Chapter 4
Complex Vector Spaces and Quantum Mechanics

In our complex vector space, we can define a scalar product. The properties of orthogonality and completeness lead to the important concept of a complete orthonormal system. The measurement process can be formulated by means of suitable projection operators.

Up to now, we have occasionally used the terms ‘vector space’ or ‘state space’. In this chapter, we will address this concept in more detail. For reasons of simplicity, we will rely heavily on the example of *polarization*, where the basic formulations are of course independent of the specific realization and are valid for all two-dimensional state spaces (such as polarization states, electron spin states, a double-well potential, the ammonia molecule, etc.). Moreover, the concepts introduced here retain their meaning in higher-dimensional state spaces, as well. Therefore, we can introduce and discuss many topics by using the example of the simple two-dimensional state space. From the technical point of view, this chapter is about the discussion of some of the elementary facts of complex vector spaces. The basic definitions are given in Appendix G, Vol. 1.[1]

In Chap. 2, we introduced the polarization states which also apply to single photons[2]

$$\begin{array}{c} |h\rangle \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |v\rangle \cong \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ |r\rangle \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}; \quad |l\rangle \cong \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \end{array}. \tag{4.1}$$

These vectors are obviously elements of a two-dimensional complex vector space $\mathcal{V}$. In fact, one can convince oneself that all the axioms which apply to a vector space are satisfied; see Appendix G, Vol. 1. To put it simply, these axioms state in the end that one can perform all operations as usual—one can add vectors

[1] Of course, we treat these technical aspects not as an end in themselves, but because they are of fundamental importance for the physical description of natural phenomena in the context of quantum mechanics.

[2] For the notation $\cong$, see Chap. 2.

J. Pade, *Quantum Mechanics for Pedestrians 1*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-030-00464-4_4

and multiply them by a number, subject to the familiar rules such as the distributive law, etc. We note in this context that products of numbers and vectors commute, so that $c \cdot |z\rangle = |z\rangle \cdot c$ holds. Although the notation $|z\rangle \cdot c$ is perhaps unfamiliar, it is nevertheless absolutely correct.

Especially important is the fact that the elements of a vector space can be *superposed*—if $|x\rangle$ and $|y\rangle$ are elements of the vector space, then so is $\lambda |x\rangle + \mu |y\rangle$ with $\lambda, \mu \in \mathbb{C}$. In our example of polarization, this means that *each* vector (except the zero vector) represents a viable physical state.[3] This superposition principle[4] is anything but self-evident—just think for example of the state space which consists of all positions that are reachable in a chess game beginning from the starting position. Obviously, here the superposition principle does not hold, since the multiplication of such a state with a number or the addition or linear combination of states is simply not meaningful. Another example is the phase space of classical mechanics, in which the states are denoted by points—the addition of these points or states is not defined.

We will repeatedly come across the central importance of the superposition principle in quantum mechanics in the following sections and chapters.

4.1 Norm, Bra-Ket Notation

The familiar visual space $\mathbb{R}^3$ has the pleasant property that one can calculate the length of a vector and the angle between two vectors, namely by means of the scalar product. We want to implement these concepts also in the complex vector space, at least to some extent.

Following the familiar formula, the length L of the vector $\begin{pmatrix} a \\ b \end{pmatrix}$ would be $L^2 = a^2 + b^2$. But this is wrong, since the vector space is *complex*. Accordingly, due to $1 + i^2 = 0$, the vector $\begin{pmatrix} 1 \\ i \end{pmatrix}$ would have zero length, which evidently makes no sense.[5] Instead, the correct formula reads

$$L = |a|^2 + |b|^2 = aa^* + bb^*. \tag{4.2}$$

Making use of the usual rules of matrix multiplication, we can write this as the product of a row vector with a column vector[6]:

[3] Later on, we will meet vector spaces where this is no longer the case; keyword 'identical particles' or 'superselection rules'.

[4] We note that the superposition principle contains three pieces of information: (1) The multiplication of a state by a scalar is meaningful. (2) The addition of two states is meaningful. (3) Every linear combination of two states is again an element in the vector space.

[5] As we know, only the zero vector has length zero.

[6] We recall that * means complex conjugation.

$$L^2 = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \tag{4.3}$$

The space of the row vectors is called the dual space to $\mathcal{V}$. One obtains the vector $\begin{pmatrix} a^* & b^* \end{pmatrix}$ from the corresponding column vector by complex conjugation and inverting the roles of column and row (= transposing, symbol ^{T}). By this process, we obtain the *adjoint*,[7] which is denoted by a kind of superscripted cross

$$\begin{pmatrix} a^* & b^* \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}^{*T} = \begin{pmatrix} a \\ b \end{pmatrix}^{\dagger}. \tag{4.4}$$

The operation is analogously defined for general $n \times m$-matrices: the adjoint is always obtained by complex conjugation and transposition. We note that the adjoint is a *very* important term in quantum mechanics.

We have denoted the elements of the vector space using the short-hand notation $|\ \rangle$. Analogously, we choose for the elements of the dual space the notation $\langle\ |$. The symbols are defined as follows:

$$\begin{aligned} &|\ \rangle \quad \text{is called a } \textit{ket} \\ &\langle\ | \quad \text{is called a } \textit{bra}. \end{aligned} \tag{4.5}$$

This is the so-called *bra-ket notation* (from bracket = bra-(c)-ket), or *Dirac notation*, named after P.A.M. Dirac who first introduced it.[8] We have for example

$$\begin{aligned} |h\rangle^{\dagger} = \langle h| \ \text{ or } \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{\dagger} &= \begin{pmatrix} 1 & 0 \end{pmatrix} \\ \langle r|^{\dagger} = |r\rangle \ \text{ or } \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \end{pmatrix}^{\dagger} &= \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}. \end{aligned} \tag{4.6}$$

With these concepts we can now define the length L of a vector $|z\rangle$ as $L^2 = \langle z|\, z\rangle$ (actually, one would expect to write $\langle z|\,|z\rangle$, but the double bar is omitted). Instead of length, the term *norm* is generally used. The designations are $\|\ \|$ or equivalently $|\ \ |$. For example, we have

$$\| |h\rangle \|^2 = \langle h|\, h\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot 1 + 0 \cdot 0 = 1 \tag{4.7}$$

and correspondingly for $|r\rangle$

---

7 Strictly speaking, there are two adjoints. The one considered here is called *Hermitian adjoint*; it applies so to say in non-relativistic considerations. In the relativistic case, there is another kind, called *Dirac adjoint* which is defined differently. The bulk of the book is devoted to non-relativistic considerations; here adjoint means always Hermitian adjoint.

8 In the bra-ket notation, one cannot identify the dimension of the corresponding vector space (the same holds true for the familiar vector notations **v** or $\overrightarrow{v}$, by the way). If necessary, this information must be given separately.

$$\| |r\rangle \|^2 = \langle r|\, r\rangle = \frac{1}{2}\begin{pmatrix} 1 & -i \end{pmatrix}\begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{2}(1 \cdot 1 - i \cdot i) = 1. \tag{4.8}$$

Both vectors have the length 1. Such vectors are called *unit vectors*; they are *normalized*. The term $\langle z|\, z\rangle$ is a *scalar product* (also called inner product or dot product); more about this topic is to be found in Chap. 11 and in Appendix G, Vol. 1. We remark that we can use an equals sign in (4.7) and (4.8) instead of $\cong$, since scalar products are *independent of the representation*.

A comment on the nomenclature: A complex vector space in which a scalar product is defined is called a *unitary space*.

4.2 Orthogonality, Orthonormality

Now that we know how to calculate the length of a vector, the question of the angle between two vectors still remains open. First, we note that we can also form inner products of different vectors, for example

$$\langle v|\, r\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{i}{\sqrt{2}} \tag{4.9}$$

Note: As with any scalar product, $\langle a|\, b\rangle$ is a (generally complex) *number*. For the adjoint of an inner product, for example, we have

$$(\langle v|\, r\rangle)^\dagger = \langle r|\, v\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{i}{\sqrt{2}}. \tag{4.10}$$

So to form the adjoint of an expression we follow the procedure: (1) A number is replaced by its complex conjugate, $c^\dagger = c^*$; (2) A ket is replaced by the corresponding bra and *vice versa*; (3) The order of terms is reversed, for example $\langle a|\, b\rangle^\dagger = \langle a|\, b\rangle^* = \langle b|\, a\rangle$.

Regarding the question of the angle, in the following only one particular angle plays a role (apart from the angle zero), namely the right angle. One says that two vectors are *orthogonal* if their scalar product vanishes (this convention of terminology is valid also for non-intuitive higher-dimensional complex vector spaces). An example:

$$\langle v|\, h\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0. \tag{4.11}$$

More generally and in short form: $\langle a|\, b\rangle = 0 \leftrightarrow |a\rangle \perp |b\rangle$.

Note that the zero vector is orthogonal to itself and to all other vectors. Just as in the trivial solution of the SEq, it does not describe a physical state and is therefore in general not taken into account in considerations concerning orthogonality, etc.

Systems of vectors, all of which are normalized and pairwise orthogonal, play a special role.[9] Such a system of vectors is called an *orthonormal system* (*ONS*). Two-dimensional examples are the systems $\{|h\rangle, |v\rangle\}$ and $\{|r\rangle, |l\rangle\}$; an example from three-dimensional visual space $\mathbb{R}^3$ are the three unit vectors lying on the coordinate axes. The general formulation reads: $\{|\varphi_n\rangle, n = 1, 2, \ldots\}$ is an ONS if and only if

$$\langle\varphi_i|\,\varphi_j\rangle = \delta_{ij} \tag{4.12}$$

where the *Kronecker delta* (Kronecker symbol) is defined as usual by

$$\delta_{ij} = \begin{cases} 1 & \text{for} \quad i = j \\ 0 & \phantom{\text{for}} \quad i \neq j \end{cases}. \tag{4.13}$$

4.3 Completeness

We can write any vector $|z\rangle$ from our two-dimensional complex vector space as

$$|z\rangle = a\,|h\rangle + b\,|v\rangle \tag{4.14}$$

where $|h\rangle$ and $|v\rangle$ are orthonormal. Due to $\langle h|\,z\rangle = a\,\langle h|\,h\rangle + b\,\langle h|\,v\rangle = a \cdot 1 + b \cdot 0 = a$ (analogously for $\langle v|\,z\rangle$), this property leads to

$$\langle h|\,z\rangle = a \text{ and } \langle v|\,z\rangle = b. \tag{4.15}$$

We insert this and find[10]

$$\begin{aligned} |z\rangle &= \langle h|\,z\rangle\,|h\rangle + \langle v|\,z\rangle\,|v\rangle \\ &= |h\rangle\,\langle h|\,z\rangle + |v\rangle\,\langle v|\,z\rangle \\ &= \{|h\rangle\,\langle h| + |v\rangle\,\langle v|\}\,|z\rangle\,, \end{aligned} \tag{4.16}$$

or in other words (by comparing the left and right sides)[11]:

---

[9] In the two-dimensional vector space that we are currently addressing, such a system consists of course of two vectors; as stated above, the zero vector is excluded *a priori* from consideration.

[10] We repeat the remark that for products of numbers and vectors, it holds that $c \cdot |z\rangle = |z\rangle \cdot c$. Because $\langle h|\,z\rangle$ is a number, we can therefore write $\langle h|\,z\rangle\,|h\rangle$ as $|h\rangle\,\langle h|\,z\rangle$.

[11] In equations such as (4.17), the 1 on the right side is not necessarily the number 1, but is generally something that works like a multiplication by 1, i.e. a *unit operator*. For instance, this is the unit matrix when working with vectors. The notation 1 for the unit operator (which implies writing simply 1 instead of $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, for instance) is of course quite lax. On the other hand, as said before, the effect of multiplication by the unit operator and by 1 is identical, so that the small inaccuracy is generally accepted in view of the economy of notation. If necessary, 'one knows' that 1 means the unit operator. But there are also special notations for it, such as $\mathbb{E}$, I_n (where n indicates the

$$|h\rangle\langle h| + |v\rangle\langle v| = 1. \tag{4.17}$$

A term like $|x\rangle\langle y|$ is called a *dyadic product*. To get an idea of the meaning of such products, we choose the representation of row and column vectors. With

$$|h\rangle \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix};\ \langle h| \cong \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{4.18}$$

we have, according to the rules of matrix multiplication,

$$|h\rangle\langle h| \cong \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.19}$$

As can be seen, dyadic products are matrices or, more generally, operators which can be applied to states and usually change them.

With

$$|v\rangle\langle v| \cong \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.20}$$

it follows that

$$|h\rangle\langle h| + |v\rangle\langle v| \cong \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ (unit matrix).} \tag{4.21}$$

This equation, or (4.17), indicates that the ONS $\{|h\rangle, |v\rangle\}$ is *complete*, i.e. it spans the whole space. Consequently, $\{|h\rangle, |v\rangle\}$ is a *complete orthonormal system* (*CONS*). Another one is, for example, $\{|r\rangle, |l\rangle\}$ (see exercises). The terminology is transferred readily to n-dimensional vector spaces: A CONS consists of states $\{|\varphi_n\rangle, n = 1, 2, \ldots\}$ which are normalized and pairwise orthogonal (orthonormality), and which span the whole space (completeness)[12]:

$$\begin{array}{l} \langle \varphi_n | \varphi_m \rangle = \delta_{nm} \text{ (orthonormality)} \\ \sum_n |\varphi_n\rangle\langle\varphi_n| = 1 \text{ (completeness)} \end{array}. \tag{4.22}$$

With the methods developed so far, we can easily calculate the fractions of vertically—and horizontally-polarized light which are found e.g. in right circularly-polarized light. Of course, the example is simple enough to read off the answer directly from (4.1). But here, we are concerned with setting up a procedure that

dimension) and others. An analogous remark applies to the zero operator. By the way, we recall that in the case of vectors we write quite naturally $\vec{a} = 0$ and not $\vec{a} = \vec{0}$.

[12]For the summation we use almost exclusively the abbreviation $\sum_n$ (instead of $\sum_{n=1}^{\infty}$ or $\sum_{n=1}^{N}$ etc.). In the shorthand notation, the range of values of n must follow from the context of the problem at hand, if necessary.

works in any space. Basically, it is a multiplication by 1—but with 1 in a special notation. We have:

$$\begin{aligned} |r\rangle = 1 \cdot |r\rangle &\overset{4.17}{=} (|h\rangle \langle h| + |v\rangle \langle v|) \cdot |r\rangle \\ &= |h\rangle \langle h|\, r\rangle + |v\rangle \langle v|\, r\rangle = \frac{1}{\sqrt{2}} |h\rangle + \frac{i}{\sqrt{2}} |v\rangle, \end{aligned} \tag{4.23}$$

where we have used $\langle h|\, r\rangle = \frac{1}{\sqrt{2}}$ and $\langle v|\, r\rangle = \frac{i}{\sqrt{2}}$ in the final step.

With (4.23), we have formulated the state $|r\rangle$ in the basis $\{|h\rangle, |v\rangle\}$. This being the case, the coefficients $1/\sqrt{2}$ and $i/\sqrt{2}$ are none other than the coordinates of $|r\rangle$ with respect to $|h\rangle$ and $|v\rangle$. However, the term *coordinate* is used quite rarely in quantum mechanics; instead, one speaks of *projection*,[13] which is perhaps an even more descriptive term.

For higher dimensions, the following applies: Given a vector space $\mathcal{V}$ and a CONS $\{|\varphi_n\rangle, n = 1, 2, \ldots\} \in \mathcal{V}$, any vector $|\psi\rangle \in \mathcal{V}$ can be represented as

$$|\psi\rangle = 1 \cdot |\psi\rangle = \sum_n |\varphi_n\rangle \langle \varphi_n\, |\psi\rangle = \sum_n c_n |\varphi_n\rangle\, ; \;\; c_n = \langle \varphi_n\, |\psi\rangle \in \mathbb{C}. \tag{4.24}$$

The coefficients (coordinates) c_n are the projections of $|\psi\rangle$ onto the basis vectors $|\varphi_n\rangle$.

4.4 Projection Operators, Measurement

4.4.1 Projection Operators

As mentioned above, expressions like $|h\rangle \langle h|$ or $|\varphi_n\rangle \langle \varphi_n|$ act on states and are therefore *operators*. They are different from those that we met up with in the analytical approach of Chap. 3 (e.g. the derivative $\frac{\partial}{\partial x}$), but this is actually not surprising, since the states defined in the algebraic and the analytical approaches are quite different. We note, however, that there is a structure common to both approaches: in each case the states are elements of a vector space, and changes of these states are produced by operators.

The term $|h\rangle \langle h|$ is a particularly simple example of a *projection operator* (or projector). If P is a projection operator, we have[14]

$$P^2 = P. \tag{4.25}$$

---

[13]For the connection between inner product and projection, see Appendix F, Vol. 1.

[14]As we shall see in Chap. 13, a projection operator in quantum mechanics must meet a further condition (self-adjointness).

In fact, in the specific example $P = |h\rangle \langle h|$, we have, due to the normalization $\langle h | h\rangle = 1$, the equality

$$P^2 = |h\rangle \langle h | h\rangle \langle h| = |h\rangle \langle h| = P. \tag{4.26}$$

A further example of a projection operator is $|h\rangle \langle h| + |v\rangle \langle v|$, namely the projection onto the total space (because of $|h\rangle \langle h| + |v\rangle \langle v| = 1$).

The property $P^2 = P$ is actually very intuitive: if one filters out (= projects) a component of a total state by means of P, then a second projection does not change this component. In the matrix representation (4.19) with $P \cong \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, this reads as follows:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix};\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}. \tag{4.27}$$

4.4.1.1 Projection Operators and Measurement

Projection operators gain special importance from the fact that they can be used for the modelling of the measurement process. To see this, we start with a simple example, namely a right circularly-polarized state. Using (4.1), we write it as a superposition of linearly-polarized states:

$$|r\rangle = \frac{|h\rangle + i\,|v\rangle}{\sqrt{2}}. \tag{4.28}$$

We send this state $|r\rangle$ through an analyzer which can detect linearly-polarized states, e.g. a polarizing beam splitter (PBS). *Before* the measurement, we cannot say with certainty which one of the two linearly-polarized states we will measure. According to the considerations of Chap. 2, we can specify only the probabilities of measuring one of the states—in our example they are $\left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$ and $\left|\frac{i}{\sqrt{2}}\right|^2 = \frac{1}{2}$. We extend this idea to the more general state

$$|z\rangle = a\,|h\rangle + b\,|v\rangle;\ |a|^2 + |b|^2 = 1, \tag{4.29}$$

for which the probabilities of obtaining the vertically or horizontally polarized state are given by $|b|^2$ or $|a|^2$.

After the measurement, we have a different state from before the measurement, namely either $|h\rangle$ or $|v\rangle$.[15] Since states can be changed only by the action of operators, we have to model this transition by an operator. This modelling should be as simple and universal as possible, in order to be independent of the specific experimental details. Let us assume that we have the state $|h\rangle$ after the measurement.

[15] In other words, due to the process of measuring, a superposition such as $|z\rangle = a\,|h\rangle + b\,|v\rangle$ 'collapses' e.g. into the state $|h\rangle$.

Then we can describe this process by applying $|h\rangle\langle h|$ to $|z\rangle$, i.e. the projection of $|z\rangle$ onto $|h\rangle$, which leads to $|h\rangle\langle h|z\rangle = a|h\rangle$ (with an analogous formulation for $|v\rangle$). As a result of this 'operation', we obtain the desired state $|h\rangle$, but multiplied by a factor a, the absolute square of which gives the probability of obtaining that state in a measurement.

Hence, we can model the measurement process $|z\rangle \to |h\rangle$ as follows:

$$\underset{\text{before measurement}}{|z\rangle} = a|h\rangle + b|v\rangle \overset{\text{projection}}{\to} |h\rangle\langle h|(a|h\rangle + b|v\rangle)$$
$$= a|h\rangle \overset{\text{normalization}}{\to} \underset{\text{after measurement}}{\frac{|a|}{a}|h\rangle} \quad , \tag{4.30}$$

where we obtain the final result with probability $|a|^2$. Occasionally, it is assumed that one can set the normalization factor equal to 1 after the measurement, which formally means $\frac{|a|}{a} = 1$. As we said above, an analogous formulation applies to the measurement result $|v\rangle$.

4.4.1.2 Extension to Higher Dimensions

The generalization to dimensions $N > 2$ is straightforward. Before the measurement, the state is a superposition of different states, i.e. $|\psi\rangle = \sum c_n |\varphi_n\rangle$, where $\{|\varphi_n\rangle, n = 1, \ldots\}$ is a CONS. After the measurement, we have just one of the states, e.g. $|\varphi_i\rangle$. The measurement process is modelled by the projection operator $P_i = |\varphi_i\rangle\langle\varphi_i|$. With a slightly different notation, we have

$$|\psi\rangle_{\text{before}} = \sum_n c_n |\varphi_n\rangle \to |\varphi_i\rangle\langle\varphi_i|\psi\rangle_{\text{before}} = |\varphi_i\rangle\langle\varphi_i|\sum_n c_n|\varphi_n\rangle = c_i|\varphi_i\rangle\,. \tag{4.31}$$

The probability of measuring this state is thus given by $|c_i|^2 = |\langle\varphi_i|\psi\rangle|^2$. After the measurement, we have again a normalized state, namely

$$|\psi\rangle_{\text{after, normalized}} = \frac{P_i|\psi\rangle}{|P_i|\psi\rangle|} = \frac{c_i|\varphi_i\rangle}{|c_i|}. \tag{4.32}$$

We emphasize that the measurement process itself is *not* modelled, but only the situation immediately before and after the measurement. As an example, the situation is sketched in Fig. 4.1.

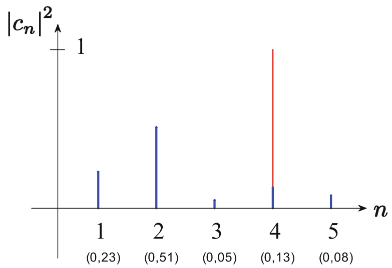

Fig. 4.1 Example sketch of the coefficients c_n in (4.31) and (4.32) before the measurement (blue) and after the measurement (red)

4.4.1.3 The Measurement Problem

Of course one may ask at this point, which mechanism picks out precisely the state $|\varphi_i\rangle$ from the superposition $\sum_n c_n |\varphi_n\rangle$, and not some other state. There is still no satisfactory answer to this question in spite of the advanced age of quantum mechanics. In fact, it is still an open problem, called the *measurement problem*. It is perhaps *the* conceptual problem of quantum mechanics. We shall meet it repeatedly in the following chapters. The different interpretations of quantum mechanics, at which we look closer in the last chapter of Vol. 2, are in some sense simply different ways of dealing with the measurement problem.

We note that the measurement problem has nothing to do with the extension to arbitrary dimensions, but applies even to the simplest systems. An example already treated in Chap. 2 and above is the right circularly-polarized photon, which we examine with respect to its possible linear polarization states. If we send

$$|r\rangle = \frac{|h\rangle + i\,|v\rangle}{\sqrt{2}} \tag{4.33}$$

through e.g. a PBS, we find either a horizontal or a vertical linearly-polarized photon, with probabilities $\frac{1}{2}$ in each case. Before the measurement, we cannot say which polarization we will obtain.

The key question is whether there is *in principle* such a selection mechanism. We have two alternatives. The first one: Yes, there is such a mechanism, although we do not know either the process (at present?) or the variables which it acts upon, the so-called *hidden variables*. If we knew these, we could describe the selection process that occurs during the measurement without any use of probabilities. The other alternative: No, there is no such mechanism. The selection of states in the course of the measurement is purely random—one speaks of *objective chance*.

The choice of the alternative in question must be decided experimentally. We have already noted in Chap. 2 that, all in all, relevant experiments do not support the existence of hidden variables. Therefore, we hereafter assume the existence of objective chance, but we will take up the measurement problem again in later chapters.

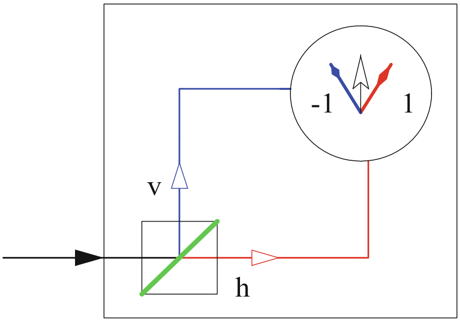

Fig. 4.2 Measurement of the linear polarization of a photon

4.4.2 Measurement and Eigenvalues

To arrive at a more compact description of the measurement, we imagine that, after the PBS, we have a detector which is connected to a display. For vertical polarization, a pointer shows '-1', for horizontal polarization '$+1$'; see Fig. 4.2. This means that after the measurement on the state $|z\rangle = a\,|h\rangle + b\,|v\rangle$, the value '$-1$' or '$+1$' is displayed with the probabilities $|b|^2$ or $|a|^2$.

We now want to describe the measured physical quantity 'horizontal/vertical polarization', encoded by ± 1. To this end, we choose a linear combination of the projection operators $|h\rangle\,\langle h|$ and $|v\rangle\,\langle v|$. The simplest non-trivial combination is clearly the polarization operator P_L

$$P_L = |h\rangle\,\langle h| - |v\rangle\,\langle v| \cong \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z, \tag{4.34}$$

where σ_z is one of the three *Pauli matrices* (more on the Pauli matrices in the exercises). We note that the Pauli matrices, and thus P_L, are not projection operators; here the P stands for 'polarization'.

The properties which are relevant for the measurement follow now by considering the *eigenvalue problem* $P_L\,|z\rangle = \mu\,|z\rangle$ (where we have now introduced the eigenvalue problem, treated in the analytical approach already in Chap. 3, into the algebraic approach considered here). As is easily verified (see exercises), P_L has the eigenvalues $\mu = +1$ and $\mu = -1$ and the eigenvectors $|z_1\rangle = |h\rangle$ and $|z_{-1}\rangle = |v\rangle$. This means that the eigenvectors describe the possible states and the eigenvalues the possible pointer positions (measurement results) after the measurement—the pointer position $+1$ tells us, for example, that after the measurement we have the state $|h\rangle$.

Similarly, we can imagine a measuring apparatus for circular polarization, in which the physical quantity 'right/left circular polarization' is encoded by ± 1. We describe this by

$$P_C = |r\rangle\,\langle r| - |l\rangle\,\langle l| \cong \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_y. \tag{4.35}$$

The eigenvalues of P_C are also ± 1, where the eigenvalue $+1$ belongs to the eigenvector $|r\rangle$ and the eigenvalue -1 to the eigenvector $|l\rangle$.

Finally, we treat a linear polarization state rotated by 45°. For the rotated state, we have (see exercises):

$$\left|h'\right\rangle = \frac{|h\rangle + |v\rangle}{\sqrt{2}} \; ; \; \left|v'\right\rangle = \frac{-|h\rangle + |v\rangle}{\sqrt{2}}. \tag{4.36}$$

We describe the corresponding measuring apparatus by the operator

$$P_{L'} = \left|h'\right\rangle\left\langle h'\right| - \left|v'\right\rangle\left\langle v'\right| \cong \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x. \tag{4.37}$$

This operator has the eigenvalues ± 1, also. (For the determination of the eigenvalues and eigenvectors of the three Pauli matrices, see the exercises.)

We learn from these three examples that the information about possible measurement results lies in the eigenvalues of certain operators. For this purpose we have constructed three examples that yield information about certain polarization states. The question of how to extend these findings and how to represent general physically-measurable variables will be treated in the following chapters.

4.4.3 Summary

We summarize with the help of the example $|z\rangle = a\,|h\rangle + b\,|v\rangle$, with $|a|^2 + |b|^2 = 1$ and $ab \neq 0$. *Before* the measurement, we can say only that: (1) The pointer will display position '+1' with probability $|a|^2 = |\langle h\,|z\rangle|^2$, and position '−1' with probability $|b|^2$; and (2) Just one of the eigenvalues of $P_L = |h\rangle\,\langle h| - |v\rangle\,\langle v|$ will be realized, with the corresponding probability. *After* the measurement, one of the two eigenvalues is realized (the pointer displays one of the two possible values), and the photon is in the corresponding state (the associated normalized eigenstate of P_L), e.g. $\frac{\langle h\,|z\rangle}{|\langle h\,|z\rangle|}\,|h\rangle$. We cannot discern which mechanism leads to this choice, but can only specify the probabilities for the possible results. The process is irreversible—the initial superposition no longer exists, and it cannot be reconstructed from the measurement results from a single photon.[16] This is possible at most by measuring an ensemble of photons in the state $a\,|h\rangle + b\,|v\rangle$ many times. From the relative frequencies of occurrence of the pointer values ± 1, we can infer the quantities $|a|^2$ and $|b|^2$.

[16]In order to make it clear once more: If, for example, we measure an arbitrarily-polarized state $|z\rangle = a\,|h\rangle + b\,|v\rangle$ with $|a|^2 + |b|^2 = 1$ and $ab \neq 0$, we find with probability $|a|^2$ a horizontal linearly-polarized photon. This does *not* permit the conclusion that the photon was in that state before the measurement. It simply makes no sense in this case to speak of a of a well-defined value of the linear polarization (+1 or −1) before the measurement.

4.5 Exercises

1. Find examples for state spaces which

   (a) have the structure of a vector space,
   (b) do not have the structure of a vector space.

2. Polarization: Determine the length of the vector $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$.
3. Given $\langle y| = i\begin{pmatrix}1 & -2\end{pmatrix}$ and $\langle z| = \begin{pmatrix}2 & i\end{pmatrix}$, determine $\langle y|\, z\rangle$.
4. The Pauli matrices are

$$\sigma_x = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix};\ \sigma_y = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix};\ \sigma_z = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \tag{4.38}$$

   In addition to $\sigma_x, \sigma_y, \sigma_z$, the notation $\sigma_1, \sigma_2, \sigma_3$ is also common.

   (a) Show that $\sigma_i^2 = 1$, $i = x, y, z$.
   (b) Determine the commutators $\left[\sigma_i, \sigma_j\right] = \sigma_i\sigma_j - \sigma_j\sigma_i$ and the anticommutators $\left\{\sigma_i, \sigma_j\right\} = \sigma_i\sigma_j + \sigma_j\sigma$ $(i \neq j)$.
   (c) Calculate the eigenvalues and eigenvectors for each Pauli matrix.

5. Determine the eigenvalues and eigenvectors of the matrix

$$M = \begin{pmatrix}1 & 4\\2 & -1\end{pmatrix}. \tag{4.39}$$

   Normalize the eigenvectors. Are they orthogonal?
6. Given the CONS $\{|a_1\rangle, |a_2\rangle\}$, determine the eigenvalues and eigenvectors of the operator

$$M = |a_1\rangle\langle a_1| - |a_2\rangle\langle a_2|\,. \tag{4.40}$$

7. Given a CONS $\{|\varphi_n\rangle\}$ and a state $|\psi\rangle = \sum_n c_n |\varphi_n\rangle$, $c_n \in \mathbb{C}$, calculate the coefficients c_n.
8. Show in bra-ket notation: The system $\{|r\rangle, |l\rangle\}$ is a CONS. Use the fact that $\{|h\rangle, |v\rangle\}$ is a CONS.
9. Given the operator $|h\rangle\langle r|$:

   (a) Is it a projection operator?
   (b) How does the operator appear in the representation (4.1)?
   (c) Given the state $|z\rangle$ with the representation $|z\rangle \cong \begin{pmatrix}z_1\\z_2\end{pmatrix}$, apply the operator $|h\rangle\langle r|$ to this state (calculation making use of the representation).

(d) Use the concrete representation to prove the equality

$$\left(|h\rangle \langle r|\ z\rangle\right)^{\dagger} = \langle z|\ r\rangle \langle h| \,. \tag{4.41}$$

10. We choose the following representation for the states $|h\rangle$ and $|v\rangle$:

$$|h\rangle \cong \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix};\ |v\rangle \cong \frac{a}{\sqrt{2}\,|a|}\begin{pmatrix} 1 \\ i \end{pmatrix}. \tag{4.42}$$

(a) Show that the representing vectors form a CONS.
(b) Determine $|r\rangle$ and $|l\rangle$ in this representation. Specialize to the cases of $a = 1$, $-1, i, -i$.

11. Show that the three vectors

$$\mathbf{a} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix};\ \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix};\ \mathbf{c} = -\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} \tag{4.43}$$

form a CONS. Do the same for

$$\mathbf{a} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix};\ \mathbf{b} = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix};\ \mathbf{c} = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}. \tag{4.44}$$

12. A three-dimensional problem: Given the CONS $\{|u\rangle, |v\rangle, |w\rangle\}$ and the operator[17]

$$L = |v\rangle \langle u| + (|u\rangle + |w\rangle) \langle v| + |v\rangle \langle w|. \tag{4.45}$$

(a) Determine the eigenvalues and eigenvectors of L.
(b) Show that the three eigenvectors form a CONS.

[17] Essentially, this operator is the x component of the orbital angular momentum operator for the angular momentum 1; see Chap. 16 Vol. 2.