15.5 Diagonalization of Operators

When we have operators we wish to understand, it can be useful to find a basis on the vector space for which the operators are represented by matrices that take a simple form. Diagonal matrices are those where all nondiagonal entries vanish. If we can find a set of basis vectors for which the matrix representing an operator is diagonal, the operator is said to be **diagonalizable**.

If an operator T is diagonal in some basis $(u_1, \ldots, u_n)$ of the vector space V, its matrix takes the form diag $(\lambda_1, \ldots, \lambda_n)$, with constants λ_i, and we have

$$Tu_1 = \lambda_1 u_1, \ldots, \quad Tu_n = \lambda_n u_n. \tag{15.5.1}$$

The basis vectors are thus eigenvectors with eigenvalues given by the diagonal elements. It follows that *a matrix is diagonalizable if and only if it possesses a set of eigenvectors that span the vector space*.

Recall that all operators T on finite-dimensional complex vector spaces have at least one eigenvalue and thus at least one eigenvector. But even in complex vector spaces, not all operators have enough eigenvectors to span the space. Those operators cannot be diagonalized. The simplest example of such an operator is provided by the 2×2 matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{15.5.2}$$

The only eigenvalue of this matrix is $\lambda = 0$, and the associated eigenvector is $(1, 0)$, or any multiple thereof. One basis vector cannot span a two-dimensional vector space. As a result, this matrix cannot be diagonalized.

Suppose we have a vector space V, and we have chosen a basis $(v_1, \ldots, v_n)$ such that a linear operator has a matrix representation $T_{ij}(\{v\})$ that is not diagonal. We can change the basis to a new one $(u_1, \ldots, u_n)$ using a linear operator A that acts as follows:

$$u_k = A\, v_k. \tag{15.5.3}$$

As we learned in section 13.5, the matrix representation $T_{ij}(\{u\})$ of the operator in the new basis takes the form

$$T(\{u\}) = A^{-1} T(\{v\}) A \quad \text{or} \quad T_{ij}(\{u\}) = (A^{-1})_{ik}\, T_{kp}(\{v\})\, A_{pj}, \tag{15.5.4}$$

where the matrix A_{ij} is the representation of A in either the original v-basis or the new u-basis. T is diagonalizable if there is an operator A such that $T_{ij}(\{u\})$ is diagonal. The matrices $T(\{u\})$ and $T(\{v\})$ are related by similarity.

There are two equivalent ways of thinking about the diagonalization of T:

1. The matrix representation of T is diagonal when using the u-basis obtained by acting with A on the original v-basis. Thus, the u_i are the eigenvectors of T.

2. The operator $A^{-1}TA$ is diagonal in the *original* v-basis.

The second viewpoint requires justification. Since T is diagonal in the u-basis, $Tu_i = \lambda_i u_i$ (i is not summed). This implies that $TA\, v_i = \lambda_i A v_i$. Acting with A^{-1} we find $(A^{-1}TA)\, v_i = \lambda_i v_i$, which confirms that $A^{-1}TA$ is represented by a diagonal matrix in the original v-basis. Both viewpoints are valuable.

Using the second viewpoint, we write the following matrix equation in the *original basis*:

$$\boxed{A^{-1}TA = D_T.} \tag{15.5.5}$$

Here D_T is a diagonal matrix, and we say that A is a matrix that diagonalizes T by similarity. It is useful to note the following:

$$\boxed{\text{The columns of the matrix } A \text{ are the eigenvectors of } T.} \tag{15.5.6}$$

We see this as follows. Recall that the eigenvectors of T are the u_k and therefore,

$$u_k = Av_k = \sum_i A_{ik} v_i = \begin{pmatrix} A_{1k} \\ \vdots \\ A_{nk} \end{pmatrix}. \tag{15.5.7}$$

In the last step, we noted that the basis vector v_i is represented by a column vector of zeroes with a single unit entry at the ith position. This confirms that the kth column of A is the kth eigenvector of T.

While not all operators on complex vector spaces can be diagonalized, the situation is much improved for Hermitian operators. Hermitian operators can be diagonalized and so can unitary operators. But even more is true: these operators take the diagonal form in an orthonormal basis!

An operator M is said to be **unitarily diagonalizable** if there is an *orthonormal* basis $\{\tilde{e}_i\}$ in which its matrix representation is a diagonal matrix. The basis $\{\tilde{e}_i\}$ is therefore an *orthonormal basis of eigenvectors*. This uses the first viewpoint on diagonalization.

Alternatively, start with an arbitrary orthonormal basis $(e_1, \ldots, e_n)$, where the matrix representation of M is written simply as the matrix M. We saw at the end of section 14.9 that orthonormal bases are mapped onto each other by unitary operators. It follows that there is a unitary operator U that maps the $\{e_i\}$ basis vectors to the orthonormal basis $\{\tilde{e}_i\}$ of eigenvectors. With $U^{-1} = U^\dagger$ and using the $\{e_i\}$ basis, we have the following matrix equation, with D_M a diagonal matrix:

$$U^\dagger M U = D_M. \tag{15.5.8}$$

This equation holds for the same reason that equation (15.5.5) holds. We say that the matrix M has been diagonalized by a unitary transformation, thus the terminology *unitarily diagonalizable*.

Exercise15.2. *Consider the operators T_1, T_2, T_3, and define $\tilde{T}_i = A^{-1}T_iA$, $i = 1, 2, 3$ obtained by similarity transformation with A. Show that $T_1T_2 = T_3$ implies $\tilde{T}_1\tilde{T}_2 = \tilde{T}_3$.*