Operators and Matrices

An **observable** is represented by an **operator**, which is a linear transformation acting on a vector of the space to produce another vector:

$$|\beta\rangle = \hat{Q}\,|\alpha\rangle\ . \tag{1}$$

With respect to a particular discrete basis $\{|e_n\rangle\}$, vectors are represented by components:

$$|\alpha\rangle = \sum_n a_n|e_n\rangle, \quad |\beta\rangle = \sum_n b_n|e_n\rangle\ . \tag{2}$$

For an orthonormal basis,

$$\langle e_m|\alpha\rangle = \langle e_m|\left(\Sigma a_n|e_n\rangle\right) = \Sigma a_n\langle e_m|e_n\rangle = \Sigma a_n\delta_{mn} = a_m\ , \tag{3}$$

so the components are

$$a_n = \langle e_n|\alpha\rangle\ , \quad b_n = \langle e_n|\beta\rangle\ . \tag{4}$$

From (1) and (2) we obtain

$$|\beta\rangle = \Sigma b_b|e_n\rangle = \hat{Q}\,\Sigma a_n|e_n\rangle = \Sigma a_n\hat{Q}\,|e_n\rangle\ , \tag{5}$$

and, using (4) ,

$$b_m = \langle e_m|\beta\rangle = \langle e_m|\left\{\Sigma a_n\hat{Q}\,|e_n\rangle\right\} = \Sigma\, a_n\langle e_m|\hat{Q}\,|e_n\rangle = \Sigma\, Q_{mn}\, a_n\ , \tag{6}$$

where $Q_{mn} \equiv \langle e_m|\hat{Q}\,|e_n\rangle$ is called the **matrix element** of the operator $\hat{Q}$ between the basis vectors $|e_m\rangle$ and $|e_n\rangle$.

The reason for this name is obvious if we recall the definition of matrix multiplication. The vectors $|\alpha\rangle$ and $|\beta\rangle$ can be represented by the **column matrices** of their components:

$$|\alpha\rangle \Rightarrow \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}\ , \quad |\beta\rangle \Rightarrow \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix}\ , \tag{7}$$

and the linear transformation of Eq. (6) is given by

$$|\beta\rangle \Rightarrow \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} & \cdots \\ Q_{21} & & \\ \vdots & & \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \Sigma Q_{1n}\, a_n \\ \Sigma Q_{2n}\, a_n \\ \vdots \end{pmatrix} \Rightarrow \hat{Q}|\alpha\rangle\ . \tag{8}$$

We see that the "matrix elements" $Q_{mn} \equiv \langle e_m|\hat{Q}\,|e_n\rangle$ constitute the elements of the matrix which performs the linear transformation corresponding to operator $\hat{Q}$.

We also need to define the **inner product** in matrix notation. Representing the vectors by their components in an orthonormal basis, and remembering the defining feature of the inner product,

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*\ , \tag{9}$$

the inner product can be written

$$\langle\alpha|\beta\rangle = \sum_n a_n^* b_n = \begin{pmatrix} a_1^* & a_2^* & \cdots \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix}\ . \tag{10}$$

Thus the inner product is given by the product of the **row matrix** of the complex conjugates of the components of the first vector by the **column matrix** of the components of the second vector.

Before proceeding further, we should review some matrix definitions:
The **transpose** $\tilde{\mathbf{T}}$ and the **hermitian conjugate** (or hermitian adjoint) $\mathbf{T}^\dagger$ of the matrix $\mathbf{T}$ are defined by

$$\mathbf{T} = \begin{pmatrix} T_{11} & T_{12} & \cdots \\ T_{21} & T_{22} & \cdots \\ \vdots \end{pmatrix}, \quad \tilde{\mathbf{T}} = \begin{pmatrix} T_{11} & T_{21} & \cdots \\ T_{12} & T_{22} & \cdots \\ \vdots \end{pmatrix}, \quad \mathbf{T}^\dagger = \begin{pmatrix} T_{11}^* & T_{21}^* & \cdots \\ T_{21}^* & T_{22}^* & \cdots \\ \vdots \end{pmatrix}, \tag{11}$$

i.e. $\tilde{T}_{ij} = T_{ji}$ and $T^\dagger{}_{ij} = T_{ji}{}^*$.
Note that the inner product (10) can be written $\langle \alpha \,|\, \beta \rangle = \mathbf{a}^\dagger\mathbf{b}$, where $\mathbf{a}$ and $\mathbf{b}$ are the column matrices representing the two vectors.

A few additional definitions:

symmetric matrix	$\tilde{\mathbf{T}} = \mathbf{T}$	$(T_{ij} = T_{ji})$
antisymmetric matrix	$\tilde{\mathbf{T}} = -\mathbf{T}$	$(T_{ij} = -T_{ji}$; all diagonal elements $T_{ii} = 0)$
hermitian matrix	$\mathbf{T}^\dagger = \mathbf{T}$	$(T_{ij} = T_{ji}{}^*)$

Hermitian operators and hermitian matrices

A hermitian operator $\hat{Q}$ is defined by the condition

$$\langle f|\hat{Q}g\rangle = \langle \hat{Q}f|g\rangle \text{ for any vectors } f \text{ and } g.$$

Represent f and g by the column matrices of their components:

$$|f\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}, \quad |g\rangle = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix}.$$

Then

$$\left|\hat{Q}f\right\rangle = \hat{Q}\left|f\right\rangle = \begin{pmatrix} Q_{11} & Q_{12} & \cdots \\ Q_{21} & & \\ \vdots & & \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \end{pmatrix}, \quad \text{or} \quad \alpha_i = \sum_j Q_{ij}\, a_j \tag{12}$$

$$\left|\hat{Q}\,g\right\rangle = \hat{Q}\left|g\right\rangle = \begin{pmatrix} Q_{11} & Q_{12} & \cdots \\ Q_{21} & & \\ \vdots & & \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \end{pmatrix}, \quad \text{or} \quad \beta_i = \sum_j Q_{ij}\, b_j \tag{13}$$

$$\left\langle f \,|\, \hat{Q}\,g\right\rangle = \sum_i a_i^*\, \beta_i = \sum_i a_i^* \left(\sum_j Q_{ij} b_j\right) = \sum_i \sum_j Q_{ij}\, a_i^*\, b_j \tag{14}$$

$$\left\langle \hat{Q}\,f \,|\, g\right\rangle = \sum_i \alpha_i^*\, b_i = \sum_i \left(\sum_j Q_{ij}^* a_j^*\right) b_i = \sum_i \sum_j Q_{ij}^*\, a_j^*\, b_i = \sum_i \sum_j Q_{ji}^*\, a_i^*\, b_j \tag{15}$$

If $\hat{Q}$ is a hermitian operator, then expressions (14) and (15) are equal, which requires $Q_{ij} = Q_{ji}{}^*$.
This proves the desired result: In an orthonormal basis, **hermitian operators are represented by hermitian matrices**.

Eigenvectors of a linear transformation (see Griffiths Appendix A.5)

Every linear transformation has some special vectors such that

$$\hat{\mathbf{T}}|\alpha\rangle = \lambda|\alpha\rangle \qquad \text{(eigenvalue equation)} \tag{16}$$

When this is true we say that "$|\alpha\rangle$ is an eigenvector of $\hat{\mathbf{T}}$ with eigenvalue λ."
We can write this as a matrix equation, with $\mathbf{T}$ and $\mathbf{a}$ denoting the matrices representing the transformation $\hat{\mathbf{T}}$ and the vector $|\alpha\rangle$ respectively: $\mathbf{Ta} = \lambda\mathbf{a}$, or

$$(\mathbf{T} - \lambda\mathbf{I})\mathbf{a} = \mathbf{0} \qquad \text{(eigenvalue equation)} \tag{17}$$

where $\mathbf{I}$ is the identity matrix (1's on the diagonal) and $\mathbf{0}$ is the zero matrix (all elements 0).

If the matrix $(\mathbf{T} - \lambda\mathbf{I})$ had an inverse, we could multiply (17) by $(\mathbf{T} - \lambda\mathbf{I})^{-1}$ and obtain $\mathbf{a} = \mathbf{0}$, which cannot be true for a normalizable vector. Thus $(\mathbf{T} - \lambda\mathbf{I})$ is a singular matrix (a matrix with no inverse). Since the requirement for a matrix to possess an inverse is that its determinant is non-zero (see Griffiths [A.57]), a necessary condition for (17) is

$$\det(\mathbf{T} - \lambda\mathbf{I}) = \begin{vmatrix} (T_{11}-\lambda) & T_{12} & \cdots & T_{1n} \\ T_{21} & (T_{22}-\lambda) & \cdots & T_{2n} \\ \vdots & \vdots & & \vdots \\ T_{n1} & T_{n2} & \cdots & (T_{nn}-\lambda) \end{vmatrix} = 0 \ . \tag{18}$$

This gives an n^{th} degree polynomial equation in λ, the "characteristic equation" of the matrix, whose solutions λ_i are the eigenvalues of $\hat{\mathbf{T}}$. The number of eigenvectors is equal to the dimension n of the vector space; the number of distinct eigenvalues may be smaller because of degeneracy – some of the λ_i may be equal. Once the eigenvalues are known, the components a_n of the eigenvectors $|\alpha\rangle$ can be found by substituting into the individual equations $\mathbf{Ta} = \lambda_i\,\mathbf{a}$.

If the eigenvectors $\{f_i\}$ of $\hat{\mathbf{T}}$ span the space, we may use them as a basis. The n individual eigenvalue equations are of the form

$$\hat{\mathbf{T}}|f_i\rangle = \lambda_i|f_i\rangle. \tag{19}$$

In this basis, the matrices representation takes a very simple form. The column matrices representing the eigenvectors in the basis of eigenvectors are just the "unit vectors":

$$|f_1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}, \quad |f_2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \end{pmatrix}, \quad \ldots \tag{20}$$

Comparing (19) and (20), we see that row i of the matrix representing $\hat{\mathbf{T}}$ can have a non-0 entry only in column i, and that value must be the eigenvalue λ_i :

$$\mathbf{T} = \begin{pmatrix} \lambda_1 & 0 & \cdots \\ 0 & \lambda_2 & \cdots \\ \vdots & & \ddots \end{pmatrix}. \tag{21}$$

Thus the matrix $\mathbf{T}$ representing $\hat{\mathbf{T}}$ in the basis of the eigenvectors of $\hat{\mathbf{T}}$ is a **diagonal** matrix. For such a matrix, (18) is satisfied trivially: the matrix $(\mathbf{T} - \lambda\mathbf{I})$ is diagonal, and its determinant is just the product of the diagonal elements. For λ equal to any eigenvalue λ_i , the diagonal element

$(T_{ii} - \lambda_i)$ is zero, making the determinant zero.

For this reason, the process of finding the eigenvectors and eigenvalues of a linear transformation is often referred to as "diagonalizing the matrix." There is a rather elaborate general procedure for diagonalizing a matrix by means of a so-called "similarity transformation", which changes from basis $\{|f_i\rangle\}$ (in which **T** is not diagonal) to a new basis $\{|e_i\rangle\}$ in which **T** is diagonal. It can be shown that both the **determinant** of the matrix and its **trace** (the sum of the diagonal elements) are unchanged by a similarity transformation. This procedure is described in Griffiths Appendix A.4, and will not be discussed further in this course.

For the purpose of solving Example 3.8, we note that the process for finding the eigenvalues and eigenvectors of a transformation can be summarized in the following steps:
1. Find the eigenvalues of $\mathbf{T}^{\mathbf{f}}$ (the matrix in the original basis $\{|f_i\rangle\}$) by solving the characteristic equation (18).
2. For each eigenvalue λ_i, write the eigenvalue equation (16) for a vector with arbitrary components $\{a_i\}$, and solve to find the components of the eigenvectors in the old basis, normalizing as needed.

Solution of a two-state system (Griffiths Example 3.8)

Consider a system with two linearly independent basis states $|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. (22)

The general (normalized) vector in this basis is

$$|\psi(t)\rangle = a|1\rangle + b|2\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \text{with } |a|^2 + |b|^2 = 1, \tag{23}$$

which satisfies the time-dependent Schrödinger equation

$$\hat{H}|\psi(t)\rangle = i\hbar \frac{d}{dt}|\psi(t)\rangle$$

for a given Hamiltonian. The coefficients $a(t)$ and $b(t)$, which include the time-dependent factors, specify all that can be known about the state of the system.

If the basis states $|1\rangle$ and $|2\rangle$ are eigenstates of the Hamiltonian, then $|a|^2$ and $|b|^2$ are constant, so the probability of finding the particle in a particular energy eigenstate does not vary with time. In this case, the matrix which represents the Hamiltonian is **diagonal**,

$$\mathbf{H} = \mathbf{H_0} = \begin{pmatrix} E_1^0 & 0 \\ 0 & E_2^0 \end{pmatrix}.$$

Suppose, instead, that there is some small interaction which **mixes** the basis states in some way:

$$\langle 1|\hat{T}|2\rangle \neq 0.$$

Then (as you will see next semester) this interaction can **perturb** the Hamiltonian,

$$\hat{H}_{total} = \hat{H}_0 + \hat{H}_{pert},$$

where the two terms on the right are represented by diagonal and off-diagonal terms in the Hamiltonian, and the basis states $|1\rangle$ and $|2\rangle$ are no longer eigenstates of the full Hamiltonian.

In this case the Hamiltonian becomes something like

$$\boldsymbol{H} = \begin{pmatrix} h & g \\ g & h \end{pmatrix} , \tag{25}$$

and solving the problem consists of finding its eigenvalues and eigenvectors, using the ideas of the previous sections of this note. In a typical application, the off-diagonal elements g are small compared to the diagonal elements h, but the derivation below is valid for arbitrary g and h. (The equality of the diagonal elements in this example implies that the unperturbed system is degenerate, with energies $E_1^0 = E_2^0 = h$.)

1) Finding the eigenvalues λ: $$\boldsymbol{H}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix} . \tag{26}$$

The characteristic equation is $\begin{vmatrix} h-\lambda & g \\ g & h-\lambda \end{vmatrix} = (h-\lambda)^2 - g^2 = \lambda^2 - 2h\lambda + h^2 - g^2 = 0$, with solutions

$$\lambda_1 = h + g = E_1 \quad \text{and} \quad \lambda_2 = h - g = E_2 \tag{27}$$

(since the eigenvalues of the Hamiltonian operator are the stationary-state energies.) Note that if $g \ll h$, the energies E_1 and E_2 are nearly equal.

2) Finding the eigenvectors.

For $\lambda_1 = E_1 = h + g$: $\begin{pmatrix} h & g \\ g & h \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = (h+g)\begin{pmatrix} a \\ b \end{pmatrix}$ becomes 2 equations:
$ha+gb = (h+g)a \;\Rightarrow\; b=a$. (The second equation gives the same result.)

Thus the normalized eigenvector corresponding to eigenvalue λ_1 is

$$\left|e_1\right\rangle = \begin{pmatrix} a \\ a \end{pmatrix} = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \tfrac{1}{\sqrt{2}}\left\{\left|1\right\rangle + \left|2\right\rangle\right\}. \tag{28}$$

For $\lambda_2 = E_2 = h - g$: $\begin{pmatrix} h & g \\ g & h \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = (h-g)\begin{pmatrix} a \\ b \end{pmatrix} \Rightarrow ha+gb = (h-g)a \;\Rightarrow b=-a$.

$$\left|e_2\right\rangle = \begin{pmatrix} a \\ -a \end{pmatrix} = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \tfrac{1}{\sqrt{2}}\left\{\left|1\right\rangle - \left|2\right\rangle\right\}. \tag{29}$$

Note that either eigenvector can be multiplied by -1 or i (or any complex number of modulus 1) without changing its validity as a normalized eigenvector of $\boldsymbol{H}$.

3) Time-dependent solution

The system starts in state $\left|1\right\rangle$ at $t = 0$. Thus, taking linear combinations of (28) and (29),

$$\left|\psi(0)\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \tfrac{1}{\sqrt{2}}\left(\left|e_1\right\rangle + \left|e_2\right\rangle\right). \tag{30}$$

Adding the time dependence in the usual way,

$$\left|\psi(t)\right\rangle = \frac{1}{\sqrt{2}}\left(\left|e_1\right\rangle e^{-iE_1 t/\hbar} + \left|e_2\right\rangle e^{-iE_2 t/\hbar}\right). \tag{31}$$

Using (22), (27) and a little algebra, we obtain

$$\left|\psi(t)\right\rangle = e^{-iht/\hbar}\begin{pmatrix}\cos(gt/\hbar)\\ -i\sin(gt/\hbar)\end{pmatrix} = e^{-iht/\hbar}\left[\cos(gt/\hbar)\begin{pmatrix}1\\0\end{pmatrix} - i\sin(gt/\hbar)\begin{pmatrix}0\\1\end{pmatrix}\right], \tag{32}$$

so the state function oscillates between the basis states $\left|1\right\rangle$ and $\left|2\right\rangle$ with frequency $g/\hbar$.

Griffiths points out that this is a crude model for various physical phenomena including neutrino oscillations. The existence of the off-diagonal term g in the Hamilonian allows the two basis states to transform spontaneously into one another and back, even if g is much smaller than h. Note that $\left|\psi(t)\right\rangle$ remains normalized at all times, and satisfies the time-dependent Schrödinger equation (show this!).