v}_1 \cdot \vec{y})\vec{v}_1 + (\vec{v}_2 \cdot \vec{y})\vec{v}_2 + \cdots + (\vec{v}_k \cdot \vec{y})\vec{v}_k$$

Note that

$$U^T\vec{y} = \begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vdots \\ \vec{u}_k^T \end{bmatrix} \vec{y} = \begin{bmatrix} \vec{u}_1^T\vec{y} \\ \vec{u}_2^T\vec{y} \\ \vdots \\ \vec{u}_k^T\vec{y} \end{bmatrix} = \begin{bmatrix} \vec{u}_1 \cdot \vec{y} \\ \vec{u}_2 \cdot \vec{y} \\ \vdots \\ \vec{u}_k \cdot \vec{y} \end{bmatrix}$$

Then

$$UU^T\vec{y} = [\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_k] \begin{bmatrix} \vec{u_1} \cdot \vec{y} \\ \vec{u_2} \cdot \vec{y} \\ \vdots \\ \vec{u_k} \cdot \vec{y} \end{bmatrix} = \text{Proj}_W(\vec{y})$$

Example. Find the orthogonal projection

$$\text{Proj}_W : \mathbb{R}^3 \to \mathbb{R}^3$$

where W is the plane $x_1 + x_2 + x_3 = 0$
By inspection, the following two vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \text{ and } \vec{v}_2 \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

form an orthogonal basis of W. Then

$$\text{Proj}_W(\vec{y}) = \frac{\vec{v}_1 \cdot \vec{y}}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{v}_2 \cdot \vec{y}}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 =$$

$$\frac{y_1 - y_2}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{y_1 + y_2 - 2y_3}{6} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} =$$

$$\frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Example. Find the standard matrix of the orthogonal projection

$$\text{Proj}_W : \mathbb{R}^3 \leftarrow \mathbb{R}^3$$

where

$$W = \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\}$$

The following two vectors

$$\vec{u_1} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \vec{u_2} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}$$

form of an orthonormal basis of W. Then the standard matrix of Proj_W is the product

$$\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}$$

which results the matrix

$$\begin{bmatrix} 5/6 & -1/6 & 1/3 \\ -1/6 & 5/6 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

Alternatively the matrix can be found by computing the orthogonal projection:

$$\text{Proj}_W(\vec{y}) = \frac{y_1 + y_2 + y_3}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{y_1 - y_2}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} =$$

$$= \frac{1}{6} \begin{bmatrix} 5y_1 - y_2 + 2y_3 \\ -y_1 - 5y_2 + 2y_3 \\ 2y_1 - 2y_2 + 2y_3 \end{bmatrix} = \begin{bmatrix} 5/6 & -1/6 & 1/3 \\ -1/6 & 5/6 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Example.

Let's consider the state

$$|\psi\rangle = \alpha |\phi_1\rangle + \beta |\phi_2\rangle, \quad \text{where } |\phi_1\rangle, |\phi_2\rangle \text{ forms an orthonormal basis}$$

- we want to decompose this state into a set of basis vectors. There is some experiment that define that basis states (SternGerlach experiment, spectroscopy experiment, etc.)

Let's get the α coefficient. Let's muliply both sides of $|\psi\rangle$ by $\langle\phi_1|$

$$\langle\phi_1|\psi\rangle = \alpha \langle\phi_1|\phi_1\rangle + \beta \langle\phi_1|\phi_2\rangle$$

-then

$$\begin{aligned}\langle\phi_1|\phi_1\rangle = 1 \quad &\text{due to orthonormality}\\ \langle\phi_1|\phi_2\rangle = 0 \quad &\text{due to orthogonality}\end{aligned}$$

- then

$$\alpha = \langle\phi_1|\psi\rangle$$

Hence, we can rewrite the $|\psi\rangle$ state without α and β coefficients:

$$\begin{aligned}&|\psi\rangle = \langle\phi_1|\psi\rangle\,|\phi_1\rangle + \langle\phi_2|\psi\rangle\,|\phi_2\rangle = \\ &|\phi_1\rangle\,\langle\phi_1|\psi\rangle + |\phi_2\rangle\,\langle\phi_2|\psi\rangle = \\ &[|\phi_1\rangle\,\langle\phi_1| + |\phi_2\rangle\,\langle\phi_2|]\,|\psi\rangle\end{aligned}$$

- i.e, the $[|\phi_1\rangle\,\langle\phi_1| + |\phi_2\rangle\,\langle\phi_2|] = \mathbb{1}$, it is an idendity matrix.

And in general:

$$\sum_i |\phi_i\rangle\,\langle\phi_i| = \mathbb{1} \tag{122}$$

and according to the (114), the $|\phi_i\rangle\,\langle\phi_i|$ it is a projection operator. It projects the $|\psi\rangle$ state, onto the $|\phi\rangle$ state.

The (122) is called **Completeness Relation**.

Example.

Let's go back to a two level system - "spin up" and "spin down"

$$|\psi\rangle = \frac{3}{5}\,|\uparrow\rangle + \frac{4}{5}\,|\downarrow\rangle$$

And let's introduce the notation for the projection operator: $\hat{P}_\uparrow \equiv |\uparrow\rangle\,\langle\uparrow|$

- according to the orthogonality and orthonormality (57):

$$\hat{P}_\uparrow\,|\psi\rangle = |\uparrow\rangle\,\langle\uparrow|\left(\frac{3}{5}\,|\uparrow\rangle + \frac{4}{5}\,|\downarrow\rangle\right) = |\uparrow\rangle\,\frac{3}{5} \tag{123}$$

- it is a projection of $|\psi\rangle$ onto the $|\uparrow\rangle$, but it is not normizied vector.

The general case of the (123) is:

$$\begin{aligned}&\hat{P}_n\,|\varphi_m\rangle = \delta_{nm}\,|\varphi_n\rangle\\ &\hat{P}_n\,|\psi\rangle = \hat{P}_n\sum_m |\varphi_m\rangle\,\langle\varphi_m|\psi\rangle = \sum_m \hat{P}\,|\varphi_m\rangle\,\langle\varphi_m|\psi\rangle\\ &= \sum_m \delta_{nm}\,|\varphi_m\rangle\,\langle\varphi_m|\psi\rangle = |\varphi_n\rangle\,\langle\varphi_n|\psi\rangle\end{aligned} \tag{124}$$

Important. The result of a projection operator, in general is not normilized state. There can be cases when you project a pure state onto a pure state, then the result can be a normilized state, but not always.

By the definition, a projection $\hat{P}$ is **idempotent** (i.e. $\hat{P}^2 = \hat{P}$)

$$\hat{P}_n^2\,|\varphi_m\rangle = \hat{P}_n\{\hat{P}_n\,|\varphi_m\rangle\} = \delta_{nm}\hat{P}_n\,|\varphi_m\rangle = \delta_{nm}^2\,|\varphi_m\rangle$$

A projection is orthogonal if and only if it is self-adjoint.

4.10. Projection matrix idempotence

Let's denote a projection as $\hat{P}$.

Definition. Let $\hat{P} : F^n \to F^n$ be a linear transformation. We say that $\hat{P}$ is a **projection** if $\hat{P} \circ \hat{P} = \hat{P}$

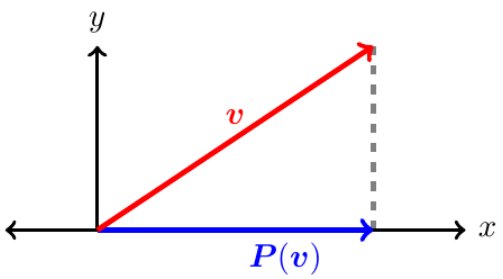
Let A be the standsrd matrix of $\hat{P}$. Then the property that $\hat{P} \circ \hat{P} = \hat{P}$ translate to mean that $A^2 = A$. Any matrix (necessarily a square) which satisfies this identity is called an **idempotent** matrix and is necessarily the standard matrox of a projection.

Let $\vec{x} \in F^n$. Then $\vec{y} = \hat{P}(\vec{x})$ is an arbitrary element of the range of $\hat{P}$. By definition, we have that

$$\hat{P}(\vec{y}) = \hat{P}(\hat{P}(\vec{x})) = \hat{P} \circ \hat{P}(\vec{x}) = \hat{P}(\vec{x}) = \vec{y}, \tag{125}$$

that is, a projection is exactly a linear transformation that fixes its image. Essentially this means that while some of the coordinates of $\vec{x}$ are unaltered the other coordinates are forgotten.

Any point already on the x-axis has the form $(x, 0)$ and is unaffected by the projection.



Example. For example, the matices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are easily to seen to be idempotent matrices and hence corresponds to projections.

The first matrix is projection in $\mathbb{R}^3$ onto the x-axis where the y-coordinate is forgotten and replaced with 0. Any point already on the x-axis has the form $(x, 0)$ and is unaffected by the projection.

The second matrix is a projection in $\mathbb{R}^3$ onto y-axis, where the x- and z-coordinates are discarded.

The third matrix is a projection in $\mathbb{R}^3$ onto the xy-plane, where the z-coordinate is the inly information forgotten.

Example.

Find a standard matrix of orthogonal projection of $\mathbb{R}^2$ onto the line spanned by $(2, -1)$, that is, the line $y = -1/2x$.

A norm of the $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ vector is $1/\sqrt{5}$. Therefore we have next matrix of orthogonal projection onto

the line $y = -1/2x$:

$$\vec{u} \otimes \vec{u} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \otimes \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \end{bmatrix} = \begin{bmatrix} 0.8944 \\ -0.4472 \end{bmatrix} \begin{bmatrix} 0.8944 & -0.4472 \end{bmatrix} =$$
$$= \begin{bmatrix} 0.8 & -0.4 \\ -0.4 & 0.2 \end{bmatrix}$$

Let's check that it is idempotent:

$$\begin{bmatrix} 0.8 & -0.4 \\ -0.4 & 0.2 \end{bmatrix} \begin{bmatrix} 0.8 & -0.4 \\ 0.4 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.4 \\ -0.4 & 0.2 \end{bmatrix}$$

Let's try to project a $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ vector:

$$\begin{bmatrix} 0.8 & -0.4 \\ -0.4 & 0.2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.4 \end{bmatrix}$$

If you draw a plot you will see it - it is an orthogonal projection.

Definition Let A be an $n \times n$ matrix. We say that A is **nilpotent** if $A^n = 0$, the zero matrix.

Theorem Let $\vec{u}, \vec{v} \in F^n$ such that $\vec{u}$ and $\vec{v}$ are orthogonal. Then $A = \vec{u} \otimes \vec{v}$ is a nilpotent matrix.

Theorem. All strictly triangular matrices are nilponent.

4.11. Gram-Schmidt process.

Let W be a subspace of an inner product space V. Let $\mathcal{B} = \{\vec{v_1}, \vec{v_2}, \cdots, \vec{v_k}\}$ be a basis of W, not necessarily orthogonal.

An orthogonal basis $\mathcal{B}'$ may be constructed from $\mathcal{B}$ as follows

$$\begin{aligned}
&\vec{w_1} = \vec{v_1}, && W_1 = \text{Span}\,\{\vec{w_1}\} \\
&\vec{w_2} = \vec{v_2} - \text{Proj}_{W_1}(\vec{v_2}), && W_2 = \text{Span}\,\{\vec{w_1}, \vec{w_2}\} \\
&\vec{w_3} = \vec{v_3} - \text{Proj}_{W_2}(\vec{v_3}), && W_3 = \text{Span}\,\{\vec{w_1}, \vec{w_2}, \vec{w_3}\} \\
&\vdots \\
&\vec{w}_{k-1} = \vec{v}_{k-1} - \text{Proj}_{W_{k-1}}(\vec{v_{k-1}}), && W_{k-1=\text{Span}}\,\{\vec{w_1}, \cdots \vec{w}_{k-1}\}, \\
&\vec{w_k} = \vec{v_k} - \text{Proj}_{W_{k-1}}(\vec{v_k})
\end{aligned} \tag{126}$$

More precisely

$$\begin{aligned}
\vec{w}_1 &= \vec{v}_1, \\
\vec{w}_2 &= \vec{v}_2 - \frac{\langle \vec{w}_1 | \vec{v}_2 \rangle}{\langle \vec{w}_1 | \vec{w}_1 \rangle} \vec{w}_1 \\
\vec{w}_3 &= \vec{v}_3 - \frac{\langle \vec{w}_1 | \vec{v}_3 \rangle}{\langle \vec{w}_1 | \vec{w}_1 \rangle} \vec{w}_1 - \frac{\langle \vec{w}_2 | \vec{v}_3 \rangle}{\langle \vec{w}_2 | \vec{w}_2 \rangle} \vec{w}_2 \\
&\vdots \\
\vec{w}_3 &= \vec{v}_3 - \frac{\langle \vec{w}_1 | \vec{v}_3 \rangle}{\langle \vec{w}_1 | \vec{w}_1 \rangle} \vec{w}_1 - \frac{\langle \vec{w}_2 | \vec{v}_3 \rangle}{\langle \vec{w}_2 | \vec{w}_2 \rangle} \vec{w}_2 - \cdots - \frac{\langle \vec{w}_{k-1} | \vec{v}_k \rangle}{\langle \vec{w}_{k-1} | \vec{w}_{k-1} \rangle} \vec{w}_{k-1}
\end{aligned} \tag{127}$$

The method of constructing the orthogonal vector $\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_k$ is known as the **Gram-Schmidt process**.

4.12. Some summary.

The ket $|\psi\rangle$ is a columnt vector.

The bra $\langle\phi|$ is a row vector (transpose and complex conjugate).

The bra-ket $\langle\phi|\psi\rangle$ is a comparison (inner product, dot product)

The ket-bra $|\psi\rangle\langle\phi|$ is a transformation (matrix, projection operator).

The ket $\langle\phi|\langle\psi|$ is meaningless, due to linear algebra rules.

5. Measurements.

5.1. Observables.

Observables are the things you measure. For example, we can make direct measurements of the coordinates of a particle; the energy, momentum, or angular momentum of a system; or the electric field at a point in space. Observables are also associated with a vector space, but they are not state-vectors. They are the things you measure. They are represented by linear operators.

John Wheeler liked to call such mathematical objects machines. He imagined a machine with two ports: an input port and an output port.

In the input port you insert a vector, such as $|\psi\rangle$. The gears turn and the machine delivers a result in the output port. This result is another vector, say $|\phi\rangle$

$$\hat{A}|\psi\rangle = |\phi\rangle \tag{128}$$

- the $\hat{A}$ - is called an and **operator.**

Let's write it in a component form (where $|j\rangle$ - it is a basis):

$$|\psi\rangle = \sum_j \alpha_j |j\rangle$$

$$|\phi\rangle = \sum_j \beta_j |j\rangle$$

$$\sum_j \hat{A}|j\rangle \alpha_j = \sum_j \beta_j |j\rangle$$

Now let's take the inner product of both sides with a particular basis vector $\langle k|$

$$\sum_j \langle k| \hat{A} |j\rangle \alpha_j = \sum_j \beta_j \langle k|j\rangle$$

$$\sum_j \langle k| \hat{A} |j\rangle \alpha_j = \sum_j \beta_j \delta_{kj}$$

- $\langle k|j\rangle$ is zero if j and k are not equal, and 1 if they are equal. That means that the sum on the right side collapses to a single term, β_k. We can abbreviate $\langle k| \hat{A} |j\rangle$ with the symbol m_{kj}. Each m_{kj} is just a complex number.

The quantities m_{kj} are called the **matrix elements** of $\hat{A}$.

$$\sum_j m_{kj}\alpha_j = \beta_k \tag{129}$$

In matrix form it becomes:

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$\begin{aligned} \beta_1 &= m_{11}\alpha_1 + m_{12}\alpha_2 + m_{13}\alpha_3 \\ \beta_2 &= m_{21}\alpha_1 + m_{22}\alpha_2 + m_{23}\alpha_3 \\ &\dots \end{aligned}$$

Next.

(1) Outcomes of the measurements must be real numbers

(2) In the measurements we are interested in basis states. For example for a Stern-Gerlach experiment we have "spin up" - $|\uparrow\rangle$ and "spin down" - $|\downarrow\rangle$ as a basis states. They are mutually exclusive. And any state in a SternGerlach experiment can be represented as a superposition of basis states.

(3) "spin up" - $|\uparrow\rangle$ and "spin down" - $|\downarrow\rangle$ in a SternGerlach experiment associated with all the possible values of observable S_z - z component of the spin.

An operator that satisfies all these requirements is called **Hermitian operator**.

Properties of a Hermitian Operator	**Properties of Observable S_z - z component of the spin.**
The eigenvalues of a Hermitean operator are all real.	Value of observable S_z measured to be real numbers ± 1 (spin-up/spin-down)
Eigenvectors belonging to different eigenvalues are orthogonal.	States $\|\uparrow\downarrow\rangle$ associated with different values of the observable are mutually exclusive.
The eigenstates form a complete set of basis states for the state space of the system. **Completeness Relation.**	The states $\|\uparrow\downarrow\rangle$ associated with all the possible values of observable S_z form a complete set of basis states for the state space of the system.

Therefore, we associate with the observable a Hermitian operator $\hat{A}$, such that has **eigenstates (eigenvectors)** $|\psi_i\rangle$ and associate **eigenvalues** α_i.

$$\hat{A}\,|\psi_i\rangle = \lambda_i\,|\psi_i\rangle \tag{130}$$

5.2. Linear operators in unitary space

5.2.1. Adjoint operators.

Riesz Representation Theorem.

Suppose V is finite-dimensional and ℓ is a linear functional on V. Then there is a unique vector $u \in V$ such that

$$\ell(v) = \langle v, u\rangle$$

for every $v \in V$.

Proof. We show there exists a vector $u \in V$ such that $\ell(v) = \langle v, u\rangle$ for every $v \in V$. Let $e_1, \cdots, e_n$ be an orthonormal basis of V. Then

$$\begin{aligned}\ell(v) &= \ell(x_1e_1 + \cdots x_ne_n)\\ &= \ell\left((e_1, v)e_1 + \cdots + (e_1, v)e_1\right)\\ &= (e_1, v)\ell(e_1) + \cdots + (e_n, v)\ell(e_n)\\ &= (v, \ell(e_1)e_1 + \cdots + \ell(e_n)e_n)\end{aligned}$$

- for every $v \in V$. Thus setting

$$u = \ell(e_1)e_1 + \cdots + \ell(e_n)e_n$$

- we have

$$\ell(v) = (v, u)$$

Also, only one vector $u \in V$ has the desired behavior.

Adjoint operators mimic the behaviour of the transpose matrix on real Euclideaan space. Recall that the transpose A^T of a real $m \times n$ matrix satisfies

$$\langle Ax, y\rangle = \langle x, A^T y\rangle$$

- for all $x \in \mathbb{R}^n$ and $y \in \mathbb{m}^n$, where $\langle\cdot,\cdot\rangle$ is the Euclidean inner product.

If T is a *bounded linear operator* from a Hilbert space H_1 into a Hilbert space H_2, i.e $T : H_1 \to H_2$, then for fixed $y \in H_2$ the linear functional ℓ defined on H_1 by

$$\ell(x) = \langle Ax, y\rangle$$

- is bounded and hence by the Reisz Representation Theorem

$$\langle Ax, y\rangle = \langle x, z\rangle$$

- for some $z \in H_1$. This z is uniquely determined by y, via A, and we denote it by $A^\dagger y$, that is

$$\langle Ax, y\rangle = \langle x, A^\dagger y\rangle$$

The $A^\dagger$ is called **adjoint** of an A operator.

We have that:

$$\langle Ax, y\rangle = \langle y, Ax\rangle^\dagger = \langle A^\dagger y, x\rangle^\dagger = \langle x, A^\dagger y\rangle$$

If V is finite-dimensional, the the adjoint operator $A^\dagger$ always exists.

Theorem. Given a matrix A with complex entries, its **adjoint matrix** is $A^\dagger = \bar{A^T}$

Given a orthonormal basis $e_1, \cdots, e_n$, and let

$$\begin{aligned} e_i A &= \alpha_{i1} e_1 + \alpha_{i2} e_2 + \cdots + \alpha_{in} e_n, \\ e_i A^\dagger &= \beta_{i1} e_1 + \beta_{i2} e_2 + \cdots + \beta_{in} e_n, \quad (i = 1, \cdots, n) \end{aligned}$$

- multiplying these equalities scalarly by e_j and use orthonormality of $e_1, \cdots, e_n$ we will get

$$\begin{aligned} \langle e_i A, e_j\rangle &= \alpha_{ij}, \\ \langle e_i A^\dagger, e_j\rangle &= \beta_{ij}, \end{aligned}$$

- hence:

$$\alpha_{ij} = \langle e_i A, e_j\rangle = \langle e_i, e_j A^\dagger\rangle = \langle e_j A^\dagger, e_i\rangle^* = \bar{\beta}_{ji}$$

Theorem If A is the matrix of a linear operator $\hat{A}$ relative to an orthonormal basis β, then the matrix of $\hat{A}^\dagger$ relative to the same basis is $A^\dagger$.

In quantum mechanics ajdoint operators defined as:

Ket vector $\hat{A}\,|\phi\rangle$ corresponding bra vector $\langle\phi|\,\hat{A}^\dagger$

$$\begin{aligned} &\hat{A}\,|\phi\rangle = \left|\hat{A}\phi\right\rangle \\ &\left\langle\hat{A}\phi\right| = \langle\phi|\,\hat{A}^\dagger \\ &\left\langle\hat{A}^\dagger\phi\middle|\psi\right\rangle = \left\langle\phi\middle|\hat{A}\psi\right\rangle \\ &(\hat{A}^\dagger)^\dagger = \hat{A}, \\ &(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \end{aligned}$$

For the matrix of operators in n-dimensional space:

$$\begin{aligned} \langle\psi|\,\hat{A}\,|\phi\rangle &= \sum_{m,n} \psi_m^* A_{mn} \phi_n, \\ (\langle\phi|\,\hat{A}^\dagger\,|\psi\rangle)^* &= \sum_{m,n} \phi_m A_{mn}^{\dagger *} \psi_n^* \end{aligned}$$

If we swap summation indices, we will find relationships between matrices of conjugated operators:

$$\begin{aligned} &\sum_{m,n} \psi_m^* A_{mn} \phi_n = \sum_{m,n} \psi_m^* (A_{nm}^\dagger)^* \phi_n, \\ &A_{mn}^* = A_{nm}^\dagger \end{aligned}$$

- i.e. given a matrix A with complex entries, its **adjoint matrix** is $A^\dagger = \bar{A^T}$

Hence:

$$\hat{A}^\dagger |\phi\rangle \equiv \left(\langle\phi| \hat{A}\right)^*$$
$$\langle\psi| \hat{A} |\phi\rangle = (\langle\phi| \hat{A}^\dagger |\psi\rangle)^* = (\langle\psi| (\hat{A}^\dagger)^\dagger |\phi\rangle)^{**} = \langle\psi| \hat{A}^{\dagger\dagger} |\phi\rangle ,$$
$$\langle\psi| \hat{A}^\dagger |\phi\rangle = (\langle\phi| \hat{A} |\psi\rangle)^*$$
$$\langle\psi| \alpha\hat{A} |\phi\rangle = (\alpha \langle\phi| \hat{A}^\dagger |\psi\rangle)^* = \alpha^* \langle\phi| \hat{A}^\dagger |\psi\rangle ,$$
$$(\alpha\hat{A})^\dagger = \alpha^* \hat{A}^\dagger$$

Any linear operator $\hat{A} : V \to V$ acts naturally on *both* bras and kets. Since a ket is a vector, its action is the usual one: $\hat{A} |\psi\rangle$ means a vector $A\psi \in V$.

Recall that bra $\langle\phi|$ is a map $V \to \mathbb{C}$, so the map $\langle\phi| \hat{A}$ is the map $V \to \mathbb{C}$ which first applies $\hat{A}$, then applies $\langle\phi|$. We get

$$\langle\phi| \hat{A} |\psi\rangle = \langle\phi| \hat{A}(|\psi\rangle) = (\langle\phi|)\hat{A} |\psi\rangle$$

- i.e the expression $\langle\phi| \hat{A} |\psi\rangle$ is consistent with both these interpretations (of couce it is, it's just composition of operators).

Example.

$$\text{Given} \quad A = \begin{bmatrix} 1 & -2i \\ 3 & i \end{bmatrix} \quad , \text{then} \quad A^\dagger = \begin{bmatrix} 1 & 3 \\ 2i & -i \end{bmatrix}$$

Adjoint operators as linear maps.

A **linear map** from V to W is a function $T : V \to W$ with the following properties:

additivity:

$T(u + v) = Tu + Tv$ for all $u, v \in V$;

homogeneity:

$T(\lambda v) = \lambda(Tv)$ for all $\lambda \in \mathbb{F}$ and all $v \in V$.

Some mathematicians use the term **linear transformation**, which means the same as linear map. Note that for linear maps often the notation Tv used as well as more standard functional notation $T(v)$.

Some times, when $V = W$ linear map is called **linear operator.** I.e. a linear operator is a linear endomorphism, that is, a linear map with the same domain and codomain.

Notation. $\mathcal{L}(V, W)$ - is a set of all linear maps from V to W.

A linear map is invertible if it is injective and surjective.

Notation. $\mathcal{L}(V)$ - is called an operator. It denotes the set of all operators on V. In other words, $\mathcal{L}(V) = \mathcal{L}(V, V)$.

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivelent

(a) T is invertible

(b) T is injective

(c) T is surjective

Definition. Suppose $T \in \mathcal{L}(V, W)$. The **adjoint** of T is the function $T^\dagger : W \to V$ such that

$\langle Tv, w\rangle = \langle v, T^\dagger w\rangle$

- for every $v \in V$ and every $w \in W$

To see why the definition above makes sence, suppose $T \in \mathcal{L}(V, W)$. Fix $w \in W$. Consider linear functional $\ell(v)$ on V that maps $v \in V$ to $\langle Tv, w\rangle$; this linear functional dependents on T and w. By Riesz Reprezentation Theorem, there exists a unique vector in V such that this linear functional is given by taking product with it. In other words, $T^\dagger$ is the unique vector in V such that $\langle Tv, w\rangle = \langle v, T^\dagger w\rangle$ for every $v \in V$.

Example.

Define T: $\mathbb{R}^3 \to \mathbb{R}^2$ by

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

- find a formulat form $T^\dagger$

Solution. Here $\hat{A}$ will be a function from $\mathbb{R}^2 \to \mathbb{R}^3$. To compute $T^\dagger$, fix a point $(y_1, y_2) \in \mathbb{R}^2$. Then for every $(x_1, x_2, x_3) \in \mathbb{R}^3$ we have

$$\begin{aligned}\langle (x_1, x_2, x_3), T^\dagger(y_1, y_2)\rangle &= \langle T(x_1, x_2, x_3), (y_1, y_2)\rangle \\ &= \langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\rangle \\ &= \langle x_2 y_1 + 3x_3 y_1 + 2x_1 y_2\rangle \\ &= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\rangle\end{aligned}$$

Thus

$$T^\dagger(y_1, y_2) = (2y_2, y_1, 3y_1)$$

Example

Fix $u \in V$ and $x \in W$. Define $T \in \mathcal{L}(V, W)$ by

$Tv = \langle v, u\rangle x$

for every $v \in V$. Find a formula for $T^\dagger$.

Solution. Fix $w \in W$. Then for every $v \in V$ we have

$$\begin{aligned}\langle v, T^\dagger w\rangle &= \langle Tv, w\rangle \\ &= \langle \langle v, u\rangle x, w\rangle \\ &= \langle v, u\rangle \langle x, w\rangle \\ &= \langle v, \langle w, x\rangle u\rangle\end{aligned}$$

Thus

$$T^\dagger w = \langle w, x\rangle u$$

5.2.2. Normal operators operators.

A linear operator $\hat{A}$ in an inner product space V is called normal if it commutes with its adjoint. That if the adjoint operator $\hat{A}^*$ exists and:

$$\hat{A} \circ \hat{A}^\dagger = \hat{A}^\dagger \circ \hat{A}$$

There are special classes of normal operators important for applications:

The operator $\hat{A}$ is **self-adjoint - (Hermitian)** if:

$$\hat{A}^\dagger = \hat{A} \quad \text{or equivalently:}$$
$$\langle \hat{A}x, y\rangle = \langle x, \hat{A}y\rangle$$

The operator $\hat{A}$ is **unitary** if:

$$\hat{A}^\dagger = \hat{A}^{-1}$$

The operator $\hat{A}$ is **anti-selfadjoint - (skew-Hermitian, anti-Hermitian)** if:

$$\hat{A}^\dagger = -\hat{A}$$

Definition. A square matrix A with real or complex entries is **normal** if $AA^\dagger = A^\dagger A$.

Example

$$\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$$

Show that matrix above is normal.

Solution

$$TT^\dagger = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 0 \\ 3 & 13 \end{bmatrix}$$

$$T^\dagger T = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 0 \\ 3 & 13 \end{bmatrix}$$

Because $TT^\dagger$ and $T^\dagger T$ have the same matrix, we see that $TT^\dagger = T^\dagger T$. Thus T is normal.

Other examples of normal matrices:

$$T = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \Rightarrow T^\dagger T = TT^\dagger \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$T = \begin{bmatrix} 4i & -1+i \\ 1-i & 4i \end{bmatrix} \Rightarrow T^\dagger T = TT^\dagger \begin{bmatrix} 18 & 8i \\ -8i & 18 \end{bmatrix}$$

$$T = \begin{bmatrix} -i & 2+3i \\ -2+3i & 0 \end{bmatrix} \Rightarrow T^\dagger T = TT^\dagger \begin{bmatrix} 14 & -3+2i \\ -3-2i & 13 \end{bmatrix}$$

$$T = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix} \Rightarrow T^\dagger T = TT^\dagger \begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix} \quad \text{This matrix is Hermitian also.}$$

$$T = \begin{bmatrix} 3 & 3-2i \\ 3+2i & 2 \end{bmatrix} \Rightarrow T^\dagger T = TT^\dagger \begin{bmatrix} 9 & 13 \\ 13 & 9 \end{bmatrix} \quad \text{This matrix is Hermitian also.}$$

Properties of normal operators.

Proposition. An operator $T \in \mathcal{L}(V)$ is normal if and only if

$$\|Tv\| = \|T^\dagger v\| \quad \text{for all } v \in V$$

Proof

$$\begin{aligned} & T^\dagger T - TT^\dagger = 0 \\ \iff & \langle (T^\dagger T - TT^\dagger)v, v\rangle = 0 \\ \iff & \langle T^\dagger Tv, v\rangle = \langle TT^\dagger v, v\rangle \\ \iff & \langle Tv, Tv\rangle = \langle T^\dagger v, T^\dagger v\rangle \\ \iff & \|Tv\|^2 = \|T^\dagger v\|^2 \end{aligned}$$

Proposition. For T normal, T and T^* have the same eigenvectors. Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is an eigenvector of T with eigenvalue λ. Then v is also eigenvector of T^* with eigenvalue $\bar{\lambda}$.

Proof.

Because T is normal, so is $T - \lambda I$. We have

$$0 = \|(T-\lambda I)v\| = \|(T-\lambda I)^\dagger v\| = \|(T^\dagger - \bar{\lambda} I)v\|$$

Hence v is an eigenvector of $T^\dagger$ with eigenvalue $\bar{\lambda}$, as desired.

I.e if T is normal, then

$$Tv = \lambda v \quad \Longrightarrow \quad T^\dagger x = \bar{\lambda} v$$

Notice, that if $T^\dagger = T$ then $\bar{\lambda} = \lambda$, hence the λ is a real number.

Proposition. If v_1 and v_2 are eigenvectors of T belonging to distinct eigenvalues λ_1 and λ_2 then $\langle v_1, v_2\rangle = 0$.

Proof.

We have $Tv_1 = \lambda_1 v_1$ and $T^\dagger v_2 = \bar{\lambda_2} v_2$. Then

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle Tv_1, v_2 \rangle =$$
$$\langle v_1, T^\dagger v_2 \rangle = \langle v_1, \bar{\lambda_2} v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

It follows that $(\lambda_1 - \lambda_2)\langle v_1, v_2 \rangle = 0$. Since $\lambda_1 \neq \lambda_2$, we obtain $\langle v_1, v_2 \rangle = 0$.

5.2.3. Self-adjoint (Hermitian) operators.

The operator $\hat{A}$ is **self-adjoint - (Hermitian)** if:

$$\hat{A}^\dagger = \hat{A} \quad \text{or equivalently:}$$
$$\langle \hat{A}x, y \rangle = \langle x, \hat{A}y \rangle$$

Hermitian operator is normal operator:

$$\hat{A} \circ \hat{A}^\dagger = \hat{A}^\dagger \circ \hat{A}$$

An Hermitian operator is the physicist's version of an object that mathematicians call a self-adjoint operator.

Hermitian Matrix: A square matrix is said to be Hermitian if it is equal to its conjugate transpose matrix.

How to find whether a matrix is Hermitian or not?

- Find the conjugate matrix of the given matrix by replacing every element with its conjugate.

- Find the transpose of the resultant matrix.

- If the original matrix is equal to its conjugate transpose matrix, then the given matrix is Hermitian.

Examples:

$$\begin{bmatrix} 3 & 3-2i \\ 3+2i & 2 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 2+i & 5-4i \\ 2-i & 4 & 6i \\ 5+4i & -6i & 2 \end{bmatrix}$$

Hermitian Matrix Formula

From the above two matrices, it is clear that the diagonal elements of a Hermitian matrix are **always real**. Also, the elements in the position (i, j) is the complex conjugate of the element in the position (j, i).

Hence, a 2×2 Hermitian matrix is on the form

$$\begin{bmatrix} x & y+zi \\ y-zi & w \end{bmatrix}$$

- where x, y, z, w are real numbers.

Similar, we can construct a 3×3 Hermitian matrix using the formula

$$\begin{bmatrix} a & b+ci & c+di \\ b-ci & e & g+hi \\ c-di & g-hi & k \end{bmatrix}$$

Properties of Hermitian Matix

- Hermitian matrices have real eigenvalues:

$A = A^\dagger = \bar{A}^T$
$Ax = \lambda x = A^\dagger x$
$\bar{x}^T Ax = \lambda \|x\|$
And if we take the complex transpose in both sides we get
$\bar{x}^T A^\dagger x = \bar{\lambda}\|x\|$
And since$A = A^\dagger$, we get that
$\lambda\|x\| = \bar{\lambda}\|x\|$
and so $\lambda \in \mathbb{R}$

- Hermitian matrices have real eigenvalues (2):

Let(λ, x)be an eigenvalue-eigenvector pair of$A$
$\bar{x}^T x = \sum_{i=1}^{n} |x_i|^2 \geq 0$
Since x is an eigenvector, x_i cannot be 0 for all i. Therefore, $\bar{x}^T x > 0$
$\bar{x}^T Ax = \bar{x}^T(\lambda x) = \lambda \bar{x}^T x$
$\bar{x}^T Ax = \bar{x}^T A^\dagger x = (Ax)^\dagger x = (\lambda x)^\dagger x = \bar{\lambda}\bar{x}^T x$
$\lambda \bar{x}^T x = \bar{\lambda}\bar{x}^T x \Longrightarrow (\lambda - \bar{\lambda})\bar{x}^T x = 0 \Longrightarrow (\lambda - \bar{\lambda}) = 0 \Longrightarrow \lambda = \bar{\lambda}$
Since $\lambda = \bar{\lambda}, \quad \lambda$ is real.

- Elements of the principal diagonal of a hermitian matrix are all real numbers.
- The non-diagonal elements of a hermitian matrix are complex numbers.
- **Every hermitian matrix is a normal matrix.**
- The *sum* of any two hermitian matrices is hermitian.
- The *inverse* of a hermitian matrix is a hermitian.
- The *product* of two hermitian matrices is hermitian.
- The *determinant* of a hermitian matrix is real.

Every Hermitian matrix can be written as $B + iC$, where B is real symmetric and C is real skew symmetric.

Proof. We have that $A^\dagger = A$ - A is Hermitian.

$$A = \frac{1}{2}(A + \bar{A}) + i\frac{1}{2i}(A - \bar{A})$$
$$A = B + iC$$

where

$$B = \frac{1}{2}(A + \bar{A}),\ C = \frac{1}{2i}(A - \bar{A})$$

- where B and C are real 'numbers' - matrices

Hence

$$\bar{A} = B - iC$$

Hence

$$A + \bar{A} = 2B$$
$$B = \frac{1}{2}(A + \bar{A})$$
$$A - \bar{A} = 2iC$$
$$C = \frac{1}{2i}(A - \bar{A})$$
$$B^T = \frac{1}{2}(A + \bar{A})^T$$
$$= \frac{1}{2}(A^T + A^\dagger) = \frac{1}{2}(A^T + A), \text{ sice A is Hermitian}$$

Let's compare B with B^T:

$$B = \frac{1}{2}(A + \bar{A}), \quad B^T = \frac{1}{2}(A + A^T)$$

And since we have:

$$A^\dagger = A$$
$$(A^\dagger)^T = A^T$$
$$((\bar{A})^T)^T = A^T$$
$$\bar{A} = A^T$$
$$B^T = \frac{1}{2}(\bar{A} + A)$$
$$B^T = B \implies B \textbf{ is symmetric.}$$

Similarly

$$
\begin{aligned}
C &= \frac{}{2i}(A - \bar{A}) \\
C^T &= \frac{1}{2i}(A - \bar{A})^T \\
&= \frac{1}{2i}(A^T - (\bar{A})^T) \\
&= \frac{1}{2i}(\bar{A} - A^\dagger) \\
&= \frac{1}{2i}(\bar{A} - A) = -\frac{1}{2i}(A - \bar{A}) \\
C^T &= -C \quad \Longrightarrow C \textbf{ is skew-symmetric.}
\end{aligned}
$$

Example.

$$
A = \begin{bmatrix} 2 & 1+i & -i \\ 1-i & 0 & -3-i \\ i & -3+i & -1 \end{bmatrix}
$$

$$
\bar{A} = \begin{bmatrix} 2 & 1-i & i \\ 1+i & 0 & -3+i \\ -i & -3-i & -1 \end{bmatrix}
$$

$$
A + \bar{A} = \begin{bmatrix} 2 & 1+i & -i \\ 1-i & 0 & -3-i \\ i & -3+i & -1 \end{bmatrix} + \begin{bmatrix} 2 & 1-i & i \\ 1+i & 0 & -3+i \\ -i & -3-i & -1 \end{bmatrix}
$$

$$
= \begin{bmatrix} 4 & 2 & 0 \\ 2 & 0 & -6 \\ 0 & -6 & -2 \end{bmatrix}
$$

$B = \frac{1}{2}(A + \bar{A}) \quad \Longrightarrow$ it is a symmetric matrix.

Next

$$A - \bar{A} = \begin{bmatrix} 0 & 2i & -2i \\ -2i & 0 & -2i \\ 2i & 2i & 0 \end{bmatrix}$$

$$\frac{1}{2i}(A - \bar{A}) = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$C = \frac{1}{2i}(A - \bar{A})$$

Symmetric Matrix: A matrix is said to be a symmetric matrix if the transpose of a matrix is equal to the given matrix: $A^T = A$

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

Skew-Symmetric Matrix: A matrix is said to be a skew-symmetric matrix if the transpose of a matrix is equal to the negative given matrix: $A^T = -A$

$$\begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}$$

Any square matrix can be written as sum of symmertic and skew symmetric matrix.

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

$$B = \left(\frac{A + A^T}{2}\right), \quad C = \left(\frac{A - A^T}{2}\right)$$

Let's show that $B = B^T$ - symmetric matrix:

$$B^T = \left(\frac{A + A^T}{2}\right)^T = \left(\frac{A^T + (A^T)^T}{2}\right) = \left(\frac{A^T + A}{2}\right) = B$$

Let's show that $-C = C^T$ - skew-symmetric matrix:

$$C^T = \left(\frac{A - A^T}{2}\right)^T = \left(\frac{A - (A^T)^T}{2}\right) = -\left(\frac{A - A^T}{2}\right)^T = -C$$

Example.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$B = \frac{1}{2}\left[\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right)\right] + \left(\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\right) = \frac{1}{2}\left(\begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}\right)$$

$$C = \frac{1}{2}\left[\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right)\right] - \left(\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\right) = \frac{1}{2}\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right)$$

$$A = B + C$$

General square matrix.

If T is not normal operator, but the eigenvalues are distinct, there are N linearly independent eigenvectors but they are not, in general, orthogonal. The eigenvectors of T and it's Hermitian conjugate, $T^\dagger$, are not the same. Howewer, if the eigenvalues are distinct, the eigenvectors of $T^\dagger$ are orthogonal to those of T. Hence the eigenvectors of $T^\dagger$ are *reciprocal* of those of T

A matrix whose eigenvectors are linearly dependent is *defective*.

Simultaneous eigenvectors

Two normal matrices have the same eigenvectors if and only if they commute.

When extended to Hermitian operators, this is related to the requirement in quantum mechanics that for two quantities to be simultaneously observable the operators must commute.

5.2.4. Unitary operators.

The operator $\hat{U}$ is **unitary** if:

$$\hat{U}^{-1} = \hat{U}^\dagger$$

Let's recollect **inverse** operator definition:

$$\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \mathbb{1}$$

Then:

$$\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \mathbb{1}$$

Unitary operator is normal operator:

$$\hat{A} \circ \hat{A}^\dagger = \hat{A}^\dagger \circ \hat{A}$$

Lemma. Let U be the matrix of change of basis: $\vec{v} = \vec{u}U$, where $\vec{u}$ is orthonormal. Then $\vec{v}$ is orthonormal if and only if U is a unitary matrix.

I.e unitary matrix change orthonormal basis to orthonormal one. I.e it does not changes the lenghs of a vectors, and does not change angles between them (is it just a roration?) It means, that the inner product have to remain unchanged:

$$\langle Ux, Uy\rangle = \langle x, y\rangle$$

Change of basis:

$$Uu_i = v_i$$

Let $x = a_1u_1 + \cdots a_nu_n$ and $y = b_1u_1 + \cdots + b_nu_n$.

Since $\{u_i\}$ is an orthonormal basis, $\langle u_i, u_i\rangle = \delta_{ij}$ the Kronecker delta, Thus, by expanding out the inner product $\langle x, y\rangle$, we see that:

$$\langle x, y\rangle = \sum_j \bar{a}_j b_j$$

Now compute $\langle Ux, Uy\rangle$:

$$\begin{aligned}
&\langle Ux, Uy\rangle = \langle U(a_1u_1 + \cdots + a_nu_n), U(b_1u_1 + \cdots b_nu_n)\rangle \\
&= \langle a_1Uu_1 + \cdots + a_nUu_n, b_1Uu_1 + \cdots b_nUu_n\rangle \\
&= \langle a_1v_1 + \cdots + a_nv_n, b_1v_1 + \cdots + b_nv_n\rangle \\
&= \sum_i \sum_j \bar{a_i} b_j \langle v_i, v_j\rangle
\end{aligned}$$

Since $\{v_i\}$ is orthonormal, $\langle v_i, v_j\rangle = \delta_{ij}$, and so the above is exactly $\langle x, y\rangle$.

Structurally, unitary matrices are rotations and reflections.

Properties of unitary operators.

(prop 1)

A matrix A is unitary if and only if its columns form an orthonormal set.
A matrix A is unitary if and only if its rows form an orthonormal set.

Proof (1)

Recall that a set of vectors $x_1, \cdots, x_k \quad \in \mathbb{C}_n$ is called **orthogonal** if $x_j^* x_m = \langle x_m, x_j\rangle = 0$ for $1 \leq j \neq m \leq k$.

The set is called **orthonormal** if

$$x_j^* x_m = \delta_{mj} = \begin{cases} 1 & j = m \\ 0 & j \neq m \end{cases}$$

By the definitioin, we have that:

$$A^*A = \mathbb{1}$$

- note that (j, k)-th entry of the identity matrix is

$$\mathbb{1}_{ik} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

- the matrix product: the (j,k)-th enty of the product A^*A is the product between the j-th row of A^* (denoted by $A^*_{j\bullet}$) and teh k-th column of A (denoted by $A_{\bullet k}$):

$$(A^*A)_{jk} = (A^*)_{j\bullet}A_{\bullet k}$$

- the conjugate transpose: the j-th row of A^* is equal to the conjgate transpose of the j-th column of A. Therefore, we have that

$$(A^*A)_{jk} = A^*_{j\bullet}A_{\bullet k} = \langle A_{\bullet k}, A_{\bullet j}\rangle$$

Suppose thet the columns of A form an orthonormal set. Then,

$$\langle A_{\bullet k}, A_{\bullet j}\rangle = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

- which implies

$$(A^*A)_{jk} = \langle A_{\bullet k}, A_{\bullet j}\rangle = \mathbb{1}_{jk}$$

AS a consequence, the columns of A are orthonormal.

Not only the columns but also the rows of a unitary matrix are orthonormal.

The rows of A are the columns of A^T, which is unitary.
1) if and only if it has orthonormal columns;
2) if and only if A is a unitary.

Proof (2)

Let's say we've got a unitary transformation $\hat{U}$ and an orthonormal basis $\{e_i\}_{i=1}^n$.

We can write:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

- each column is a vector. In particular, it's the result of transforming a bsis vector e_i by $\hat{U}$

Recall basis bransformation equations (32):

$$\left.\begin{aligned} \boldsymbol{e_1'} &= a_{11}\boldsymbol{e_1} + a_{21}\boldsymbol{e_2} + a_{31}\boldsymbol{e_3} \\ \boldsymbol{e_2'} &= a_{12}\boldsymbol{e_1} + a_{22}\boldsymbol{e_2} + a_{32}\boldsymbol{e_3} \\ \boldsymbol{e_3'} &= a_{13}\boldsymbol{e_1} + a_{23}\boldsymbol{e_2} + a_{33}\boldsymbol{e_3} \end{aligned}\right\}$$

$$\boldsymbol{\hat{U}(e_i)} = a_{1i}e_1 + a_{2i}e_1 + a_{3i}e_n \quad \text{- vector}$$

- it represents vectors, with coefficients a_{1i}, a_{2i}, a_{3i} which corresponds to columns of $\hat{U}$.

Let's take their inner product, taking into account that $\hat{U}$ preserves norms:

$$\langle \hat{U}(e_i), \hat{U}(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{i,j}$$

- that is the collection of columns of the matrix of $\hat{U}$ form another orthonormal basis.

On the other hand, let's consider some other orthonormal bais $\{e_j\}_{j=1}^n$:

$$\hat{A}(e_j) = \boldsymbol{e_1'} = a_{11}\boldsymbol{e_1} + a_{21}\boldsymbol{e_2} + a_{31}\boldsymbol{e_3}$$

- the $\hat{A}$ transformation will be unitary:

$$\langle x, y \rangle = \langle x^i e_i, y^j e_j \rangle = \bar{x^i} y^j \langle e_i, e_j \rangle = \bar{x^i} y^j \delta_{ij}$$
$$\langle \hat{A}(x), \hat{A}(y) \rangle = \langle x^i \hat{A}(e_i), y^j \hat{A}(e_j) \rangle = \bar{x^i} y^j \langle e_i', e_j' \rangle = \bar{x^i} y^j \delta_{ij}$$

- this is another evidence that only preserving norm transormations changes orthonormal basis to orthonormal one.

To sum up: with respect to an orthonormal basis, the columns of a unitary matrix form another orthonormal basis. Conversely, writing any other orthonormal basis in terms the original basis and using these coefficiants as the columns of a matrix gives a unitary matrix.

The same hold true for orthogonal matrices, with reasoning all the way through. And both of these are parallel to the situation for general linear transformations: **the columns of an invertible matrix with respect to any basis form another basis, and conversely.**

(prop 2) Product of unitary operators is an unitary operator.

$$\left.\begin{aligned} \hat{U}^{-1} &= \hat{U}^\dagger \\ \hat{V}^{-1} &= \hat{V}^\dagger \end{aligned}\right\} \quad \hat{U}\hat{V}$$

Proof

$$(\hat{U}\hat{V})^\dagger(\hat{U}\hat{V}) = \hat{V}^\dagger\hat{U}^\dagger\hat{U}\hat{V} = \hat{V}^\dagger\hat{V} = \mathbb{1}$$
$$(\hat{U}\hat{V})(\hat{U}\hat{V})^\dagger = \mathbb{1}$$

- hence, $(\hat{U}\hat{V})$ - unitary.

(prop 3) The eigenvalues of a unitary matrix are complex numbers with magnitude equal to 1.

Proof

$\hat{U}\,|\lambda\rangle = \lambda\,|\lambda\rangle$
assuming that the eigenstates are normilized $\Rightarrow \langle\lambda|\lambda\rangle = 1$ then:
$\|\hat{U}\,|\lambda\rangle\,\|^2 = \|\lambda\,|\lambda\rangle\,\|^2 = \lambda^*\lambda\,\langle\lambda|\lambda\rangle = |\lambda|^2$
also, because $\hat{U}$ is unitary:
$\|\hat{U}\,|\lambda\rangle\,\|^2 = \langle\lambda|\,\hat{U}^\dagger\hat{U}\,|\lambda\rangle = \langle\lambda|\lambda\rangle = 1$

- hence $|\lambda|^2 = 1$ and

$$|\lambda| = 1 \Longrightarrow \lambda = e^{i\varphi_\lambda}, \quad \varphi_\lambda \in \mathbb{R}$$

Hence the eigenvalues of a unitary matrix all lie on the unit circle in the complex plane.

(proof 4) Eigen states of the unitary operator coresponding to different eigenvalues are orthogonal.

Proof

Let's show that:

$$\hat{U}\left|\lambda\right\rangle = \lambda\left|\lambda\right\rangle \quad \Longrightarrow \quad \left\langle\lambda\right|\hat{U} = \lambda\left\langle\lambda\right|$$

-we have that:

$$\begin{aligned}
&\left|\lambda\right\rangle = \mathbb{1}\left|\lambda\right\rangle = \hat{U}^{\dagger}\hat{U}\left|\lambda\right\rangle = \lambda\hat{U}^{\dagger}\left|\lambda\right\rangle = e^{i\varphi_{\lambda}}\hat{U}^{\dagger}\left|\lambda\right\rangle \\
&\hat{U}^{\dagger}\left|\lambda\right\rangle = e^{-i\varphi_{\lambda}}\left|\lambda\right\rangle = \lambda^{*}\left|\lambda\right\rangle \\
&\left\langle\lambda\right|\hat{U} = \left\langle\lambda\right|e^{i\varphi_{\lambda}} = \left\langle\lambda\right|\lambda
\end{aligned}$$

Now we are ready to show that the eigenstates are orthogonal

$$\begin{aligned}
&\hat{U}\left|\lambda\right\rangle = \lambda\left|\lambda\right\rangle, \quad \hat{U}\left|\mu\right\rangle = \mu\left|\mu\right\rangle \Rightarrow \left\langle\mu\right|\hat{U} = \mu\left\langle\mu\right| \\
&\left.\begin{aligned}
&\left\langle\mu\right|\hat{U}\left|\lambda\right\rangle = \lambda\left\langle\mu|\lambda\right\rangle \\
&\left\langle\mu\right|\hat{U}\left|\lambda\right\rangle = \mu\left\langle\mu|\lambda\right\rangle
\end{aligned}\right\} \quad 0 = (\lambda-\mu)\left\langle\mu|\lambda\right\rangle \\
&\lambda \neq \mu \quad \Longrightarrow \left\langle\mu|\lambda\right\rangle = 0
\end{aligned}$$

Unitary transformation.

A key reason why unitary operators are so important in quantum mechanics is that the unitary transformations conserve the scalar product.

Example of unitary transformation between the state $\left|\psi_1\right\rangle$ and $\left|\psi_1^{'}\right\rangle$:

$$\left|\psi_1^{'}\right\rangle = \hat{U}\left|\psi_1\right\rangle, \quad \text{and} \quad \left|\psi_2^{'}\right\rangle = \hat{U}\left|\psi_2\right\rangle$$

Let's proof the unitary transformation conserve a inner product

$$\left\langle\psi_1^{'}\middle|\psi_2^{'}\right\rangle = \left\langle\psi_1\right|\hat{U}^{\dagger}\hat{U}\left|\psi_2\right\rangle = \left\langle\psi_1\right|\mathbb{1}\left|\psi_2\right\rangle = \left\langle\psi_1|\psi_2\right\rangle$$

- Unitary transformation conserve the scalar product

- Unitary transformation conserve the norm (the norm is the scalar product of the ket with itself).

It means that applying a unitary transformation allows us to change the state without changing its normalization.

A well known example is the **time evolution** of a quantum state. Which is a process that conserves the norm and can be described by unitary operator - called the time evolution operator.

An orthogonal matrix is the real specialization of a unitary matrix, and thus always a normal matrix. Orthogonal matrices satisfy:

$$\langle Ax, y\rangle = \langle x, A^{-1}y\rangle$$

-i.e. for orthogonal matrices $A^{-1} = A^{T}$

So a real unitary matrix is the same as orthogonal.

Unitary transformations of operators.

$$\hat{A} \Longrightarrow \hat{A}'$$

First let's consider what does **matrix similarity** means:

$\|\phi\rangle = \hat{A}\|\psi\rangle$	$\|\phi\rangle = \hat{A}\|\psi\rangle$
$\{\|e_j\rangle\} \sim \{\|j\rangle\}$ - basis	$\{\|e_j'\rangle\} \sim \{\|j'\rangle\}$ - basis
$\|\psi\rangle = \alpha_1\|1\rangle + \cdots \alpha_n\|n\rangle$	$\|\psi\rangle = \alpha_1'\|1'\rangle + \cdots \alpha_n'\|n'\rangle$
$\|\psi\rangle = \sum_j \alpha_j\|j\rangle$	$\|\psi\rangle = \sum_j \alpha_j'\|j'\rangle$
$\|\phi\rangle = \beta_1\|1\rangle + \cdots \beta_n\|n\rangle$	$\|\phi\rangle = \beta_1'\|1'\rangle + \cdots \beta_n'\|n'\rangle$
$\|\phi\rangle = \sum_j \beta_j\|j\rangle$	$\|\phi\rangle = \sum_j \beta_j'\|j'\rangle$
$\sum_j \hat{A}\|j\rangle\alpha_j = \sum_j \beta_j\|j\rangle$	$\sum_j \hat{A}'\|j'\rangle\alpha_j' = \sum_j \beta_j'\|j'\rangle$
$\sum_j \langle k\|\hat{A}\|j\rangle\alpha_j = \sum_j \beta_j\langle k\|j\rangle = \sum_j \beta_j\delta_{kj}$	$\sum_j \langle k'\|\hat{A}'\|j'\rangle\alpha_j' = \sum_j \beta_j'\langle k'\|j'\rangle = \sum_j \beta_j'\delta_{kj}$
$\sum_{j=n}^{n} a_{kj}\alpha_j = \beta_k$	$\sum_{j=n}^{n} a_{kj}'\alpha_j' = \beta_k'$
$\beta = A\alpha$	$\beta' = A'\alpha'$

- here we used (129)

Let us write the coordinate transformation formulas in matrix form:

$$\alpha = T\alpha', \quad \beta = T\beta'$$

- then:

$$\beta' = T^{-1}\beta = T^{-1}A\alpha = T^{-1}AT\alpha' = A'\alpha'$$
$$A' = T^{-1}AT$$

Two matrices A and B related to each other by the relation:

$$B = T^{-1}AT$$

- are called **similar**.

Thus, two matrices corresponding to the same linear operator $\hat{A}$, with different bases, are similar to each other, and the matrix T connecting these matrices coincides with the coordinate transformation matrix during the transition from the first basis to another.

In other words, a linear operator corresponds to a whole class of matrices that are similar to each other: these matrices represent a given operator in various bases.

By studying the properties of a linear operator, we thereby study the properties of matrices that are simultaneously inherent in the entire class of similar matrices, i.e. We study the properties of matrices that remain unchanged (invariant) when moving from a given matrix to a matrix similar to it.

Let's consider two bases:

$$\{|j\rangle\} \longrightarrow \{|j'\rangle = \hat{U}\,|j\rangle\}$$

- $j \in 1 \cdots m$, $\hat{U}$ is an unitary operator.

We then say that the $\hat{A}'$ operator is a unitary transformation of $\hat{A}$ operator if matrix elements of $\hat{A}'$ in $\{|j'\rangle\}$ basis equals to matrix elements of $\hat{A}$ in $\{|j\rangle\}$ basis:

$$\langle j'|\,\hat{A}'\,|j'\rangle = \langle j|\,\hat{A}\,|j\rangle$$

-then:

$$\langle j'|\,\hat{A}'\,|j'\rangle = \langle j|\,\hat{U}^\dagger\hat{A}'\hat{U}\,|j\rangle$$

- it follows that:

$$\hat{A} = \hat{U}^\dagger\hat{A}'\hat{U}$$
$$\hat{U}\hat{A}\hat{U}^\dagger = \hat{U}\hat{U}^\dagger\hat{A}'\hat{U}\hat{U}^\dagger = \hat{A}'$$

- then we say that operator $\hat{A}'$ is a unitary transformation of operator $\hat{A}$

Properties of unitary transformation applied to operators:

(1) If $\hat{A}$ is a hermitian operator, then it's unitary transformation is also unitary operator:

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger$$
$$\hat{A}^\dagger = \hat{A} \Longrightarrow (\hat{A}')^\dagger = \hat{A}'$$

Proof:

$$(\hat{A}')^\dagger = (\hat{U}\hat{A}\hat{U}^\dagger)^\dagger = \hat{U}\hat{A}^\dagger\hat{U}^\dagger = (\hat{A}^\dagger)' = \hat{A}'$$

(2) n-th power of the unitary transformation of A

$$(\hat{A}')^n = (\hat{U}\hat{A}\hat{U}^\dagger)^n = \hat{U}\hat{A}\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\cdots\hat{U}\hat{A}\hat{U}^\dagger = \hat{U}\hat{A}^n\hat{U}^\dagger = (\hat{A}^n)'$$
$$[f(\hat{A})]' = f(\hat{A}')$$

(3) Eigenvalue equation under unitary transformation.

$$\hat{A}\,|\alpha\rangle = \alpha\,|\alpha\rangle$$
$$\hat{A}'\,|\alpha'\rangle = \hat{U}\hat{A}\hat{U}^\dagger\hat{U}\,|\alpha\rangle = \hat{U}\underbrace{\hat{A}\,|\alpha\rangle}_{\alpha|\alpha\rangle} = \alpha\underbrace{\hat{U}\,|\alpha\rangle}_{|\alpha'\rangle} = \alpha\,|\alpha'\rangle$$

- it means that if we have an eigenvalue equation for an opertor $\hat{A}$ then the unitary transformation means that **eigenvalues state the same.**

(4) Relation of unitary operators to hermitian operators.

Given some hermitian operator $\hat{A}^\dagger = \hat{A}$, and we define the new operator $\hat{T} = e^{i\hat{A}}$:

$$\hat{A}^\dagger = \hat{A} \quad \Longrightarrow \quad \hat{T} = e^{i\hat{A}}$$

- where $\hat{T}$ is a unitary operator.

To see this, consider first:

$$\hat{T}^\dagger = (e^{i\hat{A}})^\dagger = \left(\sum_{n=0}^{\infty} \frac{1}{n!}(i\hat{A})^n\right)^\dagger = \sum_{n=0}^{\infty} \frac{1}{n!}(-i\hat{A}^\dagger)^n = e^{-i\hat{A}^\dagger} = e^{-i\hat{A}}$$

- then:

$$\hat{T}^\dagger\hat{T} = e^{-i\hat{A}}e^{i\hat{A}} = \mathbb{1}$$
$$\hat{T}\hat{T}^\dagger = e^{i\hat{A}}e^{-i\hat{A}} = \mathbb{1}$$
$$\hat{T}^{-1} = \hat{T}^\dagger$$

- which means that the operator $\hat{T}$ **is unitary operator.**

(5) Infinitessimal unitary operators.

Consider unitary operator $\hat{U}(\epsilon)$ that depends on the small parameter $\epsilon \in \mathbb{R}$. We then say that the $\hat{U}(\epsilon)$ is an infinitizemal operator and all we mean is that $\hat{U}(\epsilon)$ tends to $\mathbb{1}$ as $\epsilon \to 0$:

$$\hat{U}(\epsilon), \epsilon \in \mathbb{R}, \quad \hat{U}(\epsilon) \to \mathbb{1} \text{ as } \epsilon \to 0$$

Let's expand $\hat{U}(\epsilon)$ in a power series wth the first term bein the identity matrix $\mathbb{1}$, the second $\epsilon\hat{G}$, where $\hat{G}$ some other operator, and so on:

$$\hat{U}(\epsilon) = \mathbb{1} + \epsilon\hat{G} + \cdots$$
$$\hat{U}^\dagger(\epsilon) = \mathbb{1} + \epsilon\hat{G}^\dagger + \cdots$$

-then

$$\hat{U}(\epsilon)\hat{U}^\dagger(\epsilon) = \mathbb{1} + \epsilon(\hat{G} + \hat{G}\dagger) + \cdots = \mathbb{1} \quad \text{because } \hat{U} \text{ is unitary}$$

This means that

$$\hat{G} + \hat{G}^\dagger = 0 \quad \Longrightarrow \quad \hat{G} = -\hat{G}^\dagger$$

- so the $\hat{G}$ is anti-hermitian. Let's define it as:

$$\hat{F} = i\hat{G}: \quad \hat{G} = -\hat{G}^\dagger \quad \Rightarrow \quad \hat{F}^\dagger = \hat{F}$$

- so the $\hat{F}$ is hermitian.

It means that we can always write $\hat{U}(\epsilon)$ in the form:

$$\hat{U}(\epsilon) = \mathbb{1} - i\epsilon\hat{F} + \cdots$$

- if $\hat{F}$ is hermitian, then $\hat{U}$ is unitary.

5.2.5. The Spectral Theorem.

Matrix diagonalization.

An $n \times n$ matrix A is diagonalizable if it is **similar** to diagonal matrix:

$$P\Lambda P^{-1} = A$$

- where P is some invertible matrix, and λ - diagonal matrix.

An operator $\hat{A}$ is called diagonalizable if the operator has a diagonal matrix with respect to some basis.

An $n \times n$ matrix A is **diagonalizable** if and only if A has n linearly independent eigenvectors, and each geometric multiplicity is equal to its algebraic multiplicity. A matrix P consists of eigenvectors of A, and matrix Λ consists of eigenvalues of A.

An $n \times n$ matrix A is **unitary(orthogonaly) diagonalizable** if A has n orthonormal eigenvectors, and each geometric multiplicity is equal to its algebraic multiplicity. In this case, there exists an Hermitian(symmetric) matrix P and diagonal martix D such that $A = PDP^{\dagger}$.

A matrix A is normal if and only if there exists a *diagonal matrix* Λ and unitary matrix U such that $A = U\Lambda U^{\dagger}$.

As we proof before , for normal matrices, if v_1 and v_2 are eigenvectors of A belonging to distinct eigenvalues λ_1 and λ_2 then $\langle v_1, v_2 \rangle = 0$. And because the matrix P consists of vectors v_n it is unitary as we showed before.

Therefore, normal matrices are always **unitary** diagonalizable.

Some other definitions:

(1) A linear operator $T : H \longrightarrow H$ is said to be diagonalizable if there exists a basis of H consists of eigenvectors of T

(2) A linear operator $T : H \longrightarrow H$ is said to be **unitary** diagonalizable if there is an **orthonormal** basis of H in which T is represented by a diagonal matrix:

There exists an orthonormal basis $\{e_1, \cdots, e_n\}$ of H and values $\lambda_1, \cdots, \lambda_2 \in \mathbb{C}$ such that:

$$T(x) = \sum_{i=1}^{n} \lambda_i \{e_i, x\} e_i$$

Spectral decomposition.

Let A be a hermitian(symmetric) matrix. Hence it is a normal matrix. Then it has an orthogonal diagonalization given as:

$$A = PDP^{\dagger} = \begin{bmatrix} \mathbf{u_1} & \cdots & \mathbf{u_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u_1^T} \\ \vdots \\ \boldsymbol{u_n^T} \end{bmatrix} = \begin{bmatrix} \lambda_1 \boldsymbol{u_1} & \cdots & \lambda_1 \boldsymbol{u_1} \end{bmatrix} \begin{bmatrix} \boldsymbol{u_1^T} \\ \vdots \\ \boldsymbol{u_n^T} \end{bmatrix} \tag{131}$$

$$= \lambda_1 \boldsymbol{u_1} \boldsymbol{u_1}^T + \cdots + \lambda_1 \boldsymbol{u_n} \boldsymbol{u_n}^T = \lambda_1 |\boldsymbol{u_1}\rangle \langle \boldsymbol{u_1}| + \cdots \lambda_n |\boldsymbol{u_n}\rangle \langle \boldsymbol{u_n}|$$

- where $\boldsymbol{u_n}$ is the eigenvectors and λ_n is the eigenvalues of A.

The last equation is called a **spectral decomposition** of A.

Now compare it with (110). You'll see that if the matrix $\mathbb{A}$ is diagonal, the (110) equation will be simplified to (131)

Let's write an eigenvalue equation for $\hat{A}$:

$$\hat{A} |u_j\rangle = \lambda_j |u_j\rangle$$

- then the operator $\hat{A}$ is decomposed as follows:

$$\hat{A} = \sum_i \lambda_i |u_i\rangle \langle u_i| \tag{132}$$

- the (132) is also called **spectral decomposition of the observable $\hat{A}$.** The eigenvectors of an observable constitute a set of basis states for the state pace of the associated quantum system.

5.2.6. Simultaneous Diagonalization

Theorem. Two diagonalizable operators A, B are simulataneously diagonalizable if and only if $[A, B] = 0$, where $[\cdot, \cdot]$ is the commutator.

Proof. Let there be a basis $\{|e_i\rangle\}$ fro which the diabonalizible operators A, B have

$$A|e_i\rangle = \alpha_i |e_i\rangle, \quad A|e_i\rangle = \beta_i |e_i\rangle$$

This is a basis in which A and B are simultaneously diagonal. In this case, we see that

$$AB|e_i\rangle = \alpha_i\beta_i |e_i\rangle = BA|e_i\rangle$$

- so $AB = BA$. This prooves the forward direction.

Now consider two diagonalizable operators A, B that commute, $AB = BA$, and let $\{|e_i\rangle\}$ be a basis of eigenvectors of A,

$$A|e_i\rangle = \alpha_i |e_i\rangle$$

- then we have:

$$A(B|e_i\rangle) = AB|e_i\rangle = BA|e_i\rangle = B\alpha_i |e_i\rangle = \alpha_i(B|e_i\rangle)$$

- so $B|\alpha_i\rangle$ is an eigenket of A with eigenvalue α_i. In general, this means that B is block diagonal in the basis $\{|e_i\rangle\}$, with each block corresponding to a single eigenvalue of A.

5.3. Stern and Gerlach experiment

The result of a measurement is generally statistically uncertain. However, for any given observable, there are particular states for which the result is absolutely certain.

For example, if the spin-measuring apparatus S is oriented along the z axis the state $|\uparrow\rangle$ always leads to the value $\sigma_z = +1$

The result of measurement is unambiguously λ_i if the system was in eigenstate $|\psi_i\rangle$ of the operator $\hat{A}$. Here is another way to say it: if the system is in the eigenstate $|\psi_i\rangle$, the result of a measurement is **guaranteed** to be λ_i

Unambiguously distinguishable states are represented by orthogonal vectors. Two states are physically distinct if there is a measurement that can tell them apart without ambiguity. For example, $|\uparrow\rangle$ and $|\downarrow\rangle$ can be distinguished by measuring σ_z

But suppose that you are told the spin is in one of the two states, $|\uparrow\rangle$ and $|\leftarrow\rangle$ - they are not orthogonal. There is nothing you can measure that will unambiguously thell you the spin's true state. Measuring S_z won't do it. If you get $S_z = +1$, it is possible that the initial state was $|\rightarrow\rangle$ since there is a 50 percent probability of getting this answer in the state $|\rightarrow\rangle$. For that reason, $|\uparrow\rangle$ and $|\downarrow\rangle$ are said to be physically distriguishable, but $|\uparrow\rangle$ and $|\rightarrow\rangle$ are not. One might say that the inner product of two states is a measure of the inability to distinguish them with certainty.

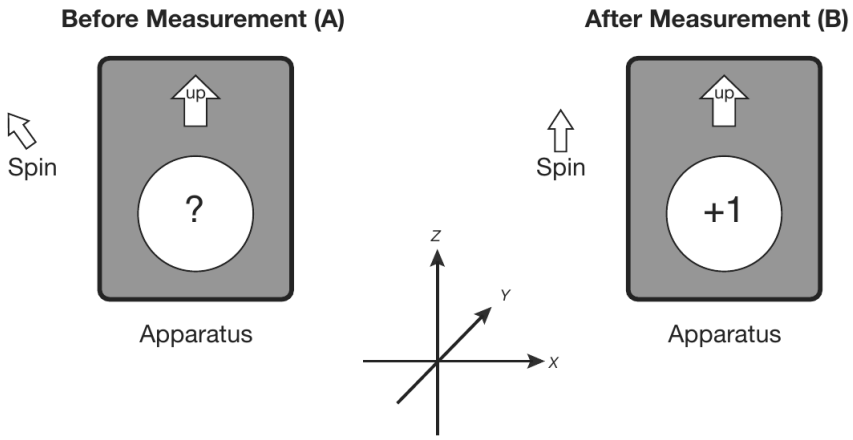


Figure 6. Stern and Gerlach experiment

Sometimes this inner product is called the overlap. This principle requires physically distinct states to be represented by orthogonal state-vectors, that is, vectors with no overlap.

We begin by pointing apparatus along the z axis. Initially we have no knowledge of whether observable $S_z = +1$ or $S_z = -1$. S_z - z component of the spin.

Before the apparatus interfacts with the spin, the window is blank (labaled with a question mark in our diagrams). After it measures observable S_z, the window shows a $+1$ of a -1. By looking at the apparatus, we determine the value of observable S_z.

Now that we've measured S_z, let's reset the apparatus to neutral and, without disturbing the spin, measure S_z again. We will get the same answer as we did the first time. It is because in the first experiment we prepared a system in one of it's eigenstate, and we already know that: if the system is in the eigenstate $|\psi_i\rangle$, the result of a measurement is guaranteed to be λ_i.

The result $S_z = +1$ will be followed by $S_z = +1$. Likewise for $S_z = -1$.

The same will be true for any number of repetitions. This is good because it allows us to confirm the result of an experiment.

We can also say this in the following way: The first interaction with the apparatus $\mathcal{S}$ **prepares** the system in one of the two states. Subsequent experiments **confirm** that state.

So far, there is no difference between classical and quantum physics.

Now lets do something new.

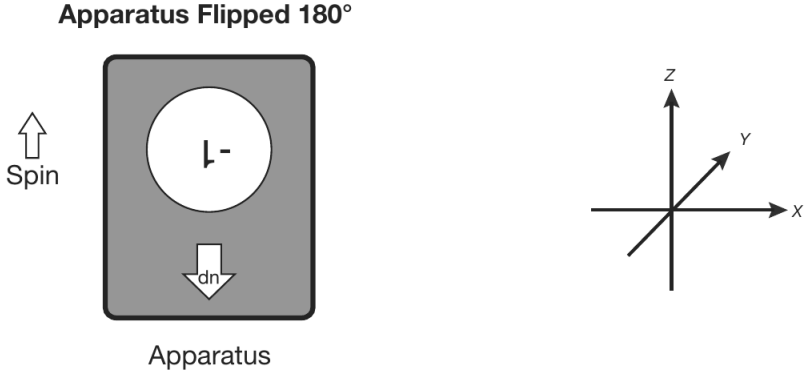


Figure 7. The apparatus is flipped without disturbing the previously measured spin. A new measurements results in $S_z = -1$.

After preparing the spin by measuring it with $\mathcal{S}$, we trun the apparatus upside down, and then measure the S_z again. What we find is that if we originally prepared $S_z = +1$, then upside down apparatus records $S_z = -1$. Similarly, if we originally prepared $S_z = -1$, the upside down apparatus records $S_z = +1$. In orther words, turning the apparatus over interchanges $S_z = +1$ and $S_z = -1$.

From these results, we might conclude that S_z is a degree of freedom that is associated with a sence of direction in space.

For example, if S_z were an oriented vector of some sort, then it would be natural to expect that turning the apparatus over would reverse the reading. A simple explanation is that the apparatus measures the component of the vector along an axis embedded in the apparatus. Is this explanation correct for all configurations?

If we are convinced that the spin is a vector, we would naturaly describe it by three componentes: S_z, S_x, S_y. When the apparatus is upright along the z axis, it is positioned to measure S_z.

Let's denote entire vector of the spin as S_σ

So far, there is still no difference between classical physics and quantum physics.

The difference only becomes apparent, when we rotate the apparatus through an arbitrary angle, say $\frac{\pi}{2}$ radians (90 degrees).

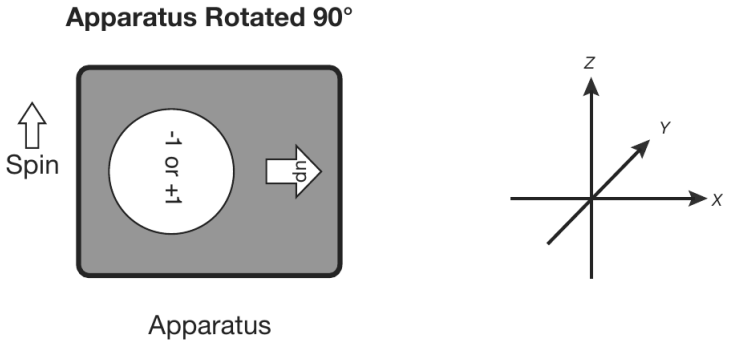


Figure 8. The apparatus rotated by 90°. An new measurement results in $S_z = -1$ with 50 percent probability.

The apparatus begins in the upright position (with the up-arrow along the z axis). A spin is prepared along the $S_z = +1$ axis. Next, rotate $\mathcal{S}$ so that the up-arrow points along the x axis. And then make a measurement of what is presumable the x component of the spin, S_x.

In fact S_z really represents the component of a vector along the up-arrow, one would expect to get zero. Why? Initially, we confirmed that S_z was directed along the z axis, suggesting that its component along x must be zero.

But we get a surprise when we measure S_x: Instead of giving $S_x = 0$, the apparatus gives either $S_x = +1$ or $S_x = -1$. $\mathcal{S}$ is very stubborn - no matter which way it is oriented, it refuses to give any answer other than $S_\sigma = \pm 1$. If the spin really is a vector, it is a very peculiar one indeed.

Nevertheless, we do find something interesting. Suppose we repeat the operation many times, each time following the same procedure, that is:

- Beginig with $\mathcal{S}$ along the z axis, prepare $S_z = +1$

- Rotate the apparatus so that it is oriented along the x axis.

- Measure S_x.

The repeated experiment spits out a random series of plus-ones and minus-ones. **Determinism has broken down, but in particular way.**

If we do many repetitions, we will find that the numbers of $S_x = +1$ events and $S_x = -1$ events

are statistically equal. In other words, the average value of observable S_σ is zero. Instead of the classical result - namely, that the component of S_σ along the axis x is zero - we find that the **average of these repeated measurements** is zero.

Now let's do the whole thing over again, but instead of rotating $\mathcal{S}$ to line on the x axis, rotate it to an arbitrary direction along the unit vector $\hat{n}$. Prepare a spin so that the apparatus reads $+1$. Then, without disturbing the spin, rotate the apparatus to the direction $\hat{n}$, as shwoen in fig(9). A new experiment on the same spin will give

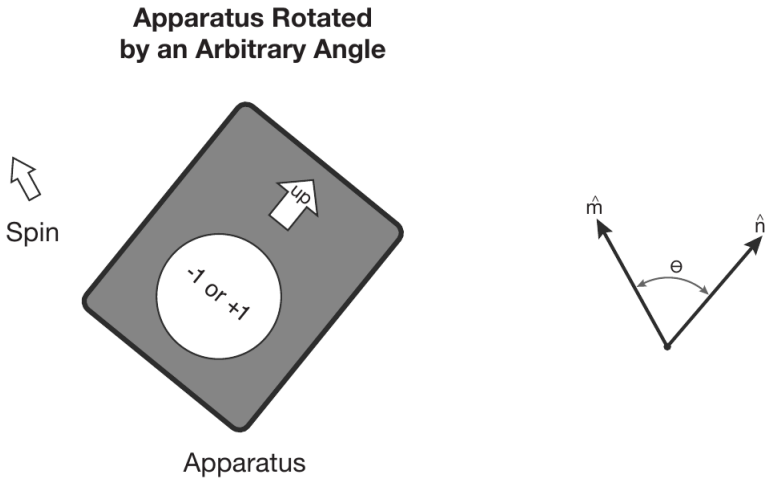


Figure 9. The apparatus rotated by an arbitrary angle within the $x-z$ plane. Average measurement result is $\hat{n}\cdot\hat{m}$

Classically, if S_σ were a vector, we would expect the result of the experiment to be the component of S_σ along the $\hat{n}$ axis. If $\hat{n}$ lies at an angle θ with respect to z, the classical answer would be $S_\sigma = \cos\theta$. But as you might guess, each time we do the experiment we get $S_\sigma = +1$ or $S_\sigma = -1$. Howerwer, the result is statistically biased so that the average value is $\cos\theta$.

The situation is of cource more general. We did not have to start with S oriented along z. Pick any direction $\hat{m}$ and start with the up-arrow pointing along $\hat{m}$. Prepare a spin so that the apparatus reads $+1$. Then, without disturbing the spin, rotate the apparatus to the direction $\hat{n}$, as shown in Figure(9). A new experiment on the same spin will give random results ± 1, **but with an average value equal to the cosine of the angle between** $\hat{n}$ **and** $\hat{m}$. In other words, the **average** will be $\hat{n}\cdot\hat{m}$.

The quantum mechanical notation for the statistical average of quantity $\mathcal{A}$ is Dirac's bracket notation $\langle\mathcal{A}\rangle$

We may summarize the results of our experimantal investigation as follows: If we begin with $\mathcal{A}$ oriented along $\hat{m}$ and confirm that $S_\sigma = +1$, then subsequent measurement with $\mathcal{S}$ oriented along $\hat{n}$ gives the statistical result

$$\langle S_\sigma\rangle = \hat{n}\cdot\hat{m}.$$

What we are learning is that quantum mechanical systems are not deterministic - the results of experiments can be statistically random - but if we repeat an experiment many times, average quantities can follow the expectations of classical physics, at least up to a point.

Experiments Are Never Gentle. *Any* interaction that is strong enough to measure some aspect of a system is necessarily strong enough to distupt some other aspect of the same system. One might say that measuring one component of the spin destroys the information about another component. In fact one symply cannot simultaneously know the components of the spin along two different axes, not in a reproducible way in any case. There is some fundamentally different about the state of a quantum system and the state of a classical system.

Let's label the possible spin states along the three coordinate axes.

If $\mathcal{S}$ os oriented along the z axis, the two possible states that can be prepared correspond to $S_z \pm 1$. Let's call them *up* and *down* and denote them by ket-vectors $|u\rangle$ and $|d\rangle$.

Thus, when the apparatus is oriented along the z axis and registers $+1$, the state $|u\rangle$ has been prepared. If the $\mathcal{S}$ is along the y axis, it can prepare the states $|i\rangle$ and $|o\rangle$ (*in* and *out*).

The idea that there are no hidden variables has a very simple mathematical representation: the space of states for a single spin has only two dimensions:

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

We could choose $|u\rangle$ and $|d\rangle$ as two basis vectors and write *any* state as a linear superposition of these two states.

Let's use the symbol $|\psi\rangle$ for a generic state. We can write this as an equation,

$$|\psi\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle$$

As we already know from the (108) the components of $|\psi\rangle$:

$$\begin{aligned} \alpha_u &= \langle u|\psi\rangle \\ \alpha_d &= \langle d|\psi\rangle \end{aligned} \tag{133}$$

The ψ can represent any state of the spin, prepared in any manner:

- Given that the spin has been prepared in the state ψ, and that the apparatus is oriented along z, the quantity $\alpha_u^* \alpha_u$ is the probability that the spin would be measured as $S_z = +1$. In other words, it is the probability of the spin being *up* if measured along the z axis.

- Likewise, $\alpha_d^* \alpha_d$ is the probability that S_z would be *down* if measured.

The α values, or quivalently $\langle u|\psi\rangle$ and $\langle d|\psi\rangle$, are called **probability amplitudes**. They are themselves not probabilities. The probabilities for measurement *up* and *down* are given by

$$\begin{aligned} P_u &= \langle \psi|u\rangle \langle u|\psi\rangle \\ P_d &= \langle \psi|d\rangle \langle d|\psi\rangle \end{aligned} \tag{134}$$

Before the measurement, all we have is the vector ψ, which represents the potential possibilities but not the actual values of our measurements.

Also:

$$\begin{aligned} \langle u|d\rangle &= 0 \\ \langle d|u\rangle &= 0 \\ \alpha_u^* \alpha_u + \alpha_d^* \alpha_d &= 1 \end{aligned}$$

This is equivalent to saying that the vector $|\psi\rangle$ is normilized to a unit vector:

$$\langle \psi|\psi\rangle = 1$$

Along the x axis.

We said before that we can represent any spin state as a linear combination of the basis vectors $|u\rangle$ and $|d\rangle$. We start with $|r\rangle$ (right) vector.

If $\mathcal{A}$ initially prepares $|r\rangle$ and then rotated to measure S_z, there will be equal probabilities for *up* and *down* Thus, $\alpha_u^* \alpha_u$ and $\alpha_d^* \alpha_d$ must both be equal to $\frac{1}{2}$. A simple vector that satisfies this rule

is

$$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle$$

When the spin has been prepared in the *left* configuration, the probabilities for S_z are again equal to $\frac{1}{2}$. And since:

$$\langle r|l\rangle = 0$$
$$\langle l|r\rangle = 0$$

Then:

$$|l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle$$

There is some ambiguity in the coice of $|l\rangle$. This is called the phase ambiguity. We can multiply the $|l\rangle$ by any complex number z. That will have no effect on whether it is orthogonal to $|r\rangle$, though in general the result will no longer be normilized. But if we choose $z = e^{i\theta}$ (where θ can be any real number), then there will be no effect on the normalization, because $e^{i\theta}$ has unit magnitude. In other words, $\alpha_u^*\alpha_u$ will remain equal to 1. Since a number of the form $z = e^{i\theta}$ is called the phase-factor, the ambiguity is called the phase ambiguity.

Along the y axis.

$$\langle i|o\rangle = 0$$

Again, physically, this means that if the spin is *in*, it is definitely not *out*.

There are additional restrictions on the vectors $|i\rangle$ and $|o\rangle$. Using the relationships expressed in (133) and (134) and the statistical results of our experiments, we can write the following:

$$\langle o|u\rangle\langle u|o\rangle = \frac{1}{2}$$
$$\langle o|d\rangle\langle d|o\rangle = \frac{1}{2}$$
$$\langle i|u\rangle\langle u|i\rangle = \frac{1}{2}$$
$$\langle i|d\rangle\langle d|i\rangle = \frac{1}{2}$$

In the first two equations, $|o\rangle$ takes the role of ψ from Eqs. (133) and (134). In the secoond two, $|i\rangle$ takes that role.

These conditiions state that if the spin in oriented along y, and is then mesured along z, it is equally likely to be *up* and *down*.

We should also expect that if the spin were measured along the x axis, it would be equally likely to be *right* or *left*. This leads to additional conditions:

$$\langle o|r\rangle\langle r|o\rangle = \frac{1}{2}$$
$$\langle o|l\rangle\langle l|o\rangle = \frac{1}{2}$$
$$\langle i|r\rangle\langle r|i\rangle = \frac{1}{2}$$
$$\langle i|l\rangle\langle l|i\rangle = \frac{1}{2}$$

These conditions are sufficient to determine the form of the vectors $|i\rangle$ and $|o\rangle$, apart from the phase ambiguity. Here is the result:

$$|i\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle$$
$$|o\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle$$

5.4. Commutation relations

5.5. Pauli matrices

In quantum mechanics, any physical observable is associated with a linear operator.

The eigenvalues of the linear opeerator, of the eigenvalues of the matrix, are the expectation values of the physical quantity, i.e., the values we expect to find if we measure the physical quantiity in an experiment. Spin is a physical observable since the associated angular momentum can be measured, as was done by Stern and Gerlach. Consequently, there must be a quantum mechanical operator associated with spin. Pauli derived the quantum mechanical operators for the spin components along three orthogonal axes - $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$. They are 2×2 complex matrices that came to be known as the Pauli spin matrices.

Pauli's approach was based on the premise that:

- (1) The measurement of the spin angular momentum component **along any coordinate axis** for an electron should give the results $+\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$

- (2) The operators for spin components along three mutually orthogoonal axes should obey commutation rules similar to those obeyed by the operators associated with components of the orbital angular momentum:

$$\begin{aligned} \hat{S}_y\hat{S}_z - \hat{S}_z\hat{S}_y &= i\hbar\hat{S}_x \\ \hat{S}_z\hat{S}_x - \hat{S}_x\hat{S}_z &= i\hbar\hat{S}_y \\ \hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x &= i\hbar\hat{S}_z \end{aligned} \tag{135}$$

Now in the Stern-Gelrach experiment, assuming that the **z-axis** this the axis joining south to norht pole of the magnet, the observation of two traces on the photographic plate **was interpreted as** being caused by a spin angular momentum $\vec{S}$ whose **z-component has two values** $\pm\frac{\hbar}{2}$.

Therefore, the matrix of operator $\hat{S}_z$ must be

(1): a 2×2 matrix, because such a matrix has *two* eigenvalues, and

(2): these eigenvalues must be $\pm\frac{\hbar}{2}$.

A 2×2 hermitian matrix that has eigenvalues of $\pm\frac{\hbar}{2}$ is the matrix

$$M_{2\times 2} = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Next, Pauli realized that since the choice of the **z-axis** as the axis joining the north and south poles of the magnet is completely arbitrary, the result of the Stern-Gerlach measurement should

not be affected if he had chosen this axis to be the $x-axis$ or $y-axis$, instead:

$$\begin{aligned} S_x &= \frac{\hbar}{2}\sigma_x \\ S_y &= \frac{\hbar}{2}\sigma_y \\ S_z &= \frac{\hbar}{2}\sigma_z \end{aligned} \tag{136}$$

Since S_x, S_y, S_z must have eigenvalues of $\pm\frac{\hbar}{2}$ it is obvious that the σ matrices must have eigenvalues of ± 1. Furthermore:

$$\begin{aligned} \sigma_y\sigma_z - \sigma_z\sigma_y &= i\hbar\sigma_x \\ \sigma_z\sigma_x - \sigma_x\sigma_z &= i\hbar\sigma_y \\ \sigma_x\sigma_y - \sigma_y\sigma_x &= i\hbar\sigma_z \end{aligned} \tag{137}$$

We have already found the σ_z: it is:

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Then Pauli needed to pick two matrices σ_x and σ_z such that:

(1): have eigenvalues of ± 1

(2): obey equation (137)

(3): must be hermitian

Pauli supposed hermitian matrices that have off-diagonal elements only, i.e.:

$$\sigma_x = \begin{bmatrix} 0 & a \\ a^* & 0 \end{bmatrix}$$

$$\sigma_y = \begin{bmatrix} 0 & b \\ b^* & 0 \end{bmatrix}$$

Since the eigenvalues of these matrices are ± 1, we must have $|a|^2 = |b|^2 = 1$, which leads to the possible choices for a and $b = \pm 1$ or $\pm i$.

Next, we must satisfy equations (137) and that mandates

$$Im(ab^*) = 1$$

Therefore, if we select $a = \pm 1$, then we must choose $b = -i$, and this yields

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and

$$\sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

This is how Pauli came up with expressions for $\sigma_x, \sigma_y, \sigma_z$. These matrices are called **Pauli spin matrices** and serve as operators for the spin components according to equation (136).

It is obvious that Pauli's choice was by no means unique. There are other legitimate choices (e.g., we could have choosen $a = -i$ and $b = \pm 1$), but Pauli's choice is now a history and universally adopted.

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The square of each of the Pauli matrices is the 2×2 **unit matrix**. Hence:

$$|S|^2 = S_x^2 + S_y^2 + S_z^2 = \frac{3}{4}\hbar^2[\mathbb{1}] = \bar{s}(\bar{s}+1)\hbar^2[\mathbb{1}]$$

-with $\bar{s} = \frac{1}{2}$. This should be compared with the equivalent relation for the orbital angular momentum operator.

Eigenvectors of the Pauli matrices: spinors.

The eigenvalues of the Pauli spin matrices are ± 1. We now evaluate the corresponding eigenvectors that we denote as $|\pm\rangle$

Matix σ_z: The eigenvectors of σ_z must satisfy

$$\sigma_z |\pm\rangle_z = \pm 1 |\pm\rangle_z$$

$$Det(\sigma_z - \lambda\mathbb{1}) = 0 \quad \Longrightarrow \quad \begin{bmatrix} 1-\lambda & 0 \\ 0 & -1-\lambda \end{bmatrix} = (1-\lambda)(-1-\lambda) - 0 = 0$$

$$\lambda_1 = 1, \lambda_2 = -1$$

$$\sigma_z |\psi\rangle = \lambda |\psi\rangle \quad \Longrightarrow \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} \psi_1 \\ -\psi_2 \end{bmatrix}$$

Let's consider eigenvalue $\lambda = +1$ and find corresponding eigenstate:

$$\begin{bmatrix} \psi_1 \\ -\psi_2 \end{bmatrix} = 1 \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \quad \Longrightarrow \quad \psi_1 = 1, \psi_2 = 0$$

$$|\psi\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and we donote it} \quad |+\rangle_z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Let's consider eigenvalue $\lambda = -1$ and find corresponding eigenstate:

$$\begin{bmatrix} \psi_1 \\ -\psi_2 \end{bmatrix} = -1 \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \implies \psi_1 = 0, \psi_2 = 1$$

$$|\psi\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and we donote it} \quad |-\rangle_z = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The $|+\rangle_z$ and $|-\rangle_z$ already unit vectors, i.e they are already normilized by default.
Matrix σ_y: the eigenvectors of σ_y must satisfy

$$\sigma_y |\pm\rangle_y = \pm 1 |\pm\rangle_y$$
$$\lambda_1 = 1, \lambda_2 = -1$$
$$\sigma_y |\psi\rangle = \lambda |\psi\rangle \iff \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} -i\psi_2 \\ i\psi_1 \end{bmatrix}$$

Let's consider eigenvalue $\lambda = +1$ and find corresponding eigenstate:

$$\begin{bmatrix} -i\psi_2 \\ i\psi_1 \end{bmatrix} = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

(1) If we rewrite ψ_2 in terms of ψ_1 e.g. $\psi_2 = i\psi_1$ then

$$|\psi\rangle = \begin{bmatrix} \psi_1 \\ i\psi_2 \end{bmatrix}$$

Examples of eigenvectors: $\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} 2 \\ 2i \end{bmatrix}, \begin{bmatrix} 30 \\ 30i \end{bmatrix}$, etc

(2) If we rewrite ψ_1 in terms of ψ_2 e.g. $\psi_1 = -i\psi_2$ then

$$|\psi\rangle = \begin{bmatrix} -i\psi_2 \\ \psi_2 \end{bmatrix}$$

Examples of eigenvectors: $\begin{bmatrix} -i \\ 1 \end{bmatrix}, \begin{bmatrix} -2i \\ 2 \end{bmatrix}, \begin{bmatrix} -30i \\ 30 \end{bmatrix}$, etc

But the generally accepted case is case 1.

So lets normilize eigenvector from the case (1): $|+\rangle_y = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$ - it is for eigenvalue $\lambda = 1$

Lets consider eigenvalue $\lambda = -1$ and find corresponding eigenstate:

$$\begin{bmatrix} -i\psi_2 \\ i\psi_1 \end{bmatrix} = -1 \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

Let's rewrite ψ_2 in terms of ψ_1 e.g. $\psi_2 = -i\psi_1$ then

$$|\psi\rangle = \begin{bmatrix} \psi_1 \\ -i\psi_2 \end{bmatrix}$$

Examples of eigenvectors: $\begin{bmatrix} 1 \\ -i \end{bmatrix}, \begin{bmatrix} 2 \\ -2i \end{bmatrix}, \begin{bmatrix} 30 \\ -30i \end{bmatrix}$, etc

Normalized eigenvector: $|-\rangle_y = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$

Matix σ_x: The eigenvectors of σ_x must satisfy

$$\sigma_x |\pm\rangle_x = \pm 1 |\pm\rangle_x$$

$$\begin{bmatrix} \psi_2 \\ \psi_1 \end{bmatrix} = 1 \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \implies \psi_1 = \psi_2 \quad \text{and} \quad \psi_2 = \psi_1$$

$$\begin{bmatrix} \psi_2 \\ \psi_1 \end{bmatrix} = -1 \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \implies \psi_2 = -\psi_1$$

- hence we have:

$$|+\rangle_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$|-\rangle_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The eigenvectors of the Puli matrices are examples of "spinprs" which are 2×1 column vectors that represent the spin state of an electron. If we know the spinor associated with an electron in a give state, we can deduce the electraon's spin orientation, i.e., find the quantuties $\langle \hat{S}_x \rangle, \langle \hat{S}_y \rangle, \langle \hat{S}_z \rangle$, where the angular brackets denote expectation values. We will see this later.

5.6. Bloch sphere

5.7. Born rule.

The fundamental assumption underlying the Born rule is that a measurement of the observable $\hat{A}$ will produce one of it's eigenvalues λ_i as a result. In what follows, $\psi \in H$ is a unit vector and hence a (pure) state in the usual sence. Then the Born rule states:

The result of a measurement of a quantity A on a single system is one of the eigenvalues λ_i of the observable $\hat{A}$.

Aafter the measurement, the state vector of the system is in the eigenspace corresponding to the eigenvalue that has been obtained.

Eigenvectors(eigenstates) of $\hat{A}$ constitute an orthonormal basis for the Hilbert space associated with the system.

Other examples of obseervables are: coordinate, momentum, energy... etc.

Now we can rewrite the (53) in terms of eigenvectors:

$$P(\lambda_i) = |\langle\psi_i|\psi\rangle|^2 = \langle\psi|\psi_i\rangle\langle\psi_i|\psi\rangle \tag{138}$$

- it is the probabilty to observe value λ_i, ψ_i - is an eigenvectors and ψ is the state-vector of a system.

Since eigenvectors of a Hermitian operator form an complete and orthogonal system, any state of a system can be represented as

$$|\psi\rangle = \sum_i \alpha_i |\psi_i\rangle = \sum_i |\psi_i\rangle\langle\psi_i|\psi\rangle$$

λ_i it is **eigenvalues** and α_i it is **probablity amplitudes**.

Possible results of a **measurement are the eigenvalues** of the operator that represents the observable.

The (138) it is a **Born rule** (53).

Example.

State of the system:

$$|\psi\rangle = \frac{1}{2}|\psi_1\rangle + \frac{1}{2}|\psi_2\rangle$$

- it is expanded in some base. If we are talking about the $\hat{A}$ basis then $|\psi_1\rangle$ and $|\psi_2\rangle$ are eigenvectors of the $\hat{A}$ operator:

$$\hat{A}|\psi_1\rangle = \lambda_1|\psi_1\rangle$$
$$\hat{A}|\psi_2\rangle = \lambda_2|\psi_2\rangle$$

- now λ_1 and λ_2 are eivenvalues of these eivenvectors of $\hat{A}$. With this information we know following:

A measurement of the observable $\hat{A}$ on our system will give the numerical value λ_1 with the probability $|\frac{1}{2}|^2 = \frac{1}{2}$ and λ_2 again with the probability $|\frac{1}{2}|^2 = \frac{1}{2}$. The fact that the total probability is equal to 1 tells us that this state is normilized.

After the measurement, the system is in a state corresponding to the eigenvalue. Hence the square of the probability amplitude tells us the probability of finding the system in corresponding eigenstate.

Eigenvalues are a property of operators and amplitudes are a property of states (or the decomposition of a state in a given basis).

5.8. The Von Neuman measurement postulate.

If on measuring $\hat{A}$ for a system in state $|\psi\rangle$, a result is λ_i is obtained, then the state of the system immediately after the measurement is $|\psi_i\rangle$.

This postulate can be rewritten by making use of the **projection operators** (124). Then the state of the system after the mesurement, for which the result α_i was obtained, is:

$$\frac{\hat{P}_i |\psi\rangle}{\sqrt{\langle\psi| \hat{P}_i |\psi\rangle}} = \frac{\hat{P}_i |\psi\rangle}{\sqrt{|\langle\psi_i|\psi\rangle|^2}} \tag{139}$$

- where the term in the denominator is there to guarentee that the state after the measurement is normalized to unity. Compate it with (138).

This postulate almost states the obvious in that we *name* a state according to the information that we obtain about it is as result of a measurement.

But it can also be argued that if, after performing a measurement that yields a particular result, we immediately repeat the measurement, it is reasonable to expect that there is a 100% chance that the same result be regained, which tells us that the system must have been in the associated **eigenstate**. This was, in fact, the main argument given by von Neumann to supportr this postulate. Thus, von Neumann argued that the fact that the value has a stable result upon repeated measurement indicates that the system really has that value after measurement.

The von Neumann postulate is quite clearly stating that as a consequence of a measurement, the state of the system undergoes a discontinuous change of state, i.e $|\psi\rangle \to |\psi_i\rangle$ if the result λ_i is obtained on performing a measurement of an obaservable $\hat{A}$.

This instantaneous change in state is known as **the collapse of the state vector**. This conjures up the impression that the process of measurement necessarily involves a physical interaction with the system, that moreover, results in a major physical disruption of the state of a system − one moment the system is in a state $|\psi\rangle$, the next it is forced into a state $|\psi_i\rangle$

5.9. Von Neumann and Luders postulates and quantum information theory

0905.4398.pdf

5.10. Probability amplitudes

5.11. Expectation(average) values

5.12. Pure and mixed states

5.13. Density matrix