EIGENSYSTEMS OF A LINEAR OPERATORS

BRANKO ĆURGUS

1. Algebra of linear operators

In this section we consider a vector space $\mathscr{V}$ over $\mathbb{F}$. Here $\mathbb{F}$ is either $\mathbb{F}$ or $\mathbb{C}$. For the most important results we assume that $\mathbb{F} = \mathbb{C}$. We do not have time to develop the theory for $\mathbb{R}$. By $\mathscr{L}(\mathscr{V})$ we denote the vector space $\mathscr{L}(\mathscr{V}, \mathscr{V})$ of all linear operators on $\mathscr{V}$. The vector space $\mathscr{L}(\mathscr{V})$ with the composition of operators as an additional binary operation is an algebra in the sense of the following definition.

Definition 1.1. A vector space $\mathscr{A}$ over a field $\mathbb{F}$ is an *algebra* over $\mathbb{F}$ if the following conditions are satisfied:

ME. There exist a binary operation $\cdot : \mathscr{A} \times \mathscr{A} \to \mathscr{A}$.
MA. (*associativity*) For all $x, y, z \in \mathscr{A}$ we have $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
MD. (*right-distributivity*) For all $x, y, z \in \mathscr{A}$ we have $(x+y)\cdot z = x\cdot z+y\cdot z$.
MD. (*left-distributivity*) For all $x, y, z \in \mathscr{A}$ we have $z\cdot(x+y) = z\cdot x+z\cdot y$.
MS. (*respect for scaling*) For all $x, y \in \mathscr{A}$ and all $\alpha \in \mathbb{F}$ we have $\alpha(x\cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y)$.

This binary operation in an algebra is often referred to as *multiplication*. As usual with multiplication we drop the dot, and just write xy instead of $x \cdot y$. $\Diamond$

The multiplicative identity in the algebra $\mathscr{L}(\mathscr{V})$ is the identity operator $I_{\mathscr{V}}$.

For $T \in \mathscr{L}(\mathscr{V})$ we recursively define nonnegative integer powers of T by $T^0 = I_{\mathscr{V}}$ and for all $n \in \mathbb{N}$ we define $T^n = T \circ T^{n-1}$.

For $T \in \mathscr{L}(\mathscr{V})$, set

$$\mathscr{A}_T = \operatorname{span}\Big\{T^k : k \in \mathbb{N} \cup \{0\}\Big\}.$$

Clearly $\mathscr{A}_T$ is a subspace of $\mathscr{L}(\mathscr{V})$. Moreover, we will see below that $\mathscr{A}_T$ is a commutative subalgebra of $\mathscr{L}(\mathscr{V})$.

Recall that by the definition of a span, a nonzero $S \in \mathscr{L}(\mathscr{V})$ belongs to $\mathscr{A}_T$ if and only if there exist $m \in \mathbb{N} \cup \{0\}$ and scalars $\alpha_0, \alpha_1, \ldots, \alpha_m \in \mathbb{F}$ such that $a_m \neq 0$ and

$$S = \sum_{k=0}^{m} \alpha_k T^k. \qquad (1) \quad \texttt{eq-lcTs}$$

Date: February 20, 2025 at11:47.

The last expression reminds us of a polynomial over $\mathbb{F}$. Recall that by $\mathbb{F}[z]$ we denote the vector space of all polynomials over $\mathbb{F}$. That is

$$\mathbb{F}[z] = \left\{ \sum_{j=0}^{n} \alpha_j z^j \;:\; n \in \mathbb{N} \cup \{0\},\ (\alpha_0, \ldots, \alpha_n) \in \mathbb{F}^{n+1} \right\}.$$

The multiplication in $\mathbb{F}[z]$ is introduced in the following definition.

def-mp **Definition 1.2** (5.16 page 138 in the textbook)**.** Let $m, n \in \mathbb{N} \cup \{0\}$ and

$$p(z) = \sum_{i=0}^{m} \alpha_i z^i \in \mathbb{F}[z] \quad \text{and} \quad q(z) = \sum_{j=0}^{n} \beta_j z^j \in \mathbb{F}[z]. \tag{2}$$ eq-pq

We set

$$(pq)(z) \stackrel{\text{def}}{=} \sum_{k=0}^{m+n} \left(\sum_{(i,j)\in \mathbb{I}_k} \alpha_i \beta_j \right) z^k,$$

where, for all $k \in \{0, \ldots, m+n\}$, we set

$$\mathbb{I}_k = \Big\{ (i,j) \in \{0, \ldots, m\} \times \{0, \ldots, n\} : i + j = k \Big\}.$$

$\Diamond$

It is straightforward to verify that the multiplication in Definition 1.2 satisfies the axioms **ME**, **MA**, **MD**, **MD**, and **MS**. Hence, $\mathbb{F}[z]$ is an algebra over $\mathbb{F}$. Since the multiplication in $\mathbb{F}$ is commutative, it follows that $pq = qp$ for all $p, q \in \mathbb{F}[z]$. That is, $\mathbb{F}[z]$ is a commutative algebra.

The clear alikeness of the expression (1) and the expression for the polynomial p in (2) is the motivation for the following definition. For a fixed $T \in \mathscr{L}(\mathscr{V})$ we define

$$\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$$

by setting

$$\Xi_T(p) = \sum_{i=0}^{m} \alpha_i T^i \qquad \text{where} \qquad p(z) = \sum_{i=0}^{m} \alpha_i z^i \in \mathbb{F}[z]. \tag{3}$$ eq-Xi

It is common to write $p(T)$ for $\Xi_T(p)$.

th-Xah **Theorem 1.3** (5.17 page 138 in the textbook)**.** *Let $T \in \mathscr{L}(\mathscr{V})$. The function $\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$ defined in* (3) *is an algebra homomorphism. The range of Ξ_T is $\mathscr{A}_T$.*

Proof. It is straightforward to prove that $\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$ is linear. We will prove that $\Xi_T : \mathbb{F}[z] \to \mathscr{L}(\mathscr{V})$ is multiplicative. That is, for all $p, q \in \mathbb{F}[z]$ we have $\Xi_T(pq) = \Xi_T(p)\Xi_T(q)$. To prove this, let $p, q \in \mathbb{F}[z]$ be arbitrary and given in (2). Then, first applying the by definition in (3),

$$\Xi_T(p)\Xi_T(q) = \left(\sum_{i=0}^{m} \alpha_i T^i \right) \left(\sum_{j=0}^{n} \beta_j T^j \right)$$

$$\begin{aligned}
\boxed{\mathscr{L}(\mathscr{V}) \text{ is an algebra}} &= \sum_{i=0}^{m}\sum_{j=0}^{n} \alpha_i \beta_j T^{i+j} \\
\boxed{\mathscr{L}(\mathscr{V}) \text{ is a vector space}} &= \sum_{k=0}^{m+n} \Biggl(\sum_{(i,j)\in \mathbb{I}_k} \alpha_i \beta_j \Biggr) T^k \\
\boxed{\text{by Definition 1.2}} &= \Xi_T(pq).
\end{aligned}$$

This proves the multiplicative property of Ξ_T.

The fact that $\mathscr{A}_T$ is the range of Ξ_T is straightforward. □

Corollary 1.4. *Let $T \in \mathscr{L}(\mathscr{V})$. The subspace $\mathscr{A}_T$ of $\mathscr{L}(\mathscr{V})$ is a commutative subalgebra of $\mathscr{L}(\mathscr{V})$.*

Proof. Let $Q, S \in \mathscr{A}_T$. Since $\mathscr{A}_T$ is the range of Ξ_T there exist $p, q \in \mathbb{F}[z]$ such that $Q = \Xi_T(p)$ and $S = \Xi_T(q)$. Then, since Ξ_T is an algebra homomorphism we have

$$QS = \Xi_T(p)\Xi_T(q) = \Xi_T(pq) = \Xi_T(qp) = \Xi_T(q)\Xi_T(p) = SQ.$$

This sequence of equalities shows that $QS \in \operatorname{ran}(\Xi_T) = \mathscr{A}_T$ and $QS = SQ$. That is $\mathscr{A}_T$ is closed with respect to the operator composition and the operator composition on $\mathscr{A}_T$ is commutative. □

co-linfact **Corollary 1.5.** *Let $\mathscr{V}$ be a complex vector space and let $T \in \mathscr{L}(\mathscr{V})$ be a nonzero operator. Then for every $p \in \mathbb{C}[z]$ such that $m = \deg p \geq 1$ there exist a nonzero $\alpha \in \mathbb{C}$ and $z_1, \ldots, z_m \in \mathbb{C}$ such that*

$$\Xi_T(p) = p(T) = \alpha(T - z_1 I) \cdots (T - z_m I).$$

Proof. Let $p \in \mathbb{C}[z]$ such that $m = \deg p \geq 1$. Then there exist $\alpha_0, \ldots, \alpha_m \in \mathbb{C}$ such that $\alpha_m \neq 0$ such that

$$p(z) = \sum_{k=0}^{m} \alpha_j z^j.$$

By the Fundamental Theorem of Algebra there exist nonzero $\alpha \in \mathbb{C}$ and $z_1, \ldots, z_m \in \mathbb{C}$ and

$$p(z) = \alpha(z - z_1) \cdots (z - z_m).$$

Here $\alpha = \alpha_m$ and $z_1, \ldots, z_m$ are the roots of p. Since by Theorem 1.3 the operator Ξ_T is an algebra homomorphism, we have

$$\begin{aligned}
p(T) &= \Xi_T(p) \\
&= \Xi_T\bigl(\alpha(z - z_1) \cdots (z - z_m)\bigr) \\
&= \alpha\, \Xi_T\bigl((z - z_1)\bigr) \cdots \Xi_T\bigl((z - z_m)\bigr) \\
&= \alpha(T - z_1 I) \cdots (T - z_m I).
\end{aligned}$$

This completes the proof. □

2. Existence of an eigenvalue

We will need the following lemma about injections.

le-inj-n **Lemma 2.1.** *Let $n \in \mathbb{N}$, let A be a nonempty set and let $f_1, \dots, f_n \in A^A$. If $f_k : A \to A$ is an injection for all $k \in \{1, \dots, n\}$, then the composition $f_1 \circ \cdots \circ f_n$ is also an injection.*

Proof. We proceed by Mathematical Induction. The base step is trivial. It is useful to prove the implication for $n = 2$. Assume that $f, g \in A^A$ are injections. Let $s, t \in A$ be such that $s \neq t$. Then, since g is an injection, $g(s) \neq g(t)$. Since f is injective, $f(g(x)) \neq f(g(t))$. Thus, $f \circ g$ is injective.

Next we prove the inductive step. Let $m \in \mathbb{N}$ and assume that $f_1 \circ \cdots \circ f_m$ is an injection whenever $f_1, \dots, f_m \in A^A$ are all injections. (This is the inductive hypothesis.) Now assume that $f_1, \dots, f_m, f_{m+1} \in A^A$ are all injections. By the inductive hypothesis the function $f = f_1 \circ \cdots \circ f_m$ is an injection. Since by assumption $g = f_{m+1}$ is an injection, the already proved claim for $n = 2$ yields that

$$f \circ g = f_1 \circ \cdots \circ f_m \circ f_{m+1}$$

is an injection. This completes the proof. $\square$

Definition 2.2. Let $\mathscr{V}$ be a vector space over $\mathbb{F}$, $T \in \mathscr{L}(\mathscr{V})$. A scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of T if there exists $v \in \mathscr{V}$ such that $v \neq 0$ and $Tv = \lambda v$. The subspace $\operatorname{nul}(T - \lambda I)$ of $\mathscr{V}$ is called the *eigenspace* of T corresponding to λ. $\lozenge$

Definition 2.3. Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$. Let $T \in \mathscr{L}(\mathscr{V})$. The set of all eigenvalues of T is denoted by $\sigma(T)$. It is called the *spectrum* of T. $\lozenge$

th-ev-ex **Theorem 2.4** (5.19 page 143 in the textbook)**.** *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{C}$. Let $T \in \mathscr{L}(\mathscr{V})$. Then there exists a $\lambda \in \mathbb{C}$ and $v \in \mathscr{V}$ such that $v \neq 0_{\mathscr{V}}$ and $Tv = \lambda v$.*

Proof. The claim of the theorem is trivial if $T = 0_{\mathscr{L}(\mathscr{V})}$. So, assume that $T \in \mathscr{L}(\mathscr{V})$ is a nonzero operator.

Let $n = \dim \mathscr{V}$ and let $u \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$. Now consider the vectors

$$u, Tu, T^2u, \dots, T^n u. \tag{4}$$ eq-uTu

If two of these vectors coincide, say $k, l \in \{0, \dots, n\}$, $k < l$ are such that $T^k u = T^l u$, setting $\alpha_j = 0$ for $j \in \{0, \dots, n\} \setminus \{k, l\}$ and $\alpha_k = 1$ and $\alpha_l = -1$ we obtain a nontrivial linear combination of the vectors in (4).

If the vectors in (4) are distinct, since $n = \dim \mathscr{V}$, it follows from the Steinitz Exchange Lemma that the vectors in (4) are linearly dependent.

Hence, in either case, there exist $\alpha_0, \dots, \alpha_n \in \mathbb{C}$ and $k \in \{0, \dots, n\}$ such that

$$\alpha_0 u + \alpha_1 Tu + \alpha_2 T^2 u + \cdots + \alpha_n T^n u = 0_{\mathscr{V}} \quad \text{and} \quad \alpha_k \neq 0. \tag{5}$$ eq-lin-com

Since $u \neq 0_{\mathscr{V}}$ it is not possible that $\alpha_j = 0$ for all $j \in \{1, \dots, n\}$. Therefore, there exists $k \in \{1, \dots, n\}$ such that $\alpha_k \neq 0$.

Set
$$p(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \cdots + \alpha_n z^n.$$
Since there exists $k \in \{1, \dots, n\}$ such that $\alpha_k \neq 0$, we have that $m = \deg p \geq k > 0$.

Thus we have constructed a polynomial p of positive degree for which, by (5),
$$p(T)u = 0_{\mathscr{V}} \quad \text{with} \quad u \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}.$$

By the Fundamental Theorem of Algebra there exist $\alpha \neq 0$ and $z_1, \dots, z_m \in \mathbb{C}$ such that
$$p(z) = \alpha(z - z_1) \cdots (z - z_m).$$
Here $\alpha = \alpha_m$ and $z_1, \dots, z_m$ are the roots of p.

Since Ξ_T is an algebra homomorphism we have
$$\begin{aligned} p(T) &= \Xi_T(p) \\ &= \alpha\, \Xi_T(z - z_1) \,\cdots\, \Xi_T(z - z_m) \\ &= \alpha(T - z_1 I) \,\cdots\, (T - z_m I). \end{aligned}$$

Equality (5) yields that the operator $p(T)$ is not an injection. Lemma 2.1 now implies that there exists $j \in \{1, \dots, m\}$ such that $T - z_j I$ is not injective. That is, there exists $v \in \mathscr{V}$, $v \neq 0_{\mathscr{V}}$ such that
$$(T - z_j I)v = 0.$$
Setting $\lambda = z_j$ completes the proof. □

le-poatev **Lemma 2.5.** *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{F}$, let $T \in \mathscr{L}(\mathscr{V})$ and $Tv = \lambda v$ with $\lambda \in \mathbb{F}$ and $v \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$. For $p \in \mathbb{F}[z]$ we have*
$$p(T)v = p(\lambda)v.$$
That is, if λ is an eigenvalue of T with a corresponding eigenvector v, then $p(\lambda)$ is an eigenvalue of $p(T)$ with the same eigenvector v.

Proof. The equality is obvious if the polynomial p is constant. Assume that $\deg p = m \in \mathbb{N}$ and let
$$p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_m z^m \tag{6}$$ eq-pcal

with $\alpha_0, \dots, \alpha_m \in \mathbb{F}$. Let $Tv = \lambda v$ with $\lambda \in \mathbb{F}$ and $v \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$. Then for every $k \in \mathbb{N}$ we have
$$T^k v = T^{k-1}(Tv) = T^{k-1}(\lambda v) = \lambda T^{k-1} v = \cdots = \lambda^k v.$$
Further, we calculate, starting with the definition of $p(T)$,
$$\begin{aligned} p(T)v &= \big(\Xi_T(p)\big)v \\ \boxed{\text{definition of } \Xi} &= \left(\sum_{k=0}^{m} \alpha_k T^k\right) v \end{aligned}$$

$$
\begin{aligned}
\boxed{\mathscr{L}(\mathscr{V}) \text{ is a vector space}} &= \sum_{k=0}^{m} \alpha_k (T^k v) \\
\boxed{\text{definition of } T^k \text{ and } Tv = \lambda v} &= \sum_{k=0}^{m} (\alpha_k \lambda^k) v \\
\boxed{\text{axioms of } \mathscr{V}} &= \left(\sum_{k=0}^{m} \alpha_k \lambda^k \right) v \\
\boxed{\text{by (6)}} &= p(\lambda) v.
\end{aligned}
$$

Thus $p(T)v = p(\lambda)v$, that is $p(\lambda)$ an eigenvalue of $p(T)$. ☐

In the preceding lemma, we proved that $p\bigl(\sigma(T)\bigr) \subseteq \sigma\bigl(p(T)\bigr)$. The converse inclusion does not hold when $\mathbb{F} = \mathbb{R}$. For example, let T be the rotation by $\pi/2$ in $\mathbb{R}^2$. Then $T^2 = -I_2$, so

$$p\bigl(\sigma(T)\bigr) = p\bigl(\emptyset\bigr) = \emptyset \subset \{-1\} = \sigma\bigl(-I_2\bigr).$$

However, the equality $p\bigl(\sigma(T)\bigr) = \sigma\bigl(p(T)\bigr)$ does hold when $\mathbb{F} = \mathbb{C}$, as established by the Spectral Mapping Theorem, which follows.

th-smt **Theorem 2.6.** *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{C}$, let $T \in \mathscr{L}(\mathscr{V})$ and $p \in \mathbb{C}[z]$. Then*

$$\sigma\bigl(p(T)\bigr) = p\bigl(\sigma(T)\bigr).$$

Proof. The equality is obvious if the polynomial p is constant. Assume that $\deg p = m \in \mathbb{N}$ and let the coefficients of p be $\alpha_0, \cdots, \alpha_m \in \mathbb{C}$.

To prove $\sigma\bigl(p(T)\bigr) \subseteq p\bigl(\sigma(T)\bigr)$, let $\mu \in \sigma\bigl(p(T)\bigr)$ be arbitrary. Then there exists $w \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ such that

$$p(T)w = \mu w.$$

Set $q(z) = p(z) - \mu$. Then $q(T)w = 0_{\mathscr{V}}$ and since $w \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ the operator $q(T)$ is not an injection. By the Fundamental Theorem of Algebra there exist $\alpha, z_1, \ldots, z_m \in \mathbb{C}$ such that $\alpha \neq 0$ and

$$q(z) = \alpha(z - z_1) \cdots (z - z_m).$$

Since Ξ_T is an algebra homomorphism we have

$$
\begin{aligned}
q(T) &= \Xi_T(q) \\
&= \alpha\, \Xi_T(z - z_1) \,\cdots\, \Xi_T(z - z_m) \\
&= \alpha(T - z_1 I) \,\cdots\, (T - z_m I).
\end{aligned}
$$

That is, $q(T)$ is a composition of $m + 1$ operators. Since $q(T)$ is not an injection, Lemma 2.1 yields that there exists $k \in \{1, \ldots, m\}$ such that the operator $T - z_k I$ is not an injection. This implies that $z_k \in \sigma(T)$. Set $\lambda = z_k \in \sigma(T)$. Then $q(\lambda) = 0$, that is $p(\lambda) - \mu = 0$. Thus we have proved that for arbitrary $\mu \in p\big(\sigma(T)\big)$ there exists $\lambda \in \sigma(T)$ such that $\mu = p(\lambda)$. This proves

$$\sigma\big(p(T)\big) \subseteq p\big(\sigma(T)\big). \qquad \square$$

Using the method of the proof of the preceding theorem one can prove.

Proposition 2.7. *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{C}$ with $n = \dim \mathscr{V}$ and let $T \in \mathscr{L}(\mathscr{V})$. Then $\sigma(T) = \{0\}$ if and only if $T^n = 0_{\mathscr{L}(\mathscr{V})}$.*

The next theorem can be stated in English simply as: Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Theorem 2.8 (5.11 page 136 in the textbook)**.** *Let $\mathscr{V}$ be a vector space over $\mathbb{F}$, $T \in \mathscr{L}(\mathscr{V})$ and $n \in \mathbb{N}$. Assume*

(a) *$\lambda_1, \dots, \lambda_n \in \mathbb{F}$ are such that $\lambda_i \neq \lambda_j$ for all $i, j \in \{1, \dots, n\}$ such that $i \neq j$,*

(b) *$v_1, \dots, v_n \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ are such that $Tv_k = \lambda_k v_k$ for all $k \in \{1, \dots, n\}$.*

Then $\{v_1, \dots, v_n\}$ is linearly independent.

Proof. Let $\lambda_1, \dots, \lambda_n$ be distinct eigenvalues of T and let $v_1, \dots, v_n$ be corresponding eigenvectors:

$$Tv_k = \lambda_k v_k, \quad \text{for all} \quad k \in \{1, \dots, n\}. \tag{7}$$

For each $k \in \{1, \dots, n\}$ define the polynomial

$$q_k(z) = \prod \Big\{ (z - \lambda_j) \,:\, j \in \{1, \dots, n\} \setminus \{k\} \Big\}.$$

Then q_k has exactly $n-1$ distinct roots $\{\lambda_1, \dots, \lambda_n\} \setminus \{\lambda_k\}$ and

$$q_k(\lambda_k) = \prod \Big\{ (\lambda_k - \lambda_j) \,:\, j \in \{1, \dots, n\} \setminus \{k\} \Big\} \neq 0.$$

That is

$$q_k(\lambda_j) = \begin{cases} 0 & j \neq k, \\ q_k(\lambda_k) \neq 0 & j = k, \end{cases} \qquad \text{for all} \quad j, k \in \{1, \dots, n\}. \tag{8}$$

eq-qk1

By Lemma 2.5 we have

$$q_k(T)v_j = q_k(\lambda_j)v_j \qquad \text{for all} \quad j, k \in \{1, \dots, n\}. \tag{9}$$

eq-qk2

Now we are ready to prove the linear independence of $v_1, \dots, v_n$. Let $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ be such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_{\mathscr{V}}. \tag{10}$$

eq-Vaughn-li

Let $k \in \{1, \dots, n\}$ be arbitrary. Apply the operator $q_k(T)$ to both sides of (10) to obtain

$$\alpha_1 q_k(T) v_1 + \cdots + \alpha_n q_k(T) v_n = 0_{\mathscr{V}}. \tag{11}$$

eq-Vaughn-li1

By (9) we have

$$\alpha_1 q_k(\lambda_1) v_1 + \cdots + \alpha_n q_k(\lambda_n) v_n = 0_{\mathscr{V}}.$$

By (8) the last equality simplifies to

$$\alpha_k q_k(\lambda_k) v_k = 0_{\mathscr{V}}.$$

Since $v_k \neq 0_{\mathscr{V}}$ and $q_k(\lambda_k) \neq 0$ we deduce

$$\alpha_k = 0.$$

Since $k \in \{1, \dots, n\}$ was arbitrary the proof of linear independence is complete. □

Corollary 2.9. *Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$ and let $T \in \mathscr{L}(\mathscr{V})$. Then T has at most $n = \dim \mathscr{V}$ distinct eigenvalues.*

Proof. Let $\mathscr{B}$ be a basis of $\mathscr{V}$ where $\mathscr{B} = \{u_1, ..., u_n\}$. Then $\#\mathscr{B} = n$ and $\operatorname{span} \mathscr{B} = \mathscr{V}$. Let $\mathscr{C} = \{v_1, ..., v_m\}$ be eigenvectors corresponding to m distinct eigenvalues. Then $\mathscr{C}$ is a linearly independent set with $\#\mathscr{C} = m$. By the Steinitz Exchange Lemma, $m \leq n$. Consequently, T has at most n distinct eigenvalues. □

3. Existence of an upper-triangular matrix representation

Definition 3.1. A matrix $A \in \mathbb{F}^{n \times n}$ with entries a_{ij}, $i, j \in \{1, \dots, n\}$ is called *upper triangular* if $a_{ij} = 0$ for all $i, j \in \{1, \dots, n\}$ such that $i > j$. ♢

Definition 3.2. Let $\mathscr{V}$ be a vector space over $\mathbb{F}$ and $T \in \mathscr{L}(\mathscr{V})$. A subspace $\mathscr{U}$ of $\mathscr{V}$ is called an *invariant subspace* under T if $T(\mathscr{U}) \subseteq \mathscr{U}$. ♢

The following proposition is straightforward.

Proposition 3.3. *Let $S, T \in \mathscr{L}(\mathscr{V})$ be such that $ST = TS$. Then each eigenspaces of S is invariant under T and each eigenspace of T is invariant under S.*

Definition 3.4. Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$ with $n = \dim \mathscr{V} \in \mathbb{N}$. Let $T \in \mathscr{L}(\mathscr{V})$. A sequence of nontrivial subspaces $\mathscr{U}_1, \dots, \mathscr{U}_n$ of $\mathscr{V}$ such that

$$\mathscr{U}_1 \subsetneq \mathscr{U}_2 \subsetneq \cdots \subsetneq \mathscr{U}_n \tag{12}$$

eq-fan-sss

and

$$T\mathscr{U}_k \subseteq \mathscr{U}_k \qquad \text{for all} \qquad k \in \{1, \dots, n\}$$

is called a *fan* for T in $\mathscr{V}$. A basis $\{v_1, \dots, v_n\}$ of $\mathscr{V}$ is called a *fan basis* corresponding to T if the subspaces

$$\mathscr{V}_k = \operatorname{span}\{v_1, \dots, v_k\}, \qquad k \in \{1, \dots, n\},$$

form a fan for T. ♢

Notice that (12) implies

$$1 \leq \dim \mathscr{U}_1 < \dim \mathscr{U}_2 < \cdots < \dim \mathscr{U}_n \leq n.$$

Consequently, if $\mathscr{U}_1, \dots, \mathscr{U}_n$ is a fan for T we have $\dim \mathscr{U}_k = k$ for all $k \in \{1, \dots, n\}$. In particular $\mathscr{U}_n = \mathscr{V}$.

th-utc **Theorem 3.5** (textbook 5.39 page 156)**.** *Let $n \in \mathbb{N}$ and let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n$ and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B} = (v_1, \dots, v_n)$ be a basis of $\mathscr{V}$ and set*

$$\mathscr{V}_k = \operatorname{span}\{v_1, \dots, v_k\}, \qquad k \in \{1, \dots, n\}.$$

The following statements are equivalent.

i-utc-1 (a) *$M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.*
i-utc-2 (b) *For all $k \in \{1, \dots, n\}$ we have $Tv_k \in \mathscr{V}_k$.*
i-utc-3 (c) *For all $k \in \{1, \dots, n\}$ we have $T\mathscr{V}_k \subseteq \mathscr{V}_k$.*
i-utc-4 (d) *$\mathscr{B}$ is a fan basis corresponding to T.*

Proof. (a) ⇒ (b). Assume that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular. That is

$$M_{\mathscr{B}}^{\mathscr{B}}(T) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{kk} & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Let $k \in \{1, \dots, n\}$ be arbitrary. Then, by the definition of $M_{\mathscr{B}}^{\mathscr{B}}(T)$,

$$C_{\mathscr{B}}(Tv_k) = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Consequently, by the definition of $C_{\mathscr{B}}$, we have

$$Tv_k = a_{1k}v_1 + \cdots + a_{kk}v_k \in \operatorname{span}\{v_1, \dots, v_k\} = \mathscr{V}_k.$$

Thus, (b) is proved.

(b) ⇒ (a). Assume that $Tv_k \in \mathscr{V}_k$ for all $k \in \{1, \dots, n\}$. Let a_{ij}, $i, j \in \{1, \dots, n\}$, be the entries of $M_{\mathscr{B}}^{\mathscr{B}}(T)$. Let $j \in \{1, \dots, n\}$ be arbitrary. Since $Tv_j \in \mathscr{V}_j$ there exist $\alpha_1, \dots, \alpha_j \in \mathbb{F}$ such that

$$Tv_j = \alpha_1 v_1 + \cdots + \alpha_j v_j.$$

By the definition of $C_{\mathscr{B}}$ we have

$$C_{\mathscr{B}}(Tv_j) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

On the other side, by the definition of $M_{\mathscr{B}}^{\mathscr{B}}(T)$, we have

$$C_{\mathscr{B}}(Tv_j) = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{jj} \\ a_{j+1,j} \\ \vdots \\ a_{nj} \end{bmatrix}.$$

The last two equalities, and the fact that $C_{\mathscr{B}}$ is a function, imply $a_{ij} = 0$ for all $i \in \{j+1, \dots, n\}$. This proves (a).

(b) $\Rightarrow$ (c). Suppose $Tv_k \in \mathscr{V}_k = \operatorname{span}\{v_1, \dots, v_k\}$ for all $k \in \{1, \dots, n\}$. Let $v \in \mathscr{V}_k$. Then $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$. Applying T, we get $Tv = \alpha_1 Tv_1 + \cdots + \alpha_k Tv_k$. Thus,

$$Tv \in \operatorname{span}\{Tv_1, \dots, Tv_k\}. \tag{13}$$

eq-Tv-span1

Since

$$Tv_j \in \mathscr{V}_j \subseteq \mathscr{V}_k \qquad \text{for all} \qquad j \in \{1, \dots, k\},$$

we have

$$\operatorname{span}\{Tv_1, \dots, Tv_k\} \subseteq \mathscr{V}_k.$$

Together with (13), this proves (c).

(c) $\Rightarrow$ (b). Suppose $T\mathscr{V}_k \subseteq \mathscr{V}_k$ for all $k \in \{1, \dots, n\}$. Then since $v_k \in \mathscr{V}_k$, we have $Tv_k \in \mathscr{V}_k$ for each $k \in \{1, \dots, n\}$.

(c) $\Leftrightarrow$ (d) follows from the definition of a fan basis corresponding to T. $\square$

th-ex-up

Theorem 3.6 (textbook 5.47 page 160)**.** *Let $\mathscr{V}$ be a finite-dimensional complex vector space with $\dim \mathscr{V} = n \in \mathbb{N}$ and let $T \in \mathscr{L}(\mathscr{V})$. Then there exists a basis $\mathscr{B}$ of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.*

Proof. We proceed by the principle of Mathematical Induction on $n = \dim(\mathscr{V})$.

The base case is trivial. Assume $\dim \mathscr{V} = 1$ and $T \in \mathscr{L}(\mathscr{V})$. Set $\mathscr{B} = \{u\}$, where $u \in \mathscr{V} \backslash \{0_{\mathscr{V}}\}$ is arbitrary. Then there exists $\lambda \in \mathbb{C}$ such that $Tu = \lambda u$. Thus, $M_{\mathscr{B}}^{\mathscr{B}}(T) = \big[\lambda\big]$.

Now we prove the inductive step. Let $m \in \mathbb{N}$ be arbitrary. The inductive hypothesis is

For every $k \in \{1, \dots, m\}$ the following implication holds: If $\dim \mathscr{U} = k$ and $S \in \mathscr{L}(\mathscr{U})$, then there exists a basis $\mathscr{A}$ of $\mathscr{U}$ such that $M_{\mathscr{A}}^{\mathscr{A}}(S)$ is upper-triangular.

To complete the inductive step, we need to prove the implication:

If $\dim \mathscr{V} = m+1$ and $T \in \mathscr{L}(\mathscr{V})$, then there exists a basis $\mathscr{B}$ of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular.

To prove the red implication assume that $\dim \mathscr{V} = m+1$ and $T \in \mathscr{L}(\mathscr{V})$. By Theorem 2.4 the operator T has an eigenvalue. Let λ be an eigenvalue of T. Set $\mathscr{U} = \operatorname{ran}(T - \lambda I)$. Because $(T - \lambda I)$ is not injective, it is not surjective, and thus $k = \dim(\mathscr{U}) < \dim(\mathscr{V}) = m+1$. That is $k \in \{1, \ldots, m\}$.

Moreover, $T\mathscr{U} \subseteq \mathscr{U}$. To show this, let $u \in \mathscr{U}$. Then $Tu = (T-\lambda I)u+\lambda u$. Since $(T - \lambda I)u \in \mathscr{U}$ and $\lambda u \in \mathscr{U}$, $Tu \in \mathscr{U}$. Denote by S the restriction of T to $\mathscr{U}$. That is, $Su = Tu$ for all $u \in \mathscr{U}$. Since $T\mathscr{U} \subseteq \mathscr{U}$, we have $S \in \mathscr{L}(\mathscr{U})$.

By the inductive hypothesis (the green box), there exists a basis $\mathscr{A} = (u_1, \ldots, u_k)$ of $\mathscr{U}$ such that $M_{\mathscr{A}}^{\mathscr{A}}(S)$ is upper-triangular. This, by Theorem 3.5, implies

$$Tu_j = Su_j \in \operatorname{span}\{u_1, \ldots, u_j\} \qquad \text{for all} \qquad j \in \{1, \ldots, k\}.$$

Extend $\mathscr{A}$ to a basis $\mathscr{B} = \{u_1, \ldots, u_k, v_1, \ldots, v_{m+1-k}\}$ of $\mathscr{V}$. Since

$$Tv_j = (T - \lambda I)v_j + \lambda v_j, \qquad j \in \{1, \ldots, m+1-k\},$$

where $(T - \lambda I)v_j \in \mathscr{U}$, for all $j \in \{1, \ldots, m+1-k\}$ we have

$$Tv_j \in \operatorname{span}\{u_1, \ldots, u_k, v_j\} \subseteq \operatorname{span}\{u_1, \ldots, u_k, v_1, \ldots, v_j\}.$$

By Theorem 3.5 $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper-triangular. □

Remark 3.7. Theorem 3.6 is stated as Theorem 5.47 in the textbook. Since the textbook covers both $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$ the book's proof is different from ours. To some extend our proof is more direct. ♢

th-invc **Theorem 3.8.** *Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$ such that $\dim \mathscr{V} = n \in \mathbb{N}$, and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B} = (v_1, \ldots, v_n)$ be a basis of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular with the diagonal entries a_{jj}, $j \in \{1, \ldots, n\}$. Then T is isomorphism if and only if for all $j \in \{1, \ldots, n\}$ we have $a_{jj} \neq 0$.*

Proof. In this proof we set

$$\mathscr{V}_k = \operatorname{span}\{v_1, ..., v_k\}, \qquad k \in \{1, ..., n\}.$$

Then

$$\mathscr{V}_1 \subsetneq \mathscr{V}_2 \subsetneq \ldots \subsetneq \mathscr{V}_n \tag{14}$$ eq-fan-sub

and by Theorem 3.5, $T\mathscr{V}_k \subseteq \mathscr{V}_k$.

We first prove the contrapositive of the "if" part. Assume that T is not an isomorphism. Then T is not injective. Consider the set

$$\mathbb{K} = \Big\{ k \in \{1, ..., n\} \, : \, T\mathscr{V}_k \subsetneq \mathscr{V}_k \Big\}$$

Since T is not injective, $\operatorname{nul} T \neq \{0_{\mathscr{V}}\}$. Thus by the Rank-Nullity Theorem, $\operatorname{ran} T \subsetneq \mathscr{V} = \mathscr{V}_n$. Since $T\mathscr{V}_n = \operatorname{ran} T$, it follows that $T\mathscr{V}_n \subsetneq \mathscr{V}_n$. Therefore $n \in \mathbb{K}$. Hence the set $\mathbb{K}$ is a nonempty set of positive integers. Hence, by the Well-Ordering principle $\min \mathbb{K}$ exists. Set $j = \min \mathbb{K}$.

If $j = 1$, then $\dim \mathscr{V}_1 = 1$, but since $T\mathscr{V}_1 \subsetneq \mathscr{V}_1$ it must be that $\dim(T\mathscr{V}_1) = 0$. Thus $T\mathscr{V}_1 = \{0_{\mathscr{V}}\}$, so $Tv_1 = 0_v$. Hence $C_{\mathscr{B}}(Tv_1) = [0 \cdots 0]^\top$ and so

$a_{11} = 0$. If $j > 1$, then $j-1 \in \{1, \ldots, n\}$ but $j - 1 \notin \mathbb{K}$. By Theorem 3.5, $T\mathscr{V}_{j-1} \subseteq \mathscr{V}_{j-1}$ and, since $j-1 \notin \mathbb{K}$, $T\mathscr{V}_{j-1} \subsetneq \mathscr{V}_{j-1}$ is not true. Hence $T\mathscr{V}_{j-1} = \mathscr{V}_{j-1}$. Since $j \in \mathbb{K}$, we have $T\mathscr{V}_j \subsetneq \mathscr{V}_j$. Now we have

$$\mathscr{V}_{j-1} = T\mathscr{V}_{j-1} \subseteq T\mathscr{V}_j \subsetneq \mathscr{V}_j.$$

Consequently,

$$j - 1 = \dim \mathscr{V}_{j-1} \leq \dim\bigl(T\mathscr{V}_j\bigr) < \dim \mathscr{V}_j = j,$$

which implies $\dim\bigl(T\mathscr{V}_j\bigr) = j - 1$ and therefore $T\mathscr{V}_j = \mathscr{V}_{j-1}$. This implies that there exist $\alpha_1, \ldots, \alpha_{j-1} \in \mathbb{F}$ such that

$$Tv_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1}.$$

By the definition of $M_{\mathscr{B}}^{\mathscr{B}}(T)$ this implies that $a_{jj} = 0$.

Next we prove the "if" part. Assume that there exists $j \in \{1, ..., n\}$ such that $a_{jj} = 0$. Then

$$Tv_j = a_{1j} v_1 + \cdots + a_{j-1,j} v_{j-1} + 0 v_j \in \mathscr{V}_{j-1}. \tag{15}$$ eq-inv-if1

By Theorem 3.5 and (14) we have

$$Tv_i \in \mathscr{V}_i \subseteq \mathscr{V}_{j-1} \qquad \text{for all} \qquad i \in \{1, \ldots, j-1\}. \tag{16}$$ eq-inv-if2

Now (15) and (16) imply $Tv_i \in \mathscr{V}_{j-1}$ for all $i \in \{1, \ldots, j\}$ and consequently $T\mathscr{V}_j \subseteq \mathscr{V}_{j-1}$. To complete the proof, we apply the Rank-Nullity theorem to the restriction $T|_{\mathscr{V}_j}$ of T to the subspace $\mathscr{V}_j$:

$$\dim \operatorname{nul}\bigl(T|_{\mathscr{V}_j}\bigr) + \dim \operatorname{ran}\bigl(T|_{\mathscr{V}_j}\bigr) = j.$$

Since $T\mathscr{V}_j \subseteq \mathscr{V}_{j-1}$ implies $\dim \operatorname{ran}\bigl(T|_{\mathscr{V}_j}\bigr) \leq j - 1$, we conclude

$$\dim \operatorname{nul}\bigl(T|_{\mathscr{V}_j}\bigr) \geq 1.$$

Thus $\operatorname{nul}\bigl(T|_{\mathscr{V}_j}\bigr) \neq \{0_{\mathscr{V}}\}$, that is, there exists $v \in \mathscr{V}_j$ such that $v \neq 0$ and $Tv = T|_{\mathscr{V}_j} v = 0$. This proves that T is not invertible. □

co-invc **Corollary 3.9.** *Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n \in \mathbb{N}$, and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B}$ be a basis of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular with diagonal entries a_{jj}, $j \in \{1, \ldots, n\}$. The following statements are equivalent.*

i-invc-1 (a) *T is not injective.*
i-invc-2 (b) *T is not invertible.*
i-invc-3 (c) *0 is an eigenvalue of T.*
i-invc-4 (d) $\prod_{i=1}^{n} a_{ii} = 0$.
i-invc-5 (e) *There exists $j \in \{1, \ldots, n\}$ such that $a_{jj} = 0$.*

Proof. The equivalence (a) ⇔ (b) follows from the Rank-nullity theorem and it has been proved earlier. The equivalence (a) ⇔ (c) is almost trivial. The equivalence (a) ⇔ (e) was proved in Theorem 3.8 and The equivalence (d) ⇔ (e) is should have been proved in high school. □

th-sp-di **Theorem 3.10** (textbook 5.41 p. 157)**.** *Let $\mathscr{V}$ be a finite-dimensional vector space over $\mathbb{F}$ with $\dim \mathscr{V} = n \in \mathbb{N}$, and let $T \in \mathscr{L}(\mathscr{V})$. Let $\mathscr{B}$ be a basis of $\mathscr{V}$ such that $M_{\mathscr{B}}^{\mathscr{B}}(T)$ is upper triangular with diagonal entries a_{jj}, $j \in \{1, \ldots, n\}$. Then*

$$\sigma(T) = \Big\{ a_{jj} : j \in \{1, ..., n\} \Big\}.$$

Proof. Notice that $M_{\mathscr{B}}^{\mathscr{B}} : \mathscr{L}(V) \to \mathbb{F}^{n\times n}$ is an isomorphism of algebras. Therefore

$$M_{\mathscr{B}}^{\mathscr{B}}(T - \lambda I) = M_{\mathscr{B}}^{\mathscr{B}}(T) - \lambda M_{\mathscr{B}}^{\mathscr{B}}(I) = M_{\mathscr{B}}^{\mathscr{B}}(T) - \lambda I_n.$$

Here I_n denotes the identity matrix in $\mathbb{F}^{n\times n}$. As $M_{\mathscr{B}}^{\mathscr{B}}(T)$ and $M_{\mathscr{B}}^{\mathscr{B}}(I) = I_n$ are upper triangular, $M_{\mathscr{B}}^{\mathscr{B}}(T - \lambda I)$ is upper triangular as well with the diagonal entries $a_{jj} - \lambda$ with $j \in \{1, ..., n\}$.

To prove the set equality

$$\sigma(T) = \big\{ a_{jj} : j \in \{1, ..., n\} \big\}.$$

in the theorem we need to prove two inclusions.

First we prove $\subseteq$. Let $\lambda \in \sigma(T)$. Because λ is an eigenvalue, $T - \lambda I$ is not injective. Because $T - \lambda I$ is not injective. By Theorem 3.8 one of the diagonal entries of the upper triangular matrix

$$M_{\mathscr{B}}^{\mathscr{B}}(T - \lambda I) = M_{\mathscr{B}}^{\mathscr{B}}(T) - \lambda I_n$$

is zero. That is, there exists $i \in \{1, ..., n\}$ such that $a_{ii} - \lambda = 0$. Thus $\lambda = a_{ii}$. So $\sigma(T) \subseteq \big\{ a_{jj} : j \in \{1, ..., n\} \big\}$.

Next we prove $\supseteq$. Let $j \in \{1, ..., n\}$ be arbitrary. Then the j-th diagonal entry of the matrix

$$M_{\mathscr{B}}^{\mathscr{B}}(T - a_{jj} I) = M_{\mathscr{B}}^{\mathscr{B}}(T) - a_{jj} I_n$$

is equal to $a_{jj} - a_{jj} = 0$. By Theorem 3.8 the operator $T - a_{jj} I$ is not injective. This implies that a_{jj} is an eigenvalue of T. Thus $a_{jj} \in \sigma(T)$. This completes the proof. ☐

4. Existence of the Minimal Polynomial

Here I present a different proof of Theorem 5.22 in the textbook. The proof in the book uses the Mathematical Induction on the dimension of the space. The proof below uses the fact that every bounded nonempty set of positive integers has a maximum.

le-pv **Lemma 4.1.** *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(\mathscr{V}) \setminus \{0_{\mathscr{L}(\mathscr{V})}\}$. For every $v \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ there exists a unique positive integer $k_v \in \mathbb{N}$ and a unique monic polynomial $p_v \in \mathbb{F}[z]$ such that*

le-pv-i1 (i) $1 \leq \deg p_v = k_v \leq \dim \mathscr{V}$,
le-pv-i2 (ii) $v, \ldots, T^{k_v - 1} v$ *are linearly independent,*
le-pv-i2a (iii) $v, \ldots, T^{k_v - 1} v, T^{k_v} v$ *are linearly dependent,*
le-pv-i3 (iv) $v, \ldots, T^{k_v - 1} v \in \operatorname{nul}\big(p_v(T)\big)$,

le-pv-i4 (v) $\deg p_v \le \dim\bigl(\operatorname{nul} p_v(T)\bigr)$.

Proof. Let $v \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$ be arbitrary and let $k \in \mathbb{N}$ be the smallest positive integer such that

$$T^k v \in \operatorname{span}\Bigl\{T^{j-1}v : j \in \{1,\ldots,k\}\Bigr\}.$$

With $k_v = k$ we have $k \le \dim \mathscr{V}$, (ii) holds, and there exist unique $\alpha_0,\ldots,\alpha_{k-1} \in \mathbb{F}$ not all zero such that

$$T^k v = \alpha_0 v + \cdots + \alpha_{k-1}T^{k-1}v.$$

Set

$$p_v(z) = -\alpha_0 - \cdots - \alpha_{k-1}z^{k-1} + z^k.$$

Then $\bigl(p_v(T)\bigr)v = 0_{\mathscr{V}}$ and (i) holds. We deduce (iv) since for all $j \in \{1,\ldots,k\}$ we have

$$\bigl(p_v(T)T^{j-1}\bigr)v = \bigl(T^{j-1}p_v(T)\bigr)v = T^{j-1}\bigl(p_v(T)v\bigr) = 0_{\mathscr{V}}.$$

As a consequence of (ii) and (iv) we obtain (v). □

Theorem 4.2. *Let $\mathscr{V}$ be a nontrivial finite-dimensional vector space over $\mathbb{F}$. For every $T \in \mathscr{L}(\mathscr{V})$, there exists a unique monic polynomial $p \in \mathbb{F}[z]$ with $p \neq 0$ such that $p(T) = 0_{\mathscr{L}(\mathscr{V})}$ and p is minimal with this property, meaning that if $q \in \mathbb{F}[z]$ satisfies $q(T) = 0_{\mathscr{L}(\mathscr{V})}$ and $\deg q \le \deg p$, then q is a scalar multiple of p. Furthermore, we have $\deg p \le \dim \mathscr{V}$.*

Proof. If $T = 0_{\mathscr{L}(\mathscr{V})}$, then $p(z) = z$ has desired properties. All polynomials that we consider in this proof are monic polynomials. Assume that $T \neq 0_{\mathscr{L}(\mathscr{V})}$ and set

$$\mathscr{P}_T = \Bigl\{q \in \mathbb{F}[z] : \deg q \le \dim\bigl(\operatorname{nul} q(T)\bigr)\Bigr\}.$$

The set $\mathscr{P}_T$ is not empty since by Lemma 4.1 the polynomial $p_v \in \mathscr{P}_T$ for every $v \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$.

By the definition of $\mathscr{P}_T$, for every polynomial $q \in \mathscr{P}_T$ we have $\deg q \le \dim \mathscr{V}$. Therefore, the following maximum exists,

$$m = \max\Bigl\{\deg q : q \in \mathscr{P}_T\Bigr\},$$

and $m \le \dim \mathscr{V}$.

Let $q \in \mathscr{P}_T$. We will prove the implication

$$q(T) \neq 0_{\mathscr{L}(\mathscr{V})} \quad \Rightarrow \quad \deg q < m. \tag{17}$$ eq-dnm

The contrapositive of the implication in (17) proves the existence of p such that $p(T) = 0_{\mathscr{L}(\mathscr{V})}$.

To prove (17) assume $q(T) \neq 0_{\mathscr{L}(\mathscr{V})}$; that is, there exists $u \in \mathscr{V}$ such that $w = q(T)u \neq 0_{\mathscr{V}}$. Let k_w and p_w be the positive integer and the polynomial associated with w in Lemma 4.1. To prove (17), we will prove that $p_w q$, the product of the polynomials p_w and q, is in $\mathscr{P}_T$, and $\deg q < \deg(p_w q)$.

First, by Lemma 4.1 we have

$$\deg(p_w q) = \deg p_w + \deg q \geq 1 + \deg q > \deg q.$$

Second, to prove $p_w q \in \mathscr{P}_T$, let $u_1, \ldots, u_l$ be a basis for $\operatorname{nul} q(T)$. Then, since by Lemma 4.1 the vectors $w, \ldots, T^{k_w-1}w$ are linearly independent, we have that the vectors

$$u, \ldots, T^{k_w-1}u,\ u_1, \ldots, u_l,$$

are linearly independent and they all belong to $\operatorname{nul}\big(p_w(T)q(T)\big)$. (Prove this as an exercise.) Therefore, by Lemma 4.1,

$$\begin{aligned}\dim\big(\operatorname{nul}\big(p_w(T)q(T)\big)\big) &\geq k_w + \dim\big(\operatorname{nul} q(T)\big)\\ &\geq k_w + \deg q\\ &\geq \deg p_w + \deg q\\ &= \deg(p_w q).\end{aligned}$$

Hence, $p_w q \in \mathscr{P}_T$. □

Remark 4.3. I like the above proof since it gives an algorithmic way of finding a polynomial q such that $q(T) = 0_{\mathscr{L}(\mathscr{V})}$. What is the algorithm? Take $v_1 \in \mathscr{V} \setminus \{0_{\mathscr{V}}\}$. Find p_{v_1} from Lemma 4.1. If $p_{v_1}(T) = 0_{\mathscr{L}(\mathscr{V})}$, we are done. If not, set $v_2 = p_{v_1}(T)u_1 \neq 0_{\mathscr{V}}$. Then

$$\deg(p_{v_1}p_{v_2}) \leq \dim\big(\operatorname{nul}\big(\big(p_{v_1}p_{v_2}\big)(T)\big)\big) \leq \dim \mathscr{V}$$

and

$$\dim\big(\operatorname{nul}\big(p_{v_1}(T)\big)\big) < \dim\big(\operatorname{nul}\big(\big(p_{v_1}p_{v_2}\big)(T)\big)\big) \leq \dim \mathscr{V}.$$

Again, if $\big(p_{v_1}p_{v_2}\big)(T) \neq 0_{\mathscr{L}(\mathscr{V})}$, then set $v_3 = \big(p_{v_1}p_{v_2}\big)(T)u_2 \neq 0_{\mathscr{V}}$. Then

$$\deg(p_{v_1}p_{v_2}p_{v_3}) \leq \dim\big(\operatorname{nul}\big(\big(p_{v_1}p_{v_2}p_{v_3}\big)(T)\big)\big) \leq \dim \mathscr{V}$$

and

$$\begin{aligned}\dim\big(\operatorname{nul}\big(p_{v_1}(T)\big)\big) &< \dim\big(\operatorname{nul}\big(\big(p_{v_1}p_{v_2}\big)(T)\big)\big)\\ &< \dim\big(\operatorname{nul}\big(\big(p_{v_1}p_{v_2}p_{v_3}\big)(T)\big)\big)\\ &\leq \dim \mathscr{V}.\end{aligned}$$

This algorithm stops after at most $(\dim \mathscr{V}) - 1$ steps. ♢