"Thinking Quantum": Lectures on Quantum Theory for High-School Students

Barak Shoshany

Perimeter Institute for Theoretical Physics
31 Caroline St. N., Waterloo, ON, Canada, N2L 2Y5

bshoshany@perimeterinstitute.ca

v2: October 24, 2017

Abstract

We present a conceptually clear introduction to quantum theory at a level suitable for high-school students attending the International Summer School for Young Physicists (ISSYP) at Perimeter Institute. It is entirely self-contained and no university-level background knowledge is required. The lectures we given over four days, four hours each day. On the first day the students were given all the relevant mathematical background from linear algebra and probability theory. On the second day, we used the acquired mathematical tools to define the full quantum theory in the case of a finite Hilbert space and discuss some consequences such as entanglement, Bell's theorem and the uncertainty principle. Finally, in days three and four we presented an overview of advanced topics related to infinite-dimensional Hilbert spaces, including canonical and path integral quantization, the quantum harmonic oscillator, quantum field theory, the Standard Model, and quantum gravity. Any typos or errors should be reported to the email address above.

Contents

Day 1: Background Material

Quantum theory is the theoretical framework believed to describe all aspects of our universe at the most fundamental level. Mathematically, it is surprisingly simple. However, conceptually, it is very

hard to understand using the intuition we have from living in a "classical" world. In these lectures we will learn to develop quantum intuition.

First we shall learn some basic mathematical tools of linear algebra and probability theory. These tools will be used extensively later on, so pay attention and make sure to solve all of the exercises.

1 Linear Algebra

1.1 Vectors

We shall mostly deal with n-dimensional complex vectors, which are simply an ordered list of n complex numbers, also known as an n-tuple. For example, a 2-dimensional complex vector is written as:

$$|\Psi\rangle = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}, \qquad \Psi_1, \Psi_2 \in \mathbb{C}, \tag{1.1}$$

where the notation $\Psi_1, \Psi_2 \in \mathbb{C}$ means that the *components* Ψ_1, Ψ_2 are complex numbers. The notation $|\Psi\rangle$ is called *Dirac (or "bra-ket") notation* and Ψ is simply a label used to refer to the vector. We can write anything as the label, including English/Greek letters, numbers, or even words, for example:

$$|A\rangle, |\beta\rangle, |3\rangle, |\text{cat}\rangle, \ldots \tag{1.2}$$

Addition of vectors is defined by simply adding the components. That is, if

$$|\Phi\rangle = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}, \tag{1.3}$$

then

$$|\Psi\rangle + |\Phi\rangle = \begin{pmatrix} \Psi_1 + \Phi_1 \\ \Psi_2 + \Phi_2 \end{pmatrix}. \tag{1.4}$$

We also define multiplication of the whole vector by a *scalar*, that is, a complex number $\lambda \in \mathbb{C}$, by multiplying all of the components:

$$\lambda |\Psi\rangle = \begin{pmatrix} \lambda\Psi_1 \\ \lambda\Psi_2 \end{pmatrix}. \tag{1.5}$$

λ is called a scalar because it "scales" the vector. All the n-dimensional complex vectors together make up a *vector space* called $\mathbb{C}^n$.

Exercise: Show that addition and multiplication by a scalar are distributive, that is, $\lambda(|\Psi\rangle + |\Phi\rangle) = \lambda|\Psi\rangle + \lambda|\Phi\rangle$ and $(\alpha + \beta)|\Psi\rangle = \alpha|\Psi\rangle + \beta|\Psi\rangle$.

1.2 Dual Vectors, Inner Product and Norm

A *dual vector* is defined by writing the vector as a row instead of a column, and replacing each component with its complex conjugate. We denote the dual vector of $|\Psi\rangle$ as follows:

$$\langle\Psi| = \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix}, \tag{1.6}$$

where the $*$ denotes the *complex conjugate*, defined for any complex number z as follows:

$$z = a + \mathrm{i}\, b \quad \Longrightarrow \quad z^* = a - \mathrm{i}\, b, \qquad a, b \in \mathbb{R}. \tag{1.7}$$

Addition and multiplication by a scalar is defined as for vectors, but you may not add vectors and dual vectors together! Note that the dual of a dual vector is again a vector, since $(z^*)^* = z$.

Using the dual vectors, we may define the *inner product*. This product allows us to take two vectors and produce one complex number out of them. The inner product only works for one vector and one dual vector, and to calculate it we multiply the components of both vectors one by one and add them up:

$$\langle \Psi | \Phi \rangle = \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \Psi_1^* \Phi_1 + \Psi_2^* \Phi_2. \tag{1.8}$$

In bra-ket notation, vectors $|\Psi\rangle$ are "kets" and dual vectors $\langle\Psi|$ are "bras". Then the notation for $\langle \Psi | \Phi \rangle$ is called a "bracket".

Note that this inner product resembles the dot product defined for vectors in $\mathbb{R}^n$, but here one must be careful to only multiply vectors by dual vectors. Finally, we may define the *norm-squared* of a vector by taking its inner product with itself ("squaring" it):

$$\|\Psi\|^2 = \langle \Psi | \Psi \rangle = \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = |\Psi_1|^2 + |\Psi_2|^2, \tag{1.9}$$

where the magnitude-squared of a complex number z is defined to be

$$|z|^2 = z^* z. \tag{1.10}$$

Then we can define the *norm* as the square root of the norm-squared:

$$\|\Psi\| = \sqrt{\|\Psi\|^2} = \sqrt{\langle \Psi | \Psi \rangle}. \tag{1.11}$$

A vector space with an inner product satisfying the properties you will now prove in the exercise is called a *Hilbert space*. Thus, $\mathbb{C}^n$ is a Hilbert space.

Exercise:

1. Prove that the norm-squared $\|\Psi\|^2$ is always non-negative, and it is zero if and only if $|\Psi\rangle$ is the zero vector, that is, the vector whose components are all zero.
2. Prove that $\langle \Phi | \Psi \rangle = \langle \Psi | \Phi \rangle^*$, that is, if we swap the vectors in the product we get the complex conjugate. (Thus the inner product is non-commutative.)
3. Prove that if $\alpha, \beta \in \mathbb{C}$ and $|\Psi\rangle, |\Phi\rangle, |\Theta\rangle \in \mathbb{C}^n$ then

   $$\langle \Psi | \left(\alpha |\Phi\rangle + \beta |\Theta\rangle \right) = \alpha \langle \Psi | \Phi \rangle + \beta \langle \Psi | \Theta \rangle, \tag{1.12}$$

   that is, the inner product is linear in its second argument.

1.3 Orthonormal Bases

An *orthonormal basis* for $\mathbb{C}^n$ is a set of n non-zero vectors $|B_i\rangle$ such that:

1. They are *linearly independent*, which means that if

   $$\sum_{i=1}^{n} \lambda_i |B_i\rangle = 0, \qquad \lambda_i \in \mathbb{C} \text{ for all } i, \tag{1.13}$$

   then necessarily $\lambda_i = 0$ for all i.
2. They span $\mathbb{C}^n$, which means that any vector $|\Psi\rangle \in \mathbb{C}^n$ can be written as a *linear combination* of the basis vectors:

   $$|\Psi\rangle = \sum_{i=1}^{n} \lambda_i |B_i\rangle, \tag{1.14}$$

   for a unique choice of $\lambda_i \in \mathbb{C}$.

3. They are all *orthogonal* to each other, that is, the inner product of any two different vectors evaluates to zero:
$$\langle B_i|B_j\rangle = 0 \text{ for all } i \neq j. \tag{1.15}$$

4. They are all *unit vectors*, that is, have a norm of 1:
$$\||B_i\rangle\|^2 = \langle B_i|B_i\rangle = 1 \text{ for all } i. \tag{1.16}$$

In fact, properties 3 and 4 may be expressed more compactly as:
$$\langle B_i|B_j\rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j, \end{cases} \tag{1.17}$$
where δ_{ij} is called the *Kronecker delta*.

Exercise:

1. Show that property 1 means that no vector in the basis can be written as a linear combination of the other vectors in the basis.

2. Any basis which is orthogonal but not orthonormal, that is, does not satisfy property 4, can be made orthonormal by normalizing each basis vector, i.e. dividing it by its norm: $|B_i\rangle \mapsto |B_i\rangle \,/\, \|B_i\|$. Assume a basis which satisfies properties 1-3 and show that after normalizing it, properties 1-3 are still satisfied.

These requirements become much simpler in $n = 2$ dimensions. Then an orthonormal basis for $\mathbb{C}^2$ is a set of 2 non-zero vectors $|B_1\rangle, |B_2\rangle$ such that:

1. They are *linearly independent*, which means that we cannot write one in terms of a scalar times the other, i.e.:
$$|B_1\rangle \neq \lambda\,|B_2\rangle, \qquad \lambda \in \mathbb{C}. \tag{1.18}$$

2. They span $\mathbb{C}^2$, which means that any vector $|\Psi\rangle \in \mathbb{C}^2$ can be written as a *linear combination* of the basis vectors:
$$|\Psi\rangle = \lambda_1\,|B_1\rangle + \lambda_2\,|B_2\rangle, \tag{1.19}$$
for a unique choice of $\lambda_1, \lambda_2 \in \mathbb{C}$.

3. They are *orthogonal* to each other, that is, the inner product between them evaluates to zero:
$$\langle B_1|B_2\rangle = 0. \tag{1.20}$$

4. They are *unit vectors*, that is, have a norm of 1:
$$\|B_1\|^2 = \langle B_1|B_1\rangle = 1, \qquad \|B_2\|^2 = \langle B_2|B_2\rangle = 1. \tag{1.21}$$

A very important basis, the *standard basis* of $\mathbb{C}^2$, is defined as:
$$|B_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |B_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.22}$$

Exercise: Show that the standard basis vectors satisfy the four properties above.

1.4 Matrices and Adjoint

Matrices in n dimensions are $n \times n$ arrays of complex numbers[1]. In $n = 2$ dimensions we have

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad A_{11}, A_{12}, A_{21}, A_{22} \in \mathbb{C}. \tag{1.23}$$

A matrix acts on a vector from the left by taking the inner product of each row of the matrix with the vector:

$$A|\Psi\rangle = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = \begin{pmatrix} A_{11}\Psi_1 + A_{12}\Psi_2 \\ A_{21}\Psi_1 + A_{22}\Psi_2 \end{pmatrix}. \tag{1.24}$$

A matrix can also act on a dual vector from the right by taking the inner product of the dual vector with each column of the matrix:

$$\langle\Psi|A = \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} \Psi_1^* A_{11} + \Psi_2^* A_{21} & \Psi_1^* A_{12} + \Psi_2^* A_{22} \end{pmatrix}. \tag{1.25}$$

Note that the dual vector $\langle\Psi|A$ is not the dual of the vector $A|\Psi\rangle$. However, we can define the *adjoint* of a matrix by both transposing rows into columns and taking the complex conjugate of the components:

$$A^\dagger = \begin{pmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{pmatrix}, \tag{1.26}$$

where the notation † for the adjoint is called "dagger". Then the vector dual to $A|\Psi\rangle$ is $\langle\Psi|A^\dagger$.
The *identity matrix*, written simply as 1, is:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.27}$$

Acting with it on any vector or dual vector does not change it: $1|\Psi\rangle = |\Psi\rangle$.
Exercise:

1. Show that the vector dual to $A|\Psi\rangle$ is indeed $\langle\Psi|A^\dagger$.
2. Show that $\left(A^\dagger\right)^\dagger = A$.
3. Show that the action of a matrix on a vector is *linear*, that is, $A(\alpha|\Psi\rangle + \beta|\Phi\rangle) = \alpha A|\Psi\rangle + \beta A|\Phi\rangle$.

1.5 The Outer Product

We have seen that vectors and dual vectors may be combined to find a complex number using the inner product. We can similarly combine a vector and a dual vector to find a matrix, using the *outer product*. For the vectors defined above:

$$\langle\Psi| = \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix}, \qquad |\Phi\rangle = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}, \tag{1.28}$$

we define the outer product as the matrix whose component at row i, column j is given by multiplying the component at column j of $|\Psi\rangle$ and row i of $\langle\Phi|$:

$$|\Phi\rangle\langle\Psi| = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix} = \begin{pmatrix} \Psi_1^*\Phi_1 & \Psi_2^*\Phi_1 \\ \Psi_1^*\Phi_2 & \Psi_2^*\Phi_2 \end{pmatrix}. \tag{1.29}$$

[1] In fact, matrices don't have to be square, they can have a different number of rows and columns, but non-square matrices will not interest us here.

Exercise: Calculate the outer product $|\Psi\rangle\langle\Phi|$ for

$$|\Psi\rangle = \begin{pmatrix} 1 \\ 2+\mathrm{i} \end{pmatrix}, \qquad |\Phi\rangle = \begin{pmatrix} 3-\mathrm{i} \\ 4\,\mathrm{i} \end{pmatrix}. \tag{1.30}$$

Remember that when writing the dual vector, the components are complex-conjugated!

1.6 The Completeness Relation

When writing the vector $|\Psi\rangle$ as a linear combination of basis vectors,

$$|\Psi\rangle = \sum_{i=1}^{n} \lambda_i \, |B_i\rangle \,, \tag{1.31}$$

we may in fact calculate and find that $\lambda_i = \langle B_i|\Psi\rangle$ for all i.

Exercise: Perform this calculation by taking the inner product of the above equation with $\langle B_j|$ and using the fact that the basis vectors are orthonormal.

Now, since λ_i is a scalar we can move it to the right:

$$|\Psi\rangle = \sum_{i=1}^{n} |B_i\rangle \, \lambda_i. \tag{1.32}$$

Then, replacing λ_i with $\langle B_i|\Psi\rangle$ we get

$$|\Psi\rangle = \sum_{i=1}^{n} |B_i\rangle\langle B_i|\Psi\rangle. \tag{1.33}$$

To make this even more suggestive, let us add parentheses:

$$|\Psi\rangle = \left(\sum_{i=1}^{n} |B_i\rangle\langle B_i| \right) |\Psi\rangle. \tag{1.34}$$

Note that what we did here is go from a vector $|B_i\rangle$ times a complex number $\langle B_i|\Psi\rangle$ to a matrix $|B_i\rangle\langle B_i|$ times a vector $|\Psi\rangle$, for each i. The fact that they are, in fact, equal to one another (as you will prove in the exercise) is made intuitive by using the bra-ket notation. The notation now suggests that

$$\sum_{i=1}^{n} |B_i\rangle\langle B_i| = 1, \tag{1.35}$$

where $|B_i\rangle\langle B_i|$ is the outer product defined above, and the 1 on the right-hand side is the identity matrix.

In $\mathbb{C}^2$, we simply have

$$|B_1\rangle\langle B_1| + |B_2\rangle\langle B_2| = 1. \tag{1.36}$$

Exercise:

1. Provide a more rigorous proof that

$$\sum_{i=1}^{n} |B_i\rangle\langle B_i|\Psi\rangle = \left(\sum_{i=1}^{n} |B_i\rangle\langle B_i| \right) |\Psi. \tag{1.37}$$

2. Given the basis

$$|B_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad |B_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{1.38}$$

first show that it is indeed an orthonormal basis, and then show that it satisfies the completeness relation $|B_1\rangle\langle B_1| + |B_2\rangle\langle B_2| = 1$.

1.7 Multiplication of Matrices

The product of two matrices is calculated by taking the inner product of each row of the left matrix with each column on the right matrix:

$$AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}. \tag{1.39}$$

Exercise:

1. Calculate the products AB and BA where:

$$A = \begin{pmatrix} 1+2\,\mathrm{i} & 3-4\,\mathrm{i} \\ 5-6\,\mathrm{i} & 7+8\,\mathrm{i} \end{pmatrix}, \qquad B = \begin{pmatrix} 9-8\,\mathrm{i} & 7+6\,\mathrm{i} \\ 5+4\,\mathrm{i} & 3-2\,\mathrm{i} \end{pmatrix}. \tag{1.40}$$

2. Show that $(AB)^\dagger = B^\dagger A^\dagger$.

3. Show that matrix multiplication is not commutative, unless at least one of the matrices is diagonal, that is, of the form

$$D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}. \tag{1.41}$$

4. Show that multiplying by a scalar is the same as multiplying by a diagonal matrix with the scalar at each components of the diagonal, that is:

$$\lambda A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} A. \tag{1.42}$$

1.8 Inner Products with Matrices

Since $A\,|\Phi\rangle$ is itself a vector, we may now calculate the inner product of that vector with the dual vector $\langle\Psi|$, which as usual gives us a complex number:

$$\langle\Psi|A|\Phi\rangle = \begin{pmatrix} \Psi_1^* & \Psi_2^* \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \Psi_1^* A_{11}\Phi_1 + \Psi_2^* A_{21}\Phi_1 + \Psi_1^* A_{12}\Phi_2 + \Psi_2^* A_{22}\Phi_2. \tag{1.43}$$

If we take the dual of $A\,|\Phi\rangle$ we get $\langle\Phi|\,A^\dagger$, as you proved above. Thus, inverting the order of the inner product, we get

$$\langle\Phi|A^\dagger|\Psi\rangle = \begin{pmatrix} \Phi_1^* & \Phi_2^* \end{pmatrix} \begin{pmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = \Psi_1 A_{11}^*\Phi_1^* + \Psi_2 A_{21}^*\Phi_1^* + \Psi_1 A_{12}^*\Phi_2^* + \Psi_2 A_{22}^*\Phi_2^*. \tag{1.44}$$

This is, of course, the complex conjugate of $\langle\Psi|A|\Phi\rangle$, since inverting the order of the inner product results in the complex conjugate. In other words, we have

$$\langle\Psi|A|\Phi\rangle^* = \langle\Phi|A^\dagger|\Psi\rangle. \tag{1.45}$$

Taking the complex conjugate thus reverses the order of the inner product, and also replaces the matrix with its adjoint.

Exercise: Calculate the inner product $\langle\Psi|A|\Phi\rangle$ where

$$|\Psi\rangle = \begin{pmatrix} 5+2\,\mathrm{i} \\ 4-3\,\mathrm{i} \end{pmatrix}, \qquad A = \begin{pmatrix} 9 & 8\,\mathrm{i} \\ 7+6\,\mathrm{i} & 5-4\,\mathrm{i} \end{pmatrix}, \qquad |\Phi\rangle = \begin{pmatrix} 3+4\,\mathrm{i} \\ 2-5\,\mathrm{i} \end{pmatrix}. \tag{1.46}$$

1.9 Eigenvalues and Eigenvectors

If the matrix A, acting on the vector $|\Psi\rangle$, results in a scalar multiple of $|\Psi\rangle$:

$$A|\Psi\rangle = \lambda|\Psi\rangle, \tag{1.47}$$

then we call $|\Psi\rangle$ an *eigenvector* of A and λ its *eigenvalue*. For example, if

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.48}$$

then it's easy to see that

$$|\Psi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.49}$$

is an eigenvector with eigenvalue $+1$ and

$$|\Phi\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.50}$$

is an eigenvector with eigenvalue -1.

Exercise:

1. The matrix
$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \tag{1.51}$$
has two eigenvectors. Find them and their corresponding eigenvalues.

2. Prove that, if $|\Psi\rangle$ is an eigenvector, then $\alpha|\Psi\rangle$ is also an eigenvector for any $\alpha \in \mathbb{C}$, and it has the same eigenvalue.

1.10 Hermitian Matrices

There is a special kind of matrix, called a *Hermitian* matrix, which is equal to its own adjoint:

$$A = A^\dagger. \tag{1.52}$$

For such a matrix, we have $\langle\Psi|A|\Phi\rangle^* = \langle\Phi|A|\Psi\rangle$.

Exercise:

1. Find the most general 2×2 Hermitian matrix by demanding that $A = A^\dagger$ and finding conditions on the components of A.

2. Prove that the eigenvalues of a Hermitian matrix are always real.

3. Prove that two eigenvectors $|\Psi\rangle, |\Phi\rangle$ of a Hermitian matrix with different eigenvalues are necessarily *orthogonal*, that is, $\langle\Phi|\Psi\rangle = 0$.

4. Prove that the eigenvectors of a Hermitian matrix, when properly normalized, make up an orthonormal basis of $\mathbb{C}^n$. Assume that all eigenvectors have distinct eigenvalues.[2]

[2] If two eigenvectors correspond to the same eigenvalue, they will not necessarily be orthogonal. This case is called *degenerate eigenvalues*. (Give an example of a Hermitian matrix which has degenerate eigenvalues!) We will not deal with degenerate eigenvalues in these lectures. However, it is in fact possible to find an orthonormal basis even if there are degenerate eigenvalues, using the *Gram-Schmidt process*, which the student is encouraged to look up and study.

1.11 Unitary Matrices

A *unitary matrix* U is a matrix which, when multiplied by its adjoint on either side, results in the identity matrix:

$$UU^\dagger = U^\dagger U = 1. \tag{1.53}$$

Exercise: Find the most general 2×2 unitary matrix by demanding that $UU^\dagger = U^\dagger U = 1$ and finding conditions on the components of U.

1.12 The Spectral Decomposition Theorem

The *spectral decomposition theorem* says that any Hermitian matrix A with (real) eigenvalues λ_i and a corresponding orthonormal basis of eigenstates $|B_i\rangle$ may be *diagonalized* in terms of an outer product of the basis eigenstates:

$$A = \sum_i \lambda_i |B_i\rangle \langle B_i| . \tag{1.54}$$

This will be referred to as the *spectral decomposition* of A. We also call this *diagonalization* because this means the operator may be represented as a diagonal matrix with the eigenvalues λ_i on the diagonal (but we are not going to give the details here because they are not important for what follows).

Exercise:

1. Write the spectral decomposition of the following matrix:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.55}$$

2. In fact, the theorem is true for any *normal matrix*, which satisfies $AA^\dagger = A^\dagger A$. Prove that a Hermitian matrix is a special case of a normal matrix. Prove that a unitary matrix is also a special case of a normal matrix.

3. Prove the spectral decomposition theorem.

1.13 The Cauchy-Schwarz Inequality

An important result in the theory of Hilbert spaces is the Cauchy-Schwarz inequality, which states that for any two vectors $|\Psi\rangle$ and $|\Phi\rangle$, we have

$$|\langle\Psi|\Phi\rangle| \leq \|\Psi\| \, \|\Phi\| . \tag{1.56}$$

Exercise: Prove the Cauchy-Schwarz inequality.

2 Probability Theory

2.1 Random Variables

A *random variable* X is a function which assigns a real value to each possible outcome of an experiment or process. Sometimes these values will be the actual measured value in some way: for example, the value of the random variable X for rolling a 6-sided die will simply be the number on the die. Other times, the value of the random variable will be just a numerical label assigned to each outcome: for example, for a coin toss we can assign 1 to heads and 0 to tails.

These examples were of *discrete* random variables, but we can also have *continuous* random variables, such as the position of a particle along a line, which in principle can take any real value. There are some subtleties to continuous random variables, which will not concern us here.

Exercise: Think of more examples of discrete and continuous random variables.

2.2 Probability Distributions

A (discrete) *probability distribution* assigns a probability to each value of a random variable. For example, for the coin toss we have

$$P(X=1)=\frac{1}{2}, \qquad P(X=0)=\frac{1}{2}, \tag{2.1}$$

and for the die roll we have

$$P(X=1)=\frac{1}{6}, \qquad P(X=2)=\frac{1}{6}, \qquad P(X=3)=\frac{1}{6}, \tag{2.2}$$

$$P(X=4)=\frac{1}{6}, \qquad P(X=5)=\frac{1}{6}, \qquad P(X=6)=\frac{1}{6}. \tag{2.3}$$

The sum of probabilities for all possible values of X is always 1.

These probability distributions are *uniform*, since they assign the same probability to each value of X. However, probability distributions need not be uniform. For example, if we toss two coins X_1 and X_2 and add the results, $X = X_1 + X_2$, we can get any of the following 4 outcomes:

$$0+0=0, \qquad 0+1=1, \qquad 1+0=1, \qquad 1+1=2. \tag{2.4}$$

The probability for each outcome is

$$\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}, \tag{2.5}$$

but the outcome 1 appears twice; thus

$$P(X=0)=\frac{1}{4}, \qquad P(X=1)=\frac{1}{2}, \qquad P(X=2)=\frac{1}{4}. \tag{2.6}$$

Exercise: Calculate the probability distribution for the sum of two rolls of a 6-sided die (a 2d6 in D&D terms, since we define $n\mathrm{d}N$ to be the sum of n rolls of an N-sided die).

2.3 Expected Values

The *expected value* (or *expectation value* or *mean*) of a random variable is the average value of many repetitions of the same experiment represented by the random variable. We denote it with angle brackets $\langle X\rangle$. To calculate it, we simply add up all the possible values of X, each weighted by its assigned probability:

$$\langle X\rangle = \sum_{i=1}^{N} p_i x_i, \tag{2.7}$$

where N is the total number of possible outcomes, x_i is the value of outcome number i and p_i is its probability. In the example of the coin toss, we have:

$$\langle X\rangle = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 1 = \frac{1}{2}, \tag{2.8}$$

and for the die roll, we have:

$$\langle X \rangle = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = \frac{7}{2} = 3.5. \tag{2.9}$$

Note that the expected value is often not an actual value the random variable can take.

Exercise: Calculate the expected value for the sum of two coin tosses and for the 2d6. What do you learn from that?

2.4 Standard Deviation

The *standard deviation* measures how far the outcomes are expected to be from the expected value[3]. To calculate the standard deviation, we take the expected value of $(X - \langle X \rangle)^2$, that is, of the difference between the actual value of X and its expected value. Then, we take the square root of the result:

$$\begin{aligned} \sigma_X &= \sqrt{\left\langle (X - \langle X \rangle)^2 \right\rangle} \\ &= \sqrt{\langle X^2 \rangle - \langle X \rangle^2}. \end{aligned}$$

Exercise: Prove that the second line follows from the first. Use the fact that the expected value is linear, $\langle X + Y \rangle = \langle X \rangle + \langle Y \rangle$ and $\langle \lambda X \rangle = \lambda \langle X \rangle$ where λ is a constant, and that $\langle X \rangle$ itself is a constant.

For example, for the coin toss we have from before

$$\langle X \rangle = \frac{1}{2}, \tag{2.10}$$

and we also calculate:

$$\left\langle X^2 \right\rangle = \frac{1}{2} \cdot 0^2 + \frac{1}{2} \cdot 1^2 = \frac{1}{2}, \tag{2.11}$$

which gives us

$$\sigma_X = \sqrt{\frac{1}{2} - \frac{1}{4}} = \frac{1}{2}. \tag{2.12}$$

This is to be expected, since the two actual values of the outcomes, 0 and 1, lie exactly $1/2$ away from the mean in each direction.

For the die roll, we have from before

$$\langle X \rangle = \frac{7}{2}, \tag{2.13}$$

and we also calculate:

$$\left\langle X^2 \right\rangle = \frac{1}{6} \left(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 \right) = \frac{91}{6}, \tag{2.14}$$

which gives us

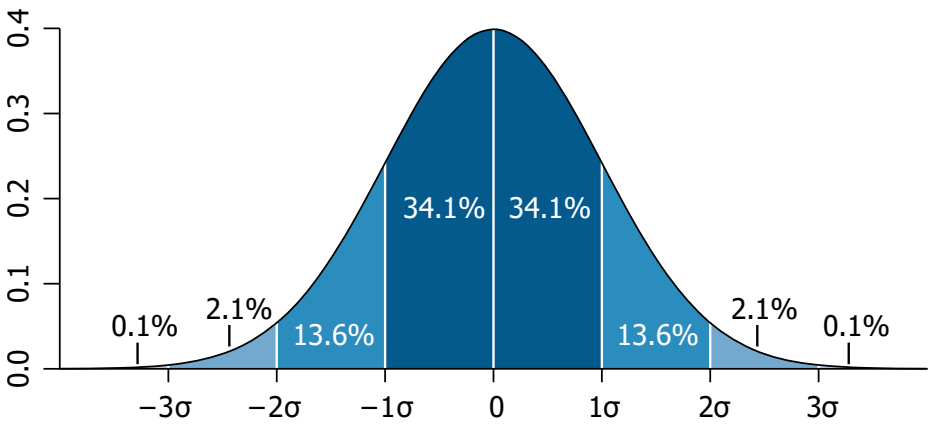
$$\sigma_X = \sqrt{\frac{91}{6} - \frac{49}{4}} = \sqrt{\frac{35}{12}} \approx 1.7. \tag{2.15}$$

Exercise: Calculate the standard deviation for the sum of two coin tosses and for the 2d6.

[3] The square of the standard deviation is called the *variance*, but it will not interest us here.

2.5 Normal (Gaussian) Distributions

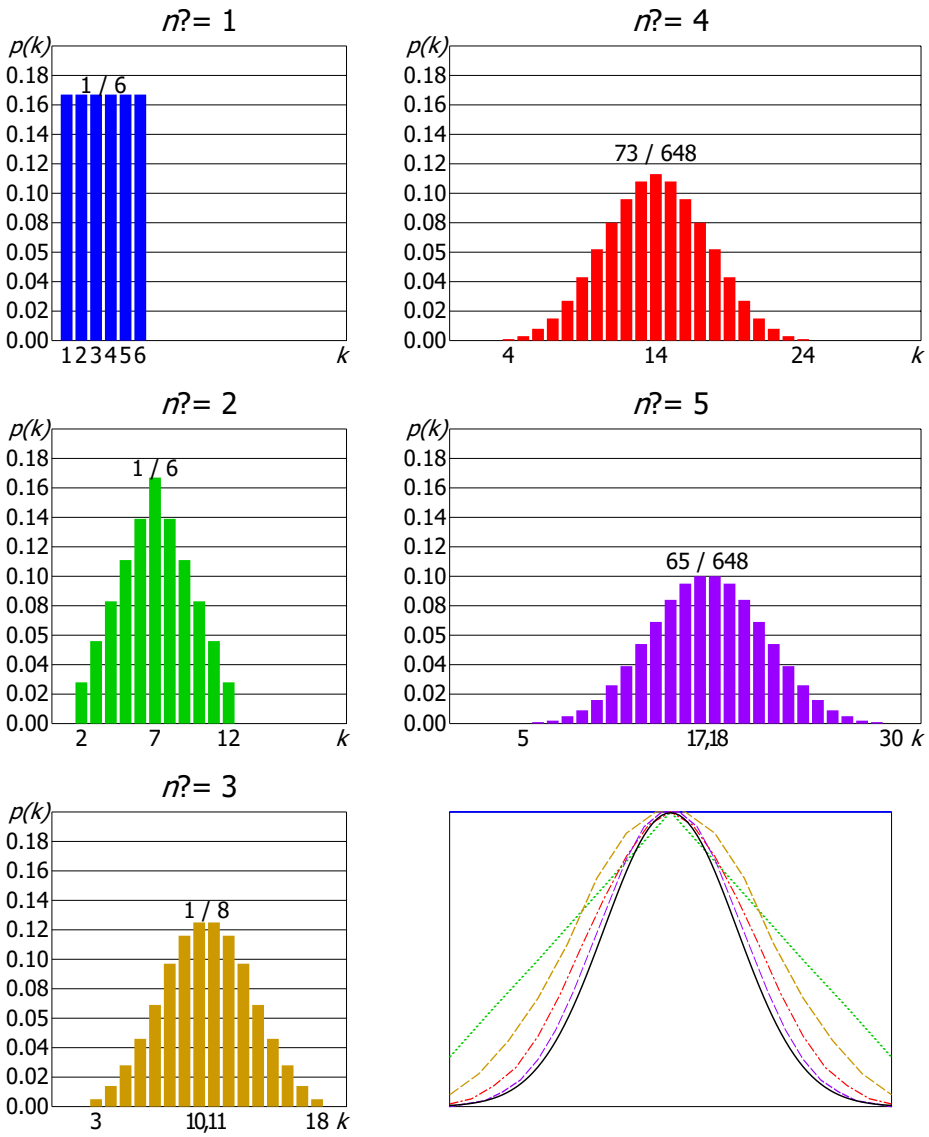
The *normal (or Gaussian) distribution* looks like this (source: Wikipedia):



It is a "bell curve" centered on some mean value μ (equal to 0 in the plot) and with a standard deviation σ. The percentage tells us the fraction of outcomes which lie within 1, 2 and 3 standard deviations from the mean.

The normal distribution is the most common probability distribution you will encounter. The reason for that is that there is a theorem, the *central limit theorem*, which states that whenever we add independent random variables, the probability distribution of the sum will look like a normal distribution. As we add more and more variables, the sum will be closer and closer to a normal distribution.

This can already be seen in the case of the die roll: for 1d6 we have a uniform distribution, but for 2d6, the sum of two die rolls, we get a triangular "bell curve" centered at the mean value of 7. For nd6, that is, the sum of n rolls of a 6-sided die, we will get closer and closer to a normal distribution. This is shown in the plot below up to $n = 5$ (source: Wikipedia).



At the limit $n \to \infty$, we will precisely obtain a normal distribution, but even for small values of n, the approximation is close enough for all practical purposes.

Exercise: Plot the probability distributions of the sum of n coin tosses up to whatever value of n satisfies you. Note how the distribution looks more and more like a normal distribution.

Day 2: Quantum Theory as a Mathematical Framework

Now that we have obtained the required background knowledge, we can finally present quantum theory. This theory provides the correct fundamental mathematical framework for most of known physics, with the notable exception of gravity. We will see that its fundamental ingredients are Hilbert spaces with states and operators. These universal ingredients are then used to create particular mod-

els describing specific physical systems.

3 Quantum Theory with Finite Hilbert Space

3.1 A Note about Units and Constants

Some constants in physics are *dimensionless*, and thus their numerical value has a physical significance. An example for that is the *fine-structure constant*, which represents the strength of the electromagnetic interaction:

$$\alpha \approx \frac{1}{137}. \tag{3.1}$$

This constant is not specified in any particular units (e.g. meters or second); it is just a number. We call constants that do not depend on the chosen system of units *dimensionless*.

In contrast, some constants in physics are *dimensionful*, that is, have dimensions. This means that they are specified in certain units, and their numerical value thus depends on the chosen system of units. For example, the speed of light c has the following values in different systems of units:

$$\begin{aligned} c &\approx 3.0 \times 10^8 \,\text{meters/second} \\ &\approx 1.1 \times 10^7 \,\text{miles/minute} \\ &\approx 170 \,\text{astronomical units/day} \\ &\approx 3.5 \times 10^{-5} \,\text{parsecs / hour} \\ &\approx 1 \,\text{light year / year}. \end{aligned}$$

The numerical value thus does not have any physical meaning whatsoever! It is merely a consequence of choosing to work with one system of units (e.g. meters and seconds) and not another. For this reason, we can simply choose to work in *Planck units*, where:

$$c = 8\pi G = \hbar = \frac{1}{4\pi\varepsilon_0} = k_B = 1. \tag{3.2}$$

Here c is the speed of light, G is the *gravitational constant*, $\hbar$ is the *(reduced) Planck constant* used in quantum mechanics, $1/4\pi\varepsilon_0$ is the *Coulomb constant* used in electromagnetism, and k_B is the *Boltzmann constant* used in statistical mechanics.

Planck units are commonly used in quantum gravity research, and we will also use them in these lecture notes. This means that $\hbar$ will not appear in any of our equations!

3.2 States and Operators

A *system* in a quantum theory is the mathematical representation of a physical system (such as a particle) as a Hilbert space. The type and dimension of the Hilbert space depend on the particular system; note that the dimension of the Hilbert space is unrelated to the dimension of spacetime. In the finite-dimensional case, the Hilbert space will usually be $\mathbb{C}^n$ for some n, such as $\mathbb{C}^2$, which was used in the examples above. In the infinite-dimensional case, it will usually be a space of functions, which is much more complicated, and we will not describe it in detail here.

An *operator* on a quantum system is a matrix in the appropriate Hilbert space. It represents an action performed on the system, such as a measurement, a transformation, or time evolution. In the continuous case, where the states are functions, this "matrix" will in fact correspond to derivatives acting on the functions.

A *state* of a quantum system is a vector with norm 1 in the appropriate Hilbert space, that is, a vector $|\Psi\rangle$ which satisfies

$$\|\Psi\| = \sqrt{\langle\Psi|\Psi\rangle} = 1. \tag{3.3}$$

The state represents the configuration of the system, and encodes the possible outcomes of measurements performed on that system. It is important to stress that, although there are many vectors in a Hilbert space, only vectors which have norm equal to 1 represent states. However, if we have a vector with non-unit norm, we can simply divide it by its norm to obtain a unit vector, which then represents a state.

3.3 Hermitian Operators

An operator corresponding to a Hermitian matrix is a *Hermitian operator*. Above you have proved some interesting properties of these operators. In particular, their eigenvalues are real and their eigenvectors form an orthonormal basis.

In quantum theory, Hermitian operators correspond to *observables*, that is, properties of the system that can be measured. The eigenvalues of these operators correspond to the different possible outcomes of the measurement. This makes sense because we always measure real numbers; there are no measurement devices that measure complex numbers!

Examples of observables are position, momentum, angular momentum, energy and spin. All of these may be represented as Hermitian operators on an appropriate Hilbert space.

3.4 Superposition

Let the state of a quantum system be $|\Psi\rangle$. Once we have chosen a Hermitian operator to be our observable, there is a basis of states $|B_i\rangle$ corresponding to the eigenvectors of the operators. We may then write the vector $|\Psi\rangle$ as a linear combination of the basis vectors, as we did above:

$$|\Psi\rangle = \sum_{i=1}^{n} |B_i\rangle\langle B_i|\Psi\rangle. \tag{3.4}$$

Remember that $\langle B_i|\Psi\rangle$ is a scalar (a complex number). So this is a sum over the basis vectors $|B_i\rangle$ with a complex number attached to each of them. Such a linear combination of states is called a *superposition*.

3.5 Probability Amplitudes

In quantum theory, the inner product $\langle B_i|\Psi\rangle$ is called the *probability amplitude* to measure the eigenvalue λ_i corresponding to the eigenvector $|B_i\rangle$, given the state $|\Psi\rangle$. When we take the magnitude-squared of a probability amplitude, we get the corresponding probability. Thus

$$|\langle B_i|\Psi\rangle|^2 \tag{3.5}$$

is the probability to measure the eigenvalue λ_i corresponding to the eigenvector $|B_i\rangle$, given the state $|\Psi\rangle$.

Why is this a probability? It is easy to see that it behaves like a probability is expected to. It's obvious that $|\langle B_i|\Psi\rangle|^2$ is always non-negative, because it is a magnitude of a complex number. In addition, the sum of all possible probabilities for distinct outcomes has to equal one. Let us check that this is indeed the case.

Exercise: Verify that $\sum_{i=1}^{n} |\langle B_i|\Psi\rangle|^2 = 1$.

So the magnitudes-squared $|\langle B_i|\Psi\rangle|^2$ are non-negative numbers which sum to 1, and therefore may represent probabilities. Why they actually represent probabilities is a question that has no good answer except that this is how quantum theory works and it can be verified experimentally.

3.6 Qubits in the Computational Basis

Let us now introduce the simplest possible quantum system, the *qubit* (quantum bit). This is a system that can represent, for example, the spin of an electron, the polarization of a photon, and many other common real-life physical systems. However, we will not discuss the real-life applications, since we are more interested in studying the properties of the general mathematical framework and what we can learn from them.

The standard basis of $\mathbb{C}^2$, which in this case is called the *computational basis*, consists of the following two basis states:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.6}$$

Note that the numbers inside the kets $|\cdot\rangle$ are just labels, as always. Any state of this system is thus represented as a superposition of these two basis states:

$$|\Psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \tag{3.7}$$

where $a, b \in \mathbb{C}$ are the probability amplitudes:

$$a = \langle 0|\Psi\rangle, \qquad b = \langle 1|\Psi\rangle. \tag{3.8}$$

Exercise:

1. Show that a and b indeed take these values (note that we have proven this for a general basis above).
2. Show that the requirement that the state $|\Psi\rangle$ is normalized means that $|a|^2 + |b|^2 = 1$, that is, the probabilities sum to one (again, this follows from the general case we've discussed above).
3. Show that we can write the state $|\Psi\rangle$ explicitly as a vector in the Hilbert space:

$$|\Psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \langle 0|\Psi\rangle \\ \langle 1|\Psi\rangle \end{pmatrix}. \tag{3.9}$$

The fact that $|0\rangle$ and $|1\rangle$ are orthogonal hints that they are eigenstates of a Hermitian operator with different eigenvalues. Let us introduce the *Pauli matrices*:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.10}$$

It is easy to see that $|0\rangle$ and $|1\rangle$ are, in fact, eigenstates of σ_z, with eigenvalues $+1$ and -1 respectively.

Exercise:

1. Find the eigenstates and eigenvalues of σ_x and σ_y.
2. Prove that σ_x, σ_y and σ_z are Hermitian.
3. Prove that σ_x, σ_y and σ_z are also unitary.

3.7 Measuring Qubits

Since the Pauli matrices are Hermitian, they should correspond to an observable. This observable is called the *spin* of the qubit along each axis, x, y and z. Specifically for the z axis, the eigenstate $|0\rangle$ with eigenvalue $+1$ represents spin "up", that is, along the positive direction of the z axis, while the eigenstate $|1\rangle$ with eigenvalue -1 represents spin "down", that is, along the negative direction of the z axis.

It's important to note that a qubit does not actually have to represent physical spin around an axis! It can represent many different quantum systems, as long as they can be in two possible states. For example, we can use $|0\rangle$ to describe a state of no particles (a vacuum) and $|1\rangle$ to describe a state with exactly one particle (see the *quantum harmonic oscillator*, below). These two states create a perfectly good qubit system, even though they have nothing to do with spin.

We denote the two basis states of a qubit by the labels 0 and 1 because classical bits of information can take the values 0 or 1, and a qubit is a way to generalize classical bits using quantum theory. Qubits have many interesting uses in quantum information and computing, which we will not describe here.

Therefore, we conclude that given a state $|\Psi\rangle = a\,|0\rangle + b\,|1\rangle$, the probability to measure the value of the qubit as 0 is

$$|a|^2 = |\langle 0|\Psi\rangle|^2\,, \tag{3.11}$$

while the probability to measure the value of the qubit as 1 is

$$|b|^2 = |\langle 1|\Psi\rangle|^2\,. \tag{3.12}$$

As a simple example, consider the state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{3.13}$$

This state has equal probability of 50% to be either 0 or 1, since $\left|1/\sqrt{2}\right|^2 = 1/2$.

Exercise: Consider the vector

$$|\Psi\rangle = \begin{pmatrix} 3 \\ 4 \end{pmatrix}. \tag{3.14}$$

First normalize this vector so that it can represent a state, and then find out the probabilities to measure 0 and 1. Verify that the probabilities add up to 1.

3.8 Inner Products with Matrices and the Expectation Value

Consider a Hermitian operator A with an orthonormal bases of eigenstates $|B_i\rangle$ and eigenvalues λ_i. This means that

$$A\,|B_i\rangle = \lambda_i\,|B_i\rangle\,, \qquad \langle B_i|B_j\rangle = \delta_{ij}, \tag{3.15}$$

where δ_{ij} is the *Kronecker delta*, defined above:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{3.16}$$

Let $|\Psi\rangle$ be the state of the system. We may write $|\Psi\rangle$, as usual, as a superposition of the basis states:

$$|\Psi\rangle = \sum_{i=1}^{n} |B_i\rangle\langle B_i|\Psi\rangle. \tag{3.17}$$

The dual vector to $|\Psi\rangle$ is similarly written as:

$$\langle\Psi| = \sum_{i=1}^{n} \langle\Psi|B_i\rangle\langle B_i|. \tag{3.18}$$

Also note that

$$\begin{aligned}\langle B_i|A|B_j\rangle &= \langle B_i|\left(A|B_j\rangle\right)\\ &= \langle B_i|\lambda_j|B_j\rangle\\ &= \lambda_j\langle B_i|B_j\rangle\\ &= \lambda_j\delta_{ij}.\end{aligned}$$

Now we can calculate the following inner product:

$$\begin{aligned}\langle\Psi|A|\Psi\rangle &= \left(\sum_{i=1}^{n}\langle\Psi|B_i\rangle\langle B_i|\right) A \left(\sum_{j=1}^{n}|B_j\rangle\langle B_j|\Psi\rangle\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\langle\Psi|B_i\rangle\langle B_i|A|B_j\rangle\langle B_j|\Psi\rangle\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_j\delta_{ij}\langle\Psi|B_i\rangle\langle B_j|\Psi\rangle.\end{aligned}$$

When taking the sum over j, the δ_{ij} is always 0 except when $j = i$. Therefore the sum over j always gives us just one element, the one where $j = i$. We get:

$$\begin{aligned}\langle\Psi|A|\Psi\rangle &= \sum_{i=1}^{n}\lambda_i\langle\Psi|B_i\rangle\langle B_i|\Psi\rangle\\ &= \sum_{i=1}^{n}\lambda_i\left|\langle\Psi|B_i\rangle\right|^2.\end{aligned}$$

Since $|\langle\Psi|B_i\rangle|^2$ is the probability to measure the eigenvalue λ_i (associated with the eigenstate $|B_i\rangle$) given the state $|\Psi\rangle$, we have obtained, by definition, the expectation value for the measurement of A! For this reason, we sometimes simply write $\langle A\rangle$ instead of $\langle\Psi|A|\Psi\rangle$, as long as it is clear that the expected value is taken with respect to the state $|\Psi\rangle$. Sometimes we also use the notation

$$\langle A\rangle_\Psi = \langle\Psi|A|\Psi\rangle. \tag{3.19}$$

Exercise: Calculate $\langle\sigma_z\rangle_\Psi$ where σ_z is the Pauli matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.20}$$

for the following three states:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \tag{3.21}$$

$$|\Psi\rangle = \frac{1}{\sqrt{5}}|0\rangle + \frac{2}{\sqrt{5}}|1\rangle, \tag{3.22}$$

$$|\Psi\rangle = \frac{3}{\sqrt{13}}|0\rangle + \frac{2}{\sqrt{13}}|1\rangle. \tag{3.23}$$

4 The Lessons of Quantum Theory

4.1 Schrödinger's Cat and Interpretations of Quantum Mechanics

Suppose that, inside a box, there is a cat and a qubit in a superposition of 0 and 1:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle . \tag{4.1}$$

A measurement apparatus measures the qubit. If it measures 0 (with 50% probability), the cat dies (in a humane way). If it measures 1 (with 50% probability), the cat stays alive. The state of the cat is thus now a superposition of dead and alive:

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}|\text{dead}\rangle + \frac{1}{\sqrt{2}}|\text{alive}\rangle . \tag{4.2}$$

Before we open the box and measure the state of the cat, is it "actually" dead, or alive? Maybe the cat is both dead and alive at the same time? A qubit being in a superposition of 0 and 1 might not be intuitive (since it does not exist for a classical qubit, which can be either 0 or 1), but it is nevertheless an experimental fact. The thought of an animal being simultaneously dead and alive, on the other hand, seems absurd.

This thought experiment (not to be performed in real life!) was suggested by Schrödinger in the early days of quantum mechanics to illustrate this discrepancy between the quantum world (of elementary particles, atoms and qubits) and the classical world (of cats and everything else we know from our daily life).

Exercise: There are many *interpretations* of quantum mechanics. These interpretations are not scientifically testable, since there is no way to distinguish one from the other, and thus they are in the realm of philosophy, not science. For this reason, we will not talk about them in these lectures. However, you are encouraged to read about the different interpretations (the Wikipedia article "Interpretations of Quantum Mechanics" is a good place to start) and try to think what each of them means for the state of the poor cat before we open the box. What is your favorite interpretation? It is recommended to do this exercise with friends.

4.2 Entanglement of Two Qubits

Consider now a system of two qubits. Each of the qubits has two eigenstates, $|0\rangle$ and $|1\rangle$, as we have seen before. Let us name the first qubit A and the second qubit B. There are four possible states for the system of both qubits:

$$|0\rangle \otimes |0\rangle , \qquad |0\rangle \otimes |1\rangle , \qquad |1\rangle \otimes |0\rangle , \qquad |1\rangle \otimes |1\rangle , \tag{4.3}$$

where in each of these, the first state is the state of qubit A and the second is the state of qubit B. Thus $|0\rangle \otimes |0\rangle$ corresponds to $|0\rangle$ for both qubits, $|0\rangle \otimes |1\rangle$ corresponds to $|0\rangle$ for qubit A and $|1\rangle$ for qubit B, $|1\rangle \otimes |0\rangle$ corresponds to $|1\rangle$ for qubit A and $|0\rangle$ for qubit B, and $|1\rangle \otimes |1\rangle$ corresponds to $|1\rangle$ for both qubits.

The product $\otimes$ is called a *tensor product*. We will not go into its definition and properties; what's important is that it simply represents a way to combine the states of two separate systems into one system. Note that it is not commutative, since it matters which system is the first and which system is the second. It is, however, distributive.

The most general state of both qubits is described as a superposition of all possible combinations:

$$|\Psi\rangle = \alpha_{00}|0\rangle \otimes |0\rangle + \alpha_{01}|0\rangle \otimes |1\rangle + \alpha_{10}|1\rangle \otimes |0\rangle + \alpha_{11}|1\rangle \otimes |1\rangle , \tag{4.4}$$

where $\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11} \in \mathbb{C}$. We may now ask the question: when do the two qubits depend on each other? Another way to phrase this question is that we would like to know whether qubit A can be $|0\rangle$ or $|1\rangle$ independently of the state of qubit B, and vice versa. This depends on the coefficients α_{ij}, as we will now see.

Let us define a *separable state*: this is a state which can be written as just one tensor product instead of a sum of tensor products. That is, a separable state is of the form

$$|\Psi\rangle = |\Psi_A\rangle \otimes |\Psi_B\rangle\,, \tag{4.5}$$

where $|\Psi_A\rangle$ is the state of qubit A and $|\Psi_B\rangle$ is the state of qubit B. A simple example of a separable state would be:

$$|\Psi\rangle = |0\rangle \otimes |0\rangle\,. \tag{4.6}$$

This just means that both qubits are, with 100% probability, in the state $|0\rangle$. A more interesting separable state is:

$$|\Psi\rangle = \frac{1}{2}\left(|0\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\right). \tag{4.7}$$

To see that it is separable, all we need to do is simplify using the distributive property, and get:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right). \tag{4.8}$$

Exercise: Use the distributive property of the tensor product, that is,

$$|\Psi\rangle \otimes \left(|\Phi\rangle + |\Theta\rangle\right) = |\Psi\rangle \otimes |\Phi\rangle + |\Psi\rangle \otimes |\Theta\rangle\,, \tag{4.9}$$

$$\left(|\Psi\rangle + |\Phi\rangle\right) \otimes |\Theta\rangle = |\Psi\rangle \otimes |\Theta\rangle + |\Phi\rangle \otimes |\Theta\rangle\,, \tag{4.10}$$

to prove the last equation.

This means that both qubits are in the state where either 0 or 1 is possible with probability 50%, that is:

$$|\Psi_A\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \qquad |\Psi_B\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right). \tag{4.11}$$

(Remember that the $1/\sqrt{2}$ comes from the requirements that the state is normalized.) In particular, the states of the two qubits are completely independent and do not depend on each other.

Exercise: Find three more separable states of two qubits.

A state which is not separable is called an *entangled state*. Here is an example of an entangled state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle\right). \tag{4.12}$$

No matter how much we try, we can never write it as just one tensor product; it is always going to be the sum of two products! This means that the states of both qubits are no longer independent. Indeed, if qubit A is in the state $|0\rangle$ then qubit B must be in the state $|1\rangle$ (due to the first term), and if qubit A is in the state $|1\rangle$ then qubit B must be in the state $|0\rangle$ (due to the second term). This is precisely what is means for two systems to be entangled.

Exercise: Find three more entangled states.

4.3 Entanglement Does Not Transmit Information

Imagine the following scenario. Alice and Bob create an entangled pair of qubits, for example in the entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle\right). \tag{4.13}$$

Alice then stays on Earth, while Bob flies out to the Andromeda galaxy, about 2.5 million light years away. When Bob gets there, he measures the qubit. He has a 50% chance to observe 0 and a 50% chance to observe 1. However, if he observes 0 he knows that Alice will surely observe 1, and if he observes 1 he knows that Alice will surely observe 0, since the qubits must have opposite values.

So it seems that Bob now knows something about Alice's qubit that he did no know before. Furthermore, he knows that instantly, even though Alice is 2.5 million light years away and this means, according to relativity, that information will take at least 2.5 million years to travel between them. But has any information actually been transferred between them?

The answer is a very definite NO; all Bob did was observe a random event. Bob cannot control which value he observes when he measures the qubit, 0 or 1; he can only observe it, and get whatever he gets. He gains information about Alice's qubit, which is completely random, but he does not receive any specific message from Alice, nor can he transmit any specific information to Alice by observing his qubit.

In fact, there is a theorem called the *no-communication theorem* which rigorously proves that no information can be transmitted using quantum entanglement, whether faster than light or otherwise. Whatever you measure, it will be completely random. (The proof of this theorem uses more advanced tools we have not learned, so we will not present it here.)

The fact that a measurement of one qubit determines the measurement of another qubit might seem like it indicates that some information must be transmitted between the qubits themselves, so they "know" about each other's states. However, even if such information is transmitted internally between the entangled qubits, it cannot be used to transmit information between observers observing the qubits (due to the no-communication theorem), and thus does not violate relativity. It might violate the principle of *locality*, which will be discussed in the next section.

Furthermore, one may argue that there isn't any actual need to transmit information between the two entangled qubits in order for them to match their measurements. After all, the state $|\Psi\rangle$ does not depend on the distance between the qubits, whether in time or in space; it is simply the combined state of the two qubits, wherever or whenever they might be.

Exercise: Make up different scenarios in which you think you might be able to transmit information using quantum entanglement, and then figure out why your scenarios cannot possibly work, which must be the case due to the no-communication theorem. It is recommended to do this exercise with friends.

4.4 Bell's Inequality

Let's say I write 0 on one piece of paper and 1 on another piece of paper. I then put each piece of paper in a separate sealed envelope, and randomly give one envelope to Alice and the other to Bob. When Bob gets to Andromeda, he opens his envelope. If he sees 0 he knows that Alice's envelope says 1, and if he sees 1 he knows that Alice's envelope says 0.

Obviously, this does not allow any information to be transmitted between Alice and Bob, nor does each envelope need to "know" what's inside the other envelope in order for the measurements to match. If Bob sees 0, then the piece of paper saying 0 was inside the envelope all along, and the piece of paper saying 1 was inside Alice's envelope all along – and vice versa. The envelopes are classically *correlated* and nothing weird is going on. What, then, is the difference between this classical correlation and quantum entanglement? The answer to this question can be made precise using *Bell's inequality*.

Consider the following experiment. I prepare two qubits, and give one to Alice and another to Bob. Alice can measure one of two different physical properties of her qubit, Q or R, both having two possible outcomes, $+1$ or -1. Similarly, Bob can measure one of two different physical properties of his qubit, S or T, both having two possible outcomes, $+1$ or -1.

We now make two important assumptions:

- *Locality*: Both Alice and Bob measure their qubits at the same time in different places, so that their measurements cannot possibly disturb or influence each other without sending information faster than light.
- *Realism*: The values of the physical properties Q, R, S, T exist independently of observation, that is, they have certain definite values q, r, s, t which are already determined before any measurements took place (like in the envelope scenario).

Together, these two assumptions form the principle of *local realism*.

Now, whatever the values of q, r, s, t are, we must always have

$$rs + qs + rt - qt = (r + q)\, s + (r - q)\, t = \pm 2. \tag{4.14}$$

To see that, note that since $r = \pm 1$ and $q = \pm 1$, we must either have $r + q = 0$ if they have opposite signs, or $r - q = 0$ if they have the same sign. In the first case we have $(r - q)\, t = \pm 2$ because $t = \pm 1$ and in the second case we have $(r + q)\, s = \pm 2$ because $s = \pm 1$.

Using this information, we can calculate the mean value of this expression. To do that, we assign the probability $p\,(q, r, s, t)$ to each outcome of q, r, s, t. For example, we might simply assign a probability distribution where all probabilities are equal:

$$p\,(q, r, s, t) = \frac{1}{16}, \tag{4.15}$$

for any values of q, r, s, t. However, the probability distribution can be anything. Even though we don't know the probabilities in advance, we we can nonetheless still calculate the mean value:

$$\begin{aligned} \langle RS + QS + RT - QT \rangle &= \sum_{q,r,s,t \in \{-1,+1\}} p\,(q, r, s, t)\,(rs + qs + rt - qt) \\ &\leq 2 \sum_{q,r,s,t \in \{-1,+1\}} p\,(q, r, s, t) \\ &= 2. \end{aligned}$$

To go to the second line we used the fact that $rs + qs + rt - qt = \pm 2$ and thus it is always less than or equal to 2, and to go to the third line we used the fact that the sum of all possible probabilities is 1. Also, since the expected value function is linear, we have

$$\langle RS + QS + RT - QT \rangle = \langle RS \rangle + \langle QS \rangle + \langle RT \rangle - \langle QT \rangle\,. \tag{4.16}$$

Exercise: Prove this.

We thus obtain the (in)famous *Bell inequality*:

$$\langle RS \rangle + \langle QS \rangle + \langle RT \rangle - \langle QT \rangle \leq 2. \tag{4.17}$$

Now we are going to see that quantum entanglement violates this inequality. To see that, assume that I prepared the following entangled state of two qubits:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle \right). \tag{4.18}$$

Alice gets the first qubit in the tensor product, and Bob gets the second qubit. We define the observables Q, R, S, T in terms of the Pauli matrices. Alice has

$$Q = \sigma_z, \qquad R = \sigma_x, \tag{4.19}$$

while Bob has

$$S = -\frac{1}{\sqrt{2}}(\sigma_x + \sigma_z), \qquad T = -\frac{1}{\sqrt{2}}(\sigma_x - \sigma_z). \tag{4.20}$$

Exercise: Show, by explicit calculation, that

$$\langle RS\rangle = \langle QS\rangle = \langle RT\rangle = \frac{1}{\sqrt{2}}, \qquad \langle QT\rangle = -\frac{1}{\sqrt{2}}. \tag{4.21}$$

We thus get:

$$\langle RS\rangle + \langle QS\rangle + \langle RT\rangle - \langle QT\rangle = 2\sqrt{2} > 2, \tag{4.22}$$

which violates the Bell inequality. This means that our assumptions, either locality or realism (or both), must be incorrect!

Exercise:

1. Think about the consequences of letting go of each assumption. Which one would you rather lose, locality or realism?
2. The precise statement of Bell's theorem is that theories of *local hidden variables* cannot reproduce all the predictions of quantum mechanics. You are encouraged to look up theories of local hidden variables and read about them for a deeper understanding of the theorem.

4.5 Non-Commuting Observables

Let us define the *commutator* of two matrices or operators:

$$[A, B] = AB - BA. \tag{4.23}$$

If the operators commute, then $AB = BA$ and thus the commutator vanishes: $[A, B] = 0$. Otherwise, $AB \neq BA$ and the commutator is non-zero: $[A, B] \neq 0$. The commutator thus tells us if the operators commute or not. Note that any operator commutes with itself: $[A, A] = 0$ for any A.

Recall now the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.24}$$

The Pauli matrices do not commute with each other, as you will find in the exercise:

Exercise: Calculate the commutators $[\sigma_x, \sigma_y]$, $[\sigma_y, \sigma_z]$ and $[\sigma_z, \sigma_x]$.

4.6 The Uncertainty Principle

When two quantum observables do not commute, we get an *uncertainty relation*. The most well-known such relation is the position-momentum uncertainty relation[4]:

$$\Delta x \Delta p \geq \frac{1}{2}. \tag{4.25}$$

On the left-hand side we have the *uncertainty*, or more precisely the *standard deviation*, in position x, which is labeled Δx, multiplied by the uncertainty in momentum p, which is labeled Δp. This relation can be derived from the commutator of the operators x and p:

$$[x, p] = \mathrm{i}. \tag{4.26}$$

[4]Recall that we are using units where $\hbar = 1$!

Let us prove this relation for the general case of two Hermitian operators, A and B, which do not commute, that is,

$$[A, B] \neq 0. \tag{4.27}$$

Recall that the standard deviation ΔA of A is given by

$$(\Delta A)^2 = \left\langle (A - \langle A \rangle)^2 \right\rangle. \tag{4.28}$$

We have seen that expectation values in quantum theory are calculated using the inner product

$$\langle A \rangle = \langle \Psi | A | \Psi \rangle, \tag{4.29}$$

where $|\Psi\rangle$ is the state of the system with respect to which the expectation value is calculated. However, the uncertainty relation does not depend on the specific choice of state, so we will work with a general state $|\Psi\rangle$. The standard deviation is thus

$$\begin{aligned}(\Delta A)^2 &= \langle \Psi | \left(A - \langle A \rangle\right)^2 | \Psi \rangle \\ &= \langle \Psi | \left(A - \langle A \rangle\right) \left(A - \langle A \rangle\right) | \Psi \rangle.\end{aligned}$$

Let us now define a new vector:

$$|a\rangle = \left(A - \langle A \rangle\right) | \Psi \rangle. \tag{4.30}$$

Then we simply have

$$(\Delta A)^2 = \langle a | a \rangle = \| a \|^2. \tag{4.31}$$

Similarly, for B we define

$$|b\rangle = \left(B - \langle B \rangle\right) | \Psi \rangle, \tag{4.32}$$

and get

$$(\Delta B)^2 = \langle b | b \rangle = \| b \|^2. \tag{4.33}$$

The product of the standard deviations in A and B is thus

$$(\Delta A)^2 (\Delta B)^2 = \| a \|^2 \| b \|^2. \tag{4.34}$$

Using the Cauchy-Schwarz inequality (given in the linear algebra chapter), we have

$$\begin{aligned}(\Delta A)^2 (\Delta B)^2 &= \| a \|^2 \| b \|^2 \\ &\geq |\langle a | b \rangle|^2 \\ &= \left(\mathrm{Re}\langle a | b \rangle\right)^2 + \left(\mathrm{Im}\langle a | b \rangle\right)^2 \\ &\geq \left(\mathrm{Im}\langle a | b \rangle\right)^2 \\ &= \left(\frac{\langle a | b \rangle - \langle b | a \rangle}{2i} \right)^2 \\ &= -\frac{1}{4} \left(\langle a | b \rangle - \langle b | a \rangle \right)^2.\end{aligned}$$

Next, we note that

$$\begin{aligned}\langle a | b \rangle &= \langle \Psi | \left(A - \langle A \rangle\right) \left(B - \langle B \rangle\right) | \Psi \rangle \\ &= \langle \left(A - \langle A \rangle\right) \left(B - \langle B \rangle\right) \rangle \\ &= \langle AB - A\langle B \rangle - \langle A \rangle B + \langle A \rangle \langle B \rangle \rangle \\ &= \langle AB \rangle - \langle A \langle B \rangle \rangle - \langle \langle A \rangle B \rangle + \langle \langle A \rangle \langle B \rangle \rangle \\ &= \langle AB \rangle - \langle A \rangle \langle B \rangle - \langle A \rangle \langle B \rangle + \langle A \rangle \langle B \rangle \\ &= \langle AB \rangle - \langle A \rangle \langle B \rangle,\end{aligned}$$

and similarly

$$\langle b|a\rangle = \langle BA\rangle - \langle A\rangle\langle B\rangle. \tag{4.35}$$

Thus

$$\begin{aligned}\langle a|b\rangle - \langle b|a\rangle &= (\langle AB\rangle - \langle A\rangle\langle B\rangle) - (\langle BA\rangle - \langle A\rangle\langle B\rangle)\\ &= \langle AB\rangle - \langle BA\rangle\\ &= \langle [A,B]\rangle,\end{aligned}$$

and so we get

$$(\Delta A)^2 (\Delta B)^2 \geq -\frac{1}{4}\langle [A,B]\rangle^2 . \tag{4.36}$$

Finally, since $-1/4 = (1/2\,\mathrm{i})^2$, we can write

$$(\Delta A)^2 (\Delta B)^2 \geq \left\langle \frac{1}{2\,\mathrm{i}}\,[A,B]\right\rangle^2 . \tag{4.37}$$

Exercise:

1. Inequalities are only defined for real numbers (why?). Prove that $\langle [A,B]\rangle$ must always be an imaginary number, and thus $\left\langle \frac{1}{2\mathrm{i}}\,[A,B]\right\rangle$ is always real, so the inequality is well-defined.

2. For the specific case of the operators x and p with commutator

   $$[x,p] = \mathrm{i}, \tag{4.38}$$

   show that we get the familiar result

   $$\Delta x \Delta p \geq \frac{1}{2}. \tag{4.39}$$

3. Calculate the uncertainty relation for σ_x and σ_y given the most general state of a qubit:

   $$|\Psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \qquad |a|^2 + |b|^2 = 1, \tag{4.40}$$

   that is, find the right-hand side of

   $$\Delta\sigma_x \Delta\sigma_y \geq ? \tag{4.41}$$

   Comment on the consequences of the relation you found for choices of different states, that is, different values of a and b.

4.7 Dynamics: Discrete Time Evolution

So far, we have only talked about *kinematics*: what the states and operators are, but not how they evolve in time. Now we will say a few things about *dynamics*, which is how states evolve in time. This is the last ingredient we need to complete the full definition of the quantum theory framework.

We will be using the so-called *"Schrödinger picture"*, where states depend on time and operators do not. This is in contrast to the *"Heisenberg picture"*, which we will not discuss here, where states are independent of time and operators get the time dependence instead. Note that the two pictures are completely equivalent.

Let $|\Psi(t)\rangle$ be the state of the system at time t. Then the two states $|\Psi(t_1)\rangle$ at time t_1 and $|\Psi(t_2)\rangle$ at time t_2 must be related by some unitary operator U, which is generally a function of t_1 and t_2:

$$|\Psi(t_2)\rangle = U(t_1,t_2)\,|\Psi(t_1)\rangle\,. \tag{4.42}$$

As always, the exact form of $U(t_1, t_2)$ is determined by the specific quantum system. All quantum theory tells us is that it must be a unitary operator, just like an observable must be described by a Hermitian operator.

Exercise:

1. Recall that the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ are unitary. Take $U(t_1, t_2)$ to be each of these matrices, and describe how the general qubit state,

$$|\Psi\rangle = a|0\rangle + b|1\rangle, \qquad |a|^2 + |b|^2 = 1, \tag{4.43}$$

   evolves using each choice. Note that this is a *discrete* evolution, that is, there is no explicit dependence of the state on a continuous time variable t. We simply evolve from one state to another by applying the unitary operator.

2. Another unitary operator which acts on qubits is the *Hadamard gate*:

$$G = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{4.44}$$

   Repeat the previous exercise for this operator.

3. One might wonder why the evolution operator must be unitary. Show that given two states $|\Psi\rangle$ and $|\Phi\rangle$, the probability amplitude $\langle\Psi|\Phi\rangle$ remains the same even after we evolve the states to $U|\Psi\rangle$ and $U|\Phi\rangle$, where U is any unitary operator. Thus unitary operators *conserve probabilities*.

4.8 Dynamics: Continuous Time Evolution

Th equation we gave in the previous section relates two states at two different times by a unitary operator. However, in cases where the states depend on a continuous time variable t, we may write a *differential equation* giving the state $|\Psi\rangle$ for any time t. This equation is the *Schrödinger equation*[5]:

$$\frac{\mathrm{d}|\Psi(t)\rangle}{\mathrm{d}t} = -\mathrm{i}H|\Psi(t)\rangle. \tag{4.45}$$

It is called a differential equation because it involves derivatives. The equation simply tells us how $|\Psi(t)\rangle$ changes when t changes, in a precise way.

On the right-hand side, the operator H is a Hermitian operator called the *Hamiltonian*. This Hermitian operator corresponds to an observable: the *energy* of the system. The evolution of the state $|\Psi(t)\rangle$ is thus dictated solely by the action of the Hamiltonian operator H on the state.

Since the Hamiltonian is Hermitian, it has eigenvalues E_k corresponding to eigenstates $|E_k\rangle$ which make up an orthonormal basis[6]:

$$H|E_k\rangle = E_k|E_k\rangle. \tag{4.46}$$

The basis eigenstate $|E_k\rangle$ is simply a state in which the system has energy E_k, and is called an *energy eigenstate*. There will always be a state of lowest energy, that is, a state $|E_0\rangle$ for which the eigenvalue E_0 is the lowest among all the eigenvalues:

$$E_0 < E_k \text{ for all } k > 0. \tag{4.47}$$

Such a state is called the *ground state*.

[5] Again, recall that we are using units where $\hbar = 1$!

[6] Note that here, E_k is analogous to λ_i and $|E_k\rangle$ is analogous to $|B_i\rangle$ in the previous notation; we are using k instead of i as the index which enumerates the states.

Now, we are going to integrate the Schrödinger equation from time t_1 to time t_2, assuming that H does not explicitly depend on time. We will not go over the details of how to do this, since they are beyond the scope of our lectures. We will simply present the solution:

$$|\Psi_2\rangle = \mathrm{e}^{-\mathrm{i}\,H(t_2-t_1)}\,|\Psi_1\rangle\,. \tag{4.48}$$

The perceptive student will surely notice that this equation is of the form presented in the previous section, with the unitary operator

$$U\,(t_1,t_2) = \mathrm{e}^{-\mathrm{i}\,H(t_2-t_1)}\,. \tag{4.49}$$

Therefore, the Schrödinger equation is equivalent to the discrete time evolution equation presented above. Note that this also explains where the i in the Schrödinger equation comes from!

Exercise: Show that, if H is any Hermitian operator, then $U = \mathrm{e}^{\mathrm{i}\,H\alpha}$ is a unitary operator for any real number $\alpha \in \mathbb{R}$.

4.9 Summary: Quantum Theory and Models

We have seen that quantum theory is a fundamental mathematical framework for describing physical systems. The framework itself is not enough on its own; one must use the framework to define a *model*, which relates the theory to a specific physical system. Without a model, we are just doing math! A model is a specific choice of the following ingredients:

- A Hilbert space describing a specific physical system,
- Hermitian operators corresponding to specific physical observables that may be measured for the system,
- Unitary operators corresponding to time evolution of the system,
- Specific states on which these operators act, which correspond to different configurations of the system.

Of course, not every possible model we can make will actually correspond to a physical system that we can find in nature. However, amazingly, the opposite statement does seem to be true: every physical system that we find in nature can be precisely described by a model built using the ingredients of quantum theory!

We can think of quantum theory as a sort of language. Just like English is a language with rules such as grammar and spelling, so is quantum theory a language with its own rules: observables must be Hermitian operators, measurements are given by eigenvalues of these operators, and so on. And just like we can use English to make any sentence we want, both true and false, we can use quantum theory to make any model we want, both models that correspond to real physical systems and those that do not.

Days 3 & 4: Advanced Topics

We have presented the fundamental mathematical framework of quantum theory in terms of the simplest possible systems: those which can be described using the finite Hilbert space $\mathbb{C}^2$. That still turned out to be enough to discuss many interesting aspects of quantum theory, such as entanglement, Bell's theorem and the uncertainty principle, in full mathematical detail, and even perform explicit calculations.

Now we are going to discuss more complicated systems, which correspond to infinite-dimensional Hilbert spaces. Naturally, the math becomes much more sophisticated and beyond the required background of these lectures; thus, we will often merely present the material conceptually, without the full mathematical details. More details may be found in any quantum mechanics or quantum field theory textbook.

5 Infinite-Dimensional Hilbert Spaces and Quantization

5.1 Introduction

We have learned that quantum mechanics is a theory where a physical system is modeled using a Hilbert space with states and operators. One might ask: how do we know what Hilbert space, states and operators to use in the first place for a given physical system?

This is often a hard question to answer, and sometimes even impossible! For example, we currently do not know which Hilbert space, states and operators describe general relativity as a quantum theory. If we could only figure that out, then we would have essentially solved the problem of quantum gravity.

The simplest cases are where a Hilbert space is derived directly from an already known classical theory by a process known as *quantization*.

5.2 Hamiltonian Mechanics and Canonical Quantization

Classical mechanics can be reformulated using a quantity called the *Hamiltonian*. This is basically the total energy of the system, usually written as *kinetic energy* plus *potential energy* and in terms of the *canonical coordinates*, q and p, which in the simplest cases represent position and momentum respectively.

Since we have limited time, and we are interested in quantum mechanics and not classical mechanics, we will no go over the Hamiltonian formulation in detail. We will instead just note that the Hamiltonian is generally of the form

$$H = T(p) + V(q), \qquad T(p) = \frac{p^2}{2m}, \tag{5.1}$$

where T is the kinetic energy, V is the potential energy, and m is the mass of the particle. Using this expression we can then get the time evolution of the system, that is, the *derivatives* of q and p with respect to time t:

$$\frac{\mathrm{d}q}{\mathrm{d}t} = +\frac{\partial H}{\partial p}, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}. \tag{5.2}$$

These equations are called *Hamilton's equations*.

In *canonical quantization* of a classical system described by a Hamiltonian, we *promote* the variables q and p, as well as the Hamiltonian, to Hermitian operators. All three operators now represent observables in the quantum theory; they have eigenstates and eigenvalues, and these represent measurements. This means that the values of q, p and H are no longer uniquely determined from some initial conditions, as in the classical theory; they become *probabilistic*. In addition, the time evolution of the system is no longer described by Hamilton's equations, but rather, by the Schrödinger equation.

We will not perform canonical quantization in detail in these lecture notes, except for the quantum harmonic oscillator (see below). However, many examples can be found in any quantum mechanics textbook.

5.3 Wavefunctions and the Position Basis

Sometimes, especially in non-relativistic quantum mechanics, you will see a description in terms of *wavefunctions* instead of states. Recall that when quantizing a system, the position function x becomes a Hermitian operator. We usually denote this operator with a hat, $\hat{x}$, to distinguish it from x, which is just a number specifying the position of the particle, and is the eigenvalue of a corresponding eigenstate $|x\rangle$:

$$\hat{x}\,|x\rangle = x\,|x\rangle\,, \qquad x \in \mathbb{R}. \tag{5.3}$$

Since $\hat{x}$ is a Hermitian operator, the eigenstates $|x\rangle$ form an orthonormal basis[7]. The wavefunction ψ for a particle described by the state $|\Psi\rangle$ is given by:

$$\psi\,(x) = \langle x|\Psi\rangle, \tag{5.4}$$

where $|x\rangle$ is an eigenstate of position. Wavefunctions are, therefore, functions on space which return the probability amplitudes to find the particle in each point in space[8].

It should be noted that wavefunctions are not fundamental entities in modern quantum theory. The fundamental entities are the states. Wavefunctions only exist for particular systems where it is meaningful to define them, such as the quantized particle, and even in those cases we lose much of the rich mathematical toolbox of linear algebra, presented above, when we describe the system using wavefunctions instead of states. Furthermore, some quantum systems, such as the qubit, can only be described using states.

5.4 Phases and the Wave-Particle Duality

It is important to note that since the probability, as predicted by quantum mechanics, only depends only on the magnitude (or absolute value) of the probability amplitude, we can multiply the amplitude by any complex number with magnitude 1, and the probabilities will remain the same. For example, the following states all have probability 50% to measure either 0 or 1:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,, \tag{5.5}$$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{\mathrm{i}}{\sqrt{2}}\,|1\rangle\,, \tag{5.6}$$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,|0\rangle - \frac{\mathrm{i}}{\sqrt{2}}\,|1\rangle\,, \tag{5.7}$$

$$|\Psi\rangle = \frac{\mathrm{i}}{\sqrt{2}}\,|0\rangle + \frac{1+\mathrm{i}}{2}\,|1\rangle\,. \tag{5.8}$$

This means that there is something "extra" in quantum mechanics, more than just probability; the probability amplitude is the fundamental quantity, and the probability is just a consequence of it!

Exercise: Write down four different states which all have a probability of $1/3$ to measure 0 and $2/3$ to measure 1. Remember that the states must be normalized to 1.

The phase of the amplitude is irrelevant if we have just one probability amplitude, but if we add up several probability amplitudes and only then take the magnitude-squared, the different amplitudes will *interfere* with each other. For example, if we have two amplitudes $a = 1/\sqrt{2}$ and $b = 1/\sqrt{2}$, then

$$|a|^2 + |b|^2 = 1, \qquad \text{and } |a+b|^2 = 1. \tag{5.9}$$

[7] Or, more precisely, the infinite-dimensional equivalent of an orthonormal basis, which we will not discuss here.

[8] Equivalently, wavefunctions may be defined on momentum basis: $\psi\,(p) = \langle p|\Psi\rangle$.

The amplitudes are then said to have *interfered constructively*. However, if we have $a = 1/\sqrt{2}$ and $b = -1/\sqrt{2}$, then

$$|a|^2 + |b|^2 = 1, \qquad \text{but } |a+b|^2 = 0! \tag{5.10}$$

The amplitudes are then said to have *interfered destructively*. If this reminds you of constructive and destructive interference of waves, that is not a coincidence; this is exactly the source of the so-called *wave-particle duality*!

This interference effect is responsible, for example, for the interference pattern in the *double-slit experiment*[9]. If both slits are open, then the particle's state is a superposition of passing through slit A and passing through slit B, or, schematically,

$$|\Psi\rangle = a\,|\Psi_A\rangle + b\,|\Psi_B\rangle\,, \qquad |a|^2 + |b|^2\,. \tag{5.11}$$

The probability amplitude to measure the particle at a specific position x_0 is, of course,

$$\langle x_0|\Psi\rangle = a\langle x_0|\Psi_A\rangle + b\langle x_0|\Psi_B\rangle. \tag{5.12}$$

The probability is then, as usual, the magnitude squared:

$$\begin{aligned} |\langle x_0|\Psi\rangle|^2 &= |a\langle x_0|\Psi_A\rangle + b\langle x_0|\Psi_B\rangle|^2 \\ &= |a|^2\,|\langle x_0|\Psi_A\rangle|^2 + |b|^2\,|\langle x_0|\Psi_B\rangle|^2 + 2\,\mathrm{Im}\left(a^* b\langle\Psi_A|x_0\rangle\langle x_0|\Psi_B\rangle\right), \end{aligned}$$

where Im stands for "the imaginary part of, that is, if $a, b \in \mathbb{R}$ then $\mathrm{Im}\,(a + b\,\mathrm{i}) = b$.

Exercise: Derive explicitly the expression given above for $|\langle x_0|\Psi\rangle|^2$.

The terms $|a|^2\,|\langle x_0|\Psi_A\rangle|^2$ and $|b|^2\,|\langle x_0|\Psi_B\rangle|^2$ are always positive. However, the third term is a real number which can be either positive or negative, depending on the specific values of a, b, and the phases of the probability amplitudes. This term will either increase or decrease the probability to find the particle at x_0, and it is precisely what is responsible for the interference pattern in the double-slit experiment.

We have thus seen that wave-particle duality is not really that mysterious: it's simply the consequence of the particle in quantum mechanics having a probability amplitude to be in every possible position, instead of just one unique position as in classical mechanics!

5.5 The Quantum Harmonic Oscillator

The *quantum harmonic oscillator* is a particular case of a quantum system which turns out to describe many different physical systems, and in particular, is essential to quantum field theory. We start with the Hamiltonian for a classical non-relativistic massive particle which has a particular potential energy:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{5.13}$$

Here, m is again the mass of the particle, and ω is some numerical constant (sometimes called the *frequency* or *angular frequency*). We quantize the system by promoting x, p and H to operators. We are interested in finding the energy eigenstates of this quantized system. Instead of finding them by solving a complicated differential equation, we will use an easier (and more intuitive) method.

To find the energy eigenstates, we define the *ladder operators*:

$$a = \sqrt{\frac{m\omega}{2}}\left(x + \frac{\mathrm{i}}{m\omega}p\right), \qquad a^\dagger = \sqrt{\frac{m\omega}{2}}\left(x - \frac{\mathrm{i}}{m\omega}p\right), \tag{5.14}$$

where $a^\dagger$ is called the *creation operator* and a is called the *annihilation operator*.

Exercise:

[9] This is a very important experiment in quantum mechanics. If you're not familiar with it, look it up!

1. Show that $a^\dagger$ is, indeed, the adjoint of a.

2. Show that these definitions may be inverted to get

$$x = \sqrt{\frac{1}{2m\omega}}\left(a^\dagger + a\right), \qquad p = \mathrm{i}\sqrt{\frac{m\omega}{2}}\left(a^\dagger - a\right). \tag{5.15}$$

3. Show, using $[x, p] = \mathrm{i}$, that the Hamiltonian can now be expressed as

$$H = \omega\left(a^\dagger a + \frac{1}{2}\right). \tag{5.16}$$

The last result of the exercise is very important: the form of the Hamiltonian has been simplified considerably! This is what allows us to solve for the energy eigenstates of the system (that is, the eigenstates of H) easily. We define a new operator called the *number operator*:

$$N = a^\dagger a. \tag{5.17}$$

Now the Hamiltonian may be written as

$$H = \omega\left(N + \frac{1}{2}\right). \tag{5.18}$$

Since both ω and $1/2$ are just numbers, the problem of finding the eigenvalues of H now reduces to finding the eigenvalues of N.

Exercise:

1. Show that if n is an eigenvalue of N then $\omega\left(n + \frac{1}{2}\right)$ is an eigenvalue of H.

2. Show that N is Hermitian.

3. Calculate, using $[x, p] = \mathrm{i}$, the following commutators:

$$[a, a^\dagger] = 1, \qquad [N, a^\dagger] = a^\dagger, \qquad [N, a] = -a. \tag{5.19}$$

Now, let $|n\rangle$ be an eigenstate of N with eigenvalue n. Since N is Hermitian, we know that n must be a real number. In fact, we can do more than that. Let us calculate the expectation value:

$$\langle N\rangle_n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = \||a|n\rangle\|^2, \tag{5.20}$$

where we used the fact that $\langle n|\, a^\dagger$ is the dual vector to $a\,|n\rangle$. On the other hand, we have

$$\langle N\rangle_n = \langle n|N|n\rangle = n\langle n|n\rangle = n, \tag{5.21}$$

where we used the fact that n is the eigenvalue of N corresponding to the eigenstate $|n\rangle$, that is, $N\,|n\rangle = n\,|n\rangle$, and that the state $|n\rangle$ is normalized to 1, like all states, so $\langle n|n\rangle = 1$. By comparing the two equations, we see that

$$n = \||a|n\rangle\|^2 \geq 0, \tag{5.22}$$

that is, n is not only real but non-negative.

Next, we act with Na and $Na^\dagger$ on $|n\rangle$. In the exercise you showed that

$$Na - aN = [N, a] = -a, \tag{5.23}$$

$$Na^\dagger - a^\dagger N = [N, a^\dagger] = a^\dagger, \tag{5.24}$$

so we have

$$Na = aN - a = a\,(N-1)\,, \qquad Na^\dagger = a^\dagger N + a^\dagger = a^\dagger\,(N+1)\,, \tag{5.25}$$

and thus

$$Na\,|n\rangle = a\,(N-1)\,|n\rangle = (n-1)\,a\,|n\rangle\,, \tag{5.26}$$

$$Na^\dagger\,|n\rangle = a^\dagger\,(N+1)\,|n\rangle = (n+1)\,a^\dagger\,|n\rangle\,, \tag{5.27}$$

where we used the fact that $N\,|n\rangle = n\,|n\rangle$ and that since $n \pm 1$ is a number, it commutes with operators and can be moved to the left. Writing this result in a different way, we see that

$$N(a\,|n\rangle) = (n-1)\,(a\,|n\rangle), \tag{5.28}$$

$$N(a^\dagger\,|n\rangle) = (n+1)\,(a^\dagger\,|n\rangle), \tag{5.29}$$

or in other words, $a\,|n\rangle$ is an eigenstate of N with eigenvalue $n-1$, and $a^\dagger\,|n\rangle$ is an eigenstate of N with eigenvalue $n+1$! However, by definition, the normalized eigenstates of N with eigenvalues $n-1$ and $n+1$ are $|n-1\rangle$ and $|n+1\rangle$ respectively. Thus, we conclude that $a\,|n\rangle$ is proportional to $|n-1\rangle$ and $a^\dagger\,|n\rangle$ is proportional to $|n+1\rangle$. The proportionality factors must be chosen so that the states are normalized. Let us therefore calculate the norms. The norm $\|a\,|n\rangle\|^2$ was already calculated above:

$$\|a\,|n\rangle\|^2 = \langle n|a^\dagger a|n\rangle = \langle n|N|n\rangle = n. \tag{5.30}$$

To calculate $\|a^\dagger\,|n\rangle\,\|^2$, we recall from the exercise that

$$aa^\dagger - a^\dagger a = [a, a^\dagger] = 1, \tag{5.31}$$

and thus

$$aa^\dagger = a^\dagger a + 1 = N + 1. \tag{5.32}$$

We therefore get

$$\|a^\dagger\,|n\rangle\,\|^2 = \langle n|aa^\dagger|n\rangle = \langle n|\,(N+1)\,|n\rangle = \langle n|N|n\rangle + \langle n|n\rangle = n+1. \tag{5.33}$$

To summarize, the norms are

$$\|a\,|n\rangle\| = \sqrt{n}, \qquad \|a^\dagger\,|n\rangle\,\| = \sqrt{n+1}. \tag{5.34}$$

The normalized eigenstates are now obtained, as usual, by dividing by the norm:

$$|n-1\rangle = \frac{1}{\sqrt{n}}a\,|n\rangle\,, \qquad |n+1\rangle = \frac{1}{\sqrt{n+1}}a^\dagger\,|n\rangle\,. \tag{5.35}$$

Another way to write this, from a different point of view, is as the action of the operators a and $a^\dagger$ on the state $|n\rangle$:

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \qquad a^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{5.36}$$

We see that a reduces the energy eigenvalue by 1, while $a^\dagger$ increases the energy eigenvalue by 1. A more fancy way to describe this is by saying that $a^\dagger$ gets us to the state of next higher energy (it "creates a quantum of energy") while a gets us to the state of next lower energy (it "annihilates a quantum of energy"). For this reason, we call $a^\dagger$ the *creation operator* and a the *annihilation operator*. We also call them the *ladder operators* because they let us "climb the ladder" of energy eigenstates.

Going back to the definition of the Hamiltonian in terms of the number operator, we see that

$$H\,|n\rangle = \omega\left(n + \frac{1}{2}\right)|n\rangle\,, \tag{5.37}$$

and thus $|n\rangle$ is an energy eigenstate with eigenvalue

$$E_n = \omega\left(n + \frac{1}{2}\right). \tag{5.38}$$

In particular, since we showed above that n must be non-negative, and since we now also see that it has to be an integer (as it can only be increased or decreased by 1), the possible eigenstates are found to be

$$|0\rangle, |1\rangle, |2\rangle, |3\rangle, \ldots. \tag{5.39}$$

The state of lowest energy is $|0\rangle$, also called the *ground state*, which has energy

$$E_0 = \frac{1}{2}\omega. \tag{5.40}$$

We say that $a^\dagger$, which takes us from $|0\rangle$ to $|1\rangle$, *excites* the harmonic oscillator from the ground state to the *first excited state*, which has exactly one *quantum*. Yes, this is what the "quantum" in "quantum mechanics" means! In general, the state $|n\rangle$ has exactly n quanta, while the ground state $|0\rangle$ has no quanta.

We found that the energy of the harmonic oscillator is discrete, or *quantized*, and the system can only have energy which differs from $\omega/2$ by equal steps of ω.

Exercise: Prove that

$$|n\rangle = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}}|0\rangle. \tag{5.41}$$

This means that, once we know the ground state, we can create any energy eigenstate by simply applying n times the operator $a^\dagger$ and normalizing.

5.6 The Path Integral Formulation

So far, we have been working exclusively with one way to define quantum mechanics, using Hilbert spaces, states and operators. There is another way, which is equivalent but very different, which is using *path integrals*.

The path integral formulation of quantum mechanics tells us that, in order to find the probability amplitude for the particle to go from the point x_1 at time t_1 to the point x_2 at time t_2, we need to take into account all the different paths that the particle can take between these points. We will thus sum (or *integrate*) over all the paths, and each path will get a certain weight in this sum, given by a *phase factor*. The equation describing this process is:

$$\langle x_2, t_2 | x_1, t_1\rangle = \int_{x(t_1)=x_1}^{x(t_2)=x_2} \mathcal{D}x\,(t)\ \Phi\left(x\,(t)\right), \tag{5.42}$$

where:

- $\langle x_2, t_2 | x_1, t_1\rangle$ is the probability amplitude for the particle to go from the point x_1 at time t_1 to the point x_2 at time t_2.
- $\int_{x(t_1)=x_1}^{x(t_2)=x_2} \mathcal{D}x\,(t)$ means "sum over all the paths $x\,(t)$ such that $x\,(t_1) = x_1$ and $x\,(t_2) = x_2$". In other words, sum over all the paths the particle can take from the point x_1 at time t_1 to the point x_2 at time t_2.
- $\Phi\left(x\,(t)\right)$ is the phase factor, which is a complex number with magnitude 1 assigned to each path $x\,(t)$.

When we sum over the phase factors for all the different paths, they will interfere constructively and destructively, as discussed previously. Thus, we see that important results of quantum theory such as the wave-particle duality follow naturally from this approach, without ever needing to define a Hilbert space, or assign states and operators to various aspects of the system. Everything is defined using only the paths $x(t)$ and the phase factor $\Phi(x(t))$, which is calculated using tools from classical mechanics.

5.7 Quantum Field Theory

We will not go too much into the details of quantum field theory here, especially not the mathematical details, which are much beyond the level of our lectures; in fact, they are usually only taught in graduate school since you need an entire undergraduate degree in math and physics to understand them! We will just provide some basic information and concepts.

Quantum field theory is, roughly speaking, a way to construct models using quantum theory which are also compatible with special relativity. It's not really a different theory from quantum theory; it still has all the basic elements, namely a Hilbert space with states and operators. The difference is in the ingredients used to define the models; these ingredients are chosen carefully to ensure compatibility with special relativity. (Yes, this only applies to special relativity; we still do not know how to combine quantum theory with general relativity, which would yield a theory of *quantum gravity*.)

In the models we have seen of non-relativistic quantum mechanics, there is an important difference between space and time. The position x of a particle is given by a Hermitian operator, whose eigenvalues are the possible positions. In contrast, the time t is just a parameter along which the system evolves using a unitary operator. Time is just a *label*: we have one state at time t_1 and another state at time t_2, they are related by a unitary operator, and that's all we have to say about the role of time.

This contradicts what is perhaps the most important lesson of special relativity: that we do not live in 3-dimesional space, we live in 4-dimensional *spacetime*, and time is just another dimension, on equal footing as the dimensions of space. In a relativistic theory, we still evolve systems in time, so in the relativistic quantum theory, we will still relate states at different times by a unitary operator; but we must treat space the same way we treat time.

There are two ways to do that. One is to promote time to an operator, just like we did for the position operator. This works to some extent, but turns out not to be very convenient as a mathematical framework. The other option is to demote position from an operator to a mere label of states, just like time is a label of the state $|\Psi(t)\rangle$. The states of the system will now have, as a label, not only time but also position, and there is no longer a position operator (and thus also no momentum operator, since the two go hand in hand).

Roughly speaking, the way it works mathematically is by placing a quantum harmonic oscillator at each point in space. Then we define creation operators $a^\dagger(x)$ and annihilation operators $a(x)$ at each point x. These operators create or destroy quanta, which we call *particles*. The *vacuum state* $|0\rangle$ is simply the state with no particles anywhere. Acting with $a^\dagger(x)$ on the vacuum state creates one particle at position x:

$$a^\dagger(x)|0\rangle = |1_x\rangle. \tag{5.43}$$

Acting with $a^\dagger(y)$ on this state creates another particle, this time at position y:

$$a^\dagger(y)|1_x\rangle = |1_x 1_y\rangle. \tag{5.44}$$

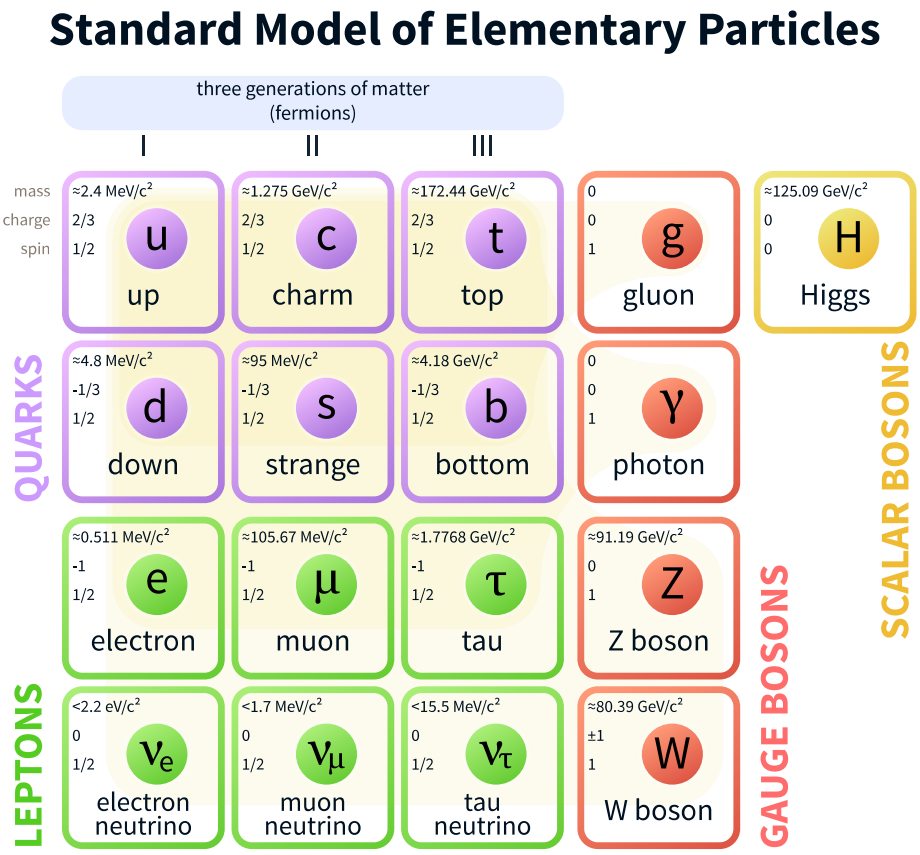
And so on. The states with one or more particles are called *excited states*. In this way, we can build excited states with arbitrary numbers of particles at arbitrary points in space. The actual state of the system is going to be, as usual, a superposition of all the possible states – that is, all the possible numbers of particles at all the possible positions!

This is called a *field* because it is exactly what you get when you quantize a classical theory describing a classical field, such as the *electromagnetic field*. We will not go over the details here. A nice way to visualize it is to think about a "field" of harmonic oscillators spread throughout space, each having its own "ladder" of eigenstates corresponding to different numbers of quanta.

5.8 The Standard Model

The *Standard Model of particle physics* is a particular model built using quantum field theory. This model is the most successful model in all of science, and has been tested experimentally to extremely high precision. The Large Hadron Collider, the most powerful particle accelerator ever built (located at CERN, Switzerland), was intended to both test the Standard Model and also find so-called "new physics" beyond the Standard Model. So far, all it has done is provide further verification for the Standard Model to much higher precision than any previous particle accelerator. This verification included the detection of the Higgs boson in 2012, which is an excited state of the Higgs field and a crucial part of the Standard Model.

The Standard Model contains the following fields (source: Wikipedia):



Each of these fields can be excited to obtain a particle. For example, by exciting the electron field using the creation operator at a particular point, we get an electron at that point. The mass of the field is simply the amount of energy needed to put into the field in order to excite it. So it is much easier to create an electron, with mass 0.5 MeV, than a Higgs boson, with mass 125,000 MeV. Mega-electron-

volt or MeV is[10] a unit of mass roughly equal to 1.8×10^{-30} kg, and $1\,\text{GeV} = 1,000\,\text{MeV}$.

There are many fields in this table, but they can be divided into three basic types:

- *Scalar fields:* They have spin 0 and are represented mathematically simply as some complex number ϕ. The only scalar in the Standard Model is the *Higgs field*, which gives the other fields mass through the *Higgs mechanism*.
- *Spinor fields:* They have spin $1/2$ and are represented mathematically as a vector in $\mathbb{C}^2$, that is, a vector (ψ_1, ψ_2) with two complex components $\psi_1, \psi_2 \in \mathbb{C}$. Note that this is an abstract vector, not a vector in spacetime itself! All particles of matter are described using spinors. This includes the quarks u, d, c, s, t, b and the leptons $e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau$. The leptons come in three generations, each more massive than the previous one (and thus harder to create).
- *Vector fields:* Also known as *gauge fields*, they have spin 1 and are represented mathematically as a vector in $\mathbb{R}^4$, that is, a vector (t, x, y, z) in spacetime itself. Gauge fields are used to mediate interactions between the particles. The photon mediates the *electromagnetic interaction*, the gluon mediates the *strong interaction*, and the W and Z bosons mediate the *weak interaction*.

Fields with integer spin (0 or 1 in the Standard Model) are called *bosons* while fields with half-integer spins ($1/2$ in the Standard Model) are called *fermions*. The meaning of spin can be interpreted as follows: if a field has spin S then it means if we rotate space around it, its mathematical representation will come back to its original configuration after $1/S$ full rotations. A scalar field is just a number, so it always stays the same; a vector field is just like a usual vector in space, so it comes back to itself after one full rotation; and a spinor field is a bit weird because it gains a minus sign after a full rotation, which means that it goes back to its original state only after two full rotations.

6 Quantum Gravity

As we have alluded above, the mathematical framework of quantum theory has been successfully applied to most of known physics, with the notable exception of gravity. *General relativity* describes gravity using the notion of *spacetime curvature*, and it is our best and most precise theory of gravity. In the terms we have used so far, quantizing gravity would mean finding a quantum system, that is, a Hilbert space with appropriate states and operators, which reduces to general relativity when quantum effects are neglected.

In naive attempts to quantize gravity, one tries to approximate curved spacetime as a flat spacetime with tiny "bumps" of curvature. These "bumps" are called *gravitons*. Quantizing them then leads to a *low-energy effective theory*, which means that the theory is only valid as long as we are dealing with large distance scales (or equivalently, low energy scales). For small distances (or high energies), which are what we are actually interested in when formulating a quantum theory, the graviton theory breaks down and can no longer produce any predictions. Thus, it is not useful as a scientific theory.

There are several speculative approaches to solving the problem of quantizing gravity. Some approaches, like *string theory*, hypothesize a completely new quantum theory, with new fundamental entities (strings, in this case), which produces general relativity in the appropriate limit. Other approaches, such as *loop quantum gravity*, start with general relativity itself, with a fully curved spacetime, and try to quantize it like any other quantum field theory. However, at this point all of these approaches are purely hypothetical, with no complete theory available and definitely no experimental verification.

In order to formulate a complete and mathematically consistent theory of quantum gravity, fresh ideas that have not been tried before are needed. Perhaps the reader of these lecture notes will contribute to this endeavor?

[10] The table uses units of MeV$/c^2$, but in our Planck units, where $c = 1$, this is the same as MeV.