Notes for “Understanding quantum mechanics 3”

Edwin Steiner

May 2021

1 Introduction

These are the notes for both parts of episode 3 of the series “Understanding quantum mechanics”.

2 WARNING: Notes in progress!

Eventually these notes will cover both parts of episode 3. As the second part of the video is not done at the time of this writing, this document is incomplete, too. Please expect changes at least until the second part of episode 3 has been published.

3 Remarks about the independence of quantum mechanics from classical physics

I claim that quantum mechanics does not depend on classical physics in its formulation. I am aware that there are expositions of quantum mechanics claiming the direct opposite of this statement. These expositions emphasize, for example, that a measurement apparatus, which may be needed to give operational meaning to a particular physical quantity, must be described in terms of “classical physics” in order to be able to talk about observable quantities at all.

My position is that these expositions try to express something true but by a very unfortunate choice of language. Here is my take one the same conceptual point: The formulation of quantum mechanics requires us to speak of the *known* as well as of the *unknown*. Observations are made at the interface[1] of the *known* with the *unknown*. Possibilities of unknown quantities become actual values of known quantities by observation. In order to define more and more refined observables, we need to rely on some things to be *known*, for example the configuration of our measurement apparatus. (Otherwise we would have to include the apparatus as part of the *a priori* unknown physical system we are analyzing.) This assumption of the apparatus as something *known* is what some expositions of quantum mechanics express by stating that the apparatus is defined “classically”.

To summarize, the language used by expositions claiming a dependence of quantum mechanics on classical physics can be translated to the language I use in this series by making the replacements:

classical	$\rightarrow$	known,
quantum	$\rightarrow$	unknown.

This turns the statement about the dependence on classical physics into the true statement that in order to define specific observables we must rely on some known physics and on known configurations of a measurement apparatus (as long as we are not willing to treat it as part of the system under analysis).

4 Remarks about physical information in the wave function

In the video, I claim that there is no physical information in an *individual* wave function beyond the recording of the observational result. This remark is somewhat provocative, with the intention of shifting the listeners’ focus from the wave function (which is a basis-dependent notion) to the mutual relationships

[1] The definition of this interface is often called the “Heisenberg cut”.

of the eigenbases of different observables and ultimately (in part 2) to the algebraic relations of these observables with each other as the main structure of the physical description.

The remark can be defended but maybe not to your full satisfaction, and that is fair enough, because introducing the wave function really does not make a lot of sense unless one also introduces the full Hilbert space structure, so it is questionable whether an *individual* wave function is even something we can or should discuss in isolation.

What I meant is the following: Suppose you are observing an electron in an atom and you measure its energy.[2] Given the result, you can write down the wave function of the electron *in the energy basis*, i.e. in an eigenbasis of the Hamiltonian operator. *However*, as long as you do not know the expression of the Hamiltonian in terms of other operators, like the position and its conjugate momentum, this wave function tells you nothing except what you have measured: the energy (difference). That's what I meant with the knowledge encoded by an *individual* wave function.

The situation is different, if you know the full algebra of observables on the Hilbert space, including relations like the expression of the Hamiltonian in terms of position and momentum and the canonical commutation relation between position and momentum. With this complete knowledge of the Hilbert space structure and the algebra of observables, you can find representations of the observables on the Hilbert space and consequently you can convert wave functions back and forth between arbitrary bases. *Then* you can convert the measured wave function in the energy eigenbasis to the position eigenbasis, say, and you can read off physical information like the spatial structure of the electron shell, etc.

What I wanted to emphasize, is that this information is primarily encoded in the mutual relations of operators, not in the coordinates of vectors, even if you can read it off the coordinates of a particular vector once you know how to convert between all the interesting bases in question.

5 Remarks about the distinction between physical and mathematical concepts

Before we dive into technical details, I want to clarify a point that might not have been formulated as clearly as possible in the video. I make the following two claims in the video, which might at first sight appear contradictory:

1. It is important to distinguish between mathematical and physical concepts.
2. There is no meaningful conceptual distinction between (maximal) knowledge and reality in a fundamental physical theory.

One might see some tension between these two points, because one usually makes the following rough associations:

mathematical	∼	knowledge
physical	∼	reality

If we turned these associations into exact replacements, the above two points would indeed directly contradict each other. However, the insights of quantum mechanics teach us to reconsider these conceptual associations, especially the latter one between the physical world and reality.

Here is a sketch of associations which are better suited for a conceptual understanding of quantum mechanics:

mathematical	∼	knowledge	∼	actual, reality	∼	observed
physical	∼	not fully knowable	∼	possible, not fully real(ized)	∼	observable

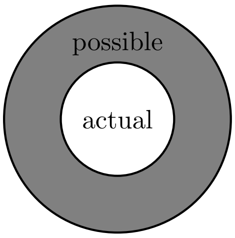
Note that in the context of this discussion, we define "reality" as *that which can be consistently and reproducibly observed*, which is, as far as I can tell, the only useful way to define reality in a scientific context.

Importantly, these two lines of associated concepts are not disjoint. Rather the second contains the first[3]. This inclusion is represented in quantum mechanics in the fact that sets of commuting observables are

[2] Practically, you can only measure an energy difference, say the difference to a known ground state.

[3] I avoid the terms subset/superset here, because—as explained in the video—the totality of physical possibilities cannot be regarded as a set in the mathematical sense.

embedded in larger, non-commutative algebra of observables.[4] What makes quantum mechanics subtle and difficult to teach is that there is no canonical choice of a maximal set of commuting observables from the full algebra.



6 Mathematical prerequisites: Linear algebra (mostly)

In order to truly understand quantum mechanics, you need to learn some mathematics. The good news is that a basic, correct understanding does not require a whole lot of mathematical knowledge[5], although it does require a bit more than what is usually taught in high school. Most importantly, you need to understand complex numbers and the basic building blocks of linear algebra: vector spaces, bases, matrices and linear maps between vector spaces, eigenvalues, eigenvectors, and the inner product.

I will not try to replicate this material here, because any dedicated introduction to linear algebra will do a better job of that than I could do in a few paragraphs. If you want to know more, I recommend "Linear Algebra Done Right"[1] by Sheldon Axler. Of course, you can also find introductions to linear algebra on YouTube, for example by **3Blue1Brown**[2].

You will also need some basic knowledge of probability theory, including:

- The axiomatic definition of probabilities as number in the interval $[0, 1]$,
- the rules for addition and multiplication of probabilities to express logical OR and AND relations between independent events,
- the concept of a probability distribution,
- the definition of expectation values (and possibly those of higher-order moments of probability distributions).

Notation and terms

We denote *a priori* unknown physical quantities, i.e. "observables" and the operators representing them by upper-case latin letters (X, P, A, B, $\dots$). Observational results, i.e. eigenvalues of an observable are denoted by the corresponding lower-case letter (x, p, a, b, $\dots$).

Time-dependence of an observable X will sometimes be indicated by writing $X(t)$ explicitly while in other cases it will be suppressed to make the notation less cumbersome.

We use an upper-case greek letter (Φ, Ψ, $\dots$) to denote the abstract Hilbert space vector represented by the wave function denoted by the corresponding lower-case greek letter (ϕ, ψ, $\dots$). For example the vector $|\Psi\rangle \in \mathscr{H}$ is the vector with wave function ψ with respect to the chosen basis. The basis is implied by the context or by the argument to the wave function, as in $\psi(x)$ for a wave function referring to the position basis.

Usually vectors will be written as "ket"-symbols ($|\Phi\rangle$, $|\Psi\rangle$, $\dots$) but in some cases, for example in the arguments of the inner product, when writing norms of vectors, and in the subscripts of Born's rule, the ket notation is suppressed and only the upper-case greek letter is used in order to make the notation less cumbersome.

[4] At the level of human language, the same inclusion is expressed in the rather trivial fact that the *actual* is necessarily also *possible*.

[5] It's a different story if you want to apply quantum mechanics to solve physical problems on your own. For that you will also need to know advanced calculus, how to solve differential equations, etc.

For the purposes of this document (and the video), the term *basis* always refers to an orthonormal basis which is also a common eigenbasis of a complete set of compatible observables. We will have no need to deal with more general bases of the Hilbert space and we will therefore drop the above adjectives and assume them implicitly for the bases we use.

Furthermore, we use the following notations:

$\overline{z}$ denotes the complex conjugate of z,
$\Im(z)$ denotes the imaginary part of z,
$(\ldots \pm \text{c.c.})$ denotes addition/subtraction of the complex conjugate of the preceeding term,
$\mathrm{Tr}(A)$ denotes the trace of a matrix or operator A,
$P_{x|\psi}$ is the probability that given the wave function ψ with respect to an eigenbasis of the observable X we observe the value of X to be x (the observable X will be clear from the contexts in which this notation is used),
$P_{\Phi|\Psi}$ is the probability that given the knowledge represented by the vector $|\Psi\rangle$ we observe a physical situation represented by the vector $|\Phi\rangle$,
$E(A)|_\Psi$ is the expectation value of the observable A given the knowledge represented by the vector $|\Psi\rangle$.

7 Properties of the inner product

The inner product $\langle \Phi \mid \Psi \rangle$ on a Hilbert space $\mathscr{H}$ has the following properties:

1. **Hermitian symmetry:** For any two vectors $|\Phi\rangle, |\Psi\rangle \in \mathscr{H}$, we have:

$$\langle \Psi \mid \Phi \rangle = \overline{\langle \Phi \mid \Psi \rangle},$$

   i.e. exchanging the arguments of the inner product complex-conjugates its value.

2. **Linearity in the right argument:** For any vectors $|\Phi\rangle, |\Psi_1\rangle, |\Psi_2\rangle \in \mathscr{H}$ and scalars $\lambda_1, \lambda_2 \in \mathbb{C}$, we have:

$$\langle \Phi \mid \lambda_1 \Psi_1 + \lambda_2 \Psi_2 \rangle = \lambda_1 \langle \Phi \mid \Psi_1 \rangle + \lambda_2 \langle \Phi \mid \Psi_2 \rangle .$$

3. **Positive definiteness:** This is a combination of **positivity:** For any vector $|\Psi\rangle \in \mathscr{H}$, we have $\langle \Psi \mid \Psi \rangle \geq 0$ and **definiteness:** $(\langle \Psi \mid \Psi \rangle = 0) \Longrightarrow (\Psi = 0)$.

These properties have many immediate and important consequences in quantum mechanics: The Hermitian symmetry implies a **symmetry of Born's rule**:

$$P_{\Phi|\Psi} = |\langle \Phi \mid \Psi \rangle|^2 = |\langle \Psi \mid \Phi \rangle|^2 = P_{\Psi|\Phi}.$$

It also implies that the inner product of any vector with itself is a real number, which, together with positive definiteness, allows us to induce a norm using this inner product.

Together with the Hermitian symmetry, linearity in the right argument implies anti-linearity in the left argument. (Both linearity properties together are referred to as *sesquilinearity*, "one and a half"-linearity.) Linearity is crucial for the probabilistic interpretation of quantum mechanics. For example, it implies that Pythagoras's theorem holds for orthogonal vectors, which is necessary for the interpretation of the addition of orthogonal vectors as corresponding to a logical "OR" of mutually exclusive alternatives.

It is a standard fact in linear algebra, the proof of which you can find in any comprehensive textbook, that an inner product with these properties can always be brought into the standard form shown in the video by choosing a suitable (orthonormal) basis. Search for the "Gram-Schmidt process" for orthonormalization for a constructive proof of this fact.

8 Equivalent forms of Born's rule

In the course of episode 3 we encounter three different forms of Born's rule:

1. $P_{x|\psi} = |\psi(x)|^2$,
2. $P_{\Phi|\Psi} = |\langle \Phi \mid \Psi \rangle|^2$,

3. $E(A)|_\Psi = \langle \Psi \,|\, A \,|\, \Psi \rangle$.

We will now show that all these forms are equivalent.

Implication 1 ⇒ 2. Given two unit vectors $|\Phi\rangle$, $|\Psi\rangle$, we choose an observable A such that $A\,|\Phi\rangle = 0$ and all eigenvalues a of A have multiplicity[6] 1. We choose an eigenbasis $\{|\Phi_a\rangle\}$ of A such that $|\Phi_a\rangle$ is an eigenvector of A with eigenvalue a. We note that according to our choices, $|\Phi_0\rangle = |\Phi\rangle$, i.e. $|\Phi\rangle$ represents the knowledge that the observable A has the value 0. The vector $|\Psi\rangle$ can be expressed in this basis as

$$|\Psi\rangle = \sum_a \psi(a)\,|\Phi_a\rangle\,.$$

for a suitable wave function $\psi(a)$. The wave function $\psi(a)$ can then be written as

$$\langle \Phi_a|\Psi\rangle = \langle \Phi_a|\,(\sum_{a'} \psi(a')\,|\Phi_{a'}\rangle) = \sum_{a'} \psi(a')\,\langle \Phi_a|\Phi_{a'}\rangle = \psi(a).$$

The probability $P_{\Phi|\Psi}$ that given the knowledge $|\Psi\rangle$ we observe the physical situation represented by $|\Phi\rangle$ is equal to the probability that we observe the value 0 for the observable A:

$$P_{\Phi|\Psi} = P_{0|\psi} = |\psi(0)|^2 = |\,\langle \Phi_0|\Psi\rangle\,|^2 = |\,\langle \Phi|\Psi\rangle\,|^2.$$

Implication 2 ⇒ 3. For an index set I, let $\{|\Phi_i\rangle\}_{i\in I}$ be an eigenbasis of the observable A such that $|\Phi_i\rangle$ is an eigenvector of A with eigenvalue a_i. $|\Phi_i\rangle$ represents certain knowledge that the value of A is a_i (and possibly additional knowledge about observables compatible with A). The operator A is diagonal with respect to this basis and can be written as

$$A = \sum_{i\in I} |\Phi_i\rangle\, a_i\, \langle \Phi_i|\,.$$

(For continuous parts of the spectrum of A, this sum shall be interpreted as the appropriate integral. We will not make this distinction explicit here.) The expectation value $E(A)|_\Psi$ is defined as

$$E(A)|_\Psi = \sum_{i\in I} a_i P_{\Phi_i|\Psi} = \sum_{i\in I} a_i|\,\langle \Phi_i\,|\,\Psi\rangle\,|^2 = \sum_{i\in I} \langle \Psi\,|\,\Phi_i\rangle\, a_i\,\langle \Phi_i\,|\,\Psi\rangle = \langle \Psi|(\sum_{i\in I} |\Phi_i\rangle\, a_i\,\langle \Phi_i|)|\Psi\rangle = \langle \Psi\,|\,A\,|\,\Psi\rangle\,.$$

Implication 3 ⇒ 1. For brevity, we will show this implication for the case of all eigenvalues of the observable X having multiplicity 1. (If this is not the case, the following proof can be generalized by adding compatible observables in order to get a complete set of compatible observables, conducting the analogous proof, and finally summing over the values of observables one is not interested in).

Let $\{|x\rangle\}$ be an eigenbasis of the observable X such that $|x\rangle$ is an eigenvector of X with eigenvalue x. We define the projection operator[7]

$$P_x = |x\rangle\,\langle x|\,.$$

The operator P_x represents an observable that has the value 1 if $X = x$ and zero otherwise. Given the wave function $\psi(x)$ with respect to the basis $\{|x\rangle\}$, we define the vector $|\Psi\rangle = \sum_x \psi(x)\,|x\rangle$. The wave function $\psi(x)$ can then be written as

$$\langle x|\Psi\rangle = \langle x|\,(\sum_{x'} \psi(x')\,|x'\rangle) = \sum_{x'} \psi(x')\,\langle x|x'\rangle = \psi(x).$$

We have

$$P_{x|\psi} = E(P_x)|_\Psi = \langle \Psi|P_x|\Psi\rangle = \langle \Psi|(|x\rangle\,\langle x|)|\Psi\rangle = \langle \Psi|x\rangle\,\langle x|\Psi\rangle = |\,\langle x|\Psi\rangle\,|^2 = |\psi(x)|^2.$$

References

[1] Axler, S., *Linear Algebra Done Right*, Springer, 3rd edition, 2015, ISBN-13: 978-3319110790.

[2] 3Blue1Brown, *Linear Alg*,
`https://www.youtube.com/playlist?list=PLO-GT3co4r2y2YErbmuJw2L5tW4Ew2O5B`.

[6] As observables are required to be represented by diagonalizable operators, the notion of geometric and algebraic multiplicity coincide.

[7] A *projection operator* is a hermitian operator P that is *idempotent*, i.e. $P^2 = P$.