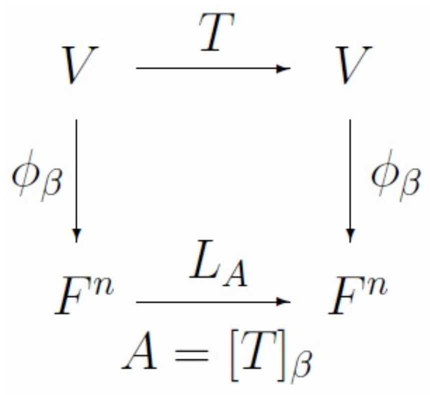
- **characteristic polynomial** of a linear operator $T : V \to V$:det($T - tI$)=det($[T]_\beta - tI$) for any basis β for V
- Finding eigenvalues and eigenvectors of T,
  $T \to [T]_\beta \to$det($[T]_\beta - \lambda I$) = 0 $\to \lambda$
  $\lambda \to ([T]_\beta - \lambda I)[v]_\beta = 0 \to [v]_\beta \to v$
- **eigenspace** E_λ
  - $E_\lambda = N(T - \lambda I) = N(L_{A-\lambda I_n})$ for a matrix $A \in M_{n\times n}$
    E_λ is a subspace of F^n.



- T-invariant set
  - $T-$**invariant set** $S : T(S) \subseteq S$
  - $T-$**invariant subspace** $W : W$ is a subspace and $T(W) \subseteq W$.
  - E_λ, $\{0\}$, $N(T)$, $R(T)$ and V are $T-$ invariant subspaces.
- Theorem 5.5 : $T : V \to V$ is a linear operator; $\lambda_1, \cdots, \lambda_k$ are "distinct" eigenvalues of T; and $v_1, \cdots, v_k$ are respective eigenvectors. Then the eigenvectors are linearly independent.
- Corollary 5.5: $T : V \to V$ is a linear operator; dim(V)$= n$. Then if T is n distinct eigenvalues, T is diagonalizable.
- The converses of Theorem 5.5 and corollary 5.5 are false.
- A polynomial $f(t)$ is said to **split** over the field F if there are $c, a_1, \cdots, a_n \in F$ such that $f(t) = c(t - a_1)(t - a_2)\cdots(t - a_n)$.

- Theorem 5.6; The characteristic polynomial of a diagonalizable linear operator splits.
  - Its converse is not true, i.e., splitting characteristic polynomial does not guarantee diagonalizability.

- **algebraic multiplicity** of an eigenvalue λ :
  a-mul(λ)=k if $f(t) = (t - \lambda)^k g(t)$ and $g(\lambda) \neq 0$.

- **geometric multiplicity** of an eigenvalue λ : g-mul(λ)=dim(E_λ)
  - g-mul(λ)=nullity($T - \lambda I$)
  - identity operator I : a-mul(1)g-mul(1)=n

- Theorem 5.7; $1 \leq$ g-mul(λ) $\leq$ a-mul(λ) .

- example: present an operator with $\lambda = 1$, a-mul(1)$=$ n,g-mul(1)$=$ $n - 1$

- Theorem 5.9: dim(V)$< \infty$; $T : V \to V$ is a linear operator; its characteristic polynomial splits; and $\lambda_1, \cdots, \lambda_k$ are distinct eigenvalues of T. Then
  1. T is diagonalizable $\Leftrightarrow$ a-mul(λ_i)=g-mul(λ_i), $i = 1, \cdots, k$.
  2. T is diagonalizable; β_i is an ordered basis for E_{λ_i} $\Rightarrow \beta = \beta_1 \cup \cdots \cup \beta_k$ is an ordered basis for V.

- characteristic polynomial splits
  $\Rightarrow \sum_{i=1}^{k}$ a-mul(λ_i)$= n =$dim(V)
  $\Rightarrow$ a-mul(λ_i)=g-mul(λ_i),$i = 1, \cdots, k$.

  [End of Review]

- example: $T : \mathbf{R^3} \to \mathbf{R^3}$, $T\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\right) = \begin{pmatrix} -2a_2 - 3a_3 \\ a_1 + 3a_2 + 3a_3 \\ a_3 \end{pmatrix}$;

β is the standard basis. $A = [T]_\beta = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ 0 & 0 & 1 \end{pmatrix}$

$f(t)$=det$\begin{pmatrix} -t & -2 & -3 \\ 1 & 3-t & 3 \\ 0 & 0 & 1-t \end{pmatrix} = -t(3-t)(1-t) + 2(1-t)$

$= -(t-1)^2(t-2)$

$\lambda = 1 \Rightarrow (A - 1I)v = \begin{pmatrix} -1 & -2 & -3 \\ 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix} v = 0$, rank($A - 1I$)=1

$E_1 = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} : v1 + 2v_2 + 3v_3 = 0 \right\}, \beta_1 = \left\{ \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \right\}$

$\Rightarrow$ a-mul(1)=g-mul(1)=2

$\lambda = 2 \Rightarrow (A-2I)v = \begin{pmatrix} -2 & -2 & -3 \\ 1 & 1 & 3 \\ 0 & 0 & -1 \end{pmatrix} v = 0$, rank($A-2I$)=2

$$E_2 = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} : v1 + v_2 = 0 \right\}, \beta_2 = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$$

$\Rightarrow$ a-mul(2)=g-mul(2)=1

Let $\beta' = \beta_1 \cup \beta_2$.

$$Q = [I]_{\beta'}^{\beta} = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, Q^{-1} = [I]_{\beta}^{\beta'} = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 0 & -1 \\ -1 & -2 & -3 \end{pmatrix}$$

$$Q^{-1}AQ = [I]_{\beta}^{\beta'}[T]_{\beta}[I]_{\beta'}^{\beta} = Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = D = [T]_{\beta'}$$

- Diagonalization can be used to efficiently compute the nth power of a matrix.
  $A = QDQ^{-1}$
  $\Rightarrow A^m = QDQ^{-1}QDQ^{-1} \cdots QDQ^{-1} = QD^mQ^{-1}$
  $$D^m = \begin{pmatrix} \lambda_1^m & & O \\ & \ddots & \\ O & & \lambda_k^m \end{pmatrix}$$

- example: system of linear differential equations:
  $x_1' = -2x_2 - 3x_3$
  $x_2' = x_1 + 3x_2 + 3x_3$
  $x_3' = x3$
  $$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, A = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ 0 & 0 & 1 \end{pmatrix}, x' = Ax$$
  $x' = QDQ^{-1}x \Rightarrow y' = Dy$,where $y = Q^{-1}x$.

$\Rightarrow y_1' = y_1, y_2' = y_2, y_3' = 2y_3$ [prev example]

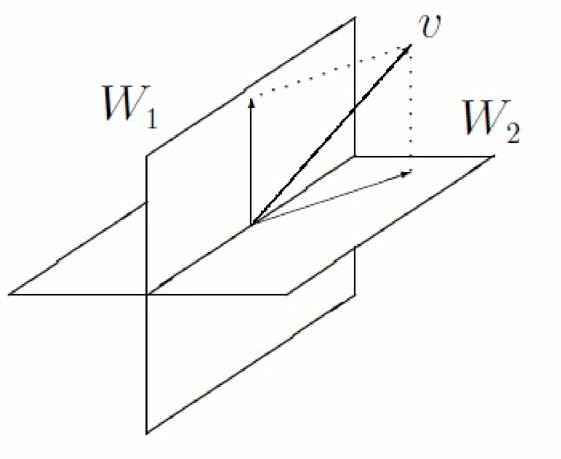
$\Rightarrow y_1 = c_1 e^t, y_2 = c_2 e^t, y_3 = c_3 e^{2t}$

$$x = Qy = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} (2c_1 + 3c_2)e^t + c_3 e^{2t} \\ -c_1 e^t - c_3 e^{2t} \\ -c_2 e^t \end{pmatrix}$$

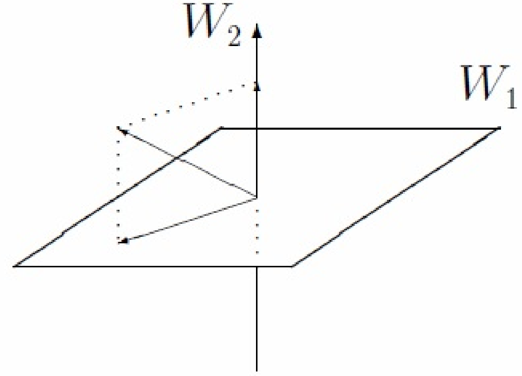
$$= e^t \left[c_1 \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \right] + e^{2t} \left[c_3 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right]$$

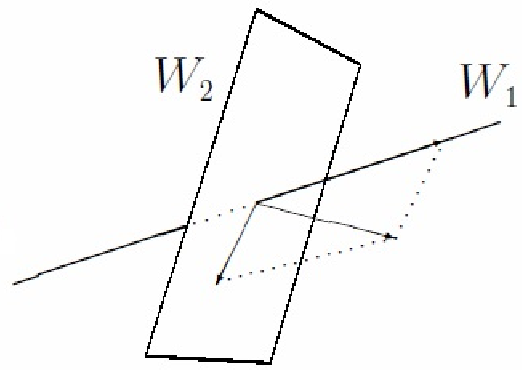
- **sum** of subspaces: $\sum_{i=1}^{k} W_i = \{\sum_{i=1}^{k} v_i : v_i \in W_i, i = 1, \cdots, k\}$
  - $W_1 + W_2 = \{v_1 + v_2 : v_1 \in W_1, v_2 \in W_2\}$
  - **In $\mathbf{R^3}$**,
    $W_1 = \{(a, b, 0) : a, b \in \mathbf{R}\}$
    $W_2 = \{(a, 0, c) : a, c \in \mathbf{R}\}$
    $W_1 + W_2 = \mathbf{R^3}$

$W_1 \cap W_2 = \{(a, 0, 0) : b \in \mathbf{R}\}$
$W_1 \cup W_2 = \{(a, b, c) : a, b, c \in \mathbf{R}, b = 0 \text{ or } c = 0\}$
For a $v \in W_1 + W_2$ there are many pairs $(v_1, v_2), v_1 \in W_1, v_2 \in W_2$ such that $v = v_1 + v_2$. Imagine any plane that include v; it determines v_1 and v_2.
When will v_1 and v_2 be unique?



- **direct sum** of subspaces: $\sum_{i=1}^{k} W_i$ when $W_j \cap \sum_{i \neq j} W_i = \{0\}, j = 1, \cdots, k\}$
  - $W_1 \oplus W_2 = \{v_1 + v_2 : v_1 \in W_1, v_2 \in W_2\}, W_1 \cap W_2 = \{0\}$
  - In $\mathbf{R^3}$,
    $W_1 = \{(a, b, 0) : a, b \in \mathbf{R}\}$
    $W_2 = \{(0, 0, c) : c \in \mathbf{R}\}$
    $W_1 \oplus W_2 = \mathbf{R^3}, W_1 \cap W_2 = \{0\}$
    For $v \in W_1 + W_2 = W_1 \oplus W_2 = \mathbf{R^3}$ the pairs $(v_1, v_2), v_1 \in W_1, v_2 \in W_2$ such that $v = v_1 + v_2$ is unique





  - $\beta = \{v_1, \cdots, v_n\}$, W_1=span($\{v_1, v_2\}$), W_2=span($\{v_3, \cdots, v_n\}$)

- Theorem 5.10: $\dim(V) < \infty$; $W_1, \cdots, W_k$ are subspaces of V. Then the following are equivalent.
  1. $V = W_1 \oplus \cdots \oplus W_k$
  2. $V = \sum_{i=1}^{k} W_i$; and if $v_1 + \cdots + v_k = 0$ for $v_i \in W_i$, $i = 1, \cdots, k$ then $v_i = 0, i = 1, \cdots, k$.
  3. $\forall v \in V, v = v_1 + \cdots + v_k$,where $v_i \in W_i, i = 1, \cdots, k$ are unique.
  4. If γ_i is an ordered basis for $W_i, i = 1, \cdots, k$ then $\gamma_1 \cup \cdots \cup \gamma_k$ is an ordered basis for V.

  proof: (i) Assume 1 and prove 2.
  Let $v_1 + \cdots + v_k = 0$ for $v_i \in W_i, i = 1, \cdots, k$
  $\Rightarrow -v_j = \sum_{i \neq j} v_i, -v_j \in W_j, \sum_{i \neq j} v_i \in \sum_{i \neq j} W_i$
  $\Rightarrow -v_j \in W_j \cap \sum_{i \neq j} W_i = \{0\}$ [1]
  $\Rightarrow -v_j = 0$, Repeat this for $j = 1, \cdots, k$.
  (ii) Assume 2 and prove 3.

Let $v = v_1 + \cdots + v_k = w_1 + \cdots + w_k$, where $v_i, w_i \in W_i$
$i = 1, \cdots, k$
$\Rightarrow (v_1 - w_1) + \cdots + (v_k - w_k) = 0, v_i - w_i \in W_i, i = 1, \cdots, k$
$\Rightarrow v_i - w_i = 0, i = 1, \cdots, k$ [2]
(iii) Assume 3 and prove 4.
Let $\gamma_i = \{v_{ij}, j = 1, \cdots, m_i\}$ be an ordered basis for W_i,
$i = 1, \cdots, k$,
$\Rightarrow \forall v \in V, v = v_1 + \cdots + v_k = \sum_{i=1}^{k} \sum_{j=1}^{m_i} a_{ij} v_{ij}$,for some a_{ij}
[3]
$\Rightarrow \beta = \gamma_1 \cup \cdots \cup \gamma_k$ generates V. We now show lin indep
Let $\sum_{i=1}^{k} \sum_{j=1}^{m_i} a_{ij} v_{ij} = 0$ for scalars a_{ij} and let $w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}$.
$\Rightarrow w_1 + \cdots + w_k = 0, w_i \in W_i, i = 1, \cdots, k$
$\Rightarrow w_i = 0, i = 1, \cdots, k[0 + \cdots + 0 = 0]$ and 3: uniqueness]
$\Rightarrow a_{ij} = 0, i = 1, \cdots, k; j = 1, \cdots, m_i[\gamma_i$ is a basis]
$\Rightarrow v_{ij}, i = 1, \cdots, k; j = 1, \cdots, m_i$ are linearly independent.

(iv) Assume 4 and prove 1.

Let γ_i be and ordered basis for $W_i, i = 1, \cdots, k$.

$\Rightarrow V$ =span($\gamma_1 \cup \cdots \cup \gamma_k$) [4]

=span(γ_1)+ $\cdots$ +span(γ_1)= $W_1 + \cdots + W_k$

$[v \in$ span($\gamma_1 \cup \cdots \cup \gamma_k$) $\Leftrightarrow v = \sum_{i=1}^{k}(\sum_{j=1}^{m_i} a_{ij}v_{ij})]$

Let $v \in W_l \cap \sum_{i \neq l} W_i$

$\Rightarrow v \in$ span(γ_l) $\cap$ span($\cap_{i \neq l}\gamma_i$)

$\Rightarrow v \in$ span(γ_l) and $v \in$span($\cap_{i \neq l}\gamma_i$)

$\Rightarrow v = \sum_{j=1}^{m_i} a_{lj}v_{lj}$ and $v = \sum_{i=1,i \neq l}^{k}(\sum_{j=1}^{m_i} a_{ij}v_{ij})]$

$\Rightarrow (0, \cdots, 0, a_{l1}, \cdots, a_{lm_l}, 0, \cdots, 0)$

$= (a_{11}, \cdots, a_{(l-1)m_{l-1}}, 0, \cdots, 0, a_{1m_{l+1}}, \cdots, a_{km_k})$ [uniq rep]

$\Rightarrow v = 0 \Rightarrow W_l \cap \sum_{i \neq l} W_i = 0$

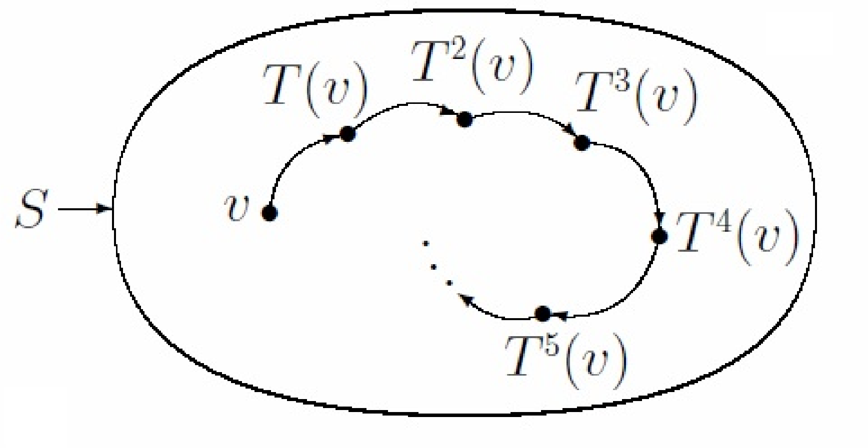
- dim(V)=dim(W_1)$+\cdots+$dim(W_k)
- If $V = W_1 \oplus W_2$, we say W_2 is a **compliment** of W_1, and vice versa.
- So if $V = W_1 \oplus \cdots \oplus W_k$, $\sum_{i \neq j} W_i$ is a compliment of W_j.

- Theorem 5.11: dim(V) $< \infty$. Then
$T : V \to V$ is diagonalizable.
$\Leftrightarrow V$ is the direct sum of eigenspaces of T.

Proof: By Theorem 5.9, T is diagonalizable; β_i is an ordered basis for $E_{\lambda_i} \Rightarrow \beta = \beta_1 \cup \cdots \cup \beta_k$ is an ordered basis for V. Then the theorem is proved by 4. of Theorem 5.10.

Invariant subspace and Cayley-Hamilton theorem

- Recall that a set S is said to be T-invariant if $T(S) \subseteq S$.
  - W is a T-invariant subspace if W is a subspace and also T-invariant.
  - $\{0\}, N(T), R(T), E_\lambda$, and V are T-invariant subspaces.
- T-orbit of v: T-orb(v)$= \{v, T(v), T^2(v), \cdots\}$
  - T-orb(v) is a T-invariant set.
  - W is a T-invariant subspace, and $v \in W$.
    $\Rightarrow T$-orb$(v) \subseteq W$ [T-inv]
    $\Rightarrow$ span(T-orb(v))$\subseteq W$ [subspace]
    $\Rightarrow$ span(T-orb(v)) is the ”smallest” T-invariant subspace containing v.



- example:
  $T : P(\mathbf{R}) \to P(\mathbf{R}), T(f) = f'$
  T-orb(x^2)= $\{x^2, 2x, 2, 0, 0, \cdots\}$
  span(T-orb(x^2))=span($\{x^2, 2x, 2, 0, 0, \cdots\}$)=$P_2$($\mathbf{R}$)

- Recall that given a function $f : X \to Y$, the **restriction** of f to S is $f_s : S \to Y$ such that $\forall x \in S, f_s(x) = f(x)$.

  - $T : V \to V$ is a linear operator; W is a T-invariant subspace. Then T_w is a linear operator on W, ie, $T_w : W \to W$.

■ Theorem 5.21: dim(V)$< \infty$; $T; V \to V$ is a linear operator; W is a T-invariant subspace of V. Then the characteristic polynomial $g(t)$ of T_w divides the characteristic polynomial $f(t)$ of T.

proof: Let $\gamma = \{v_1, \cdots, v_k\}$ be a basis for W.

$\Rightarrow \beta = \{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ is a basis for V. [extension]

Let $A = [T]_\beta$ and $B_1 = [T_W]_\gamma$

$\Rightarrow A = ([T(v_1)]_\beta, \cdots, [T(v_k)]_\beta, \cdots, [T(v_n)]_\beta)$ $[n \times n]$

$B_1 = ([T_W(v_1)]_\gamma, \cdots, [T_W(v_k)]_\gamma)$ $[k \times k]$

$$v_i \in W, i = 1, \cdots, k \Rightarrow [T(v_i)]_\beta = [T_W(v_i)]_\beta = \begin{pmatrix} [T_W(v_i)]_\gamma \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where there are $n - k$ trailing zeros.

$$\Rightarrow A = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix} \text{ with size } \begin{pmatrix} k \times k & k \times (n-k) \\ (n-k) \times k & (n-k) \times (n-k) \end{pmatrix}$$

$$\Rightarrow f(t) = \det(A - tI_n) = \det\begin{pmatrix} B_1 - tI_k & B_2 \\ O & B_3 - tI_{n-k} \end{pmatrix}$$
$$= \det(B_1 - tI_k)\det(B_3 - tI_{n-k}) = g(t)\det(B_3 - tI_{n-k})$$

- So each T-invariant subspace E_λ corresponds to a factor $g_\lambda(t)$ of $f(t)$.

- dim(E_λ)=g-mul(λ)$=$nullity$(T-\lambda I)$$=n-$rank$(T-\lambda I)$
- Diagonalization can be used to efficiently compute the nth power of a matrix.
  $A = QDQ^{-1}$
  $\Rightarrow A^m = QDQ^{-1}QDQ^{-1}\cdots QDQ^{-1} = QD^mQ^{-1}$
  $D^m = \begin{pmatrix} \lambda_1^m & & O \\ & \ddots & \\ O & & \lambda_k^m \end{pmatrix}$
- **sum** of subspaces: $\sum_{i=1}^k W_i = \{\sum_{i=1}^k v_i : v_i \in W_i, i = 1, \cdots, k\}$
- **direct sum** of subspaces: $\sum_{i=1}^k W_i$ when $W_j \cap \sum_{i\neq j} W_i = \{0\}, j = 1, \cdots, k\}$; For example,
  $W_1 \oplus W_2 = \{v_1 + v_2 : v_1 \in W_1, v_2 \in W_2\}, W_1 \cap W_2 = \{0\}$

- Theorem 5.10: $\dim(V) < \infty$; $W_1, \cdots, W_k$ are subspaces of V. Then the following are equivalent.
  1. $V = W_1 \oplus \cdots \oplus W_k$
  2. $V = \sum_{i=1}^{k} W_i$; and if $v_1 + \cdots + v_k = 0$ for $v_i \in W_i$, $i = 1, \cdots, k$ then $v_i = 0, i = 1, \cdots, k$.
  3. $\forall v \in V, v = v_1 + \cdots + v_k$,where $v_i \in W_i, i = 1, \cdots, k$ are unique.
  4. If γ_i is an ordered basis for $W_i, i = 1, \cdots, k$ then $\gamma_1 \cup \cdots \cup \gamma_k$ is an ordered basis for V.

- dim(V)=dim(W_1)+ $\cdots$ +dim(W_k)

- If $V = W_1 \oplus W_2$, we say W_2 is a **compliment** of W_1, and vice versa.

- So if $V = W_1 \oplus \cdots \oplus W_k$, $\sum_{i \neq j} W_i$ is a compliment of W_j.

- Theorem 5.11: dim(V) $< \infty$. Then
  $T : V \to V$ is diagonalizable.
  $\Leftrightarrow$ V is the direct sum of eigenspaces of T.

- T-orbit of v: T-orb(v)$= \{v, T(v), T^2(v), \cdots\}$

- T-orb(v) is a T-invariant set.

- W is a T-invariant subspace, and $v \in W$.
  $\Rightarrow$ T-orb$(v) \subseteq W$ [T-inv]
  $\Rightarrow$ span(T-orb(v))$\subseteq W$ [subspace]
  $\Rightarrow$ span(T-orb(v)) is the "smallest" T-invariant subspace containing v.

- For given a function $f : X \to Y$, the **restriction** of f to S is $f_s : S \to Y$ such that $\forall x \in S, f_s(x) = f(x)$.

- Theorem 5.21: dim(V)$< \infty$; $T; V \to V$ is a linear operator; W is a T-invariant subspace of V. Then the characteristic polynomial $g(t)$ of T_w divides the characteristic polynomial $f(t)$ of T.
proof: Let $\gamma = \{v_1, \cdots, v_k\}$ be a basis for W.
$\Rightarrow \beta = \{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ is a basis for V.
$\Rightarrow A = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix}$ with size $\begin{pmatrix} k \times k & k \times (n-k) \\ (n-k) \times k & (n-k) \times (n-k) \end{pmatrix}$
$\Rightarrow f(t) =$det$(A - tI_n)$=det$\begin{pmatrix} B_1 - tI_k & B_2 \\ O & B_3 - tI_{n-k} \end{pmatrix}$
=det$(B_1 - tI_k)$det$(B_3 - tI_{n-k})$$= g(t)det(B_3 - tI_{n-k})$

- Each T-invariant subspace E_λ corresponds to a factor $g_\lambda(t)$ of $f(t)$.

- [End of Review]

- Theorem 5.22: dim(V)$< \infty$; $T : V \to V$ is a linear operator; $v \neq 0$;W =span(T-orb(v)); dim(W)= k. Then
  1. $\{v, T(v), T^2(v), \cdots, T^{k-1}(v)\}$ is a basis for W.
  2. $a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$
  $\Rightarrow$The characteristic polynomial of T_W is
  $g(t) = (-1)^k (a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$.

  - Express $T^k(v) = -(a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v))$.
    proof: "1": $v \neq 0 \Rightarrow \{v\}$ is linearly independent.
    Find the maximum $j \geq 1$ such that $\beta = \{v, T(v), \cdots, T^{j-1}(v)\}$ is linearly independent, and let Z =span(β).
    $\Rightarrow$ β is a basis for Z.[lin indep, generating]
    $\Rightarrow \{v, T(v), \cdots, T^{j-1}(v), T^j(v)\}$ is linearly dependent.
    $\Rightarrow T^j(v) \in Z$
    We will show that $Z = W$, ie, that none of $\beta \cup \{T^i(v)\}, i \geq j$ are linearly independent.

(i) $\beta \subseteq T$-orb(v)$\Rightarrow$ span(β) $\subseteq$ span(T-orb(v))$\Rightarrow$ $Z \subseteq W$

(ii) Since W =span(T-orb(v)) is the smallest T-invariant subspace containing v, we simply show that Z is T-invariant to show $W \subseteq Z$.

$w \in Z \Rightarrow w = b_0 v + b_1 T(v) + \cdots + b_{j-1} T^{j-1}(v)$

$\Rightarrow T(w) = b_0 T(v) + b_1 T^2(v) + \cdots + b_{j-1} T^j(v)$

$\Rightarrow T(w) \in Z$ [$T^j(v) \in Z$]

$\Rightarrow T(Z) \subseteq Z$, ie, Z is T-invariant. [$j = k$]

"2": Let $a_0 v + a_1 T(v) + \cdots a_{k-1} T^{k-1}(v) + T^k(v) = 0$.

$[T_w]_\beta = ([T_W(v)]_\beta, [T_W(T(v))]_\beta, \cdots, [T_W(T^{k-1}(v))_\beta])$

$= ([T(v)]_\beta, [T^2(v))]_\beta, \cdots, [T^{k-1}(v))]_\beta, [T^k(v)]_\beta)$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix} \quad [\beta = \{v, T(v), \cdots, T^{k-1}(v)\}]$$

$g(t)$ =det($T_W - tI$)=det($[T_W]_\beta - tI_k$)

$$=\det\begin{pmatrix} -t & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & -t & \cdots & 0 & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -t & -a_{k-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{k-1} - t \end{pmatrix}$$

We use induction in k to evaluate the determinant.

(i) If $k = 1, g(t)$ =det($-a_0 - t$)= $(-1)^k(a_0 + t)$

(ii) Assume for dim(W)= $k - 1$,

$g(t) = (-1)^{k-1}(a_0 + a_1 t + \cdots + a_{k-2}t^{k-2} + t^{k-1})$.

(iii) For dim(W)= k, expand the determinant along the first row.

$$g(t) = (-1)^{1+1}(-t)\mathbf{det}\begin{pmatrix} -t & 0 & \cdots & 0 & 0 & -a_1 \\ 1 & -t & \cdots & 0 & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -t & -a_{k-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{k-1} - t \end{pmatrix}$$

$$+(-1)^{1+k}(-a_0)\mathbf{det}\begin{pmatrix} 1 & -t & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -t \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$= (-t)(-1)^{k-1}(a_1 + a_2 t + \cdots + a_{k-1}t^{k-2} + t^{k-1}) \text{ [(ii)]}$$

$$+(-1)^{1+k}(-a_0)$$

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

- Theorem 5.23(Cayley-Hamilton): dim(V)$< \infty$; $T : V \to V$ is a linear operator; $f(t)$ is the characteristic polynomial of T.
Then $f(T) = T_0$, the zero transformation.

  - For example, $T : P_2(\mathbf{R}) \to P_2(\mathbf{R}), T(g) = g', \beta = \{1, x, x^2\}$
  $$[T]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, f(t) = (-t)^3, -T^3(a_0 + a_1 x + a_2 x^2) = 0$$

proof: We show that $\forall v \in V, f(T)(v) = 0$.
(i) If $v = 0, f(T)(0) = 0$.
(ii) Assume $v \neq 0$. Find the maximum k
such that $\beta = \{v, T(v), \cdots, T^{k-1}(v)\}$ is linearly independent.
Let W =span(β)=span(T-orb(v)). [Thm 5.22]
Let $T^k(v) = -a_0 v - a_1 T(v) - \cdots - a_{k-1} T^{k-1}(v)$ [lin comb]
and $g(t)$ be the characteristic polynomial of T_W.

$\Rightarrow g(t) = (-1)^k(a_0 + a_1 t + \cdots + a_{k-1}t^{k-1} + t^k)$ [Thm 5.22]
$\Rightarrow g(T)(v) = (-1)^k(a_0 I + a_1 T + \cdots + a_{k-1}T^{k-1} + T^k)(v)$
$= (-1)^k(a_0 v + a_1 T(v) + \cdots + a_{k-1}T^{k-1}(v) + T^k(v)) = 0$
$f(t) = h(t)g(t)$ [W is T-invariant; Thm 5.21]
$\Rightarrow f(T)(v) = (h(T)g(T))(v) = h(T)(g(T)(v))$
$= h(T)(0) = 0$

- example: $V = \mathbf{R^3}; T : \mathbf{R^3} \to \mathbf{R^3}$
  $T((a, b, c)) = (-b + c, a + c, 3c)$
  Let's begin an orbit with $e_1 = (1, 0, 0)$.
  $T(e_1) = (0, 1, 0), T^2(e_1) = (-1, 0, 0), T^3(e_1) = (0, -1, 0)$
  $\Rightarrow \{e_1, T(e_1)\} = \{e_1, e_2\}$ is linearly independent but
  $\{e_1, T(e_1), T^2(e_1)\}$ is not.
  $\Rightarrow W$ =span(T-orb(e_1))=span($\{e_1, e_2\}$)
  $\Rightarrow \beta = \{e_1, e_2\}$
  $T^2(e_1) = -e_1 = -a_0 e_1 - a_1 T(e_1)$

$\Rightarrow a_0 = 1, a_1 = 0$

$\Rightarrow g(t) = (-1)^2(a_0 + a_1 t + t^2) = 1 + t^2$

Let's confirm it.

$[T_W]_\beta = ([T_W(e_1)]_\beta, [T_W(e_2)]_\beta) = ([e_2]_\beta, [-e_1]_\beta) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

$\Rightarrow g(t) =$det$\begin{pmatrix} -t & -1 \\ 1 & -t \end{pmatrix} = t^2 + 1$

β is extended to the standard basis $\gamma = \{e_1, e_2, e_3\}$ for $\mathbf{R^3}$.

$\Rightarrow A = [T]_\gamma = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 3 \end{pmatrix} \Rightarrow f(t) =det\begin{pmatrix} -t & -1 & 1 \\ 1 & -t & 1 \\ 0 & 0 & 3-t \end{pmatrix}$

$\Rightarrow f(t) = -(t^2 + 1)(t - 3)$, so $g(t)$ divides $f(t)$.

$\Rightarrow f(T) = -(T^2 + I)(T - 3I), g(T) = T^2 + I$

$A = [T]_\gamma = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 3 \end{pmatrix} A^2 = [T]_\gamma^2 = [T^2]_\gamma = \begin{pmatrix} -1 & 0 & 2 \\ 0 & -1 & 4 \\ 0 & 0 & 9 \end{pmatrix}$

$[f(T)]_\gamma = [-(T^2+I)(T-3I)]_\gamma = -(A^2+I)(A-3I)$

$$= -\begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 10 \end{pmatrix} \begin{pmatrix} -3 & -1 & 1 \\ 1 & -3 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = [T_0]_\gamma$$

$\Rightarrow f(T) = T_0$ confirms the Cayley-Hamilton theorem.

[Uniqueness of representation]

Note also that $\forall v \in W, g(T)(v) = (T^2+I)(v) = 0$ because

$$[g(T)]_\gamma = [(T^2+I)]_\gamma = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 10 \end{pmatrix}$$

$\Rightarrow g(T_W) = T_0$

Let's begin an orbit with another vector $v = (1,1,1)$.

$T(v) = (0,2,3), T^2(v) = (1,3,9), T^3(v) = (6,10,27)$

$\Rightarrow \{v, T(v), T^2(v)\}$ is linearly independent but

$\{v, T(v), T^2(v), T^3(v)\}$ is not.

$\Rightarrow W$ =span(T-orb(v))$= \mathbf{R^3} = V$

$\Rightarrow \beta = \{v, T(v), T^2(v)\}$

$T^3(v) = 3v - T(v) + 3T^2(v) = -a_0 v - a_1 T(v) - a_2 T^2(v)$

$\Rightarrow a_0 = -3, a_1 = 1, a_2 = -3$

$\Rightarrow g(t) = (-1)^3(a_0 + a_1 t + a_2 t^2 + t^3) = 3 - t + 3t^2 - t^3 = f(t)$

So the starting vector v of an orbit

determines the resulting T-invariant subspace W =span(T-orb(v))

and $g(t)$

W_1 =span($\{e_1, e_2\}$) $\Rightarrow g_1(t) = t^2 + 1$

$W_2 = \mathbf{R^3} \Rightarrow g_2(t) = 3 - t + 3t^2 - t^3$

If we begin with (1,2,5), the eigenvector corresponding to $\lambda = 3$

the

W_3 =span($\{$(1,2,5)$\}$)$= E_3 \Rightarrow g_3(t) = 3 - t$

- Corollary 5.23 (Cayley-Hamilton theorem for matrices): $A \in M_{n\times n}$; $f(t)$ is the characteristic polynomial of A Then $f(A) = O$, the $n \times n$ zero matrix.

  - Computation of a matrix polynomial $p(A)$ of a high degree: $p(t) = f(t)q(t) + r(t) \Rightarrow p(A) = f(A)q(A) + r(A) = r(A)$, where $r(t)$ is a polynomial of a much lower degree.

  - If A is invertible,
    $f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots a_1 t + a_0$
    $\Rightarrow f(A) = a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = O$
    $\Rightarrow f(A)A^{-1} = a_n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I + a_0 A^{-1} = O$
    $\Rightarrow A^{-1} = -\frac{1}{a_0}(a_n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I)$
    where $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1}tr(A)$, and $a_0 =$det(A)

- Theorem 5.24: dim(V)$< \infty$;
  $T : V \to V$ is a linear operator;
  $f(t)$ is the characteristic polynomial of T;
  W_i is a T-invariant subspace, $i = 1, \cdots, k$;
  $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$;
  $f_i(T)$ is the characteristic polynomial of T_{W_i} Then
  $f(t) = f_1(t) f_2(t) \cdots f_k(t)$
  - Is the converse true? From $f_i(t)$ to T_{W_i} ?

- Theorem 5.25: In addition to the above,
  β_i is a basis for W_i;
  $\beta = \beta_1 \cup \cdots \cup \beta_k$ is a basis for V. [Thm 5.10]

$$\Rightarrow [T]_\beta = \begin{pmatrix} [T_{W_1}]_{\beta_1} & O_{12} & \cdots & O_{1k} \\ O_{21} & [T_{W_2}]_{\beta_2} & \cdots & O_{2k} \\ \vdots & \vdots & & \vdots \\ O_{k1} & O_{k2} & \cdots & [T_{W_k}]_{\beta_k} \end{pmatrix}$$

- example: $T : \mathbf{R^4} \to \mathbf{R^4}$ is defined by
  $T(a,b,c,d) = (2a-b, a+b, c-d, c+d)$.
  $W_1 = \{(a,b,0,0) : a,b \in \mathbf{R}\}$ $W_2 = \{(0,0,c,d) : c,d \in \mathbf{R}\}$
  $\Rightarrow W_1$ and W_2 are T-invariant, and $\mathbf{R^4} = W_1 \oplus W_2$.
  $\Rightarrow T_{W_1} : W_1 \to W_1$ is such that $T(a,b,0,0) = (2a-b, a+b, 0, 0)$
  $T_{W_2} : W_2 \to W_2$ is such that $T(0,0,c,d) = (0,0,c-d,c+d)$
  Let $\beta_1 = \{e_1, e_2\}$, $\beta_2 = \{e_3, e_4\}$ and $\beta = \beta_1 \cup \beta_2$
  $\Rightarrow [T_{W_1}]_{\beta_1} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$ $[T_{W_2}]_{\beta_2} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

$$\Rightarrow [T_{W_1}]_{\beta_1} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} [T_{W_1}]_{\beta_1} & O_{12} \\ O_{21} & [T_{W_2}]_{\beta_2} \end{pmatrix}$$

$$\Rightarrow f(t) = \textbf{det}\begin{pmatrix} 2-t & -1 & 0 & 0 \\ 1 & 1-t & 0 & 0 \\ 0 & 0 & 1-t & -1 \\ 0 & 0 & 1 & 1-t \end{pmatrix}$$

$$= \textbf{det}\begin{pmatrix} 2-t & -1 \\ 1 & 1-t \end{pmatrix} \textbf{det}\begin{pmatrix} 1-t & -1 \\ 1 & 1-t \end{pmatrix} = f_1(t) f_2(t)$$