On common eigenbases of commuting operators

Paolo Glorioso

In this note we try to answer the question: "*Given two commuting Hermitian operators $\hat{A}$ and $\hat{B}$, is each eigenbasis of $\hat{A}$ also an eigenbasis of $\hat{B}$?*" We take this occasion to review the mathematical results needed to explore the answer to such question. Moreover, we will assume that the reader is familiar with the concepts of vector space, vector subspace, linear combination, linear independence, diagonalization, inner product, and basis. These concepts can be found in Sections 1.1, 1.2 and 1.4 in [1]. A less specific treatment of the following is given in Section 1.8 therein.
Consider an operator $\hat{A}$, acting on vectors belonging to a vector space $\mathbb{V}$. We will make use of the following definitions:

i) **Eigenvalue:** A constant $\lambda \in \mathbb{C}$ is called an eigenvalue of $\hat{A}$ if it satisfies the following equation:

$$\hat{A}v = \lambda v, \tag{1}$$

for some *nonvanishing* vector $v \in \mathbb{V}$.

ii) **Eigenvector:** A nonvanishing vector $v \in \mathbb{V}$ is an eigenvector of $\hat{A}$ if it satisfies Equation (1) for some $\lambda \in \mathbb{C}$. Note that v is also called eigenstate, or eigenfunction, depending on the context.

iii) **Nondegenerate eigenvalue/eigenvector:** An eigenvalue λ of $\hat{A}$ is called nondegenerate if Equation (1) is satisfied by only one vector v, up to an overall complex number, i.e. all the solutions of Equation (1) are of the form αv, where $\alpha \in \mathbb{C}$. This is equivalent to say that the eigenvalue λ corresponds to only one eigenstate of $\hat{A}$. Similarly, an eigenvector is called nondegenerate if it is the only vector, up to an overall complex number, that satisfies Equation (1) for some λ. Degenerate eigenvalues/eigenvectors are those which don't satisfy this uniqueness property. Note that some references use the expression "degenerate state" with respect to Equation (1) but referring only to the energy operator, see for example [2].

$$\hat{E}v = \varepsilon v,$$

see below for more details.

iv) **Eigenbasis:** A set of vectors

$$\mathcal{V} = \{v_i\},$$

such that each v_i is an eigenvector of $\hat{A}$ and $\mathcal{V}$ is a basis of $\mathbb{V}$, is called an eigenbasis of $\mathbb{V}$ with respect to $\hat{A}$.

In this note we will refer to Hermitian operators, where $\hat{A}$ is Hermitian if, for any $u, v \in \mathbb{V}$,

$$(u, \hat{A}v) = (\hat{A}u, v),$$

and (u, v) is the scalar product in $\mathbb{V}$. There are two reasons why we consider Hermitian operators. First, because in Quantum Mechanics all observables are postulated to be Hermitian.[1] Second, because Hermitian operators are *diagonalizable*, i.e. they admit a basis in which they have a diagonal form, which is then an eigenbasis. See Theorem 10 in Chapter 1 of [1] for this point.

Proposition 1. *Let $\hat{A}$ be a Hermitian operator with only nondegenerate eigenvalues, and $\mathcal{V} = \{v_i\}$ and $\mathcal{W} = \{w_i\}$ two eigenbases of $\hat{A}$. Then $\mathcal{V}$ is obtained from $\mathcal{W}$ by permutations and multiplications by complex numbers of the eigenvectors of $\mathcal{W}$, i.e., for each $v_i \in \mathcal{V}$, there is $w_j \in \mathcal{W}$ and $\alpha \in \mathbb{C}$ such that*

$$v_i = \alpha w_j.$$

In other words, $\mathcal{V}$ and $\mathcal{W}$ contain the same eigenstates.

Proof. Let λ_i and μ_i be the eigenvalues of v_i and w_i, respectively, i.e.

$$\hat{A}v_i = \lambda_i v_i, \qquad \hat{A}w_i = \mu_i w_i. \tag{2}$$

Since $\mathcal{W}$ is a basis, we can write any $v_i \in \mathcal{V}$ as a linear combination of the w_i's,

$$v_i = \sum_j \alpha_j w_j, \tag{3}$$

where $\alpha_j \in \mathbb{C}$. Then,

$$\lambda_i v_i = \hat{A}v_i = \hat{A}\sum_j \alpha_j w_j = \sum_j \alpha_j \hat{A}w_j = \sum_j \alpha_j \mu_j w_j.$$

where we used the linearity of $\hat{A}$. Comparing the first and the last members of the above Equation with Equation (3), we get

$$\sum_j \lambda_i \alpha_j w_j = \sum_j \mu_j \alpha_j w_j \quad \Longrightarrow \quad \sum_j (\mu_j - \lambda_i)\, \alpha_j w_j = 0,$$

and using the fact that the w_j's are linearly independent, we obtain,

$$(\mu_j - \lambda_i)\, \alpha_j = 0.$$

This is a set of equations, labelled by j. Each of them has two solutions: either $\alpha_j = 0$, or $\mu_j = \lambda_i$. Since by hypothesis, the μ_j's are all different from each other, being nondegenerate eigenvalues, only one of the above set of equations can satisfy $\mu_j = \lambda_i$. Thus all α_j's but one are zero, and we obtain, from Equation (3),

$$v_i = \alpha_j w_j,$$

[1] Recall that this is implied by the requirement of having *real* eigenvalues.

for one value of j. This tells us that each vector of $\mathcal{V}$ is also contained in $\mathcal{W}$, up to an overall complex number, which in the above Equation is given by α_j. Of course, we can also revert the whole reasoning and show that each vector of $\mathcal{W}$ is contained in $\mathcal{V}$. Therefore we can say that the two eigenbases are the same, up to permutations and multiplications by complex numbers of their vectors, and we are done. □

Proposition 2. *Suppose that $\hat{A}$ and $\hat{B}$ are Hermitian operators with vanishing commutator, i.e.*

$$[\hat{A}, \hat{B}] = 0.$$

Then $\hat{A}$ and $\hat{B}$ share a common eigenbasis.

Proof. Consider an eigenbasis of $\hat{A}$, $\mathcal{V} = \{v_i\}$, with λ_i the eigenvalue associated to v_i. Then, for any v_i,

$$\hat{A}\hat{B}v_i = \hat{B}\hat{A}v_i = \lambda_i \hat{B}v_i, \tag{4}$$

i.e., if $\hat{B}v_i \neq 0$, $\hat{B}v_i$ is an eigenvector of $\hat{A}$ associated to the same eigenvalue as v_i, λ_i. We have two cases to consider:

- λ_i **nondegenerate:** $\hat{B}v_i$ can differ from v_i only by a constant factor, or

  $$\hat{B}v_i = \mu_i v_i,$$

  and thus v_i is also an eigenvector of $\hat{B}$, with eigenvalue μ_i.

- λ_i **degenerate:** in this case there are more vectors associated to λ_i, which we denote by w_j, $j = 1, \ldots, N$, where N is the degeneracy of λ_i. Since $\hat{B}w_j$ is still an eigenvector of $\hat{A}$, we can write it as a linear combination of the w_j's. For this reason, the operator $\hat{B}$ can be seen as acting "internally" in the subspace spanned by the w_j's. Since $\hat{B}$ is Hermitian, it is Hermitian in particular in this subspace. Indeed, for any u_1, u_2 belonging to such subspace,

  $$(u_1, \hat{B}u_2) = (\hat{B}u_1, u_2),$$

  just because u_1, u_2 belong also to the "big" vector space $\mathbb{V}$.
  Now, since $\hat{B}$ is Hermitian in this subspace we can diagonalize it, or in other words we can choose a basis of eigenvectors of $\hat{B}$ which span this subspace, and we call them w'_j. These w'_j are still eigenvectors of $\hat{A}$, and thus the Proposition is proved.

□

To understand better the last proof, we can view $\hat{B}$ as an "infinite matrix", in the sense explained by the following picture:

$$\hat{B} = \begin{pmatrix} (v_1, \hat{B}v_1) & (v_1, \hat{B}v_2) & \cdots \\ (v_2, \hat{B}v_1) & (v_2, \hat{B}v_2) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

where the v_i's belong to the basis with respect to which $\hat{B}$ is represented. If we view $\hat{B}$ in the eigenbasis $\mathcal{V}$ introduced at the beginning of the proof, we have

$$\hat{B} = \begin{pmatrix} \boxed{B_1} & & & \\ & \boxed{B_2} & & \\ & & \boxed{B_3} & \\ & & & \ddots \end{pmatrix}, \tag{5}$$

where each block B_i is a submatrix representing the internal action of $\hat{B}$ on the subspaces similar to the one spanned by the w_j's. Each block refers to an eigenvalue λ_i of $\hat{A}$, and if λ_i is nondegenerate the block will be just a 1×1 matrix. If λ_i is degenerate with degeneracy N then the block will be $N \times N$. What we did in the degenerate case of the proof was just to show that the corresponding block B_i is a Hermitian matrix, and thus diagonalizable.
Finally, note that if we know that $\hat{A}$ and $\hat{B}$ share a common eigenbasis, then their commutator is zero. Indeed, sharing a common eigenbasis means that in such basis they are *both* represented as diagonal operators, and thus they commute. This consideration allows us to state a more powerful statement than the above Preposition:
Proposition 3. *Let $\hat{A}$ and $\hat{B}$ be two Hermitian operators. Then the following two statements are equivalent:*

i) $\hat{A}$ and $\hat{B}$ possess a common eigenbasis.

ii) $\hat{A}$ and $\hat{B}$ commute.

Aimed of the mathematical results we have found, we shall now answer the following question:

Given two commuting Hermitian operators $\hat{A}$ and $\hat{B}$, is each eigenbasis of $\hat{A}$ also an eigenbasis of $\hat{B}$?

The short answer is: it depends. Consider the case where both $\hat{A}$ and $\hat{B}$ have only nondegenerate eigenvalues. Then, by virtue of Proposition 1 and 2, each eigenbasis of $\hat{A}$ is also an eigenbasis of $\hat{B}$. Indeed, by Proposition 2 we can consider a common eigenbasis of $\hat{A}$ and $\hat{B}$, which we denote by $\mathcal{V}$. By Proposition 1 we know that we would exhaust all the eigenbases of $\hat{A}$ by permuting and multiplying by complex numbers the vectors of $\mathcal{V}$, and the same for the eigenbases of $\hat{B}$. Thus, in this case, the answer to the above question is YES. In all the other cases, the answer is NO.
Consider e.g. the case where $\hat{A}$ has some degenerate eigenvalues. Then, in some eigenbasis of $\hat{A}$, $\hat{B}$ would look like in Equation (5), which is *not* in diagonal form if some of the blocks B_i are nondiagonal. For a neat example, we can consider the following matrices:

$$A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 2 \end{pmatrix}, \qquad B = \left(\begin{array}{c|cc} 3 & & \\ \hline & 1 & 2i \\ & -2i & 1 \end{array}\right),$$

note that $[A, B] = 0$, essentially because the 2×2 lower diagonal block of A is a scalar matrix,[2] and thus it must commute with the corresponding lower diagonal block of B. Moreover, B is Hermitian,

[2]A scalar matrix is a matrix proportional to the identity.

and then diagonalizable. Note that A and B are represented in terms of an eigenbasis of A, and that 2 is a degenerate eigenvalue of A. We denote this eigenbasis by $\mathcal{V} = \{e_1, e_2, e_3\}$, where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Following the proof of Proposition 2, we only need to find two linear combinations of e_2, e_3 such that the lower block of B assumes a diagonal form. Working out the standard diagonalization procedure, we find the common eigenbasis $\mathcal{W} = \{e_1, v_2, v_3\}$, where

$$v_2 = \frac{1}{\sqrt{2}}(e_2 + ie_3), \qquad v_3 = \frac{1}{\sqrt{2}}(e_2 - ie_3),$$

and the above matrices assume the following form:

$$A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & & \\ & -1 & \\ & & 3 \end{pmatrix}.$$

For a more physical example, consider the hydrogen atom. The energy eigenstates are commonly labelled by

$$|n, \ell, m\rangle, \tag{6}$$

where n is the energy level, ℓ is the label associated to $\hat{L}^2$, and m is the label associated to $\hat{L}_z$ through

$$\hat{L}_z \psi = \hbar m \psi.$$

In this basis, $\hat{E}, \hat{L}^2$ and $\hat{L}_z$ are all diagonal. Using the fact that the states (6) are degenerate, we want to construct a new basis where $\hat{L}_z$ is nondiagonal. One example is

$$\mathcal{V} = \{|n, \ell, s\rangle\}, \qquad \text{where} \qquad |n, \ell, s\rangle = \frac{1}{\sqrt{\ell + s + 1}} \sum_{m=-\ell}^{s} |n, \ell, m\rangle,$$

and where $s = -\ell, \ldots, \ell$. Clearly, $\mathcal{V}$ is a complete set of eigenstates of $\hat{E}$ and $\hat{L}^2$, but they are not eigenstates of $\hat{L}_z$, indeed

$$\hat{L}_z|n, \ell, s\rangle = \frac{1}{\sqrt{\ell + s + 1}} \sum_{m=-\ell}^{s} \hat{L}_z|n, \ell, m\rangle = \frac{1}{\sqrt{\ell + s + 1}} \sum_{m=-\ell}^{s} \hbar m|n, \ell, m\rangle,$$

which clearly doesn't correspond to Equation (1). Another simple construction we can make with the hydrogen atom is the following. Consider, instead of (6), the basis given by the eigenvectors of $\hat{E}, \hat{L}^2$, and $\hat{L}_x$, instead of $\hat{L}_z$, and denote it by

$$\mathcal{W} = \{|n, \ell, m_x\rangle\}. \tag{7}$$

Thus we have that

$$\hat{L}_x|n, \ell, m_x\rangle = \hbar m_x|n, \ell, m_x\rangle.$$

However, $\hat{L}_z$ will not be diagonal in this basis, indeed suppose that, for all m_x, there exist α_{m_x} such that

$$\hat{L}_z|n,\ell,m_x\rangle = \alpha_{m_x}|n,\ell,m_x\rangle.$$

Then, on any eigenstate of the form (7),

$$i\hat{L}_y|n,\ell,m_x\rangle = [\hat{L}_z,\hat{L}_x]|n,\ell,m_x\rangle = [\hbar m,\alpha_{m_x}]|n,\ell,m_x\rangle = 0,$$

which we know to be false. Therefore $\mathcal{W}$ is not an eigenbasis for $\hat{L}_z$, but it keeps being an eigenbasis for $\hat{E}$ and $\hat{L}^2$. In the basis given by the states (6), the operator $\hat{L}_z$ assumes the following matrix form for the first 4 eigenstates:

$$\hat{L}_z = \hbar \begin{pmatrix} 0 & & & & \\ & \boxed{\begin{matrix} -1 & & \\ & 0 & \\ & & 1 \end{matrix}} & \\ & & \ddots \end{pmatrix},$$

whereas, in the basis (7),

$$\hat{L}_z = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & & \\ & \boxed{\begin{matrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{matrix}} & \\ & & \ddots \end{pmatrix}.$$

References

[1] R. Shankar, *Principles of Quantum Mechanics* (Kluwer, 1994).

[2] D.J. Griffiths, *Introduction to Quantum Mechanics* (Pearson Prentice Hall, 2005).