15.7 Simultaneous Diagonalization of Hermitian Operators

We say that two operators S and T in a vector space V can be **simultaneously diagonalized** if there is some basis of V in which both the matrix representation of S and the matrix representation of T are diagonal. It then follows that each vector in this basis is an eigenvector of S *and* an eigenvector of T. There is therefore an entire basis of *shared* eigenvectors.

A necessary condition for simultaneous diagonalization is that the operators S and T commute. Indeed, if they can be simultaneously diagonalized, there is a basis where both are diagonal, and they manifestly commute. If the operators don't commute, this is a basis-independent statement, and therefore a simultaneous diagonal presentation cannot exist. Since arbitrary linear operators S and T on a complex vector space cannot be diagonalized, the vanishing of $[S, T]$ does not guarantee simultaneous diagonalization. But if the operators are Hermitian, it does, as we show now.

Theorem 15.7.1. *If S and T are commuting Hermitian operators, they can be simultaneously diagonalized.*

Proof. The main complication is that degeneracies in the spectrum require some discussion. Either both operators have degeneracies or at least one has no degeneracies. Without loss of generality, we can assume there are two cases to consider:

1. There is no degeneracy in the spectrum of T, or
2. both T and S have degeneracies in their spectrum.

Case (1). Since T is nondegenerate, there is a basis $(u_1, \ldots, u_n)$ of eigenvectors of T with different eigenvalues:

$$Tu_i = \lambda_i u_i, \quad i \quad \text{not summed}, \quad \lambda_i \neq \lambda_j \quad \text{for} \quad i \neq j. \tag{15.7.1}$$

We now want to understand what kind of vector Su_i is. For this we act with T on it:

$$T(Su_i) = S(Tu_i) = S(\lambda_i u_i) = \lambda_i (S\, u_i). \tag{15.7.2}$$

It follows that Su_i is also an eigenvector of T with eigenvalue λ_i, thus it must equal u_i, up to scale:

$$Su_i = \omega_i u_i, \tag{15.7.3}$$

showing that u_i is also an eigenvector of S, this time with eigenvalue ω_i. Thus, any eigenvector of T is also an eigenvector of S, showing that these operators are simultaneously diagonalizable.

Case (2). Since T has degeneracies, as explained in the previous section, we have a decomposition of V in T-invariant subspaces U_k spanned by eigenvectors:

$$\begin{gathered}
U_k \equiv \{u \,|\, Tu = \lambda_k u\}, \quad \dim U_k = d_k, \qquad V = U_1 \oplus \cdots U_m, \\
\text{orthonormal basis for } V: \quad (u_1^{(1)}, \ldots, u_{d_1}^{(1)}, \ \ldots \ , u_1^{(m)}, \ldots, u_{d_m}^{(m)}). \\
T = \text{diag}\,\big(\underbrace{\lambda_1, \ldots, \lambda_1}_{d_1 \text{ times}}, \ \ldots \ , \underbrace{\lambda_m, \ldots, \lambda_m}_{d_m \text{ times}} \big) \quad \text{in this basis.}
\end{gathered} \tag{15.7.4}$$

We also explained that the alternative orthonormal basis of V given by

$$\big(V_1 u_1^{(1)}, \ldots, V_1 u_{d_1}^{(1)}, \ \ldots \ , V_m u_1^{(m)}, \ \ldots \ , V_m u_{d_m}^{(m)}\big) \tag{15.7.5}$$

leads to the same matrix for T when each V_k is a unitary operator on U_k.

We now claim that the U_k are also S-invariant subspaces! To show this, let $u_k \in U_k$, and examine the vector Su_k. We have

$$T(Su_k) = S(Tu_k) = \lambda_k Su_k \quad \Rightarrow \quad Su_k \in U_k. \tag{15.7.6}$$

It follows that in the basis (15.7.4) the matrix for S takes *block-diagonal* form, with blocks on each of the U_k subspaces. We cannot guarantee, however, that S is diagonal within each square block; we only know that $Su_i^{(k)} \in U_k$.

Since S, restricted to each S-invariant subspace U_k, is Hermitian, we can find an orthonormal basis of U_k in which the matrix S is diagonal. This new basis is unitarily related to the original basis $(u_1^{(k)}, \ldots, u_{d_k}^{(k)})$ and thus takes the form $(V_k u_1^{(k)}, \ldots, V_k u_{d_k}^{(k)})$ with V_k a unitary operator in U_k. Note that the eigenvalues of S in this block need not be degenerate. Doing this for each block, we find a basis of the form (15.7.5) in which S is diagonal. But T is still diagonal and unchanged in this new basis, so both S and T have been simultaneously diagonalized.

□

Remarks:

1. The above proof gives an algorithmic way to produce the common list of eigenvectors. If one of the operators is nondegenerate, we are in case (1), and its eigenvectors will do the job even if the other operator is degenerate. If both operators are degenerate, we are in case (2). One diagonalizes one of the operators and constructs the second matrix in the basis of eigenvectors of the first. This second matrix is block diagonal, where the blocks are organized by the degeneracies in the spectrum of the first matrix. One must then diagonalize within the blocks of the second matrix to select new eigenvectors in each degenerate subspace. It is guaranteed that the new basis that works for the second matrix also works for the first.
2. If we had to simultaneously diagonalize three different commuting Hermitian operators S_1, S_2, and S_3, all of which have degenerate spectra, we would proceed as follows. We would diagonalize S_1 and fix a basis in which S_1 is diagonal. In this basis we must find that S_2 and S_3 have exactly the same block structure. The corresponding block matrices are simply the matrix representations of S_2 and S_3 in each of the invariant spaces U_k appearing in the diagonalization of S_1. Since S_2 and S_3 commute, their restrictions to U_k commute. These restrictions can be diagonalized simultaneously, as guaranteed by our theorem that works for two matrices. The new basis in U_k that makes the restriction of both S_2 and S_3 diagonal will not disturb the diagonal form of S_1 in this block. This would be repeated for each block, until we got a common basis of eigenvectors.
3. An inductive argument is now apparent. If we know how to simultaneously diagonalize n commuting Hermitian operators, we can diagonalize $n + 1$ of them, $S_1, \ldots, S_{n+1}$, as follows. We diagonalize S_1 and then consider the remaining n operators in the basis that makes S_1 diagonal. We are guaranteed a common block structure for the n operators. The problem becomes one of simultaneous diagonalization of n commuting Hermitian block matrices, which is possible by the induction assumption. We have thus proven the corollary below.

Corollary. If $\{S_1, \ldots, S_n\}$ is a set of mutually commuting Hermitian operators, they can all be simultaneously diagonalized.

Example 15.4. *Simultaneous diagonalization of two matrices.*
Consider two Hermitian matrices A_1 and A_2 that commute:

$$A_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \frac{5}{4} & \frac{1}{2\sqrt{2}} & -\frac{1}{4} \\ \frac{1}{2\sqrt{2}} & \frac{3}{2} & -\frac{1}{2\sqrt{2}} \\ -\frac{1}{4} & -\frac{1}{2\sqrt{2}} & \frac{5}{4} \end{pmatrix}. \tag{15.7.7}$$

We wish to find simultaneous eigenvectors for the two matrices. This takes some effort because both matrices have a degenerate spectrum: the A_1 eigenvalues are (2, 0, 0), and the A_2 eigenvalues are (2, 1, 1). So we are not in the simple situation in which one of the matrices is nondegenerate, and its eigenvectors automatically work for the other matrix. We must then follow the route summarized in remark (1) above. We pick the matrix A_1, which looks simpler, and we quickly find its eigenvalues and orthonormal eigenvectors:

$$\lambda_1 = 2,\ u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_2 = 0,\ u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_3 = 0,\ u_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{15.7.8}$$

Since $\lambda_1 = 2$ is nondegenerate, its eigenvector u_1 must be an eigenvector of A_2. One quickly finds that $A_2 u_1 = u_1$, so it has eigenvalue one. We now build the matrix for the operator A_2 in the basis of A_1 eigenvectors u_1, u_2, u_3. Because u_1 is an A_2 eigenvector of eigenvalue one, the first column in the matrix, which we call A_2', is (1, 0, 0). Thus, we have

$$A_2' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a'_{22} & a'_{23} \\ 0 & a'_{32} & a'_{33} \end{pmatrix}, \tag{15.7.9}$$

since the matrix must be Hermitian. Now we need to determine the rest of the entries. A short calculation gives

$$A_2u_2 = \frac{3}{2}u_2 - \frac{1}{2}u_3,$$
$$A_2u_3 = -\frac{1}{2}u_2 + \frac{3}{2}u_3. \tag{15.7.10}$$

The left-hand sides have the matrix A_2 in (15.7.7) acting on the second and third eigenvectors in (15.7.8). The results are easily written as superpositions of u_2 and u_3 by inspection. More systematically, you could write the u_i's in terms of the original basis vectors e_1, e_2, e_3 used to describe A_1 and A_2 and then solve for the e_i in terms of the u_i, which allows you to write any column vector in terms of u_i's. At any rate the above relations imply that the full matrix A_2' is

$$A_2' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}. \tag{15.7.11}$$

As a good check on our arithmetic, we confirm the invariance of the trace: $\mathrm{tr}A_2 = \mathrm{tr}A_2' = 4$. We now must diagonalize the 2×2 block that, as expected, is associated with the degenerate subspace of A_1 spanned by the eigenvectors u_2 and u_3. Calling the eigenvalues t_2, t_3 and the associated eigenvectors w_2, w_3, we have

$$\begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix}, \quad t_2 = 2, \; w_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad t_3 = 1, \; w_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{15.7.12}$$

Since this matrix is relative to the $\{u_2, u_3\}$ basis, the eigenvectors are in fact

$$w_2 = \frac{1}{\sqrt{2}}(-u_2 + u_3), \quad w_3 = \frac{1}{\sqrt{2}}(u_2 + u_3). \tag{15.7.13}$$

These are the missing common eigenvectors of the two matrices. All in all, the three simultaneous eigenvectors of A_1 and A_2 are u_1, w_2, w_3:

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \; (2, 1); \quad w_2 = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \; (0, 2); \quad w_3 = \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \; (0, 1). \tag{15.7.14}$$

Following each eigenvector, $(\#_1, \#_2)$ denote the A_1 and A_2 eigenvalues, respectively.