

4.4: Change of Basis, and Matrix Diagonalization

From the discussion of the last section, it may look that the matrix language is fully similar to, and in many instances more convenient than the general bra-ket formalism. In particular, Eqs. (54)-(55) and (63)-(64) show that any part of any bra-ket expression may be directly mapped on the similar matrix expression, with the only slight inconvenience of using not only columns but also rows (with their elements complex-conjugated), for state vector representation. This invites the question: why do we need the bra-ket language at all? The answer is that the elements of the matrices depend on the particular choice of the basis set, very much like the Cartesian components of a usual geometric vector depend on the particular choice of reference frame orientation (Fig. 4), and very frequently, at problem solution, it is convenient to use two or more different basis sets for the same system. (Just a bit more patience numerous examples will follow soon.)

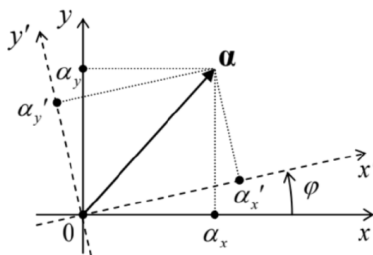


Fig. 4.4. The transformation of components of a 2D vector at a reference frame's rotation.

With this motivation, let us explore what happens at the transform from one basis, $\{u\}$, to another one, $\{v\}$ - both full and orthonormal. First of all, let us prove that for each such pair of bases, and an arbitrary numbering of the states of each base, there exists such an operator \hat{U} that, first,

Unitary operator:

$$|v_j\rangle = \hat{U} |u_j\rangle, \quad (4.4.1)$$

and, second,

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}. \quad (4.4.2)$$

(Due to the last property,¹⁶ \hat{U} is called a unitary operator, and Eq. (75), a unitary transformation.)

A very simple proof of both statements may be achieved by construction. Indeed, let us take

Unitary operator: construction

$$\hat{U} \equiv \sum_{j'} |v_{j'}\rangle \langle u_{j'}|, \quad (4.4.3)$$

- an evident generalization of Eq. (44). Then, using Eq. (38), we obtain

$$\hat{U} |u_j\rangle = \sum_{j'} |v_{j'}\rangle \langle u_{j'} | u_j\rangle = \sum_{j'} |v_{j'}\rangle \delta_{jj'} = |v_j\rangle, \quad (4.4.4)$$

so that Eq. (75) has been proved. Now, applying Eq. (31) to each term of the sum (77), we get

Conjugate unitary operator

$$\hat{U}^\dagger \equiv \sum_{j'} |u_{j'}\rangle \langle v_{j'}|, \quad (4.4.5)$$

so that

$$\hat{U}\hat{U}^\dagger = \sum_{j,j'} |v_j\rangle \langle u_j | u_{j'}\rangle \langle v_{j'}| = \sum_{j,j'} |v_j\rangle \delta_{jj'} \langle v_{j'}| = \sum_j |v_j\rangle \langle v_j| \quad (4.4.6)$$

But according to the closure relation (44), the last expression is just the identity operator, so that one of Eqs. (76) has been proved. (The proof of the second equality is absolutely similar.) As a by-product of our proof, we have also got another important expression - Eq. (79). It implies, in particular, that while, according to Eq. (75), the operator \hat{U} performs the transform from the "old" basis u_j to the "new" basis v_j , its Hermitian adjoint \hat{U}^\dagger performs the reciprocal transform:

$$\hat{U}^\dagger |v_j\rangle = \sum_{j'} |u_{j'}\rangle \delta_{jj'} = |u_j\rangle \quad (4.4.7)$$

Now let us see how do the matrix elements of the unitary transform operators look like. Generally, as was discussed above, the operator's elements may depend on the basis we calculate them in, so let us be specific - at least initially. For example, let us calculate the desired matrix elements $U_{jj'}$ in the "old" basis $\{u\}$, using Eq. (77):

$$U_{ij'}|_{\text{in } u} \equiv \langle u_j | \hat{U} | u_{j'} \rangle = \left\langle u_j \left| \left(\sum_{j''} |v_{j''}\rangle \langle u_{j''}| \right) \right| u_{j'} \right\rangle = \left\langle u_j \left| \sum_{j''} |v_{j''}\rangle \right. \right\rangle \delta_{j'j''} = \langle u_j | v_{j'} \rangle. \quad (4.4.8)$$

Now performing a similar calculation in the "new" basis $\{v\}$, we get

$$U_{ij'}|_{\text{in } v} \equiv \langle v_j | \hat{U} | v_{j'} \rangle = \left\langle v_j \left| \left(\sum_{j''} |v_{j''}\rangle \langle u_{j''}| \right) \right| v_{j'} \right\rangle = \sum_{j''} \delta_{jj''} \langle u_{j''} | v_{j'} \rangle = \langle u_j | v_{j'} \rangle. \quad (4.4.9)$$

Surprisingly, the result is the same! This is of course true for the Hermitian conjugate (79) as well:

$$U_{ij'}^\dagger|_{\text{in } u} = U_{ij'}^\dagger|_{\text{in } v} = \langle v_j | u_{j'} \rangle. \quad (4.4.10)$$

These expressions may be used, first of all, to rewrite Eq. (75) in a more direct form. Applying the first of Eqs. (41) to any state v_j of the "new" basis, and then Eq. (82), we get

$$|v_{j'}\rangle = \sum_j |u_j\rangle \langle u_j | v_{j'} \rangle = \sum_j U_{jj'} |u_j\rangle. \quad (4.4.11)$$

Similarly, the reciprocal transform is

$$|u_{j'}\rangle = \sum_j |v_j\rangle \langle v_j | u_{j'} \rangle = \sum_j U_{ij'}^\dagger |v_j\rangle. \quad (4.4.12)$$

These formulas are very convenient for applications; we will use them already in this section.

Next, we may use Eqs. (83)-(84) to express the effect of the unitary transform on the expansion coefficients α_j of the vectors of an arbitrary state α , defined by Eq. (37). As a reminder, in the "old" basis $\{u\}$ they are given by Eqs. (40). Similarly, in the "new" basis $\{v\}$,

$$\alpha_j|_{\text{in } v} = \langle v_j | \alpha \rangle. \quad (4.4.13)$$

Again inserting the identity operator in its closure form (44) with the internal index j' , and then using Eqs. (84) and (40), we get

$$\alpha_j|_{\text{in } v} = \left\langle v_j \left| \left(\sum_{j'} |u_{j'}\rangle \langle u_{j'}| \right) \right| \alpha \right\rangle = \sum_{j'} \langle v_j | u_{j'} \rangle \langle u_{j'} | \alpha \rangle = \sum_{j'} U_{ij'}^\dagger \langle u_{j'} | \alpha \rangle = \sum_{j'} U_{ij'}^\dagger \alpha_{j'} \Big|_{\text{in } u}. \quad (4.4.14)$$

The reciprocal transform is performed by matrix elements of the operator \hat{U} :

$$\alpha_j|_{\text{in } u} = \sum_{j'} U_{jj'} \alpha_{j'} \Big|_{\text{in } v}. \quad (4.4.15)$$

So, if the transform (75) from the "old" basis $\{u\}$ to the "new" basis $\{v\}$ is performed by a unitary operator, the change (88) of state vectors components at this transformation requires its Hermitian conjugate. This fact is similar to the transformation of components of a usual vector at coordinate frame rotation. For example, for a 2D vector whose actual position in space is fixed (Fig. 4):

$$\begin{pmatrix} \alpha_x' \\ \alpha_y' \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \alpha_x \\ \alpha_y \end{pmatrix}, \quad (4.4.16)$$

but the reciprocal transform is performed by a different matrix, which may be obtained from that participating in Eq. (90) by the replacement $\varphi \rightarrow -\varphi$. This replacement has a clear geometric sense: if the "new" reference frame $\{x', y'\}$ is obtained from the "old" frame $\{x, y\}$ by a counterclockwise rotation by angle φ , the reciprocal transformation requires angle $-\varphi$. (In this analogy, the unitary property (76) of the unitary transform operators corresponds to the equality of the determinants of both rotation matrices to 1.)

Due to the analogy between expressions (88) and (89) on one hand, and our old friend Eq. (62) on the other hand, it is tempting to skip indices in these new results by writing

$$|\alpha\rangle_{\text{in } v} = \hat{U}^\dagger |\alpha\rangle_{\text{in } u}, \quad |\alpha\rangle_{\text{in } u} = \hat{U} |\alpha\rangle_{\text{in } v}. \quad (\text{SYMBOLIC ONLY!}) \quad (4.4.17)$$

Since the matrix elements of \hat{U} and \hat{U}^\dagger do not depend on the basis, such language is not too bad and is mnemonically useful. However, since in the bra-ket formalism (or at least its version presented in this course), the state vectors are basis-independent, Eq. (91) has to be treated as a symbolic one, and should not be confused with the strict Eqs. (88)-(89), and with the rigorous basis-independent vector and operator equalities discussed in Sec. 2.

Now let us use the same trick of identity operator insertion, repeated twice, to find the transformation rule for matrix elements of an arbitrary operator:

$$A_{jj'}|_{\text{in } v} \equiv \langle v_j | \hat{A} | v_{j'} \rangle = \left\langle v_j \left| \left(\sum_k |u_k\rangle \langle u_k| \right) \hat{A} \left(\sum_{k'} |u_{k'}\rangle \langle u_{k'}| \right) \right| v_{j'} \right\rangle = \sum_{k,k'} U_{jk}^\dagger A_{kk'} \Big|_{\text{in } u} U_{kj'} \quad (4.4.18)$$

$$A_{jj'}|_{\text{in } u} \equiv \sum_{k,k'} U_{jk} A_{kk'} \Big|_{\text{in } v} U_{kj'}^\dagger \quad (4.4.19)$$

In the spirit of Eq. (91), we may represent these results symbolically as well, in a compact form:

$$\hat{A}|_{\text{in } v} = \hat{U}^\dagger \hat{A}|_{\text{in } u} \hat{U}, \quad \hat{A}|_{\text{in } u} = \hat{U} \hat{A}|_{\text{in } v} \hat{U}^\dagger. \quad (\text{SYMBOLIC ONLY!}) \quad (4.4.20)$$

As a sanity check, let us apply Eq. (93) to the identity operator:

$$\hat{I}|_{\text{in } v} = (\hat{U}^\dagger \hat{I} \hat{U})|_{\text{in } u} = (\hat{U}^\dagger \hat{U})|_{\text{in } u} = \hat{I}|_{\text{in } u} \quad (4.4.21)$$

- as it should be. One more (strict rather than symbolic) invariant of the basis change is the trace of any operator, defined as the sum of the diagonal terms of its matrix:

$$\text{Tr } \hat{A} \equiv \text{Tr } A \equiv \sum_j A_{jj}. \quad (4.4.22)$$

The (easy) proof of this fact, using previous relations, is left for the reader's exercise.

So far, I have implied that both state bases $\{u\}$ and $\{v\}$ are known, and the natural question is where does this information come from in quantum mechanics of actual physical systems. To get a partial answer to this question, let us return to Eq. (68), which defines the eigenstates and the eigenvalues of an operator. Let us assume that the eigenstates a_j of a certain operator \hat{A} form a full and orthonormal set, and calculate the matrix elements of the operator in the basis $\{a\}$ of these states, at their arbitrary numbering. For that, it is sufficient to inner-multiply both sides of Eq. (68), written for some index j' , by the bra-vector of an arbitrary state a_j of the same set:

$$\langle a_j | \hat{A} | a_{j'} \rangle = \langle a_j | A_{j'} | a_{j'} \rangle. \quad (4.4.23)$$

The left-hand side of this equality is the matrix element $A_{jj'}$, we are looking for, while its right-hand side is just $A_{j'} \cdot \delta_{ij}$. As a result, we see that the matrix is diagonal, with the diagonal consisting of the operator's eigenvalues:

$$A_{ij'} = A_{j'} \delta_{ij} \quad (4.4.24)$$

In particular, in the eigenstate basis (but not necessarily in an arbitrary basis!), A_{jj} means the same as A_j . Thus the important problem of finding the eigenvalues and eigenstates of an operator is equivalent to the diagonalization of its matrix,¹⁷ i.e. finding the basis in which the operator's matrix acquires the diagonal form (98); then the diagonal elements are the eigenvalues, and the basis itself is the desirable set of eigenstates.

To see how this is done in practice, let us inner-multiply Eq. (68) by a bra-vector of the basis (say, $\{u\}$) in that we have happened to know the matrix elements A_{jj} :

$$\langle u_k | \hat{A} | a_j \rangle = \langle u_k | A_j | a_j \rangle. \quad (4.4.25)$$

On the left-hand side, we can (as usual :-)) insert the identity operator between the operator \hat{A} and the ket-vector, and then use the closure relation (44) in the same basis $\{u\}$, while on the right-hand side, we can move the eigenvalue A_j (a c -number) out of the bracket, and then insert a summation over the same index as in the closure, compensating it with the proper Kronecker delta symbol:

$$\left\langle u_k \left| \hat{A} \sum_{k'} | u_{k'} \rangle \langle u_{k'} | \right. \right\rangle \langle u_{k'} | a_j \rangle = A_j \sum_{k'} \langle u_{k'} | a_j \rangle \delta_{kk'}. \quad (4.4.26)$$

Moving out the signs of summation over k' , and using the definition (47) of the matrix elements, we get

$$\sum_{k'} (A_{kk'} - A_j \delta_{kk'}) \langle u_{k'} | a_j \rangle = 0. \quad (4.4.27)$$

But the set of such equalities, for all N possible values of the index k , is just a system of linear, homogeneous equations for unknown c -numbers $\langle u_k | a_j \rangle$. According to Eqs. (82)-(84), these numbers are nothing else than the matrix elements U_{kj} of a unitary matrix providing the required transformation from the initial basis $\{u\}$ to the basis $\{a\}$ that diagonalizes the matrix A . This system may be represented in the matrix form:

$$\begin{pmatrix} A_{11} - A_j & A_{12} & \dots \\ A_{21} & A_{22} - A_j & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} U_{1j} \\ U_{2j} \\ \dots \end{pmatrix} = 0 \quad (4.4.28)$$

and the condition of its consistency,

Characteristic	$\begin{vmatrix} A_{11} - A_j & A_{12} & \dots \\ A_{21} & A_{22} - A_j & \dots \\ \dots & \dots & \dots \end{vmatrix} = 0,$
equation	
for	
eigenvalues	

plays the role of the characteristic equation of the system. This equation has N roots A_j — the eigenvalues of the operator \hat{A} ; after they have been calculated, plugging any of them back into the system (102), we can use it to find N matrix elements $U_{kj} (k = 1, 2, \dots, N)$ corresponding to this particular eigenvalue. However, since the equations (103) are homogeneous, they allow finding U_{kj} only to a constant multiplier. To ensure their normalization, i.e. enforce the unitary character of the matrix U , we may use the condition that all eigenvectors are normalized (just as the basis vectors are):

$$\langle a_j | a_j \rangle \equiv \sum_k \langle a_j | u_k \rangle \langle u_k | a_j \rangle \equiv \sum_k |U_{kj}|^2 = 1, \quad (4.4.29)$$

for each j . This normalization completes the diagonalization.¹⁸

Now (at last!) I can give the reader some examples. As a simple but very important case, let us diagonalize each of the operators described (in a certain two-function basis $\{u\}$, i.e. in two-dimensional Hilbert space) by the so-called Pauli matrices

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.4.30)$$

Though introduced by a physicist, with a specific purpose to describe electron's spin, these matrices have a general mathematical significance, because together with the 2×2 identity matrix, they provide a full, linearly-independent system - meaning that an arbitrary 2×2 matrix may be represented as

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = b\mathbf{I} + c_x\sigma_x + c_y\sigma_y + c_z\sigma_z, \quad (4.4.31)$$

with a unique set of four c -number coefficients b , c_x , c_y , and c_z .

Since the matrix σ_z is already diagonal, with the evident eigenvalues ± 1 , let us start with diagonalizing the matrix σ_x . For it, the characteristic equation (103) is evidently

$$\begin{vmatrix} -A_j & 1 \\ 1 & -A_j \end{vmatrix} = 0, \quad \text{i.e. } A_j^2 - 1 = 0, \quad (4.4.32)$$

and has two roots, $A_{1,2} = \pm 1$. (Again, the state numbering is arbitrary!) So the eigenvalues of the matrix σ_x are the same as of the matrix σ_z . (The reader may readily check that the eigenvalues of the matrix σ_y are also the same.) However, the eigenvectors of the operators corresponding to these three matrices are different. To find them for σ_x , let us plug its first eigenvalue, $A_1 = +1$, back into equations (101) spelled out for this particular case ($j = 1; k, k' = 1, 2$) :

$$\begin{aligned} -\langle u_1 | a_1 \rangle + \langle u_2 | a_1 \rangle &= 0, \\ \langle u_1 | a_1 \rangle - \langle u_2 | a_1 \rangle &= 0. \end{aligned}$$

These two equations are compatible (of course, because the used eigenvalue $A_1 = +1$ satisfies the characteristic equation), and any of them gives

$$\langle u_1 | a_1 \rangle = \langle u_2 | a_1 \rangle, \quad \text{i.e. } U_{11} = U_{21}. \quad (4.4.33)$$

With that, the normalization condition (104) yields

$$|U_{11}|^2 = |U_{21}|^2 = \frac{1}{2}. \quad (4.4.34)$$

Although the normalization is insensitive to the simultaneous multiplication of U_{11} and U_{21} by the same phase factor $\exp\{i\varphi\}$ with any real φ , it is convenient to keep the coefficients real, for example taking $\varphi = 0$, to get

$$U_{11} = U_{21} = \frac{1}{\sqrt{2}}. \quad (4.4.35)$$

Performing an absolutely similar calculation for the second characteristic value, $A_2 = -1$, we get $U_{12} = -U_{22}$, and we may choose the common phase to have

$$U_{12} = -U_{22} = \frac{1}{\sqrt{2}}, \quad (4.4.36)$$

so that the whole unitary matrix for diagonalization of the operator corresponding to σ_x is ¹⁹

$$U_x = U_x^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (4.4.37)$$

For what follows, it will be convenient to have this result expressed in the ket-relation form - see Eqs. (85)-(86):

$$|a_1\rangle = U_{11}|u_1\rangle + U_{21}|u_2\rangle = \frac{1}{\sqrt{2}}(|u_1\rangle + |u_2\rangle), \quad |a_2\rangle = U_{12}|u_1\rangle + U_{22}|u_2\rangle = \frac{1}{\sqrt{2}}(|u_1\rangle - |u_2\rangle), \quad (4.4.38)$$

$$|u_1\rangle = U_{11}^\dagger|a_1\rangle + U_{21}^\dagger|a_2\rangle = \frac{1}{\sqrt{2}}(|a_1\rangle + |a_2\rangle), \quad |u_2\rangle = U_{12}^\dagger|a_1\rangle + U_{22}^\dagger|a_2\rangle = \frac{1}{\sqrt{2}}(|a_1\rangle - |a_2\rangle). \quad (4.4.39)$$

Now let me show that these results are already sufficient to understand the Stern-Gerlach experiments described in Sec. 1 - but with two additional postulates. The first of them is that the interaction of a particle with the external magnetic field, besides that due to its orbital motion, may be described by the following vector operator of its spin dipole magnetic moment: ²⁰

$$\hat{\mathbf{m}} = \hat{\mathbf{S}}, \quad (4.4.40)$$

where the constant coefficient γ , specific for every particle type, is called the gyromagnetic ratio, ²¹ and $\hat{\mathbf{S}}$ is the vector operator of spin, with three Cartesian components:

$$\hat{\mathbf{S}} = \mathbf{n}_x\hat{S}_x + \mathbf{n}_y\hat{S}_y + \mathbf{n}_z\hat{S}_z \quad (4.4.41)$$

Here $\mathbf{n}_{x,y,z}$ are the usual Cartesian unit vectors in the 3D geometric space (in the quantum-mechanics sense, just c -numbers, or rather " c -vectors"), while $\hat{S}_{x,y,z}$ are the "usual" (scalar) operators.

For the so-called spin $-1/2$ particles (including the electron), these components may be simply, as

$$\hat{S}_{x,y,z} = \frac{\hbar}{2} \hat{\sigma}_{x,y,z} \quad (4.4.42)$$

Spin-1/2

operator expressed via those of the Pauli vector operator $\hat{\sigma} \equiv \mathbf{n}_x \hat{\sigma}_x + \mathbf{n}_y \hat{\sigma}_y + \mathbf{n}_z \hat{\sigma}_z$, so that we may also write

$$\hat{S} = \frac{\hbar}{2} \hat{\sigma} \quad (4.4.43)$$

In turn, in the so-called z -basis, each Cartesian component of the latter operator is just the corresponding Pauli matrix (105), so that it may be also convenient to use the following 3D vector of these matrices:

$$\boldsymbol{\sigma} \equiv \mathbf{n}_x \sigma_x + \mathbf{n}_y \sigma_y + \mathbf{n}_z \sigma_z \equiv \begin{pmatrix} \mathbf{n}_z & \mathbf{n}_x - i\mathbf{n}_y \\ \mathbf{n}_x + i\mathbf{n}_y & -\mathbf{n}_z \end{pmatrix} \quad (4.4.44)$$

The z -basis, in which such matrix representation of $\hat{\sigma}$ is valid, is defined as an orthonormal basis of certain two states, commonly denoted \uparrow and \downarrow , in that the matrix of the operator $\hat{\sigma}_z$ is diagonal, with eigenvalues, respectively, $+1$ and -1 , and hence the matrix $S_z \equiv (\hbar/2)\sigma_z$ of \hat{S}_z is also diagonal, with the eigenvalues $+\hbar/2$ and $-\hbar/2$. Note that we do not "understand" what exactly the states \uparrow and \downarrow are, ²² but loosely associate them with some internal rotation of a spin- $1/2$ particle about the z -axis, with either positive or negative angular momentum component S_z . However, attempts to use such classical interpretation for quantitative predictions runs into fundamental difficulties - see Sec. 6 below.

The second necessary postulate describes the general relation between the bra-ket formalism and experiment. Namely, in quantum mechanics, each real observable A is represented by a Hermitian operator $\hat{A} = \hat{A}^\dagger$, and the result of its measurement, ²³ in a quantum state α described by a linear superposition of the eigenstates a_j of the operator,

$$|\alpha\rangle = \sum_j \alpha_j |a_j\rangle, \quad \text{with } \alpha_j = \langle a_j | \alpha \rangle, \quad (4.4.45)$$

may be only one of the corresponding eigenvalues A_j . ²⁴ Specifically, if the ket (118) and all eigenkets $|a_j\rangle$ are normalized to 1,

$$\langle \alpha | \alpha \rangle = 1, \quad \langle a_j | a_j \rangle = 1, \quad (4.4.46)$$

then the probability of a certain measurement outcome A_j is ²⁵

$$W_j = |\alpha_j|^2 \equiv \alpha_j^* \alpha_j \equiv \langle \alpha | a_j \rangle \langle a_j | \alpha \rangle, \quad (4.4.47)$$

This relation is evidently a generalization of Eq. (1.22) in wave mechanics. As a sanity check, let us assume that the set of the eigenstates a_j is full, and calculate the sum of the probabilities to find the system in one of these states:

$$\sum_j W_j = \sum_j \langle \alpha | a_j \rangle \langle a_j | \alpha \rangle = \langle \alpha | \hat{I} | \alpha \rangle = 1. \quad (4.4.48)$$

Now returning to the Stern-Gerlach experiment, conceptually the description of the first (z - oriented) experiment shown in Fig. 1 is the hardest for us, because the statistical ensemble describing the unpolarized electron beam at its input is mixed ("incoherent"), and cannot be described by a pure ("coherent") superposition of the type (6) that have been the subject of our studies so far. (We will discuss such mixed ensembles in Chapter 7.) However, it is intuitively clear that its results are compatible with the description of the two output beams as sets of electrons in the pure states \uparrow and \downarrow , respectively. The absorber following that first stage (Fig. 2) just takes all spin-down electrons out of the picture, producing an output beam of polarized electrons in the definite \uparrow state. For such a beam, the probabilities (120) are $W_\uparrow = 1$ and $W_\downarrow = 0$. This is certainly compatible with the result of the "control" experiment shown on the bottom panel of Fig. 2: the repeated SG (z) stage does not split such a beam, keeping the probabilities the same.

Now let us discuss the double Stern-Gerlach experiment shown on the top panel of Fig. 2. For that, let us represent the z -polarized beam in another basis - of the two states (I will denote them as \rightarrow and \leftarrow) in that, by definition, the matrix S_x is diagonal. But this is exactly the set we called $a_{1,2}$ in the σ_x matrix diagonalization problem solved above. On the other hand, the states \uparrow and \downarrow are

exactly what we called $u_{1,2}$ in that problem, because in this basis, we know matrix σ explicitly - see Eq. (117). Hence, in the application to the electron spin problem, we may rewrite Eqs. (114) as

$$\begin{aligned} |\rightarrow\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle), & |\leftarrow\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle), \\ |\uparrow\rangle &= \frac{1}{\sqrt{2}}(|\rightarrow\rangle + |\leftarrow\rangle), & |\downarrow\rangle &= \frac{1}{\sqrt{2}}(|\rightarrow\rangle - |\leftarrow\rangle), \end{aligned}$$

Currently for us the first of Eqs. (123) is most important, because it shows that the quantum state of electrons entering the SG (x) stage may be represented as a coherent superposition of electrons with $S_x = +\hbar/2$ and $S_x = -\hbar/2$. Notice that the beams have equal probability amplitude moduli, so that according to Eq. (122), the split beams \rightarrow and \leftarrow have equal intensities, in accordance with experimental results. (The minus sign before the second ket-vector is of no consequence here, but it may have an impact on outcomes of other experiments - for example, if coherently split beams are brought together again.)

Now, let us discuss the most mysterious (from the classical point of view) multi-stage SG experiment shown on the middle panel of Fig. 2. After the second absorber has taken out all electrons in, say, the \leftarrow state, the remaining electrons, all in the state \rightarrow , are passed to the final, SG (z), stage. But according to the first of Eqs. (122), this state may be represented as a (coherent) linear superposition of the \uparrow and \downarrow states, with equal probability amplitudes. The final stage separates electrons in these two states into separate beams, with equal probabilities $W_{\uparrow} = W_{\downarrow} = 1/2$ to find an electron in each of them, thus explaining the experimental results.

To conclude our discussion of the multistage Stern-Gerlach experiment, let me note that though it cannot be explained in terms of wave mechanics (which operates with scalar de Broglie waves), it has an analogy in classical theories of vector fields, such as the classical electrodynamics. Indeed, let a plane electromagnetic wave propagate normally to the plane of the drawing in Fig. 5, and pass through the linear polarizer 1.

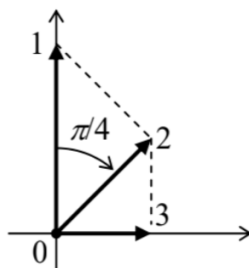


Fig. 4.5. A light polarization sequence similar to the three-stage Stern-Gerlach experiment shown on the middle panel of Fig. 2.

Similarly to the output of the initial SG (z) stages (including the absorbers) shown in Fig. 2, the output wave is linearly polarized in one direction - the vertical direction in Fig. 5. Now its electric field vector has no horizontal component - as may be revealed by the wave's full absorption in a perpendicular polarizer 3. However, let us pass the wave through polarizer 2 first. In this case, the output wave does acquire a horizontal component, as can be, again, revealed by passing it through polarizer 3. If the angles between the polarization directions 1 and 2, and between 2 and 3, are both equal to $\pi/4$, each polarizer reduces the wave amplitude by a factor of $\sqrt{2}$, and hence the intensity by a factor of 2, exactly like in the multistage SG experiment, with the polarizer 2 playing the role of the SG (x) stage. The "only" difference is that the necessary angle is $\pi/4$, rather than by $\pi/2$ for the SternGerlach experiment. In quantum electrodynamics (see Chapter 9 below), which confirms classical predictions for this experiment, this difference may be interpreted by that between the integer spin of electromagnetic field quanta (photons) and the half-integer spin of electrons.

¹⁶ An alternative way to express Eq. (76) is to write $\hat{U}^\dagger = \hat{U}^{-1}$, but I will try to avoid this language.

¹⁷ Note that the expression "matrix diagonalization" is a very common but dangerous jargon. (Formally, a matrix is just a matrix, an ordered set of c -numbers, and cannot be "diagonalized".) It is OK to use this jargon if you remember clearly what it actually means - see the definition above.

¹⁸ A possible slight complication here is that the characteristic equation may give equal eigenvalues for certain groups of different eigenvectors. In such cases, the requirement of the mutual orthogonality of these degenerate states should be additionally enforced.

¹⁹ Note that though this particular unitary matrix is Hermitian, this is not true for an arbitrary choice of phases φ .

²⁰ This was the key point in the electron spin's description, developed by W. Pauli in 1925-1927.

²¹ For the electron, with its negative charge $q = -e$, the gyromagnetic ratio is negative: $\gamma_e = -g_e e / 2m_e$, where $g_e \approx 2$ is the dimensionless g -factor. Due to quantum-electrodynamic (relativistic) effects, this g -factor is slightly higher than 2 : $g_e = 2(1 + \alpha/2\pi + \dots) \approx 2.002319304\dots$, where $\alpha \equiv e^2/4\pi\epsilon_0\hbar c \equiv (E_H/m_e c^2)^{1/2} \approx 1/137$ is the so-called fine structure constant. (The origin of its name will be clear from the discussion in Sec. 6.3.)

²² If you think about it, the word "understand" typically means that we can express a new, more complex notion in terms of those discussed earlier and considered "known". In our current case, we cannot describe the spin states by some wavefunction $\psi(\mathbf{r})$, or any other mathematical notion discussed in the previous three chapters. The bra-ket formalism has been invented exactly to enable mathematical analyses of such "new" quantum states we do not initially "understand". Gradually we get accustomed to these notions, and eventually, as we know more and more about their properties, start treating them as "known" ones.

²³ Here again, just like in Sec. 1.2, the statement implies the abstract notion of "ideal experiments", deferring the discussion of real (physical) measurements until Chapter 10.

²⁴ As a reminder, at the end of Sec. 3 we have already proved that such eigenstates corresponding to different values A_j are orthogonal. If any of these values is degenerate, i.e. corresponds to several different eigenstates, they should be also selected orthogonal, in order for Eq. (118) to be valid.

²⁵ This key relation, in particular, explains the most common term for the (generally, complex) coefficients α_j , the probability amplitudes.

This page titled [4.4: Change of Basis, and Matrix Diagonalization](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Konstantin K. Likharev](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.