Decomposition of Operators

Throughout V is a finite-dimensional vector space over a field F and T is a linear endomorphism of V. Recall that $A = \mathrm{End}(V)$ is an associative algebra with unity I_V over F under composition. The endomorphism T is *semisimple* if V is spanned by eigenvectors of T and is *nilpotent* if $T^m = 0$ for some $m \geq 0$. Observe that T is both semisimple and nilpotent if and only if $T = 0$.

A subspace W of V is T-*invariant* if $T(W) \subseteq W$. Suppose that S is an endomorphism of V which commutes with T. Then $\mathrm{Im}\, S$ and $\mathrm{Ker}\, S$ are T-invariant subspaces of V.

We will show that $T = S + N$ has a unique decomposition as the sum of commuting endomorphisms, where S is semisimple and N is nilpotent, if and only if $f(T) = 0$ for some $f(x) \in F[x]$ which splits into linear factors over F. This requires very little linear algebra and just a few very basic facts about polynomials.

1 Preliminaries

Let $F[x]$ be the algebra of polynomials in indeterminate x over F. Suppose that A is an associative algebra with unity over F. For all $a \in A$ there is an algebra map

$$\pi_a : F[x] \longrightarrow A$$

determined by $\pi_a(x) = a$. Thus the image of $f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_r x^r \in F[x]$ under π_a is

$$f(a) = \pi_a(f(x)) = \alpha_0 + \alpha_1 a + \cdots + \alpha_r a^r,$$

the result of substitution of a for x in $f(x)$. In particular, if $f(x) = g(x)h(x) + \ell(x)$ then $f(T) = g(T) \circ h(T) + \ell(T)$.

The LaGrange polynomials, see (1) below, play a basic role in the analysis of a linear endomorphism of a finite dimensional vector space over F. Let $r \geq 1$ and $\lambda_1, \ldots, \lambda_r \in F$ be distinct. For all $1 \leq i \leq r$ set

$$e_i(x) = \prod_{j \neq i} \left(\frac{x - \lambda_j}{\lambda_i - \lambda_j} \right) \tag{1}$$

and set $f(x) = (x - \lambda_1) \cdots (x - \lambda_r)$. Observe that the polynomial $e_i(x)$ has degree $r - 1$ and that

$$e_i(\lambda_j) = \delta_{i\,j} \tag{2}$$

for all $1 \leq i, j \leq r$. The preceding equation implies that $\{e_1(x), \ldots, e_r(x)\}$ is linearly independent. Consequently this set of polynomials is a basis for the r-dimensional subspace $\mathcal{V}$ of $F[x]$ consisting of all polynomials of degree at most $r - 1$. In particular $1 = a_1 e_1(x) + \cdots + a_r e_r(x)$ for some $a_1, \ldots, a_r \in F$. Substituting λ_i for x in the preceding equation we conclude that $a_i = 1$ for all $1 \leq i \leq r$ by (2). Thus

$$1 = e_1(x) + \cdots + e_r(x), \tag{3}$$

$$f(x) \quad \text{divides} \quad (x-\lambda_i)e_i(x), \quad \text{and therefore divides} \quad e_i(x)e_j(x) \tag{4}$$

whenever i and j are different.

Let $a_1, \ldots, a_r \in F$ and set $h(x) = a_1e_1(x) + \cdots + a_re_r(x)$. Then $h(\lambda_i) = a_i$ for all $1 \leq i \leq r$ which follows by (2).

One more bit of terminology concerning polynomials. We say that $g(x) \in F[x]$ *splits into linear factors over* F if $g(x) = a(x-\rho_1)\cdots(x-\rho_s)$ for some non-zero $a \in F$ and $\rho_1, \ldots, \rho_s \in F$ which are not necessarily distinct.

Exercise 1 Suppose that $\lambda_1, \ldots, \lambda_r \in F$ are distinct and $a_1, \ldots, a_r \in F$. Show that there is a unique polynomial $h(x) \in F[x]$ of degree at most $r-1$ such that $h(\lambda_i) = a_i$ for all $1 \leq i \leq r$. [Hint: Suppose $\ell(x) \in F[x]$ also satisfies the condition as well and consider the roots of the difference $d(x) = h(x) - \ell(x)$.]

2 Decompositions into Semisimple and Nilpotent Parts

Suppose that $g(T) = 0$, where $g(x) \in F[x]$ splits into a product of linear factors over F. Write $g(x) = a(x-\lambda_1)^{n_1}\cdots(x-\lambda_r)^{n_r}$, where $a \in F$ is not zero, $\lambda_1, \ldots, \lambda_r \in F$ are distinct, and $n_1, \ldots, n_r > 0$. Set $f(x) = (x-\lambda_1)\cdots(x-\lambda_r)$ and let n be equal to greater than the maximum of $n_1, \ldots, n_r$. Since $g(x)$ divides $f(x)^n$, it follows that $f(T)^n = 0$.

Let $e_1(x), \ldots, e_r(x)$ be defined by (1) and set $m = (n-1)r + 1$. Using (3) we calculate

$$1 = (e_1(x) + \cdots + e_r(x))^m = \sum_{1 \leq i_1, \ldots, i_m \leq r} e_{i_1}(x)\cdots e_{i_m}(x).$$

Each summand must have at least n factors which are $e_i(x)$ for some $1 \leq i \leq r$; otherwise $m \leq (n-1)r$. Thus there are $E_1(x), \ldots, E_r(x) \in F[x]$ which satisfy

$$1 = E_1(x) + \cdots + E_r(x), \tag{5}$$

$$f(x)^n \quad \text{divides} \quad (x-\lambda_i)^nE_i(x), \quad \text{and divides} \quad E_i(x)E_j(x) \tag{6}$$

whenever i and j are different. Substituting T for x we deduce from (5) and (6) that

$$\mathrm{I}_V = E_1(T) + \cdots + E_r(T) \quad \text{and} \quad E_i(T)\circ E_j(T) = 0 \tag{7}$$

whenever i and j are different. Thus from (7) we conclude that

$$E_i(T) = E_i(T)\circ\mathrm{I}_V = E_i(T)\circ E_1(T) + \cdots + E_i\circ E_r(T) = E_i(T)\circ E_i(T)$$

which means

$$E_i(T)\circ E_j(T) = \delta_{i\,j}E_i(T) \tag{8}$$

for all $1 \leq i, j \leq r$. Set

$$S = \lambda_1E_1(T) + \cdots + \lambda_rE_r(T).$$

We will show that S is semisimple. Let $v \in V$. Then $v = E_1(T)(v) + \cdots + E_r(T)(v)$ by (7). The calculation

$$S(E_i(T)(v)) = \lambda_1(E_1(T)\circ E_i(T))(v) + \cdots + \lambda_r(E_r(T)\circ E_i(T))(v) = \lambda_iE_i(T)(v),$$

which follows by (8), shows that V is spanned by eigenvectors of S. By definition S is semisimple. Since the $E_i(T)$'s are polynomials in T it follows that S is a polynomial in T.

Since $T = T{\circ}\mathrm{I}_V = T{\circ}E_1(T) + \cdots + T{\circ}E_r(T)$ by (7), the difference $N = T - S$ can be written

$$N = (T - \lambda_1 \mathrm{I}_V){\circ}E_1(T) + \cdots + (T - \lambda_r \mathrm{I}_V){\circ}E_r(T).$$

Now the $E_i(T)$'s commute with T since they are polynomials in T. Using (8) we calculate

$$N^n = (T - \lambda_1 \mathrm{I}_V)^n{\circ}E_1(T) + \cdots + (T - \lambda_r \mathrm{I}_V)^n{\circ}E_r(T).$$

Since $f(x)^n$ divides $(x - \lambda_i)^n E_i(x)$ by (6), and $f(T)^n = 0$, each of the summands in the preceding equation is zero. Therefore N is nilpotent. Observe that N is a polynomial in T since S is.

Theorem 1 *Let $T : V \longrightarrow V$ be a linear endomorphism of a finite-dimensional vector space V over the field F. Then:*

(a) *$T = S + N$, where S and N are commuting endomorphisms of V, S is semisimple, and N is nilpotent. Furthermore S and N are polynomials in T. These polynomials can be chosen to have constant term zero.*

(b) *Suppose that $T = S' + N'$, where S' and N' are commuting endomorphisms of V, S' is semisimple, and N' is nilpotent. Then $S = S'$ and $N = N'$.*

Proof: We have shown part (a), except for the last sentence. This is established in Exercise 3. We sketch the proof of part (b), leaving the details as an exercise for the reader.

Suppose that S' and N' satisfy the hypothesis of part (b). Since S' and N' commute with each other they commute with T and therefore with any polynomial in T. As a consequence S' and N' commute with S and N. Since the difference of commuting semisimple operators is semisimple, and the difference of commuting nilpotent operators is nilpotent, the difference $S - S' = N' - N$ is both semisimple and nilpotent. Therefore $S - S' = N' - N = 0$. $\square$

As a corollary of the proof of the preceding theorem:

Corollary 1 *Let $T : V \longrightarrow V$ be a linear endomorphism of a finite-dimensional vector space V over the field F. Then the following are equivalent:*

(a) *T is semisimple.*

(b) *V has a basis of eigenvectors of T.*

(c) *$f(T) = 0$ for some $f(x) \in F[x]$ which splits into a product of distinct linear factors.*

Proof: Since any spanning set of V contains a basis for V, part (a) implies part (b). To show that part (b) implies part (c), suppose that $\{v_1, \ldots, v_r\}$ is a basis for V consisting of eigenvectors for T and let $f(x) = (x - \lambda_1) \cdots (x - \lambda_s) \in F[x]$, where $\lambda_1, \ldots, \lambda_s \in F$ are the distinct eigenvalues for T. Let $1 \leq i \leq r$. Then $T(v_i) = \lambda_\ell v$, or $(T - \lambda_\ell \mathrm{I}_V)(v_i) = 0$, for some $1 \leq \ell \leq s$. Substituting T for x in $f(x)$ we obtain

$$f(T)(v_i) = \Big((T - \lambda_1 \mathrm{I}_V){\circ}\cdots{\circ}\widehat{(T - \lambda_\ell \mathrm{I}_V)}{\circ}\cdots{\circ}(T - \lambda_s \mathrm{I}_V)\Big)\,((T - \lambda_\ell \mathrm{I}_V)(v_i)) = 0,$$

where " $\widehat{\quad}$ " means factor omitted. Therefore $f(T)$ vanishes on a spanning set for V which means that $f(T) = 0$. We have shown part (b) implies part (c). The proof will be concluded once we show part (c) implies part (a).

Assume the hypothesis of part (c). Then $f(T) = 0$ for some $f(x) = (x-\lambda_1)\cdots(x-\lambda_r) \in F[x]$ where $\lambda_1, \ldots, \lambda_r \in F$ are distinct. We retrace the first part of the proof of Theorem 1.

Observe that we may take $n = 1$ in (6). Since since $f(T) = 0$ it follows that $(T-\lambda_i \mathrm{I}_V)\circ E_i(T) =$ for all $1 \leq i \leq r$. Let $v \in V$. The preceding equation implies that $E_i(T)(v)$ is an eigenvector of T belonging to λ_i. As $v = E_1(T)(v) + \cdots + E_r(T)(v)$ by (6), it follows that v is the sum of eigenvectors of T. Therefore T is semisimple by definition. $\square$

Suppose that $v \in V$ is an eigenvector for T belonging to λ. Since $T^\ell(v) = \lambda^\ell v$ for all $\ell \geq 0$ it follows that $f(T)(v) = f(\lambda)v$ for all $f(x) \in F[x]$. Therefore $f(T) = 0$ implies that the eigenvalues of T are roots of $f(x)$.

Let W be a T-invariant subspace of V. Then W is $f(T)$-invariant for all $f(x) \in F[x]$ and $f(T|_W) = f(T)|_W$ holds for the restrictions. Let $\overline{T} : V/W \longrightarrow V/W$ be the endomorphism defined by $\overline{T}(v + W) = T(v) + W$ for all $v \in V$. Note that $f(\overline{T}) = \overline{f(T)}$ for all $f(x) \in F[x]$ as well. Thus if T is semisimple, by virtue of the preceding corollary the restriction $T|_W : W \longrightarrow W$ and the induced endomorphism $\overline{T} : V/W \longrightarrow V/W$ are semisimple also.

Exercise 2 Suppose that $f(T) = 0$, where $f(x) = (x-\lambda_1)^{n_1}\cdots(x-\lambda_r)^{n_r}$, $n_1, \ldots, n_r > 0$, and $\lambda_1, \ldots, \lambda_r \in F$ are distinct. Let $E_i(x)$ satisfy (5) and (6), set $V_i = \operatorname{Im} E_i(T)$, and let S be as in part a) of Theorem 1. Show that:

(a) $V = V_1 \oplus \cdots \oplus V_r$;

(b) $V_i = \operatorname{Ker}(T - \lambda_i \mathrm{I}_V)^m$ for all $m \geq n_i$, $S(v) = \lambda_i v$ for all $v \in V_i$; and

c) When T is semisimple the non-zero V_i's are the eigenspaces for T.

[Hint: For part (b) observe that $(x - \lambda_j)^n$ divides $E_i(x)$ whenever $j \neq i$.]

Exercise 3 Show that the polynomials mentioned in part (a) of Theorem 1 can be chosen to have constant term zero. [Hint: We may assume that $\lambda_r = 0$. In this case $f(x)$ has constant term zero and therefore $(x - \lambda_i)E_i(x)$ does as well for all $1 \leq i \leq r$ by (6).]

Exercise 4 Show that T is semisimple if and only if its minimal polynomial splits into a product of distinct linear factors over F.

3 A Necessary and Sufficient Condition for the Decomposition

We have shown that if $f(T) = 0$, where $f(x) \in F[x]$ splits into linear factors over F, then T is the sum of commuting operators, one of which is semisimple and the other nilpotent. This vanishing condition is necessary.

Proposition 1 *Let $T : V \longrightarrow V$ be a linear endomorphism of a finite-dimensional vector space V over the field F. Then the following are equivalent:*

(a) $f(T) = 0$ *for some polynomial* $f(x) \in F[x]$ *which splits into linear factors over* F.

(b) $T = S + N$, *where* S, T *are commuting endomorphisms of* V, S *is semisimple, and* N *is nilpotent.*

PROOF: We need only show that part (b) implies part (a). Let $\lambda_1, \dots, \lambda_r \in F$ list the distinct eigenvalues of S and let $V_i = \mathrm{Ker}(S - \lambda_i I)$ be the S-invariant subspace of all eigenvectors of S belonging to λ_i for all $1 \le i \le r$. Since S and N commute, S and T commute. In particular $T(V_i) \subseteq V_i$ for all $1 \le i \le r$.

Since N is nilpotent $N^n = 0$ for some $n > 0$. Let $1 \le i \le r$ and $v \in V_i$. Since $T(V_i) \subseteq V_i$ and $S|_{V_i} = \lambda_i \mathrm{I}_V|_{V_i}$, for $v \in V_i$ we can easily deduce

$$0 = N^n(v) = (T - S)^n(v) = (T - \lambda_i \mathrm{I}_V)^n(v).$$

Thus $g(T)^n(v) = 0$, where $g(x) = (x - \lambda_1) \cdots (x - \lambda_r)$. Since $V = V_1 + \cdots + V_r$ necessarily $g(T)^n = 0$. Take $f(x) = g(x)^n$. $\square$

Exercise 5 Show that $T = S + N$, where S, T are commuting endomorphisms of V, S is semisimple, and N is nilpotent, if and only if the characteristic polynomial of T over F splits into linear factors over F. [Hint: The characteristic polynomial divides a power of the minimal polynomial.]

4 When V is an algebra and T is a Derivation

Suppose that A is a finite-dimensional algebra over F and D is a derivation of A with a decomposition into the sum of commuting semisimple and nilpotent endomorphisms. These summands are derivations also. Two preliminary lemmas will explain the essential reasons for this.

For $\lambda \in F$ let $V_{(\lambda)} = \bigcup_{\ell=0}^{\infty} \mathrm{Ker}(D - \lambda \mathrm{I}_V)^\ell$. Since

$$\mathrm{Ker}(D - \lambda \mathrm{I}_V) \subseteq \mathrm{Ker}(D - \lambda \mathrm{I}_V)^2 \subseteq \mathrm{Ker}(D - \lambda \mathrm{I}_V)^3 \subseteq \dots$$

is a chain of subspaces of V, their union $V_{(\lambda)}$ is a subspace of V. Observe that the sum $\sum_{\lambda \in F} V_{(\lambda)}$ is direct.

Lemma 1 *Let* D *be a derivation of an algebra* A *over* F *and suppose that* $A = \bigoplus_{\lambda \in F} A_{(\lambda)}$. *Then* $A_{(\lambda)} A_{(\rho)} \subseteq A_{(\lambda+\rho)}$ *for all* $\lambda, \rho \in F$.

PROOF: Let $a, b \in A$. Here it is convenient to regard the product on A as a linear map $m : A \otimes_F A \longrightarrow A$; thus $m(a \otimes b) = ab$. Now let $\lambda, \rho \in F$. Since δ is a derivation of A we calculate

$$D(ab) - (\lambda + \rho)ab = aD(b) + D(a)b - (\lambda + \rho)ab = a((D - \rho \mathrm{I}_A)(b)) + ((D - \lambda \mathrm{I}_A)(a))b.$$

Thus

$$(D - (\lambda + \rho)\mathrm{I}_A) \circ m = m \circ (\mathrm{I}_A \otimes (D - \rho \mathrm{I}_A) + (D - \lambda \mathrm{I}_A) \otimes \mathrm{I}_A).$$

Now the summands in the right hand side of the preceding equation are commuting elements in the algebra $\mathrm{End}(A) \otimes \mathrm{End}(A)$. Therefore by the binomial theorem

$$\begin{aligned} (D - (\lambda + \rho)\mathrm{I}_A)^\ell \circ m &= m \circ (\mathrm{I}_A \otimes (D - \rho \mathrm{I}_A) + (D - \lambda \mathrm{I}_A) \otimes \mathrm{I}_A)^\ell \\ &= m \circ \left(\sum_{i=0}^{\ell} \binom{\ell}{i} (D - \lambda \mathrm{I}_A)^{\ell - i} \otimes (D - \rho \mathrm{I}_V)^i \right) \end{aligned}$$

for all $\ell \geq 0$.

Suppose that $a \in A_{(\lambda)}$ and $b \in A_{(\rho)}$. Then $(D - \lambda \mathrm{I}_A)^{\ell'}(a) = 0 = (D - \rho \mathrm{I}_A)^{\ell''}(b)$ for some $\ell', \ell'' > 0$. Let $\ell = \ell' + \ell'' - 1$ and $0 \leq i \leq \ell$. Then $\ell \geq 0$ and either $\ell' \leq \ell - i$ or $\ell'' \leq i$. Therefore one of $(D - \lambda \mathrm{I}_V)^{\ell - i}(a)$ or $(D - \rho \mathrm{I}_V)^{i}(b)$ is zero. Using the preceding calculation we deduce

$$(D - (\lambda + \rho)\mathrm{I}_A)^{\ell}(ab) = \sum_{i=0}^{\ell} \binom{\ell}{i} (D - \lambda \mathrm{I}_A)^{\ell - i}(a)(D - \rho \mathrm{I}_A)^{i}(b) = 0$$

which means $ab \in A_{(\lambda + \rho)}$. □

Lemma 2 *Let $A = \bigoplus_{\lambda \in F} A(\lambda)$ be an algebra over F which is the direct sum of subspaces $A(\lambda)$ such that $A(\lambda)A(\rho) \subseteq A(\lambda + \rho)$ for all $\lambda, \rho \in F$. Then the endomorphism S of A defined by $S(a) = \lambda a$ for all $\lambda \in F$ and $a \in A_\lambda$ is a derivation of A.*

Proof: We need only that $S(ab) = aS(b) + S(a)b$ holds for all a, b in some spanning set of A. Let $a \in A(\lambda)$ and $b \in A(\rho)$. Since $ab \in A(\lambda + \rho)$ by assumption, $S(ab) = (\lambda + \rho)ab$. On the other hand $S(a)b + aS(b) = (\lambda a)b + a(\rho b) = (\lambda + \rho)ab$. □

Proposition 2 *Suppose that $D : A \longrightarrow A$ is a derivation of a finite-dimensional algebra over the field F and $D = S + N$, where S, N are commuting endomorphisms of A, S is semisimple and N is nilpotent. Then S and N are derivations of A.*

Proof: Since $\mathrm{Der}(A)$ is a subspace of $\mathrm{End}(A)$ we need only show that S is derivation of A. Using Proposition 1 we see there is a polynomial $f(x) = (x - \lambda_1) \cdots (x - \lambda_r) \in F[x]$, where $\lambda_, \ldots, \lambda_r \in F$ are distinct, and $n > 0$ such that $f(T)^n = 0$. By parts (a) and (b) of Exercise 2 we conclude that $A = A_{(\lambda_1)} \oplus \cdots \oplus A_{(\lambda_r)}$ and $S(a) = \lambda_i a$ for all $1 \leq i \leq r$ and $a \in A_{(\lambda_i)}$.

Suppose that $\lambda \in F$ is not one of $\lambda_1, \ldots, \lambda_r$. Then $\mathsf{f}(T)^n = 0$, where

$$\mathsf{f}(x) = f(x)(x - \lambda) = (x - \lambda_1) \cdots (x - \lambda_r)(x - \lambda).$$

By parts (a) and (b) of Exercise 2 again we have

$$A = A_{(\lambda_1)} \oplus \cdots \oplus A_{(\lambda_r)} \oplus A_{(\lambda)} = A \oplus A_{(\lambda)}.$$

Therefore $A_{(\lambda)} = (0)$. We have shown that $A = \bigoplus_{\lambda \in F} A_{(\lambda)}$ and $S(a) = \lambda a$ for all $\lambda \in F$ and $a \in A_{(\lambda)}$. Thus S is a derivation of A by Lemmas 1 and 2. □