# 18.700 LECTURE NOTES, 11/15/04

Contents



## 1. The generalized eigenspace decomposition

Let $T : V \to V$ be a linear operator on a finite dimensional vector space. Denote the factorization of $c_T(X)$ by $f_1(X)^{e_1} \cdots f_s(X)^{e_s}$, where every $f_i(X)$ is an irreducible, monic polynomial of degree $d_i > 0$,

$$f_i(X) = X^{d_i} + a_{i,d_i-1}X^{d_i-1} + \cdots + a_{i,1}X + a_{i,0},$$

where $f_1, \ldots, f_s$ are distinct and where $e_1, \ldots, e_s$ are positive integers.

**Remark 1.1.** The case we are most interested in is where each $f_i(X) = X - \lambda_i$, for distinct elements $\lambda_1, \ldots, \lambda_s \in \mathbb{F}$. This is always the case if the field $\mathbb{F}$ is algebraically closed, e.g., if $\mathbb{F} = \mathbb{C}$.

**Lemma 1.2.** *If $s \geq 2$, then for every $1 \leq i < j \leq 2$, the polynomials $f_i^{e_i}$ and $f_j^{e_j}$ are coprime.*

*Proof.* Because $f_i$ and $f_j$ are irreducible, the only factors of each are scalar multiples of 1 and $f_i$, resp. 1 and $f_j$. Since $f_i \neq f_j$ and each are monic, they are not proportional (i.e., they are linearly independent). Therefore 1 is a greatest common factor of $f_i$ and $f_j$. By the Chinese Remainder Theorem, Theorem 3.2 from 11/10/04, there exist polynomials $g_i, g_j$ such that $1 = g_i f_i + g_j f_j$. Let $e = \max_{e_i,e_j}$. Then $1^{2e-1} = (g_i f_i + g_j f_j)^{2e-1}$. Expanding the right-hand-side using the multinomial formula, each term is divisible by $f_i^k f_j^{2e-1-k}$ for some $0 \leq k \leq 2e-1$. At least one of $k, 2e-1-k$ is at least $e$. Gathering terms, the equation is $1 = 1^{2e-1} = h_i f_i^e + h_j f_j^e$ for some polynomials $h_i, h_j$. Therefore $1 = h'_i f_i^{e_i} + h'_j f_j^{e_j}$, where $h'_i = h_i f_i^{e-e_i}$, $h'_j = h_j f_j^{e-e_j}$. By the Chinese Remainder Theorem again, $f_i^{e_i}$ and $f_j^{e_j}$ are coprime. □

**Notation 1.3.** For every $i = 1, \ldots, s$, denote $E_i = E_{T,f_i(X)^{e_i}} = \ker(f_i^{e_i}(T))$, as in Definition 2.3 from 11/10/04. Denote $\mathcal{W} = (E_1, \ldots, E_s)$ as in Definition 1.1 from 11/10/04. If $s \geq 2$, for every $i = 1, \ldots, s$, denote,

$$r_i(X) = \prod_{1 \leq j \leq s, j \neq i} f_j(X)^{e_j},$$

as in Corollary 3.3 from 11/10/04.

**Proposition 1.4.** *If $s \geq 2$, there exist polynomials $g_1(X), \dots, g_s(X)$ such that $1 = g_1(X)r_1(X) + \cdots + g_s(X)r_s(X)$.*

*Proof.* This follows from Corollary 3.3 from 11/10/04 and Lemma 1.2. □

**Theorem 1.5.** *The $s$-tuple of subspace $(E_1, \dots, E_s)$ is a direct sum decomposition of $V$. Moreover, if $s \geq 2$ then $g_i(T) \circ r_i(T) : V \to V$ has image $E_i$ and is the unique linear operator whose restriction to $E_i$ is the identity and whose restriction to $E_j$ is 0 for $1 \leq j \leq s, j \neq i$.*

*Proof.* First consider the case that $s = 1$. Then $c_T(X) = f_1(X)^{e_1}$. By the Cayley-Hamilton Theorem, Corollary 6.5 from 11/12/04, $c_T(T)$ is the zero operator. Therefore $E_1 = \ker(f_1^{e_1}(T))$ is all of $V$. So $\mathcal{W} = (V)$ is clearly a direct sum decomposition of $V$.

Next assume $s \geq 2$. By the same arguments from Section 4 from 11/10/04, the restriction of $g_i(T) \circ r_i(T)$ to $E_j$ is the identity if $j = i$ and is the zero transformation if $j \neq i$. By the same argument as in Theorem 4.4 from 11/10/04, $\mathcal{W}$ is linearly independent. The claim is that $\text{Image}(g_i(T) \circ r_i(T))$ is contained in $E_i$. This is the same as saying that $f_i^{e_i}(T) \circ g_i(T) \circ r_i(T)$ is the zero linear operator. Of course this is $(f_i^{e_i} \cdot g_i \cdot r_i)(T) = (g_i \cdot f_i^{e_i} \cdot r_i)(T) = g_i(T) \circ c_T(T)$. By the Cayley-Hamilton theorem, $c_T(T)$ is the zero operator, thus $g_i(T) \circ c_T(T)$ is the zero operator, so $\text{Image}(g_i(T) \circ r_i(T)) \subset E_i$.

Because $1 = g_1 \cdot r_1 + \cdots + g_s \cdot r_s$, there is an equation of linear operators $\text{Id}_V = g_1(T) \circ r_1(T) + \cdots + g_s(T) \circ r_s(T)$. For every $\mathbf{v} \in V$, this gives,

$$\mathbf{v} = g_1(T) \circ r_1(T)(\mathbf{v}) + \cdots + g_s(T) \circ r_s(T)(\mathbf{v}).$$

By the last paragraph, each $\mathbf{v}_i := g_i(T) \circ r_i(T)(\mathbf{v})$ is in $E_i$. Therefore $\mathcal{W}$ is spanning. Since it is both linearly independent and spanning, $\mathcal{W}$ is a direct sum decomposition.

To see that $g_i(T) \circ r_i(T)$ is the unique linear operator whose restriction to every $E_j$ is either $\text{Id}_{E_i}$ if $j = i$ or else the zero operator if $j \neq i$, assume $\pi_i : V \to V$ is also such an operator. Then $\pi_i - g_i(T) \circ r_i(T)$ is a linear operator whose restriction to every $E_j$ is the zero operator, i.e., $\ker(\pi_i - g_i(T) \circ r_i(T))$ contains $E_j$ for every $j = 1, \dots, s$. Since these subspaces span $V$, the kernel contains all of $V$, i.e., $\pi_i - g_i(T) \circ r_i(T)$ is the zero operator. Therefore $\pi_i = g_i(T) \circ r_i(T)$. □

## 2. Operators commuting with $T$

**Definition 2.1.** Linear operators $T, T' : V \to V$ *commute* if $T \circ T' = T' \circ T$, and the pair $(T, T')$ is a *commuting pair.*

**Lemma 2.2.**
- (i) *For every commuting pair $(T_1, T')$ and every commuting pair $(T_2, T')$, $(T_1 + T_2, T')$ is a commuting pair.*
- (ii) *For every commuting pair $(T, T')$, for every $a \in \mathbb{F}$, $(a \cdot T, T')$ is a commuting pair.*
- (iii) *For every commuting pair $(T_1, T')$ and every commuting pair $(T_2, T')$, $(T_2 \circ T_1, T')$ is a commuting pair.*
- (iv) *For every commuting pair $(T, T')$ of linear operators on $V$, and for every $f(X) \in \mathbb{F}[X]$, $(f(T), T')$ is a commuting pair of linear operators on $V$.*

*Proof.* **(i)** By distributivity of composition and addition, $(T_1 + T_2) \circ T' = T_1 \circ T' + T_2 \circ T'$. By hypothesis, $T_1 \circ T' = T' \circ T_1$ and $T_2 \circ T' = T' \circ T_2$ so that $T_1 \circ T' + T_2 \circ T' = T' \circ T_1 + T' \circ T_2$. By distributivity of composition and addition, $T' \circ T_1 + T' \circ T_2 = T' \circ (T_1 + T_2)$. Therefore $(T_1 + T_2) \circ T' = T' \circ (T_1 + T_2)$, i.e., $(T_1 + T_2, T')$ is a commuting pair.

**(ii)** By distributivity of scalar multiplication and composition, $(a \cdot T) \circ T' = a \cdot (T \circ T')$. By hypothesis, $T \circ T' = T' \circ T$, thus $a \cdot (T \circ T') = a \cdot (T' \circ T)$. By distributivity of scalar multiplication and composition, $a \cdot (T' \circ T) = T' \circ (a \cdot T)$. Therefore $(a \cdot T) \circ T' = T' \circ (a \cdot T)$, i.e., $(a \cdot T, T')$ is a commuting pair.

**(iii)** By associativity of composition, $(T_1 \circ T_2) \circ T' = T_1 \circ (T_2 \circ T')$. By hypothesis, $T_2 \circ T' = T' \circ T_2$, so that $T_1 \circ (T_2 \circ T') = T_1 \circ (T' \circ T_2)$. By associativity of composition, this is $(T_1 \circ T') \circ T_2$. By hypothesis, $T_1 \circ T' = T' \circ T_1$, so that $(T_1 \circ T') \circ T_2 = (T' \circ T_1) \circ T_2$. By associativity of composition, this is $T' \circ (T_1 \circ T_2)$. Therefore $(T_1 \circ T_2) \circ T' = T' \circ (T_1 \circ T_2)$, i.e., $(T_1 \circ T_2, T')$ is a commuting pair.

**(iv)** The first claim is that for every integer $n \geq 0$, $(T^n, T')$ is a commuting pair. This is proved by induction on $n$. For $n = 0$ this is trivial because $T^0 = \mathrm{Id}_V$ and clearly $\mathrm{Id}_V \circ T' = T' = T' \circ \mathrm{Id}_V$. By way of induction, assume $n > 0$ and the result is known for $n - 1$, i.e., $(T^{n-1}, T')$ is a commuting pair. By hypothesis, $(T, T')$ is a commuting pair. By (iii), $(T \circ T^{n-1}, T')$ is a commuting pair, i.e., $(T^n, T')$ is a commuting pair, proving the claim by induction on $n$.

Now let $f(X) = a_n X^n + \cdots + a_1 X + a_0$. Then $f(T) = a_n T^n + \cdots + a_1 T + a_0 \mathrm{Id}_V$. By the last paragraph, each $(T^k, T')$ is a commuting pair. Repeatedly applying (i) and (ii), $(f(T), T')$ is a commuting pair. □

**Proposition 2.3.** *For every commuting pair $(T, T')$ of linear operators on $V$, for every $f(X) \in \mathbb{F}[X]$, $T'(E_{T,f(X)}) \subset E_{T,f(X)}$. In particular, $T(E_{T,f(X)}) \subset E_{T,f(X)}$ because $(T, T)$ is a commuting pair.*

*Proof.* By Lemma 2.2(iv), $(f(T), T')$ is a commuting pair. Therefore, for every $\mathbf{v} \in E_{T,f(X)}$, $f(T)(T'(\mathbf{v}) = (f(T) \circ T')(\mathbf{v})$ equals $(T' \circ f(T))(\mathbf{v}) = T'(f(T)(\mathbf{v}))$. Because $\mathbf{v} \in E_{T,f(X)}$, $f(T)(\mathbf{v}) = \mathbf{0}$. Thus $f(T)(T'(\mathbf{v})) = T'(\mathbf{0}) = \mathbf{0}$. Therefore $T'(\mathbf{v}) \in E_{T,f(X)}$, i.e., $T'(E_{T,f(X)}) \subset E_{T,f(X)}$. □

**Notation 2.4.** For every commuting pair $(T, T')$ and every $f(X) \in \mathbb{F}[X]$, denote by $T'_{f(X)} : E_{T,f(X)} \to E_{T,f(X)}$ the unique linear operator that agrees with the restriction of $T'$ to $E_{T,f(X)}$.

The relevance is the following. Let $c_T(X) = f_1(X)^{e_1} \cdots f_s(X)^{e_s}$. For every $i = 1, \dots, s$, let $\mathcal{B}_i$ be an ordered basis for $E_i = E_{T,f_i^{e_i}}$ of size $n_i = \dim(E_i)$. Let $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_s$. By Proposition 1.5 from 11/10/04 and by Theorem 1.5, $\mathcal{B}$ is an ordered basis for $V$, in particular $n = n_1 + \cdots + n_s$. Let $\underline{n} = (0, n_1, n_1 + n_2, \dots, n_1 + \cdots + n_i, \dots, n_1 + \cdots + n_s)$ be the partition of $n$ associated to $n = n_1 + \cdots + n_s$. Denote by $A$ the $n \times n$-matrix, $[T]_{\mathcal{B},\mathcal{B}}$. For every $i = 1, \dots, s$, denote $T_i = T_{f_i^{e_i}}$, and denote by $A_i$ the $n_i \times n_i$-matrix, $[T_i]_{\mathcal{B}_i,\mathcal{B}_i}$.

**Definition 2.5.** Let $\underline{j}$ be a partition of $n$. A $(\underline{j}, \underline{j})$-partitioned $n \times n$-matrix, $B$, is a *diagonal block matrix* if $B_{i,j}$ is the zero matrix unless $i = j$.

**Proposition 2.6.** *The* $(\underline{n}, \underline{n})$*-partitioned matrix* $A$ *has blocks,*

$$A_{i,j} = \begin{cases} A_i, & i = j, \\ 0_{n_i,n_j}, & i \neq j \end{cases}$$

*In other words,* $A$ *is a diagonal block matrix and the diagonal block* $A_{i,i}$ *equals* $A_i$*.*

*Proof.* For every $i = 1, \dots, s$ and every $1 \leq l \leq n_i$, the $(n_1 + \cdots + n_s) - n_s + l$ column of $A$ is the $\mathcal{B}$-coordinate vector of $T(\mathbf{v}_{i,l})$, where $\mathbf{v}_{i,l}$ is the $l^{\text{th}}$ vector in $\mathcal{B}_i$. By Proposition 2.3, $T(\mathbf{v}_{i,l}) \in E_{T,f_i^{e_i}}$, thus it is a linear combination of vectors in $\mathcal{B}_i$. Therefore the only nonzero entries of the $(n_1 + \cdots + n_s) - n_s + l$ column occur in the $(i,i)$-block. This proves that $A_{i,j} = 0_{n_i,n_j}$ if $i \neq j$. Moreover, since $T_i$ agrees with the restriction of $T$ to $E_{T,f_i^{e_i}}$, the $\mathcal{B}$-coordinate vector of $T(\mathbf{v}_{i,l})$ equals the $\mathcal{B}$-coordinate vector of $T_i(\mathbf{v}_{i,l})$. This proves $A_{i,i} = A_i$. □

**Corollary 2.7.** *The characteristic polynomial of* $T$ *is* $c_T(X) = c_{T_1}(X) \cdots c_{T_s}(X)$*.*

*Proof.* For a block matrix $A$ as above, clearly $c_A(X) = c_{A_1}(X) \cdots c_{A_s}(X)$, by cofactor expansion. □

**Corollary 2.8.** *Assume* $c_T(X) = (X - \lambda_1)^{e_1} \cdots (X - \lambda_s)^{e_s}$*. For every* $i = 1, \dots, s$*,* $c_{T_i}(X) = (X - \lambda_i)^{e_i}$*. In particular, dim*$(E^{(e_i)}_{T,\lambda_i}) = e_i$*, the algebraic multiplicity of* $\lambda_i$*.*

*Proof.* Let $X - \mu$ be a linear factor of $c_{T_i}(X)$, i.e., $\mu$ is an eigenvalue for $T_i$. There exists a nonzero $\mu$-eigenvector $\mathbf{v}$ for $T_i$. The claim is that for every integer $e \geq 0$, $(T_i - \lambda_i \mathrm{Id})^e(\mathbf{v}) = (\mu - \lambda_i)^e \cdot \mathbf{v}$. For $e = 0$ this is obvious. By way of induction, assume $e > 0$ and the result is true for $e - 1$. Then, by definition of $(T_i - \lambda_i \mathrm{Id})^e$, $(T_i - \lambda_i \mathrm{Id})^e(\mathbf{v}) = (T_i - \lambda_i \mathrm{Id})^{e-1}(T_i(\mathbf{v}) - \lambda_i \cdot \mathbf{v})$. By hypothesis, $T_i(\mathbf{v}) = \mu \cdot \mathbf{v}$. So this is $(T_i - \lambda_i \mathrm{Id})^{e-1}((\mu - \lambda) \cdot \mathbf{v})$, which by linearity equals $(\mu - \lambda_i) \cdot (T_i - \lambda_i \mathrm{Id})^{e-1}(\mathbf{v})$. By the induction hypothesis, this is $(\mu - \lambda_i) \cdot ((\mu - \lambda_i)^{e-1} \cdot \mathbf{v})$. By distributivity of multiplication of scalars and scalar multiplication, this is $((\mu - \lambda_i)(\mu - \lambda_i)^{e-1}) \cdot \mathbf{v} = (\mu - \lambda_i)^e \cdot \mathbf{v}$. So the claim is proved by induction on $e$.

By definition, $(T_i - \lambda_i \mathrm{Id})^{e_i}$ is the zero operator. Thus, $(\mu - \lambda_i)^{e_i} \cdot \mathbf{v} = (T_i - \lambda_i \mathrm{Id})^{e_i}(\mathbf{v}) = \mathbf{0}$. Since $\mathbf{v} \neq \mathbf{0}$, $(\mu - \lambda_i)^{e_i} = 0$. Therefore $\mu = \lambda_i$, i.e., the only linear factor of $c_{T_i}(X)$ is $X - \lambda_i$.

By Corollary 2.7, $c_{T_i}(X)$ factors $c_T(X)$, in particular it is a product of linear factors. Since the only linear factor of $c_{T_i}(X)$ is $X - \lambda_i$, $c_{T_i}(X) = (X - \lambda_i)^{n_i}$, where $n_i = \dim(E^{(e_i)}_{T,\lambda_i})$. Therefore $c_T(X) = c_{T_1}(X) \cdots c_{T_s}(X) = (X - \lambda_1)^{n_1} \cdots (X - \lambda_s)^{n_s}$. Since also $c_T(X) = (X - \lambda_1)^{e_1} \cdots (X - \lambda_s)^{e_s}$, for every $i = 1, \dots, s$, $n_i = e_i$, i.e., $\dim(E^{e_i}_{T,\lambda_i})$ equals the algebraic multiplicity $e_i$. □

**Remark 2.9.** It is true, more generally, that if $c_T(X) = f_1(X)^{e_1} \cdots f_s(X)^{e_s}$ is the the irreducible decomposition, then for every $i = 1, \dots, s$, $c_{T_i}(X) = f_i(X)^{e_i}$. The simplest proof uses "base-change to the algebraic closure" and Corollary 2.8.

Let $(T, T')$ be a commuting pair. Let $c_T(X) = f_1^{e_1} \cdots f_s^{e_s}$. Let $\mathcal{W} = (E_1, \dots, E_s)$, where $E_i = E_{T,f_i^{e_i}}$. Let $f(X) \in \mathbb{F}[X]$ and for every $i = 1, \dots, s$, define $\mathcal{W}_{T',f(X)} = (E_{T',f(X),1}, \dots, E_{T',f(X),s})$ by $E_{T',f(X),i} = E_{T',f(X)} \cap E_i$.

**Proposition 2.10.** *The sequence* $\mathcal{W}_{T',f(X)}$ *is a direct sum decomposition of* $E_{T',f(X)}$*.*

*Proof.* Because $\mathcal{W}$ is linearly independent, and because each $E_{T',f(X),i}$ is contained in $E_i$, also $\mathcal{W}_{T',f(X)}$ is linearly independent. Let $\mathbf{v} \in E_{T',f(X)}$. By Theorem 1.5, there exists an ordered $s$-tuple of vectors in $\mathcal{W}$, $(\mathbf{v}_1, \dots, \mathbf{v}_s)$ such that $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_s$.

The claim is that for each $i = 1, \dots, s$, $\mathbf{v}_i \in E_{T',f(X)}$. By Lemma 2.2(iv), $(T, f(T'))$ is a commuting pair. For every $i = 1, \dots, s$, denote $\mathbf{v}'_i = f(T')(\mathbf{v}_i)$. By Proposition 2.3, $\mathbf{v}'_i \in E_i$. And,

$$\mathbf{v}'_1 + \cdots + \mathbf{v}'_s = f(T')(\mathbf{v}_1) + \cdots + f(T')(\mathbf{v}_s) = f(T')(\mathbf{v}_1 + \cdots + \mathbf{v}_s) = f(T')(\mathbf{v}) = \mathbf{0},$$

where the first inequality is the definition, the second is by linearity, the third is by definition, and the last is the hypothesis that $\mathbf{v} \in E_{T',f(X)}$. Because $\mathcal{W}$ is linearly independent, $\mathbf{v}'_1 = \cdots = \mathbf{v}'_s = \mathbf{0}$. Therefore each $\mathbf{v}_i \in E_{T',f(X)}$, i.e., $\mathbf{v}_i \in E_{T',f(X),i}$. So $\mathcal{W}_{T',f(X)}$ spans $E_{T',f(X)}$. Thus $\mathcal{W}_{T',f(X)}$ is a direct sum decomposition of $E_{T',f(X)}$. □

## 3. The semisimple-nilpotent decomposition

**Definition 3.1.** A linear operator $N : V \to V$ is *nilpotent of index $e$* if $N^e$ is the zero operator. A linear operator $N : V \to V$ is *nilpotent* if there exists an integer $e > 0$ such that $N$ is nilpotent of index $e$. A linear operator $S : V \to V$ is *semisimple* if there exists a finite ordered basis $\mathcal{B}$ for $V$ such that $[S]_{\mathcal{B},\mathcal{B}}$ is a diagonal matrix.

For a linear operator $T : V \to V$, a *semisimple-nilpotent decomposition* is a pair $(S, N)$ of a semisimple and nilpotent matrix such that

(i) $T = S + N$,
(ii) $(T, S)$ is a commuting pair, and
(iii) $(T, N)$ is a commuting pair.

**Lemma 3.2.** *Let $(S, T)$ be a commuting pair where $S$ is diagonalizable. Then for every $f(X) \in \mathbb{F}[X]$, $E_{T,f(X)}$ has a basis of $S$-eigenvectors.*

*Proof.* Let $c_S(X) = (X - \lambda_1)^{e_1} \cdots (X - \lambda_s)^{e_s}$. Choosing a basis with respect to which $S$ is diagonal, it is clear that for every $i = 1, \dots, s$, $E^{(e_i)}_{S,\lambda_i} = E_{S,\lambda_i}$, the $\lambda_i$-eigenspace of $S$. So the sequence $\mathcal{W} = (E^{(e_1)}_{S,\lambda_1}, \dots, E^{(e_s)}_{S,\lambda_s})$ is the sequence of $\lambda_i$-eigenspaces for $S$. By Proposition 2.10, $\mathcal{W}_{T,f(X)} = (E^{(e_1)}_{S,\lambda_i} \cap E_{T,f(X)}, \dots, E^{(e_s)}_{S,\lambda_i} \cap E_{T,f(X)})$ is a direct sum decomposition of $E_{T,f(X)}$. For every $i = 1, \dots, s$, let $\mathcal{B}_i$ be an ordered basis for $E^{(e_i)}_{S,\lambda_i} \cap E_{T,f(X)}$. Because these vectors are contained in $E^{(e_i)}_{S,\lambda_i}$, they are $\lambda_i$-eigenvectors of $S$. By Proposition 1.5 from 11/10/04, the concatenation $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_s$ is an ordered basis for $E_{T,f(X)}$ consisting of eigenvectors for $S$. □

**Theorem 3.3.** *Assume $c_T(X) = (X - \lambda_1)^{e_1} \cdots (X - \lambda_s)^{e_s}$. There exists a unique semisimple-nilpotent decomposition $(S, N)$ for $T$.*

*Proof.* **Existence:** By Theorem 1.5, $\mathcal{W} = (E^{(e_1)}_{T,\lambda_1}, \dots, E^{(e_s)}_{T,\lambda_s})$ is a direct sum decomposition of $V$. Define $S : V \to V$ to be the unique linear operator such that for every $i = 1, \dots, s$, the restriction of $S$ to $E^{(e_i)}_{T,\lambda_i}$ is $\lambda_i \mathrm{Id}$. For the basis $\mathcal{B}$ of Proposition 2.6, $S$ is diagonal.

For every $i = 1, \dots, s$, the restriction of $T \circ S$ to $E^{(e_i)}_{T,\lambda_i}$ is $T_i \circ (\lambda_i \mathrm{Id}) = \lambda \cdot (T_i \circ \mathrm{Id}) = \lambda_i \cdot T_i = (\lambda_i \mathrm{Id}) \circ T_i$, which is the restriction of $S \circ T$ to $E^{(e_i)}_{T,\lambda_i}$. So for every $i = 1, \dots, s$, the restriction of $T \circ S - S \circ T$ to $E^{(e_i)}_{T,\lambda_i}$ is the zero operator, i.e., $E^{(e_i)}_{T,\lambda_i}$ is in the kernel of $T \circ S - S \circ T$. Since $\mathcal{W}$ spans $V$, $T \circ S - S \circ T$ is the zero operator, i.e., $T \circ S = S \circ T$.

Define $N = T - S$. Because $(S, T)$ is a commuting pair, and because $(T, T)$ is a commuting pair, by Lemma 2.2, $(N, S)$ is a commuting pair. Therefore $N(E^{(e_i)}_{T,\lambda_i}) \subset E^{(e_i)}_{T,\lambda_i}$ by Proposition 2.3. Moreover, denoting by $N_i$ the restriction of $N$ to $E^{(e_i)}_{T,\lambda_i}$, by definition $N_i^{e_i} = (T_i - \lambda_i \mathrm{Id})^{e_i}$ is the zero linear operator. Denote $e = \max(e_1, \dots, e_s)$. Then for every $i = 1, \dots, s$, $N_i^e$ is the zero linear operator, i.e., $E^{(e_i)}_{T,\lambda_i}$ is in the kernel of $N^e$. Since $\mathcal{W}$ spans $V$, $N^e$ is the zero operator, i.e., $N$ is nilpotent of index $e$.

**Uniqueness:** Let $(S', N')$ be a semisimple-nilpotent decomposition of $T$. By Proposition 2.3, for every $i = 1, \dots, s$, $S'(E^{(e_i)}_{T,\lambda_i}) \subset E^{(e_i)}_{T,\lambda_i}$ and $N'(E^{(e_i)}_{T,\lambda_i}) \subset E^{(e_i)}_{T,\lambda_i}$. Denote by $S'_i$ and $N'_i$ the restrictions of $S'$ and $N'$ to $E^{(e_i)}_{T,\lambda_i}$. Then $S'_i \circ (\lambda_i \mathrm{Id}) = (\lambda_i \mathrm{Id}) \circ S'_i$, i.e., $S'_i \circ S_i = S_i \circ S'_i$ for every $i = 1, \dots, s$. Thus $(S', S)$ is a commuting pair.

Since $(S', S)$ is a commuting pair and $(T, S)$ is a commuting pair, by Lemma 2.2, $(N', S)$ is a commuting pair. Since also $(N', T)$ is a commuting pair, by Lemma 2.2 $(N', N)$ is a commuting pair. Let $N$ be nilpotent of index $e$ and let $N'$ be nilpotent of index $e'$. Because $N$ and $N'$ commute, the binomial theorem applies and $(N - N')^{e+e'-1} = B \circ N^e + C \circ (N')^{e'}$ for some linear operators $B$ and $C$. Because $N$ is nilpotent of index $e$ and $N'$ is nilpotent of index $e'$, $B \circ N^e + C \circ (N')^{e'} = 0 + 0$, i.e., $(N - N')^{e+e'-1}$ is the zero operator. By hypothesis, $S + N = T = S' + N'$, so $S' - S = N - N'$. Thus $(S' - S)^{e+e'-1}$ is the zero operator. In particular, for every $i = 1, \dots, s$, $(S'_i - \lambda_i \mathrm{Id})^{e+e'-1}$ is the zero operator.

By Lemma 3.2, there is a basis of $E^{(e_i)}_{T,\lambda_i}$ of $S'_i$-eigenvectors. If $\mathbf{v}$ is a $\mu$-eigenvector for $S'_i$, by the same argument as in the proof of Corollary 2.8, $\mu = \lambda_i$. Thus $E^{(e_i)}_{T,\lambda_i}$ has a basis of $\lambda_i$-eigenvectors for $S'_i$, i.e., $S'_i = \lambda_i \mathrm{Id}$ for every $i = 1, \dots, s$. Therefore $S' = S$, and so also $N' = T - S' = T - S$ equals $N$. So $(S, N)$ is the unique semisimple-nilpotent decomposition of $T$. □

**Corollary 3.4.** *Let $(S, N)$ be the semisimple-nilpotent decomposition of $T$. For every linear operator $T'$, $(T, T')$ is a commuting pair iff*

(i) $T'(E^{(e_i)}_{T,\lambda_i}) \subset E^{(e_i)}_{T,\lambda_i}$ *for every $i = 1, \dots, s$, and*
(ii) $(N, T')$ *is a commuting pair.*

*Proof.* If $(T, T')$ is a commuting pair, then by Proposition 2.3, $T'(E^{(e_i)}_{T,\lambda_i}) \subset E^{(e_i)}_{T,\lambda_i}$. Denote by $T'_i$ the restriction of $T$ to $E^{(e_i)}_{T,\lambda_i}$. Then for every $i = 1, \dots, s$, $T'_i$ commutes with $\lambda_i \mathrm{Id}$, i.e., $T'_i$ commutes with $S_i$. Therefore $T'$ commutes with $S$, i.e., $(S, T')$ is a commuting pair. By Lemma 2.2, since $(T, T')$ and $(S, T')$ are commuting pairs, also $(N, T')$ is a commuting pair.

Conversely, suppose that $T'$ satisfies (i) and (ii). By the argument above, $T'_i$ commutes with $S_i$ for every $i = 1, \dots, s$, i.e., $T'$ commutes with $S$. Since $(S, T')$ and $(N, T')$ are commuting pairs, by Lemma 2.2, also $(T, T')$ is a commuting pair. □

Let $N : V \to V$ be a nilpotent operator. For every integer $e \geq 0$, define $E^{(e)} = E^{(e)}_{N,0} = \ker(N^e)$.

**Lemma 4.1.** *For every integer $e \geq 0$, $N(E^{(e+1)}) \subset E^{(e)}$. For every integer $e \geq 1$, $E^{(e-1)} \subset E^{(e)}$.*

*Proof.* If $\mathbf{v} \in E^{(e+1)}$, then $N^e(N(\mathbf{v})) = N^{e+1}(\mathbf{v}) = \mathbf{0}$, therefore $N(\mathbf{v}) \in E^{(e)}$. Therefore $N(E^{(e+1)}) \subset E^{(e)}$. Clearly $E^{(e-1)} \subset E^{(e)}$. □

**Notation 4.2.** Denote $F^{(0)} = N(E^{(1)}) \subset E^{(0)}$. For every integer $e \geq 1$, denote $F^{(e)} = N(E^{(e+1)}) + E^{(e-1)} \subset E^{(e)}$.

**Lemma 4.3.** *For every integer $e \geq 0$, there exists a vector subspace $G^{(e)} \subset E^{(e)}$ so that $(F^{(e)}, G^{(e)})$ is a direct sum decomposition of $E^{(e)}$.*

*Proof.* For every integer $e$, let $\mathcal{B}$ be a basis for $F^{(e)}$. This is a linearly independent set of vectors in $E^{(e)}$. By the basis extension theorem, there exists a collection of vectors $\mathcal{B}'$ in $E^{(e)}$ so that $\mathcal{B} \cup \mathcal{B}'$ is a basis for $E^{(e)}$. Define $G^{(e)} = \mathrm{span}(\mathcal{B}')$. By Proposition 1.5 from 11/10/04, $(F^{(e)}, G^{(e)})$ is a direct sum decomposition of $E^{(e)}$. □

**Notation 4.4.** For every integer $e \geq 0$, denote by $\mathcal{A}_e = (\mathbf{v}_{e,1}, \dots, \mathbf{v}_{e,r_e})$ an ordered basis for $G^{(e)}$, possibly empty (in which case define $r_e = 0$). For every vector $\mathbf{v}_{e,j}$, for every $i = 1, \dots, e$, define $\mathbf{v}_{e,j,i} = N^{e-i}(\mathbf{v}_{e,j})$, and define $\mathcal{B}_{e,j} = (\mathbf{v}_{e,j,1}, \dots, \mathbf{v}_{e,j,e})$. Define $E_{e,j} = \mathrm{span}(\mathcal{B}_{e,j})$. Define $\mathcal{W}$ to be the sequence of all nonzero subspaces $E_{e,j}$.

**Theorem 4.5.** *For every $e \geq 0$ and every $1 \leq j \leq r_e$, $N(E_{e,j}) \subset E_{e,j}$. Moreover $\mathcal{W}$ is a direct sum decomposition of $V$.*

*Proof.* By construction $N$ maps each element of $\mathcal{B}_{e,j}$ either to $\mathbf{0}$ or to another element of $\mathcal{B}_{e,j}$. Therefore $N$ maps $E_{e,j}$ into itself.

For every $e \geq 0$ and every $1 \leq j \leq r_e$, let $\mathbf{v}_{(e,j)} \in E_{e,j}$ be a vector such that $\sum_{(e,j)} \mathbf{v}_{(e,j)} = \mathbf{0}$. Let $l \geq 0$ be the largest integer such that at least one element $N^l(\mathbf{v}_{(e,j)}) \neq \mathbf{0}$, if such an integer exists. For every $(e,f)$, define $\mathbf{v}'_{(e,j)} = N^l(\mathbf{v}_{(e,j)})$. By the first paragraph, each $\mathbf{v}'_{(e,j)} \in E_{e,j}$, and of course $\sum_{(e,j)} \mathbf{v}'_{(e,j)} = N^l(\mathbf{0}) = \mathbf{0}$. Moreover, $\mathbf{v}'_{(e,j)} \in E^{(1)}$ for every $(e,j)$. Therefore, for every $(e,j)$, $\mathbf{v}'_{(e,j)} = a_{(i,j)} \cdot \mathbf{v}_{(e,j,1)}$ for some $a_{(e,j)} \in \mathbb{F}$.

Let $e \geq 0$ be the least integer such that for some $j$, $a_{(e,j)} \neq 0$. Denote,

$$\mathbf{w}_e = \sum_{1 \leq j \leq r_e} a_{(e,j)} \mathbf{v}_{(e,j)},$$

and denote,

$$\mathbf{w}'_e = \sum_{e' > e} \sum_{1 \leq j \leq r_{e'}} N^{e'-e}(a_{e',j} \cdot \mathbf{v}_{e',j}).$$

Then $\mathbf{w}'_e \in N(E^{(e+1)}) \subset F^{(e)}$, and $N^{e-1}(\mathbf{w}_e + \mathbf{w}'_e)$ equals $\sum_{(e,j)} a_{(e,j)} \mathbf{v}_{(e,j,1)} = \mathbf{0}$. Thus $\mathbf{w}_e + \mathbf{w}'_e \in E^{(e-1)} \subset F^{(e)}$ So $\mathbf{w}_e = (\mathbf{w}_e + \mathbf{w}'_e) - \mathbf{w}'_e$ is in $F^{(e)}$. Of course also $\mathbf{w}_e \in G^{(e)}$. Since $(F^{(e)}, G^{(e)})$ is linearly independent by construction, $\mathbf{w}_e = \mathbf{0}$. But by construction, $(\mathbf{v}_{(e,1)}, \ldots, \mathbf{v}_{(e,r_e)})$ is linearly independent. Therefore $a_{(e,j)} = 0$ for every $1 \leq j \leq r_e$. This is a contradiction, proving the integer $l$ does not exist.

Since there is no integer $l \geq 0$ such that for some $(e,j)$, $N^l(\mathbf{v}_{(e,j)}$ is nonzero, in particular for $l = 0$, for every $(e,j)$, $\mathbf{v}_{(e,j)} = N^0(\mathbf{v}_{(e,j)})$ is the zero vector. This proves $\mathcal{W}$ is linearly independent. By construction $\mathcal{W}$ is clearly spanning, thus $\mathcal{W}$ is a direct sum decomposition of $V$. □

**Notation 4.6.** For every integer $e \geq 0$ and every $1 \leq j \leq r_e$, denote by $N_{e,j} : E_{e,j} \to E_{e,j}$ the restriction of $N$ to $E_{e,j}$.

**Definition 4.7.** For every integer $n \geq 1$, the *nilpotent Jordan block of size* $n$ is the $n \times n$ matrix $J_{0,n}$ with,

$$J_{0,n}(i,j) = \begin{cases} 1, & 1 \leq i \leq n-1, j = i+1, \\ 0, & \text{otherwise} \end{cases}$$

For every integer $n \geq 1$ and every element $\lambda \in \mathbb{F}$, the *Jordan block of size* $n$ *and eigenvalue* $\lambda$ is the $n \times n$ matrix $J_{\lambda,n} = \lambda I_n + J_{0,n}$.

**Proposition 4.8.** *For every* $e \geq 0$ *and every* $1 \leq j \leq r_e$*, the matrix representative* $[N_{e,j}]_{\mathcal{B}_{e,j},\mathcal{B}_{e,j}}$ *is the nilpotent Jordan block of length* $e$*,* $J_{0,e}$*.*

*Proof.* Denote $J = [N_{e,j}]_{\mathcal{B}_{e,j},\mathcal{B}_{e,j}}$. By construction, $N_{e,j}(\mathbf{v}_{e,j,1}) = \mathbf{0}$, therefore the first column of $J$ is the zero vector. And for every $i = 1, \ldots, e-1$, $N_{e,j}(\mathbf{v}_{e,j,i+1}) = \mathbf{v}_{e,j,i}$. Therefore the $(i+1)^{\text{st}}$ column of $J$ is the $i^{\text{th}}$ standard basis vector. This is precisely the definition of $J_{0,e}$. □

**Corollary 4.9.** *There exists a basis* $\mathcal{B}$ *for* $V$ *so that the matrix representative* $[N]_{\mathcal{B},\mathcal{B}}$ *is a diagonal block matrix whose diagonal blocks are all nilpotent Jordan blocks.*

*Proof.* By Theorem 4.5, and by Proposition 1.5 from 11/10/04, the concatenation $\mathcal{B}$ of all the sets $\mathcal{B}_{e,j}$ is a basis for $V$. By Theorem 4.5, the matrix representative $[N]_{\mathcal{B},\mathcal{B}}$ breaks up into diagonal blocks, $[T_{e,j}]_{\mathcal{B}_{e,j},\mathcal{B}_{e,j}}$. By Proposition 4.8, each of these blocks is a nilpotent Jordan block. □

**Corollary 4.10.** *Let* $T : V \to V$ *be a linear operator such that* $c_T(X) = (X - \lambda)^n$*. There exists an ordered basis* $\mathcal{B}$ *for* $V$ *such that* $[T]_{\mathcal{B},\mathcal{B}}$ *is a diagonal block matrix whose diagonal blocks are all Jordan blocks of eigenvalue* $\lambda$*.*

*Proof.* Define $N = T - \lambda \text{Id}_V$. By the Cayley-Hamilton theorem, $N^n$ is the zero operator, i.e., $N$ is nilpotent. By Corollary 4.9, there exists a basis $\mathcal{B}$ for $V$ such that $[N]_{\mathcal{B},\mathcal{B}}$ is a diagonal block matrix whose diagonal blocks are all nilpotent Jordan blocks $J_{0,e}$. Of course $[\lambda \text{Id}_V]_{\mathcal{B},\mathcal{B}} = \lambda I_n$. Therefore $[T]_{\mathcal{B},\mathcal{B}} = \lambda I_n + [N]_{\mathcal{B},\mathcal{B}}$ is a diagonal block matrix whose diagonal blocks are all $\lambda I_e + J_{0,e}$, i.e., a Jordan block with eigenvalue $\lambda$. □

**Theorem 4.11.** *Let* $T : V \to V$ *be a linear operator with* $c_T(X) = (X-\lambda_1)^{e_1} \cdots (X-\lambda_s)^{e_s}$*. There exists an ordered basis* $\mathcal{B}$ *for* $V$ *such that* $[T]_{\mathcal{B},\mathcal{B}}$ *is a diagonal block matrix, whose diagonal blocks are*

$$(J_{\lambda_1,e_{1,1}}, \ldots, J_{\lambda_1,e_{1,m_1}}, J_{\lambda_2,e_{2,1}}, \ldots, J_{\lambda_2,e_{2,m_2}}, \ldots, J_{\lambda_s,e_{s,1}}, \ldots, J_{\lambda_s,e_{s,m_s}})$$

*where every for every* $i = 1, \ldots, s$*,* $e_{i,1} \geq \cdots \geq e_{i,m_i} \geq 1$ *and* $e_{i,1} + \cdots + e_{i,m_i} = e_i$*. The matrix* $[T]_{\mathcal{B},\mathcal{B}}$ *is unique up to reordering of* $(\lambda_1, \ldots, \lambda_s)$*, but the ordered basis* $\mathcal{B}$ *is typically not unique.*

*Proof.* Existence of an ordered basis $\mathcal{B}$ such that $[T]_{\mathcal{B},\mathcal{B}}$ is a diagonal matrix of Jordan blocks follows from Proposition 2.6 and Corollary 4.10. To compute the number of Jordan blocks $J_{\lambda_i,e}$, and form $N_i : E^{(e_i)}_{T,\lambda_i}$, define $E^{(e)}, F^{(e)} \subset E^{(e_i)}_{T,\lambda_i}$ as above. Then the number of $J_{\lambda_i,e}$-blocks is $\dim(E^{(e)}) - \dim(F^{(e)})$. This depends only on $T$, not on a choice of basis. Therefore the sequence of lengths of $J_{\lambda_i,e}$-blocks is canonically determined by $T$. Reorder the subbases putting $T$ into diagonal block form so that $e_{i,1} \geq \cdots \geq e_{i,m_i} \geq 1$ for every $i = 1, \dots, s$. □

**Example 4.12.** Let $\lambda \in \mathbb{F}$ and let $T = T_A : \mathbb{F}^3 \to \mathbb{F}^3$ where,

$$A = \begin{pmatrix} \lambda & 0 & 1 \\ 0 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}.$$

One basis putting $A$ into Jordan canonical form is the set of columns of the matrix,

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Another basis putting $A$ into Jordan canonical form is the set of columns of the matrix,

$$P_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

In both cases, $AP_i = P_iJ$, i.e., $A = P_iJP_i^{-1}$, where $J$ is the Jordan normal form,

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

The Jordan normal form $J$ is uniquely determined, but the change-of-basis matrices $P_1$ and $P_2$ are not uniquely determined.