6. EIGENVALUE EQUATIONS

Many important problems in physics can be cast as equations of the generic form

$$H\,\psi = \lambda\,\psi\,. \qquad (6.1)$$

where *H* is a linear operator whose domain and range is a Hilbert space, ψ is a function in the space, and λ is a constant. The operator *A* is known, but both ψ and λ are unknown and the task at hand is to solve Eq. (6.1). Because the solutions to an equation of this type yield functions that are unchanged by the operator (except for multiplication by a scale factor λ), they are termed **eigenvalue equations**: **Eigen** is German for "[its] own." A *function* that solves an eigenvalue equation is called an **eigenfunction**, and the value of λ that goes with an eigenfunction is called an **eigenvalue**.

If H is a 2x2 matrix (and I is 2x2 identity matrix), then the matrix representation of the operator and the eigenfunction is

$$(H - I\,\lambda)\,\psi = 0\,. \qquad (6.2)$$

The condition for a non trivial solution (ψ *not equal to 0*) is

$$\det(\mathsf{H} - \lambda\mathbf{1}) = \begin{vmatrix} h_{11} - \lambda & h_{12} \\ h_{21} & h_{22} - \lambda \end{vmatrix} = 0. \qquad (6.3)$$

Expanding the determinant, which is sometimes called a **secular determinant** (the name arising from early applications in celestial mechanics), we have an algebraic equation, the **secular equation**

$$(h_{11} - \lambda)(h_{22} - \lambda) - h_{12}h_{21} = 0, \qquad (6.5)$$

which can be solved for λ.

The left hand side of Eq. (6.5) is also called the **characteristic polynomial** (in λ) of H, and Eq. (6.5) is for that reason is also known as the **characteristic equation** of H.

* Once a value of λ that solves Eq. (6.5) has been obtained, we can return to the homogeneous equation system, Eq. (6.2), and solve it for the vector ψ.
* This can be repeated for all λ that are solutions to the secular equation, thereby giving a set of eigenvalues and the associated eigenvectors.

Example: For the Operator matrix

$$\mathsf{H} = \begin{pmatrix} -2 & \sqrt{5} \\ \sqrt{5} & -6 \end{pmatrix},$$

the secular equation takes the form

$$\det(\mathsf{H} - \lambda\mathbf{1}) = \begin{vmatrix} -2 - \lambda & \sqrt{5} \\ \sqrt{5} & -6 - \lambda \end{vmatrix} = \lambda^2 + 8\lambda + 7 = 0.$$

$$(\lambda + 1)(\lambda + 7) = 0.$$

we see that the secular equation has as solutions the eigenvalues λ= -1 and λ = -7. To get the eigenvector corresponding to λ= -1, we return to Eq. (6.2), which, written in great detail, is

$$(\mathsf{H} - \lambda\mathbf{1})\mathbf{r} = \begin{pmatrix} -2 - (-1) & \sqrt{5} \\ \sqrt{5} & -6 - (-1) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & \sqrt{5} \\ \sqrt{5} & -5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0,$$

which expands into a linearly dependent pair of equations:

$$-x + \sqrt{5}\, y = 0$$

$$\sqrt{5}\, x - 5y = 0.$$

This is, of course, the intention associated with the secular equation, because if these equations were linearly independent they would inexorably lead to the solution . $x = y = 0.$

Instead, from either equation, we have $x = \sqrt{5}\, y;$ so we have the eigenvalue/eigenvector pair

$$\lambda_1 = -1, \quad \mathbf{r}_1 = C \begin{pmatrix} \sqrt{5} \\ 1 \end{pmatrix},$$

where C is a constant that can assume any value. Thus, there is an infinite number of x; y pairs that define a **direction** in the 2-D space, with the magnitude of the displacement in that direction arbitrary. The arbitrariness of scale is a natural consequence of the fact that the equation system was homogeneous; any multiple of a solution of a linear homogeneous equation set will also be a solution.

We now consider the possibility that $\lambda = -7.$ This leads to a different eigenvector, obtained by solving

$$(\mathsf{H} - \lambda \mathbf{1})\mathbf{r} = \begin{pmatrix} -2+7 & \sqrt{5} \\ \sqrt{5} & -6+7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 & \sqrt{5} \\ \sqrt{5} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0,$$

corresponding to $y = -x\sqrt{5}.$ This defines the eigenvalue/eigenvector pair

$$\lambda_2 = -7, \quad \mathbf{r}_2 = C' \begin{pmatrix} -1 \\ \sqrt{5} \end{pmatrix}.$$

- ***Nondegenerate Eigenvalues***

Find the eigenvalues and eigenvectors of

$$\mathsf{H} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{6.7}$$

Writing the secular equation and expanding in minors using the third row, we have

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 - 1) = 0. \tag{6.8}$$

We see that the eigenvalues are 2, $+1$, and -1.

To obtain the eigenvector corresponding to $\lambda = 2$, we examine the equation set $[\mathsf{H} - 2(1)]\,\mathbf{c} = 0$:

$$-2c_1 + c_2 = 0,$$

$$c_1 - 2c_2 = 0,$$

$$0 = 0.$$

The first two equations of this set lead to $c_1 = c_2 = 0$. The third obviously conveys no information, and we are led to the conclusion that c_3 is arbitrary. Thus, at this point we have

$$\lambda_1 = 2, \quad \mathbf{c}_1 = \begin{pmatrix} 0 \\ 0 \\ C \end{pmatrix}. \tag{6.9}$$

Taking next $\lambda = +1$, our matrix equation is $[\mathsf{H} - 1(\mathbf{1})]\mathbf{c} = 0$, which is equivalent to the ordinary equations

$$-c_1 + c_2 = 0,$$
$$c_1 - c_2 = 0,$$
$$c_3 = 0.$$

We clearly have $c_1 = c_2$ and $c_3 = 0$, so

$$\lambda_2 = +1, \quad \mathbf{c}_2 = \begin{pmatrix} C \\ C \\ 0 \end{pmatrix}. \tag{6.10}$$

Similar operations for $\lambda = -1$ yield

$$\lambda_3 = -1, \quad \mathbf{c}_3 = \begin{pmatrix} C \\ -C \\ 0 \end{pmatrix}. \tag{6.11}$$

Collecting our results, and normalizing the eigenvectors (often useful, but not in general necessary), we have

$$\lambda_1 = 2, \quad \mathbf{c}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \lambda_2 = 1, \quad \mathbf{c}_2 = \begin{pmatrix} 2^{-1/2} \\ 2^{-1/2} \\ 0 \end{pmatrix}, \qquad \lambda_3 = -1, \quad \mathbf{c}_3 = \begin{pmatrix} 2^{-1/2} \\ -2^{-1/2} \\ 0 \end{pmatrix}.$$

Note that because H was block-diagonal, with an upper-left 2×2 block and a lower-right 1×1 block, the secular equation separated into a product of the determinants for the two blocks, and its solutions corresponded to those of an individual block, with coefficients of value zero for the other block(s). Thus, $\lambda = 2$ was a solution for the 1×1 block in row/column 3, and its eigenvector involved only the coefficient c_3. The λ values ± 1 came from the 2×2 block in rows/columns 1 and 2, with eigenvectors involving only coefficients c_1 and c_2. ■

➢ *Degenerate Eigenvalues*

14. Find the eigenvalues and eigenvectors: $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 2 & 2 & 3 \end{bmatrix}$

Solution: $A - \lambda I = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 2 & 2 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & 0 \\ 2 & 2 & 3-\lambda \end{bmatrix}$

$\mathrm{Det}\,(A - \lambda I) = \begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & 0 \\ 2 & 2 & 3-\lambda \end{vmatrix} = (2-\lambda)(2-\lambda)(3-\lambda) - (3-\lambda) = 0$

$(3-\lambda)\,[(2-\lambda)(2-\lambda) - 1] = (3-\lambda)\,[(\lambda\lambda - 4\lambda + 4 - 1]$

$(3-\lambda)\,[(\lambda^2 - 4\lambda + 4 - 1] = (3-\lambda)\,[(\lambda^2 - 4\lambda + 3] = (\lambda - 3)^2(\lambda - 1) = 0$

λ = 3, 3 and 1. As the root 3 is repeated twice it is doubly degenerate.

- *Considering the root*: 1

$$(A - 3I)X = \begin{bmatrix} 2-1 & -1 & 0 \\ -1 & 2-1 & 0 \\ 2 & 2 & 3-1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$= \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 2 & 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} x_1 & -x_2 & 0 \\ -x_1 & x_2 & 0 \\ 2x_1 & 2x_2 & 2x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow x_2 = x_1 \text{ and } x_3 = -x_1 - x_2 = -2x_2,\ V_1 = \begin{bmatrix} x_1 \\ x_1 \\ -2x_1 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

- *Next considering the root*: 3

$$(A - 3I)X = \begin{bmatrix} 2-3 & -1 & 0 \\ -1 & 2-3 & 0 \\ 2 & 2 & 3-3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} -1 & -1 & 0 \\ -1 & -1 & 0 \\ 2 & 2 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \Rightarrow -x_1 - x_2 = 0 \text{ and } 2x_1 + 2x_2 = 0$$

$$\Rightarrow x_2 = -x_1$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ is } \begin{bmatrix} x_1 \\ -x_1 \\ x_3 \end{bmatrix} \Rightarrow x_1\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \text{ and } x_3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

[*Hint*: Write $\begin{bmatrix} x_1 \\ -x_1 \\ x_3 \end{bmatrix}$ as $\begin{bmatrix} x_1 \\ -x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix}$ and then take $x_1 = 1, x_3 = 0$ and $x_1 = 0 \quad x_3 = 1$]

IDENTITY, INVERSE, ADJOINT

- **Inverse Matrix → A^{-1}**

Some, but not all operators will have an inverse, namely an operator that will "undo" its effect.

$$A^{-1}A = AA^{-1} = 1. \tag{5.46}$$

- **Adjoint matrix → $A^{\dagger}$**

Associated with many matrices (operators) will be another operator, called its **adjoint** and denoted $A^{\dagger}$ which will be such that for all functions *f* and *g* in the Hilbert space,

$$\langle f|Ag\rangle = \langle A^{\dagger} f|g\rangle. \tag{5.47}$$

Equation (5.47) is, in essence, the defining equation for $A^{\dagger}$. Thus, we see that $A^{\dagger}$ is an operator that, applied to the left member of **any** scalar product, produces the same result as is obtained if *A* is applied to the right member of the same scalar product.

- **Hermitian** Matrix

If $A = A^{\dagger}$, *A* is referred to as **self-adjoint**, or equivalently, **Hermitian**.

Any real symmetric matrix is Hermitian.

- *A***nti-Hermitian** matrix

If $A = -A^{\dagger}$, A is called **anti-Hermitian**.

- *U***nitary** matrix

If the adjoint $U^{\dagger}$, of an operator (matrix) U is equal to its inverse U^{-1}, the operator is called **unitary**.

If $U^{\dagger} = U^{-1}$, U is unitary. (5.49)

In the special case that U is both real and unitary, it is called **orthogonal.**

The column (or row) vectors of a unitary matrix are *orthonormal*, i.e. they are both orthogonal and normalized.

- ***Normal*** matrix

A square matrix is *normal* if it commutes with its conjugate transpose.

If A is real then A^T is normal.

Obviously unitary matrices, Hermitian matrices , and skew-Hermitian matrices are all normal.

But there exist normal matrices not belonging to any of these.

MATRIX DIAGONALIZATION

Consider the matrix U^{-1} formed by *normalized* eigenvectors corresponding to the nondegenerate eigenvalues 2 , 1, and -1 of the symmetric-real (Hermetian) matrix $H = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

(*Check if H is Hermitian*),

$$U^{-1} = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix}.$$

- *Then verify if* $U = (U^{-1})^{\dagger} = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}$. *Show also that* U is a *unitary matrix* ➔ $U^{-1} = U^{\dagger}$

We can rewrite the eigenvalue equation (6.1) :<>: $H\,\psi = \lambda\,\psi$ (6.1)

as $H\,U^{-1}\,U\,\psi = \lambda\,\psi$

operating from left on both sides by U :<>: $(U\,H\,U^{-1})\,(U\,\psi) = \lambda\,(U\,\psi)$

The above equation can be rewritten as :<>: $H'\,\psi' = \lambda\,\psi'$

where :<>: $H' = U\,H\,U^{-1}$ and $\psi' = U\,\psi$

Assignment: *Show that* $H' = U\,H\,U^{-1}$ is a diagonal matrix $H' = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

- Note that the diagonal elements directly yield the eigenvalues of the original Matrix H.

CAYLEY-HAMILTON THEOREM

This theorem is named after its founder mathematicians Arthur Cayley and William Hamilton.

Cayley-Hamilton theorem is an important theorem used in matrix theory.

Cayley-Hamilton theorem states that: "A square matrix satisfies its own characteristic equation."

Explanation: Let A be a square matrix of order n×n and if its characteristic polynomial is defined as: $P(\lambda) = | A- \lambda I_n |$,where I_n , is the identity matrix of same order as A.

Then, according to Cayley-Hamilton theorem :

P(A) = O , where O represents the zero matrix of same order as A.

That is if we replace λ by matrix A , then the relation would be equal to zero.

Hence matrix A annihilates its own characteristic equation.

Proof :

Let us assume a square matrix A of dimension $n \times n$. If $P(\lambda)$ be its characteristic polynomial, then by the definition of characteristic polynomial:

$$P(\lambda) = |A - \lambda I| = P_0 + P_1\lambda + P_2\lambda^2 + \ldots + P_n\lambda^n$$

Also, let us suppose that $Q(\lambda)$ be the adjoint matrix of $A - \lambda I$, such that:

$$Q(\lambda) = Q_0 + Q_1\lambda + Q_2\lambda^2 + \ldots + Q_k\lambda^k$$

We have the formula:

$$(adjA)A = (detA)I$$

$$Q(\lambda)(A - \lambda I) = P(\lambda)I$$

$$Q(\lambda)(A - \lambda I) = P_0I + P_1\lambda I + P_2\lambda^2 I + \ldots + P_n\lambda^n I \quad \text{\_\_\_\_\_\_\_\_\_(1)}$$

$$Q(\lambda)(A - \lambda I) = Q_0A + (Q_1A - Q_0)\lambda + (Q_2A - Q_1)\lambda^2 + \ldots + (Q_kA - Q_{k-1})\lambda^k - Q_k\lambda^{k+1} \quad \text{\_\_(2)}$$

On comparing (1) and (2), we get:

$k = n - 1$

and

$$Q_0A = P_0I$$
$$Q_1A - Q_0 = P_1I$$
$$Q_2A - Q_1 = P_2I$$

.

.

.

$-Q_k = P_nI$ or $-Q_{n-1} = P_nI$

On multiplying ascending powers of A in each equation,

$$Q_0A = P_0$$
$$Q_1A^2 - Q_0A = P_1A$$
$$Q_2A^3 - Q_1A^2 = P_2A^2$$

.

.

.

$$-Q_{n-1}A^n = P_nA^n$$

On adding all the equations together, Everything on left hand side cancels out and we obtain,

0 = $P_0 + p_1 A + P_2 A^2 + \ldots + P_n A^n$

or $P(A) = O$

Hence, the statement of Cayley Hamilton theorem is proved.

Example 1 : Prove Cayley-Hamilton theorem for the following matrix.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Solution: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

Let us find characteristic polynomial of given matrix.

$$A - \lambda I = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{bmatrix}$$

$$= (1-\lambda)(4-\lambda) - 6 \quad = 4 - 5\lambda + \lambda^2 - 6$$

$$P(\lambda) = \lambda^2 - 5\lambda - 2$$

Substituting A for λ in above equation

$$P(A) = A^2 - 5A - 2$$

$$= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 5 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - 5 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ 15 & 22 \end{bmatrix} - \begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$P(A) = O$, hence Cayley Hamilton theorem for given matrix A is proved.

Example 2 : If Cayley Hamilton theorem holds for the matrix, then find its inverse.

$$A = \begin{bmatrix} 2 & 0 & 1 \\ -2 & 3 & 4 \\ -5 & 5 & 6 \end{bmatrix}$$

Solution : $A = \begin{bmatrix} 2 & 0 & 1 \\ -2 & 3 & 4 \\ -5 & 5 & 6 \end{bmatrix}$

Its characteristic polynomial is -

$$P(\lambda) = |A - \lambda I| = \begin{vmatrix} 2-\lambda & 0 & 1 \\ -2 & 3-\lambda & 4 \\ -5 & 5 & 6-\lambda \end{vmatrix}$$

$P(\lambda) = -\lambda^3 + 11\lambda^2 - 21\lambda + 1$

According to Cayley Hamilton theorem -

P(A) = O

or $A^3 - 11A^2 + 21A - I = O$

$I = A^3 - 11A^2 + 21A$

Multiplying by A^{-1}, we get

$A^{-1} = A^2 - 11A + 21I$ ....**(1)**

$$A^2 = \begin{bmatrix} -1 & 5 & 8 \\ -30 & 29 & 34 \\ -50 & 45 & 51 \end{bmatrix}$$

Therefore, equation (1) becomes -

$$A^{-1} = \begin{bmatrix} -1 & 5 & 8 \\ -30 & 29 & 34 \\ -50 & 45 & 51 \end{bmatrix} - 11 \begin{bmatrix} 2 & 0 & 1 \\ -2 & 3 & 4 \\ -5 & 5 & 6 \end{bmatrix} + 21 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 5 & 8 \\ -30 & 29 & 34 \\ -50 & 45 & 51 \end{bmatrix} - \begin{bmatrix} 22 & 0 & 11 \\ -22 & 33 & 44 \\ -55 & 55 & 66 \end{bmatrix} + \begin{bmatrix} 21 & 0 & 0 \\ 0 & 21 & 0 \\ 0 & 0 & 21 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -2 & 5 & -3 \\ -8 & 17 & -10 \\ 5 & -10 & 6 \end{bmatrix}$$

=====xx====== ====xx==== ====xx==== ====xx====== ====xx=====