CHAPTER 1

Topological vector spaces

1. Vector spaces

In the following F is a field.

DEFINITION 1.1. *A vector space V over F is given by a commutative group $(V, +, 0)$ and a map $m : F \times V \to V$ such that*

$$m(\lambda, x + y) \;=\; m(\lambda, x) + m(\lambda, y) \;.$$

DEFINITION 1.2. *Let V be a vector space. A system $S \subset V$ is said to be linear independent, if for every finite family (λ_s) of scalars*

$$\sum_{s \in S} \lambda_s s \;=\; 0$$

implies $\lambda_s = 0$ for all $s \in S$.

EXAMPLE 1.3. *$C[0,1]$ is a vector space. The polynomials $\{p_k \,:\, k \geq 0\}$ are independent ($p_k(t) = t^k$).*

EXAMPLE 1.4. *The space $F(\mathbb{R}) \;=\; \mathbb{R}^{\mathbb{R}}$ is a vector space over $\mathbb{R}$.*

EXAMPLE 1.5. *Let I be an index set. Then $F^I \;=\; \{f : I \to F \,:\, f function\}$ is a vector space over F.*

EXAMPLE 1.6. *$\mathbb{R}$ is a vector space over $\mathbb{Q}$.*

DEFINITION 1.7. *A subspace W of V is a subset $W \subset V$ such that $x, y \in W$ and $\lambda \in F$ implies*

$$x + \lambda y \in W \;.$$

EXAMPLE 1.8. *Let I be an index set. Consider $F(I) \subset F^I$ defined by*

$$F(I) \;=\; \{f : I \to F \,:\, \exists_{S \subset I finite}(i \notin S \Rightarrow f(i) = 0)\} \;.$$

$F(I)$ is called the free vector space over I.

PROPOSITION 1.9. *Let $W \subset V$ be a subspace. Define $x \sim y$ by $x - y \in W$. Then $\sim$ is an equivalence relation and $V/\sim$ is a vector space. This space is denoted by V/W.*

PROOF. Obviously, $\sim$ is an equivalence relation. We define the operations

$$[x] + \lambda[y] \; = \; [x + \lambda y] \, .$$

If $x - x' \in W$ and $y - y' \in W$, then $x - x' + \lambda(y - y') \in W$. Thus $V/\sim$ is a vector space. ■

DEFINITION 1.10. *Let V and W be vector spaces. The direct sum $V \oplus W$ is the vector space $V \times W$ with the operation*

$$(x, y) + \lambda(x', y') \; = \; (x + \lambda x', y + \lambda y') \, .$$

DEFINITION 1.11. *1) Let $S \subset V$. Then the span of S is defined as*

$$\operatorname{span}(S) \; = \; \{\sum_{s \in S'} \lambda_s s \; : \; \lambda_s \in F, S' \subset S \textit{ finite}\} \, .$$

2) $B \subset V$ is called a basis if B is linear independent and $\operatorname{span}(B) = V$.

b0 THEOREM 1.12. *Every vector space has a basis.*

PROOF. Consider the collection L of linear independent subsets of V. We say that $S_1 \leq S_2$ of $S_1 \subset S_2$. Let $(S_i)_{i \in I}$ be a chain, i.e. for every i and $j \in I$ there exists $k \in I$ such that $S_i \subset S_k$ and $S_j \subset S_k$. We claim that $S = bigcup_i S_i$ is linear independent. For let $(\lambda_s)_{s \in S'}$ be finite family such that

$$\sum_s \lambda_s s \; = \; 0 \, .$$

For every $s \in S'$ we find $i(s)$ such that $s \in S_{i(s)}$. Since S' is finite we may find $k \in I$ such that $S_{i(s)} \subset S_k$ for all $s \in S'$. Hence, $\{s \; : \; S'\} \subset S_k$. By linear independence, we deduce $\lambda_s = 0$ for all $s \in S'$. By Zorn's Lemma, we find a maximal element B in L. This means for no $x \in V$ $B \cup \{x\} \in L$. We claim that $V = \operatorname{span}(B)$. Let us show that $x \notin \operatorname{span}(B)$ implies that $B \cup \{x\}$ is linear independent. Indeed, let λ_x and $(\lambda_s)_{s \in S}$, $S \subset B$ finite such that

$$\lambda_x x + \sum_s \lambda_s s \; = \; 0 \, .$$

If $\lambda_x = 0$, we deduce $0 = \sum_s \lambda_s s$ and hence $\lambda_s = 0$ for all $s \in S$. If $\lambda_x \neq 0$ we find

$$x \; = \; \sum_s -\lambda_s/\lambda_x s \, .$$

Since $x \notin \operatorname{span}(B)$ we deduce $\lambda_x = 0$ and the assertion follows. ■

EXAMPLE 1.13. $\{1, \sqrt{2}\}$ *is linear independent over* $\mathbb{Q}$*. Let* B *be a basis for* $\mathbb{R}$ *over* $\mathbb{Q}$*. Then we find an injective map*

$$\Phi : \mathbb{R} \to \bigcup_{S' \subset B \text{ finite}} \mathbb{Q}^{S'}$$

Since $\mathbb{Q}^{S'}$ *is countable and for an infinite* B *set the collection of all finite subsets of* B *has the same cardinality as* B*, we deduce that* B *has the same cardinality as* $\mathbb{R}$*.*

2. Linear transformations

DEFINITION 2.1. *Let* V *and* W *vector spaces over* F*. A map* $T : V \to W$ *is called linear, if*

$$T(x + \lambda y) \;=\; T(x) + \lambda T(y)$$

holds for all $x, y \in V$*. A linear map* T *is called an isomorphism if* T *is bijective.*

REMARK 2.2. *1) If* $T(V)$ *is a subspace of* W*.*
2) If $T : V \to W$ *is bijective, then* T^{-1} *is linear.*
3) If $T : V \to W$ *and* $S : W \to Z$ *are linear maps, then the composition* $ST : V \to Z$ *is linear.*

EXAMPLE 2.3. *Let* $\phi : I \to J$ *be a map. Then*

$$T_\phi : F(J) \to F(I)\;,\; T(f) \;=\; f \circ \phi$$

is a linear map. If ϕ *is bijective, then* T_ϕ *is an isomorphism.*

THEOREM 2.4. *Every vector space is isomorphic to a free vector space.*

PROOF. Let B be a basis. We define $T : F(B) \to V$ by

$$T(f) \;=\; \sum_{b \in B} f(b)b\,.$$

Note that $T(f)$ is well-defined because only finitely many coordinates are non-zero. It is easily checked that T is an isomorphism. ■

PROPOSITION 2.5. *Let* V *be a vector space and* W *be a subspace. Then* V *isomorphic to* $W \oplus V/W$*.*

PROOF. Let B be a basis of V/W. Let $f : B \to V$ such that $f(b) \in b$ holds for all $b \in B$. Let us show that $\{f(b) \;:\; b \in B\}$ is linearly independent. Indeed, if we assume

$$0 \;=\; \sum_b \lambda_b f(b)$$

then $0 = [0] = \sum_b \lambda_b[f(b)] = \sum_b \lambda_b b$. Hence $\alpha_b = 0$. Therefore, we may define $T_2 : V/W \to V$ by

$$T_2(\sum_b \lambda_b b) = \sum_b \lambda_b f(b) .$$

Then we define $T : W \oplus V/W \to V$ by

$$T(w, x) = w + T_2(x) .$$

Obviously, T is linear. Let us show that T is injective. This is to show $T(w, x) = 0$ implies $w = 0$ and $x = 0$. Let $x = \sum_b \lambda_b f(b)$. $0 = T(w, x)$ implies $T_2(x) \in W$. Hence

$$0 = \sum_b \lambda_b[f(b)] = \sum_b \lambda_b b = x .$$

Thus $x = 0$, $T_2(x) = 0$ and $w = 0$. Let $y \in V$. We write $[y] = \sum_b \lambda_b b$. Then we have

$$[y - \sum_b \lambda_b f(b)] = [y] - \sum_b \lambda_b[f(b)] = [y] - \sum_b b = 0 .$$

Thus $y \sim \sum_b \lambda_b f(b)$. We get

$$T(y - \sum_b \lambda_b f(b), [y]) = y .$$

Hence T is bijective. ■

Definition 2.6. *Let $T : V \to W$. We define the kernel*

$$\ker(T) = \{x \in V \ : \ T(x) = 0\}$$

and $\mathrm{rg}(V) = T(V)$ the range.

Proposition 2.7. *Let $V \to W$ be a linear map and $q : V \to V \ker(T)$ be the quotient map $q(x) = [x]$. There exists a unique injective linear map $\hat{T} : V/\ker(T) \to W$ such that $q\hat{T} = T$.*

Proof. We have to show that $\hat{T}([x]) = T(x)$ is well-defined. However $x - x' \in \ker(T)$ implies $T(x) - T(x') = T(x - x') = 0$. Now, we assume $\hat{T}([x_1]) = \hat{T}([x_2])$. Then $T(x_1) = T(x_2)$. This implies $T(x_1 - x_2) = 0$. In particular, $x_1 \sim x_2$. ■

Definition 2.8. *We define $\dim(V)$ to be the smallest cardinal number given by a basis. It is easy to show that*

$$\dim(V \oplus W) = \dim(V) + \dim(W) .$$

COROLLARY 2.9. *Let $T : V \to W$. Then*

$$\dim(\ker(T)) + \dim(\mathrm{rg}(T)) \; = \; \dim(V) \, .$$

PROOF. Since $\hat{T} : V/\ker(T) \to \mathrm{rg}(T)$ is an isomorphism, we see that

$$V \cong \ker(T) \oplus V/\ker(T) \cong \ker(T) \oplus \mathrm{rg}(T) \, .$$

The formula for the dimensions follows. ∎

COROLLARY 2.10. *Let V be a finite dimensional vector space and $T : V \to V$. Then T is injective if and only if T is surjective.*

PROOF. T is injective iff $\dim(\ker(T)) = 0$ iff $\dim(\mathrm{rg}(T)) = n$. ∎

COROLLARY 2.11. *Let W_1 and W_2 be finite dimensional subspaces of a vector space V. Then*

$$\dim(W_1) + \dim(W_2) \; = \; \dim(W_1 + W_2) + \dim(W_1 \cap W_2) \, .$$

PROOF. W.l.o.g. we may assume $V = W_1 + W_2$. Let us first assume that $W_1 \cap W_2 = \{0\}$. Then the map $\iota : W_2 \to V/W_1$ is bijective and hence

k0 (2.1) $$\dim(W_1 + W_2) \; = \; \dim(V/W_1) + \dim(W_1) \; = \; \dim(W_2) + \dim(W_1) \, .$$

Now, we consider the general case. Since $W_1 \cap W_2$ is a subspace of V, W_1 and W_2 we have

$$\begin{aligned} \dim(W_1 + W_2) &= \dim(W_1 + W_2/W_1 \cap W_2) + \dim(W_1 \cap W_2) \, , \\ \dim(W_1) &= \dim(W_1/W_1 \cap W_2) + \dim(W_1 \cap W_2) \, , \\ \dim(W_2) &= \dim(W_2/W_1 \cap W_2) + \dim(W_1 \cap W_2) \, . \end{aligned}$$

The assertion follows from

$$\dim(W_1 + W_2/W_1 \cap W_2) \; = \; \dim(W_1/W_1 \cap W_2) + \dim(W_2/W_1 \cap W_2)$$

which is particular of our preliminary observation because $W_1/W_1 \cap W_2 \cap W_2/W_1 \cap W_2 = \{0\}$. ∎

DEFINITION 2.12. *Let B be a basis for V and $S \subset V$. Then we find scalars $\lambda_{b,s}$ such that*

$$s \; = \; \sum_{b \in B} \lambda_{s,b} b \, .$$

The matrix $m_{S,B} = [\lambda_{s,b}]_{s \in S, b \in B}$ is called transition matrix.

LEMMA 2.13. *Let B and S be a basis for V with transition matrices $m_{B,S}$ and $m_{S,B}$. Then*

$$id \;=\; m_{S,B}m_{B,S} \quad and \quad id \;=\; m_{B,S}m_{S,B}\;.$$

PROOF. Let $b \in B$. Then we have

$$\begin{aligned} b \;&=\; \sum_{s\in S}\mu_{b,s}s \;=\; \sum_{s\in S}\mu_{b,s}\sum_{b'\in B}\lambda_{s,b'}b' \\ &=\; \sum_{b'\neq b}(\sum_{s}\mu_{b',s}\lambda_{s,b})b' + (\sum_{s}\mu_{b,s}\lambda_{s,b})b \end{aligned}$$

By linear independence, we get

$$\sum_{s}\mu_{b',s}\lambda_{s,b} \;=\; \delta_{b',b}\;.$$

This shows $m_{S,B}m_{B,S} = id$. Starting with elements in s yields the first assertion. ∎

Let $T : V \to W$ be a linear map and C a basis for V and B be a basis for W. We may write

$$T(c) \;=\; \sum_{b\in B}\lambda_{c,b}b\;.$$

Then $m^T_{C,B} = [\lambda_{b,c}]$ is the matrix associated to T (with respect to C and B).

LEMMA 2.14. *Let B and S be a matrix for V and C, D be a basis for W. Then*

$$m^T_{D,S} \;=\; m_{D,C}m^T_{C,B}m_{B,S}\;.$$

PROOF. Assume $D = C$ first. Then

$$T(c) \;=\; \sum_{b\in B}\lambda_{c,b}b \;=\; \sum_{b\in B}\lambda_{c,b}\sum_{s\in S}\mu_{b,s}s\;.$$

This yields

$$m^T_{C,S} \;=\; m^T_{C,B}m_{B,S}\;.$$

The equation

$$m^T_{D,B} \;=\; m_{D,C}m^T_{C,B}$$

is proved similarly. ∎

COROLLARY 2.15. *Let $T : V \to V$ a linear map. Let B and S be a basis. Then*

$$m^T_{S,S} \;=\; m_{S,B}m^T_{B,B}m_{B,S} \;=\; m^{-1}_{B,S}m^T_{B,B}m_{B,S}\;.$$

3. Determinant and adjacent matrix

We recall that on the space of $n \times n$ matrices over F the determinant is given by

$$\det(A) \;=\; \sum_{\pi \in S_n} (-1)^{I(\pi)} \prod_{i=1}^{n} a_{i\pi(i)} \,.$$

Here S_n is the space of all permutation of the set $\{1, ..., n\}$ and

$$I(\pi) \;=\; \#\{i < j \;:\; \pi(i) > \pi(j)\}$$

is the number of inversions. It is easily checked that det is multi-linear, antisymmetric and satisfies $\det(I) = 1$. Using the Gauss-elimination method one can then show that $\det(A) \neq 0$ if and only if A is invertible. We will use a different approach.

DEFINITION 3.1. *Given a matrix $A = (a_{ij})$ we denote by B_{ij} the matrix obtained by deleting the i and j-th column. We define*

$$b_{ij} \;=\; (-1)^{i+j} \det(X_{ji})$$

Then $B = adj(A)$ is called the adjacent matrix to A

LEMMA 3.2. *$A \,\mathrm{adj}(A) \;=\; \det(A) I$.*

PROOF. Consider $C = A \,\mathrm{adj}(A)$. Then we deduce from the well-known determinant expansion

$$c_{ii} \;=\; \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(X_{ij}) \;=\; \det(A) \,.$$

For $i \neq k$ we get

$$c_{ik} \;=\; \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(X_{kj}) \;=\; \det(\tilde{A}_{ik})$$

Here $\tilde{A}$ repeats the i-th column in the k-column. Thus $\det(\tilde{A}_{ik}) \;=\; 0$ by antisymmetry. ■

COROLLARY 3.3. *If $\det(A) \neq 0$, then $A^{-1} \;=\; \det(A)^{-1} \,\mathrm{adj}(A)$.*

LEMMA 3.4. *Let F be $\mathbb{R}$ or $\mathbb{C}$ and A_j be $m \times m$ matrices such that*

$$\sum_{k=0}^{n} A_k \lambda^k \;=\; 0$$

for all $\lambda \in F$. Then $A_0 = A_1 = \cdots = A_n = 0$.

Proof. Proof by induction $n = 0$. By assumption $A_0 = 0$. We assume the assertion is true n. Let $\sum_{k=0}^{n+1} A_k \lambda_k = 0$. Inserting $\lambda = 0$ we get $A_0 = 0$ and hence

$$\lambda(\sum_{k=0}^{n} A_{k+1}\lambda^k) = 0 .$$

This for $\lambda \neq 0$ we must have

$$\sum_{k=0}^{n} A_{k+1}\lambda^k = 0$$

By continuity this also holds for $\lambda = 0$ and the induction hypothesis implies $A_1 = A_2 = \cdots = A_{n+1} = 0$. ■

For a polynomial $p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ and a matrix A we define

$$p(A) = \sum_{k=0}^{n} a_k A^k .$$

Theorem 3.5. *Let A be an $m \times m$ matrix and*

$$p_A(\lambda) = \det(\lambda I - A)$$

the characteristic polynomial. Then

$$p_A(A) = 0 .$$

Proof. We know that

$$(\lambda I - A)\,\mathrm{adj}(\lambda I - A) = \det(\lambda I - A) I = p_A(\lambda) I .$$

We write

$$\mathrm{adj}(\lambda I - A) = \sum_{k=0}^{n-1} B_k \lambda^k$$

Let us write

$$p_A(\lambda) = \sum_{k=0}^{n} a_k \lambda^k .$$

Here $a_n = 1$. This gives

$$\begin{aligned}\sum_{k=0}^{n} a_k \lambda^k I = p_A(\lambda) I &= (\lambda I - A)\,\mathrm{adj}(\lambda I - A) \\ &= (\lambda I - A)\sum_{k=0}^{n-1} B_k \lambda^k \\ &= -AB_0 + \sum_{k=1}^{n} (B_{k-1} - AB_k)\lambda^k .\end{aligned}$$

Comparing coefficients we get $B_{n-1} = I$, $-AB_0 = a_0$ and

$$B_{k-1} - AB_k \;=\; a_k I$$

for $k = 1, ..., n-2$. Multiplying with A^{k-1} we deduce

$$A^k B_{k-1} - A^{k+1} B_k \;=\; a_k A^k \,.$$

Therefore

$$\sum_{k=0}^{n} a_k A^k = A^n - AB_0 + \sum_{k=0}^{n-1} (A^{k-1} B_{k-1} - A^k B_k) \;=\; A^n - A^n \;=\; 0\,. \qquad \blacksquare$$

Definition 3.6. *Let us consider the ideal*

$$I_A \;=\; \{p \in \mathbb{C}[X] \;:\; p(A) = 0\}$$

of polynomials. Since the integral domain of polynomials admits a factorization algorithm there exists a polynomial m_A with minimal degree and leading coefficients 1 *such that*

$$I_A \;=\; C[X] m_A \,.$$

In particular, the minimal polynomial divides the characteristic polynomial p_A.

4. Eigenvalues and eigenvectors

Definition 4.1. *Let $T : V \to V$ be a linear map. A number $\lambda \in F$ is is called an eigenvalue if there exists $v \neq 0$ such that*

$$T(v) \;=\; \lambda v$$

In this case v is called eigenvector. The space K_λ of eigenvectors is given by

$$E_\lambda \;=\; \ker(T - \lambda id)\,.$$

Example 4.2. On $F^{\mathbb{N}}$ we define

$$T(f)(n) \;=\; f(n+1)\,.$$

Then $T(f) \;=\; \lambda f$ implies

$$\lambda f(n) \;=\; f(n+1)$$

Therefore every eigenvector for $T(f) \;=\; \lambda f$ is given by $f(n) \;=\; \lambda^n f(0)$ with $f(0) \neq 0$. Note that for $\lambda \neq 0$ the function $f_\lambda(n) = \lambda^n$ does not belong to the free vector space $F(\mathbb{N})$. For $\lambda = 0$ we must have $0 = f(2) = f(3) = \cdots$. Hence on f given by $f(1) = 1$ and 0 else is an eigenvector for T. We may also consider the subspace $\ell_2 \subset \mathbb{R}^{\mathbb{N}}$ given by

$$\ell_2 \;=\; \{f : \mathbb{N} \to \mathbb{R} \;:\; \sum_n |f(n)|^2 < \infty\} \,.$$

Then $f_\lambda \in \ell_2$ if and only if $|\lambda| < 1$.

Lemma 4.3. *Let $F \in \{\mathbb{C}, \mathbb{R}\}$, $A \in M_n(F)$ and λ an eigenvalue. Then $m_A(\lambda) \;=\; 0$.*

Proof. Let $v \neq 0$ such that $A(v) = \lambda v$. Then $A^k v \;=\; \lambda^k v$ implies

$$0 \;=\; m(A)v \;=\; m(\lambda)v \,.$$

Thus $m(\lambda) = 0$. ∎

exeigen

Proposition 4.4. *Let V be a complex finite dimensional vector space of positive dimension. Let $T : V \to V$ be a a linear map. Then T has an eigenvalue.*

Proof. After fixing a basis, we may associate with T a matrix A. Then $p_A(\lambda) \;=\; \det(\lambda I - A)$ is a polynomial with leading coefficient $\lambda^{\dim(V)}$. Thus $\dim(V) > 0$ implies with the fundamental theorem that there exists λ with $p_A(\lambda) \;=\; 0$. This implies $\det(\lambda I - A) = 0$ and hence there exists $0 \neq v \in \ker(\lambda I - A)$. Using the transition matrix, we deduce that T has an eigenvector. ∎

Remark 4.5. *The eigenvalues are exactly the roots of the characteristic polynomial. Indeed, λ is an eigenvalue iff $\ker(\lambda I - A) \neq 0$ iff $\det(\lambda I - A) = 0$.*

5. Jordan normal form

Definition 5.1. *We say that A is similar to B if there exists an invertible map S such that $A = S^{-1}BS$.*

Theorem 5.2. *Let $A \in M_n(\mathbb{C})$ a complex matrix with*

$$p_A(x) \;=\; (x - \lambda_1)^{r_1} \cdots (x - \lambda_p)^{r_p}$$

and

$$m_A(x) \;=\; (x - \lambda_1)^{s_1} \cdots (x - \lambda_p)^{s_p} \,.$$

Then A is similar to a block matrix B with blocks

$$B_i \;=\; \begin{pmatrix} \lambda_i & 1 & 0 & 0\cdots & \\ 0 & \lambda_i & 1 & 0\cdots & \\ \vdots & & & & \\ 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & \cdots & 0 & 0 & \lambda_i \end{pmatrix}$$

Here not more than s_i blocks occur.

DEFINITION 5.3. *Let $T : V \to V$ be a linear map. We define*

$$\rho(T) = \{\lambda \in \mathbb{C} \ : \ (\lambda I - T)^{-1} \text{ exists } \}$$

and the spectrum $\sigma(T) \ = \ \mathbb{C}\rho(T)$.

REMARK 5.4. *$A \in M_n(\mathbb{C})$. The $\lambda \in \sigma(A)$ iff $p_A(\lambda) = 0$ if $m_A(\lambda) = 0$.*

LEMMA 5.5. *Let $A \in M_n(\mathbb{C})$. Let $\lambda \in \sigma(A)$. Then the seqeunce of subspaces Consider*

$$M^j \ = \ \ker(A - \lambda)^j$$

is ordered by inclusion. Moreover, the exists a minimal k such that $M^k = M^{k+1}$.

PROOF. Consider $d_j \ = \ \dim(M^j)$. Then d_j are integers and (d_j) is bounded by n. Thus $d = \lim_j d_j$ converges. Using $\varepsilon = \frac{1}{2}$ we see that for some $k \geq k_0$ we must have $d_j = d$. Hence, we define $k = \min\{j : d_j = d\}$. ∎

DEFINITION 5.6. *For two subspace V and W of $\mathbb{C}^n$ we write $V \oplus W = \mathbb{C}^n$ if $V + W = \mathbb{C}^n$ and $V \cap W = \{0\}$.*

dec LEMMA 5.7. *$M^k \oplus \mathrm{rg}(A - \lambda)^k \ = \ \mathbb{C}^n$ and $\mathrm{rg}(A - \lambda)^j = \mathrm{rg}(A - \lambda)^k$ for all $j \geq k$.*

PROOF. Note that

$$n \ = \ \dim(\ker(\lambda - A)^j) + \dim(\mathrm{rg}(\lambda - A)^j)) \ .$$

Moreover, the sequence $W_j = \mathrm{rg}((A-\lambda)^j)$ is decreasing. Thus $d_j \ = \ d_k$ for all $j \geq k$ implies $\dim(\mathrm{rg}(\lambda - A)^j)) = \dim(\mathrm{rg}(\lambda - A)^k))$ for all $j \ \geq \ k$. Thus $W_j = W_k$.
Now, let $v \in \mathbb{C}^n$. Define $w = (A - \lambda)^k v$. Note that $(A - \lambda)(W_k) \ = \ (A - \lambda)^k(A - \lambda)(\mathbb{C}^n) \ \subset \ W_k$. Moreover, $\dim(A - \lambda)(W_k) \ = \ \dim W_{k+1} \ = \ \dim W_k$ implies that $\lambda I - A$ is injective on W_k. Hence $(A - \lambda)^k$ is injective and surjective when restricted to W_k. Therefore we find $v_0 \in W_k$ such that $(A - \lambda)^k(v_0) \ = \ w$. Equivalently $v_0 - v \in \ker((A - \lambda)^k)$. Hence

$$v \ = \ v - v_0 + v_0 \in M^k + W_k \ .$$

This implies

$$\mathbb{C}^n \ = \ M^k + W_k$$

Using the dimension formula we must have $\dim(M^k \cap W_k) = 0$. ∎

In the following we use $\sigma(A) = \{\lambda_1, ..., \lambda_p\}$

$$M_i \;=\; \bigcup_l \ker((A-\lambda)^l) \quad , \quad W_i \;=\; \bigcap_l \mathrm{rg}((A-\lambda)^l)\;.$$

We denote by k_i the smallest integer from Lemma dec 5.7.

inj LEMMA 5.8. *$i \neq j$ implies $M_i \subset W_j$.*

PROOF. Let us show that $(A-\lambda_j)$ leaves M_i invariant. Indeed, let $x \in M_i$. Then

$$(A-\lambda_j)(x) \;=\; (\lambda_j-\lambda_i)(x) + (A-\lambda_i)(x) \in M_i\;.$$

Recall that for $x \in \ker((A-\lambda)^k)$ we know that

$$(A-\lambda)^k(\lambda_i I - A)x \;=\; \lambda(\lambda_i I - A)(A-\lambda_i)^k(x) \;=\; 0\;.$$

Thus we get $(\lambda_j - A)(M_i) \subset M_i$. Now, we want to show

$$\ker(\lambda_j - A) \cap M_i \;=\; \{0\}\;.$$

Indeed, let $x \in \ker((\lambda_j - A)) \cap M_i$. Then we have

$$\begin{aligned}(\lambda_j-\lambda_i)^k(x) &= ((A-\lambda_i)-(A-\lambda_j))^k(x)\\ &= (A-\lambda_i)^k(x) + \sum_{l=1}^{k-1}\binom{k}{l}(-1)^l((A-\lambda_i))^{k-l}(A-\lambda_j)^l(x) \;=\; 0\;.\end{aligned}$$

Therefore $(A-\lambda_j)$ is an isomorphism when restricted to M_i. Thus for every $x \in M_i$ we may find $y \in M_i$ such that

$$x \;=\; (A-\lambda_j)^k(y) \in W_i\;.$$

The assertion is proved. ■

LEMMA 5.9. *$\mathbb{C}^n \;=\; M_1 \oplus M_2 \oplus \cdots \oplus M_p$.*

PROOF. We have

$$\mathbb{C}^n \;=\; M_1 \oplus W_1$$

Every element $x \in W_1$ can be written in a unique way $x = z + y$, $z \in W_2$, $y \in M_2$. Since $y \in M_2$, we get $z \in W_2 \cap W_1$:

$$W_1 \;=\; W_1 \cap W_2 \oplus M_2\;.$$

Hence

$$\mathbb{C}^n \;=\; M_1 \oplus M_2 \oplus W_1 \cap W_2\;.$$

Since $M_3 \subset W_1 \cap W_2$, we continue and get

$$\mathbb{C}^n \;=\; M_1 \oplus \cdots \oplus M_p \oplus W_1 \cap \cdots \cap W_p \;.$$

Let $W = \bigcap_i W_i$. Note that

$$(A-\lambda_j)(W_i) \subset [(\lambda_j - \lambda_i)W_i + (A-\lambda_i)(W_i) \subset W_i \;.$$

Therefore $(A-\lambda_j)$ maps W into W for all $j = 1, ..., p$. In particular $A(W) \subset W$. Let us denote the induced linear map on W by T. If λ is an eigenvalue for T, then λ is an eigenvalue for A. Thus $\sigma(T) \subset \{\lambda_1, ..., \lambda_p\}$. However, for every $\lambda = \lambda_i$ every eigenvector x is contained in $\ker((A-\lambda_i)^{k_i})$. This yields $x \in M_i \cap W \subset M_i \cap W \;=\; \{0\}$. Therefore we have shown that T has no eigenvalues. According to Proposition exeigen 4.4 we must have $\dim(W) = 0$. ■

LEMMA 5.10. *The dimensions* $k_i = \dim(M_i)$ *coincide with degree* s_i *corresponding to* λ_i *in the minimal polynomial.*

PROOF. Let us recall that $M_i \;=\; \ker((A-\lambda_i)^{k_i})$ and hence

$$(A-\lambda_i)^{k_i}(M_i) \;=\; 0 \;.$$

Since $\mathbb{C}^n$ is a direct sum of the M_i's and the $(A-\lambda_i)^{k_i}$'s commute we get

$$(A-\lambda_1)^{k_1} \cdots (A-\lambda_p)^{k_p}(\mathbb{C}^n) \;=\; \{0\} \;.$$

This implies that the minimal polynomial m_A divides $q(x) \;=\; \prod_{i=1}^p (\lambda_i - x)$. Thus $s_i \leq k_i$ for all $i = 1, ..., p$. Suppose there exists i such that $s_i < k_i$. Then there exists an $0 \neq x \in M_i$ such that $(A-\lambda_i)^{s_i}(x) \neq 0$. We have shown in Lemma inj 5.8 that that for every $j \neq i$ the map $(A-\lambda_j)$ maps M_i to M_i and is injective. Therefore $\prod_{j\neq i}(A-\lambda_j)^{s_j}$ is injective on M_i. We get

$$m_A(A)(x) = [\prod_{j\neq i}(A-\lambda_j)^{s_j}](A-\lambda_i)^{s_i} x \neq 0 \;.$$

Thus $m_A(A) \neq 0$ and this contradiction concludes the proof. ■

The next result concludes our proof of the existence of the Jordan normal form.

PROPOSITION 5.11. *Let* $\lambda_i \in \sigma(A)$ *and for* $j = 1, ..., k = s_i$ *we define*

$$\Delta_j \;=\; \dim(\ker(A-\lambda_i)^j) - \dim(\ker(A-\lambda_i)^{j-1})$$

and $\partial_k = \Delta_k$ *and*

$$\partial_j \;=\; \Delta_j - \Delta_{j+1} \;.$$

Then the restriction of A to M_i is similar to a direct sum of ∂_j Jordan blocks of length j.

PROOF. In the following $\lambda = \lambda_i$ and $k = s_i$. We consider $M = \ker((A-\lambda)^k))$. We decompose

$$M = \ker((A-\lambda)^{k-1} \oplus \ker((A-\lambda)^k)/\ker((A-\lambda)k-1)$$

Let $[b_1], ..., [b_m]$ be a basis for the quotient space. Note that $m = \Delta_k = \partial_k$. Let us show that

$$S = \{(A-\lambda)^j(b_i) \,:\, 0 \leq j < k, 1 \leq i \leq m\}$$

is a system of linear independent vectors. Indeed, assume

$$\sum_{ij} a_{ij}(A-\lambda)^j(b_i) = 0$$

Since $(A-\lambda)^{k-1}(A-\lambda)^j = 0$ for $j \geq 1$, we get

$$\sum_{i=1}^{m} a_{i0}(A-\lambda)^{k-1}(b_i) = 0 \,.$$

This implies $\sum_i a_{i,0}b_i \in \ker((A-\lambda)^{k-1}$. By linear independence of the $[b_1], ..., [b_m]$ we deduce $a_{i,0} = 0$ for $i = 1, ..., m$. Similarly, we assume

$$\sum_{i=1}^{m}\sum_{j=1}^{k-1} a_{ij}(A-\lambda)^j(b_i) = 0$$

and deduce $a_{11}, ..., a_{m1} = 0$. Inductively, we find $a_{ij} = 0$. For a fixed $1 \leq i \leq m$ we consider

$$x_j = (A-\lambda)^j(b_i) \,.$$

Note that

$$A(x_j) = (A-\lambda)(x_j) + \lambda x_j = x_{j+1} + \lambda x_j$$

for $j = 0, ..., k-1$ but $(A-\lambda)(x_{k-1}) = (A-\lambda)^k(b_1) = 0$. Hence x_{k-1} is an eigenvector. We see that A leaves

$$F_i = \operatorname{span}\{x_0, ..., x_{k-1}\}$$

invariant. In this basis we finally find the Jordan block

$$A|_{F_i} \cong \begin{pmatrix} \lambda & 1 & 0 & 0\cdots & \\ 0 & \lambda & 1 & 0\cdots & \\ \vdots & & & & \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} .$$

Using the reversed order $v_{(i-1)m+j} = x_{k-j}$ we obtain (finally) a Jordan block and a basis $(v_j)_{1\leq j\leq mk}$ for m blocks of length k. Now, we proceed inductively and consider $\ker((A-\lambda)^{k-1})/\ker((A-\lambda)^{k-1})$. Then vectors $c_1, ..., c_m = [(A-\lambda)(b_i)]$ are already linearly independent. Thus we may complete this system with linearly independent vectors $[B_1], ..., [B_l]$ where

$$l \;=\; \Delta_{k-1} - \Delta_k \;=\; \partial k \;.$$

The descendants of the B_i form l Jordan normal blocks of size $k-1$. In general, we have

$$\partial_j \;=\; \Delta_j - (\partial_k + \cdots + \partial_{j+1})$$

many Jordan blocks of size j. For example, we have

$$\partial_k + \partial_{k-1} \;=\; \Delta_k + \Delta_{k-1} - \Delta_k \;=\; \Delta_{k-1} \;.$$

By induction, we get

$$\partial_k + \cdots + \partial_{j+1} \;=\; \Delta_{j+1} \;.$$

Our claim is proved. ■

Remark 5.12. *The JNF is uniquely determined by the dimensions ∂_j. It is unique up to permutation of the blocks.*

Corollary 5.13. $\Delta_k \leq \Delta_j$ *for* $j = 1, ..., k$ *and*

$$k_i = \min\{j : \dim(\ker(A-\lambda_i)^j) = \dim(\ker(A-\lambda_i)^{j+}) \;.$$

Proof. In fact we have seen that for $b_1, ..., b_m$ such that $[b_i] = \ker((A-\lambda)^k)/\ker((A-\lambda)^{k-1})$ are linearly independent we have $(A-\lambda)^{k-j}(b_i)$ are linearly independent. This yields $\Delta_k \leq \Delta_k$. Since $\Delta_k > 1$ by definition, we deduce $d_{j+1} > d_j$ for all $j = 1, ..., k-1$. ■

Let us consider an example

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & -1 & -1 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

We have ${}_A(x) = \det(A-x)^2 = (2-x)^5$. Thus the eigen value 2. Consider

$$(A-2)^2 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = 0\,.$$

The minimal polynomial is $m_A(x) = (2-x)^2$. The kernel of $(A-2)$ is the span$\{e_1, e_2 - e_3, e_4 - e_5\}$. Thus A is similar to

$$B = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

In fact I constructed A as

$$A = S^{-1}BS$$

where

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$