Recall that a linear operator T on a vector space V of dim n is diagonalizable if $\exists$ a basis consisting of eigenvectors of T.

If $\{v_1, \ldots, v_n\}$ are linearly ind. eigenvectors of T, then T is a diagonalizable.

Theorem: Let T be a linear operator on a finite dimensional vector space V of dimension n. If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of T and suppose v_i is an eigenvector of T corresponding to λ_i for each i. Then $\{v_1, v_2, \ldots, v_k\}$ are linearly independent.

Proof: Let $a_1, \ldots, a_k$ be s.t

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0$$

$$(T - \lambda_k I) v_j = (\lambda_j - \lambda_k) v_j$$

$$\therefore \quad (T - \lambda_k I)(a_1 v_1 + \cdots + a_k v_k) = (T - \lambda_k I) 0 = 0$$

$$\Rightarrow \quad a_1 (\lambda_1 - \lambda_k) v_1 + \cdots + a_{k-1} (\lambda_{k-1} - \lambda_k) v_{k-1} = 0$$

Consider $(T - \lambda_{k-1} I) v_j = (\lambda_j - \lambda_{k-1}) v_j$

By a similar arguement as above, we have

$$a_1 (\lambda_1 - \lambda_k)(\lambda_1 - \lambda_{k-1}) v_1 + \cdots + a_{k-2} (\lambda_{k-2} - \lambda_k)(\lambda_{k-2} - \lambda_{k-1}) v_{k-1} = 0$$

After $k-1$ such steps, we have

$$a_1(\lambda_1-\lambda_k)(\lambda_1-\lambda_{k-1})\cdots(\lambda_1-\lambda_2)v_1 = 0$$

$$\Rightarrow \quad a_1 = 0 \quad \left(\text{since } (\lambda_1-\lambda_k)\cdots(\lambda_1-\lambda_2)\neq 0\right).$$

Hence we have

$$a_2v_2 + \cdots + a_kv_k = 0$$

By a similar argument, $a_2 = 0$. & so on.

Therefore

$\{v_1,\ldots,v_k\}$ are linearly independent. —■

Corollary: If an $n \times n$ matrix A has n distinct eigenvalues then A is diagonalizable.

Example: Consider $A = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}$

The characteristic polynomial of A is given by $f(\lambda) = (1-\lambda)(4-\lambda) + 2$

$$= \lambda^2 - 5\lambda + 6 . = (\lambda - 2)(\lambda - 3).$$

Example 1 : $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

The characterstic polynomial of A is

$f(\lambda) = (\lambda - 2)^2$

Example 2 : $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$

Characteristic polynomial of A is given by

$$f(\lambda) = (\lambda-2)^2.$$

Hence 2 is the only eigenvalue of A.

Let $v = (x, y)$ be an eigenvector.

$L_A v = Av = 2v$ i.e $(2x+y, 2y) = 2(x, y)$

$\Rightarrow$ $2x+y = 2x$ & $2y = 2y$ $\Rightarrow$ $y = 0$.

i.e $E_2 \subseteq W = \{(x, y) : y = 0\}$.

$\dim(W) = 1$.

$$\Rightarrow \quad \dim(E_2) \leq \dim(W) = 1.$$

$$\Rightarrow \quad \dim(E_2) = 1.$$

Hence there does not exist a basis of V consisting of eigenvectors of T.

Therefore T is not diagonalizable.

<u>Example 3</u> : Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

The characteristic poly of A is

$$f(\lambda) = (\lambda-2)^2(\lambda-3).$$

Hence 2 and 3 are eigenvalues of A.

Let $v \in E_2$. Then $L_A v = 2v$.

$$\begin{pmatrix} 2x \\ 2y+z \\ 3z \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}.$$

$$\Rightarrow 2y + z = 2y \quad \text{or} \quad z = 0.$$

$$E_2 \subset W_1 = \{ (x, y, z) : z = 0 \}.$$

$v_1 := (1,0,0)$, $v_2 := (0,1,0)$ is a basis of W_1

$$Av_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \& \text{ hence } v_1 \in E_2.$$

$$Av_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \& \text{ hence } v_2 \in E_2.$$

$$E_3 \subseteq W_2 = \{(x,y,z) : 2x = 3x,\ 2y + z = 3y,\ 3z = 3z\}.$$
$$= \{(x,y,z) : x = 0 \ \&\ y = z\}.$$
$$= \text{Span}\{(0,1,1)\}.$$

& note that $v_3 = (0,1,1) \in E_3$.

$\beta = ((1,0,0), (0,1,0), (0,1,1))$ is a basis of $\mathbb{R}^3$ consisting of eigenvectors of A.

Definition of (algebraic) multiplicity of an eigenvalue:
Let T be a linear operator on a finite dimensional vector space V of dimension n. Let λ_0 be an eigenvalue of T. Then the (algebraic) multiplicity of λ_0 is the largest

positive integer k s.t. $(\lambda-\lambda_0)^k$ divides the characteristic polynomial $f(\lambda)$ of T.

Theorem: Let T be a linear operator on a finite dim. vector space V. Let λ_0 be an eigenvalue of T. Then
$1 \leqslant \dim(E_{\lambda_0}) \leqslant$ multiplicity of λ_0.

Proof: Let $k = \dim(E_{\lambda_0})$

Then $\exists \{v_1, \ldots, v_k\}$ which is a basis of E_{λ_0}.

Extending this we get a basis $\beta = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$

$$[T]_\beta^\beta = \begin{pmatrix} \lambda_0 I_k & C_{k\times(n-k)} \\ O_{(n-k)\times k} & B_{(n-k)\times(n-k)} \end{pmatrix}$$

Then the characteristic poly of T is given

$$f(\lambda) = \det\left([T]_\beta^\beta - \lambda I_n\right) = \det\begin{pmatrix} (\lambda_0-\lambda)I_k & C \\ O & B-\lambda I_{n-k} \end{pmatrix}$$

$$= \det\left((\lambda_0 - \lambda) I_k\right) \det\left(B - \lambda I_{n-k}\right).$$

$$= (\lambda_0 - \lambda)^k g(\lambda)$$

Hence $(\lambda - \lambda_0)^k$ divides $f(\lambda)$.

$\Rightarrow \quad k \leq$ multiplicity of λ_0.

i.e $\dim(E_{\lambda_0}) \leq$ " ——— $\blacksquare$

Proposition: Let T be a linear operator on a finite dimensional vector space V. Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T and suppose $v_i \in E_{\lambda_i}$ for $i = 1, 2, \ldots, k$. If

$$v_1 + v_2 + \cdots + v_k = 0, \quad \text{then} \quad v_i = 0.$$

Proof: Exercise.

Theorem: Let T be a linear operator on V & $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T. Suppose S_j be a linearly

independent set consisting of eigenvectors with eigenvalue λ_j.
Then $S = S_1 \cup S_2 \cup \dots \cup S_k$ is linearly independent.

Proof: Let $S_j = \{ v_{j1}, v_{j2}, \dots, v_{jn_j} \}$.

Let a_{ij} be such that

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} a_{ji} v_{ji} = 0$$

Define $w_j = \sum_{i=1}^{n_j} a_{ji} v_{ji}$

then $w_j \in E_{\lambda_j}$ (since $w_j \in \text{span}(S_j)$ & $S_j \subseteq E_{\lambda_j}$)

Also $\displaystyle\sum_{j=1}^{k}\sum_{i=1}^{n_j} a_{ji} v_{ji} = \sum_{j=1}^{k} w_j = 0$

By the previous proposition, we have

$w_j = 0 \quad \forall\, j.$

$\Rightarrow \displaystyle\sum_{i=1}^{n_j} a_{ji} v_{ji} = 0 \quad \forall\, j$

$\Rightarrow \quad a_{ji} = 0 \quad \forall\, j \ \& \ \forall\, i$

& hence S is linearly independent.

Theorem: Let T be a linear operator on a finite dimensional vector space V such that the characteristic polynomial of T splits. Then T is diagonalizable iff the $\dim(E_{\lambda_i})$ = multiplicity of λ_i for each eigenvalue λ_i of T.

Proof: Let $n = \dim(V)$ & $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of T.

Assume that T is diagonalizable.

Let β be a basis of V consisting of eigenvectors of T.

Let $\beta_i = \beta \cap E_{\lambda_i}$ & $n_i = \# \beta_i$

Let $d_i = \dim(E_{\lambda_i})$ and m_i = multiplicity of λ_i

By a previous theorem, $d_i \leq m_i$

& the fact that a linearly ind. set in a vector space of dim d_i has size atmost d_i implies that

$n_i \leq d_i$

Notice that $\sum n_i = n$ (Since β is a basis)

Also $\sum m_i = \deg(f(\lambda)) = n$

$$n = \sum n_i \leq \sum d_i \leq \sum m_i = n$$

$$\left.\begin{aligned} &\Rightarrow \quad \sum (m_i - d_i) = 0 \\ &\text{But} \quad m_i - d_i \geq 0 \quad \forall\, i \end{aligned}\right\} *$$

$$(*) \Rightarrow \quad m_i - d_i = 0 \quad \forall\, i$$

Let us now assume that $d_i = m_i \quad \forall\, i$

Let β_i be a basis of E_{λ_i}.

By the previous theorem $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is linearly independent.

$$\# \beta = \sum \# \beta_i = \sum d_i = \sum m_i = n = \dim(V)$$

Hence β is a basis consisting of eigenvectors of T.

Hence T is a diagonalizable.

Definition: Let T be a linear transformation from a vector space to itself. A subspace W is said to be invariant under T or T-invariant if $Tv \in W$ $\forall$ $v \in W$ (i.e. $T(W) \subseteq W$).

Examples: 1) $\{0\}$ is a T-invariant for every linear operator T

2) V is T-invariant for every linear operator T.

3) $W = \mathrm{Null}(T)$ is a T-invariant subspace.

4) $W = R(T)$ is a T-invariant subspace.

5) Let v be an eigenvector with eigenvalue λ.

Let $W = \mathrm{span}(\{v\})$

$$T(av) = a\lambda v = \lambda a v \in \mathrm{span}(\{v\}).$$

Hence W is T-invariant.

6) E_λ the eigenspace corresponding to λ is T-invariant.

If $T|_W$ is the restriction of T to the subspace W, then $T|_W$ is a linear operator on W.

Theorem: Let T be a linear operator on a finite dimensional v.s. V and W be a T-invariant subspace. Then the characteristic polynomial of $T|_W$ divides the char. poly. of T.

Proof: Let $\alpha = (v_1, \ldots, v_k)$ be an ordered basis of W.

and $\beta = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ be an ordered basis of V.

$$[T|_W]_\alpha^\alpha = B \quad \text{a } k \times k \text{ matrix}$$

$$A = [T]_\beta^\beta = \begin{pmatrix} B & C_{k\times(n-k)} \\ O_{(n-k)\times k} & D_{(n-k)\times(n-k)} \end{pmatrix}$$

Let $f(\lambda)$ be the char. poly of T & $g(\lambda)$ be the char. poly of $T|_W$. i.e $f(\lambda) = \det(A - \lambda I_n)$ & $g(\lambda) = \det(B - \lambda I_k)$.

$$f(\lambda) = \det(A - \lambda I_n) = \det \begin{pmatrix} B - \lambda I_k & C \\ 0 & D - \lambda I_{n-k} \end{pmatrix}$$

$$= \det(B - \lambda I_k) \det(D - \lambda I_{n-k}).$$

$$= g(\lambda)\ p(\lambda).$$

& hence $g(\lambda) \mid f(\lambda)$. —— ■

Another <u>example of a T-invariant</u> subspace

Let $v \in V$ and define

$$W = \text{span}\{v, Tv, T^2v, \dots\}$$

where $T^k v = \underbrace{T(T(T\dots}_{k\text{-times}}(Tv))$

<u>Exercise</u>: W is invariant under T.

<u>Definition</u>: The subspace W is called the T-cyclic subspace generated by v.

Theorem: Let W be a T-cyclic subspace generated by a non-zero vector v. Suppose $\dim(W) = k$. Then $\{v, Tv, \dots, T^{k-1}v\}$ is a basis of W. Moreover, if $a_0 v + a_1 Tv + \cdots + a_{k-1}T^{k-1}v + T^k v = 0$, then the characteristic polynomial of $T|_W$ is given by

$$g(\lambda) = (-1)^k \left(a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k \right).$$

Proof: Let j be the largest positive integer such that $\{v, Tv, \dots, T^{j-1}v\}$ is a linearly independent set.

Then $T^j v \in \text{span}\{v, \dots, T^{j-1}v\}$.

<u>Claim</u>: $T^l v \in$ " for all $l \geq j$.

Let us assume that the claim is proved for upto $l-1$.

(Base case is $l=j$).

$T^l(v) = T(T^{l-1}v)$

But $T^{l-1}v \in \text{span}\{v, \dots, T^{j-1}v\}$.

$\Rightarrow T^{l-1}v = b_0 v + \dots + b_{j-1} T^{j-1}v$

$\Rightarrow T^l v = \left(b_0 Tv + \dots + b_{j-2} T^{j-1}v\right) + b_{j-1} T^j v$

$\in \text{span}(\{v, Tv, \dots, T^{j-1}v\})$.

i.e $\quad T^{l} v \in \operatorname{span}(v, \ldots, T^{j-1} v) \quad \forall\, l.$

$\Rightarrow \quad W \subset \operatorname{span}(v, \ldots, T^{j-1} v) \subseteq W$

$\operatorname{span}(v, \ldots, T^{j-1} v) = W$

Hence $\{v, \ldots, T^{j-1} v\}$ is a basis of W.

Since $\dim(W) = k$, we have $j = k$.

Let $\quad -T^{k} v = a_0 v + a_1 T v + \cdots + a_{k-1} T^{k-1} v.$

Then for $\beta = (v, Tv, \ldots, T^{k-1} v)$,

$$[T|_W]_\beta^\beta = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & 0 & & & \vdots \\ \vdots & \vdots & & 0 & \\ 0 & 0 & & 1 & -a_{k-1} \end{pmatrix}$$

Then the characteristic polynomial $g(\lambda)$ is given by

$$\det\left([T|_W]_\beta^\beta - \lambda I_k\right) = \det \begin{pmatrix} -\lambda & 0 & \cdots & 0 & -a_0 \\ 1 & -\lambda & & \vdots & \vdots \\ 0 & 1 & & \vdots & \vdots \\ \vdots & 0 & & & \\ \vdots & \vdots & & & \\ 0 & 0 & & -\lambda & -a_{k-1} \\ & & & 1 & -a_{k-1}-\lambda \end{pmatrix}$$

$$= -\lambda \det \begin{pmatrix} -\lambda & 0 & \cdots & & -a_1 \\ 1 & -\lambda & & & \vdots \\ 0 & 1 & & & \\ \vdots & 0 & & & \\ 0 & 0 & & & -a_{k-2} \\ & & & -a_{k-1} & -\lambda \end{pmatrix} + (-1)^{k+1}(-a_0) \underbrace{\det \begin{pmatrix} 1 & -\lambda & & 0 \\ & 1 & -\lambda & \\ & & \ddots & -\lambda \\ 0 & & & 1 \end{pmatrix}}_{= 1.}$$

$$= -\lambda \left((-1)^{k-1} \left(a_1 + a_2 \lambda + \cdots + a_{k-1} \lambda^{k-2} + \lambda^{k-1} \right) \right) + (-1)^k a_0$$

$$= (-1)^k \left(a_0 + a_1 \lambda + \cdots + a_{k-1} \lambda^{k-1} + \lambda^k \right).$$

Given a polynomial $p(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n$

we define $p(T) := a_0 I + a_1 T + \cdots + a_n T^n$.

Thus $p(T)v = a_0 v + a_1 Tv + \cdots + a_n T^n v$.

<u>Exercise</u>: $p(T)q(T)v = q(T)p(T)v \quad \forall\, v \in V$

& polynomials p, q.

Cayley - Hamilton Theorem :

Let V be a finite dim. vector space & T be a linear operator on V with characteristic polynomial $f(\lambda)$. Then $f(T)$ is the zero operator.

Proof: Enough to show that given a vector $v \in V$, $f(T)v = 0$.

Let W be the T-cyclic subspace generated by v.

Then W is invariant under T

& suppose $a_0 v + a_1 v + \cdots + a_{k-1} T^{k-1} v + T^k v = 0 \longrightarrow (*)$

then $g(\lambda) = (-1)^k (a_0 + a_1 \lambda + \cdots + \lambda^k)$ is the char. poly of $T/_W$.

$(*)$ can be rewritten as $g(T)v = 0$.

Let $f(\lambda)$ be the char. poly of T. By a theorem above, $\exists$ a polynomial $p(\lambda)$ s.t

$$f(\lambda) = p(\lambda)\, g(\lambda)$$

Hence $\quad f(T)v = p(T)\big(g(T)v\big)$

$$= p(T)(0)$$

$$= 0$$

Hence $f(T)$ is the zero operator. — $\blacksquare$