6 Measurements and uncertainties

6.1 Probabilities and measurements

In many respects, Quantum Mechanics is a mere mathematical algorithm for predicting the outcome of experiments. (Remember from the first lecture that the word "experiment" should be understood in a broad sense, here: e.g., detecting the arrival of a bunch of photons with the naked eye is an experiment. Remember, also, that a theoretical prediction is what a theory says about the results of an experiment, whether this experiment will be done in the future, or has been done in the past, or is a thought experiment which will never be done.)

In practice, a well controlled experiment would typically consist in preparing a system in a certain state (e.g., exciting an atom in a certain way), making a measurement or some test on this system (e.g., testing whether the atom is in a state of spin up or a state of spin down), and recording the result of this experiment (a number, or a yes/no result). From a Quantum Theory standpoint, the initial state of this system would typically be described by a vector belonging to a certain Hilbert space,[1] the measurement would be described by a theoretical model formulated in terms of operators acting in that Hilbert space, and the theoretical predictions about the outcome of the measurement would normally be probabilistic. There is certainty about a specific outcome only if the probability of this outcome is 0 or 1. E.g., if the probability for a certain result is zero, this result will never be found — it is impossible. A very small but non-zero probability only makes the result unlikely, it does not make it impossible.

We have already seen that, depending on the system, the vector describing the state of a quantum system could be, e.g., a function or one or several variables, a column vector, or a ket vector. Functions and ket vectors describing quantum states are usually referred to as wave functions and state vectors, respectively.

The probability of each possible outcome of the experiment is entirely determined by the wave function or state vector describing the state of the system at the time of the experiment. Hence, all what one can say about what can be observed on a given quantum system is entirely determined by its wave function or state vector. In this sense, *the wave function or state vector contains all the information that can be known about the state of the system.* In fact, often no distinction is made between the state of a system and the state vector describing it. E.g., one would say "the atom is in the state $|\psi\rangle$" instead of "the atom is in the state described by the state vector $|\psi\rangle$".

It is important to understand that what is measured in experiments is not the

[1] See Section 4.1 of these notes. Alternative formulations exist but are not addressed in this course.

wave function or state vector describing a quantum state (although the results obtained in measurements may of course reveal a lot about the state of a system). Namely, there is no "spin-meter" which would return the values of the coefficients a and b in the ket vector

$$|\chi\rangle = a|\uparrow\rangle + b|\downarrow\rangle \tag{6.1}$$

representing the spin state of a given atom of silver. A measurement of the spin in the z-direction would only say whether each atom measured is found to be in a state of spin up ($|\uparrow\rangle$) or a state of spin down ($|\downarrow\rangle$). It may be that information about the values of a and b could be obtained by making well-chosen measurements on many atoms all prepared in the same state, but no single measurement will return values for a and b. The wave functions and state vectors are not measureable as such; they are theoretical constructs which are used to calculate quantities that can be measured.[2]

The Born rule

Suppose that one would consider checking whether or not a quantum system prepared in a state $|\psi\rangle$ is in a state $|\phi\rangle$ (for example, cheking whether or not an atom prepared in the spin state of Eq. (6.1) is in the state of spin up, $|\uparrow\rangle$.) It might seem bizarre that a system in a state $|\psi\rangle$ could be found to be in a different state $|\phi\rangle$ in a careful experiment; however, see Section 4.4, about the Principle of Superposition, and note that we can always write $|\psi\rangle$ as a linear combination of $|\phi\rangle$ and another vector:

$$|\psi\rangle = |\phi\rangle + (|\psi\rangle - |\phi\rangle). \tag{6.2}$$

Whether or not a measurement would find the system to be in state $|\phi\rangle$ cannot (normally) be predicted with certainty, but it is possible to calculate the probability $\Pr(|\phi\rangle; |\psi\rangle)$ of finding it in that state, given that it was in state $|\psi\rangle$ just before the measurement. It is a fundamental principle of Quantum Mechanics, often referred to as the Born rule or Born postulate, that this probability is given by the following equation:

$$\Pr(|\phi\rangle; |\psi\rangle) = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle}. \tag{6.3}$$

[To keep the notation simple, $\Pr(|\phi\rangle; |\psi\rangle)$ will usually be written as $\Pr(|\phi\rangle)$ in the lectures.]

Note that the numerator and denominator of Eq. (6.3) would both be zero if $|\psi\rangle$ or $|\phi\rangle$ was the zero vector (the zero vector has zero norm and its inner product

[2]This situation is not unique to Quantum Mechanics: e.g., in Classical Mechanics, it is not possible to measure the Lagrangian of a system of particle; the Lagrangian is merely an auxiliary function introduced for helping with the calculation of positions and velocities.

with any other vector is zero). As Eq. (6.3) would then be meaningless, the zero vector never describes a possible state of a quantum system.

The probability $\Pr(|\phi\rangle; |\psi\rangle)$ is exactly the same whether the state of the system is described by the vector $|\psi\rangle$ or by the vector $|c\psi\rangle = c|\psi\rangle$, where c is a non-zero complex number. Indeed, $|\langle\phi|c\psi\rangle|^2 = |c|^2|\langle\phi|\psi\rangle|^2$ since $\langle\phi|c\psi\rangle = c\langle\phi|\psi\rangle$, and $\langle c\psi|c\psi\rangle = |c|^2\langle\psi|\psi\rangle$ since $|c\psi\rangle = c|\psi\rangle$ and $\langle c\psi| = c^*\langle\psi|$. Thus

$$\Pr(|\phi\rangle; |c\psi\rangle) = \frac{|\langle\phi|c\psi\rangle|^2}{\langle\phi|\phi\rangle\langle c\psi|c\psi\rangle} = \frac{|c|^2|\langle\phi|\psi\rangle|^2}{|c|^2\langle\phi|\phi\rangle\langle\psi|\psi\rangle} = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle} = \Pr(|\phi\rangle; |\psi\rangle). \tag{6.4}$$

The same applies for the calculation of any probability, and therefore the vectors $|\psi\rangle$ and $|c\psi\rangle$ describe the same state, in the sense that there is no difference in the predictions from the theory whether we describe the state by $|\psi\rangle$ or by $|c\psi\rangle$: Multiplying a state vector by a non-zero complex number never changes its physical content.

Since the norm of the state vectors is irrelevant as far as calculating probabilities is concerned, the choice of this norm is arbitrary. It is convenient to work only with normalized state vectors — i.e., with state vectors $|\psi\rangle$ and $|\phi\rangle$ such that $\langle\psi|\psi\rangle = 1$ and $\langle\phi|\phi\rangle = 1$, as Eq. (6.3) then takes the simpler form

$$\Pr(|\phi\rangle; |\psi\rangle) = |\langle\phi|\psi\rangle|^2. \tag{6.5}$$

From now on we will work only with normalized state vectors and eigenvectors.

The inner product $\langle\phi|\psi\rangle$ is called the probability amplitude of finding the system in the state $|\phi\rangle$ when it is initially prepared in the state $|\psi\rangle$. (Don't be confused: A probability amplitude is not a probability! It is, in general, a complex number. The probability $\Pr(|\phi\rangle; |\psi\rangle)$ is is a real number, equal to the square of the modulus of the probability amplitude.)

☞ Even though they describe the same quantum state, $|\psi\rangle$ and $c|\psi\rangle$ are nonetheless two different vectors (unless of course $c = 1$). The value of c may thus matter. For example, the linear combinations $|\psi_1\rangle + |\psi_2\rangle$ and $|\psi_1\rangle - |\psi_2\rangle$ may describe very different states although the kets $|\psi_2\rangle$ and $-|\psi_2\rangle$ describe the same state.

☞ Since multiplying the state vector by a non-zero overall factor does not change anything to the predictions of the theory, as long as probabilities are correctly calculated, it is sometimes said that quantum states are represented by rays rather than by vectors (given a vector $|\psi\rangle$, the corresponding ray is the 1-dimensional subspace spanned by $|\psi\rangle$, excluding the zero vector).

We make three important observations at this point:

1. The probability $\Pr(|\phi\rangle; |\psi\rangle)$ is zero if, and only if, the two states $|\psi\rangle$ and $|\phi\rangle$ are orthogonal (i.e., $\langle\phi|\psi\rangle = 0$). Hence, aside from experimental errors, it is impossible that a quantum system be found in a state orthogonal to that in which it was immediately prior to the measurement.

2. The probability $\Pr(|\phi\rangle; |\psi\rangle)$ is 1 if, and only if, $|\phi\rangle = c|\psi\rangle$ with $c \neq 0$ (see proof below). The two ket vectors $|\phi\rangle$ and $|\psi\rangle$ then describe the same quantum state. As would be expected, aside from experimental errors, an experiment aiming at finding whether a quantum system is in the state in which it was prepared immediately before the measurement will certainly find it in that state.

3. In any other case, $0 < \Pr(|\phi\rangle; |\psi\rangle) < 1$: the measurement may or may not find the system to be in the state $|\phi\rangle$, and whether it will find it to be in the state $|\phi\rangle$ cannot be predicted with certainty. Referring to Eq. (6.2) above, there is a non-zero probability that it be found in the state $|\phi\rangle$ and a non-zero probability that it be found in the state $|\psi\rangle - |\phi\rangle$; in that sense, one can say that the system is simultaneously in the states $|\psi\rangle$, $|\phi\rangle$ and $|\psi\rangle - |\phi\rangle$ immediately before the measurement (although one should be careful not to put too much meaning in this interpretation: the theory only predicts probabilities, nothing more).

For example, if we take the spin state given by Eq. (6.1), the probability that the atom is found to be in the state of spin up with respect to the z-direction is $|\langle\uparrow|\chi\rangle|^2$, assuming that $|\chi\rangle$ and $|\uparrow\rangle$ are normalized ($\langle\chi|\chi\rangle = \langle\uparrow|\uparrow\rangle = 1$). Since $\langle\uparrow|\downarrow\rangle = \langle\downarrow|\uparrow\rangle = 0$ and $\langle\downarrow|\downarrow\rangle = 1$, this probability is $|a|^2$, and furthermore $|b|^2 = 1 - |a|^2$. (The latter equation follows from the normalization condition $\langle\chi|\chi\rangle = 1$ and from the fact that $\langle\chi|\chi\rangle = |a|^2\langle\uparrow|\uparrow\rangle + a^*b\langle\uparrow|\downarrow\rangle + ab^*\langle\downarrow|\uparrow\rangle + |b|^2\langle\downarrow|\downarrow\rangle = |a|^2 + |b|^2$.) Thus, in the case where $a = 0$, the probability to find in the atom in the state $|\uparrow\rangle$ is zero (in fact, the atom is then in the state $|\downarrow\rangle$, which is orthogonal to the state $|\uparrow\rangle$. In the case where $a = 1$, then $b = 0$ and $|\chi\rangle = |\uparrow\rangle$: the atom will certainly be found in the state $|\uparrow\rangle$. Otherwise, there is a non-zero probability $|a|^2$ to find it in the state $|\uparrow\rangle$ and a non-zero probability $|b|^2 = 1 - |a|^2$ not to find it in that state.

> ☞ Being the square of the real number $|\langle\phi|\psi\rangle|$, it is obvious that $\Pr(|\phi\rangle; |\psi\rangle) \geq 0$. Moreover, it is clear from Schwarz inequality (Section 2.9 of these notes) that $\Pr(|\phi\rangle; |\psi\rangle) \leq \langle\phi|\phi\rangle\langle\psi|\psi\rangle$, hence, on account of $\langle\phi|\phi\rangle = \langle\psi|\psi\rangle = 1$, that $\Pr(|\phi\rangle; |\psi\rangle) \leq 1$, as befits a probability.
>
> Suppose that $|\phi\rangle = c|\psi\rangle$ with $|c| = 1$. Then $\langle\phi|\psi\rangle = c^*\langle\psi|\psi\rangle = c^*$ and therefore $\Pr(|\phi\rangle; |\psi\rangle) = |c|^2 = 1$. Proving the converse is slightly longer:

Suppose that $\Pr(|\phi\rangle;|\psi\rangle) = 1$. Then, this means that $\langle\phi|\psi\rangle$ must be a number of modulus 1. Thus $\langle\psi|\phi\rangle$ must also be a number of modulus 1. Let $|\eta\rangle = |\phi\rangle - \langle\psi|\phi\rangle|\psi\rangle$. Proceeding as in the proof of the Schwarz inequality given in Section 2.9, one obtains $\langle\eta|\eta\rangle = 1 - |\langle\phi|\psi\rangle|^2$, which is zero. Thus $|\eta\rangle$ can only be the zero vector and $|\phi\rangle$ can only be $|\psi\rangle$ times a number of modulus 1. □

6.2 Dynamical variables and observables

By dynamical variable, in Quantum Mechanics, one means a variable such as the position, momentum, energy or angular momentum of a particle or a whole quantum system — i.e., physical quantities whose value has a probability distribution which may vary in time. Constant quantities which have only one possible values for a given system, such as the charge and the mass of the electron, are not dynamical variables.

Much of Quantum Mechanics is based on a small number of fundamental principles, two of which we have already seen: the Born rule, and that each state of a quantum system can be described by a vector belonging to a Hilbert space. To those, we add the following ones:

- Given the Hilbert space of the state vectors or wave functions describing the possible states of the system of interest, each dynamical variable of this system is associated with a linear operator acting in this Hilbert space.
- The only values a dynamical variable may be found to have, when measured, are the eigenvalues of the operator with which this variable is associated.

For example, the operator associated with the internal energy of an atom of hydrogen is the Hamiltonian $\hat{H}$ of this system. This operator acts on ket vectors $|\psi\rangle$ describing the state of the atom. A measurement of the energy on an atom in a state $|\psi\rangle$ would return one of the eigenvalues of $\hat{H}$ as a result (within the experimental uncertainties, of course).

As noted in Section 1.2, what is directly measured in an actual experiment is often not the dynamical variable of interest but rather some other quantities the value of this dynamical variable can be inferred from. For example, what was directly measured in the experiment of Stern and Gerlach was the position of each of the silver atoms recorded after the magnet; however, measuring each of these positions amounted to a measurement of the component of the spin of the respective atom in the direction of the magnetic field. Hence, the operator associated with this measurement can be taken to be the relevant spin operator

rather than an operator whose eigenvalues would correspond to the possible positions at which atoms could have been found in the experiment.

We will soon see that it is important for the consistency of the theory that the operators associated with dynamical variables each have a complete set of eigenvectors. It is also important that the eigenvalues of these operators are real (not complex) since physical quantities such as the position, momentum, etc, are real. For these two reasons, dynamical variables are represented by Hermitian operator whose eigenvectors form a complete set. Recall from Sections 5.3 that Hermitian operators acting in a finite-dimensional space always have a complete set of eigenvectors. Hermitian operators acting in an infinite-dimensional space may or may not have a complete set of eigenvectors (here, by eigenvectors, we mean ordinary eigenvectors as well as the generalized eigenvectors associated with the continuous spectrum, see Part 8 of these notes).

As the term "dynamical variable" is a bit heavy and a bit undescriptive, in the following we will use the word "observable" to refer to a dynamical variable amenable (in principle) to a measurement.

> ☞ In this we follow the textbook recommended for the Term 1 Quantum Mechanics course.[3] However, in the context of Quantum Mechanics, the word "observable" is more commonly taken to mean a Hermitian operator whose eigenvectors form a complete set spanning the relevant Hilbert space, whether or not this operator corresponds to a physical quantity accessible to experiment. Within this definition, any Hermitian operator acting in a finite-dimensional Hilbert space is an observable.

The question we are thus concerned with is calculating the probability that the outcome of the measurement of a certain observable is a given eigenvalue of the Hermitian operator representing this observable in the theory. We will assume that the wave function or state vector representing the state of the system is known. How such probabilities depends on whether the eigenvalues of interest form part of the discrete spectrum of the relevant operator or part of its continuum spectrum (if there would be one). The latter situation may arise only in infinite-dimensional Hilbert spaces and is more fully addressed in Part 8 of these notes. **At this stage we exclusively consider the case of discrete eigenvalues.** Much of what we will say for the discrete case also applies to the continuous case, with the replacement of discrete summations by integrals. There is, however, an important difference: in the case of a continuum spectrum the calculation gives *densities of probability* of finding specific eigenvalues, whereas in the case of a discrete spectrum it gives *probabilities* of finding specific eigenvalues.

[3] D J Griffiths and D F Schroeter, "Introduction to Quantum Mechanics", 3rd ed., Cambridge University Press, Cambridge (2018).

Suppose that a measurement of an observable A is made on a system in the state $|\psi\rangle$. Suppose, also, that this observable is associated with a Hermitian operator $\hat{A}$. Consider a discrete eigenvalue λ_n of $\hat{A}$, which for the time being we will assume to be non-degenerate, and an eigenvector $|\psi_n\rangle$ of $\hat{A}$ corresponding to that eigenvalue (i.e., the vector $|\psi_n\rangle$ is such that $\hat{A}|\psi_n\rangle = \lambda_n|\psi_n\rangle$.) We also assume that the state vector $|\psi\rangle$ and the eigenvector $|\psi_n\rangle$ are normalized: $\langle\psi|\psi\rangle = \langle\psi_n|\psi_n\rangle = 1$. Then, the probability $\Pr(\lambda_n; |\psi\rangle)$ that the observable A is found to have the value λ_n is given by the equation

$$\Pr(\lambda_n; |\psi\rangle) = |\langle\psi_n|\psi\rangle|^2. \tag{6.6}$$

We note the following:

1. Because $\hat{A}$ is a Hermitian operator, there is no possibility that some of its eigenvalues would be complex rather than real (which would be problematic since these eigenvalues are meant to represent results of measurements).

2. Suppose that the system is in an eigenstate $|\psi_i\rangle$ of $\hat{A}$ corresponding to an eigenvalue λ_i (i.e., $|\psi\rangle = |\psi_i\rangle$). Then there is a zero probability that any other eigenvalue of $\hat{A}$ would be found in the measurement (because eigenvectors of a Hermitian operator corresponding to different eigenvalues are always orthogonal).

3. Eq. (6.6) applies only if the eigenvalue λ_n is non-degenerate. If λ_n is M-fold degenerate, then one can find a set of M orthonormal vectors $|\psi_{n_1}\rangle, |\psi_{n_2}\rangle, \ldots, |\psi_{n_M}\rangle$ such that $\hat{A}|\psi_{n_r}\rangle = \lambda_n|\psi_{n_r}\rangle$, $r = 1, \ldots, M$, and Eq. (6.6) becomes

$$\Pr(\lambda_n; |\psi\rangle) = \sum_{r=1}^{M} |\langle\psi_{n_r}|\psi\rangle|^2. \tag{6.7}$$

4. Obtaining one eigenvalue or another are mutually exclusive possibilities. Therefore the probabilities $\Pr(\lambda_n; |\psi\rangle)$ should sum to 1:

$$\sum_n \Pr(\lambda_n; |\psi\rangle) = 1, \tag{6.8}$$

   where the summation runs over all the eigenvalues of $\hat{A}$. To keep the notation simple, let us assume that the eigenvalues λ_n are all non-degenerate and that the corresponding eigenvectors $|\psi_n\rangle$ are orthonormal. Then Eq. (6.8) says that

$$1 = \sum_n |\langle\psi_n|\psi\rangle|^2 = \sum_n \langle\psi_n|\psi\rangle^* \langle\psi_n|\psi\rangle = \sum_n \langle\psi|\psi_n\rangle\langle\psi_n|\psi\rangle.$$

Since this equation must hold for any state $|\psi\rangle$, we conclude that

$$\sum_n |\psi_n\rangle\langle\psi_n| = \hat{I}, \tag{6.9}$$

where $\hat{I}$ is the identity operator. Eq. (6.9) is the completeness relation, Eq. (5.26) of Section 5.2 of these notes. For the probabilistic interpretation of the inner products $\langle\psi_n|\psi\rangle$ to make sense, it is thus necessary that the eigenvectors $|\psi_n\rangle$ form a complete set.

We know from Section 5.3 that if $\hat{A}$ is defined in a finite-dimensional space it is always possible to form an orthonormal basis of eigenvectors of $\hat{A}$ spanning that space; hence, in spaces of dimension N, the requirement that the eigenvectors of $\hat{A}$ form a complete set does not add to the requirement that $\hat{A}$ is Hermitian. The case where the space is infinite-dimensional is more complicated, though. A Hermitian operator $\hat{A}$ can represent an observable only if its has a complete set of eigenvectors (including generalized eigenvectors in the sense of Part 8 of the course).

☞ That the theory would not be consistent if these eigenvectors did not form a complete set can also be understood from the following argument. Suppose that the eigenvectors of $\hat{A}$ would not form a complete set — i.e., that there would exist a vector $|\psi\rangle$ orthogonal to all the eigenvectors of $\hat{A}$. If the system was prepared in that state, there would then be a zero probability of obtaining any of the eigenvalues of $\hat{A}$ in a measurement of the dynamical variable associated with $\hat{A}$, since the inner products $\langle\psi_n|\psi\rangle$ would be zero for all n; this would be inconsistent with the rule that the only possible values of this variable are the eigenvalues of $\hat{A}$.

☞ Important note. The above considerations generalize to the case of quantum states described by elements of an infinite-dimensional Hilbert space; however, there are many subtleties and the mathematically rigorous theory is far from straightforward. For example, we will see, in Part 8 of the course, that in the case of quantum states described by wave functions defined in 3D space, Eq. (6.9) may take on the form

$$\sum_n \psi_n(\mathbf{r})\psi_n^*(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'), \tag{6.10}$$

or the form

$$\int \phi_{\mathbf{k}}(\mathbf{r})\phi_{\mathbf{k}}^*(\mathbf{r}')\,\mathrm{d}^3k = \delta(\mathbf{r}-\mathbf{r}'), \tag{6.11}$$

or even the form

$$\sum_n \psi_n(\mathbf{r})\psi_n^*(\mathbf{r}') + \int \phi_{\mathbf{k}}(\mathbf{r})\phi_{\mathbf{k}}^*(\mathbf{r}')\,\mathrm{d}^3k = \delta(\mathbf{r}-\mathbf{r}'), \tag{6.12}$$

depending on the Hamiltonian. In these equations, $\mathbf{r}$ and $\mathbf{r}'$ represent position vectors and $\delta(\mathbf{r}-\mathbf{r}')$ is a "3D delta function". The functions $\psi_n(\mathbf{r})$ are eigenfunctions of the Hamiltonian in the ordinary sense of the word (they belong to the Hilbert space of square-integrable functions and are such that $H\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r})$, where H is the Hamiltonian and the E_n's are the eigenenergies). They correspond to discrete energy levels. Eq. (6.10) has exactly the same form as Eq. (6.9), except that the eigenfunctions $\psi_n(\mathbf{r})$ and $\psi_n^*(\mathbf{r}')$ stand for the ket and bra eigenvectors $|\psi_n\rangle$ and $\langle\psi_n|$ and that the delta function $\delta(\mathbf{r}-\mathbf{r}')$ stands for the identity operator $\hat{I}$. However, in Eqs. (6.11) and (6.12), the functions $\phi_{\mathbf{k}}(\mathbf{r})$ do not belong to that Hilbert space, and are eigenfunctions of the Hamiltonian only in a generalized sense (they correspond to continuously distributed value of the energy and of the three components of a wave vector, hence they are summed over through an integration rather than a discrete summation). For this result to be possible, though, it is important that the Hamiltonian is not just Hermitian but also self-adjoint.

We have seen, in Section 5.1, that an operator is self-adjoint when identical to its adjoint. All self-adjoint operators are Hermitian, and in finite-dimensional spaces all Hermitian operators are self-adjoint. However, in infinite-dimensional spaces self-adjointness is a stronger condition than Hermiticity: an operator may be Hermitian but not self-adjoint. Whether a Hermitian operator is or is not self-adjoint depends on whether the domain of the adjoint of this operator is or is not the same as the domain the operator itself (which is something that can be difficult to establish). In Quantum Mechanics, observables are always described by self-adjoint operators. Contrary to mere Hermiticity, self-adjointness guarantees that the probabilities derived from the theory as described above always sum up to 1, and also that the operator has a spectral decomposition, i.e., can be expressed in terms of a (discrete or continuous) sum of projectors as seen in Section 5.3 in the finite-dimensional case. Both are essential for the consistency of the theory.

Throughout the rest of this course, we will always assume that the observables of interest correspond to self-adjoint operators.

6.3 Expectation value of an observable

The expectation value of an observable A in a state $|\psi\rangle$ is the matrix element $\langle\psi|\hat{A}|\psi\rangle$, where $\hat{A}$ is the operator representing this observable. The matrix element $\langle\psi|\hat{A}|\psi\rangle$ is also called the expectation value of the operator $\hat{A}$ in the

state $|\psi\rangle$. If $\hat{A}$ has p distinct eigenvalues λ_1, λ_2,..., λ_p, one can show that

$$\langle\psi|\hat{A}|\psi\rangle = \sum_{n=1}^{p} \lambda_n \Pr(\lambda_n; |\psi\rangle), \tag{6.13}$$

where $\Pr(\lambda_n; |\psi\rangle)$ is the probability that the value λ_n be found if the corresponding physical quantity is measured on a system in the state $|\psi\rangle$.

The reason why $\langle\psi|\hat{A}|\psi\rangle$ is called an "expectation value" is perhaps best grasped from the following example: Suppose that you toss a coin with someone, repeatedly, and at each toss bet one pound that you will get head. Thus at each toss you gain one pound if you get head and lose one pound (i.e., "gain" minus one pound) if you get tail. In probabilistic terms, the "expectation value" of your gain at each toss is the amount you gain if you get head times the probability of that outcome, plus the (negative) amount you gain if you get tail times the probability of that outcome — i.e., (1 pound) $\times$ 0.5 + (-1 pound) $\times$ 0.5 (obviously, this amounts to 0 pound, assuming that the coin is fair). Eq. (6.13) says just the same in a Quantum Mechanical context: The expectation value of the observable $\hat{A}$ is the eigenvalue λ_1 of that operator times the probability that the outcome of the experiment is λ_1, plus the eigenvalue λ_2 times the probability that the outcome is λ_2, etc.

> ☞ Since observables are represented by Hermitian operators (in fact, self-adjoint operators, as noted above), it it clear that $\langle\psi|\hat{A}|\psi\rangle$ is real: by definition of a Hermitian operator, $\langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\psi\rangle^*$ if $\hat{A}$ is Hermitian. It can be shown that the converse is true for complex vector spaces (but is false for real vector spaces): for complex vector spaces, $\langle\psi|\hat{A}|\psi\rangle$ is real for any ket vector $|\psi\rangle$ only if $\hat{A}$ is Hermitian.

6.4 Probability distributions

To summarize the previous sections, Quantum Mechanics predicts that the only values a dynamical variable A can be found to have are the eigenvalues λ_n of the operator $\hat{A}$ associated with A, and that each of these eigenvalues is obtained with a probability $\Pr(\lambda_n; |\psi\rangle)$ given by Eq. (6.6) or Eq. (6.7) if the system is in the state $|\psi\rangle$ at the time of the measurement.

A is thus obtained as a random variable. Following well established traditions, we will denote the mean of this random variable by $\langle A\rangle$ and its variance by $(\Delta A)^2$:

$$\langle A\rangle = \sum_n \lambda_n \Pr(\lambda_n; |\psi\rangle), \tag{6.14}$$

$$(\Delta A)^2 = \sum_n (\lambda_n - \langle A\rangle)^2 \Pr(\lambda_n; |\psi\rangle). \tag{6.15}$$

(It is clear that the values of $\langle A \rangle$ and $(\Delta A)^2$ depend on the state of the system, $|\psi\rangle$; for simplicity, we do not specify this dependence in the notation.) $\langle A \rangle$ can also be written in terms of the operator $\hat{A}$ as

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle \tag{6.16}$$

As we have seen in the previous section, $\langle \psi | \hat{A} | \psi \rangle$ is called the expectation value of $\hat{A}$ in the state $|\psi\rangle$. The variance $(\Delta A)^2$ can be written in two equivalent ways in terms of $\hat{A}$:

$$(\Delta A)^2 = \langle \psi | (\hat{A} - \langle A \rangle)^2 | \psi \rangle \tag{6.17}$$

$$= \langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2. \tag{6.18}$$

☞ Proof: First, note that $\langle \psi | (\hat{A} - \langle A \rangle)^2 | \psi \rangle$ should really be written as $\langle \psi | (\hat{A} - \langle A \rangle \hat{I})^2 | \psi \rangle$, since $\langle A \rangle$ is a scalar, not an operator. Since $\langle \psi | \psi \rangle = 1$,

$$\begin{aligned} \langle \psi | (\hat{A} - \langle A \rangle \hat{I})^2 | \psi \rangle &= \langle \psi | \hat{A}^2 - 2\langle A \rangle \hat{A} + \langle A \rangle^2 \hat{I} | \psi \rangle \\ &= \langle \psi | \hat{A}^2 | \psi \rangle - 2\langle A \rangle \langle \psi | \hat{A} | \psi \rangle + \langle A \rangle^2 \\ &= \langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2. \end{aligned} \tag{6.19}$$

Eqs. (6.17) and (6.18) are therefore equivalent. □

Showing that Eqs. (6.17) and (6.18) are also equivalent to Eq. (6.15) is left as an exercise.

ΔA, the square of the variance $(\Delta A)^2$, is often referred to as the uncertainty in the variable A. Eqs. (6.17) and (6.18) provide a precise definition of this quantity.

Suppose we have N identical copies of the quantum system of interest (e.g., N atoms of hydrogen), that all these copies are prepared in the same state $|\psi\rangle$, and that we make the same measurement of the observable A on each of them. Experimental errors put aside, the value found for A in each of these N measurements will be one of the eigenvalues λ_n of $\hat{A}$. If the theoretical description of these measurements is correct, the probability distribution $\Pr(\lambda_n; |\psi\rangle)$ predicts the frequency distribution of these experimental results.

Let $\lambda^{(1)}$ be the value of A found for system 1, $\lambda^{(2)}$ the value found for system 2, etc., and $\lambda^{(N)}$ the value found for system N. These N results form a statistical sample of mean $\bar{\lambda}$ and standard deviation σ, with

$$\bar{\lambda} = \frac{1}{N} \sum_{j=1}^{N} \lambda^{(j)} \qquad \text{and} \qquad \sigma = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} \left(\lambda^{(j)} - \bar{\lambda}\right)^2}.$$

Quantum Mechanics predicts that $\bar{\lambda} \to \langle A \rangle$ and $\sigma \to \Delta A$ for $N \to \infty$ if experimental errors can be ignored.

The dispersion of the data about their mean $\bar{\lambda}$ is characterized by the standard deviation σ. How close to $\langle A \rangle$ we may expect each of the measured values of A to be is characterized by the uncertainty ΔA. We stress that ΔA has nothing to do with an experimental error. A non-zero value of ΔA sets a fundamental limit on what can be predicted about the value found for A if this observable was measured, irrespective of any experimental considerations.

Apart for the experimental error, there is certainty about what value A would be found to have if and only if $\Delta A = 0$. This is the case if the state vector $|\psi\rangle$ is in an eigenvector of $\hat{A}$, and only if it is in an eigenvector of $\hat{A}$ (see proof below); obviously, the value found for A would then be the eigenvalue of $\hat{A}$ this eigenvector corresponds to.

☞ Let us first show that $\Delta A = 0$ if $|\psi\rangle$ is an eigenvector of $\hat{A}$. Suppose that $\hat{A}|\psi\rangle = \lambda|\psi\rangle$. Then $\langle A \rangle = \langle\psi|\hat{A}|\psi\rangle = \lambda\langle\psi|\psi\rangle = \lambda$. Moreover, $\hat{A}^2|\psi\rangle = \lambda\hat{A}|\psi\rangle = \lambda^2|\psi\rangle$. Therefore $\langle\psi|\hat{A}^2|\psi\rangle = \lambda^2\langle\psi|\psi\rangle = \lambda^2$, and $(\Delta A)^2 = \langle\psi|\hat{A}^2|\psi\rangle - \langle\psi|\hat{A}|\psi\rangle^2 = \lambda^2 - \lambda^2 = 0$.

We now prove the converse, which is that $|\psi\rangle$ is necessarily an eigenvector of $\hat{A}$ if $\Delta A = 0$. Let us assume that $\Delta A = 0$. We note that $(\Delta A)^2 = \langle\psi|(\hat{A} - \langle A\rangle\hat{I})^2|\psi\rangle$. Let $|\eta\rangle = (\hat{A} - \langle A\rangle\hat{I})|\psi\rangle$. Since $\hat{A}$ and $\hat{I}$ are self-adjoint and $\langle A \rangle$ is real, $(\hat{A} - \langle A\rangle\hat{I} = ((\hat{A} - \langle A\rangle\hat{I})^\dagger$. Hence $(\Delta A)^2 = \langle\psi|(\hat{A} - \langle A\rangle\hat{I})^\dagger(\hat{A} - \langle A\rangle\hat{I})|\psi\rangle = \langle\eta|\eta\rangle$. Thus $\langle\eta|\eta\rangle = 0$ since $\Delta A = 0$, which means that $|\eta\rangle$ is the zero vector, therefore that $\hat{A}|\psi\rangle = \langle A\rangle|\psi\rangle$. Hence $|\psi\rangle$ is an eigenvector of $\hat{A}$. □

6.5 Uncertainty relations

Suppose that A and B are observables represented, respectively, by the operators $\hat{A}$ and $\hat{B}$. One can show that the product of the uncertainties ΔA and ΔB always obeys the following inequality, where $[\hat{A}, \hat{B}]$ is the commutator of $\hat{A}$ and $\hat{B}$:

$$(\Delta A)^2(\Delta B)^2 \geq -\frac{1}{4}\big(\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle\big)^2. \tag{6.20}$$

☞ The proof of Eq. (6.20) is similar to the one you have seen in the Term 1 QM course for a particular case of this uncertainty relation. Before starting, we recall that $\hat{A}$ and $\hat{B}$ are self-adjoint operators (see the important note at the end of Section 6.2). Let $\hat{A}' = \hat{A} - \langle A\rangle\hat{I}$ and $\hat{B}' = \hat{B} - \langle B\rangle\hat{I}$. By definition of the uncertainties ΔA and ΔB, $(\Delta A)^2 = \langle\psi|\hat{A}'^2|\psi\rangle$ and $(\Delta B)^2 = \langle\psi|\hat{B}'^2|\psi\rangle$. Introducing the real number λ, we observe that

$(A' + i\lambda B')|\psi\rangle$ is a vector and that the square of the norm of this vector is necessarily non-negative (as is the case for any vector). Thus

$$\langle\psi|(A'^{\dagger} - i\lambda B'^{\dagger})(A' + i\lambda B')|\psi\rangle \geq 0. \tag{6.21}$$

Given that the operators $\hat{A}$ and $\hat{B}$ are self-adjoint and that $\langle A\rangle$ and $\langle B\rangle$ are real, it is also true that

$$\langle\psi|(A' - i\lambda B')(A' + i\lambda B')|\psi\rangle \geq 0. \tag{6.22}$$

That is,

$$\langle\psi|A'^2|\psi\rangle + i\lambda\langle\psi|A'B' - B'A'|\psi\rangle + \lambda^2\langle\psi|B'^2|\psi\rangle \geq 0. \tag{6.23}$$

It is easy to see that $A'B' - B'A' = [A, B]$, and as noted above $\langle\psi|[A, B]|\psi\rangle$ is an imaginary number. Let us denote $\langle\psi|[A, B]|\psi\rangle$ by iC, where C is real, and rewrite Eq. (6.23) in the form

$$\lambda^2(\Delta B)^2 - \lambda C + (\Delta A)^2 \geq 0. \tag{6.24}$$

This inequation is fulfilled for $\lambda = 0$ and must be fulfilled for any real value of λ. Thus the quadratic equation $\lambda^2(\Delta B)^2 - \lambda C + (\Delta A)^2 = 0$, as an equation for λ, cannot have two distinct real solutions, which is the case only if $C^2 - 4(\Delta A)^2(\Delta B)^2 \leq 0$. Eq. (6.20) follows. □

☞ Despite the minus sign, the right-hand side of this equation is always a non-negative real number.

Proof: Since $\hat{A}$ and $\hat{B}$ are self-adjoint, $\langle\psi|\hat{A}\hat{B}|\psi\rangle = \langle\psi|\hat{B}^{\dagger}\hat{A}^{\dagger}|\psi\rangle^* = \langle\psi|\hat{B}\hat{A}|\psi\rangle^*$. Thus $\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle = \langle\psi|\hat{A}\hat{B} - \hat{B}\hat{A}|\psi\rangle = \langle\psi|\hat{B}\hat{A} - \hat{A}\hat{B}|\psi\rangle^* = \langle\psi|[\hat{B}, \hat{A}]|\psi\rangle^* = -\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle^*$. Now, $\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle$ is a number, and we have shown that this number is equal to the negative of its complex conjugate. This number must therefore be imaginary (i times a real number), which implies that $\big(\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle\big)^2$ is negative. □

The following is worth noting:

1. This inequality is a relation between the widths of two different probability distributions, under the assumption that each one is based on the same state $|\psi\rangle$. There is no assumption that the variables A and B are measured simultaneously or even measured in a same experiment.

2. A state $|\psi\rangle$ in which the product $\Delta A\Delta B$ is zero may exist if the right-hand side of this inequality is zero. This is the case, in particular, if $\hat{A}$ commutes with $\hat{B}$. (In fact, if $\hat{A}$ commutes with $\hat{B}$, these two operators have common eigenstates, and for such states the uncertainties ΔA and

ΔB are both zero.) There is no theoretical limit on how small ΔA and ΔB may both be in a same quantum state when $\hat{A}$ commutes with $\hat{B}$.

3. The time-energy uncertainty relation is not amenable to this formulation. Time is a parameter in Quantum Mechanics, not a dynamical variable associated with a Hermitian operator.

☞ One can also show that in finite-dimensional spaces, $\langle\psi|[\hat{A},\hat{B}]|\psi\rangle = 0$ if $|\psi\rangle$ is an eigenvector of $\hat{A}$ or $\hat{B}$. This result is consistent with the fact previously mentioned that $\Delta A = 0$ if $|\psi\rangle$ is an eigenvector of $\hat{A}$. However, it does not apply to the case where one would work in an infinite-dimensional space and take $|\psi\rangle$ to be a generalized eigenvector of $\hat{A}$ or $\hat{B}$ (i.e., an eigenvector belonging to the continuous spectrum of one of these operators, in the "physicists' definition" of these, as discussed in Part 8 of these notes).

6.6 The state of the system after a measurement

This discussion of measurements in Quantum Mechanics would not be complete without a mention of the influence of a measurement on the quantum state of the system. There would be enormously to say about this issue; however, it is outside the syllabus of the course and therefore we discuss it but briefly here.

The first thing to say is that it is not infrequent that a measurement made on a system destroys this system completely. In their historical experiment, Stern and Gerlach detected the atoms of silver they studied by the trace they left when absorbed on a screen they impacted on. Clearly, it would be a moot point to discuss the spin state of the atoms after such a destructive measurement. However, not all measurements are of this nature, and experiments can be done in which the system under study is measured repeatedly. The question then is: if the state of the system was described by a certain state vector $|\psi\rangle$ before the first measurement, what can then one predict about the outcome of a second measurement?

The question is answered concisely by what is often regarded as a fundamental principle of Quantum Mechanics, called the "collapse postulate" or an equivalent name. In its simplest form, this principle can be stated as follows:

- If a system is initially in a state $|\psi\rangle$ and is found to be in a state $|\phi\rangle$ in a measurement, then immediately after the measurement this system is in the state $|\phi\rangle$.

If a second measurement is made on this system, the outcome of this second measurement therefore needs to be calculated from the state vector $|\phi\rangle$, in

the same way as the outcome of the first measurement needs to be calculated from the state vector $|\psi\rangle$. Thus, if the system is found to be in the state $|\phi\rangle$ and an identical measurement would be made immediately after the first, this second measurement would also find that the system is in state $|\phi\rangle$.

In some respects, this rule may seem completely intuitive: if the system is found to be in state $|\phi\rangle$, why would it not be found to be still in the same state if checked and the state has not changed between the two measurements? Note, however, that the collapse postulate implies that the state vector of the system changes abruptly upon a measurement. Suppose that an atom would be in the spin state $|\chi\rangle = a|\uparrow\rangle + b|\downarrow\rangle$ and somehow one would measure the z-component of the spin and find it to be up, then at the point of this measurement the spin state of this atom changes from $|\chi\rangle$ to $|\uparrow\rangle$. It is often said that the measurement "collapses" the wave function (or here the state vector) from a linear combination of several states to the state found in the measurement. However, the situation is normally rather more complicated than just stated and requires a detailed analysis of what a measurement really entails. There has been much discussion about the status of the collapse postulate and the interpretation of the abrupt change over from one state to another it describes — e.g., whether it corresponds to something "real" in the physical properties of the atom, or whether it merely reflects a change in what we know about the state of the atom. In many respects, however, the issue of how the collapse postulate is best interpreted is a problem of Philosophy rather than of Quantum Theory, and is outside the scope of this course.

Let us illustrate the above by yet another thought experiment on spin states. Suppose that the spin state of a spin-1/2 atom (e.g., an atom of silver) is initially represented by the column vector

$$\chi = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}. \tag{6.25}$$

As we will see later in the course, the column vectors

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \alpha^{(x)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

are eigenvectors of, respectively, the z-component and the x-component of the spin operator, and both correspond to a state of spin "up" in the respective direction (i.e., they both correspond to an eigenvalue of $\hbar/2$ rather than $-\hbar/2$. Note that these three column vectors are normalized. Now, imagine an experiment in which one checks whether an atom initially in the state χ is or is not in the state α (thus in the state of spin up with respect to the z-direction). The probability of finding it in that state is given by Eq. (6.5) as

$$\left| \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{1/3} \\ \sqrt{2/3} \end{pmatrix} \right|^2 = \left| 1 \times \sqrt{1/3} + 0 \times \sqrt{2/3} \right|^2 = 1/3. \tag{6.26}$$

Suppose that the atom is found to be in the state α. Upon this finding, its state changes from χ to α. If this state is not perturbed by some interaction or another measurement, a subsequent check of whether it is in the state α will find it to be in that state with a probability of 1 since

$$\left|\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right|^2 = 1. \tag{6.27}$$

Instead of checking whether it is in the state α after the first measurement, let us imagine that it would be checked whether it is in the state $\alpha^{(x)}$. It would be found to be in that state with a probability of

$$\left|\begin{pmatrix}\sqrt{1/2} & \sqrt{1/2}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right|^2 = 1/2. \tag{6.28}$$

Assumint that it is found to be in that state, its state would then be $\alpha^{(x)}$ after this second measurement. If now a third measurement is made, on whether this atom is or is not in the state α, the probability of finding it in that state is now 1/2, rather than 1, since

$$\left|\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}\sqrt{1/2}\\\sqrt{1/2}\end{pmatrix}\right|^2 = 1/2. \tag{6.29}$$

The collapse postulate can be recast in terms of the projector operator $\hat{\mathcal{P}}_\phi$ introduced in Section 3.10, as we will now see. A system initially in a state $|\psi\rangle$ can be found to be in a state $|\phi\rangle$ only if $\langle\phi|\psi\rangle \neq 0$. If so, and in view of Eq. (3.103), $\hat{\mathcal{P}}_\phi|\psi\rangle \neq 0$. Given that $c|\phi\rangle$ represents the same quantum state as $|\phi\rangle$ for any non-zero value of the number c, we see that the collapse postulate can also be phrased as follows: "If a system is initially in a state $|\psi\rangle$ and is found to be in a state $|\phi\rangle$ in a measurement, then immediately after the measurement this system is in a quantum state described by the state vector $\hat{\mathcal{P}}_\phi|\psi\rangle$." Or, in other words,

- If a system is initially in a state $|\psi\rangle$, finding it to be in a state $|\phi\rangle$ projects its state vector onto the 1-dimensional subspace spanned by $|\phi\rangle$.

The same applies to the case where what is measured is the value of a certain observable. Suppose that a particular eigenvalue λ_n of the Hermitian operator $\hat{A}$ representing this observable is obtained as a result and that this eigenvalue is non-degenerate. Then, if $|\psi_n\rangle$ is a normalized eigenvector of $\hat{A}$ corresponding to this eigenvalue, the collapse postulate sets that the measurement leaves the system in the state $|\psi_n\rangle$.

This principle can be generalized to the case where the eigenvalue λ_n is degenerate. Recall that one can find a set of M orthonormal eigenvectors of $\hat{A}$ all belonging to the eigenvalue λ_n if this eigenvalue is M-fold degenerate. Let $|\psi_{n_1}\rangle, |\psi_{n_2}\rangle, \ldots, |\psi_{n_M}\rangle$ be such a set. These M orthonormal vectors span a M-dimensional subspace of the whole Hilbert space, and the operator

$$\hat{\mathcal{P}}_{\lambda_n} = \sum_{r=1}^{M} |\psi_{n_r}\rangle\langle\psi_{n_r}| \tag{6.30}$$

projects any vector of this Hilbert space onto that M-dimensional subspace. In general, the collapse postulate sets that if a system is initially in a state $|\psi\rangle$, finding the eigenvalue λ_n of $\hat{A}$ in a measurement of the corresponding dynamical variable projects $|\psi\rangle$ onto the subspace of this eigenvalue. Correspondingly, the measurement transforms the state vector of the system from $|\psi\rangle$ to $\hat{\mathcal{P}}_{\lambda_n}|\psi\rangle$ (or, in terms of a normalized vector, to $\hat{\mathcal{P}}_{\lambda_n}|\psi\rangle/||\hat{\mathcal{P}}_{\lambda_n}|\psi\rangle||$).

The case of a pair of observables represented by commuting operators is particularly noteworthy. Suppose that the observables A and B are represented by the operators $\hat{A}$ and $\hat{B}$, respectively, with $[\hat{A}, \hat{B}] = 0$, and suppose that A is measured. The measurement transforms the initial state vector into a superposition of eigenvectors of $\hat{A}$ belonging to the eigenvalue found in the measurement (λ_n, say):

$$|\psi\rangle \to \hat{\mathcal{P}}_{\lambda_n}|\psi\rangle = \sum_{r=1}^{M} c_n |\psi_{n_r}\rangle \tag{6.31}$$

with $c_n = \langle\psi_{n_r}|\psi\rangle$. A subsequent measurement of B resulting in the eigenvalue μ_m of $\hat{B}$ then further transforms this superposition into a superposition of eigenvectors of $\hat{B}$ belonging to μ_m:

$$\sum_{r=1}^{M} c_n |\psi_{n_r}\rangle \to \sum_{r=1}^{M} c_n \hat{\mathcal{P}}_{\mu_m}|\psi_{n_r}\rangle. \tag{6.32}$$

However (the proof is left as an exercise), the resulting superposition is still an eigenvector of $\hat{A}$ belonging to the eigenvalue λ_n. Therefore, if A is measured again, after the measurement of B, λ_n is found with probability 1. In other words, a measurement of B would not affect the result of a measurement of A. In view of this fact, the observables A and B are said to be compatible. Two observables are compatible when, and only when, they are representing by operators commuting with each other.

[This page is intentionally left blank.]