Proof. 依 $\mu_T(x)$ 的定義, 對任意 $\mathbf{v}\in V$, 皆有 $\mu_T(T)(\mathbf{v})=\mathbf{O}_V\in W$, 得 $\overline{\mu_T(T)(\mathbf{v})}=\overline{\mathbf{O}_V}=\mathbf{O}_{V/W}$. 故由 Lemma 4.4.7 知

$$\mu_T(\overline{T})(\overline{\mathbf{v}})=\overline{\mu_T(T)}(\overline{\mathbf{v}})=\overline{\mu_T(T)(\mathbf{v})}=\mathbf{O}_{V/W}.$$

利用 Lemma 3.3.5 (套用在 $\overline{T}$) 得 $\mu_{\overline{T}}(x)\mid\mu_T(x)$.

同理, 因 $\mu_{\mathbf{v}}(T)(\mathbf{v})=\mathbf{O}_V$, 我們得 $\mu_{\mathbf{v}}(\overline{T})(\overline{\mathbf{v}})=\mathbf{O}_{V/W}$, 故由 Lemma 4.4.3 (套用在 $\overline{\mathbf{v}}$ 以及 $\overline{T}$) 得 $\mu_{\overline{\mathbf{v}}}(x)\mid\mu_{\mathbf{v}}(x)$. □

事實上在某些情況之下有可能 $\mu_{\overline{\mathbf{v}}}(x)=\mu_{\mathbf{v}}(x)$, 例如以下的情況.

Lemma 4.4.9. 令 $T:V\to V$ 為一個 *F-linear operator.* 給定 $\mathbf{v}\in V$, 考慮 $\overline{T}:V/C_{\mathbf{v}}\to V/C_{\mathbf{v}}$ 為 *linear operator induced by* T *on* $V/C_{\mathbf{v}}$. 若 $\mathbf{w}\in V$ 滿足 $\mu_{\mathbf{w}}(x)\mid\mu_{\mathbf{v}}(x)$, 則存在 $\mathbf{u}\in V$ 滿足 $\overline{\mathbf{u}}=\overline{\mathbf{w}}\in V/C_{\mathbf{v}}$ 且 $\mu_{\mathbf{u}}(x)=\mu_{\overline{\mathbf{u}}}(x)=\mu_{\overline{\mathbf{w}}}(x)$.

Proof. 因 $\mu_{\overline{\mathbf{w}}}(\overline{T})(\overline{\mathbf{w}})=\mathbf{O}_{V/C_{\mathbf{v}}}$, 利用 Lemma 4.4.7 得 $\overline{\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})}=\overline{\mathbf{O}_V}$, 亦即 $\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})\in C_{\mathbf{v}}$. 換言之, 存在 $f(x)\in F[x]$ 使得

$$\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})=f(T)(\mathbf{v}). \tag{4.4}$$

依 $\mu_{\mathbf{w}}(x)\mid\mu_{\mathbf{v}}(x)$ 之假設, 以及由 Corollary 4.4.8 知 $\mu_{\overline{\mathbf{w}}}(x)\mid\mu_{\mathbf{w}}(x)$ 可得 $\mu_{\overline{\mathbf{w}}}(x)\mid\mu_{\mathbf{v}}(x)$, 亦即存在 $h(x)\in F[x]$ 使得 $\mu_{\mathbf{v}}(x)=h(x)\mu_{\overline{\mathbf{w}}}(x)$. 故由等式 (4.4) 得

$$\mu_{\mathbf{v}}(T)(\mathbf{w})=h(T)\circ\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})=h(T)\circ f(T)(\mathbf{v}). \tag{4.5}$$

然而 $\mu_{\mathbf{w}}(x)\mid\mu_{\mathbf{v}}(x)$, 故由 Lemma 4.4.3 與等式 (4.5) 知 $\mathbf{O}_V=\mu_{\mathbf{v}}(T)(\mathbf{w})=h(T)\circ f(T)(\mathbf{v})$. 再次利用 Lemma 4.4.3 得 $\mu_{\mathbf{v}}(x)\mid h(x)f(x)$, 亦即 $h(x)\mu_{\overline{\mathbf{w}}}(x)\mid h(x)f(x)$. 由此知 $\mu_{\overline{\mathbf{w}}}(x)\mid f(x)$, 亦即存在 $g(x)\in F[x]$ 使得

$$f(x)=\mu_{\overline{\mathbf{w}}}(x)g(x). \tag{4.6}$$

現令 $\mathbf{u}=\mathbf{w}-g(T)(\mathbf{v})$. 因 $g(T)(\mathbf{v})\in C_{\mathbf{v}}$, 我們有 $\overline{\mathbf{u}}=\overline{\mathbf{w}}\in V/C_{\mathbf{v}}$. 利用 $\mu_{\overline{\mathbf{w}}}(T)$ 為 linear operator 得

$$\mu_{\overline{\mathbf{w}}}(T)(\mathbf{u})=\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w}-g(T)(\mathbf{v}))=\mu_{\overline{\mathbf{w}}}(T)(\mathbf{w})-\mu_{\overline{\mathbf{w}}}(T)\circ g(T)(\mathbf{v}),$$

所以由等式 (4.6) 以及等式 (4.4) 得

$$\mu_{\overline{\mathbf{w}}}(T)(\mathbf{u})=\mathbf{O}_V.$$

再次利用 Lemma 4.4.3 得 $\mu_{\mathbf{u}}(x)\mid\mu_{\overline{\mathbf{w}}}(x)$. 然而 $\overline{\mathbf{u}}=\overline{\mathbf{w}}$, 故 $\mu_{\overline{\mathbf{w}}}(x)=\mu_{\overline{\mathbf{u}}}(x)$, 即 $\mu_{\mathbf{u}}(x)\mid\mu_{\overline{\mathbf{u}}}(x)$. 再加上 Lemma 4.4.8 告訴我們 $\mu_{\overline{\mathbf{u}}}(x)\mid\mu_{\mathbf{u}}(x)$, 得證 $\mu_{\mathbf{u}}(x)=\mu_{\overline{\mathbf{u}}}(x)$. □

一般來說若 $\deg(\mu_{\overline{\mathbf{w}}}(x))=d$, 雖然 $\{\overline{\mathbf{w}},\overline{T}(\overline{\mathbf{w}}),\dots,\overline{T}^{\circ d-1}(\overline{\mathbf{w}})\}$ 會是 $C_{\overline{\mathbf{w}}}$ 的一組 basis, 不過 $\{\mathbf{w},T(\mathbf{w}),\dots,T^{\circ d-1}(\mathbf{w})\}$ 就未必會是 $C_{\mathbf{w}}$ 的一組 basis. 不過在 Lemma 4.4.9 的假設條件下我們可找到 $\mathbf{u}$ 滿足 $\overline{\mathbf{u}}=\overline{\mathbf{w}}$ 且 $\{\mathbf{u},T(\mathbf{u}),\dots,T^{\circ d-1}(\mathbf{u})\}$ 和 $\{\overline{\mathbf{u}},\overline{T}(\overline{\mathbf{u}}),\dots,\overline{T}^{\circ d-1}(\overline{\mathbf{u}})\}$ 分別會是 $C_{\mathbf{u}}$ 與 $C_{\overline{\mathbf{u}}}=C_{\overline{\mathbf{w}}}$ 的一組 basis.

現在我們可以利用 primary decomposition theorem 證得以下重要的定理.

Theorem 4.4.10 (Cyclic Decomposition Theorem)**.** 假設 V 為 *finite dimensional* F*-space* 且 $T:V\to V$ 為 *linear operator.* 則 V 可以寫成一些 T*-cyclic subspaces* 的 *direct sum.* 事實上, 若 $\mu_T(x)=p_1(x)^{m_1}\cdots p_k(x)^{m_k}$, 其中 $p_i(x)\in F[x]$ 為相異的 *monic irreducible polynomial,* 則 $V=W_1\oplus\cdots\oplus W_k$, 其中

$$W_i=\mathrm{Ker}(p_i(T)^{\circ m_i})=C_{\mathbf{v}_{i,1}}\oplus\cdots\oplus C_{\mathbf{v}_{i,n_i}},$$

而且每個 $\mathbf{v}_{i,j}$ 的 T*-annihilator* 為 $p_i(x)^{m_{i,j}}$ 滿足 $m_i=m_{i,1}\geq m_{i,2}\geq\cdots\geq m_{i,n_i}>0$.

Proof. 由 primary decomposition theorem, 我們知道 $T|_{W_i}:W_i\to W_i$ 的 minimal polynomial 為 $p_i(x)^{m_i}$. 若能證得每一個 W_i 可以寫成定理所述的 T-cyclic subspaces 的 direct sum, 則由 Corollary 3.4.7 可得 V 可以寫成一些 T-cyclic subspaces 的 direct sum. 所以我們僅要證明當 $T:V\to V$ 是 F-linear operator 且 $\chi_T(x)=p(x)^m$ 其中 $p(x)\in F[x]$ 是 monic irreducible polynomial 的情形下, 存在 $\mathbf{v}_1,\ldots\mathbf{v}_n\in V$ 使得 $V=C_{\mathbf{v}_1}\oplus\cdots\oplus C_{\mathbf{v}_n}$ 且對於 $1\leq i\leq n$, $\mu_{\mathbf{v}_i}(x)=p(x)^{m_i}$ 滿足 $m=m_1\geq m_2\geq\cdots\geq m_n$.

我們利用對 $\dim(V)$ 作數學歸納法證明. 當 $\dim(V)=1$ 時, 很自然對於任意 $\mathbf{v}\neq\mathbf{O}_V$ in V, 我們有 $V=C_{\mathbf{v}}$ 且 $\mu_T(x)=\mu_{\mathbf{v}}(x)$, 所以定理成立. 現假設此定理在維度小於 $\dim(V)$ 的情形都成立, 此時依假設 $\mu_T(x)=p(x)^m$, 故存在 $\mathbf{v}_1\in V$ 滿足 $p(T)^{\circ m-1}(\mathbf{v}_1)\neq\mathbf{O}_V$. 因 $\mu_{\mathbf{v}_1}(x)\mid\mu_T(x)=p(x)^m$ 以及 $p(x)$ 為 irreducible, 所以存在 $m_1\leq m$ 使得 $\mu_{\mathbf{v}_1}(x)=p(x)^{m_1}$. 然而若 $m_1\leq m-1$, 由 $\mu_{\mathbf{v}_1}(x)\mid p(x)^{m-1}$, 得 $p(T)^{\circ m-1}(\mathbf{v}_1)=\mathbf{O}_V$. 此和當初 $\mathbf{v}_1$ 的選取相矛盾, 所以 $m_1>m-1$, 因此得 $m_1=m$.

現考慮 $\overline{T}:W/C_{\mathbf{v}_1}\to V/C_{\mathbf{v}_1}$ induced by T on $V/C_{\mathbf{v}_1}$. 注意此時 $\mu_{\overline{T}}(x)\mid\mu_T(x)$ (Corollary 4.4.8), 故知 $\mu_{\overline{T}}(x)=p(x)^{m'}$, 其中 $m'\leq m$. 因此由 $\dim(V/C_{\mathbf{v}_1})<\dim(V)$, 我們可以套用數學歸納法之假設, 即存在 $\mathbf{w}_2,\ldots,\mathbf{w}_n\in V$ 滿足

$$V/C_{\mathbf{v}_1}=C_{\overline{\mathbf{w}_2}}\oplus\cdots\oplus C_{\overline{\mathbf{w}_n}},$$

且對於 $2\leq i\leq n$, $\mu_{\overline{\mathbf{w}_i}}(x)=p(x)^{m_i}$ 滿足 $m\geq m'=m_2\geq\cdots\geq m_n$. 又由於 $m_i\leq m=m_1$, 即 $\mu_{\mathbf{w}_i}(x)\mid\mu_{\mathbf{v}_1}(x)$, 所以利用 Lemma 4.4.9 知, 存在 $\mathbf{v}_i\in V$ 使得 $\overline{\mathbf{v}_i}=\overline{\mathbf{w}_i}$ 且 $\mu_{\mathbf{v}_i}(x)=\mu_{\overline{\mathbf{v}_i}}(x)=p(x)^{m_i}$.

現若 $\deg(p(x))=d$, 由 direct sum 的性質 (Proposition 3.4.6) 以及 Theorem 4.4.4 知

$$\{\overline{\mathbf{v}_2},\overline{T}(\overline{\mathbf{v}_2}),\ldots,\overline{T}^{\circ dm_2-1}(\overline{\mathbf{v}_2}),\ldots,\overline{\mathbf{v}_n},\overline{T}(\overline{\mathbf{v}_n}),\ldots,\overline{T}^{\circ dm_n-1}(\overline{\mathbf{v}_n})\}$$

為 $V/C_{\mathbf{v}_1}=C_{\overline{\mathbf{v}_2}}\oplus\cdots\oplus C_{\overline{\mathbf{v}_n}}$ 的一組 basis. 現因 $\overline{T}^{\circ j}(\overline{\mathbf{v}_i})=\overline{T^{\circ j}(\mathbf{v}_i)}$ (Lemma 4.4.7), 以及 $\{\mathbf{v}_1,T(\mathbf{v}_1),\ldots,T^{\circ dm_1-1}(\mathbf{v}_1)\}$ 為 $C_{\mathbf{v}_1}$ 的一組 basis, 利用 Proposition 1.6.2 的證明所用的方法我們得

$$\{\mathbf{v}_1,T(\mathbf{v}_1),\ldots,T^{\circ dm_1-1}(\mathbf{v}_1),\mathbf{v}_2,T(\mathbf{v}_2),\ldots,T^{\circ dm_2-1}(\mathbf{v}_2),\ldots,\mathbf{v}_n,T(\mathbf{v}_n),\ldots,T^{\circ dm_n-1}(\mathbf{v}_n)\}$$

為 V 的一組 basis. 因為對所有 $1\leq i\leq n$, $\{\mathbf{v}_i,T(\mathbf{v}_i),\ldots,T^{\circ dm_i-1}(\mathbf{v}_i)\}$ 為 $C_{\mathbf{v}_i}$ 的一組 basis, 故由 direct sum 的性質 (Proposition 3.4.6) 得證

$$V=C_{\mathbf{v}_1}\oplus C_{\mathbf{v}_2}\oplus\cdots\oplus C_{\mathbf{v}_n}.$$

□

Question 4.20. 在 *Theorem 4.4.10* 的證明中, 為何要將 $\mathbf{w}_2,\dots,\mathbf{w}_n$ 改成 $\mathbf{v}_2,\dots,\mathbf{v}_n$?

Question 4.21. 可以用 *cyclic decomposition theorem* 說明若 $\mu_T(x)=(x-\lambda_1)\cdots(x-\lambda_k)$, 其中 $\lambda_i\neq\lambda_j$ *for* $i\neq j$, 則 T 是 *diagonalizable* 嗎?

利用 primary decomposition theorem, 我們可以找到 V 的 ordered basis β, 使得 $[T]_\beta$ 為以下的 block diagonal matrix

$$\begin{pmatrix} A_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & A_k \end{pmatrix},$$

其中每個 A_i 的 minimal polynomial 為 $p_i(x)^{m_i}$. 而 cyclic decomposition theorem 告訴我們, β 可以由一些 cyclic vectors 所形成的 cyclic bases 所組成, 此時每一個 A_i 可寫成

$$\begin{pmatrix} C_{i,1} & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & C_{i,n_i} \end{pmatrix},$$

其中每個 $C_{i,j}$ 是 the companion matrix of $p_i(x)^{m_{i,j}}$. 這也告訴我們任何的方陣都會 similar to 這樣形式的方陣, 我們稱此為 *rational form.*

Example 4.4.11. 考慮 over $\mathbb{R}$, 我們要求出 A 的 rational form, 其中

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

首先算出 $\chi_A(x)=(x^2+1)(x-1)^3$, 再求得 $\mu_A(x)=(x^2+1)(x-1)^2$.

首先考慮 primary decomposition, 求出 A^2+I_5 與 $(A-I_5)^2$ 的 null space W_1,W_2. 得 W_1,W_2 之一組 basis 分別為 $\{(1,0,0,0,0)^{\mathrm{t}},(0,1,0,0,0)^{\mathrm{t}}\}$, $\{(-1,0,0,0,1)^{\mathrm{t}},(1,0,1,0,0)^{\mathrm{t}},(-1,1,0,1,0)^{\mathrm{t}}\}$.

由 $\dim(W_1)=2$ 可知 W_1 本身是一個 cyclic space. 事實上取 $\mathbf{w}_1=(1,0,0,0,0)^{\mathrm{t}}$, 則 $A\mathbf{w}_1=(2,1,0,0,0)^{\mathrm{t}}$ (注意 $A^2\mathbf{w}_1=-\mathbf{w}_1$). 即 $W_1=C_{\mathbf{w}_1}$.

至於要將 W_2 分解成 cyclic subspaces 的 direct sum, 我們需先選出 $\mathbf{w}_2$ 滿足 $(A-I_5)\mathbf{w}_2\neq(0,0,0,0,0)^{\mathrm{t}}$. 事實上若選 $\mathbf{w}_2=(1,0,1,0,0)^{\mathrm{t}}$, 則 $A\mathbf{w}_2=(1,1,2,1,0)^{\mathrm{t}}$ (注意 $A^2\mathbf{w}_2=2A\mathbf{w}_2-\mathbf{w}_2$), 所以 $\dim(C_{\mathbf{w}_2})=2$. 由於 $\dim(W_2)=3$, 我們知道 W_2 應為 $C_{\mathbf{w}_2}$ 和另一個 dimension 為 1 的 cyclic subspace 的 direct sum. 此 cyclic subspace 應為 eigenvalue 為 1 的 eigenvector $\mathbf{w}_3$ 所形成, 而且 $\mathbf{w}_3\notin C_{\mathbf{w}_2}$. 我們選取 $\mathbf{w}_3=(-1,0,0,0,1)^{\mathrm{t}}$, 所以若令

$$P=\begin{pmatrix} 1 & 2 & 1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},\ \text{則 } P^{-1}AP=\begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

為 A 的 rational form.

事實上在 Example 4.4.11 中, 我們可以很快的判斷出 A 的 rational form. 因為 $\chi_A(x)=(x^2+1)(x-1)^3$ 所以 A^2+I_5 的 null space 僅由一個 cyclic subspace 所組成, 且其 cyclic vector 的 annihilator 為 x^2+1. 而 $\mu_A(x)=(x^2+1)(x-1)^2$ 所以由 Theorem 4.4.10 知, $(A-I_5)^2$ 的 null space 中一定有一個 cyclic subspace 其 cyclic vector 的 annihilator 為 $(x-1)^2$. 也因而我們知僅剩的 cyclic subspace 其 cyclic vector 的 annihilator 為 $x-1$. 所以 A 的 rational form 為有三個 blocks 的 diagonal matrix 其中每個 block 分別為 x^2+1, x^2-2x+1 以及 $x-1$ 的 companion matrix. 一般來說一個 matrix 的 rational form 並不能僅由其 characteristic polynomial 和 minimal polynomial 就能確定. 不過我們可以列出其所有可能的情形. 另一方面 rational form 是一個 canonical form, 也就是說兩個矩陣是 similar 的若且唯若它們可以化成同樣的 rational form. 我們會在下一節討論完 classical form 之後再探討這些課題.

4.5. Classical Form

當一個 linear operator 的 minimal polynomial 可以完全分解成一次多項式的乘積時, 除非它沒有重根 (即 diagonalizable), 此 linear operator 的 rational form 並不是 Jordan form. 此節中我們將說明如何另外選取 cyclic subspace 的一組 basis, 將其化成所謂的 classical form. 我們很容易看出 classical form 就是 Jordan form 的推廣.

當 $T:V\to V$ 是 F-linear, 給定 $\mathbf{v}\in V$, 考慮 T-cyclic subspace $C_\mathbf{v}$. 如果 $\mathbf{v}$ 的 T-annihilator 可以寫成 $\mu_\mathbf{v}(x)=p(x)^m$ (這裡 $p(x)\in F[x]$ 不需假設為 irreducible), 回顧一下若 $\deg(p(x))=d$, 則 $\{\mathbf{v},T(\mathbf{v}),T^{\circ 2}(\mathbf{v}),\dots,T^{\circ md-1}(\mathbf{v})\}$ 為 $C_\mathbf{v}$ 的一組 basis, 稱為 cyclic basis. 我們可以考慮以下一組新的 basis.

Lemma 4.5.1. 假設 $T:V\to V$ 是 F-*linear,* 給定 $\mathbf{v}\in V$. 若 $\mu_\mathbf{v}(x)=p(x)^m$, 其中 $p(x)\in F[x]$ 且 $\deg(p(x))=d$, 則

$$\begin{matrix} \mathbf{v} & T(\mathbf{v}) & \dots & T^{\circ d-1}(\mathbf{v}) \\ p(T)(\mathbf{v}) & p(T)(T(\mathbf{v})) & \dots & p(T)(T^{\circ d-1}(\mathbf{v})) \\ \vdots & \vdots & & \vdots \\ p^{m-1}(T)(\mathbf{v}) & p^{m-1}(T)(T(\mathbf{v})) & \dots & p^{m-1}(T)(T^{\circ d-1}(\mathbf{v})) \end{matrix} \tag{4.7}$$

是 $C_\mathbf{v}$ 的一組 *basis.*

Proof. 由 $\mu_\mathbf{v}(x)=p(x)^m$ 知 $\dim(C_\mathbf{v})=dm$. 因為 (4.7) 中共有 dm 個元素, 若能證明它們為 linearly independent over F, 則它們便是 $C_\mathbf{v}$ 的一組 basis.

對於 $0\le i\le m-1$, $0\le j\le d-1$, 若令 $h_{i,j}(x)=p^i(x)x^j$, 則 $p^i(T)(T^{\circ j})(\mathbf{v})=h_{i,j}(T)(\mathbf{v})$. 因為 $\deg(h_{i,j}(x))=di+j$, 我們知若 $(i,j)\neq(i',j')$, 則 $\deg(h_{i,j}(x))\neq\deg(h_{i',j'}(x))$. 換言之, 若 $c_{0,0},\dots,c_{i,j},\dots,c_{m-1,d-1}\in F$ 不全為 0, 則 $\sum_{i,j}c_{i,j}h_{i,j}(x)$ 是 $F[x]$ 中一個 nonzero polynomial.

現若存在一組不全為 0 的 $\{c_{i,j}\}$ 使得 $\sum_{i,j}c_{i,j}p^i(T)(T^{\circ j}(\mathbf{v}))=\mathbf{O}_V$, 表示 $h(x)=\sum_{i,j}c_{i,j}h_{i,j}(x)$ 這一個 nonzero polynomial 會滿足 $h(T)(\mathbf{v})=\mathbf{O}_V$. 很顯然 $\deg(h(x))<dm=\deg(\mu_\mathbf{v}(x))$, 這和 annihilator 的定義相矛盾, 故得證 (4.7) 中的元素為 linearly independent. □