Chapter 5
Operators

In this chapter, we recall the material from the previous chapters and summarise how vectors and operators play different roles in quantum mechanics. After an introduction to so-called eigenvalue problems, we discuss observables and evolution operators, as well as projectors and commutators. They provide us with the essential mathematical techniques for later chapters.

5.1 Eigenvalue Problems

Mathematically, quantum mechanics is a theory about calculating probabilities using the concepts of states and operators. In this chapter we are going to study the most important aspects of operators for quantum mechanics, and show how it fits in with the theory. First, we will address the most fundamental properties of operators, namely their eigenvalues and eigenstates.

In the previous chapter you have encountered the eigenvalue equation of the Hamiltonian

$$H|E_n\rangle = E_n|E_n\rangle\,, \tag{5.1}$$

where H is an operator, and $|E_n\rangle$ is an eigenstate of H with eigenvalue E_n. The Hamiltonian is the energy operator and the eigenvalues E_n are the energies that the system can possess. The states (or vectors) $|E_n\rangle$ describe the system when it has the energy E_n. Often we need to find out the energy states $|E_n\rangle$ given a particular Hamiltonian H. Sometimes the energy eigenvalues E_n are known straight away or can be found from applying well-known physical principles, and other times the

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-319-92207-2_5) contains supplementary material, which is available to authorized users.

© Springer International Publishing AG, part of Springer Nature 2018

P. Kok, *A First Introduction to Quantum Physics*,
Undergraduate Lecture Notes in Physics,
https://doi.org/10.1007/978-3-319-92207-2_5

eigenvalues must be calculated from a general form of H. In this section we will show how you can find both E_n and $|E_n\rangle$ given H.

Let's write the previous equation in matrix form with $|g\rangle$ and $|e\rangle$ the energy eigenstates, E_g and E_e the corresponding energies, and

$$H = \begin{pmatrix} E_g & 0 \\ 0 & E_e \end{pmatrix} . \tag{5.2}$$

It is straightforward to show that Eq. (5.1) holds. First, we define

$$|g\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |e\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{5.3}$$

Next, we substitute H, $|g\rangle$, and $|e\rangle$ into Eq. (5.1). We find that

$$\begin{aligned} H|g\rangle &= \begin{pmatrix} E_g & 0 \\ 0 & E_e \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} E_g \\ 0 \end{pmatrix} = E_g \begin{pmatrix} 1 \\ 0 \end{pmatrix} = E_g|g\rangle , \\ H|e\rangle &= \begin{pmatrix} E_g & 0 \\ 0 & E_e \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ E_e \end{pmatrix} = E_e \begin{pmatrix} 0 \\ 1 \end{pmatrix} = E_e|e\rangle . \end{aligned} \tag{5.4}$$

You see that there is a direct link between the numerical values of the vector components of $|g\rangle$, $|e\rangle$, and the fact that H is diagonal. When a matrix is diagonal, the only non-zero matrix elements are along the line from the top left to the bottom right, and the diagonal elements are the eigenvalues (which can also include zero). The eigenvectors are then simply of the form of Eq. (5.3). This is true for matrices of all dimensions, so if we define a diagonal matrix M according to

$$M = \begin{pmatrix} m_1 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 \\ 0 & 0 & m_3 & 0 \\ 0 & 0 & 0 & m_4 \end{pmatrix} , \tag{5.5}$$

then the four eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ are given by

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} , \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \quad \mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} , \tag{5.6}$$

and the eigenvalue equation is given by

$$M\mathbf{v}_j = m_j \mathbf{v}_j , \tag{5.7}$$

where $j = 1, 2, 3, 4$. You see that the generalisation to arbitrary dimension is straightforward.

We often have to deal with matrices that are not diagonal. Consider the situation where the atom interacts with a laser field, also discussed in the previous chapter. The Hamiltonian H_{total} is a combination of the free evolution Hamiltonian H and the interaction Hamiltonian H_{int} from Eq. (4.30):

$$H_{\text{total}} = \begin{pmatrix} E_g & \gamma^* \\ \gamma & E_e \end{pmatrix}, \tag{5.8}$$

If we use this Hamiltonian in Eq. (5.1) with the vectors $|g\rangle$ and $|e\rangle$, we see that it is no longer satisfied, and therefore $|g\rangle$ and $|e\rangle$ are no longer eigenstates of the Hamiltonian. In other words, and very much against our intuition, $|g\rangle$ and $|e\rangle$ are not the states of the system with a well-defined energy when there is an interaction present! The states with a particular value of the total energy change as the interaction strength γ is increased, and the allowed energy values tend to change also. So what are the new energy eigenvalues and eigenstates?

We can set up an eigenvalue equation for the new Hamiltonian with energies E'_g and E'_e, and we want to express these in terms of the original energies E_g and E_e, as well as the interaction strength γ. Lucky for us, this is a problem that has been solved a long time ago in a branch of mathematics called linear algebra, which deals with matrices and vectors of all kinds. Here we will give the recipe for finding the eigenvalues and eigenvectors without showing explicitly where it comes from, and for a full explanation where this all comes from we refer you to any introductory linear algebra book (see Further Reading).

First, we need to use a property of matrices called the *determinant*, which is a kind of measure for the "size" of a matrix. The determinant of a 2×2 matrix is given by

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc\,. \tag{5.9}$$

We take the eigenvalue equation and rearrange it so the right-hand side becomes zero:

$$H|\lambda\rangle - \lambda|\lambda\rangle = (H - \lambda\mathbb{I})\,|\lambda\rangle = 0\,, \tag{5.10}$$

where λ is the eigenvalue and $|\lambda\rangle$ the eigenvector of H. We included the identity matrix $\mathbb{I}$ when we factor out $|\lambda\rangle$. Since we don't want $|\lambda\rangle = 0$, somehow the operation of $H - \lambda\mathbb{I}$ on $|\lambda\rangle$ must give zero in order to satisfy Eq. (5.10). This can be achieved by requiring that the determinant of $H - \lambda\mathbb{I}$ is zero:

$$\begin{aligned} \det(H - \lambda\mathbb{I}) &= \det \begin{pmatrix} E_g - \lambda & \gamma^* \\ \gamma & E_e - \lambda \end{pmatrix} \\ &= (E_g - \lambda)(E_e - \lambda) - |\gamma|^2 = 0\,. \end{aligned} \tag{5.11}$$

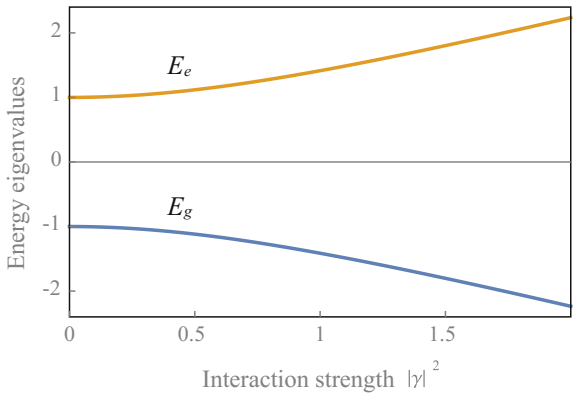


Fig. 5.1 Changes in eigenvalues with increasing interaction strength $|\gamma|^2$. The energy units on the axes are arbitrary

The eigenvalues of H are the two values of λ that make the determinant zero. Since the equation for λ in Eq. (5.11) is a quadratic equation of the form $a\lambda^2 + b\lambda + c = 0$, we can use the standard solution:

$$\lambda_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1}{2}\left(E_g + E_e \pm \sqrt{(E_e - E_g)^2 + 4|\gamma|^2}\right). \tag{5.12}$$

When $\gamma \to 0$ the eigenvalues are $\lambda_- = E_g$ and $\lambda_+ = E_e$, as you would expect. The effect of the interaction is to increase the energy gap between the energy eigenstates: E_e increases, while E_g decreases (see Fig. 5.1).

Now that we have the eigenvalues λ_+ and λ_- of the Hamiltonian H, we want to know what the corresponding eigenvectors $|\lambda_+\rangle$ and $|\lambda_-\rangle$ are. To find the eigenvectors, we set up the eigenvalue equation

$$H|\lambda_\pm\rangle = \lambda_\pm|\lambda_\pm\rangle\,. \tag{5.13}$$

We can express $|\lambda_+\rangle$ and $|\lambda_-\rangle$ in terms of the original eigenstates $|g\rangle$ and $|e\rangle$ as

$$|\lambda_+\rangle = a_+|g\rangle + b_+|e\rangle \quad \text{and} \quad |\lambda_-\rangle = a_-|g\rangle + b_-|e\rangle\,, \tag{5.14}$$

where we impose the additional normalisation condition $|a_\pm|^2 + |b_\pm|^2 = 1$. Let's take the eigenvalue λ_+ and the eigenstate $|\lambda_+\rangle$. The eigenvalue equation is

$$\begin{pmatrix} E_g & \gamma^* \\ \gamma & E_e \end{pmatrix}\begin{pmatrix} a_+ \\ b_+ \end{pmatrix} = \lambda_+ \begin{pmatrix} a_+ \\ b_+ \end{pmatrix}. \tag{5.15}$$

When we write this out, both the top and the bottom vector components on either side of the equality sign must match (this is a vector equation). We therefore get two equations:

$$E_g a_+ + \gamma^* b_+ = \lambda_+ a_+ \quad \text{and} \quad \gamma a_+ + E_e b_+ = \lambda_+ b_+ \,. \tag{5.16}$$

Rearranging these equations yields

$$\gamma^* b_+ = (\lambda_+ - E_g) a_+ \quad \text{and} \quad \gamma a_+ = (\lambda_+ - E_e) b_+ \,. \tag{5.17}$$

To take into account the normalisation condition, we take the square modulus of these equations and obtain

$$\begin{aligned} |\gamma|^2 |b_+|^2 &= (\lambda_+ - E_g)^2 |a_+|^2 \\ |\gamma|^2 |a_+|^2 &= (\lambda_+ - E_e)^2 |b_+|^2 \,, \end{aligned} \tag{5.18}$$

which, using $|b_+|^2 = 1 - |a_+|^2$ leads to

$$|\gamma|^2 - |\gamma|^2 |a_+|^2 = (\lambda_+ - E_g)^2 |a_+|^2 \tag{5.19}$$

$$|\gamma|^2 |a_+|^2 = (\lambda_+ - E_e)^2 (1 - |a_+|^2) \,, \tag{5.20}$$

Substituting Eq. (5.20) into the left-hand side of Eq. (5.19) and solving for $|a_+|^2$, we find

$$|a_+|^2 = \frac{|\gamma|^2 - (\lambda_+ - E_e)^2}{(\lambda_+ - E_g)^2 - (\lambda_+ - E_e)^2} \,, \tag{5.21}$$

and

$$|b_+|^2 = 1 - |a_+|^2 = \frac{(\lambda_+ - E_g)^2 - |\gamma|^2}{(\lambda_+ - E_g)^2 - (\lambda_+ - E_e)^2} \,. \tag{5.22}$$

When $\gamma \to 0$ you see that the values of a_+ and b_+ return to 1 and 0, respectively, since in that case $\lambda \to E_g$ and $|\lambda_+\rangle \to |E_g\rangle$.

We obtain the magnitudes of the complex numbers a_+ and b_+ by taking the positive square roots of Eqs. (5.21) and (5.22), but we still have to find the phases. We can accomplish this by setting $a_+ = |a_+| e^{i\phi_a}$ and $b_+ = |b_+| e^{i\phi_b}$, and solving for ϕ_a and ϕ_b by substituting into Eq. (5.16). Unfortunately, this is a bit lengthy and messy, so we will only give the answer and leave it up to you to check the result:

$$|\lambda_+\rangle = \frac{1}{2\mathcal{N}_+} \begin{pmatrix} E_g - E_e + \sqrt{(E_e - E_g)^2 + 4|\gamma|^2} \\ 2\gamma \end{pmatrix} \,, \tag{5.23}$$

where the normalisation $\mathcal{N}_+$ is given by

$$\mathcal{N}_+ = 4|\gamma|^2 + (E_e - E_g)^2 + \sqrt{(E_e - E_g)^4 + 4|\gamma^2 (E_e - E_g)^2} \,. \tag{5.24}$$

Similarly, we can calculate the eigenvector $|\lambda_-\rangle$, and we will find

$$|\lambda_-\rangle = \frac{1}{2\mathcal{N}_-}\begin{pmatrix} E_g - E_e - \sqrt{(E_e - E_g)^2 + 4|\gamma|^2} \\ 2\gamma \end{pmatrix}, \tag{5.25}$$

with the normalisation factor

$$\mathcal{N}_- = 4|\gamma|^2 + (E_e - E_g)^2 - \sqrt{(E_e - E_g)^4 + 4|\gamma|^2(E_e - E_g)^2}\,. \tag{5.26}$$

This is the full solution of the eigenvectors and eigenvalues of H_{total}.

We can apply this mathematical procedure to any matrix, and find the eigenvalues and eigenvectors. The eigenvalues are often most important, since these are the physically allowed values of the energy, and determine the spectral absorption and emission lines in spectroscopy. We also often want to know the ground state of a system, since this is the state the system most likely relaxes into after some time. The ground state is the eigenstate of the lowest energy level ($|\lambda_-\rangle$ in the example above). If we want to calculate the probability p_n that a particular state $|\psi\rangle$ is measured in an energy eigenstate $|E_n\rangle$, we must find the eigenstates $|E_n\rangle$ in order to evaluate the Born rule $p_n = |\langle E_n|\psi\rangle|^2$.

How do we evaluate the determinant of higher-dimensional matrices? This is determined recursively by the determinant of 2×2 matrices. For example, given a 3×3 matrix we can calculate the determinant as follows: choose any row or column (preferably with as many zero entries as possible, as you will see shortly). For each element in the row or column we define a new matrix that we obtain by removing the row and column of that element. This is the sub-matrix of that element. If we start with a 3×3 matrix, the sub-matrix of any element will be a 2×2 matrix. The determinant of that sub-matrix (which we can evaluate) is called the *cofactor* of that element.

We now evaluate the determinant of the 3×3 matrix as follows: we multiply each element in our chosen row or column by its cofactor. Depending on the position of the element in the matrix, this product picks up a minus sign:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}, \tag{5.27}$$

which for larger matrices extends in a regular checker board pattern. The determinant is the sum of these products.

As an example, consider the determinant of the 3×3 matrix A:

$$A = \begin{pmatrix} 5 & 1 & 2 \\ 1 & -8 & -4 \\ 4 & 6 & -7 \end{pmatrix}. \tag{5.28}$$

Let's pick the top row to evaluate the determinant:

$$\begin{aligned}\det A &= 5 \times \det \begin{pmatrix} -8 & -4 \\ 6 & -7 \end{pmatrix} - 1 \times \det \begin{pmatrix} 1 & -4 \\ 4 & -7 \end{pmatrix} + 2 \times \det \begin{pmatrix} 1 & -8 \\ 4 & 6 \end{pmatrix} \\ &= 5 \times 80 - (-23) + 2 \times 38 = 499\,. \end{aligned} \tag{5.29}$$

You see that you want to pick the row or column with the most zeros, since it will reduce the amount of work.

We can calculate the determinant of any matrix using the above procedure. The determinant is expressed in terms of determinants of smaller matrices, which in turn can be expressed in terms of determinants of smaller matrices, and so on, until we arrive at 2×2 matrices.

5.2 Observables

We have seen that there are two types of matrices that play an important role in quantum mechanics. The first type we encountered was the matrix describing the beam splitter, and it tells us how the state of the photon changes after we do something to it. In other words, it describes the *evolution* of the quantum state. The second type of matrix we encountered was the matrix describing the spin of an electron, and later we discussed the matrix describing the energy of a system. This type of matrix generally describes a physical property of the system (such as spin and energy), and we call these matrices *observables*.

When you think about it, it is a little bit strange that observables are matrices that are completely independent of the state of the system. You may have expected that the state is simply a list of physical properties of the system. However, we have seen in Chap. 1 that "a phenomenon is not a phenomenon until it is an observed phenomenon", which means that quantum systems have no fixed properties independent of the measurement of the system. So instead, we have an abstract vector denoting the state of the system, and a separate matrix representing the physical observable. Only using this vector and matrix can we calculate the probabilities of measurement outcomes and the average values of physical observables. In the standard version of quantum mechanics, this is all we can say about the actual physical system (more on that later in Chap. 10).

The key requirement for an observable is that its eigenvalues are the possible physical values of the corresponding physical property. For example, the eigenvalues of the spin operator are the possible values of the electron spin in a given direction (since spin is a vector), and the eigenvalues of the Hamiltonian are the possible energy values that a system can possess. Since these are real-world experimental values, they should be real numbers. After all, no physical measurement device can give you a complex number as a valid result. Our use of complex numbers is simply a mathematical convenience (quantum mechanics without complex numbers is possible, but much more complicated!).

We can view this requirement of real eigenvalues as the defining property of observables: a matrix is an observable if and only if it has real eigenvalues. We would then like to have a quick way of telling whether a matrix has real or complex eigenvalues. Of course, we can always calculate the actual eigenvalues, but this may be a lot of work. Luckily, there is another way.

First, consider the simplest possible matrix, namely a single number a. This is a 1×1 matrix. You know already how we can tell whether it is a real number: just check whether or not $a^* = a$. Next, let's try this same technique with the Pauli matrices:

$$\sigma_z^* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^* = \begin{pmatrix} 1^* & 0 \\ 0 & -1^* \end{pmatrix} = \sigma_z \tag{5.30}$$

and

$$\sigma_x^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^* = \begin{pmatrix} 0 & 1^* \\ 1^* & 0 \end{pmatrix} = \sigma_x \,, \tag{5.31}$$

so this clearly works for σ_x and σ_z. However, it does not work for σ_y :

$$\sigma_y^* = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^* = \begin{pmatrix} 0 & -i^* \\ i^* & 0 \end{pmatrix} = -\sigma_y \,. \tag{5.32}$$

Yet, the eigenvalues of σ_y are ± 1 and therefore σ_y is a proper observable since it has real eigenvalues. From a physical perspective you already know that σ_y should be an observable, because it is directly related to the spin in the y-direction.

So the complex conjugate is not a good indicator for telling whether the eigenvalues of a matrix are real. However, when we inspect the effect on σ_y, we can make a small modification that may work: In addition to taking the complex conjugate, we can also mirror the matrix in the main diagonal. In other words, the top left element swaps places with the bottom right. The diagonal elements stay where they are. This operation on a matrix is called the *transpose*, and you have already encountered this in Chap. 2, when we turned a column vector (a ket $|\cdot\rangle$) into a row vector (a bra $\langle\cdot|$) in order to calculate the scalar product. This operation also required us to take the complex conjugate in order to arrive at the normalisation condition $|a|^2 + |b|^2 = 1$. The combination of taking the transpose and the complex conjugate is called the adjoint, and is denoted by a † as a superscript.

An important mathematical theorem states that a matrix M has real eigenvalues if and only if $M = M^\dagger$. A matrix with this property is called *Hermitian*. It then follows that all physical observables are described by Hermitian matrices (or Hermitian operators, since we have a direct correspondence between matrices and operators).

In quantum mechanics, every Hermitian matrix is considered a valid physical observable, even though it may not be obvious how the matrix relates to an experimental setup. For example, in the case of the Stern–Gerlach experiment we can rotate the apparatus around the y-axis from alignment in the z-direction to the x-direction,

and we can measure any spin direction in the xz-plane. However, we cannot rotate the magnets so as to measure the y-component, because the magnets would block the path of the electrons. Nevertheless, it makes intuitive sense to talk about the spin in the y-direction, and there are other ways to measure this component, involving more complicated setups.

For the simplest case of a 2×2 matrix we can construct all possible observables in terms of the Pauli matrices and the identity matrix $\mathbb{I}$. Since σ_x, σ_y, σ_z, and $\mathbb{I}$ are Hermitian, a linear combination of them is also Hermitian:

$$(a\sigma_x + b\sigma_y + c\sigma_z + d\mathbb{I})^\dagger = a^*\sigma_x + b^*\sigma_y + c^*\sigma_z + d^*\mathbb{I}\,, \tag{5.33}$$

so if a, b, c, and d are real the linear combination in Eq. (5.32) is Hermitian. This can be written as

$$M = \begin{pmatrix} c+d & a-ib \\ a+ib & -c+d \end{pmatrix}. \tag{5.34}$$

All Hermitian 2×2 matrices can be written like this, because $c + d$ and $-c + d$ can take on any real value, and $a + ib$ can take on any complex value and fully determines $a - ib$. The Pauli matrices are thus not just important matrices for describing the spin of an electron (up to a numerical factor of $\hbar/2$), but play a crucial role in the mathematics of 2×2 matrices.

We can take the Hermitian adjoint of a vector and a matrix:

$$\begin{aligned} M &\to M^\dagger \\ |\psi\rangle &\to (|\psi\rangle)^\dagger = \langle\psi|\,, \end{aligned} \tag{5.35}$$

as well as a product of two matrices:

$$AB \to (AB)^\dagger = B^\dagger A^\dagger\,, \tag{5.36}$$

so the adjoint applied to a product of two operators changes the order of the operators! This is important, because AB is generally not the same as BA. The Hermitian adjoint of an expectation value is

$$\langle\psi|A|\psi\rangle^\dagger = \langle\psi|A^\dagger|\psi\rangle\,, \tag{5.37}$$

which is the same as the complex conjugate, since the expectation $\langle\psi|A|\psi\rangle$ is a single complex number. From this you can work out what happens when you take the adjoint twice.

5.3 Evolution

We have seen that physical observables such as spin and energy are described by Hermitian operators, and that these operators have real eigenvalues. Can we give a similarly straightforward criterion that tells us whether an operator is a valid evolution? It turns out we can, by considering how evolved state vectors must remain normalised.

First, note that the expectation value of a (possibly Hermitian) matrix A is given by

$$\langle A \rangle = \langle \psi | A | \psi \rangle \,. \tag{5.38}$$

Next, let us evolve the state $|\psi\rangle$ to $U|\psi\rangle$, where U is the operator that describes the evolution. The evolved state must again be normalised, which restricts the form U can take. The new expectation value is then

$$\langle A \rangle' = \langle \psi | U^\dagger A U | \psi \rangle \,. \tag{5.39}$$

Now look at the special case where A is the identity operator ($A = \mathbb{I}$). In this case Eq. (5.38) becomes $\langle \psi | \mathbb{I} | \psi \rangle = \langle \psi | \psi \rangle = 1$, and Eq. (5.39) gives

$$\langle \psi | U^\dagger \mathbb{I} U | \psi \rangle = \langle \psi | U^\dagger U | \psi \rangle \,. \tag{5.40}$$

However, this is the same as the scalar product of the state $U|\psi\rangle$ with itself:

$$(\langle \psi | U^\dagger)(U|\psi\rangle) = 1 \,, \tag{5.41}$$

which is equal to 1 due to the normalisation condition. More importantly, this must be true for any state $|\psi\rangle$. We can also write $\langle \psi | \psi \rangle = 1$, and we therefore have

$$\langle \psi | U^\dagger U | \psi \rangle = \langle \psi | \psi \rangle \,, \tag{5.42}$$

Since this must be true for any $|\psi\rangle$ we can choose, this means that

$$U^\dagger U = \mathbb{I} \,, \tag{5.43}$$

or $U^\dagger$ is the inverse of U. Any U that takes any valid quantum state $|\psi\rangle$ to another valid quantum state is bound to obey this rule. These are called unitary matrices, and they describe the evolution of the quantum system. The same reasoning, but with U and $U^\dagger$ swapped, will give $U U^\dagger = \mathbb{I}$.

We now derive an important mathematical property about unitary operators and vectors. If we apply a unitary evolution to two different vectors, the scalar product between the vectors before and after the evolution remains the same:

$$\langle\phi|\psi\rangle \xrightarrow{U} (\langle\phi|U^\dagger)(U|\psi\rangle) = \langle\phi|U^\dagger U|\psi\rangle = \langle\phi|\psi\rangle\,. \tag{5.44}$$

In particular, this means that orthogonal states evolve into new states that are again orthogonal. This was the assumption we physically motivated when we worked out the matrix form of the beam splitter in Chap. 2.

As we have seen in Eq. (4.29), we can write the time evolution as an exponential function of the Hamiltonian:

$$U(t) = \exp\left(-\frac{i}{\hbar}Ht\right)\,. \tag{5.45}$$

This is a function of a matrix, and it is itself a matrix. The inverse of this matrix is given by the Hermitian adjoint

$$U(t)^{-1} = U(t)^\dagger = \exp\left(\frac{i}{\hbar}H^\dagger t\right) = \exp\left(\frac{i}{\hbar}Ht\right)\,, \tag{5.46}$$

where we used that $H^\dagger = H$. Therefore, it is easy to find the inverse of a unitary matrix when it is given in the form of Eq. (5.45).

5.4 The Commutator

We have noted earlier that the product of two matrices AB is generally not the same as BA. As a quick example, we can multiply the Pauli σ_x and σ_y matrices in two ways:

$$\begin{aligned}\sigma_x\sigma_y &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} = \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix} = i\sigma_z,\\ \sigma_y\sigma_x &= \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix}-i & 0\\ 0 & i\end{pmatrix} = -i\sigma_z,\end{aligned} \tag{5.47}$$

and see that the two products $\sigma_x\sigma_y$ and $\sigma_y\sigma_x$ are not the same. They differ by a minus sign. We can capture this difference by subtracting $\sigma_y\sigma_x$ from $\sigma_x\sigma_y$:

$$\sigma_x\sigma_y - \sigma_y\sigma_x = 2i\sigma_z\,. \tag{5.48}$$

In general, for two operators A and B we define

$$[A, B] = AB - BA\,. \tag{5.49}$$

The quantity $[A, B]$ is called the commutator of A and B, and when $[A, B] = 0$, we say that the operators A and B commute: it does not matter in which order they are

written down. However, when $[A, B] \neq 0$, the order does make a difference, and you should always be careful when you calculate with operators in quantum mechanics!

For unitary operators it is not so surprising that the order of the operation matters. Suppose that operator U describes "putting on my socks", while operator V describes"putting on my shoes". It makes a difference whether I apply U first and then V, or V first and then U. The difference is between looking normal and looking silly.

However, the interpretation is not so clear for observables. These are also described by matrices and therefore also do not commute in general. As an example, the Pauli matrices above are Hermitian and represent (up to a factor $\hbar/2$) the spin of an electron in the x, y and z-direction. What does it mean that the spin observables in different directions do not commute? To answer this, we first note that two diagonal matrices of the same dimension always commute. For example,

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} ac & 0 \\ 0 & bd \end{pmatrix}. \tag{5.50}$$

In general, matrices commute if and only if they have the same eigenvectors. This is a key property of matrices that holds true in general, with a subtle modification when two or more eigenvalues of the operators are the same (but we will not consider those cases here).

The second key property of quantum mechanics is that a measurement of an observable will put the state of the quantum system in the eigenstate corresponding to the measurement outcome. If we measure the spin of an electron in the z-direction (S_z) and find spin $\uparrow$, a second measurement of S_z will always give spin $\uparrow$ again. That is, the measurement outcome $\uparrow$ has probability one, and $\downarrow$ has probability zero. So if two observables A and B commute, you can measure A and leave the system in an eigenstate of A. Then when you measure B, the system is already in an eigenstate of B (because A and B share the same eigenvectors) and you will find the corresponding measurement outcome (the eigenvalue) with probability one. Moreover, the state of the system is not disturbed by the measurement of B. Consequently, you can measure A again after B, and because the system is still in the same eigenstate you will again find the same measurement outcome for A as you did the first time.

If A and B do *not* commute, the measurement of A leaves the system in an eigenstate of A, and a subsequent measurement of B will leave the system in an eigenstate of B. A third measurement—of A again—will now no longer be guaranteed to give the same measurement outcome as the first measurement of A. Take, for example, the spin of an electron in the x and the z-direction. These observables do not commute. Suppose that we measure first spin $\uparrow$ in the z-direction. The state of the spin after the measurement is then $|\uparrow\rangle$. We can calculate the probabilities of the spin measurement outcomes $+$ and $-$ in the x-direction as

$$p_+ = |\langle + | \uparrow \rangle|^2 = \frac{1}{2} \quad \text{and} \quad p_- = |\langle - | \uparrow \rangle|^2 = \frac{1}{2}. \tag{5.51}$$

So the measurement outcome of the second observable, S_x, is uncertain. Moreover, performing a third measurement, again in the z-direction, will not yield outcome $\uparrow$ with probability one, but will rather be $p_\uparrow = p_\downarrow = 1/2$. The measurement of S_x disturbs the quantum system if it is in an eigenstate of S_z.

So when two observables commute, their measurements can be carried out in different orders, or even simultaneously, which makes no difference for the actual measurement outcomes. However, when the observables do not commute, each measurement disturbs the quantum system, and repeated measurements will generally not yield the same measurement outcomes (e.g., the first measurement of S_z gave $\uparrow$, but after S_x the second measurement of S_z gives either $\uparrow$ or $\downarrow$). This means that two observables can be measured together if they commute (they are compatible), but not if they do not commute.

5.5 Projectors

The scalar product of two state vectors $|\psi\rangle$ and $|\phi\rangle$ is given by $\langle\phi|\psi\rangle$. You can also calculate $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$. The easiest way to calculate the scalar products is via the vector form of $|\psi\rangle$ and $|\phi\rangle$. For example, if

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{and} \quad |\phi\rangle = \begin{pmatrix} c \\ d \end{pmatrix}, \tag{5.52}$$

then the scalar product $\langle\phi|\psi\rangle$ is

$$\langle\phi|\psi\rangle = \begin{pmatrix} c^* & d^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = c^* a + d^* b, \tag{5.53}$$

where a, b, c, and d are complex numbers. You should check that $\langle\psi|\phi\rangle$ is indeed the complex conjugate of $\langle\phi|\psi\rangle$.

These scalar products are used to calculate probabilities of measurement outcomes, for example when $|\psi\rangle$ is the spin state of an electron and $|\phi\rangle$ is the state $|\uparrow\rangle$ the probability of finding measurement outcome $\uparrow$ in a Stern–Gerlach experiment is given by

$$p_\uparrow = |\langle\uparrow|\psi\rangle|^2 . \tag{5.54}$$

From geometry you know that the scalar product measures the overlap of two vectors, or more precisely, how much of one vector lies in the direction of the other. It so happens that it is very convenient to describe this in terms of projection operators, or projectors for short (see Fig. 5.2).

Let's consider a vector $\mathbf{v}$ in ordinary three-dimensional space. It can be written as a column of three numbers

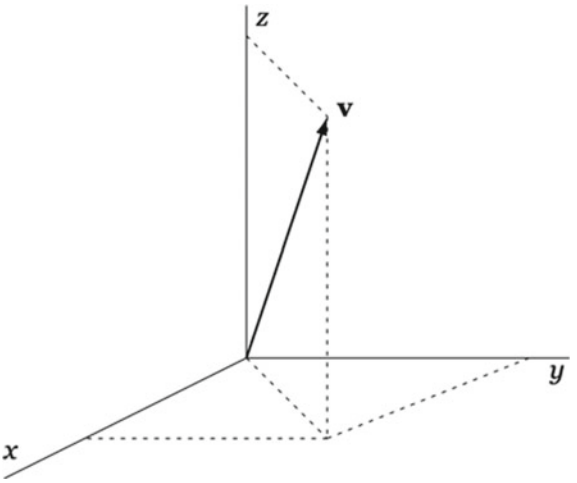


Fig. 5.2 Projections of vectors. The interactive figure is available online (see supplementary material 1)

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} , \tag{5.55}$$

which are the coordinates with respect to some coordinate system. We can project this vector onto the x-axis by setting $v_y = v_z = 0$. This is an operation on the vector, and we therefore expect that we can write this as a matrix. Indeed, it is straightforward to verify that the matrix

$$P_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{5.56}$$

does the job for any value of v_x, v_y, and v_z:

$$P_x \mathbf{v} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} v_x \\ 0 \\ 0 \end{pmatrix} . \tag{5.57}$$

Similarly, we can construct projectors onto the y- and z-axis:

$$P_y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad P_z = \begin{pmatrix} 0\,0\,0 \\ 0\,0\,0 \\ 0\,0\,1 \end{pmatrix} . \tag{5.58}$$

You should check that

$$P_y\mathbf{v} = \begin{pmatrix} 0 \\ v_y \\ 0 \end{pmatrix} \quad \text{and} \quad P_z\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ v_z \end{pmatrix} . \tag{5.59}$$

We can also project vectors onto planes instead of axes. Projecting the vector $\mathbf{v}$ onto the xy-plane is achieved by setting $z = 0$. The projector P_{xy} that accomplishes this is

$$P_{xy} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{5.60}$$

Similarly, projections onto the xz- and yz-planes are given by

$$P_{xz} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad P_{yz} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.61}$$

Finally, the projection onto the entire three-dimensional space is the identity operator.

One of the key properties of projectors is that the square of the projector is equal to itself: $P^2 = P$. You should try to prove that the eigenvalues of the matrices must therefore be 0 or 1. Indeed, the projectors we constructed above all obey this rule. This means that projectors are Hermitian operators: $P^\dagger = P$.

In the quantum mechanical example of finding the measurement outcome $\uparrow$ for the spin of an electron in a Stern–Gerlach experiment we calculate the probability by projecting the state vector of the electron onto the eigenstate $|\uparrow\rangle$ and determining the length of the resulting vector. Suppose that the state vector of the electron before the measurement is given by

$$|\psi\rangle = a|\uparrow\rangle + b|\downarrow\rangle = \begin{pmatrix} a \\ b \end{pmatrix} , \tag{5.62}$$

where we used

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{5.63}$$

Then, as we know, the probability of finding measurement outcome $\uparrow$ is calculated via

$$p_\uparrow = |\langle\uparrow|\psi\rangle|^2 = \langle\psi|\uparrow\rangle\langle\uparrow|\psi\rangle = |a|^2 . \tag{5.64}$$

Alternatively, we could use our new knowledge about projectors: We project the state $|\psi\rangle$ onto the state vector corresponding to the measurement outcome $\uparrow$ using

a projector $P_\uparrow$ and calculate the length-squared of the resulting vector $|\phi_\uparrow\rangle = P_\uparrow|\psi\rangle$. But calculating the length-squared of a vector is nothing more than taking the scalar product with itself:

$$p_\uparrow = \langle\phi_\uparrow|\phi_\uparrow\rangle = \langle\psi|P_\uparrow^\dagger P_\uparrow|\psi\rangle\,. \tag{5.65}$$

Simplifying this using the rules $P^\dagger = P$ and $P^2 = P$, and comparing to Eq. (5.64), we find that

$$p_\uparrow = \langle\psi|P_\uparrow|\psi\rangle = \langle\psi|\uparrow\rangle\langle\uparrow|\psi\rangle = |a|^2\,. \tag{5.66}$$

Since this is true for all $|\psi\rangle$, we find that

$$P_\uparrow = |\uparrow\rangle\langle\uparrow|\,. \tag{5.67}$$

We can verify that this is true by calculating $P_\uparrow$ using the vector form of $|\uparrow\rangle$ and $\langle\uparrow|$:

$$P_\uparrow = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\,. \tag{5.68}$$

Acting with $P_\uparrow$ on $|\psi\rangle$ gives

$$P_\uparrow|\psi\rangle = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}a\\0\end{pmatrix}\,, \tag{5.69}$$

projecting the state vector $|\psi\rangle$ onto the axis defined by the vector $|\uparrow\rangle$.

Using Eq. (5.67) we now see how we can construct projectors onto an arbitrary axis defined by a vector $|\phi\rangle$. We just take the "outer" product $P_\phi = |\phi\rangle\langle\phi|$. You can easily verify that

$$P_\phi^2 = P_\phi \quad\text{and}\quad P_\phi^\dagger = P_\phi\,, \tag{5.70}$$

and the operator P_ϕ is therefore a projector. For a vector $|\phi\rangle = \begin{pmatrix}a\\b\end{pmatrix}$ with $a = i/\sqrt{3}$ and $b = \sqrt{2/3}$, the projector in matrix form is given by

$$P_\phi = |\phi\rangle\langle\phi| = \begin{pmatrix}\frac{i}{\sqrt{3}}\\ \sqrt{\frac{2}{3}}\end{pmatrix}\begin{pmatrix}-\frac{i}{\sqrt{3}} & \sqrt{\frac{2}{3}}\end{pmatrix} = \begin{pmatrix}\frac{1}{3} & \frac{i\sqrt{2}}{3}\\ -\frac{i\sqrt{2}}{3} & \frac{2}{3}\end{pmatrix}\,. \tag{5.71}$$

You should explicitly verify for this matrix that $P_\phi^2 = P_\phi$ and $P_\phi^\dagger = P_\phi$.

Now that we have constructed a projector for the eigenvector $|\uparrow\rangle$ for the measurement outcome $\uparrow$, nothing prevents us from doing the same thing for the state vector $|\psi\rangle$, so that we obtain

$$P_\psi = |\psi\rangle\langle\psi| \,. \tag{5.72}$$

When we're after the overlap of the state vector with the eigenvector $|\uparrow\rangle$, we can project both the ket and the bra in P_ψ onto $|\uparrow\rangle$:

$$\begin{aligned} P_\uparrow P_\psi \; P_\uparrow^\dagger &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} |a|^2 & 0 \\ a^*b & 0 \end{pmatrix} = \begin{pmatrix} |a|^2 & 0 \\ 0 & 0 \end{pmatrix} . \end{aligned} \tag{5.73}$$

Similarly, we can calculate the projection of $P_\downarrow$ as

$$\begin{aligned} P_\downarrow P_\psi \; P_\downarrow^\dagger &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & |b|^2 \end{pmatrix} . \end{aligned} \tag{5.74}$$

Somehow we want to convert these two matrices to the numbers $|a|^2$ and $|b|^2$, respectively. You see that we can accomplish this by adding the diagonal elements. For $P_\uparrow P_\psi \; P_\uparrow^\dagger$ this gives $|a|^2 + 0 = |a|^2$, and for $P_\downarrow P_\psi \; P_\downarrow^\dagger$ this gives $0 + |b|^2 = |b|^2$.

There is a matrix operation called the trace that does exactly what we want, namely summing over all the diagonal elements. We denote the trace of a matrix M by Tr[M], so we have

$$\mathrm{Tr}\,(P_\uparrow P_\psi \; P_\uparrow^\dagger) = |a|^2 \quad \text{and} \quad \mathrm{Tr}(\, P_\downarrow P_\psi \; P_\downarrow^\dagger) = |b|^2 \,. \tag{5.75}$$

A very useful property of the trace is the cyclic property:

$$\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(CAB) \,. \tag{5.76}$$

We can use this, and the properties of the projectors, to write the probabilities of the measurement outcomes as

$$p_\uparrow = \mathrm{Tr}\,(P_\uparrow \; P_\psi) \quad \text{and} \quad p_\downarrow = \mathrm{Tr}\,(P_\downarrow \; P_\psi) \,. \tag{5.77}$$

You should verify that this is true in the example above.

Finally, let's consider the expectation value of the spin operator S_z. In terms of the eigenvectors $|\uparrow\rangle$ and $|\downarrow\rangle$ this can be written as

$$S_z = \frac{\hbar}{2}|\uparrow\rangle\langle\uparrow| - \frac{\hbar}{2}|\downarrow\rangle\langle\downarrow| = \frac{\hbar}{2}P_\uparrow - \frac{\hbar}{2}P_\downarrow \,. \tag{5.78}$$

The expectation value of S_z with respect to the state vector $|\psi\rangle$ can be written as

$$\begin{aligned}\langle S_z\rangle &= \langle\psi|S_z|\psi\rangle = \frac{\hbar}{2}\langle\psi|P_\uparrow|\psi\rangle - \frac{\hbar}{2}\langle\psi|P_\downarrow|\psi\rangle \\ &= \frac{\hbar}{2}\mathrm{Tr}\,(P_\uparrow\, P_\psi) - \frac{\hbar}{2}\mathrm{Tr}\,(P_\downarrow\, P_\psi) \\ &= \mathrm{Tr}\left(\frac{\hbar}{2}P_\uparrow - \frac{\hbar}{2}P_\downarrow\right)P_\psi = \mathrm{Tr}\,S_z\, P_\psi \,. \end{aligned} \tag{5.79}$$

This is a general rule: for any operator A the expectation value is given by

$$\langle A\rangle = \mathrm{Tr}\,(A\, P_\psi)\,. \tag{5.80}$$

You may wonder what's the point of these mathematical exercises. All we gained is a different way to calculate quantities that we already knew how to calculate. However, it turns out that it is very convenient to use projectors when we deal with decoherence in Chap. 7.

Exercises

1. Calculate the determinant of the following matrices:

$$\text{(a)}\ \begin{pmatrix}3 & 0\\ 4 & -1\end{pmatrix}, \qquad \text{(b)}\ \begin{pmatrix}2i & 5\\ 1 & 7\end{pmatrix}, \qquad \text{(c)}\ \begin{pmatrix}1 & 0 & -1\\ 2 & 3 & 8\\ 2 & 9 & 5\end{pmatrix}.$$

2. Calculate the eigenvalues and eigenvectors of the Pauli matrices σ_x, σ_y and σ_z. Do your results agree with the spin vectors in Chap. 3? Are the Pauli matrices Hermitian and/or unitary?
3. Prove that the eigenvalues of a unitary operator are complex numbers with modulus 1. In other words, they are of the form $e^{i\phi}$ with ϕ in $[0, 2\pi)$.

4. Indicate whether the following matrices are Hermitian, unitary, or both:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \qquad \begin{pmatrix}3 & 4-2i\\ 4-2i & 5\end{pmatrix}, \qquad \frac{1}{2}\begin{pmatrix}1 & e^{-3t}\\ e^{-3t} & e^{-t}\end{pmatrix}.$$

5. For an arbitrary Hamiltonian H whose square is not proportional to the identity matrix, for example the one given in Eq. (4.32), it is not so easy to calculate the unitary evolution $U = \exp(-iHt/\hbar)$ directly. We will calculate U in two steps:

   (a) Let V be the unitary matrix that diagonalises H such that $H_D = VHV^{-1}$ is a diagonal matrix. Show that

$$U = Ve^{-\frac{i}{\hbar}H_D t}V^{-1}.$$

   (b) Let $H = \frac{1}{2}\hbar\omega\sigma_z + g\hbar\sigma_x$, with g a real number. Find the matrix V and calculate U.

6. An electron with spin 'up' in the positive z-direction ($|\uparrow\rangle$) is measured using a Stern–Gerlach apparatus oriented in the x-direction.

   (a) Before the electron spin is measured, it evolves for a time T due to a uniform magnetic field in the y-direction. The Hamiltonian can be written as

$$H = -\frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B},$$

   where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the standard Pauli matrices, e is the charge of the electron, m its mass, and $\mathbf{B}$ is the magnetic field. Calculate the new spin state of the electron entering the Stern–Gerlach apparatus.

   (b) Calculate the probabilities of the measurement outcomes of the Stern–Gerlach apparatus for the spin state calculated in part (a).

   (c) Sketch the evolution of the electron spin state in part (b) in the Bloch sphere, and indicate the essential elements.

7. A photon in a polarisation state $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$ is measured in the polarisation basis $\{|H\rangle, |V\rangle\}$.

   (a) If we measure an identically prepared photon in the *circular* polarisation basis $(|H\rangle \pm i|V\rangle)/\sqrt{2}$, what will be the measurement outcomes and their corresponding probabilities?

(b) The free evolution of the photon is governed by the Hamiltonian

$$H = i\hbar\omega(|H\rangle\langle V| - |V\rangle\langle H|)\,,$$

with $\hbar$ Planck's constant and ω the frequency of the light. Find the state of the photon at time $t = T$ given that the state at $t = 0$ is given by $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$.

(c) Calculate and sketch the probability of finding the outcomes of a measurement of the photon polarisation in the basis $\{|H\rangle, |V\rangle\}$ in the time interval $0 \leq t \leq T$ when $\alpha = \beta = 1/\sqrt{2}$.

8. Calculate P^2 with P given by

$$P = \begin{pmatrix} a \\ b \end{pmatrix} (a^*, b^*) = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \quad \text{and} \quad |a|^2 + |b|^2 = 1.$$

9. Show that two operators commute *if and only if* they have the same eigenvectors.
10. Show that the projectors on eigenstates of σ_x do not commute wth the projectors on eigenstates of σ_y.
11. Consider three (possibly non-orthogonal) states $|a\rangle$, $|b\rangle$, and $|c\rangle$ for an electron spin.

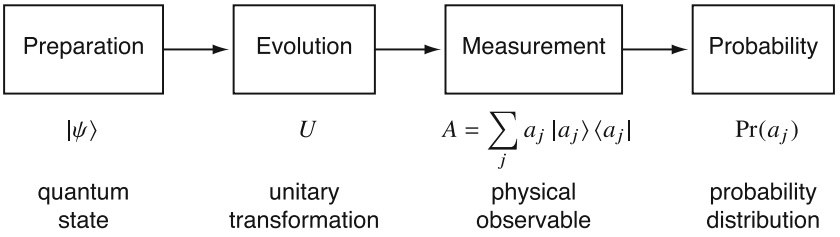
(a) Construct the projectors onto the three states, and write down the formulas for the probabilities that the electron has spin a, b, c in terms of these projectors.

(b) What the probability that the electron does *not* have spin a, and what is the projector associated with this situation?

(c) What is the probability that the electron has spin a *or* b, and what is the projector associated with this situation?

The Rules of Quantum Mechanics

We have arrived at the halfway point of this book, and it is time to take stock of what we have learned so far. We introduced the quantum state of a system, which can be seen as a mathematical short-hand for how the system was prepared. After preparation, the system may evolve to some other state, either as a change over time, or as a change in the motion, for example due to a beam splitter. Finally, we choose to measure a physical observable of the quantum system (like spin or energy), and the measurement outcome indicates the value of the observable. For example, an upward-deflected electron in a Stern–Gerlach apparatus indicates the spin value $\hbar/2$ in that direction. We find the probabilities of the measurement outcomes by taking the squared modulus of the scalar product between the state and the eigenvector associated with the measurement outcome. We can schematically represent the structure of quantum mechanics as follows:



We will capture quantum mechanics in five rules that will completely determine the theory. These are often also called the *postulates of quantum theory*:

1. The state of a quantum system is described by a vector $|\psi\rangle$ of length $\langle\psi|\psi\rangle = 1$. The vector components are complex numbers and carry no units. The state does not directly describe the physical properties of a system; for that we need observables.
2. Physical observables are described by Hermitian operators (matrices) A, which obey the rule $A^\dagger = A$. This relation ensures that the eigenvalues $a_1, \ldots, a_n$ of A are real, so they can be interpreted as the possible physical values of the observable. The eigenvalues therefore carry units. Any observable A can be written in the spectral decomposition

$$A = \sum_j a_j |a_j\rangle\langle a_j| \,,$$

   where the $|a_j\rangle$ are the eigenstates of a_j and $|a_j\rangle\langle a_j|$ are projectors onto the eigenstates. These states do not carry units.
3. Changes in the state of the quantum system—over time or otherwise—are described by unitary operators U. An operator or matrix is unitary if $U^\dagger = U^{-1}$. We can often write $U(t) = \exp(-iHt/\hbar)$, where H is the Hamiltonian, or energy operator of the system. This leads to the Schrödinger equation for the quantum state:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle\,.$$

If U is known, we can write the evolved state directly as $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ and solve this as a matrix equation, where $|\psi(t)\rangle$ and $|\psi(0)\rangle$ are vectors, and $U(t)$ is a matrix.

4. The probability of finding measurement outcome a_j during the measurement of an observable A with eigenvalues $a_1, \ldots, a_n$ for a system in the quantum state $|\psi\rangle$ is given by the square modulus of the scalar product between the state and the eigenvector:

$$\Pr(a_j) = |\langle a_j|\psi\rangle|^2\,.$$

This is called the "Born rule". We can calculate the expectation value of A for the system in state $|\psi\rangle$ by taking the weighted sum over the eigenvalues:

$$\begin{aligned}\langle A\rangle &= \sum_j a_j \Pr(a_j) = \sum_j a_j \langle\psi|a_j\rangle\langle a_j|\psi\rangle \\ &= \langle\psi|\left(\sum_j a_j|a_j\rangle\langle a_j|\right)|\psi\rangle = \langle\psi|A|\psi\rangle\,,\end{aligned}$$

5. Immediately after the measurement of an observable, the system is in the eigenstate $|a_j\rangle$ corresponding the eigenvalue a_j found in the measurement. Measuring the observable for a second time right after the first will give the same outcome a_j with probability $\Pr(a_j) = 1$. This is called the "Projection Postulate".

The material we have covered so far forms the basis of quantum mechanics, and in the five chapters that follow we will use the theory to explore more advanced topics, such as the description of composite systems, decoherence (where we lose quantum behaviour), a particle moving in space, the uncertainty relations, and how we can make sense of quantum mechanics.