Quantum Mechanics

Antonin Krpensky

Department of Physics
Faculty of Electrical Engineering
Czech Technical University in Prague
Technicka 2, Prague, Czech Republic

antonin.krpensky[at]fel.cvut.cz

These lecture notes are intended as study material for the course BV002TF2.

Version: 2025-3-26 (Draft)

Contents

Chapter 1

Hilbert Spaces and Dirac's Notation

1.1 Hilbert Space & Inner Product

Let us start by briefly defining the Hilbert space (usually denoted as $\mathscr{H}$), which plays a key role in the world of quantum mechanics. Basically, the Hilbert space is just a vector space as we know it from the course of Linear Algebra with one additional operation – the *inner product* – that we will inspect later.

First, the elements of the Hilbert space $\mathscr{H}$ are still called vectors. But here comes the first big difference in the notation. Most of you are probably familiar with vectors being denoted as something like $\boldsymbol{v}$, or $\overrightarrow{v}$. In quantum mechanics, however, we use the so called *Dirac's notation* (Paul dirac, 1937–1984) that uses the following symbol for a vector

$$|\psi\rangle \in \mathscr{H}\,. \tag{1.1}$$

To make things even weirder, we are going to call such object as a *Ket-vector* (I will explain why later in the text). As usual, for Ket-vectors, we define the two operations:

1. *addition* : $\mathscr{H} \times \mathscr{H} \to \mathscr{H}$ (this rather fancy formula states that the operation takes two Ket-vectors in $\mathscr{H}$ and produces another Ket-vector in $\mathscr{H}$)

$$|\psi\rangle + |\phi\rangle \in \mathscr{H}\,, \tag{1.2}$$

2. *multiplication by scalar* : $\mathbb{C} \times \mathscr{H} \to \mathscr{H}$, where by scalar in this context we mean a complex number

$$\alpha\,|\psi\rangle \in \mathscr{H}\,, \quad \alpha \in \mathbb{C}\,. \tag{1.3}$$

We can, of course, combine the two preceding operations to express a general *linear combination* as

$$\alpha\,|\psi\rangle + \beta\,|\phi\rangle + \cdots \in \mathscr{H}\,. \tag{1.4}$$

Note that the output of both the operations (or their combination) results always in another Ket-vector in $\mathscr{H}$. We say that $\mathscr{H}$ is *closed* under addition and scalar multiplication.

Up to this point, everything should seem quite normal (except from the weird notation, perhaps). But now comes the key difference between our Hilbert space $\mathscr{H}$ and the common vector space – the Hilbert space includes one additional operation called the *inner product* : $\mathscr{H} \times \mathscr{H} \to \mathbb{C}$, denoted as

$$\langle \phi | \psi \rangle \in \mathbb{C} \,. \tag{1.5}$$

Note that the result of this operation is not another object in $\mathscr{H}$ but rather a scalar (complex number). The inner product is then defined by the following properties

1. $\langle \psi | \psi \rangle \geq 0$,
2. $\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^*$ (* denotes the complex conjugation),
3. $\langle \phi | \psi + \xi \rangle = \langle \phi | \psi \rangle + \langle \phi | \xi \rangle$,
4. $\langle \phi | \alpha \psi \rangle = \alpha \, \langle \phi | \psi \rangle$, $\alpha \in \mathbb{C}$.

Before we move on, note the confusion in the Dirac's notation here. First, regarding the third property, the $|\psi + \xi\rangle$ term stands for $|\psi\rangle + |\xi\rangle$. Second, when we defined the multiplication operation by Eq. (1.3) we used $\alpha \, |f\rangle$ and now when dealing with inner product we write $|\alpha f\rangle$. Unfortunately, the Dirac's notation is sometimes a bit inconsistent. The best thing you can do here is just to get used to it. We will encounter a similar confusion soon when talking about operators.

The first property ensures that the norm of a Ket-vector defined as

$$\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle} \tag{1.6}$$

is always a non-negative real number while $\langle \psi | \psi \rangle = 0$ implies $|\psi\rangle = 0$. This is another ambiguity in the Dirac's notation - the 0 here stands for the zero Ket-vector, while $|0\rangle$ will usually have different meaning. Like I said, just get used to it.

By combining the second property with the third and fourth, one can then easily show that the following holds

$$\langle \psi + \xi | \phi \rangle = \langle \psi | \phi \rangle + \langle \xi | \phi \rangle \;, \quad \langle \alpha \psi | \phi \rangle = \alpha^* \, \langle \psi | \phi \rangle \;. \tag{1.7}$$

Now is the appropriate time to explain why do we uste the term Ket-vectors. According to the above definition, the inner products is an operation that takes two Ket-vectors as an input and produces a complex number at the output. But there is a different way to

look at this. We can define the so called Hermitian adjoint to the Ket-vector $|\phi\rangle$ denoted as

$$|\phi\rangle^\dagger \equiv \langle\phi| \tag{1.8}$$

which now represents a different object living in a different vector space (formally it is called a dual space to our Hilbert space and has a formal mathematical definition which I am not going to bother with here). Such an object is then called a *Bra-vector* and will be treated as follows: whenever a Bra-vector meets a Ket-vector (in this exact ordering), together they form an inner product as

$$\boxed{\langle\phi|\,|\psi\rangle \equiv \langle\phi|\psi\rangle}\,. \tag{1.9}$$

As you can see: Bra meets Ket $\rightarrow$ BraKet $\rightarrow$ *bracket*, which symbolises the inner product notation.

Up to this point everything was defined rather formally. Now let's concretize things a little bit.

1.2 Bases and Representations

The first time you encountered vectors was probably in high school where you were told that vector is something like an arrow which has a direction and length (this reminds me of the animated movie "Despicable Me"). Now look at Fig. 1.1 depicting a 2D vector $\boldsymbol{v}$ (for now, ignore the red axes). The teacher at the high school would then write

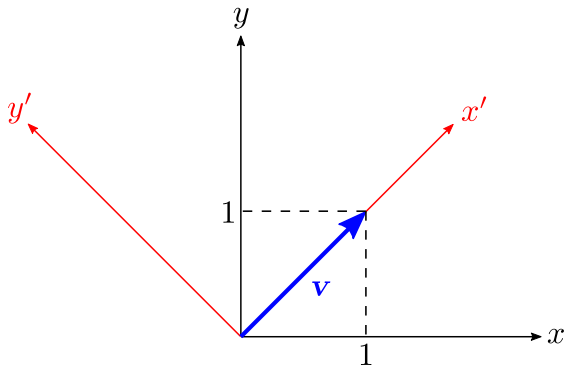

Figure 1.1: 2D vector

something like

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\,. \tag{1.10}$$

After a while you visited your first course of linear algebra and (hopefully) someone explained to you that this is not correct. A vector is not a column of numbers. This is

only the representation of the vector in a specifically chosen *basis* - a set of n linearly independent vectors where n is the dimension of the corresponding vector space. The correct way to express $\boldsymbol{v}$ would be

$$\boldsymbol{v} = 1\boldsymbol{e}_1 + 1\boldsymbol{e}_2\,, \tag{1.11}$$

where $\boldsymbol{e}_i$ stand for the basis vectors - unit vectors pointing along the x and y axes in this case. One can then represent the vector as a column of numbers but only if the basis is known a priori. If someone would then take a different basis such as $\boldsymbol{e}'_i$ (corresponding to the red x' and y' axes in Fig. 1.1), the vector $\boldsymbol{v}$ would be expressed as

$$\boldsymbol{v} = \sqrt{2}\boldsymbol{e}'_1 + 0\boldsymbol{e}'_2\,. \tag{1.12}$$

The corresponding column representation would then be

$$\boldsymbol{v} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}\,, \tag{1.13}$$

which is something completely different than the expression (1.10), even though it represents the same object (the vector $\boldsymbol{v}$)! Hence, we can (and will do) use the representations but at the same time we need to be sure to understand the difference between a vector and it's representation, otherwise QM would not make any sense to us.

Now let's get back to our favourite Hilbert space and define an appropriate basis. We must divide the text into two subsections based on the choice of either discrete or continuous basis. We start with the discrete case.

1.2.1 Discrete Basis

By a discrete basis we mean that the basis vectors can be indexed by a discrete index n (typically an integer) such as $|n\rangle$. Further, we will require the following condition to be met:

$$\boxed{\langle m|n\rangle = \delta_{mn}}\,, \tag{1.14}$$

where δ_{mn} stands for the Kronecker's delta defined as

$$\delta_{mn} = \begin{cases} 1\,, & m = n \\ 0\,, & \text{otherwise} \end{cases}\,. \tag{1.15}$$

A basis with this property is called *orthonormal* (it is trivial to show that $\||\psi\rangle\| = 1$). Any Ket-vector can now be expressed as an appropriate linear combination of such basis Ket-vectors

$$|\psi\rangle = \sum_n \psi_n\,|n\rangle\,, \tag{1.16}$$

where ψ_n are called *coefficients of the expansion*. We can now represent $|\psi\rangle$ as a column (of complex numbers) in the following way:

$$|\psi\rangle \rightarrow \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix} , \tag{1.17}$$

which will be further referred to as the *discrete basis representation* of $|\psi\rangle$. Now pause for a minute and note a few things. First, I used "$\rightarrow$" instead of "$=$" to stress out that this is only the representation of $|\psi\rangle$ in the respective basis $\{|n\rangle\}$ and not $|\psi\rangle$ itself. Second, the first component of the column definitely does not have to start with the index 1, it is just a convention to write it this way. Third, N can go (and usually will go) to infinity, so the column will have infinitely (countably) many components.

At this point just one question remains - how to determine the respective coefficients of the expansion? To answer this we proceed as follows:

$$\langle n|\psi\rangle = \langle n| \left(\sum_{n'} \psi_{n'} \left|n'\right\rangle \right) = \sum_{n'} \psi_{n'} \left\langle n\middle|n'\right\rangle = \sum_{n'} \psi_{n'} \delta_{n,n'} = \psi_n \, . \tag{1.18}$$

By comparing the LHS with the RHS we get

$$\psi_n = \langle n|\psi\rangle \ . \tag{1.19}$$

Finally, we can write down the complete formula

$$\boxed{|\psi\rangle = \sum_n \psi_n \, |n\rangle \quad \text{where} \quad \psi_n = \langle n|\psi\rangle}, \tag{1.20}$$

which we will frequently use later in the the upcoming sections.

As the Ket-vectors live in $\mathscr{H}$, we should inspect of how to add two Ket-vectors together and further how to multiply a Ket-vector by a scalar. First, we deal with the addition by writing

$$|\psi\rangle + |\phi\rangle = \sum_n \psi_n \, |n\rangle + \sum_n \phi_n \, |n\rangle = \sum_n (\psi_n + \phi_n) \, |n\rangle \ . \tag{1.21}$$

Hence, to add $|\psi\rangle + |\phi\rangle$ one simply adds the respective components $\psi_n + \phi_n$ as usual. To multiply a Ket-vector by a scalar we write

$$\alpha \, |\psi\rangle = \alpha \sum_n \psi_n \, |n\rangle = \sum_n \alpha \psi_n \, |n\rangle \ , \tag{1.22}$$

so basically we just multiply each component ψ_n by α.

Moving on, we define the corresponding Bra-vector as

$$\boxed{\langle\psi| = \sum_n \psi_n^* \langle n|}, \tag{1.23}$$

To see why, let's express $\langle\phi|\psi\rangle$ as

$$\begin{aligned}\langle\phi|\psi\rangle = \langle\phi|\,|\psi\rangle = \left(\sum_n \phi_n^* \langle n|\right)\left(\sum_{n'} \psi_{n'} \left|n'\right\rangle\right) &= \sum_{n,n'} \phi_n^* \psi_{n'} \left\langle n\middle|n'\right\rangle \\ &= \sum_{n,n'} \phi_n^* \psi_{n'} \delta_{nn'} = \sum_n \phi_n^* \psi_n\,, \end{aligned} \tag{1.24}$$

where we have used the property that once a Bra-vector meets a Ket-vector they form an inner product. Since this result is very important, let's put it in a box

$$\boxed{\langle\phi|\psi\rangle = \sum_n \phi_n^* \psi_n}. \tag{1.25}$$

It should be stressed out that this simple expression holds only due to the orthonormal property of the basis! It is pretty straightforward to show that such an expression for the inner product meets exactly the properties as defined above in the previous section, which justifies the definition (1.23). E.g., the norm of the Ket-vector is given by

$$\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle} = \sqrt{\sum_n \psi_n^* \psi_n} = \sqrt{\sum_n |\psi_n|^2}\,, \tag{1.26}$$

which is a perfectly good (and standardly used) norm. To prove the remaining properties is let up to the reader as an exercise.

Based on Eqs. (1.23) and (1.25), one can represent the respective Bra-vector as a row

$$\langle\phi| \to \begin{pmatrix} \phi_1^* & \cdots & \phi_N^* \end{pmatrix}, \tag{1.27}$$

which enables us to express the inner product only by the representations of the respective vectors as

$$\langle\phi|\psi\rangle = \begin{pmatrix} \phi_1^* & \cdots & \phi_N^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix} = \sum_n \phi_n^* \psi_n\,, \tag{1.28}$$

where the standard matrix multiplication rule is used. Note that in the above expression, there is no need to use "$\to$" anymore since these are truly equal.

1.2.2 An Aside - Dirac Distribution

Perhaps before doing the continuous case, it might be useful to define the Dirac delta function (distribution, to be mathematically correct), since it plays a very important role in QM and we are going to use it a lot in this chapter. Let's start by defining the following function

$$\delta_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon}\,, & \text{if } |x| \leq \frac{\varepsilon}{2} \\ 0\,, & \text{otherwise} \end{cases} . \tag{1.29}$$

The function is plotted in Fig. 1.2. Now there is one important property to this function:

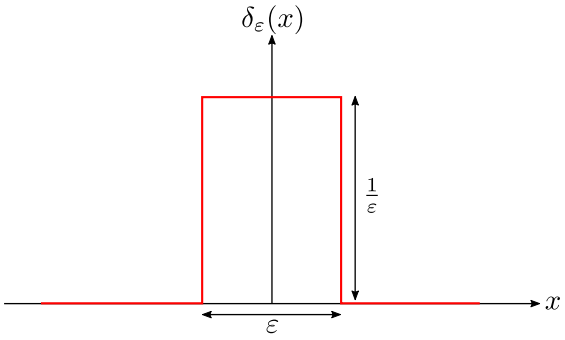

Figure 1.2: Dirac distribution

no matter what the value of ε is, the area under the curve is always equal to unity, i.e.,

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \delta_\varepsilon(x) = 1\,. \tag{1.30}$$

Now the Dirac distribution is simply defined as

$$\delta(x) = \lim_{\varepsilon \to 0} \delta_\varepsilon(x)\,. \tag{1.31}$$

Please note that this is not a correct mathematical way to define such object, but it is sufficient for the purpose of this text. Now based on that definition we can write

$$\delta(x) = \begin{cases} \infty\,, & \text{if } x = 0 \\ 0\,, & \text{otherwise} \end{cases} . \tag{1.32}$$

but the infinity value at $x = 0$ is a very specific infinity such that the following condition still holds

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x) = 1\,. \tag{1.33}$$

Strictly speaking, such object we have just defined is not a function - a function cannot have a nonzero value at only one point and at the same time nonzero area under the

curve - and therefore it is called the *Dirac distribution*.

Before moving on, there is one additional important property of the Dirac distribution. To present it we perform the following calculation

$$\int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta(x-x_0) = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x_0)\delta(x-x_0) = f(x_0)\int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x-x_0) = f(x_0)\,, \quad (1.34)$$

where in the first equality we have used the fact that the Dirac distribution is nonzero only at $x = x_0$, which enables us to replace $f(x)$ with just $f(x_0)$, and the last equality just follows from the fact that the Dirac distribution has a unity are under the curve. As this is an important property, let's put it in the box

$$\boxed{\int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta(x-x_0) = f(x_0)}\,. \quad (1.35)$$

This property is sometimes called the *sampling property* of the Dirac distribution as it takes the whole function $f(x)$ as an input and produces just one "sample" at $x = x_0$, i.e., $f(x_0)$.

1.2.3 Continuous Basis

In the continuous case, the basis Ket-vectors can not be denoted by a discrete index n anymore but rather by a continuous inded x (typically a real number). Please note that at this points x has nothing to do with position along the x axis! Hence, we will use the notation $|x\rangle$. As in the previous case, we will require the basis to be orthonormal, which can be written as

$$\boxed{\langle x|x'\rangle = \delta(x-x')}\,, \quad (1.36)$$

where $\delta(x)$ stands for the Dirac distribution and plays essentialy the same role as the Kronecker delta in the discrete case. Now to translate from the discrete case to the continuous one, we simply replace all the sums with the corresponding integrals as

$$\sum_n \quad \to \quad \int \mathrm{d}x\ . \quad (1.37)$$

Even though not explicitly stated, the integral here runs over all possible values of x. The Ket-vector $|\psi\rangle$ can thereby be expressed in the basis $\{|x\rangle\}$ as

$$\boxed{|\psi\rangle = \int \mathrm{d}x\ \psi(x)\,|x\rangle}\,. \quad (1.38)$$

Note that when dealing with QM, we usually write the differential $\mathrm{d}x$ right after the integral symbol as opposed to what you have learned in the introductory courses of

Calculus where the differential is written at the end of the expression that is being integrated. The function $\psi(x)$ should now be thought of a continuous set of expansion coefficients w.r.t. a continuous base, similarly as ψ_n in the discrete case. Since it is not possible to express all the values of a function into a column, the representation of $|\psi\rangle$ in the basis $\{|x\rangle\}$ will be just

$$|\psi\rangle \to \psi(x)\,, \tag{1.39}$$

which will be further referred to as the *continuous basis representation* of a Ket-vector (actually, this is what we are later going to call the wavefunction).

By calculating

$$\langle x|\psi\rangle = \langle x| \int \mathrm{d}x'\, \psi(x') \left|x'\right\rangle = \int \mathrm{d}x'\, \psi(x') \left\langle x\middle|x'\right\rangle = \psi(x) \int \mathrm{d}x'\, \delta(x-x') = \psi(x) \tag{1.40}$$

and comparing the LHS with the RHS we get

$$\psi(x) = \langle x|\psi\rangle\ . \tag{1.41}$$

Therefore, we can write down the general formula for expressing a Ket-vector in terms of a continuous basis as

$$\boxed{|\psi\rangle = \int \mathrm{d}x\, \psi(x)\, |x\rangle \quad \text{where} \quad \psi(x) = \langle x|\psi\rangle}\,. \tag{1.42}$$

Moving on, the Bra-vector in the continuous case is defined as

$$\boxed{\langle\phi| = \int \mathrm{d}x\, \phi^*(x)\, \langle x|}\,. \tag{1.43}$$

We can now derive the expression for an inner product in the case of a continuous basis. To do this we express the inner product $\langle\phi|\psi\rangle$ once again as

$$\begin{aligned}\langle\phi|\,|\psi\rangle = \left(\int \mathrm{d}x\, \phi^*(x)\, \langle x|\right) \left(\int \mathrm{d}x'\, \psi(x') \left|x'\right\rangle\right) &= \int \mathrm{d}x\, \mathrm{d}x'\, \phi^*(x)\psi(x') \left\langle x\middle|x'\right\rangle \\ &= \int \mathrm{d}x\, \mathrm{d}x'\, \phi^*(x)\psi(x')\delta(x-x') = \int \mathrm{d}x\, \phi^*(x)\psi(x)\,. \end{aligned} \tag{1.44}$$

Since this expression will be important later, we put it in a box

$$\boxed{\langle\phi|\psi\rangle = \int \mathrm{d}x\, \phi^*(x)\psi(x)}\,. \tag{1.45}$$

You can check once again that the above expression meets the four properties of the inner product defined in the previous section and thus the definition (1.43) is justified.

Things get just a little bit more complicated when looking for the representation of

a Bra-vector in a continuous basis. Based on Eq. (1.45), we should probably write something like

$$\langle\phi| \to \int \mathrm{d}x\, \phi^*(x) \times \,, \tag{1.46}$$

where $\times$ denotes that this is an operation that needs another complex-valued function as an input such as $\psi(x)$ before being evaluated. Now you may ask why do we not need any sum or the "$\times$" symbol when dealing with the Bra representation in the discreet case. It is simply because the sum is implied when multiplying a row with a column (in this order) by the rules of the matrix multiplication, which is not the case here.

1.3 Operators

In this section, we introduce the objects called *operators* which play one of the key roles in the context of QM. So, what exactly is an operator? An operator on a Hilbert space $\mathscr{H}$ is something similar as a function on real numbers $\mathbb{R}$. A function f takes a real number $x \in \mathbb{R}$ as an input and produces another real number $f(x) \in \mathbb{R}$ as an output. In analogy, an operator $\hat{\mathbf{A}}$ is an object that takes a Ket-vector $|\psi\rangle \in \mathscr{H}$ as an input and produces another Ket-vector $|\phi\rangle$ belonging to the same Hilbert space $\mathscr{H}$ as an output. A mathematician would describe such object as

$$\hat{\mathbf{A}} : \mathscr{H} \to \mathscr{H} \,. \tag{1.47}$$

For pedagogical reasons (since this is an introductory text), all operators will be strictly denoted as bold characters with a hat (more advanced textbooks usually omit the notation and it is left up to you to decipher what is an operator and what is just an ordinary variable).

Now as you were warn above, there is another inconsistency in the Dirac's notation. The most logical way of how to write that $\hat{\mathbf{A}}$ takes $|\psi\rangle$ as an input - hereinafter referred to as the operator $\hat{\mathbf{A}}$ *acting on* $|\psi\rangle$ - and produces $|\phi\rangle$ is

$$\hat{\mathbf{A}}\,|\psi\rangle = |\phi\rangle \,. \tag{1.48}$$

However, you will see shortly that sometimes (even though it does not make much sense) it can be useful to denote it as

$$\left|\hat{\mathbf{A}}\psi\right\rangle = |\phi\rangle \,. \tag{1.49}$$

1.3.1 Linear Operators on $\mathscr{H}$

In this text we will deal solely with the *linear* operators, so this short paragraph is devoted to the respective definition. Basically, any linear operator $\hat{\mathbf{A}}$ on $\mathscr{H}$ meets the following two conditions:

1. linearity in vector addition

$$\hat{\mathbf{A}}(|\psi\rangle + |\phi\rangle) = \hat{\mathbf{A}}\,|\psi\rangle + \hat{\mathbf{A}}\,|\phi\rangle \,,$$

2. linearity in scalar multiplication

$$\hat{\mathbf{A}}(\alpha\,|\psi\rangle) = \alpha\hat{\mathbf{A}}\,|\psi\rangle\ ,$$

where $\alpha \in \mathbb{C}$.

Both of those properties can then be formulated together as

$$\hat{\mathbf{A}}(\alpha\,|\psi\rangle + \beta\,|\phi\rangle) = \alpha\hat{\mathbf{A}}\,|\psi\rangle + \beta\hat{\mathbf{A}}\,|\phi\rangle\ , \tag{1.50}$$

where $\alpha, \beta \in \mathbb{C}$. If this reminds you of the linear transformations defined in the course of the linear algebra then you are absolutely correct. Actually, we will see shortly that in the discrete basis representation all linear operators can be represented by matrices.

1.3.2 Operators - Discrete Basis Expansion & Representation

Similarly as in the previous section where we showed how a Ket-vector can be expressed by terms of basis Ket-vectors, here we will do the same with linear operators (it only works for the linear ones) in the discrete case. Further, we will show that any linear operator can be represented by a matrix.

First, we define the *unit operator* which we are going to use the most during this section. The definition is simple - the unit operator $\hat{\mathbf{1}}$ acting on any Ket-vector $|\psi\rangle$ leaves it unchanged. Or mathematically speaking

$$\hat{\mathbf{1}}\,|\psi\rangle = |\psi\rangle\ . \tag{1.51}$$

Now let's look on how we can express the unit operator in terms of our discrete basis $\{|n\rangle\}$. For such purpose, we use Eq. (1.20) to write

$$|\psi\rangle = \sum_n \psi_n\,|n\rangle = \sum_n \langle n|\psi\rangle\,|n\rangle\ . \tag{1.52}$$

To find the formula for the unit operator, we do the following:

$$|\psi\rangle = \sum_n \langle n|\psi\rangle\,|n\rangle = \sum_n |n\rangle\,\langle n|\psi\rangle = \left(\sum_n |n\rangle\,\langle n|\right)|\psi\rangle\ . \tag{1.53}$$

Now by comparing the LHS with the RHS and employing the definition of the unit operator, we can identify

$$\hat{\mathbf{1}} = \sum_n |n\rangle\,\langle n|\ . \tag{1.54}$$

As a next step, we use the unit operator to express any (linear) operator $\hat{\mathbf{A}}$ by terms of the basis Ket-vectors $\{|n\rangle\}$. We proceed as follows:

$$\hat{\mathbf{A}} = \hat{\mathbf{1}}\hat{\mathbf{A}}\hat{\mathbf{1}} = \sum_k |k\rangle\,\langle k|\,\hat{\mathbf{A}}\sum_l |l\rangle\,\langle l| = \sum_{k,l}\langle k|\,\hat{\mathbf{A}}\,|l\rangle\,|k\rangle\,\langle l| = \sum_{k,l} A_{kl}\,|k\rangle\,\langle l|\ , \tag{1.55}$$

where we have denoted $A_{kl} \equiv \langle k| \hat{\mathbf{A}} |l\rangle$. Since this is a very important formula, we put it in a box

$$\boxed{\hat{\mathbf{A}} = \sum_{k,l} A_{kl} |k\rangle \langle l| \quad \text{with} \quad A_{kl} \equiv \langle k| \hat{\mathbf{A}} |l\rangle} . \tag{1.56}$$

There is a little subtlety in (1.55) that I should explain, as it might seem that this would work for any operator. But actually, we have secretly made use of the linearity property of $\hat{\mathbf{A}}$ (1.50). Let's inspect this more carefully. If we write

$$\hat{\mathbf{A}} |\psi\rangle = \hat{\mathbf{A}} \left(\hat{\mathbf{1}} |\psi\rangle \right) = \hat{\mathbf{A}} \left(\sum_n |n\rangle \langle n|\psi\rangle \right) , \tag{1.57}$$

then the only way to be able to shift $\hat{\mathbf{A}}$ inside the sum as

$$\hat{\mathbf{A}} \left(\sum_n |n\rangle \langle n|\psi\rangle \right) = \sum_n \hat{\mathbf{A}} |n\rangle \langle n|\psi\rangle \tag{1.58}$$

is to assume $\hat{\mathbf{A}}$ linear, which is exactly what we have done in (1.55).

Next, we show how does the operator $\hat{\mathbf{A}}$ act on a Ket-vector $|\psi\rangle$. To do so, we write

$$\begin{aligned} \hat{\mathbf{A}} |\psi\rangle = \hat{\mathbf{A}} \sum_n \psi_n |n\rangle &= \sum_{k,l} A_{kl} |k\rangle \langle l| \sum_n \psi_n |n\rangle \\ &= \sum_{k,l,n} A_{kl} \psi_n \langle l|n\rangle |k\rangle = \sum_{k,l,n} A_{kl} \psi_n \delta_{ln} |k\rangle = \sum_{k,l} A_{kl} \psi_l |k\rangle , \end{aligned} \tag{1.59}$$

where in the first equality the expansion of $|\psi\rangle$ into the basis $\{|n\rangle\}$ was used and further in the second one the above derived expression (1.56) for $\hat{\mathbf{A}}$ was employed. Due to it's importance, the result of the above calculation is rewritten here in a compact form

$$\boxed{\hat{\mathbf{A}} |\psi\rangle = \sum_{k,l} A_{kl} \psi_l |k\rangle} . \tag{1.60}$$

If we now denote

$$|\phi\rangle = \hat{\mathbf{A}} |\psi\rangle \tag{1.61}$$

and further expand $|\phi\rangle$ as

$$|\phi\rangle = \sum_k \phi_k |k\rangle , \tag{1.62}$$

then by comparing the expressions (1.62) and (1.60) we get the following relation between the expansion coefficients

$$\phi_k = \sum_l A_{kl} \psi_l , \tag{1.63}$$

which you should recognise as the matrix multiplication formula. Hence, the discrete basis representation of an operator $\hat{\mathbf{A}}$ is the following matrix

$$\hat{\mathbf{A}} \to \mathbb{A} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots \\ A_{2,1} & A_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} , \tag{1.64}$$

where once again the indexing starting here from 1 is just a conventional choice.

As an exercise, let us find a discrete basis representation of a unit operator. By employing Eq. (1.56) we find

$$1_{kl} = \langle k| \, \hat{\mathbf{1}} \, |l\rangle = \langle k|l\rangle = \delta_{kl} \, , \tag{1.65}$$

which makes sence, because any column vector multiplied by a unit matrix (from left) results in the same unit vector. (Note that in order to distinguish between the number 1 and the element of the matrix representation of $\hat{\mathbf{1}}$ I used the italics font.) Hence, the sought representation for the unit operator is

$$\hat{\mathbf{1}} \to \mathrm{diag}(1, 1, \dots) \, , \tag{1.66}$$

where "diag" stands for the diagonal matrix.

1.3.3 Operators - Continuous Basis Expansion & Representation

Following the previous subsection, here we deal with the continuous basis expansion of operators together with the corresponding representations. Let's start once again with the unit operator which (in analogy) can be expressed as

$$\boxed{\hat{\mathbf{1}} = \int \mathrm{d}x \; |x\rangle \, \langle x|} \, . \tag{1.67}$$

We can simply check this by calculating

$$\begin{aligned} \hat{\mathbf{1}} \, |\psi\rangle = \int \mathrm{d}x \; |x\rangle \, \langle x| \int \mathrm{d}x' \, \psi(x') \, |x'\rangle = \int \mathrm{d}x \, \mathrm{d}x' \; \langle x|x'\rangle \, \psi(x') \, |x\rangle \\ = \int \mathrm{d}x \, \mathrm{d}x' \, \delta(x - x') \psi(x') \, |x\rangle = \int \mathrm{d}x \, \psi(x) \, |x\rangle = |\psi\rangle \; . \end{aligned} \tag{1.68}$$

Now we focus on expressing an arbitrary operator $\hat{\mathbf{A}}$ by terms of the continuous basis $\{|x\rangle\}$. To do so we follow exactly the same procedure as in the previous case, so

$$\begin{aligned} \hat{\mathbf{A}} = \hat{\mathbf{1}}\hat{\mathbf{A}}\hat{\mathbf{1}} = \int \mathrm{d}x \; |x\rangle \, \langle x| \, \hat{\mathbf{A}} \int \mathrm{d}y \; |y\rangle \, \langle y| \\ = \int \mathrm{d}x \, \mathrm{d}y \; \langle x| \, \hat{\mathbf{A}} \, |y\rangle \, |x\rangle \, \langle y| = \int \mathrm{d}x \, \mathrm{d}y \, A(x, y) \, |x\rangle \, \langle y| \; , \end{aligned} \tag{1.69}$$

where we have denoted $A(x,y) \equiv \langle x| \hat{\mathbf{A}} |y\rangle$. (Once again, we assumed the linearity property of $\hat{\mathbf{A}}$ in the above calculation.) Note that instead of x' we used y here for the second integration variable. Of course this has absolutely no effect other than aesthetic. Due to it's importance, we put the result in a box

$$\boxed{\hat{\mathbf{A}} = \int \mathrm{d}x\, \mathrm{d}y\, A(x,y) \,|x\rangle \langle y| \quad \text{with} \quad A(x,y) \equiv \langle x| \hat{\mathbf{A}} |y\rangle}\,. \tag{1.70}$$

It is now pretty straightforward to show how does $\hat{\mathbf{A}}$ act on $|\psi\rangle$. So as with the discreet case,

$$\begin{aligned}\hat{\mathbf{A}} |\psi\rangle &= \int \mathrm{d}x\, \mathrm{d}y\, A(x,y) \,|x\rangle \langle y| \int \mathrm{d}z\, \psi(z) \,|z\rangle = \int \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z\, A(x,y)\psi(z) \langle y|z\rangle \,|x\rangle \\ &= \int \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z\, A(x,y)\psi(z)\delta(y-z) \,|x\rangle = \int \mathrm{d}x\, \mathrm{d}y\, A(x,y)\psi(y) \,|x\rangle\ . \end{aligned} \tag{1.71}$$

Hence,

$$\boxed{\hat{\mathbf{A}} |\psi\rangle = \int \mathrm{d}x\, \mathrm{d}y\, A(x,y)\psi(y) \,|x\rangle}\,. \tag{1.72}$$

If we further denote

$$|\phi\rangle = \hat{\mathbf{A}} |\psi\rangle \tag{1.73}$$

and expand $|\phi\rangle$ in the respective basis as

$$|\phi\rangle = \int \mathrm{d}x\, \phi(x) \,|x\rangle\ , \tag{1.74}$$

then by comparing with Eq. (1.72) we get the following relationship between $\phi(x)$ and $\psi(x)$:

$$\phi(x) = \int \mathrm{d}y\, A(x,y)\psi(y)\,, \tag{1.75}$$

which can be thought of as a continuous variant of the matrix multiplication. Thus, the representation of the operator $\hat{\mathbf{A}}$ in the continuous basis can be written as

$$\hat{\mathbf{A}} \to \int \mathrm{d}y\, A(x,y)\times\,, \tag{1.76}$$

where similarly as in the case of a continuous representation of $|\psi\rangle$ the symbol $\times$ indicates an operation which first takes another complex valued function $\psi(y)$ and only then evaluates the integral.

1.4 Commutator

Now that we know how does an operator act on a Ket-vector, it is time to investigate what happens when we act with two different operators on the same Ket-vector. We

will work this out only by using the discrete basis. Hence, assume

$$\hat{\mathbf{A}} = \sum_{k,l} A_{kl} \left|k\right\rangle \left\langle l\right| \quad \text{and} \quad \hat{\mathbf{B}} = \sum_{kl} B_{kl} \left|k\right\rangle \left\langle l\right| . \tag{1.77}$$

The question now arises - does the order of the operators matter? I.e., does $\hat{\mathbf{A}}\hat{\mathbf{B}} \left|\psi\right\rangle$ produce the same result as $\hat{\mathbf{B}}\hat{\mathbf{A}} \left|\psi\right\rangle$? The answer is simple: no. To show why, we compute

$$\hat{\mathbf{A}}\hat{\mathbf{B}} = \left(\sum_{k,l} A_{kl} \left|k\right\rangle \left\langle l\right| \right) \left(\sum_{m,n} B_{mn} \left|m\right\rangle \left\langle n\right| \right) = \sum_{k,n} \left(\sum_{m} A_{km} B_{mn} \right) \left|k\right\rangle \left\langle n\right| . \tag{1.78}$$

But the term in brackets can be expressed as

$$\sum_{m} A_{km} B_{mn} = (\mathbb{A}\mathbb{B})_{kn} , \tag{1.79}$$

where $\mathbb{A}, \mathbb{B}$ stand for the matrices representing the respective operators. But as we know, the matrix multiplication is not commutative, i.e.,

$$\sum_{k,l} (\mathbb{A}\mathbb{B})_{kl} \left|k\right\rangle \left\langle l\right| \neq \sum_{k,l} (\mathbb{B}\mathbb{A})_{kl} \left|k\right\rangle \left\langle l\right| , \tag{1.80}$$

and hence

$$\hat{\mathbf{A}}\hat{\mathbf{B}} \neq \hat{\mathbf{B}}\hat{\mathbf{A}} . \tag{1.81}$$

(In the continuous case the proof is similar.) We say that the operators generally do not *commute*. To determine whether two operators commute or not we define the *commutator* as

$$\boxed{\left[\hat{\mathbf{A}}, \hat{\mathbf{B}}\right] = \hat{\mathbf{A}}\hat{\mathbf{B}} - \hat{\mathbf{B}}\hat{\mathbf{A}}} . \tag{1.82}$$

The two operators then commute iff the corresponding commutator is equal to zero. We will see later that the commutator plays an important role in the context of QM.

There are three properties of commutator that will be stated without proof - that is left for the reader as an exercise:

1. Anticommutativity
$$\left[\hat{\mathbf{A}}, \hat{\mathbf{B}}\right] = - \left[\hat{\mathbf{B}}, \hat{\mathbf{A}}\right] ,$$

2. (Bi)linearity
$$\left[\alpha\hat{\mathbf{A}} + \beta\hat{\mathbf{B}}, \hat{\mathbf{C}}\right] = \alpha \left[\hat{\mathbf{A}}, \hat{\mathbf{C}}\right] + \beta \left[\hat{\mathbf{B}}, \hat{\mathbf{C}}\right] ,$$
where $\alpha, \beta \in \mathbb{C}$ and the prefix "Bi" meaning that this holds also for the second argument,

3. Jacobi identity

$$\left[\hat{\mathbf{A}}, \left[\hat{\mathbf{B}}, \hat{\mathbf{C}}\right]\right] + \left[\hat{\mathbf{B}}, \left[\hat{\mathbf{C}}, \hat{\mathbf{A}}\right]\right] + \left[\hat{\mathbf{C}}, \left[\hat{\mathbf{A}}, \hat{\mathbf{B}}\right]\right] = 0\,.$$

We then say that commutator forms the so called *Lie Algebra*. Finally, there is one additional property that might come in handy when doing algebraic operations with commutators which I am going to state here (again, without the proof):

$$\left[\hat{\mathbf{A}}\hat{\mathbf{B}}, \hat{\mathbf{C}}\right] = \hat{\mathbf{A}}\left[\hat{\mathbf{B}}, \hat{\mathbf{C}}\right] + \left[\hat{\mathbf{A}}, \hat{\mathbf{C}}\right]\hat{\mathbf{B}}\,.$$

1.5 Inverse Operator

Assume an equation in a Hilbert space of the following form

$$\hat{\mathbf{A}}\,|\psi\rangle = |\phi\rangle\ , \tag{1.83}$$

where $|x\rangle$ is the unknown Ket-vector that we would like to determine and $|b\rangle$ is the Ket-vector representing the known RHS of the above equation. To solve such equation, we define the *inverse operator* to $\hat{\mathbf{A}}$, denoted as $\hat{\mathbf{A}}^{-1}$, by the following formula

$$\hat{\mathbf{A}}^{-1}\hat{\mathbf{A}} = \hat{\mathbf{1}}\,. \tag{1.84}$$

Now - based on the previous section - we know that when acting with two operators the order generally matters. So let's take the above equation and act on both sides by $\hat{\mathbf{A}}$ - an incorrect but usual way of saying that would be that we multiply both sides by $\hat{\mathbf{A}}$ from the left. We get

$$\hat{\mathbf{A}}\hat{\mathbf{A}}^{-1}\hat{\mathbf{A}} = \hat{\mathbf{A}}\,. \tag{1.85}$$

But if we rewrite this as

$$(\hat{\mathbf{A}}\hat{\mathbf{A}}^{-1})\hat{\mathbf{A}} = \hat{\mathbf{A}}\,, \tag{1.86}$$

which we can do since acting by a linear operator can be represented by matrix multiplication which is associative, it is apparent that

$$\hat{\mathbf{A}}\hat{\mathbf{A}}^{-1} = \hat{\mathbf{1}} \tag{1.87}$$

and hence the following can be written for the inverse operator

$$\boxed{\hat{\mathbf{A}}^{-1}\hat{\mathbf{A}} = \hat{\mathbf{A}}\hat{\mathbf{A}}^{-1} = \hat{\mathbf{1}}}. \tag{1.88}$$

We can say that any (linear) operator commutes with it's own inverse, i.e.,

$$\left[\hat{\mathbf{A}}, \hat{\mathbf{A}}^{-1}\right] = 0\,. \tag{1.89}$$

The solution to Eq. (1.83) is then simply

$$|\psi\rangle = \hat{\mathbf{A}}^{-1}\,|\phi\rangle\ . \tag{1.90}$$

1.5.1 Inverse Operator - Discrete Case

To determine the form of the inverse operator in the discrete case, we want to solve the following operator equation for $\hat{\mathbf{B}}$:

$$\hat{\mathbf{B}}\hat{\mathbf{A}} = \hat{\mathbf{1}}\,. \tag{1.91}$$

The left hand side of the above equation can be expressed using (1.78) as

$$\sum_{k,l}\left(\sum_{m} B_{km}A_{ml}\right)|k\rangle\langle l| = \sum_{k}|k\rangle\langle k|\,, \tag{1.92}$$

where on the right hand side we employed the discrete basis formula for the unit operator. The above equation is then satisfied if and only if

$$\sum_{m} B_{km}A_{ml} = \delta_{kl}\,, \tag{1.93}$$

or in matrix notation

$$\mathbb{B}\mathbb{A} = \mathbb{I}\,, \tag{1.94}$$

where $\mathbb{I}$ stands for the identity matrix. But this is exactly the definition for the inverse matrix, i.e.,

$$\mathbb{B} = \mathbb{A}^{-1}\,. \tag{1.95}$$

We can then write

$$\hat{\mathbf{A}}^{-1} = \sum_{k,l}\left(\mathbb{A}^{-1}\right)_{kl}|k\rangle\langle l| \tag{1.96}$$

and

$$\hat{\mathbf{A}}^{-1} \to \mathbb{A}^{-1}\,. \tag{1.97}$$

1.5.2 Inverse Operator - Continuous Case

In the continuous case, Eq. (1.92) becomes

$$\int \mathrm{d}x\,\mathrm{d}y\,\left(\int \mathrm{d}z\,B(x,z)A(z,y)\right)|x\rangle\langle y| = \int \mathrm{d}x\;|x\rangle\langle x|\,, \tag{1.98}$$

which is then solved by imposing

$$\int \mathrm{d}z\,B(x,z)A(z,y) = \delta(x-y)\,. \tag{1.99}$$

By denoting

$$A^{-1}(x,y) \equiv B(x,y)\,, \tag{1.100}$$

(note that the $A^{-1}(x,y)$ does not mean an inverse to the function $A(x,y)$ in this case), we can now write

$$\hat{\mathbf{A}}^{-1} = \int \mathrm{d}x\,\mathrm{d}y\,A^{-1}(x,y)\,|x\rangle\langle y| \tag{1.101}$$

and

$$\hat{\mathbf{A}}^{-1} \to \int \mathrm{d}y\, A^{-1}(x,y) \times \ . \tag{1.102}$$

Note that in case where $\hat{\mathbf{A}}$ represents a differential operator, such function $A^{-1}(x,y)$ is usually called the *Green function*.

1.6 Adjoint Operator

Now it is time to move on to study how do the operators affect inner products. It is also the time when we revisit one of the confusions of the Dirac's notation.

Assume that we act with $\hat{\mathbf{A}}$ onto $|\psi\rangle$ and then form an inner product with $\langle\phi|$. Probably the most correct way of writing this would be something like

$$\langle\phi| \left(\hat{\mathbf{A}}\,|\psi\rangle\right) . \tag{1.103}$$

Of course there is no specific need for the round brackets so we can simplify this to

$$\langle\phi|\,\hat{\mathbf{A}}\,|\psi\rangle \tag{1.104}$$

but we need to keep in mind that when writing it this way the operator acts on the Ket-vector to the right, which is why some authors prefer the notation

$$\left\langle \phi \middle| \hat{\mathbf{A}}\psi \right\rangle . \tag{1.105}$$

Now let's ask the following question – at this point without a specific reasoning of why (that will come later) – is there a way to take the operator $\hat{\mathbf{A}}$ that acts on the right side of the inner product and somehow shift it to the other side such that the inner product stays unchanged? The answer is yes. We start by defining the so called adjoint operator to $\hat{\mathbf{A}}$ and denote it as $\hat{\mathbf{A}}^\dagger$ (the same notation as for the adjoint vector). Then, we impose the following condition on such adjoint operator

$$\boxed{\left\langle \hat{\mathbf{A}}^\dagger\phi \middle| \psi \right\rangle = \left\langle \phi \middle| \hat{\mathbf{A}}\psi \right\rangle}, \tag{1.106}$$

where

$$\left\langle \hat{\mathbf{A}}^\dagger\phi \right| \equiv \left(\hat{\mathbf{A}}^\dagger\,|\phi\rangle\right)^\dagger . \tag{1.107}$$

To proceed further, we need to (as always) separate the following derivation into the discrete and the continuous case.

1.6.1 Adjoint Operator - Discrete Case

We already know that $\hat{\mathbf{A}}$ acting on $|\psi\rangle$ results in

$$\hat{\mathbf{A}}\,|\psi\rangle = \sum_{k,l} A_{kl}\,\psi_l\,|k\rangle\ . \tag{1.108}$$

Then, by forming the inner product with

$$\langle\phi| = \sum_n \phi_n^*\,\langle n| \tag{1.109}$$

we get

$$\langle\phi|\,\hat{\mathbf{A}}\,|\psi\rangle = \sum_n \phi_n^*\,\langle n| \sum_{k,l} A_{kl}\,\psi_l\,|k\rangle = \sum_{k,l} \phi_k^*\,A_{kl}\,\psi_l\,, \tag{1.110}$$

which can be written by using the corresponding representations as

$$\begin{pmatrix}\phi_1^* & \cdots & \phi_N^*\end{pmatrix} \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots \\ A_{2,1} & A_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix}\psi_1 \\ \vdots \\ \psi_N\end{pmatrix}\ . \tag{1.111}$$

Now let's get back to the freshly defined adjoint operator $\hat{\mathbf{A}}^\dagger$. First, let it act on $|\phi\rangle$ such that

$$\hat{\mathbf{A}}^\dagger\,|\phi\rangle = \sum_{k,l} A_{kl}^\dagger\,\phi_l\,|k\rangle \tag{1.112}$$

and then follow by turning this into the corresponding Bra-vector:

$$\left\langle \hat{\mathbf{A}}^\dagger\phi \right| = \left(\hat{\mathbf{A}}^\dagger\,|\phi\rangle\right)^\dagger = \sum_{k,l} (A_{kl}^\dagger)^*\,\phi_l^*\,\langle k|\ . \tag{1.113}$$

Finally, the inner product with $|\psi\rangle$ ten gives us

$$\left\langle \hat{\mathbf{A}}^\dagger\phi \middle| \psi \right\rangle = \sum_{k,l} (A_{kl}^\dagger)^*\,\phi_l^*\,\langle k| \sum_m \psi_m\,|m\rangle = \sum_{k,l} \psi_k\,(A_{kl}^\dagger)^*\,\phi_l^*\,. \tag{1.114}$$

Now based on the definition we know that $A_{kl}^\dagger = \langle k|\,\hat{\mathbf{A}}^\dagger\,|l\rangle$. Hence, we can write

$$(A_{kl}^\dagger)^* = \langle k|\,\hat{\mathbf{A}}^\dagger\,|l\rangle^* = \left\langle k \middle| \hat{\mathbf{A}}^\dagger l \right\rangle^* = \left\langle \hat{\mathbf{A}}^\dagger l \middle| k \right\rangle = \left\langle l \middle| \hat{\mathbf{A}} k \right\rangle = \langle l|\,\hat{\mathbf{A}}\,|k\rangle = A_{lk}\,, \tag{1.115}$$

where in the third equality we used the second property of the inner product and in the fourth one the definition (1.106) of the adjoint operator. Finally, by comparing the LHS with the RHS of the above equation we get

$$A_{kl}^\dagger = A_{lk}^*\,. \tag{1.116}$$

With this knowledge we can then further manipulate the calculation (1.114) as

$$\sum_{k,l} \psi_k \, (A^\dagger_{kl})^* \, \phi^*_l = \sum_{k,l} \psi_k \, A_{lk} \, \phi^*_l = \sum_{k,l} \phi^*_k \, A_{kl} \, \psi_l \, , \tag{1.117}$$

where in the second equality we just appropriately renamed the indices. As you can see, the result is exactly the same as previously when we acted with $\hat{\mathbf{A}}$ to the right (see the calculation (1.110)), which is exactly what we wanted to achieve. Further, in the discrete representation the matrix corresponding to $\hat{\mathbf{A}}^\dagger$ is then

$$\mathbb{A}^\dagger = \left(\mathbb{A}^{\mathrm{T}}\right)^* \, , \tag{1.118}$$

where $^{\mathrm{T}}$ stands for the matrix transpose. Finally, by using Eq. (1.116) we can write the general formula for $\hat{\mathbf{A}}^\dagger$ as

$$\boxed{\hat{\mathbf{A}}^\dagger = \sum_{k,l} A^*_{lk} \, |k\rangle \, \langle l| \quad \text{with} \quad A_{kl} = \langle k| \, \hat{\mathbf{A}} \, |l\rangle} \, . \tag{1.119}$$

1.6.2 Adjoint Operator - Continuous Case

As always, to proceed with the continuous case we follow the same steps as in the discrete case. Hence, we start by reminding that acting by $\hat{\mathbf{A}}$ onto $|\psi\rangle$ while working with a continuous basis $\{|x\rangle\}$ results in

$$\hat{\mathbf{A}} \, |\psi\rangle = \int \mathrm{d}x \, \mathrm{d}y \, A(x,y) \psi(y) \, |x\rangle \, . \tag{1.120}$$

Then we form the corresponding inner product with

$$\langle \phi| = \int \mathrm{d}z \, \phi^*(z) \, \langle z| \, , \tag{1.121}$$

resulting in

$$\langle \phi| \, \hat{\mathbf{A}} \, |\psi\rangle = \int \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \, \phi^*(z) A(x,y) \psi(y) \, \langle z|x\rangle = \int \mathrm{d}x \, \mathrm{d}y \, \phi^*(x) A(x,y) \psi(y) \, , \tag{1.122}$$

which is an continuous analogue of (1.111).

Next, we work out the formula for $\hat{\mathbf{A}}^\dagger$ as follows:

$$\left\langle \hat{\mathbf{A}}^\dagger \phi \middle| \psi \right\rangle = \int \mathrm{d}x \, \mathrm{d}y \, [A^\dagger(x,y)]^* \phi^*(y) \, \langle x| \int \mathrm{d}z \, \psi(z) \, |z\rangle = \int \mathrm{d}x \, \mathrm{d}y \, [A^\dagger(x,y)]^* \phi^*(y) \psi(x) \, , \tag{1.123}$$

where

$$[A^\dagger(x,y)]^* = \langle x| \, \hat{\mathbf{A}}^\dagger \, |y\rangle^* = \left\langle \hat{\mathbf{A}}^\dagger y \middle| x \right\rangle = \left\langle y \middle| \hat{\mathbf{A}} x \right\rangle = \langle y| \, \hat{\mathbf{A}} \, |x\rangle = A(y,x) \, . \tag{1.124}$$

Finally, by comparing both sides of the above calculation we get

$$A^\dagger(x,y) = A^*(y,x)\,. \tag{1.125}$$

Then, inserting this back into (1.123) produces

$$\int \mathrm{d}x\,\mathrm{d}y\,[A^\dagger(x,y)]^*\phi^*(y)\psi(x) = \int \mathrm{d}x\,\mathrm{d}y\,\phi^*(y)A(y,x)\psi(x) = \int \mathrm{d}x\,\mathrm{d}y\,\phi^*(x)A(x,y)\psi(y)\,, \tag{1.126}$$

which agrees with (1.122). Therefore, the general formula for $\hat{\mathbf{A}}^\dagger$ can be written as

$$\boxed{\hat{\mathbf{A}}^\dagger = \int \mathrm{d}x\,\mathrm{d}y\,A^*(y,x)\,|x\rangle\,\langle y| \quad \text{with} \quad A(x,y) = \langle x|\,\hat{\mathbf{A}}\,|y\rangle}\,. \tag{1.127}$$

The continuous representation of $\hat{\mathbf{A}}^\dagger$ is then

$$\int \mathrm{d}y\,A^*(y,x)\times\,, \tag{1.128}$$

where the $\times$ symbol has the same meaning as before.

Lastly, we ask ourselves the question - what is the appropriate formula for an adjoint operator to a multiple of two (or more) operators? Fortunately, the answer is trivial and follows directly from the definition (1.106):

$$\left\langle \phi \middle| \hat{\mathbf{A}}\hat{\mathbf{B}}\psi \right\rangle = \left\langle \hat{\mathbf{A}}^\dagger\phi \middle| \hat{\mathbf{B}}\psi \right\rangle = \left\langle \hat{\mathbf{B}}^\dagger\hat{\mathbf{A}}^\dagger\phi \middle| \psi \right\rangle\,, \tag{1.129}$$

and therefore

$$\boxed{(\hat{\mathbf{A}}\hat{\mathbf{B}})^\dagger = \hat{\mathbf{B}}^\dagger\hat{\mathbf{A}}^\dagger}\,. \tag{1.130}$$

This can then be simply extended to a multiple of N operators as

$$(\hat{\mathbf{A}}_1\cdots\hat{\mathbf{A}}_N)^\dagger = \hat{\mathbf{A}}_N^\dagger\cdots\hat{\mathbf{A}}_1^\dagger\,. \tag{1.131}$$

To sum up, we have shown how do operators affect the inner products when acting on the Ket-vector to the right, followed by deriving formulas for the adjoint operators that preserve the inner product while acting on the Bra-vector to the left.

1.7 Eigenvalues and Eigenvectors

Each one of you were surely at least once given the task to calculate the eigenvalues and the corresponding eigenvectors to a specific matrix. No we will just generalise this task to operators. First, assume there exist $\hat{\mathbf{A}}$, $|\lambda\rangle$ and $\lambda \in \mathbb{C}$ such that

$$\hat{\mathbf{A}}\,|\lambda\rangle = \lambda\,|\lambda\rangle\ . \tag{1.132}$$

Then we call λ the *eigenvalue* and $|\lambda\rangle$ the corresponding (to that λ) *eigenvector* of $\hat{\mathbf{A}}$ (we do not use the term eigen-Ket-vector because that looks horrible). There is not much more we can say about the eigenvalue problem in general. But we will further inspect the properties of eigenvalues and the corresponding eigenvectors for a very special class of so called *hermitian operators*, in which case things are going to get much more interesting.

1.8 Hermitian Operators

Now is the time to explain why did we define the adjoint operators. Let's start with a definition. The operator $\hat{\mathbf{A}}$ with the following property:

$$\hat{\mathbf{A}}^\dagger = \hat{\mathbf{A}} \tag{1.133}$$

is called *hermitian*. Or in terms of the inner products we can write

$$\left\langle \hat{\mathbf{A}}\phi \middle| \psi \right\rangle = \left\langle \phi \middle| \hat{\mathbf{A}}\psi \right\rangle = \langle \phi | \, \hat{\mathbf{A}} \, | \psi \rangle \ . \tag{1.134}$$

In other words, it does not matter if $\hat{\mathbf{A}}$ acts on either the right or the left side of the inner product, the result will be the same. Such operators play a key role in quantum mechanics and have some amazing properties that we are going to explore now.

First, we focus on the eigenvalue problem for the hermitian operators, i.e.,

$$\hat{\mathbf{A}} \, | \lambda \rangle = \lambda \, | \lambda \rangle \ . \tag{1.135}$$

First, note that there can (and most typically will) be more then just one eigenvalue to each hermitian operator. The complete set of all the eigenvalues of $\hat{\mathbf{A}}$ is then called the *spectrum* of $\hat{\mathbf{A}}$. And such spectrum can be either discrete or continuous as usual.

Let's now calculate the following inner product:

$$\left\langle \lambda \middle| \hat{\mathbf{A}}\lambda \right\rangle = \lambda \, \langle \lambda | \lambda \rangle \ . \tag{1.136}$$

However, since the operator is hermitian, then

$$\left\langle \lambda \middle| \hat{\mathbf{A}}\lambda \right\rangle = \left\langle \hat{\mathbf{A}}\lambda \middle| \lambda \right\rangle = \lambda^* \, \langle \lambda | \lambda \rangle \ , \tag{1.137}$$

where the second equality follows from the third property of the inner product. By comparing those two results, we arrive at

$$\boxed{\lambda = \lambda^*} . \tag{1.138}$$

In other words, all the eigenvalues of a Hermitian operator are real

$$\lambda \in \mathbb{R} \, . \tag{1.139}$$

And that is a powerful statement!

Next, assume two distinct eigenvalues (and the corresponding eigenvectors) as

$$\hat{\mathbf{A}} \, | \lambda \rangle = \lambda \, | \lambda \rangle \ , \quad \hat{\mathbf{A}} \, | \mu \rangle = \mu \, | \mu \rangle \ , \tag{1.140}$$

where $\lambda \neq \mu$. We then consider the following inner product

$$\left\langle \lambda \middle| \hat{\mathbf{A}}\mu \right\rangle = \mu \left\langle \lambda | \mu \right\rangle \, . \tag{1.141}$$

Once again, we make use of the hermitian property of $\hat{\mathbf{A}}$ to express

$$\left\langle \lambda \middle| \hat{\mathbf{A}}\mu \right\rangle = \left\langle \hat{\mathbf{A}}\lambda \middle| \mu \right\rangle = \lambda \left\langle \lambda | \mu \right\rangle \, , \tag{1.142}$$

while there is no need for the * symbol anymore since λ is real. If we then take the preceding two equations and subtract the first one from the second we get

$$(\lambda - \mu) \left\langle \lambda | \mu \right\rangle = 0 \, , \tag{1.143}$$

but since $\lambda \neq \mu$, the only possibility to fulfill the above equation is that

$$\boxed{\left\langle \lambda | \mu \right\rangle = 0 \quad \text{for} \quad \lambda \neq \mu} \, . \tag{1.144}$$

In other words, all the eigenvectors of $\hat{\mathbf{A}}$ corresponding to distinct eigenvalues are mutually orthogonal. Another powerful statement here!

Let's now make use of the two found properties to express an appropriate orthonormal basis by the set of all the eigenvectors of $\hat{\mathbf{A}}$, hereinafter referred to as the *eigenbasis*.

1.8.1 Orthonormal Eigenbasis - Discrete Case

In the discrete case, i.e., the case where the spectrum of $\hat{\mathbf{A}}$ is discrete, we can label each eigenvalue and the corresponding eigenvector by a discrete index k as

$$\hat{\mathbf{A}} \left| \lambda_k \right\rangle = \lambda_k \left| \lambda_k \right\rangle \, . \tag{1.145}$$

By applying Eq. (1.144), we get

$$\left\langle \lambda_k | \lambda_l \right\rangle = 0 \quad \text{for} \quad k \neq l \, . \tag{1.146}$$

Then, it must hold that

$$\left\langle \lambda_k | \lambda_l \right\rangle = \alpha_k \delta_{kl} \, , \tag{1.147}$$

since there is no reason for the eigenvectors to be automatically normalized to unit norm. Note that even if there are two indices k, l present at the left hand side of the above equation, we only need one of them for α_k due to the presence of δ_{kl} on the right hand side. Also, due to the first property of the inner product, $\alpha_k \in \mathbb{R}$. So at this point, we have an orthogonal basis. To make it further orthonormal, we simply define

$$\left| k \right\rangle = \frac{\left| \lambda_k \right\rangle}{\sqrt{\alpha_k}} \, . \tag{1.148}$$

By calculating

$$\langle k|l\rangle = \frac{\langle \lambda_k|\lambda_l\rangle}{\sqrt{\alpha_k\alpha_l}} = \frac{\alpha_k\delta_{kl}}{\sqrt{\alpha_k\alpha_l}} = \delta_{kl}\,, \tag{1.149}$$

we come to a conclusion that such renormalized eigenvectors form an appropriate orthonormal eigenbasis. Actually, to be completely honest, there is one additional think we ought to do. Formally, we must show that the basis $\{|k\rangle\}$ is complete, which basically means that any $|\psi\rangle \in \mathscr{H}$ can be expressed just by those. But the proof is generally complicated, so you have to blindly believe me that it is true.

1.8.2 Orthonormal Eigenbasis - Continuous Case

In case of the continuous spectrum of $\hat{\mathbf{A}}$, we use a continuous index x to label the eigenvalues and the corresponding eigenvectors as

$$\hat{\mathbf{A}}\,|\lambda(x)\rangle = \lambda(x)\,|\lambda(x)\rangle\ . \tag{1.150}$$

Based on Eq. (1.144) we then have

$$\langle\lambda(x)|\lambda(y)\rangle = 0 \quad \text{for} \quad x \neq y\,, \tag{1.151}$$

which means that

$$\langle\lambda(x)|\lambda(y)\rangle = \alpha(x)\delta(x-y)\,. \tag{1.152}$$

By following the same steps as in the discrete case, we can now normalize the respective eigenvectors to

$$|x\rangle = \frac{|\lambda(x)\rangle}{\sqrt{\alpha(x)}}\,. \tag{1.153}$$

By checking

$$\langle x|y\rangle = \frac{\langle\lambda(x)|\lambda(y)\rangle}{\sqrt{\alpha(x)\alpha(y)}} = \frac{\alpha(x)\delta(x-y)}{\sqrt{\alpha(x)\alpha(y)}} = \delta(x-y)\,, \tag{1.154}$$

and therefore we have just proven that such renormalization (again) leads to an appropriate orthonormal eigenbasis $\{|x\rangle\}$.

1.8.3 Hermitian Operator in it's Eigenbasis - Discrete Case

We proceed further to show how does one express a general linear hermitian operator $\hat{\mathbf{A}}$ in it's own eigenbasis. Based on the above discussions, we assume an orthonormal basis $\{|k\rangle\}$ formed by the eigenvalues of $\hat{\mathbf{A}}$, i.e.,

$$\hat{\mathbf{A}}\,|k\rangle = \lambda_k\,|k\rangle\ . \tag{1.155}$$

By employing the general formula for expressing any operator in terms of a discrete basis (1.56) we write

$$\hat{\mathbf{A}} = \sum_{k,l}\langle k|\,\hat{\mathbf{A}}\,|l\rangle\,|k\rangle\,\langle l| = \sum_{k,l}\lambda_l\,\delta_{kl}\,|k\rangle\,\langle l| = \sum_k \lambda_k\,|k\rangle\,\langle k|\ , \tag{1.156}$$

Due to the importance of this formula we put it in a box

$$\boxed{\hat{\mathbf{A}} = \sum_k \lambda_k \left|k\right\rangle \left\langle k\right|} . \tag{1.157}$$

By acting with $\hat{\mathbf{A}}$ on $\left|\psi\right\rangle$ then results in

$$\hat{\mathbf{A}} \left|\psi\right\rangle = \sum_k \lambda_k \left|k\right\rangle \left\langle k\right| \sum_l \psi_l \left|l\right\rangle = \sum_k \lambda_k \psi_k \left|k\right\rangle \ , \tag{1.158}$$

and the discrete representation of $\hat{\mathbf{A}}$ is therefore

$$\hat{\mathbf{A}} \to \mathrm{diag}(\lambda_1, \lambda_2, \dots) . \tag{1.159}$$

1.8.4 Hermitian Operator in it's Eigenbasis - Continuous Case

In the continuous case, we assume a continuous orthonormal basis $\{\left|x\right\rangle\}$ formed by the eigenvectors of $\hat{\mathbf{A}}$, i.e.,

$$\hat{\mathbf{A}} \left|x\right\rangle = \lambda(x) \left|x\right\rangle \ . \tag{1.160}$$

Now we use the general formula (1.70) to calculate

$$\hat{\mathbf{A}} = \int \mathrm{d}x \, \mathrm{d}y \ \left\langle x\right| \hat{\mathbf{A}} \left|y\right\rangle \left|x\right\rangle \left\langle y\right| = \int \mathrm{d}x \, \mathrm{d}y \, \lambda(y) \delta(x-y) \left|x\right\rangle \left\langle y\right| = \int \mathrm{d}x \, \lambda(x) \left|x\right\rangle \left\langle x\right| \ . \tag{1.161}$$

Another important formula here – let's put it in a box:

$$\boxed{\hat{\mathbf{A}} = \int \mathrm{d}x \, \lambda(x) \left|x\right\rangle \left\langle x\right|} . \tag{1.162}$$

Now by acting with $\hat{\mathbf{A}}$ onto $\left|\psi\right\rangle$ we get

$$\hat{\mathbf{A}} \left|\psi\right\rangle = \int \mathrm{d}x \, \lambda(x) \left|x\right\rangle \left\langle x\right| \int \mathrm{d}y \, \psi(y) \left|y\right\rangle = \int \mathrm{d}x \, \lambda(x) \psi(x) \left|x\right\rangle \ . \tag{1.163}$$

Hence, the corresponding continuous representation of $\hat{\mathbf{A}}$ is just

$$\hat{\mathbf{A}} \to \lambda(x) . \tag{1.164}$$

And that basically covers all the basic properties of the Hermitian operators we need for the purpose of this text.

1.9 Spectral Decomposition

The last bit of mathematics we are going to need in order to solve the famous Schrödinger equation is called the *spectral decomposition theorem*. But first, we need to introduce a new concept called the function of an operator. All of us are familiar with functions of

the variable $x \to f(x)$, such as $\sin(x), \exp(x)$, etc. But what if I asked you to evaluate the respective function while replacing x with a operator $\hat{\mathbf{A}}$, i.e.,

$$f(\hat{\mathbf{A}}) = ? \tag{1.165}$$

Fortunately, the answer to this is simple. We just use the fact that each "well behaved" (analytic, to be precise) function can be expressed by the respective Taylor series as

$$f(x) = \sum_n c_n x^n \,. \tag{1.166}$$

The same function but now evaluated at $\hat{\mathbf{A}}$ is then defined by the same Taylor series as

$$f(\hat{\mathbf{A}}) = \sum_n c_n \hat{\mathbf{A}}^n \,. \tag{1.167}$$

If we now assume a general hermitian operator $\hat{\mathbf{A}}$ and the corresponding basis of normalised eigenvectors $\{|n\rangle\}$ (for the purpose of this text we need only the discrete version), the spectral decomposition theorem then states the following:

$$\boxed{f(\hat{\mathbf{A}}) = \sum_k f(\lambda_k) \, |k\rangle \, \langle k|} \,. \tag{1.168}$$

To prove this, we proceed as follows:

$$\begin{aligned} f(\hat{\mathbf{A}}) &= \sum_{k,l} \langle k| \, f(\hat{\mathbf{A}}) \, |l\rangle \, |k\rangle \, \langle l| = \sum_{k,l} \langle k| \left(\sum_n c_n \hat{\mathbf{A}}^n \right) |l\rangle \, |k\rangle \, \langle l| = \sum_{k,l,n} c_n \, \langle k| \, \hat{\mathbf{A}}^n \, |l\rangle \, |k\rangle \, \langle l| \\ &= \sum_{k,l} \left(\sum_n c_n \lambda_l^n \right) \langle k|l\rangle \, |k\rangle \, \langle l| = \sum_{k,l} f(\lambda_l) \delta_{kl} \, |k\rangle \, \langle l| = \sum_k f(\lambda_k) \, |k\rangle \, \langle k| \,, \end{aligned} \tag{1.169}$$

where in the first equality we have used the fact that $f(\hat{\mathbf{A}})$ is itself an operator and therefore it must be able to be expressed using (1.56).

1.10 Summary

I understand that this was probably a bit overwhelming for someone who encountered the Dirac's notation and Hilbert spaces for the first time. Therefore, here I present two tables (for the discrete and the continuous case) with the most important formulae while strictly distinguishing between the real objects and their respective representations.

Object	True Form	Representation
Ket-vector	$\|\psi\rangle = \sum_n \psi_n \|n\rangle$	$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}$
Bra-vector	$\langle\psi\| = \sum_n \psi_n \langle n\|$	$\begin{pmatrix} \psi_1^* & \psi_2^* & \cdots \end{pmatrix}$
Operator	$\hat{\mathbf{A}} = \sum_{k,l} A_{kl} \|k\rangle \langle l\|$	$\mathbb{A} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots \\ A_{2,1} & A_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$
Inverse operator	$\hat{\mathbf{A}}^{-1} = \sum_{k,l} (\mathbb{A}^{-1})_{kl} \|k\rangle \langle l\|$	$\mathbb{A}^{-1} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots \\ A_{2,1} & A_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}^{-1}$
Adjoint operator	$\hat{\mathbf{A}}^\dagger = \sum_{k,l} A_{lk}^* \|k\rangle \langle l\|$	$\mathbb{A}^\dagger = \left(\mathbb{A}^{\mathrm{T}}\right)^* = \begin{pmatrix} A_{1,1}^* & A_{2,1}^* & \cdots \\ A_{1,2}^* & A_{2,2}^* & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$
Operator in eigenbasis	$\hat{\mathbf{A}} = \sum_k \lambda_k \|k\rangle \langle k\|$	$\mathbb{A} = \operatorname{diag}(\lambda_1, \lambda_2, \dots)$

Table 1.1: Discrete basis formulas

Object	True Form	Representation
Ket-vector	$\|\psi\rangle = \int \mathrm{d}x\, \psi(x) \|x\rangle$	$\psi(x)$
Bra-vector	$\langle\psi\| = \int \mathrm{d}x\, \psi^*(x) \langle x\|$	$\int \mathrm{d}x\, \psi^*(x)\times$
Inner product	$\langle\phi\|\psi\rangle$	$\int \mathrm{d}x\, \phi^*(x)\psi(x)$
Operator	$\hat{\mathbf{A}} = \int \mathrm{d}x\, \mathrm{d}y\, A(x,y) \|x\rangle \langle y\|$	$\int \mathrm{d}y\, A(x,y)\times$
Inverse operator	$\hat{\mathbf{A}}^{-1} = \int \mathrm{d}x\, \mathrm{d}y\, A^{-1}(x,y) \|x\rangle \langle y\|$	$\int \mathrm{d}y\, A^{-1}(x,y)\times$
Adjoint operator	$\hat{\mathbf{A}}^\dagger = \int \mathrm{d}x\, \mathrm{d}y\, A^*(y,x) \|x\rangle \langle y\|$	$\int \mathrm{d}y\, A^*(y,x)\times$
Operator in eigenbasis	$\hat{\mathbf{A}} = \int \mathrm{d}x\, \lambda(x) \|x\rangle \langle x\|$	$\lambda(x)$

Table 1.2: Continuous basis formulas

Chapter 2

Postulates of Quantum Mechanics

In this chapter we focus on developing the framework of QM based on the postulates which will be stated in their general form by using the Dirac's notation. And only then, based on those postulates together with what we have learnt in the preceding chapter, we will derive the formulas in the famous wavefunction representation.

2.1 Postulate I. - Physical System

Each physical system is represented by a corresponding Hilbert space $\mathscr{H}$. The state of such system is fully described by $\|\psi\rangle \in \mathscr{H}$.

I can imagine that for someone that has never seen QM before this probably makes absolutely zero sense. Please be prepared that all the postulates of QM are just weird when you go through them for the first time and it seems like there is no correspondence to the classical physics as you have learnt it so far. But don't worry - after we deal with all the postulates, I will show you that there is actually a very direct correspondence with the Hamilton formulation of classical mechanics that we will make use of.

Let's compare this postulate with what do we know from the classical dynamics. In classical dynamics, a physical system is represented by the phase space consisting of the generalised coordinates and the conjugate momenta. The state of the system is then represented by a point in such phase space. On the other hand, in the world of QM, the system is represented by an abstract Hilbert space $\mathscr{H}$ and the state is then given by $|\psi\rangle \in \mathscr{H}$. Maybe it is worth reminding ourselves of what do we mean by the state of the system – to know the state of a system means that we possess all the necessary information to be able to calculate how will the system evolve in time.

There is not much more we can do here at this point. But before moving on, for the purpose of this text, we will require that the Ket-vectors representing the states will be always normalised, i.e.,

$$\langle\psi|\psi\rangle = 1\,. \tag{2.1}$$

This is by no means necessary, but it is a standard convention that will make our life a bit easier.

2.2 Postulate II. - Observables

The generalised coordinates and the conjugate momenta are replaced by the linear hermitian operators

$$q \to \hat{\mathbf{q}}\,, \quad p \to \hat{\mathbf{p}}\,.$$

Any dynamical variable is then replaced by the linear hermitian operator by the following rule

$$A(p,q,t) \to \hat{\mathbf{A}} = A(\hat{\mathbf{p}}, \hat{\mathbf{q}}, t)\,.$$

First, in the context of QM, dynamical variables are usually referred to as the *observables* (we can observe them by the act of measurement). Second, at this point we are not going to specify how to define the $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$ operators – that will come later, but it is important to understand the latter formula for $\hat{\mathbf{A}}$. In classical dynamics, any observable (dynamical variable) is a function of the generalised coordinates, conjugate momenta and time, i.e., $A = A(q,p,t)$. The formula $\hat{\mathbf{A}} = A(\hat{\mathbf{p}}, \hat{\mathbf{q}}, t)$ then means that to express the operator $\hat{\mathbf{A}}$ we take the the function $A(q,p,t)$ and replace $q \to \hat{\mathbf{q}}$ and $p \to \hat{\mathbf{p}}$. Some authors use the notation $\hat{\mathbf{A}}(\hat{\mathbf{q}}, \hat{\mathbf{q}}, t)$, which is perfectly correct, since (of course) the resulting object is an operator. But I wanted to stress out that the operator arises by taking the same function as in the classical case and only then replacing the arguments with the corresponding operators. So please keep that in mind. Now moving on to the next postulate.

2.3 Postulate III. - Measurements

When measuring the observable A, the only possible outcomes are the eigenvalues of the corresponding operator $\hat{\mathbf{A}}$.

Now things start to get interesting. First, we should specify more clearly what do we mean by "measuring". In the context of QM, by measuring we simply mean that we somehow determine the value of the observable a. Do not think of of some elaborate method including various expensive measurement tools. On the contrary, think of it in a more abstract way.

So the postulate states that the only allowed results of the measurement are the eigenvalues of the corresponding operator. This explains why did we require all the operators to be hermitian in the preceding postulate – they have purely real eigenvalues, and of course the measurement of a real physical quantity cannot have a complex result.

We should emphasise that this is something completely different than what we are used to in the world of classical dynamics. In QM, one of the most important operators is the

Hamilton operator $\hat{\mathbf{H}}$ representing the total energy of the system. And in many cases, the spectrum of $\hat{\mathbf{H}}$ will be discrete. Then, according to this postulate, only discrete values of energy can be measured on the system! But more on this later. Now, let's move on to the next postulate.

2.4 Postulate IV. - Mean Value of Measurements

> When repeatedly measuring the observable A on a system in state $|\psi\rangle$, the mean value is equal to $\langle\psi|\,\hat{\mathbf{A}}\,|\psi\rangle$.

This is probably the most complicated postulate to understand, so let me make a few comments here. First, I should probably explain what do we mean by repeatedly measuring on a system in state $|\psi\rangle$. But it would not make much sence until we learn the postulate 5, so let me postpone this for now. Second, the previously mentioned postulate tells us that the only possible outcomes when measuring the observable a are eigenvalues of $\hat{\mathbf{A}}$. And now this postulate actually specifies the probabilities of the individual outcomes - and I will show you how.

2.4.1 Mean Value - Discrete Case

Assume a basis $\{|n\rangle\}$ of eigenvectors of $\hat{\mathbf{A}}$, i.e.,

$$\hat{\mathbf{A}}\,|n\rangle = a_n\,|n\rangle\ , \tag{2.2}$$

where a_n denotes the corresponding eigenvalue. Since the system is in a state $|\psi\rangle$, we can expand it in our basis as

$$|\psi\rangle = \sum_n \psi_n\,|n\rangle\ . \tag{2.3}$$

According to the postulate, the mean value of the results when repeatedly measuring A is equal to

$$\bar{A} = \langle\psi|\,\hat{\mathbf{A}}\,|\psi\rangle = \sum_{m,n}\psi_m^*\psi_n\,\langle m|\,\hat{\mathbf{A}}\,|n\rangle = \sum_{m,n}\psi_m^*\psi_n a_n\delta_{mn} = \sum_n |\psi_n|^2 a_n\,. \tag{2.4}$$

Note that I have used the overline to denote a mean value. Some authors prefer to use $\langle A\rangle$, but (at least to me) this is unpractical, since we already use the pointy brackets for the inner product. From the theory of probability we can now identify that

$$\boxed{w_n \equiv |\psi_n|^2 = |\,\langle n|\psi\rangle\,|^2} \tag{2.5}$$

represents the probability that the output of the individual measurement will be a_n. Since the probabilities must always add up to 1, we can write

$$\sum_n w_n = \sum_n |\psi_n|^2 = \sum_n \psi_n^*\psi_n = \langle\psi|\psi\rangle = 1\,, \tag{2.6}$$

which is why we required all the physical states to be normalised to unity.

2.4.2 Mean Value - Continuous Case

In the continuous case, the following applies for the continuous basis $\{|x\rangle\}$ of eigenvectors

$$\hat{\mathbf{A}}\,|x\rangle = a(x)\,|x\rangle\ , \tag{2.7}$$

where $a(x)$ denotes the continuous spectrum of eigenvalues. The state $|\psi\rangle$ is now expressed as

$$|\psi\rangle = \int \mathrm{d}x\,\psi(x)\,|x\rangle \tag{2.8}$$

while the following applies for the mean value $\bar{a}$:

$$\bar{A} = \langle\psi|\,\hat{\mathbf{A}}\,|\psi\rangle = \int \mathrm{d}x\,\mathrm{d}y\,\psi^*(x)\psi(y)\,\langle x|\,\hat{\mathbf{A}}\,|y\rangle = \int \mathrm{d}x\,|\psi(x)|^2 a(x)\,, \tag{2.9}$$

whereas we can identify

$$\boxed{w(x) \equiv |\psi(x)|^2 = |\,\langle x|\psi\rangle\,|^2} \tag{2.10}$$

as the probability density for the measurement result to be equal to $a(x)$. Similarly as in the previous case, the probability density must always give 1 when integrated over all the possible values of x. Hence,

$$\int \mathrm{d}x\,w(x) = \int \mathrm{d}x\,|\psi(x)|^2 = \int \mathrm{d}x\,\psi^*(x)\psi(x) = \langle\psi|\psi\rangle = 1\,, \tag{2.11}$$

which once again explains the necessity of normalised states.

So now we know that the outputs of any measurement can be only the eigenvalues of the corresponding hermitian operator and at the same time if we know the state $|\psi\rangle$ we are able to predict the probabilities of the individual outcomes. But even though all of this sounds nice and easy, we should be shocked - the QM dictates that the world as we know id does not behave in a deterministically as Newton thought but rather stochastically! To paraphrase Forrest Gump – "Quantum mechanics is like a box of chocolates. You never know what you're gonna get."

2.5 Postulate V. - System Collapse

By measuring a specific value of the observable A, the system transitions (collapses) to the corresponding eigenstate.

Now, as promised, let me explain what do we mean by the repeated measurement on a system in the state $|\psi\rangle$ in the context of the preceding postulate. Assume we want to measure the observable A on a system in a general state $|\psi\rangle$. But this postulate states that whenever we conduct the measurement, the system collapses to the eigenstate corresponding to the measured eigenvalue. Hence, we cannot conduct the measurement again, since the system is now in a different state then before! Hence, the only way

to conduct repeated measurements is to have (infinitely) many identical copies of the system, each one prepared in the same state $|\psi\rangle$. The measurement is then conducted on each such system exactly once. The mean value of all such measurements then gives $\langle\psi|\,\hat{\mathbf{A}}\,|\psi\rangle$ as stated above.

2.6 Postulate VI. - Schrödinger Picture of Time Evolution

The the time evolution of a system is contained in the state Ket-vector $|\psi\rangle = |\psi(t)\rangle$ while the operators $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$ are time independent. The equation of motion is called the Schrödinger's equation and is expressed as

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\psi(t)\rangle = \hat{\mathbf{H}}\,|\psi(t)\rangle\ ,$$

where

$$\hat{\mathbf{H}} = H(\hat{\mathbf{q}}, \hat{\mathbf{p}}, t)$$

is the Hamilton operator corresponding to the Hamiltonian and

$$\hbar \approx 1.054571817 \cdot 10^{-34}\,\mathrm{J\,s}$$

is the reduced Planck constant.

A few notes here. First, the $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$ stand for the whole set of $\{\hat{\mathbf{q}}_k\}, \{\hat{\mathbf{p}}_k\}$, but we will use this simplified notation for the sake of readability. Second, the fact that the time evolution of a system is contained in the respective state Ket-vector is just a standardly used convention called the *Schrödinger picture*. Later in the text, we will also present a different convention called the *Heisenberg picture* where (on the contrary) the states are time independent and $\hat{\mathbf{q}} = \hat{\mathbf{q}}(t)$, $\hat{\mathbf{p}} = \hat{\mathbf{p}}(t)$ instead. Third, even though $\hat{\mathbf{q}}$, $\hat{\mathbf{p}}$ are time independent, a general observable $\hat{\mathbf{A}} = A(\hat{\mathbf{q}}, \hat{\mathbf{p}}, t)$ can depend on time explicitly as

$$\frac{\mathrm{d}\hat{\mathbf{A}}(t)}{\mathrm{d}t} = \frac{\partial A(\hat{\mathbf{q}}, \hat{\mathbf{p}}, t)}{\partial t}\,, \tag{2.12}$$

where the notation on the RHS means

$$\frac{\partial A(\hat{\mathbf{q}}, \hat{\mathbf{p}}, t)}{\partial t} \equiv \left.\frac{\partial A(q, p, t)}{\partial t}\right|_{\substack{q\to\hat{\mathbf{q}}\\ p\to\hat{\mathbf{p}}}}\,. \tag{2.13}$$

(This is why I prefer not to use the operator symbol for $A(\hat{\mathbf{q}}, \hat{\mathbf{p}}, t)$ – to stress out that e.g., when dealing with derivatives, we treat A as a standard multivariable function and only then replace q, p with the corresponding operators.) Hence, an observable $\hat{\mathbf{A}}$ depends on time iff the dynamical variable A has an explicit time dependence, which (mostly) won't be the case in this text.

Now we have covered all the postulates of QM. And even though they might seem a bit random at first sight, in the following section, I will show you that there is actually a nice correspondence between QM and the classical dynamics.

2.7 Correspondence between QM and classical dynamics

There is a different way of how to look at the time evolution in the context of QM. Instead of considering the state of the system to be time dependent, we can define the *evolution operator* in the following way

$$|\psi(t)\rangle = \hat{\mathbf{U}}(t, t_0)\,|\psi(t_0)\rangle\ . \tag{2.14}$$

where $t_0 = const.$ represents an arbitrary (let's call it initial) time and $|\psi_0\rangle$ stands for the state Ket-vector evaluated at t_0. (Note that due to the Schrödinger equation being linear, $\hat{\mathbf{U}}(t, t_0)$ must also be linear.) Hence, the time evolution is then represented by acting with this operator onto the initial state. For simplicity and without the loss of generality, we set $t_0 = 0$ and relabel

$$\hat{\mathbf{U}}(t, 0) \to \hat{\mathbf{U}}(t)\,, \quad |\psi(t_0)\rangle = |\psi(0)\rangle \to |\psi_0\rangle\ . \tag{2.15}$$

Hence, we can rewrite (2.14) as

$$|\psi(t)\rangle = \hat{\mathbf{U}}(t)\,|\psi_0\rangle\ . \tag{2.16}$$

Since we require the state vector to be normalised at any time, which we can write as

$$\langle\psi_0|\psi_0\rangle = \langle\psi(t)|\psi(t)\rangle = \left\langle \hat{\mathbf{U}}(t)\psi_0 \middle| \hat{\mathbf{U}}(t)\psi_0 \right\rangle = 1\,, \tag{2.17}$$

the time evolution operator must be unitary, i.e.,

$$\hat{\mathbf{U}}^\dagger(t)\hat{\mathbf{U}}(t) = \hat{\mathbf{1}}\,. \tag{2.18}$$

By using the formula (2.14) and substituting into the Schrödinger's equation we obtain

$$\mathrm{i}\hbar\frac{\mathrm{d}\hat{\mathbf{U}}(t)}{\mathrm{d}t} = \hat{\mathbf{H}}\hat{\mathbf{U}}(t)\,, \tag{2.19}$$

which holds since $|\psi_0\rangle$ does not depend on time. We will also need the adjoint version of the above equation, which we can write as

$$-\mathrm{i}\hbar\frac{\mathrm{d}\hat{\mathbf{U}}^\dagger(t)}{\mathrm{d}t} = \hat{\mathbf{U}}^\dagger(t)\hat{\mathbf{H}}^\dagger\,. \tag{2.20}$$

Now assume an arbitrary observable $\hat{\mathbf{A}}$. According to the fourth postulate, the mean value of A at the time t is given as

$$\langle\psi(t)|\,\hat{\mathbf{A}}\,|\psi(t)\rangle = \langle\psi_0|\,\hat{\mathbf{U}}^\dagger(t)\hat{\mathbf{A}}\hat{\mathbf{U}}(t)\,|\psi_0\rangle\ , \tag{2.21}$$

where we have used (2.16). Now we apply the total time derivative onto this expression, resulting in

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\psi_0\middle|\hat{\mathbf{U}}^\dagger(t)\hat{\mathbf{A}}\hat{\mathbf{U}}(t)\middle|\psi_0\right\rangle = \left\langle\psi_0\middle|\frac{\mathrm{d}\hat{\mathbf{U}}^\dagger(t)}{\mathrm{d}t}\hat{\mathbf{A}}\hat{\mathbf{U}}(t) + \hat{\mathbf{U}}^\dagger(t)\frac{\mathrm{d}\hat{\mathbf{A}}}{\mathrm{d}t}\hat{\mathbf{U}}(t) + \hat{\mathbf{U}}^\dagger(t)\hat{\mathbf{A}}\frac{\mathrm{d}\hat{\mathbf{U}}(t)}{\mathrm{d}t}\middle|\psi_0\right\rangle \\
&= \left\langle\psi_0\middle|\hat{\mathbf{U}}^\dagger(t)\frac{1}{\mathrm{i}\hbar}\left[\hat{\mathbf{A}}, \hat{\mathbf{H}}\right]\hat{\mathbf{U}}(t) + \hat{\mathbf{U}}^\dagger(t)\frac{\mathrm{d}\hat{\mathbf{A}}}{\mathrm{d}t}\hat{\mathbf{U}}(t)\middle|\psi_0\right\rangle = \left\langle\psi(t)\middle|\frac{1}{\mathrm{i}\hbar}\left[\hat{\mathbf{A}}, \hat{\mathbf{H}}\right] + \frac{\mathrm{d}\hat{\mathbf{A}}}{\mathrm{d}t}\middle|\psi(t)\right\rangle\ .
\end{aligned} \tag{2.22}$$

Finally, we get

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\psi(t)\right|\hat{\mathbf{A}}\left|\psi(t)\right\rangle=\left\langle\psi(t)\right|\frac{1}{\mathrm{i}\hbar}\left[\hat{\mathbf{A}},\hat{\mathbf{H}}\right]+\frac{\mathrm{d}\hat{\mathbf{A}}}{\mathrm{d}t}\left|\psi(t)\right\rangle}. \tag{2.23}$$

To clearly demonstrate the correspondence to the classical case, we further rewrite the above result as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\psi(t)\right|A(\hat{\mathbf{q}},\hat{\mathbf{p}},t)\left|\psi(t)\right\rangle=\left\langle\psi(t)\right|\frac{1}{\mathrm{i}\hbar}\left[A(\hat{\mathbf{q}},\hat{\mathbf{p}},t),H(\hat{\mathbf{q}},\hat{\mathbf{p}},t)\right]+\frac{\partial A(\hat{\mathbf{q}},\hat{\mathbf{p}},t)}{\partial t}\left|\psi(t)\right\rangle. \tag{2.24}$$

If you go now through your lecture notes from the course of classical dynamics, you should be able to find a most general equation of motion for an arbitrary dynamical variable $A(q,p,t)$ expressed by using the Poisson brackets:

$$\frac{\mathrm{d}A(q,p,t)}{\mathrm{d}t}=\{A(q,p,t),H(q,p,t)\}+\frac{\partial A(q,p,t)}{\partial t}. \tag{2.25}$$

As you can see when comparing this with Eq. (2.24), those equations are remarkably similar! Actually, Paul Dirac was one of the first people who noted this resemblance. And most importantly, based on this observation, he came up with the following idea: it seems that if we want to translate from the language of classical dynamics into QM, we shall replace

$$\boxed{\{A,B\}\rightarrow\frac{1}{\mathrm{i}\hbar}\left[\hat{\mathbf{A}},\hat{\mathbf{B}}\right]}, \tag{2.26}$$

which actually makes quite a bit of sense, since both the Poisson brackets and the commutator form the Lie algebra.

Based on Eq. (2.26), we can now derive (actually more like postulate) the three fundamental relations that determine the relationships between $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$. In classical mechanics, we have

$$\{q_k,q_l\}=0\,,\quad\{p_k,p_l\}=0\,,\quad\{q_k,p_l\}=\delta_{kl}\,. \tag{2.27}$$

Hence, the QM version of the above statement is that

$$\boxed{[\hat{\mathbf{q}}_k,\hat{\mathbf{q}}_l]=0\,,\quad[\hat{\mathbf{p}}_k,\hat{\mathbf{p}}_l]=0\,,\quad[\hat{\mathbf{q}}_k,\hat{\mathbf{p}}_l]=\mathrm{i}\hbar\delta_{kl}}. \tag{2.28}$$

These are called the *canonical quantization* formulas and they play a key role not just in QM but also in the quantum field theory. In case where we deal only with a 1D problem with $\hat{\mathbf{q}}=\hat{\mathbf{x}}$, we can write a simplified version as

$$\boxed{[\hat{\mathbf{x}},\hat{\mathbf{p}}]=\mathrm{i}\hbar}. \tag{2.29}$$

2.8 Ehrenfest Theorems

Here we make use of the above section to show another nice correspondence between QM and the Newton's second law of motion. We assume a single point particle that can move only along the x axis. Let's start by computing how does $\bar{x}$ change in time. According to (2.23), we have

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = \left\langle \psi(t) \middle| \frac{1}{\mathrm{i}\hbar} \left[\hat{\mathbf{x}}, \hat{\mathbf{H}}\right] \middle| \psi(t) \right\rangle . \tag{2.30}$$

To further expand $\left[\hat{\mathbf{x}}, \hat{\mathbf{H}}\right]$, we make use of the Postulate II. to replace

$$H(x,p) = \frac{p^2}{2m} + V(x) \quad \rightarrow \quad \hat{\mathbf{H}} = H(\hat{\mathbf{x}}, \hat{\mathbf{p}}) = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}}) \,. \tag{2.31}$$

By employing the properties of the commutator and the canonical quantisation formula (2.29) we get

$$\left[\hat{\mathbf{x}}, \hat{\mathbf{H}}\right] = \left[\hat{\mathbf{x}}, \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}})\right] = \frac{1}{2m}\left[\hat{\mathbf{x}}, \hat{\mathbf{p}}^2\right] + \left[\hat{\mathbf{x}}, V(\hat{\mathbf{x}})\right] = \frac{1}{2m}\left(\hat{\mathbf{p}}\left[\hat{\mathbf{x}}, \hat{\mathbf{p}}\right] + \left[\hat{\mathbf{x}}, \hat{\mathbf{p}}\right]\hat{\mathbf{p}}\right) = \frac{\mathrm{i}\hbar}{m}\hat{\mathbf{p}} \,. \tag{2.32}$$

To check that $[\hat{\mathbf{x}}, V(\hat{\mathbf{x}})] = 0$, we expand $V(x)$ by a Taylor series and further replace $x \rightarrow \hat{\mathbf{x}}$, resulting in

$$V(\hat{\mathbf{x}}) = \sum_n c_n \hat{\mathbf{x}}^n \,, \tag{2.33}$$

and since $[\hat{\mathbf{x}}, \hat{\mathbf{x}}^n] = 0$, the original commutator must vanish. After substituting the result of (2.31) into (2.30), we arrive at

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = \frac{\bar{p}}{m} \,, \tag{2.34}$$

which is the *first Ehrenfest theorem*.

Moving on, we calculate

$$\frac{\mathrm{d}\bar{p}}{\mathrm{d}t} = \left\langle \psi(t) \middle| \frac{1}{\mathrm{i}\hbar} \left[\hat{\mathbf{p}}, \hat{\mathbf{H}}\right] \middle| \psi(t) \right\rangle . \tag{2.35}$$

The commutator $\left[\hat{\mathbf{p}}, \hat{\mathbf{H}}\right]$ can be written as

$$\left[\hat{\mathbf{p}}, \hat{\mathbf{H}}\right] = \left[\hat{\mathbf{p}}, \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}})\right] = \left[\hat{\mathbf{p}}, V(\hat{\mathbf{x}})\right] . \tag{2.36}$$

We now show that

$$\left[\hat{\mathbf{p}}, \hat{\mathbf{x}}^n\right] = -\mathrm{i}\hbar n \hat{\mathbf{x}}^{n-1} \quad \text{for} \quad n \geq 0 \,. \tag{2.37}$$

For such purpose, we use the mathematical induction. First, for $n = 0$ we have $\left[\hat{\mathbf{p}}, \hat{\mathbf{1}}\right] = 0$. Second, we assume $[\hat{\mathbf{p}}, \hat{\mathbf{x}}^n] = -\mathrm{i}\hbar n \hat{\mathbf{x}}^{n-1}$ (the induction assumption) and need to show that

$$\left[\hat{\mathbf{p}}, \hat{\mathbf{x}}^{n+1}\right] = -\mathrm{i}\hbar(n+1)\hat{\mathbf{x}}^n \,. \tag{2.38}$$

To do so, we proceed as

$$\left[\hat{\mathbf{p}}, \hat{\mathbf{x}}^{n+1}\right] = [\hat{\mathbf{p}}, \hat{\mathbf{x}}^n \hat{\mathbf{x}}] = \hat{\mathbf{x}}^n [\hat{\mathbf{p}}, \hat{\mathbf{x}}] + [\hat{\mathbf{p}}, \hat{\mathbf{x}}^n] \hat{\mathbf{x}} = -\mathrm{i}\hbar \hat{\mathbf{x}}^n - \mathrm{i}\hbar n \hat{\mathbf{x}}^{n-1} \hat{\mathbf{x}} = -\mathrm{i}\hbar(n+1)\hat{\mathbf{x}}^n \,, \tag{2.39}$$

where in the third equality we have used the induction assumption and the canonical quantization formula. The proof by induction is now complete. It is now left for the reader to show that

$$[\hat{\mathbf{p}}, V(\hat{\mathbf{x}})] = -\mathrm{i}\hbar \frac{\mathrm{d}V(\hat{\mathbf{x}})}{\mathrm{d}x} \,. \tag{2.40}$$

By inserting the above result into (2.35), we arrive at

$$\frac{\mathrm{d}\bar{p}}{\mathrm{d}t} = -\overline{\frac{\mathrm{d}V}{\mathrm{d}x}} \,, \tag{2.41}$$

where we can interpret the term on the RHS as the mean value of the acting force, i.e.,

$$\frac{\mathrm{d}\bar{p}}{\mathrm{d}t} = \bar{F} \,. \tag{2.42}$$

This is the *second Ehrenfest theorem*.

After a careful inspection, we can see that both of the presented theorems are just the QM version of the Newton's second law statement

$$\frac{\mathrm{d}p}{\mathrm{d}t} = F \,, \quad \text{where} \quad p = m\frac{\mathrm{d}x}{\mathrm{d}t} \,. \tag{2.43}$$

Another nice correspondence here. We can now see that the QM when taken on average actually exactly reproduces the results that we know from classical physics.

2.9 Time evolution

We will now explore the time evolution in the most simple case where $\hat{\mathbf{H}}$ does not depend on time (and we will assume this contidion to hold for the rest of this text). First, let us assume an eigenbasis $\{|n\rangle\}$ corresponding to the Hamilton operator (this is mostly the discrete case) as

$$\hat{\mathbf{H}} |n\rangle = E_n |n\rangle \,, \tag{2.44}$$

where (according to the Postulate III.) E_n represent the only measureable values of energy. Further, since $\hat{\mathbf{H}}$ is time independent, the corresponding eigenvectors $|n\rangle$ also do not depend on time. Now we would like to determine how does a system that initially starts at one of the energy eigenstates, i.e.,

$$|\psi_0\rangle = |n\rangle \,, \tag{2.45}$$

evolve in time. We propose that the time evolution is given by

$$|n(t)\rangle = \alpha(t)\,|n\rangle \ , \tag{2.46}$$

where $|n(t)\rangle \equiv |\psi(t)\rangle$. By plugging the above ansatz into the SE, we get

$$\mathrm{i}\hbar\frac{\mathrm{d}\alpha(t)}{\mathrm{d}t}\,|n\rangle = E_n\alpha(t)\,|n\rangle \ . \tag{2.47}$$

Hence, we obtain an ordinary differential equation for the function $\alpha(t)$ as

$$\mathrm{i}\hbar\frac{\mathrm{d}\alpha(t)}{\mathrm{d}t} = E_n\alpha(t)\,, \tag{2.48}$$

whose solution is simply

$$\alpha(t) = \exp\left(-\mathrm{i}\frac{E_n}{\hbar}t\right) \ , \tag{2.49}$$

and therefore,

$$|n(t)\rangle = \exp\left(-\mathrm{i}\frac{E_n}{\hbar}t\right)|n\rangle \ . \tag{2.50}$$

It can be easily shown (and the reader is encouraged to do so) that a general time evolving state Ket-vector can then be expressed as

$$|\psi(t)\rangle = \sum_n \psi_{0n}\,|n(t)\rangle \ , \tag{2.51}$$

where

$$\psi_{0n} = \langle n|\psi_0\rangle \ . \tag{2.52}$$

We can further rearrange this to get

$$|\psi(t)\rangle = \sum_n \exp\left(-\mathrm{i}\frac{E_n}{\hbar}t\right)|n\rangle\,\langle n|\psi_0\rangle = \exp\left(-\mathrm{i}\frac{\hat{\mathbf{H}}}{\hbar}t\right)|\psi_0\rangle \ , \tag{2.53}$$

where in the last equality we used the spectral decomposition of $\hat{\mathbf{H}}$, see Eq. (1.168). Hence, the time evolution of an arbitrary state is simply

$$\boxed{|\psi(t)\rangle = \exp\left(-\mathrm{i}\frac{\hat{\mathbf{H}}}{\hbar}t\right)|\psi_0\rangle}\,. \tag{2.54}$$

Now pause for a while and think about what we have just done. Basically, we have demonstrated that in the case of time-independent Hamiltonian we do not need to care about the time evolution what so ever. This is an amazing simplification, since (as you probably know) to find the time evolution of a studied system is the key quest in the kingdom of classical mechanics.

Finally, note that by comparing Eqs. (2.16) and (2.54), we can identify

$$\hat{\mathbf{U}}(t) = \exp\left(-\mathrm{i}\frac{\hat{\mathbf{H}}}{\hbar}t\right) . \tag{2.55}$$

By employing the Taylor series for $\exp(x)$ it can be shown that in this case

$$\hat{\mathbf{U}}^{\dagger}(t) = \exp\left(\mathrm{i}\frac{\hat{\mathbf{H}}}{\hbar}t\right) \tag{2.56}$$

and

$$\hat{\mathbf{U}}^{\dagger}(t)\hat{\mathbf{U}}(t) = \exp\left(\mathrm{i}\frac{\hat{\mathbf{H}}}{\hbar}t - \mathrm{i}\frac{\hat{\mathbf{H}}}{\hbar}t\right) = \exp(0) = \hat{\mathbf{1}}\,, \tag{2.57}$$

by which we immediately see that $\hat{\mathbf{U}}(t)$ is truly unitary. But be careful, since the rule

$$\exp\left(\hat{\mathbf{A}}\right)\exp\left(\hat{\mathbf{B}}\right) = \exp\left(\hat{\mathbf{A}} + \hat{\mathbf{B}}\right) \tag{2.58}$$

holds only in case where

$$\left[\hat{\mathbf{A}}, \hat{\mathbf{B}}\right] = 0\,, \tag{2.59}$$

which (fortunately) is the case here (we will not prove this).

2.10 Heisenberg Picture

I will just briefly mention here what does the Heisenberg picture mean since you will need it in your future studies of QM and quantum field theory (QFT). The Heisenberg picture is basically just a convenient relabel in Eq. (2.21), where we denote

$$\hat{\mathbf{A}}_{\mathrm{H}} \equiv \hat{\mathbf{U}}^{\dagger}(t)\hat{\mathbf{A}}\hat{\mathbf{U}}(t) \tag{2.60}$$

and call this new operator as the operator $\hat{\mathbf{A}}$ in the Heisenberg picture. We can then simply rewrite Eq. (2.21) as

$$\langle\psi(t)|\,\hat{\mathbf{A}}\,|\psi(t)\rangle = \langle\psi_0|\,\hat{\mathbf{A}}_{\mathrm{H}}\,|\psi_0\rangle\ . \tag{2.61}$$

The LHS then represents the *Schrödinger picture* where the time evolution of the system is included in the state Ket-vector and the RHS represents the *Heisenberg picture* where, on the contrary, the state Ket-vector is time independent and the time evolution is included in the respective operator. Note that there is also a hybrid of the two pictures called the interaction picture, but that is beyond the scope of this text.

2.11 Heisenberg Uncertainty Principle

Last but not least, in this chapter, we will discuss a very important topic regarding the possibility of determining (measuring) two observables on the system at the same time. To be honest, I personally do not like the derivation, since it is quite a tedious one and not very elegant. Nevertheless, the resulting formula is beautiful!

The Postulate IV. tells us that the mean value of a repeated measurement of the observable A on a system in the state $|\psi\rangle$ is equal to $\langle\psi|\,\hat{\mathbf{A}}\,|\psi\rangle$. Now we would like to express a similar formula for the variance of such measurement. The classical formula for variance is given by

$$\sigma_A^2 = \overline{\left(A - \bar{A}\right)^2}\,. \tag{2.62}$$

Hence, the QM version of this must be

$$\sigma_A^2 = \left\langle \psi \middle| \left(\hat{\mathbf{A}} - \bar{A}\right)^2 \middle| \psi \right\rangle\,. \tag{2.63}$$

Assume now that we would like to measure two observables A and B on the system. It would be wise once again to explain what do we mean by measuring repeatedly on a system in the state $|\psi\rangle$ – now with the extension that we are measuring not one, but two different observables. As we have stated before, we need infinite many copies of the studied system, each one prepared in exactly the same state $|\psi\rangle$. Then, on each one of those copies we can conduct only one measurement. So if we measure e.g., the observable A and next we would like to measure B, we still need to take another copy, since the previous one already collapsed to the respective eigenstate of $\hat{\mathbf{A}}$ which is generally different than the original state $|\psi\rangle$. Our task now is to determine the lower bound of the product $\sigma_A^2\sigma_B^2$. First, we denote

$$\Delta\hat{\mathbf{A}} \equiv \hat{\mathbf{A}} - \bar{A}\,. \tag{2.64}$$

Then we proceed as follows

$$\sigma_A^2\sigma_B^2 = \langle\psi|\,\Delta\hat{\mathbf{A}}^2\,|\psi\rangle\,\langle\psi|\,\Delta\hat{\mathbf{B}}^2\,|\psi\rangle = \left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{A}}\psi \right\rangle \left\langle \Delta\hat{\mathbf{B}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle = \left\| \left| \Delta\hat{\mathbf{A}}\psi \right\rangle \right\|^2 \cdot \left\| \left| \Delta\hat{\mathbf{B}}\psi \right\rangle \right\|^2\,, \tag{2.65}$$

while it is trivial to check that $\Delta\hat{\mathbf{A}}^\dagger = \Delta\hat{\mathbf{A}}$. For the next step we need the following lemma:

$$|\,\langle\phi|\psi\rangle\,| \leq \||\phi\rangle\| \cdot \||\psi\rangle\|\,. \tag{2.66}$$

We will not prove this for a general Hilbert space, but at least show the proof for $\mathbb{R}^3$. From the introductory course of Linear Algebra we know that

$$\boldsymbol{a}\cdot\boldsymbol{b} = \|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos(\varphi)\,, \tag{2.67}$$

where φ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$. By applying the absolute value operation on both sides of the above expression and further noting that $0 \leq |\cos(\varphi)| \leq 1$, we immediately get

$$|\boldsymbol{a} \cdot \boldsymbol{b}| \leq \|\boldsymbol{a}\| \|\boldsymbol{b}\| \,, \tag{2.68}$$

which is just a special case of (2.66). We can now continue our calculation as

$$\left\| \left| \Delta\hat{\mathbf{A}}\psi \right\rangle \right\|^2 \cdot \left\| \left| \Delta\hat{\mathbf{B}}\psi \right\rangle \right\|^2 \geq \left| \left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle \right|^2 . \tag{2.69}$$

The next inequality we are going to need is as follows:

$$|z|^2 = \mathrm{Re}[z]^2 + \mathrm{Im}[z]^2 \geq \mathrm{Im}[z]^2 = \left[\frac{1}{2\mathrm{i}} (z - z^*) \right]^2 , \tag{2.70}$$

where $z \in \mathbb{C}$. By assigning

$$z = \left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle , \tag{2.71}$$

we proceed further as

$$\begin{aligned} \left| \left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle \right|^2 \geq \left[\frac{1}{2\mathrm{i}} \left(\left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle - \left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle^* \right) \right]^2 = \\ \left[\frac{1}{2\mathrm{i}} \left(\left\langle \Delta\hat{\mathbf{A}}\psi \middle| \Delta\hat{\mathbf{B}}\psi \right\rangle - \left\langle \Delta\hat{\mathbf{B}}\psi \middle| \Delta\hat{\mathbf{A}}\psi \right\rangle \right) \right]^2 = \left[\frac{1}{2\mathrm{i}} \left\langle \psi \middle| \left[\Delta\hat{\mathbf{A}}, \Delta\hat{\mathbf{B}} \right] \middle| \psi \right\rangle \right]^2 . \end{aligned} \tag{2.72}$$

It is left for the reader to check that

$$\left[\Delta\hat{\mathbf{A}}, \Delta\hat{\mathbf{B}} \right] = \left[\hat{\mathbf{A}}, \hat{\mathbf{B}} \right] . \tag{2.73}$$

Finally, by employing the above equality, equating the LHS of (2.65) with the RHS of (2.72) and taking the square root, we arrive at

$$\boxed{\sigma_A \sigma_B \geq \left| \frac{1}{2\mathrm{i}} \left\langle \psi \middle| \left[\hat{\mathbf{A}}, \hat{\mathbf{B}} \right] \middle| \psi \right\rangle \right|}, \tag{2.74}$$

which is the famous formula for the so called *Heisenberg uncertainity principle*.

The derivation process (even though quite tedious) should be pretty straightforward and clear. But now comes the trickier part: we need to interpret the result somehow. First, refer back to the Postulate IV., where we discovered that the outputs of a measurement on a quantum system are stochastic – i.e., each possible result has a certain probability of happening. But there is actually one singular case where this is not true. Assume (without the loss of generality we show this for the discrete case) that the system is in the k-th eigenstate of $\hat{\mathbf{A}}$:

$$|\psi\rangle = |k\rangle \,. \tag{2.75}$$

By employing Eq. (2.10), we get

$$w_k = \langle k|\psi\rangle = |\,\langle k|k\rangle\,|^2 = 1\,. \tag{2.76}$$

This means that if we measure A on the system, the output will be a_k. By employing the concept of repeated measurement on the system (we prepare infinitely many identical copies, each in the state $|k\rangle$), the mean value becomes

$$\bar{A} = a_k \tag{2.77}$$

and the corresponding standard deviation (the square root of the variance) is equal to

$$\sigma_A = 0\,. \tag{2.78}$$

In such case, we can say that the system has $A = a_k$, since we know that the measurement of A must always result in a_k. So e.g., when the studied system is an electron in a presence of a central Coulomb field and the electron is in the eigenstate $|k\rangle$ of $\hat{\mathbf{H}}$, we say that the electron has energy $E = E_k$. But if the electron is in a general superposition of the eigenstates, such a statement does not make sense anymore.

Now we ask ourselves the following question: is there a possibility for a specific state to exist to which we can assign two values of two distinct observables A and B? Or in other words a state for which

$$\sigma_A = \sigma_B = 0\,? \tag{2.79}$$

The answer is provided by the Heisenberg uncertainity formula (2.74), which tells us that in such case the following condition must hold:

$$0 \geq \left|\frac{1}{2\mathrm{i}}\left\langle\psi\right|\left[\hat{\mathbf{A}},\hat{\mathbf{B}}\right]\left|\psi\right\rangle\right|\,, \tag{2.80}$$

which is true only if $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ commute, i.e.,

$$\left[\hat{\mathbf{A}},\hat{\mathbf{B}}\right] = 0\,. \tag{2.81}$$

Same as before, assume $|k\rangle$ being an eigenvector of $\hat{\mathbf{A}}$:

$$\hat{\mathbf{A}}\,|k\rangle = a_k\,|k\rangle\ . \tag{2.82}$$

Now we act with $\hat{\mathbf{B}}$ on both sides:

$$\hat{\mathbf{B}}\hat{\mathbf{A}}\,|k\rangle = a_k\hat{\mathbf{B}}\,|k\rangle\ . \tag{2.83}$$

Since $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ commute, we can rearrange this as

$$\hat{\mathbf{A}}\left(\hat{\mathbf{B}}\,|k\rangle\right) = a_k\left(\hat{\mathbf{B}}\,|k\rangle\right)\,, \tag{2.84}$$

which implies that

$$\hat{\mathbf{B}}\,|k\rangle \propto |k\rangle \ . \tag{2.85}$$

So by acting with $\hat{\mathbf{B}}$ onto $|k\rangle$, we get something proportional to $|k\rangle$. But this is exactly the definition of an eigenvector of $\hat{\mathbf{B}}$. Hence, by identifying the proportionality constant with the eigenvalue b_k, we have

$$\hat{\mathbf{B}}\,|k\rangle = b_k\,|k\rangle \ . \tag{2.86}$$

We have therefore proven that if the two operators commute, they have the same set of all eigenvectors! (Actually, we have just shown that if $|k\rangle$ is an eigenvector of $\hat{\mathbf{A}}$, it must be also an eigenvector of $\hat{\mathbf{B}}$. But since the whole argument can be reversed, i.e., we start with an eigenvector of $\hat{\mathbf{B}}$ and show that it must be also an eigenvector of $\hat{\mathbf{A}}$, then the statement must be true.) Note that regardless of the same set of eigenvectors, the corresponding eigenvalues are different for each operator ($a_k \neq b_k$).

Now let's demonstrate how does this explain our previous question. Imagine that the system is prepared in one of the eigenstates $|k\rangle$ of $\hat{\mathbf{A}}$. This means that if we would measure A on the system, we will get a_k. But $|k\rangle$ is also an eigenstate of $\hat{\mathbf{B}}$, so if we measure B instead, we would get b_k. And for that reason we can say that the system has simultaneously $A = a_k$ and $B = b_k$. Once again, such a statement would not make sense anymore if the system would be in a general superposition of eigenstates.

To conclude this section, we make use of the standard nomenclature and (from now on) denote any two observables that commute as *compatible*. And note that there does not need to be just two of those. E.g., when we solve the previously mentioned system constituting of an electron in a spherically symmetric Coulomb field, there will be three compatible observables, namely the energy, the magnitude of the angular momentum and the z-component of the angular momentum.

2.12 Summary

Now that we have covered all the postulates and discovered the canonical quantization formula, we are finally ready to move on to the famous "wavefunction representation" of QM.

Chapter 3

Continuous Representation of QM

This chapter is devoted to the formulation of QM in the continuous x-representation. To make things easier, we assume a simple system consisting of a point particle (or – to be more precise – something that we would call a point particle in the context of classical mechanics) moving in one dimension that would be described by one coordinate x with the conjugate momentum p in the classical case. The extension to 3D would then be done just heuristically.

3.1 Position Operator & Wavefunction

We start by employing the Postulate II. and assume the so called *position operator* $\hat{\mathbf{x}}$ corresponding to the coordinate x. Let's continue by formulating the eigenvalue problem for such operator as

$$\hat{\mathbf{x}}\,|x\rangle = x\,|x\rangle\ . \tag{3.1}$$

Now first of all, since $\hat{\mathbf{x}}$ is hermitian, $x \in \mathbb{R}$ and based on the Postulate III. it represents all the possible outcomes when measuring the position of the particle. The Postulate V. further tells us that whenever we measure the position to be x, no matter what the initial state was, the system collapses to $|x\rangle$. Hence, the eigenstate $|x\rangle$ has the interpretation that the particle is located at x. Moreover, once again due to $\hat{\mathbf{x}}$ being hermitian, $\langle x|x'\rangle = \delta(x-x')$ (the eigenbasis can always be normalized, see the discussion in Chapter 1), and therefore we can use $\{|x\rangle\}$ as an appropriate basis to express any general state as

$$|\psi\rangle = \int \mathrm{d}x\,\psi(x)\,|x\rangle\ . \tag{3.2}$$

The function

$$\psi(x) = \langle x|\psi\rangle \tag{3.3}$$

is then called the *wavefunction* of the particle. If we now have a particle in a general state $|\psi\rangle$ then according to the Postulate IV.,

$$w(x) = |\psi(x)|^2 \tag{3.4}$$

represents the probability density that the particle can be found at x. Please note that this is something completely different from the classical mechanics, where the particle's position is strictly defined whereas in QM the position of a particle is not deterministic. This is sometimes referred to as the *non-locality*. Given the state $|\psi\rangle$, we can only determine the probability that the particle can be found in the range $\langle a, b\rangle$ as

$$P[x \in \langle a, b\rangle] = \int_a^b \mathrm{d}x \, |\psi(x)|^2 \,. \tag{3.5}$$

Moving on, by using the formula (1.162) we can express the position operator as

$$\hat{\mathbf{x}} = \int \mathrm{d}x \, x \, |x\rangle \, \langle x| \,. \tag{3.6}$$

Finally, based on the formulae given in Tab. 1.10 we will now use the continuous representation as

$$\boxed{|\psi\rangle \to \psi(x) \,, \quad \hat{\mathbf{x}} \to x} \,, \tag{3.7}$$

hereinafter referred to as the *x-representation*.

3.2 Momentum Operator

Now that we have derived the x-representation for a state Ket-vector (the wavefunction) and the position operator, it is time to determine the form and the corresponding representation for the (conjugate) *momentum operator*. And the recipe of how to do so lies in the canonical quantisation formula presented in the previous chapter. Recall that the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ must satisfy

$$[\hat{\mathbf{x}}, \hat{\mathbf{p}}] = \mathrm{i}\hbar \,. \tag{3.8}$$

Further, $\hat{\mathbf{p}}$ can be expressed in terms of the eigenbasis $\{|x\rangle\}$ as

$$\hat{\mathbf{p}} = \int \mathrm{d}x \, \mathrm{d}y \, p(x, y) \, |x\rangle \, \langle y| \,. \tag{3.9}$$

We will now show that the correct expression for $p(x, y)$ is

$$p(x, y) = -\mathrm{i}\hbar \lim_{\varepsilon \to 0} \frac{\delta(y - (x + \varepsilon)) - \delta(y - x)}{\varepsilon} \,. \tag{3.10}$$

We start by investigating how does $\hat{\mathbf{p}}$ act on $|\psi\rangle$. This can be written as

$$\hat{\mathbf{p}} \, |\psi\rangle = -\mathrm{i}\hbar \int \mathrm{d}x \, \mathrm{d}y \, \lim_{\varepsilon \to 0} \frac{\delta(y - (x + \varepsilon)) - \delta(y - x)}{\varepsilon} \, |x\rangle \, \langle y| \int \mathrm{d}z \, \psi(z) \, |z\rangle \,. \tag{3.11}$$

After a bit of algebra we get

$$\hat{\mathbf{p}} \, |\psi\rangle = -\mathrm{i}\hbar \int \mathrm{d}x \, \lim_{\varepsilon \to 0} \frac{\psi(x + \varepsilon) - \psi(x)}{\varepsilon} = \int \mathrm{d}x \, \left(-\mathrm{i}\hbar \frac{\partial \psi(x)}{\partial x} \right) |x\rangle \,. \tag{3.12}$$

The partial differentiation symbol is used since the wavefunction depends also on time, which is not denoted explicitly here for the sake of readability. The continuous representation of $\hat{\mathbf{p}}$ in the eigenbasis $\{|x\rangle\}$ is therefore

$$\boxed{\hat{\mathbf{p}} \to -\mathrm{i}\hbar\frac{\partial}{\partial x}}\,. \tag{3.13}$$

Now let's calculate the commutator between $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$.

$$[\hat{\mathbf{x}}, \hat{\mathbf{p}}]\,|\psi\rangle = \int \mathrm{d}x\, x \left(-\mathrm{i}\hbar\frac{\partial\psi(x)}{\partial x}\right)|x\rangle - \int \mathrm{d}x \left(-\mathrm{i}\hbar\frac{\partial[x\psi(x)]}{\partial x}\right)|x\rangle = \mathrm{i}\hbar\,|\psi\rangle\ . \tag{3.14}$$

By comparing the LHS with the RHS we get $[\hat{\mathbf{x}}, \hat{\mathbf{p}}] = \mathrm{i}\hbar$, which is exactly what we wanted to achieve.

3.3 Schrödinger equation in the continuous representation

Being equipped with the continuous representation for $\hat{\mathbf{x}}$, $\hat{\mathbf{p}}$ and $|\psi\rangle$ we can now work out the form of the corresponding form of the Schrödinger equation. We start with the Hamilton operator $\hat{\mathbf{H}}$. In classical non-relativistic mechanics of one particle the Hamiltonian is given as

$$H(x,p) = \frac{p^2}{2m} + V(x)\,. \tag{3.15}$$

According to the second postulate, the corresponding Hamilton operator then must possess the following form

$$\hat{\mathbf{H}} = H(\hat{\mathbf{x}}, \hat{\mathbf{p}}) = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}})\,, \tag{3.16}$$

where m stands for the mass of the particle. By using the formulae derived in the sections above we can immediately write down the continuous representation of the Hamilton operator as

$$\boxed{\hat{\mathbf{H}} \to -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)}\,. \tag{3.17}$$

Now let's focus on the LHS of the Schrödinger equation. We have

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\psi\rangle = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}x\,\psi(t,x)\,|x\rangle = \int \mathrm{d}x \left(\mathrm{i}\hbar\frac{\partial\psi(t,x)}{\partial t}\right)|x\rangle\ , \tag{3.18}$$

where we have now explicitly denoted the time dependence of the wavefunction $\psi(t,x)$. By combining the above expressions we can now write down the Schrödinger equation in the $\{|x\rangle\}$ representation as

$$\boxed{\mathrm{i}\hbar\frac{\partial\psi(t,x)}{\partial t} = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(t,x)}, \tag{3.19}$$

which is the form you will most certainly find in most of the introductory textbooks on QM.

3.4 Time evolution once again

As we have demonstrated in the previous chapter, since in this introductory text we will deal only with time independent hamiltonians, we do not need to care about the Schrödinger equation anymore as the time evolution is given simply by Eq. (2.54). However, this expression - even though being beautiful mathematically - is not particularly useful for real calculations, since (at least to my knowledge) you cannot insert an operator including derivatives into exponential function and assume it to produce some results in a typical Computer Algebra System (CAS) such as Maple or Mathematica. Hence, the expression that you would use in a real life scenario is

$$\boxed{\psi(x,t) = \sum_n \exp\left(-\mathrm{i}\frac{E_n}{\hbar}t\right)\psi_n(x)}, \tag{3.20}$$

where $\psi_n(x)$ stand for the (time independent) eigenvectors of the Hamilton operator in the x-representation satisfying

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi_n(x) = E_n\psi_n(x)\,. \tag{3.21}$$

Note that $\psi_n(x)$ is now a single variable function of x and hence we should use the total derivative symbol instead of the partial one. Nevertheless, we will stick with the partial differentiation, since in more complicated problems involving 3D and possibly more than one particle, the wavefunction will include additional coordinates. This is now a perfectly valid formula that any CAS should be able to work with.

I could have ended this section right here, since the formula (3.20) is basically all we need for the purpose of this text. However, I would like to comment on the nomenclature regarding the Schrödinger equation. Sometimes, Eq. (3.19) is referred to as the *time dependent Schrödinger equation* due to the presence of time. A standard method for solving such partial differential equation is to assume a time harmonic solution (a general time dependence can then be obtained by applying the Fourier transform) in the form

$$\psi(t,x) = \psi(x)\exp(-\mathrm{i}\omega t)\,, \tag{3.22}$$

which is a little bit confusing, since we have used the same symbol ψ for the two different functions (but at this point, I believe, you should be used to all sort of notation inconsistencies). By inserting this into (3.19) we obtain

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(x) = E\psi(x)\,, \tag{3.23}$$

where we have identified $E = \hbar\omega$. Eq. (3.23) is then sometimes called the *time independent Schrödinger equation*, as it no longer includes time. Now here comes the problem (or at least my personal problem) with this designation. In real life, people like to abbreviate long terms. So it is typical that both of Eqs. (3.19) and (3.23) are simply called

the Schrödinger equation. But even though one could be technically derived from the other (as we have shown above), they have completely different meanings! The original Schrödinger equation represents a QM version of the equation of motion, while the time independent one is actually just the x-representation of the eigenvalue problem

$$\hat{\mathbf{H}}\,|\psi\rangle = E\,|\psi\rangle \tag{3.24}$$

which we could have written right away completely without any relevance to the time dependent form. Hence, in this text the only equation that is going to be referred to as the Schrödinger equation will be the one giving the time evolution.

3.5 Generalization to 3D

Now we want to extend the above formulations for the case where the point particle exists in 3D. Such generalization is pretty simple. First, we define the three position operators corresponding to the coordinates x, y, z as

$$\hat{\mathbf{x}} \to x\,, \quad \hat{\mathbf{y}} \to y\,, \quad \hat{\mathbf{z}} \to z\,. \tag{3.25}$$

Then, based on Eq. (3.13), we define the conjugate momenta operators as

$$\hat{\mathbf{p}}_x \to -\mathrm{i}\hbar\frac{\partial}{\partial x}\,, \quad \hat{\mathbf{p}}_y \to -\mathrm{i}\hbar\frac{\partial}{\partial y}\,, \quad \hat{\mathbf{p}}_z \to -\mathrm{i}\hbar\frac{\partial}{\partial z}\,. \tag{3.26}$$

And this is basically all we need. E.g., the Hamiltonian in the 3D case can be written as

$$H = \frac{p^2}{2m} + V(x,y,z) = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + V(x,y,z)\,. \tag{3.27}$$

According to the Postulate II., this now translates into

$$\hat{\mathbf{H}} = \frac{\hat{\mathbf{p}}_x^2 + \hat{\mathbf{p}}_y^2 + \hat{\mathbf{p}}_z^2}{2m} + V(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{z}})\,. \tag{3.28}$$

By employing the above expressions (3.25), (3.26), we arrive at

$$\hat{\mathbf{H}} \to -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(x,y,z) = -\frac{\hbar^2}{2m}\nabla^2 + V(x,y,z)\,. \tag{3.29}$$

Thus, the Schrödinger equation can be formulated as

$$\boxed{\mathrm{i}\hbar\frac{\partial\psi(t,x)}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(x)\right]\psi(t,x)}\,, \tag{3.30}$$

while the equation for determining the eigenstates/eigenvalues of $\hat{\mathbf{H}}$ is

$$\boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + V(x)\right]\psi(x) = E\psi(x)}\,. \tag{3.31}$$

By this procedure, it is then straightforward to extend all other operators into the 3D formalism.

3.6 Summary

It is finally here! At this point, we have covered all the necessary theoretical basics and are now ready to get into some simple introductory problems.

Chapter 4

Simple Examples