CHANGE OF BASIS AND SIMILARITY

We have seen that a matrix (a_{ij}) represents an abstract linear transformation w.r.t. a specific basis. What then happens when we change basis? We can address this by asking the following questions:

Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ be two different bases for an n-dim. vector space V.

<u>Q1</u>: If $x = \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=1}^{n} \beta_i y_i$ what is the relation between $\{\alpha_i\}$ and $\{\beta_i\}$ (1)

<u>Q2</u>: Given a set of constants $\{\alpha_i\}$ what is the relation between

$$x = \sum_{i=1}^{n} \alpha_i x_i \quad \text{and} \quad y = \sum_{i=1}^{n} \alpha_i y_i \tag{2}$$

<u>ANSWERS</u>: Since $\{x_i\}$ and $\{y_i\}$ are both bases there must be some linear transformation relating them:

$$\boxed{y_j = Ax_j = \sum_i a_{ij} x_i} \tag{3}$$

<u>A1</u>:

$$x = \sum_j \beta_j y_j = \sum_j \beta_j \left(\sum_i a_{ij} x_i\right) = \sum_i \left(\sum_j a_{ij} \beta_j\right) x_i \equiv \sum_i \alpha_i x_i \tag{4}$$

$$\boxed{\therefore \ \alpha_i = \sum_j a_{ij} \beta_j} \tag{5}$$

<u>A2</u>:

$$y = \sum_i \alpha_i y_i = \sum_i \alpha_i A x_i = A \sum_i \alpha_i x_i = Ax \tag{6}$$

$$\boxed{\therefore \ y = Ax} \tag{7}$$

Q3: Let L be a lin. transf. on V which has the following representations in the bases X and Y F124

$X = \{x_i\}$ $\quad Y = \{y_i\}$

(8)

an abstract lin. txfm. $L = \begin{cases} B = (b_{ij}) \text{ in } X \\ C = (c_{ij}) \text{ in } Y \end{cases}$ $\Rightarrow$ What is the relationship between the matrices B and C?

A3:

$$B y_j = B(Ax_j) = B\left(\sum_k a_{kj} x_k\right) = \sum_k a_{kj} B x_k \tag{9}$$

$$= \sum_k a_{kj} \underbrace{\left(\sum_i b_{ik} x_i\right)}_{\text{matrix rep. of } B \text{ w.r.t. } \{x_i\}} = \sum_i \left(\sum_k b_{ik} a_{kj}\right) x_i \tag{10}$$

However, the definition of (c_{ij}) is such that the linear transf. B (or L) when represented in terms of y_j is given by

$$B y_j \equiv \sum_k c_{kj} y_k = \sum_k c_{kj} A x_k = \sum_k c_{kj} \sum_i a_{ik} x_i \tag{11}$$

$\uparrow$ matrix rep. of A w.r.t. $\{x_i\}$

$$= \sum_i \left(\sum_k a_{ik} c_{kj}\right) x_i \tag{12}$$

Comparing (10) and (12) $\Rightarrow \sum_k b_{ik} a_{kj} = \sum_k a_{ik} c_{kj} \Rightarrow$ (13)

$BA = AC$

w.r.t $\{x_i\}$ $\quad$ w.r.t $\{y_i\}$

$$\Rightarrow \boxed{\underbrace{C}_{\text{w.r.t}\{y_i\}} = \underbrace{A^{-1}BA}_{\text{w.r.t.}\{x_i\}}} \tag{14}$$

This relates the matrix representation (C) of the abstract linear txfm. L given in the basis $\{y_i\}$ to its matrix rep. B given in terms of $\{x_i\}$ when the ~~basis~~ bases $\{x_i\}$ and $\{y_i\}$ are related to each other by $y_i = Ax_i$ (so that A is also represented in the basis $\{x_i\}$). Thus C and B are representations of the same txfm w.r.t. different bases. Eq. (14) is a

$\boxed{\text{SIMILARITY TRANSFORMATION}}$

Q4: If (b_{ij}) is a matrix what is the relation between the linear transformations B and C defined by:

$$Bx_j = \sum_i b_{ij} x_i \qquad Cy_j = \sum_i b_{ij} y_i$$

A4:

$$\equiv CAx_j \Rightarrow$$

$$Cy_j = \sum_i b_{ij} y_i = \sum_i b_{ij} (Ax_i) = A \sum_i b_{ij} x_i = A Bx_j$$

$$\therefore\ CAx_j = ABx_j \Rightarrow \qquad CA = AB \Rightarrow \boxed{C = ABA^{-1}}$$

Invariance of Properties of Matrices Under Similarity Transformations:

Let $C = A^{-1}BA$ then $\Rightarrow$

Thm:
a) $\det C = \det B$
b) $\operatorname{tr} C = \operatorname{tr} B$

Proof
a) $\det C = \det(A^{-1}BA) = (\det A^{-1})(\det B)(\det A)$

$$= \underbrace{(\det A)(\det A^{-1})}_{\det(AA^{-1}) = \det I = 1}(\det B) = \det B \checkmark$$

b) $A = (a_{ij})$ $\qquad A^{-1} = (\bar{a}_{ij})$ $\qquad C = (c_{ij})$ $\quad B = (b_{ij})$

$$AA^{-1} = I \Rightarrow \sum_k a_{ik}\bar{a}_{kj} = \delta_{ij}$$

$$C = A^{-1}BA \Rightarrow c_{ij} = \sum_{\ell,k} \bar{a}_{i\ell} b_{\ell k} a_{kj} \Rightarrow \operatorname{Tr} C = \sum_i c_{ii} = \sum_{i,\ell,k} \bar{a}_{i\ell} b_{\ell k} a_{ki}$$

$$= \sum_{\ell,k} \Big(\underbrace{\sum_i a_{ki}\bar{a}_{i\ell}}_{\delta_{k\ell}}\Big) b_{\ell k} = \sum_\ell b_{\ell\ell} = \operatorname{Tr} B \checkmark$$

THE EIGENVALUE PROBLEM

In many physical systems, particularly QM we confront the equation

$$\text{Operator} \rightarrow \underset{\uparrow}{\mathbf{M}} x = \underset{\uparrow}{\lambda} x \equiv \lambda I x \qquad (1)$$

eigenvector — eigenvalue

The eigenvalue problem involves finding the non-trivial solutions to (1).

$$(1) \Rightarrow \underbrace{(M - \lambda I)}_{A} x = 0 = Ax \qquad (2)$$

As noted previously (2) has non-trivial solutions only when A^{-1} does not exist $\Rightarrow \det A = 0$

$$\boxed{\therefore \text{ non-trivial} \Rightarrow \det(M - \lambda I) = 0} \quad (3)$$

$$M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix} \Rightarrow \underbrace{\det(M - \lambda I) = \begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ a_{31} & & & \\ \vdots & & & \\ a_{n1} & & & a_{nn}-\lambda \end{pmatrix} = 0}_{\text{Characteristic equation}} \quad (4)$$

polynomial of order λ^n $\Rightarrow$ n roots (not ~~necessarily~~ necessarily distinct)

I $\boxed{(\lambda_1 \cdots \lambda_n) = \text{Spectrum of } M}$ (5)

II Given $\{\lambda_n\}$, only certain eigenvectors x will solve the original equation (1). These are found by

$$\mathbf{M} x_1 = \lambda_1 x_1 \qquad M x_2 = \lambda_2 x_2 \quad \cdots \qquad (6)$$

Superposition (key in QM)

$$M x_1 = \lambda_1 x_1 \qquad M x_2 = \lambda_1 x_2 \Rightarrow M(\alpha x_1 + \beta x_2) = \alpha M x_1 + \beta M x_2 \qquad (7)$$

$$= \alpha \lambda_1 x_1 + \beta \lambda_1 x_2 = \lambda_1 (\alpha_1 x_1 + \beta x_2) \Rightarrow \underline{\underline{\alpha x_1 + \beta x_2 = \text{eigenvector too}}}$$

Definition:

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

☆ Theorem: $P^{-1} A P = B \Rightarrow$ 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:

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

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

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

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

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

But we have shown previously that $\det B = \det A$; $\mathrm{Tr}B = \mathrm{Tr}A \Rightarrow$ 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 $P^{-1} A P = D$ = 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 $P^{-1} A P = D$?

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: $A X_i = \lambda_i X_i$ (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 $D = P^{-1} A P \Rightarrow AP = PD$, 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 - (PURDUE) F129.3

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

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

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) $N_1 = I$ $\qquad a_1 = -\text{tr}[A]$

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

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

$\vdots$

and in general,

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

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

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

and

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

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

From Ray DeCarlo (Purdue)

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

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

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

$\lambda_1 = +1$

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

$x_1 = x_2$
$x_2 = x_1$

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

$\lambda_1 = -1$

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

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

$\Rightarrow \text{Const} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

<u>Note:</u> $Ax = \lambda x$ is linear in $x \Rightarrow$ <u>overall const. not fixed</u> (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 $P^{-1}\sigma_x P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

<u>Check on Solutions:</u> Since $\sigma_x^2 = I \Rightarrow$

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

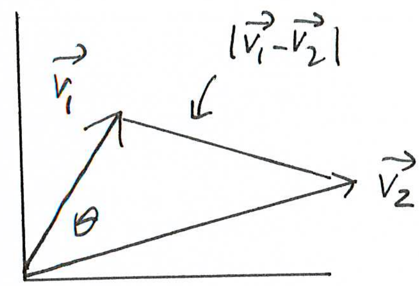
$\therefore \sigma_x^2 x = Ix = x \equiv \lambda^2 x \Rightarrow \lambda^2 = 1 \Rightarrow \lambda = \pm 1$ ✓

INNER PRODUCT SPACES:

To introduce concepts such as length and angle into a vectorspace 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: $\underbrace{\vec{V}_1 \cdot \vec{V}_2}_{\text{high school}} = \underbrace{(\vec{V}_1, \vec{V}_2)}_{\text{mathematicians}} = \langle \vec{V}_1 | \vec{V}_2 \rangle \to$ PHYSICISTS

Scalar Products of Complex Vectors:

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

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

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

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

[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 <u>orthogonal</u>

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

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

4) An orthonormal set is <u>complete</u> and <u>linearly independent</u>*:

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

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

<u>BESSEL'S INEQUALITY</u>: 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)

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

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

Proof:

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

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

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

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

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

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

Physical Interpretation of Bessel's Inequality:

Consider a 3-dim vector $V = \sum_{i=1}^{3} \alpha_i x_i$ $\alpha_i = \langle x_i | V \rangle = V_i$

$x_i = \hat{\imath}, \hat{\jmath}, \hat{k}$

Then we can also write $V = \sum_{i=1}^{3} V_i x_i$ (9)

$$\rightarrow |\vec{V}|^2 = \sum_{i=1}^{3} |V_i|^2$$

this holds if $\{x_i\}$ is a <u>complete</u> <u>orthonormal</u> (<u>CON</u>) basis.

Suppose now that $\{x_i\}$ are orthonormal, but not complete; <u>Ex</u>:

$\{x_i\} = \{\hat{\imath}, \hat{\jmath}\}$ (but not $\hat{k}$) in 3-dim. Then:

$$\sum_{i=1}^{2} |V_i|^2 = V_x^2 + V_y^2 = |\vec{V}|^2 - V_z^2 \Rightarrow \boxed{\sum_{i=1}^{2} |V_i|^2 \leq |\vec{V}|^2} \quad (10)$$

The equality would hold iff $V_z \equiv 0$.

Continuing this example, the second part of Bessel's inequality states that the vector $x' = x - \sum_j \alpha_j x_j$ is orthogonal to the other x_j's

In this example

$$\vec{x}' = \vec{V} - \sum_{i=1}^{2} V_i x_i = \vec{V} - V_x \hat{\imath} - V_y \hat{\jmath} \equiv V_3 \hat{k}$$

Obviously this vector is orthogonal to $V_x \hat{\imath} + V_y \hat{\jmath}$. ✓

COMPLETENESS:

This is a crucial concept. The following statements are equivalent descriptions of completeness:

Let $X = \{x_i\}$ be a set of m orthonormal vectors in a finite vector space V.

(1) X is complete

(2) If $\langle x_i | x \rangle = 0$ for $i = 1, \dots, m \Rightarrow x = 0$

(3) X spans V $[x \in V \Rightarrow x = \sum_i \alpha_i x_i]$

(4) If $x \in V$ then $x = \sum_i \langle x_i | x \rangle x_i$

(5) If $x, y \in V$ then $\boxed{\langle y | x \rangle = \sum_i \langle y | x_i \rangle \langle x_i | x \rangle}$ (1)

PARSEVAL'S IDENTITY

(6) If $x \in V$ then $|x|^2 = \sum_i |\langle x_i | x \rangle|^2 = \sum_i |\alpha_i|^2$ (2)

Proof: We show (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (1)

(1) $\Rightarrow$ (2) [We show not(2) $\Rightarrow$ not(1)]: If $\langle x_i | x \rangle = 0$ but $x \neq 0$ then we can adjoin $x/|x|$ to X to form a larger set X' bigger than X, so X is not complete.

(2) $\Rightarrow$ (3) If X does not span V $\Rightarrow$ $\exists$ an x which cannot be written as

$$x = \sum_i \alpha_i x_i = \sum_i \langle x_i | x \rangle x_i \Rightarrow x' = x - \sum_i \langle x_i | x \rangle x_i \neq 0 \quad (3)$$

However, we can still show that $\langle x_j | x' \rangle = 0$ even though $x' \neq 0$ (4)

$$\langle x_j | x' \rangle = \langle x_j | x \rangle - \sum_i \langle x_i | x \rangle \underbrace{\langle x_j | x_i \rangle}_{\delta_{ji}} = \langle x_j | x \rangle - \langle x_j | x \rangle = 0 \quad (5)$$

$\therefore$ not(3) $\Rightarrow$ not(2)

(3) ⇒ (4): If every x has the form $x = \sum_j \alpha_j x_j$

(with α_j not yet specified) then form

$$\langle x_i | x \rangle = \sum_j \alpha_j \underbrace{\langle x_i | x_j \rangle}_{\delta_{ij}} = \alpha_i \longrightarrow \boxed{x = \sum_i \langle x_i | x \rangle x_i} \quad (6)$$

(4) ⇒ (5): If $x = \sum_i \alpha_i x_i$ & $y = \sum_j \beta_j x_j$; $\alpha_i = \langle x_i | x \rangle \quad \beta_j = \langle x_j | y \rangle$ (7)

Hence $$\langle y | x \rangle = \langle \sum_j \beta_j x_j \,|\, \sum_i \alpha_i x_i \rangle = \sum_{i,j} \beta_j^* \alpha_i \underbrace{\langle x_j | x_i \rangle}_{\delta_{ji}} = \sum_i \beta_i^* \alpha_i \quad (8)$$

$$\therefore \boxed{\langle y | x \rangle = \sum_i \langle y | x_i \rangle \langle x_i | x \rangle} \quad (9)$$

← "inserting complete set of states"
VERY USEFUL IN QM!!

PARSEVAL'S IDENTITY

(5) ⇒ (6): Set $y = x$ in PARSEVAL: $$\langle x | x \rangle = |x|^2 = \sum_i \underbrace{\langle x | x_i \rangle}_{\alpha_i^*} \underbrace{\langle x_i | x \rangle}_{\alpha_i} = \sum_i |\alpha_i|^2 \quad (10)$$

(6) ⇒ (1): If X was not complete and was therefore contained in a larger set $X' = \{x_i, x_0\}$ where $\langle x_0 | x_i \rangle = 0$. Then

$$|x_0|^2 = \sum_i |\langle x_i | x_0 \rangle|^2 = 0 \Rightarrow x_0 = 0$$

Hence the only vector not in X which is orthogonal to all the x_i is the zero vector (which is trivial).

This completes the proof (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (5) ⇒ (6) ⇒ (1).

SCHWARZ'S INEQUALITY:

[This is a generalization of $|\cos\theta| \leq 1$]

Theorem: For $x, y, \in V$ then $\boxed{|\langle x|y\rangle| \leq \sqrt{\langle x|x\rangle\langle y|y\rangle} = |x|^2|y|^2}$ $= |x|\,|y|$

Proof: if $y=0$ both sides vanish ✓

if $y \neq 0$ then the set consisting of $y/|y|$ by itself is orthonormal and therefore satisfies Bessel's inequality:

$$|\langle x|y\rangle|^2 \leq |x|^2|y|^2 \Rightarrow$$

$$\|$$

$$|\langle x|\frac{y}{|y|}\rangle|^2 \leq |x|^2 \quad \text{or} \quad |\alpha|^2 \leq |x|^2 \quad \text{Q.E.D}$$

EXAMPLES:

1) Euclidean Space $\Rightarrow |\cos\theta| \leq 1$

2) Unitary Space $C^n \Rightarrow$ CAUCHY INEQUALITY:

for $(\alpha_1, \ldots, \alpha_n)$ & $(\beta_1, \ldots \beta_n)$

$$\boxed{\Rightarrow \left|\sum_{i=1}^n \alpha_i^* \beta_i\right|^2 \leq \sum_{i=1}^n |\alpha_i|^2 \cdot \sum_{j=1}^n |\beta_j|^2}$$

3) In the space of polynomials:

$$\left|\int_0^1 dt\; x^*(t)y(t)\right|^2 \leq \int_0^1 dt\, |x(t)|^2 \cdot \int_0^1 dt\, |y(t)|^2$$

COMPLETE ORTHONORMAL SETS:

For convenience we want a starting point in calculations where our basis vectors are not only complete, but form a complete orthonormal (CON) set. The formal method for converting a complete set to a CON set is the GRAM-SCHMIDT method:

$$\text{Complete} \longrightarrow \text{CON}$$

Let $X = \{x_1, \dots, x_n\}$ be any basis in $\mathcal{V}$ (hence complete). We want to use these to form a $Y = \{y_1, \dots, y_n\} \ni \langle y_i | y_j \rangle = \delta_{ij}$ (1)

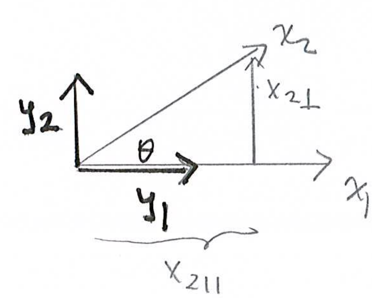
Start: ① use $x_1 \rightarrow \boxed{y_1 = x_1/|x_1|}$ (2)

② Next take x_2 and form $y_2 = (x_2 - \alpha_1 y_1)/|x_2 - \alpha_1 y_1|$ (3)

We determine $\alpha_1 \ni \langle y_1 | y_2 \rangle = 0 = \dfrac{\langle y_1 | x_2 \rangle - \alpha_1}{|x_2 - \alpha_1 y_1|} \Rightarrow \alpha_1 = \langle y_1 | x_2 \rangle$ (4)

Hence: $\boxed{y_2 = \dfrac{x_2 - y_1 \langle y_1 | x_2 \rangle}{|x_2 - y_1 \langle y_1 | x_2 \rangle|}}$ (5) *

Pictorially:



$$\vec{y}_2 = \frac{\vec{x}_2 - \vec{x}_{2\parallel}}{|\vec{x}_2 - \vec{x}_{2\parallel}|} = \frac{\vec{x}_{2\perp}}{|\vec{x}_{2\perp}|} \quad (6)$$

$$\vec{x}_{2\parallel} = \hat{y}_1 |\vec{x}_{2\parallel}| = \hat{y}_1 |\vec{x}_2| \underbrace{\cos\theta}_{\hat{x}_2 \cdot \hat{y}_1} = \hat{y}_1 \underbrace{|\vec{x}_2| \hat{x}_2}_{\vec{x}_2} \cdot \hat{y}_1 = \hat{y}_1 \, \vec{x}_2 \cdot \hat{y}_1 \quad (7)$$

$\therefore$ (6) & (7) $\Rightarrow \boxed{\vec{y}_2 = \dfrac{\vec{x}_2 - \vec{x}_2 \cdot \hat{y}_1 \hat{y}_1}{|\vec{x}_2 - \vec{x} \cdot \hat{y}_1 \hat{y}_1|}}$ (6) *

GRAM-SCHMIDT (continued)

This procedure can be repeated: At each step we start with one of the x_k and form the corresponding y_k by subtracting off the components of x_k that lie along the directions of the previously computed y_k.

Suppose that y_j vectors have been constructed, with $j = 1, 2, \ldots, r$ then

$$\boxed{|y_{r+1}\rangle = \frac{|x_{r+1}\rangle - \sum_{i=1}^{r} |y_i\rangle\langle y_i|x_{r+1}\rangle}{\left|\,|x_{r+1}\rangle - \sum_{i=1}^{r} |y_i\rangle\langle y_i|x_{r+1}\rangle\right|}}$$

GRAM-SCHMIDT (7)

Check: We can verify that $|y_{r+1}\rangle$ so constructed is indeed $\perp$ to the y_k ($k \leq r$): [Dropping the normalization]

$$\langle y_k|y_{r+1}\rangle = \langle y_k|x_{r+1}\rangle - \sum_{i=1}^{r} \underbrace{\langle y_k|y_i\rangle}_{\delta_{ki}}\langle y_i|x_{r+1}\rangle \qquad (8)$$

$$= \langle y_k|x_{r+1}\rangle - \langle y_k|x_{r+1}\rangle = 0 \;\checkmark$$

CONVENIENCE OF CON BASES

Recall: $\displaystyle A x_i = \sum_k \alpha_{ki} \overset{\text{CON}}{x_k} \Rightarrow \langle x_i|Ax_j\rangle = \langle x_i|\sum_k \alpha_{kj} x_k\rangle$

$$= \sum_k \alpha_{kj} \underbrace{\langle x_i|x_k\rangle}_{\delta_{ik}\ [\text{CON}]} = \alpha_{ij} \qquad (9)$$

$$\boxed{\begin{aligned} \therefore\ \alpha_{ij} &= \langle x_i|Ax_j\rangle \\ &= \langle x_i|A|x_j\rangle \end{aligned}} \quad (10) \longrightarrow \text{Useful in QM}$$

Self-Adjoint Transformations

↳ These are important in QM

<u>Def</u>: Adjoint Transformation : given a l.t. A on V, we define the operator $A^\dagger$ by:

$$\boxed{\langle Ax | y\rangle \equiv \langle x | A^\dagger y\rangle} \quad (1)$$

equivalent

Alternatively one can write: $\langle Ay|x\rangle^* = \langle x|Ay\rangle$

$$\langle y|A^\dagger x\rangle^* = \langle A^\dagger x|y\rangle \quad (2)$$

$$\boxed{\therefore \ \langle x|Ay\rangle = \langle A^\dagger x|y\rangle} \quad (3)$$

<u>Def</u>: A is self-adjoint if $A = A^\dagger$.

real V → <u>Symmetric</u>

Complex → <u>Hermitian</u>

Properties of the Adjoint

1) $A = A^\dagger$ & $B = B^\dagger \Rightarrow (A+B) = (A+B)^\dagger$

2) $A = A^\dagger$ & $\alpha \neq 0 \Rightarrow (\alpha A) = (\alpha A)^\dagger$ only if α = real

☆ 3) $(A)_{ij} = \alpha_{ij} \Rightarrow (A^\dagger)_{ij} = \alpha^*_{ji}$

<u>Proof</u>: $\alpha_{ij} = \langle x_i|Ax_j\rangle = \langle A^\dagger x_i|x_j\rangle = \langle x_j|A^\dagger|x_i\rangle^* = [(A^\dagger)_{ji}]^*$

$\therefore (A)_{ij} = \alpha_{ij} \Leftrightarrow \alpha_{ij} = (A^\dagger)_{ji}^* \Rightarrow \alpha^*_{ij} = (A^\dagger)_{ji}$

$\Rightarrow (A^\dagger)_{ij} = \alpha^*_{ji}$

If A is real $\Rightarrow A^\dagger = A^T$.

Adjoint Continued:

4) $(A^\dagger)^\dagger = A$

def. $A^\dagger$ — def. of [redacted] $*$ — def of $\dagger$

Proof: $\langle x | (A^\dagger)^\dagger y \rangle \stackrel{\text{def. } A^\dagger}{=} \langle A^\dagger x | y \rangle = \langle y | A^\dagger x \rangle^* = \langle A y | x \rangle^* = \langle x | A y \rangle$

$\therefore (A^\dagger)^\dagger = A$

5) $(AB)^\dagger = B^\dagger A^\dagger$

Proof: $\langle x | (AB)^\dagger y \rangle \stackrel{\text{def. } \dagger}{=} \langle AB x | y \rangle \equiv \langle A(Bx) | y \rangle = \langle Bx | A^\dagger y \rangle$

$= \langle x | B^\dagger A^\dagger y \rangle \Rightarrow (AB)^\dagger = B^\dagger A^\dagger$

Theorem: A is a l.f. on V. Then $A = 0 \Leftrightarrow \langle x | Ay \rangle = 0 \ \forall x, y$.

Proof: a) $A = 0 \Rightarrow \langle x | Ay \rangle = 0$ trivial

b) If $\langle x | Ay \rangle = 0 \ \forall x, y)$ let $x = Ay \Rightarrow \langle Ay | Ay \rangle = 0$

But by the axioms for an inner product space $\Rightarrow |Ay\rangle = 0$

But if $Ay = 0$ for all $y \Rightarrow$ $A = 0$ ✓

Theorem: Given a $\left\{ \begin{matrix} \text{self adjoint} \\ \text{arbitrary} \end{matrix} \right\}$ l.f. A on V where $V = \left\{ \begin{matrix} \text{inner product} \\ \text{unitary} \end{matrix} \right\}$ space then $A = 0 \Leftrightarrow \langle x | Ax \rangle = 0$.

Proof: not proved in class.

Theorem: If $A = A^\dagger$ & $B = B^\dagger$ then $AB = (AB)^\dagger$ or $BA = (BA)^\dagger$ only if $[A, B] = 0$.

Proof: $(AB)^\dagger = B^\dagger A^\dagger = BA \stackrel{?}{=} AB$ only if $[A, B] = 0$

<u>Definition:</u> If $A^\dagger = -A$ A is skew-symmetric or skew-Hermitian

<u>Note:</u> $A = \underbrace{\frac{1}{2}(A+A^\dagger)}_{\text{self-adjoint}} + \underbrace{\frac{1}{2}(A-A^\dagger)}_{\text{skew self-adjoint}}$

☆ <u>Theorem:</u> A is a l.t. on V ; <u>A is Hermitian iff. $\langle x|Ax\rangle$ = real</u>

<u>Proof:</u> a) If $A = A^\dagger \Rightarrow \langle x|Ax\rangle \underset{\text{def. } A^\dagger}{=} \langle A^\dagger x|x\rangle \underset{\text{def. } *}{=} \langle x|A^\dagger x\rangle^* \underset{A=A^\dagger}{=} \langle x|Ax\rangle^*$

$\therefore A = A^\dagger \Rightarrow \langle x|Ax\rangle$ is real.

b) If $\langle x|Ax\rangle$ is always real $(\forall x) \Rightarrow \langle x|Ax\rangle \underset{\text{assumed}}{=} \langle x|Ax\rangle^* \underset{\text{def. } *}{=} \langle Ax|x\rangle \underset{\text{def. } \dagger}{=} \langle x|A^\dagger x\rangle$

$\therefore \langle x|Ax\rangle = \langle x|Ax\rangle^* \Rightarrow A = A^\dagger$

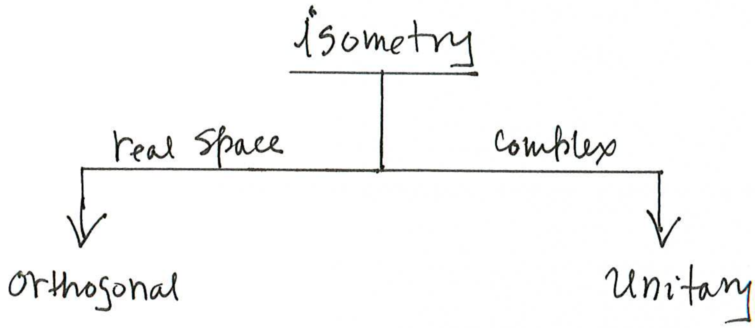
<u>Notation:</u> $\langle x|Ax\rangle \equiv$ expectation value of A in the state x.

In QM $\langle x|Ax\rangle \equiv \langle x|A|x\rangle$ is the value obtained by measuring the eigenvalue of A. Since this is a physical quantity $\langle x|Ax\rangle$ must be real. Hence, observables in QM are the (real) eigenvalues of Hermitian operators.

Isometries: Orthogonal & Unitary Transformations

Definition: A is an isometry if $A^{\dagger} = A^{-1}$ (iso = same, -metry = measure)

$A^{\dagger} = A^{-1} \Rightarrow A A^{\dagger} = 1 = A^{\dagger} A$



Summary of Transformations

$(A)_{ij} = \alpha_{ij}$

	Self Adjoint Transf. $(A^{\dagger} = A)$	Isometric Transf. $(A^{\dagger} = A^{-1})$
Unitary (complex) Space	Hermitian $(A^{\dagger})_{ij} = \alpha_{ji}^{*}$	Unitary $\sum_k \alpha_{ik} \alpha_{jk}^{*} = \delta_{ij}$
Euclidean (real) Space	Symmetric $(A^{\dagger})_{ij} = \alpha_{ji}$	Orthogonal $\sum_k \alpha_{ik} \alpha_{jk} = \delta_{ij}$

Theorem on Isometric Transformations:

The following 3 conditions are equivalent statements about unitary transformations:

a) $U^{\dagger} U = I$ (or $U^{\dagger} = U^{-1}$)

b) $\langle Ux | Uy \rangle = \langle x | y \rangle$

c) $|Ux| = |x|$

Unitary Transformations (cont'd):

Proof: We show (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1)

(1) $\Rightarrow$ (2) $\quad U^{\dagger}U = I \Rightarrow \langle Ux|Uy\rangle \underset{\uparrow \text{def } \dagger}{=} \langle x|\underbrace{U^{\dagger}U}_{I}y\rangle = \langle x|y\rangle$ ✓ (1)

(2) $\Rightarrow$ (3) $\quad \langle Ux|Uy\rangle = \langle x|y\rangle \underset{y=x}{\Rightarrow} \underbrace{\langle Ux|Ux\rangle}_{|Ux|^2} = \underbrace{\langle x|x\rangle}_{|x|^2} \Rightarrow |Ux|^2 = |x|^2$ (*)

$\Rightarrow |Ux| = |x|$ ✓ (2)

(3) $\Rightarrow$ (1) $\quad \left.\begin{array}{l} |Ux|^2 = \langle Ux|Ux\rangle = \langle x|U^{\dagger}Ux\rangle \\ \quad \hookrightarrow = \langle x|x\rangle \end{array}\right\} \Rightarrow U^{\dagger}U = I$ (3)

Isometric Transformations preserve the lengths of vectors, (*) and scalar products, and hence they also preserve the angles between vectors:

In an abstract (i.e. any) vector space

$$\cos\theta \equiv \frac{\langle x|y\rangle}{|x|\cdot|y|} \xrightarrow{U} \frac{\langle Ux|Uy\rangle}{|Ux|\cdot|Uy|} = \frac{\langle x|y\rangle}{|x|\cdot|y|} \checkmark \qquad (4)$$

Side Comment: If $\{x_i\}$ is a CON basis so is $\{Ux_i\}$.

Uniterity Constraints: $(U)_{ij} = u_{ij} \qquad (U^{\dagger})_{ij} = u^*_{ji} \Rightarrow$ (5)

$$\underset{\overset{\shortparallel}{\delta_{ik}}}{(I)_{ik}} = (UU^{\dagger})_{ik} = \sum_j u_{ij}(U^{\dagger})_{jk} = \sum_j u_{ij}u^*_{kj} \Bigg\} \boxed{\sum_j u_{ij}u^*_{kj} = \delta_{ik}} \quad (6)$$

↳ these are the familiar orthogonality relations when u_{ij} = real.

Check: 2x2 $\quad U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

$i=1 \quad k=1 \Rightarrow \delta_{11} = 1 \overset{?}{=} \sum_j u_{1j}u^*_{1j} = u_{11}u^*_{11} + u_{12}u^*_{12} = \cos^2\theta + \sin^2\theta = 1$ ✓

$i=1 \quad k=2 \Rightarrow \delta_{12} = 0 \overset{?}{=} \sum_j u_{1j}u_{2j} = u_{11}u^*_{21} + u_{12}u^*_{22} = \cos\theta(-\sin\theta) + \sin\theta\cos\theta$

$= 0$ ✓

(7)

Theorem: If A is orthogonal then $\det A = \pm 1$

Proof: $AA^+ = I \Rightarrow \det(AA^+) = (\det A)(\det A^+) = \det I = 1$ (8)

$= (\det A^{-1}) = \det A$

$\therefore (\det A)^2 = 1 \Rightarrow \boxed{\det A = \pm 1}$ (9)

Note that if A is orthogonal $\Rightarrow A^+ = A^T$ and $\det A^T = \det A$ (Recall volume argument)

$\det A = +1 \Rightarrow$ rotation

$\det A = -1 \Rightarrow$ inversion

Theorem: If U is Unitary then $|\det U| = 1 \Leftrightarrow \det U = e^{i\theta}$

Proof: Assume U is diagonalized so that $U = \begin{pmatrix} u_{11} & & 0 \\ & u_{22} & \\ 0 & & u_{33} \end{pmatrix}$ (10)

$$U^{-1} = \begin{pmatrix} 1/u_{11} & & 0 \\ & 1/u_{22} & \\ 0 & & 1/u_{33} \end{pmatrix} = U^+ \quad (11)$$

$$(\det U)(\det U^+) = (u_{11}u_{22}u_{33}) \frac{1}{u_{11}u_{22}u_{33}} = 1 \quad (12)$$

However, by definition the elements of U^+ are $\begin{pmatrix} u_{11}^* & & 0 \\ & u_{22}^* & \\ 0 & & u_{33}^* \end{pmatrix}$ (13)

Hence $U^+ = U^{-1} \Rightarrow u_{11}^* = 1/u_{11}$ etc $\Rightarrow |u_{11}|^2 = 1$

$\Rightarrow u_{11} = e^{i\theta_1}$ etc.

U must then have the form

$$U = \begin{pmatrix} e^{i\theta_1} & & 0 \\ & e^{i\theta_2} & \\ 0 & & e^{i\theta_3} \end{pmatrix} \Rightarrow \boxed{\det U = e^{i(\theta_1+\theta_2+\theta_3)} \equiv e^{i\theta} \Rightarrow |\det U| = 1} \quad (14)$$