Definition:

The multiplicity (degeneracy) of an eigenvalue is the number of linearly-independent eigenvectors corresponding to that eigenvalue.

☆ Theorem: $P^{-1}AP = B \Rightarrow$ A and B have same eigenvalues and multiplicities.

Proof: $(B-\lambda I) = P^{-1}(A)P - \lambda I = P^{-1}(A-\lambda I)P$ (1)

$\det(B-\lambda I) = \det P^{-1}\det(A-\lambda I)\cdot\det P = \det(P^{-1}P)\det(A-\lambda I)$

$\therefore \det(B-\lambda I) = \det(A-\lambda I)$ Q.E.D (2)

Explanation: The characteristic eqn is a polynomial of λ^n, with all lower order terms in general present. (2) $\Rightarrow$

$0 = \det(B-\lambda I) = \det(A-\lambda I) \Rightarrow$ coefficients of each power of λ must be the same, even though $B=(b_{ij})$ and $A=(a_{ij})$ are not.

To see how this comes about examine 2x2 case:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow \det(A-\lambda I)=0=\det\begin{pmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{pmatrix} \quad (3)$$

$$\Rightarrow (a_{11}-\lambda)(a_{22}-\lambda)-a_{12}a_{21}=0 = \boxed{\lambda^2-\lambda\underbrace{(a_{11}+a_{22})}_{Tr\,A}+\underbrace{(a_{11}a_{22}-a_{12}a_{21})}_{\det A}=0} \quad (4)$$

$\therefore$ Characteristic equation $\Rightarrow \lambda^2 - \lambda\, Tr A + \det A = 0$

Had we done the same calculation with $B = P^{-1}AP$ we would have found

$\lambda^2 - \lambda(b_{11}+b_{22}) + (b_{11}b_{22}-b_{12}b_{21}) = \lambda^2 - \lambda\, Tr B + \det B$

But we have shown Previously that $\det B = \det$; $Tr B = Tr A \Rightarrow$ Same λ's

More Generally: Consider the general case: It can be shown that the solution of $\det(\lambda I - A) = \lambda^n + \lambda^{n-1} a_1 + \cdots$ can always be expressed in terms of invariant quantities.

Diagonalizing a Matrix A: If $P^{-1}AP = D$ = Diagonal then the diagonal elements are themselves the eigenvalues:

$$\det(D - \lambda I) = \begin{pmatrix} d_{11}-\lambda & 0 & 0 & \cdots \\ 0 & d_{22}-\lambda & & \\ \vdots & & & \\ & \cdots & & d_{nn}-\lambda \end{pmatrix} \Rightarrow (d_{11}-\lambda)(d_{22}-\lambda)(d_{33}-\lambda)\cdots 0$$

$$\lambda_1 = d_{11} \quad \lambda_2 = d_{22} \quad \cdots \qquad (1)$$

Q: What is the matrix P where $P^{-1}AP = D$?

A: Let $X_i = (x_{1i}, x_{2i}, \cdots x_{ni})$ be an eigenvector of A:

$$A X_i = \lambda_i X_i$$

Then $P = (x_{ij}) = \begin{pmatrix} x_{11} & x_{12} & & x_{1n} \\ x_{21} & x_{22} & & \vdots \\ x_{31} & \vdots & & \vdots \\ \vdots & & & x_{nn} \end{pmatrix}$ (2)

→ The eigenvectors of A are the columns which form P

this column is X_1

Proof: $AX_i = \lambda_i X_i$ (no sum on i) (3)

$\Rightarrow \sum_k a_{\ell k} x_{ki} = \lambda_i x_{\ell i}$ (to understand this, just keep i fixed) (4)

If $D = P^{-1}AP \Rightarrow AP = PD$, so we compute & compare:

$$(AP)_{\ell j} = \sum_k a_{\ell k} x_{kj} = \lambda_j x_{\ell j} \quad \text{(from (4))}$$

$$(PD)_{\ell j} = \sum_k x_{\ell k} d_{kj} = \sum_k x_{\ell k}(\lambda_k \delta_{kj}) = x_{\ell j}\lambda_j \text{ (no sum)} = (AP)_{\ell j} \checkmark$$

-FROM R. De Carlo-

F129.3

Definition 12.1. Let $A(\cdot) = [a_{ij}(\cdot)]$ be $n \times n$. Then the (matrix) *trace* of $A(\cdot)$, denoted $\mathrm{tr}[A(\cdot)]$, is

$$\mathrm{tr}[A(\cdot)] = \sum_{i=1}^{n} a_{ii}(\cdot) \tag{12.1}$$

The trace is simply the sum of the diagonal entries of the matrix.

The second is a statement of the Leverrier-Souriau-Faddeeva-Frame formula [1,2] stated here without proof.

Definition 12.2. The *Leverrier-Souriau-Faddeeva-Frame formula* for computing the coefficients a_i of the characteristic polynomial of A, $\pi_A(\lambda) = \det[\lambda I - A] = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_n$, is as follows:

(i) $N_1 = I$ $\qquad a_1 = -\mathrm{tr}[A]$

(ii) $N_2 = N_1 A + a_1 I$ $\qquad a_2 = -\frac{1}{2}\mathrm{tr}[N_2 A]$

(iii) $N_3 = N_2 A + a_2 I$ $\qquad a_3 = -\frac{1}{3}\mathrm{tr}[N_3 A]$

$\vdots$

and in general,

$$N_n = N_{n-1}A + a_{n-1}I \qquad a_n = -\frac{1}{n}\mathrm{tr}[N_n A] \tag{12.2}$$

where

$$[0] = N_n A + a_n I$$

As an example of the use of this algorithm, consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then from the formula, $\pi_A(\lambda) = \det[\lambda I - A] = \lambda^2 - 2$, – i.e., $a_1 = 0$ and $a_2 = -2$. Using the above algorithm

$$a_1 = -\mathrm{tr}[A] = 0$$

and

$$a_2 = -0.5\,\mathrm{tr}[N_2 A] = -0.5\,\mathrm{tr}[A^2] = -0.5\,\mathrm{tr}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = -2$$

which are the correct coefficients. The important point, for this chapter is that $a_1 = -\mathrm{tr}[A]$. Although of theoretical interest, the algorithm itself is numerically unstable.

From Ray DeCarlo ...

Simple Example: Find the eigenvalues and eigenvectors of the Pauli matrix σ_x. Also find the diagonalizing matrix P.

Solution: $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow \det(\sigma_x - \lambda I) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = 0$

$+\lambda^2 - 1 = 0 \quad \Rightarrow \quad \boxed{\lambda = \pm 1}$ ✓

$\lambda_1 = +1$

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = +1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow$

$x_1 = x_2$
$x_2 = x_1$

$\Rightarrow$ const $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\lambda_1 = -1$

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$x_1 = -x_2$
$x_2 = -x_1$

$\Rightarrow$ const $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Note: $Ax = \lambda x$ is linear in $x \Rightarrow$ overall const. not fixed (except by other considerations e.g. boundary conditions, normalization

Then $P = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$; it is trivial to show that $P^{-1} \sigma_x P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Check on Solutions: Since $\sigma_x^2 = I \Rightarrow$

$\sigma_x x = \lambda x \quad \Rightarrow \quad \sigma_x^2 x = \sigma_x \lambda x = \lambda \sigma_x x = \lambda^2 x$

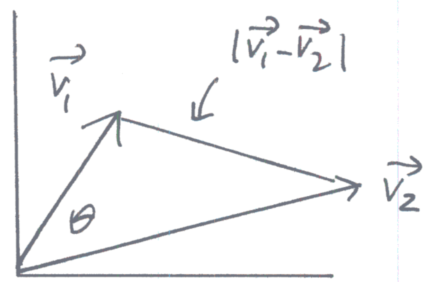
$\therefore \sigma_x^2 x = I x = x \equiv \lambda^2 x \Rightarrow \lambda^2 = 1 \Rightarrow \lambda = \pm 1$ ✓

INNER PRODUCT SPACES:

To introduce concepts such as length and angle into a vector space one must define an inner (scalar) product of two vectors in an appropriate way. When the "vectors" are themselves functions in a Hilbert space, this is not always trivial.

As a model consider 2-dim vectors $\vec{V}_1 = (x_1, y_1)$; $\vec{V}_2 = (x_2, y_2)$

Then $\vec{V}_1 \cdot \vec{V}_2 \equiv x_1 x_2 + y_1 y_2$



$|\vec{V}_1 - \vec{V}_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

Also: $|\vec{V}_1| = \sqrt{x_1^2 + y_1^2}$

Since $\vec{V}_1 \cdot \vec{V}_2$ can also be written as $\vec{V}_1 \cdot \vec{V}_2 = |\vec{V}_1||\vec{V}_2| \cos\theta \Rightarrow$

$$\boxed{\theta = \cos^{-1} \frac{\vec{V}_1 \cdot \vec{V}_2}{|\vec{V}_1||\vec{V}_2|}}$$

Notation: $\underbrace{\vec{V}_1 \cdot \vec{V}_2}_{\text{high school}} = \underbrace{(\vec{V}_1, \vec{V}_2)}_{\text{mathematicians}} = \langle \vec{V}_1 | \vec{V}_2 \rangle$ → PHYSICISTS

Scalar Products of Complex Vectors:

$\vec{V}_1 = (ix, 0) \Rightarrow |\vec{V}_1| = \sqrt{x_1^2 + y_1^2} \rightarrow \sqrt{(ix)^2} = \sqrt{-x^2} = $ imaginary!

So we need a generalized definition: $\vec{A} = (A_x, A_y)$ $\vec{B} = (B_x, B_y)$

Then

$$\boxed{\langle A | B \rangle \equiv A_x^* B_x + A_y^* B_y \Big\} \text{ vector on left is complex conjugated}}$$

This generalizes to any scalar product as follows:

Conditions on Any Scalar Product:

1) $\langle A|B\rangle = \langle B|A\rangle^*$

2) $\langle \alpha_1 A_1 + \alpha_2 A_2 | B\rangle = \alpha_1^* \langle A_1|B\rangle + \alpha_2^* \langle A_2|B\rangle$

3) $\langle A|A\rangle \geq 0$

$\langle A|A\rangle = 0 \Leftrightarrow A = 0$

This holds for any vectors, including complex functions in Hilbert space. Recall that for any complex number $z = x+iy$ then $z^* = \bar{z} = x-iy$. This corrects the problem of an imaginary length:

$$|\vec{v}_1| = |(ix,0)| = |\langle ix|ix\rangle|^{1/2} = |(ix)^*(ix)|^{1/2} = x$$

So this ensures that even complex vectors have real lengths.

Definition: An inner product space is a vector space with an inner product defined.

Definition: Real Inner Product Space $\equiv$ EUCLIDEAN

Complex Inner Product Space $\equiv$ UNITARY

Examples: [1] $\left.\begin{array}{l} x = (\alpha_1, \alpha_2, \ldots, \alpha_n) \\ y = (\beta_1, \beta_2, \ldots, \beta_n) \end{array}\right\} \langle x|y\rangle = \sum_{i=1}^{n} \alpha_i^* \beta_i$

[2] In the space of polynomials in the variable t

$$\langle x(t)|y(t)\rangle = \int_0^1 dt\, x^*(t)\, y(t) \qquad 0 \leq t \leq 1$$

↳ this is the scalar product for solutions of the Schrödinger equation.

ORTHOGONALITY & COMPLETENESS

↳ A most important property of vectors

1) $\langle x|y\rangle = 0 \Leftrightarrow$ x and y are orthogonal

2) A set of vectors $\{x_1 \dots x_n\}$ is orthonormal if $\langle x_i|x_j\rangle = \delta_{ij}$

3) Any vector can be normalized by writing $x \to \dfrac{x}{|x|}$.

4) An orthonormal set is complete and linearly independent*:

* Proof: If $\{x_1, \dots, x_n\}$ is orthonormal then

$$\sum_i \alpha_i x_i = 0 \Rightarrow \langle x_j|\sum_i \alpha_i x_i\rangle = 0 = \sum_i \alpha_i \underbrace{\langle x_j|x_i\rangle}_{\delta_{ji}} = \text{[redacted]}\ \alpha_j = 0 \checkmark$$

BESSEL'S INEQUALITY: If $X = \{x_1, \dots, x_n\}$ is a finite orthonormal set and x is any vector, then if $\boxed{\alpha_i \equiv \langle x_i|x\rangle}$ (1)

$$\boxed{\sum_i |\alpha_i|^2 \le |x|^2} \quad (2)$$

In addition $x' = x - \sum_i \alpha_i x_i$ is orthogonal to all the x_j.

Proof:

$$0 \le |x'|^2 = \langle x'|x'\rangle = \langle x - \sum_i \alpha_i x_i | x - \sum_j \alpha_j x_j\rangle = \langle x|x\rangle - \langle \sum_i \alpha_i x_i | x\rangle - \langle x|\sum_j \alpha_j x_j\rangle + \langle \sum_i \alpha_i x_i|\sum_j \alpha_j x_j\rangle \quad (3)$$

$$= |x|^2 - \sum_i \alpha_i^* \underbrace{\langle x_i|x\rangle}_{\alpha_i} - \sum_j \alpha_j \underbrace{\langle x|x_j\rangle}_{\alpha_j^*} + \sum_{i,j} \alpha_i^* \alpha_j \underbrace{\langle x_i|x_j\rangle}_{\delta_{ij}} \quad (4)$$

$$= |x|^2 - \sum_i |\alpha_i|^2 - \cancel{\sum_j |\alpha_j|^2} + \cancel{\sum_i |\alpha_i|^2} = |x|^2 - \sum_i |\alpha_i|^2 \quad (5)$$

$$\therefore \quad x^2 - \sum_i |\alpha_i|^2 \ge 0 \quad \text{or} \quad \boxed{|x|^2 \ge \sum_i |\alpha_i|^2} \quad (6)$$

Next consider $\langle x'|x_j\rangle = \langle x - \sum_i \alpha_i x_i | x_j\rangle = \langle x|x_j\rangle - \sum_i \alpha_i^* \underbrace{\langle x_i|x_j\rangle}_{\delta_{ij}}$ (7)

$$= \alpha_j^* - \sum_i \alpha_i^* \delta_{ij} = \alpha_j^* - \alpha_j^* = 0 \quad (8)$$