Chapter 15
Polynomials and the Fundamental Theorem of Algebra

In this chapter we discuss polynomials in more detail. We consider the division of polynomials and derive classical results from polynomial algebra, including the factorization into irreducible factors. We also prove the Fundamental Theorem of Algebra, which states that every non-constant polynomial over the complex numbers has a least one complex root. This implies that every complex matrix and every endomorphism on a (finite dimensional) complex vector space has at least one eigenvalue.

15.1 Polynomials

Let us recall some of the most important terms in the context of polynomials. If K is a field, then

$$p = \alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \quad \text{with} \quad n \in \mathbb{N}_0 \quad \text{and} \quad \alpha_0, \alpha_1, \ldots \alpha_n \in K$$

is a polynomial over K in the variable t. The set $K[t]$ of all these polynomials forms a commutative ring with unit (cp. Example 3.17). If $\alpha_n \neq 0$, then $\deg(p) = n$ is called the *degree* of p. If $\alpha_n = 1$, then p is called *monic*. If $p = 0$, then $\deg(p) := -\infty$, and if $\deg(p) < 1$, then p is called *constant*.

Lemma 15.1 *For two polynomials $p, q \in K[t]$ the following assertions hold:*

(1) $\deg(p + q) \leq \max\{\deg(p), \deg(q)\}$.
(2) $\deg(p \cdot q) = \deg(p) + \deg(q)$.

Proof Exercise. □

We now introduce some concepts associated with the division of polynomials.

J. Liesen and V. Mehrmann, *Linear Algebra*, Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-24346-7_15

Definition 15.2 Let K be a field.

(1) If for two polynomials $p, s \in K[t]$ there exists a polynomial $q \in K[t]$ with $p = s \cdot q$, then s is called a *divisor* of p and we write $s|p$ (read this as "s divides p").
(2) Two polynomials $p, s \in K[t]$ are called *coprime*, if $q|p$ and $q|s$ for some $q \in K[t]$ always imply that q is constant.
(3) A non-constant polynomial $p \in K[t]$ is called *irreducible* (over K), if $p = s \cdot q$ for two polynomials $s, q \in K[t]$ implies that s or q are constant. If there exist two non-constant polynomials $s, q \in K[t]$ with $p = s \cdot q$, then p is called *reducible* (over K).

Note that the property of irreducibility is only defined for polynomials of degree at least 1. A polynomial of degree 1 is always irreducible. Whether a polynomial of degree at least 2 is irreducible may depend on the underlying field.

Example 15.3 The polynomial $2 - t^2 \in \mathbb{Q}[t]$ is irreducible, but the factorization

$$2 - t^2 = \left(\sqrt{2} - t\right) \cdot \left(\sqrt{2} + t\right)$$

shows that $2 - t^2 \in \mathbb{R}[t]$ is reducible. The polynomial $1 + t^2 \in \mathbb{R}[t]$ is irreducible, but using the imaginary unit $\mathbf{i}$ we have

$$1 + t^2 = (-\mathbf{i} + t) \cdot (\mathbf{i} + t),$$

so that $1 + t^2 \in \mathbb{C}[t]$ is reducible.

The next result concerns the *division with remainder* of polynomials.

Theorem 15.4 *If $p \in K[t]$ and $s \in K[t] \setminus \{0\}$, then there exist uniquely defined polynomials $q, r \in K[t]$ with*

$$p = s \cdot q + r \quad \text{and} \quad \deg(r) < \deg(s). \tag{15.1}$$

Proof We show first the existence of polynomials $q, r \in K[t]$ such that (15.1) holds.

If $\deg(s) = 0$, then $s = s_0$ for an $s_0 \in K \setminus \{0\}$ and (15.1) follows with $q := s_0^{-1} \cdot p$ and $r := 0$, where $\deg(r) < \deg(s)$.

We now assume that $\deg(s) \geq 1$. If $\deg(p) < \deg(s)$, then we set $q := 0$ and $r := p$. Then $p = s \cdot q + r$ with $\deg(r) < \deg(s)$.

Let $n := \deg(p) \geq m := \deg(s) \geq 1$. We prove (15.1) by induction on n. If $n = 1$, then $m = 1$. Hence $p = p_1 \cdot t + p_0$ with $p_1 \neq 0$ and $s = s_1 \cdot t + s_0$ with $s_1 \neq 0$. Therefore,

$$p = s \cdot q + r \quad \text{for} \quad q := p_1 s_1^{-1}, \quad r := p_0 - p_1 s_1^{-1} s_0,$$

where $\deg(r) < \deg(s)$.

Suppose that the assertion holds for an $n \geq 1$. Let two polynomials p and s with $n+1 = \deg(p) \geq \deg(s) = m$ be given, and let $p_{n+1}(\neq 0)$ and $s_m(\neq 0)$ be the highest coefficients of p and s. If

$$h := p - p_{n+1}s_m^{-1}s \cdot t^{n+1-m} \in K[t],$$

then $\deg(h) < \deg(p) = n+1$. By the induction hypothesis there exist polynomials $\widetilde{q}, r \in K[t]$ with

$$h = s \cdot \widetilde{q} + r \quad \text{and} \quad \deg(r) < \deg(s).$$

It then follows that

$$p = s \cdot q + r \quad \text{with} \quad q := \widetilde{q} + p_{n+1}s_m^{-1}t^{n+1-m},$$

where $\deg(r) < \deg(s)$.

It remains to show the uniqueness. Suppose that (15.1) holds and that there exist polynomials $\widehat{q}, \widehat{r} \in K[t]$ with $p = s \cdot \widehat{q} + \widehat{r}$ and $\deg(\widehat{r}) < \deg(s)$. Then

$$r - \widehat{r} = s \cdot (\widehat{q} - q).$$

If $\widehat{r} - r \neq 0$, then $\widehat{q} - q \neq 0$ and thus

$$\deg(r - \widehat{r}) = \deg(s \cdot (\widehat{q} - q)) = \deg(s) + \deg(\widehat{q} - q) \geq \deg(s).$$

On the other hand, we also have

$$\deg(r - \widehat{r}) \leq \max\{\deg(r), \deg(\widehat{r})\} < \deg(s).$$

This is a contradiction, which shows that indeed $r = \widehat{r}$ and $q = \widehat{q}$. □

This theorem has some important consequences for the roots of polynomials. The first of these is known as the *Theorem of Ruffini*.[1]

Corollary 15.5 *If $\lambda \in K$ is a root of $p \in K[t]$, i.e., $p(\lambda) = 0$, then there exists a uniquely determined polynomial $q \in K[t]$ with $p = (t - \lambda) \cdot q$.*

Proof When we apply Theorem 15.4 to the polynomials p and $s = t - \lambda \neq 0$, then we get uniquely determined polynomials q and r with $\deg(r) < \deg(s) = 1$ and

$$p = (t - \lambda) \cdot q + r.$$

The polynomial r is constant and evaluating it at λ gives

$$0 = p(\lambda) = (\lambda - \lambda) \cdot q(\lambda) + r(\lambda) = r(\lambda),$$

[1] Paolo Ruffini (1765–1822).

which yields $r = 0$ and $p = (t - \lambda) \cdot q$. □

If a polynomial $p \in K[t]$ has at least degree 2 and a root $\lambda \in K$, then the linear factor $t - \lambda$ is a divisor of p and, in particular, p is reducible. The converse of this statement *does not hold*. For instance the polynomial $4-4t^2+t^4 = (2-t^2)\cdot(2-t^2) \in \mathbb{Q}[t]$ is reducible, but it does not have a root in $\mathbb{Q}$.

Corollary 15.5 motivates the following definition.

Definition 15.6 If $\lambda \in K$ is a root of $p \in K[t] \setminus \{0\}$, then its *multiplicity* is the uniquely determined nonnegative integer m, such that $p = (t - \lambda)^m \cdot q$ for a polynomial $q \in K[t]$ with $q(\lambda) \neq 0$.

Recursive application of Corollary 15.5 to a given polynomial $p \in K[t]$ leads to the following result.

Corollary 15.7 *If $\lambda_1 \ldots, \lambda_k \in K$ are pairwise distinct roots of $p \in K[t] \setminus \{0\}$ with the corresponding multiplicities $m_1, \ldots, m_k$, then there exists a unique polynomial $q \in K[t]$ with*

$$p = (t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_k)^{m_k} \cdot q$$

and $q(\lambda_j) \neq 0$ for $j = 1, \ldots, k$. In particular, the sum of the multiplicities of all pairwise distinct roots of p is at most $\deg(p)$.

The next result is known as the *Lemma of Bézout*.[2]

Lemma 15.8 *If $p, s \in K[t] \setminus \{0\}$ are coprime, then there exist polynomials $q_1, q_2 \in K[t]$ with*

$$p \cdot q_1 + s \cdot q_2 = 1.$$

Proof We may assume without loss of generality that $\deg(p) \geq \deg(s)$ (≥ 0), and we proceed by induction on $\deg(s)$.

If $\deg(s) = 0$, then $s = s_0$ for an $s_0 \in K \setminus \{0\}$, and thus

$$p \cdot q_1 + s \cdot q_2 = 1 \quad \text{with} \quad q_1 := 0, \quad q_2 := s_0^{-1}.$$

Suppose that the assertion holds for all polynomials $p, s \in K[t] \setminus \{0\}$ with $\deg(s) = n$ for an $n \geq 0$. Let $p, s \in K[t] \setminus \{0\}$ with $\deg(p) \geq \deg(s) = n + 1$ be given. By Theorem 15.4 there exist polynomials q and r with

$$p = s \cdot q + r \quad \text{and} \quad \deg(r) < \deg(s).$$

Here we have $r \neq 0$, since by assumption p and s are coprime.

Suppose that there exists a non-constant polynomial $h \in K[t]$ that divides both s and r. Then h also divides p, in contradiction to the assumption that p and s are coprime. Thus, the polynomials s and r are coprime. Since $\deg(r) < \deg(s)$, we can

[2] Étienne Bézout (1730–1783).

apply the induction hypothesis to the polynomials $s, r \in K[t] \setminus \{0\}$. Hence there exist polynomials $\widetilde{q}_1, \widetilde{q}_2 \in K[t]$ with

$$s \cdot \widetilde{q}_1 + r \cdot \widetilde{q}_2 = 1.$$

From $r = p - s \cdot q$ we then get

$$1 = s \cdot \widetilde{q}_1 + (p - s \cdot q) \cdot \widetilde{q}_2 = p \cdot \widetilde{q}_2 + s \cdot (\widetilde{q}_1 - q \cdot \widetilde{q}_2),$$

which completes the proof. □

Using the Lemma of Bézout we can easily prove the following result.

Lemma 15.9 *If $p \in K[t]$ is irreducible and a divisor of the product $s \cdot h$ of two polynomials $s, h \in K[t]$, then p divides at least one of the factors, i.e., $p|s$ or $p|h$.*

Proof If $s = 0$, then $p|s$, because every polynomial is a divisor of the zero polynomial.

If $s \neq 0$ and p is not a divisor of s, then p and s are coprime, since p is irreducible. By Lemma 15.8 there exist polynomials $q_1, q_2 \in K[t]$ with $p \cdot q_1 + s \cdot q_2 = 1$, and hence

$$h = h \cdot 1 = (q_1 \cdot h) \cdot p + q_2 \cdot (s \cdot h).$$

The polynomial p divides both terms on the right hand side, and thus also $p|h$. □

By recursive application of Lemma 15.9 we obtain the *Euclidean theorem*, which describes a prime factor decomposition in the ring of polynomials.

Theorem 15.10 *Every polynomial $p = \alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \in K[t] \setminus \{0\}$ has a unique (up to the ordering of the factors) decomposition*

$$p = \mu \cdot p_1 \cdot \ldots \cdot p_k$$

with $\mu \in K$ and monic irreducible polynomials $p_1, \ldots, p_k \in K[t]$.

Proof If $\deg(p) = 0$, and thus $p = \alpha_0$, then the assertion holds with $k = 0$ and $\mu = \alpha_0$.

Let $\deg(p) \geq 1$. If p is irreducible, then the assertion holds with $p_1 = \mu^{-1} p$ and $\mu = \alpha_n$. If p is reducible, then $p = p_1 \cdot p_2$ for two non-constant polynomials p_1 and p_2. These are either irreducible, or we can decompose them further. Every multiplicative decomposition of p that is obtained in this way has at most $\deg(p) = n$ non-constant factors. Suppose that

$$p = \mu \cdot p_1 \cdot \ldots \cdot p_k = \beta \cdot q_1 \cdot \ldots \cdot q_\ell \tag{15.2}$$

for some k, ℓ, where $1 \leq \ell \leq k \leq n$, $\mu, \beta \in K$, as well as monic irreducible polynomials $p_1, \ldots, p_k, q_1, \ldots, q_\ell \in K[t]$. Then $p_1|p$ and hence $p_1|q_j$ for some j. Since the polynomials p_1 and q_j are irreducible, we must have $p_1 = q_j$.

We may assume without loss of generality that $j = 1$ and cancel the polynomial $p_1 = q_1$ in the identity (15.2), which gives

$$\mu \cdot p_2 \cdot \ldots \cdot p_k = \beta \cdot q_2 \cdot \ldots \cdot q_\ell .$$

Proceeding analogously for the polynomials $p_2, \ldots, p_k$, we finally obtain $k = \ell$, $\mu = \beta$ and $p_j = q_j$ for $j = 1, \ldots, k$. □

15.2 The Fundamental Theorem of Algebra

We have seen above that the existence of roots of a polynomial depends on the field over which it is considered. The field $\mathbb{C}$ is special in this sense, since here the Fundamental Theorem of Algebra[3] guarantees that every non-constant polynomial has a root. In order to use this theorem in our context, we first present an equivalent formulation in the language of Linear Algebra.

Theorem 15.11 *The following statements are equivalent:*

(1) Every non-constant polynomial $p \in \mathbb{C}[t]$ has a root in $\mathbb{C}$.
(2) If $\mathcal{V} \neq \{0\}$ is a finite dimensional $\mathbb{C}$-vector space, then every endomorphism $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ has an eigenvector.

Proof

(1) $\Rightarrow$ (2): If $\mathcal{V} \neq \{0\}$ and $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$, then the characteristic polynomial $P_f \in \mathbb{C}[t]$ is non-constant, since $\deg(P_f) = \dim(\mathcal{V}) > 0$. Thus, P_f has a root in $\mathbb{C}$, which is an eigenvalue of f, so that f indeed has an eigenvector.

(2) $\Rightarrow$ (1): Let $p = \alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \in \mathbb{C}[t]$ be a non-constant polynomial with $\alpha_n \neq 0$. The roots of p are equal to the roots of the monic polynomial $\widehat{p} := \alpha_n^{-1} p$. Let $A \in \mathbb{C}^{n,n}$ be the companion matrix of $\widehat{p}$, then $P_A = \widehat{p}$ (cp. Lemma 8.4).
If $\mathcal{V}$ is an n-dimensional $\mathbb{C}$-vector space and B is an arbitrary basis of $\mathcal{V}$, then there exists a uniquely determined $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ with $[f]_{B,B} = A$ (cp. Theorem 10.16). By assumption, f has an eigenvector and hence also an eigenvalue, so that $\widehat{p} = P_A$ has a root. □

The Fundamental Theorem of Algebra cannot be proven without tools from Analysis. In particular, one needs that polynomials are *continuous*. We will use the following standard result, which is based on the continuity of polynomials.

Lemma 15.12 *Every polynomial $p \in \mathbb{R}[t]$ with odd degree has a (real) root.*

[3] Numerous proofs of this important result exist. Carl Friedrich Gauß (1777–1855) alone gave four different proofs, starting with the one in his dissertation from 1799, which contained however a gap. The history of this result is described in detail in the book [Ebb91].

Proof Let the highest coefficient of p be positive. Then

$$\lim_{t\to\infty} p(t) = +\infty, \quad \lim_{t\to-\infty} p(t) = -\infty.$$

Since the real function $p(t)$ is continuous, the *Intermediate Value Theorem* from Analysis implies the existence of a root of p. The argument in the case of a negative leading coefficient is analogous. □

Our proof of the Fundamental Theorem of Algebra below follows the presentation in the article [Der03]. The proof is by induction on the dimension of $\mathcal{V}$. However, we do not use the usual consecutive order, i.e., $\dim(\mathcal{V}) = 1, 2, 3, \ldots$, but an order that is based on the sets

$$M_j := \{2^m \cdot \ell \mid 0 \le m \le j-1, \ \ \ell \ \text{ odd}\} \subset \mathbb{N}, \quad j = 1, 2, 3, \ldots .$$

For instance,

$$M_1 = \{\ell \mid \ell \ \text{ odd}\} = \{1, 3, 5, 7, \ldots\}, \quad M_2 = M_1 \cup \{2, 6, 10, 14, \ldots\}.$$

Lemma 15.13

(1) If $\mathcal{V}$ is an $\mathbb{R}$-vector space and if $\dim(\mathcal{V})$ is odd, i.e., $\dim(\mathcal{V}) \in M_1$, then every $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ has an eigenvector.

(2) Let K be a field and $j \in \mathbb{N}$. If for every K-vector space $\mathcal{V}$ with $\dim(\mathcal{V}) \in M_j$ every $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ has an eigenvector, then two commuting $f_1, f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ have a common eigenvector. That is, if $f_1 \circ f_2 = f_2 \circ f_1$, then there exists a vector $v \in \mathcal{V} \setminus \{0\}$ and two scalars $\lambda_1, \lambda_2 \in K$ with $f_1(v) = \lambda_1 v$ and $f_2(v) = \lambda_2 v$.

(3) If $\mathcal{V}$ is an $\mathbb{R}$-vector space and if $\dim(\mathcal{V})$ is odd, then two commuting $f_1, f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ have a common eigenvector.

Proof

(1) For every $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ the degree of $P_f \in \mathbb{R}[t]$ is odd. Hence Lemma 15.12 implies that P_f has a root, and therefore f has an eigenvector.

(2) We proceed by induction on $\dim(\mathcal{V})$, where $\dim(\mathcal{V})$ runs through the elements of M_j in increasing order. The set M_j is a proper subset of $\mathbb{N}$ consisting of natural numbers that are not divisible by 2^j and, in particular, 1 is the smallest element of M_j.
If $\dim(\mathcal{V}) = 1 \in M_j$, then by assumption two arbitrary $f_1, f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ each have an eigenvector, i.e.,

$$f_1(v_1) = \lambda_1 v_1, \quad f_2(v_2) = \lambda_2 v_2.$$

Since $\dim(\mathcal{V}) = 1$, we have $v_1 = \alpha v_2$ for an $\alpha \in K \setminus \{0\}$. Thus,

$$f_2(v_1) = f_2(\alpha v_2) = \alpha f_2(v_2) = \lambda_2(\alpha v_2) = \lambda_2 v_1,$$

i.e., v_1 is a common eigenvector of f_1 and f_2.
Let now $\dim(\mathcal{V}) \in M_j$, and let the assertion be proven for all K-vector spaces whose dimensions is an element of M_j that is smaller than $\dim(\mathcal{V})$. Let $f_1, f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ with $f_1 \circ f_2 = f_2 \circ f_1$. By assumption, f_1 has an eigenvector v_1 with corresponding eigenvalue λ_1. Let

$$\mathcal{U} := \operatorname{im}(\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1), \qquad \mathcal{W} := \mathcal{V}_{f_1}(\lambda_1) = \ker(\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1).$$

The subspaces $\mathcal{U}$ and $\mathcal{W}$ of $\mathcal{V}$ are f_1-invariant, i.e., $f_1(\mathcal{U}) \subseteq \mathcal{U}$ and $f_1(\mathcal{W}) \subseteq \mathcal{W}$. For the space $\mathcal{W}$ we have shown this in Lemma 14.4 and for the space $\mathcal{U}$ this can be easily shown as well (cp. Exercise 14.1). The subspaces $\mathcal{U}$ and $\mathcal{W}$ are also f_2-invariant:
If $u \in \mathcal{U}$, then $u = (\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1)(v)$ for a $v \in \mathcal{V}$. Since f_1 and f_2 commute, we have

$$\begin{aligned} f_2(u) &= (f_2 \circ (\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1))(v) = ((\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1) \circ f_2)(v) \\ &= (\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1)(f_2(v)) \in \mathcal{U}. \end{aligned}$$

If $w \in \mathcal{W}$, then

$$\begin{aligned} (\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1)(f_2(w)) &= ((\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1) \circ f_2)(w) = (f_2 \circ (\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1))(w) \\ &= f_2((\lambda_1 \mathrm{Id}_{\mathcal{V}} - f_1)(w)) = f_2(0) = 0, \end{aligned}$$

hence $f_2(w) \in \mathcal{W}$.
We have $\dim(\mathcal{V}) = \dim(\mathcal{U}) + \dim(\mathcal{W})$ and since $\dim(\mathcal{V})$ is not divisible by 2^j, either $\dim(\mathcal{U})$ or $\dim(\mathcal{W})$ is not divisible by 2^j. Hence either $\dim(\mathcal{U}) \in M_j$ or $\dim(\mathcal{W}) \in M_j$.
If the corresponding subspace is a proper subspace of $\mathcal{V}$, then its dimension is an element of M_j that is smaller than $\dim(\mathcal{V})$. By the induction hypothesis then f_1 and f_2 have a common eigenvector in this subspace. Thus, f_1 and f_2 have a common eigenvector in $\mathcal{V}$.
If the corresponding subspace is equal to $\mathcal{V}$, then this must be the subspace $\mathcal{W}$, since $\dim(\mathcal{W}) \geq 1$. But if $\mathcal{V} = \mathcal{W}$, then every vector in $\mathcal{V} \setminus \{0\}$ is an eigenvector of f_1. By assumption also f_2 has an eigenvector, so that there exists at least one common eigenvector of f_1 and f_2.

(3) By (1) it follows that the assumption of (2) holds for $K = \mathbb{R}$ and $j = 1$, which means that (3) holds as well. □

We will now prove the Fundamental Theorem of Algebra in the formulation (2) of Theorem 15.11.

Theorem 15.14 *If $\mathcal{V} \neq \{0\}$ is a finite dimensional $\mathbb{C}$-vector space, then every $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ has an eigenvector.*

Proof We prove the assertion by induction on $j = 1, 2, 3, \ldots$ and $\dim(\mathcal{V}) \in M_j$.

We start with $j = 1$ and thus by showing the assertion for all $\mathbb{C}$-vector spaces of odd dimension. Let $\mathcal{V}$ be an arbitrary $\mathbb{C}$-vector space with $n := \dim(\mathcal{V}) \in M_1$. Let $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ and consider an arbitrary scalar product on $\mathcal{V}$ (such a scalar product always exists; cp. Exercise 12.1), as well as the set of self-adjoint maps with respect to this scalar product,

$$\mathcal{H} := \{g \in \mathcal{L}(\mathcal{V}, \mathcal{V}) \mid g = g^{ad}\}.$$

By Lemma 13.15 the set $\mathcal{H}$ forms an $\mathbb{R}$-vector space of dimension n^2. If we define $h_1, h_2 \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ by

$$h_1(g) := \frac{1}{2}(f \circ g + g \circ f^{ad}), \quad h_2(g) := \frac{1}{2\mathbf{i}}(f \circ g - g \circ f^{ad})$$

for all $g \in \mathcal{H}$, then $h_1 \circ h_2 = h_2 \circ h_1$ (cp. Exercise 15.8). Since n is odd, also n^2 is odd. By (3) in Lemma 15.13, h_1 and h_2 have a common eigenvector. Hence, there exists a $\widetilde{g} \in \mathcal{H} \setminus \{0\}$ with

$$h_1(\widetilde{g}) = \lambda_1 \widetilde{g}, \quad h_2(\widetilde{g}) = \lambda_2 \widetilde{g} \quad \text{for some} \quad \lambda_1, \lambda_2 \in \mathbb{R}.$$

We have $(h_1 + \mathbf{i}h_2)(g) = f \circ g$ for all $g \in \mathcal{H}$ and therefore, in particular,

$$f \circ \widetilde{g} = (h_1 + \mathbf{i}h_2)(\widetilde{g}) = (\lambda_1 + \mathbf{i}\lambda_2)\widetilde{g}.$$

Since $\widetilde{g} \neq 0$, there exists a $v \in \mathcal{V}$ with $\widetilde{g}(v) \neq 0$. Then

$$f(\widetilde{g}(v)) = (\lambda_1 + \mathbf{i}\lambda_2)\,(\widetilde{g}(v)),$$

which shows that $\widetilde{g}(v) \in \mathcal{V}$ is an eigenvector of f, so that the proof for $j = 1$ is complete.

Assume now that for some $j \geq 1$ and every $\mathbb{C}$-vector space $\mathcal{V}$ with $\dim(\mathcal{V}) \in M_j$, every $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ has an eigenvector. Then (2) in Lemma 15.13 implies that every two commuting $f_1, f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ have a common eigenvector.

We have to show that for every $\mathbb{C}$-vector space $\mathcal{V}$ with $\dim(\mathcal{V}) \in M_{j+1}$, every $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ has an eigenvector. Since

$$M_{j+1} = M_j \cup \{2^j q \mid q \text{ odd}\},$$

we only have to prove this for $\mathbb{C}$-vector spaces $\mathcal{V}$ with $n := \dim(\mathcal{V}) = 2^j q$ for odd q. Let $\mathcal{V}$ be such a vector space and let $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ be given. We choose an arbitrary basis of $\mathcal{V}$ and denote the matrix representation of f with respect to this basis by $A \in \mathbb{C}^{n,n}$. Let

$$\mathcal{S} := \{B \in \mathbb{C}^{n,n} \mid B = B^T\}$$

be the set of complex symmetric $n \times n$ matrices. If we define $h_1, h_2 \in \mathcal{L}(\mathcal{S}, \mathcal{S})$ by

$$h_1(B) := AB + BA^T, \quad h_2(B) := ABA^T$$

for all $B \in \mathcal{S}$, then $h_1 \circ h_2 = h_2 \circ h_1$ (cp. Exercise 15.9). By Lemma 13.16 the set $\mathcal{S}$ forms a $\mathbb{C}$-vector space of dimension $n(n+1)/2$. We have $n = 2^j q$ for an odd natural number q. Thus,

$$\frac{n(n+1)}{2} = \frac{2^j q\,(2^j q + 1)}{2} = 2^{j-1} q \cdot (2^j q + 1) \in M_j.$$

By the induction hypothesis, the commuting endomorphisms h_1 and h_2 have a common eigenvector. Hence there exists a $\widetilde{B} \in \mathcal{S} \setminus \{0\}$ with

$$h_1(\widetilde{B}) = \lambda_1 \widetilde{B}, \quad h_2(\widetilde{B}) = \lambda_2 \widetilde{B} \quad \text{for some} \quad \lambda_1, \lambda_2 \in \mathbb{C}.$$

In particular, we have $\lambda_1 \widetilde{B} = A\widetilde{B} + \widetilde{B}A^T$. Multiplying this equation from the left with A yields

$$\lambda_1 A\widetilde{B} = A^2\widetilde{B} + A\widetilde{B}A^T = A^2\widetilde{B} + h_2(\widetilde{B}) = A^2\widetilde{B} + \lambda_2 \widetilde{B},$$

so that

$$\left(A^2 - \lambda_1 A + \lambda_2 I_n\right) \widetilde{B} = 0.$$

We now factorize $t^2 - \lambda_1 t + \lambda_2 = (t - \alpha)(t - \beta)$ with

$$\alpha = \frac{\lambda_1 + \sqrt{\lambda_1^2 - 4\lambda_2}}{2}, \quad \beta = \frac{\lambda_1 - \sqrt{\lambda_1^2 - 4\lambda_2}}{2},$$

where we have used that every complex number has a square root. Then

$$(A - \alpha I_n)(A - \beta I_n)\, \widetilde{B} = 0.$$

Since $\widetilde{B} \neq 0$, there exists a $v \in \mathbb{C}^{n,1}$ with $\widetilde{B}v \neq 0$. If $(A - \beta I_n)\widetilde{B}v = 0$, then $\widetilde{B}v$ is an eigenvector of A corresponding to the eigenvalue β. If $(A - \beta I_n)\widetilde{B}v \neq 0$, then $(A - \beta I_n)\widetilde{B}v$ is an eigenvector of A corresponding to the eigenvalue α. Since A has an eigenvector, also f has an eigenvector. □

MATLAB-Minute.
Compute the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \\ 2 & 3 & 4 & 1 & 5 \\ 5 & 1 & 4 & 2 & 3 \\ 4 & 2 & 3 & 1 & 5 \end{bmatrix} \in \mathbb{R}^{5,5}$$

using the command `eig(A)`.
By definition a real matrix A can only have real eigenvalues. The reason for the occurrence of complex eigenvalues is that MATLAB interprets *every* matrix as a complex matrix. This means that within MATLAB *every* matrix can be unitarily triangulated, since every complex polynomial (of degree at least 1) decomposes into linear factors.

As a direct corollary of the Fundamental Theorem of Algebra and (2) in Lemma 15.13 we have the following result.

Corollary 15.15 *If $\mathcal{V} \neq \{0\}$ is a finite dimensional $\mathbb{C}$-vector space, then two commuting $f_1, f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ have a common eigenvector.*

Example 15.16 The two complex 2×2 matrices

$$A = \begin{bmatrix} \mathbf{i} & 1 \\ 1 & \mathbf{i} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2\mathbf{i} & 1 \\ 1 & 2\mathbf{i} \end{bmatrix}$$

commute. The eigenvalues of A are $\pm 1 + \mathbf{i}$ and those of B are $\pm 2 + \mathbf{i}$. Hence A and B do not have a common eigenvalue, while $[1,\ 1]^T$ and $[-1,\ 1]^T$ are common eigenvectors of A and B.

Using Corollary 15.15, Schur's theorem (Corollary 14.20) can be generalized as follows.

Theorem 15.17 *If $\mathcal{V} \neq \{0\}$ is a finite dimensional unitary vector space and $f_1,\ f_2 \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ commute, then f_1 and f_2 can be* simultaneously unitarily triangulated*, i.e., there exists an orthonormal basis B of $\mathcal{V}$, such that $[f_1]_{B,B}$ and $[f_2]_{B,B}$ are both upper triangular.*

Proof Exercise. □

Exercises

(In the following exercises K is an arbitrary field.)

15.1. Prove Lemma 15.1.

15.2. Show the following assertions for $p_1, p_2, p_3 \in K[t]$:

(a) $p_1|(p_1 p_2)$.
(b) $p_1|p_2$ and $p_2|p_3$ imply that $p_1|p_3$.
(c) $p_1|p_2$ and $p_1|p_3$ imply that $p_1|(p_2 + p_3)$.
(d) If $p_1|p_2$ and $p_2|p_1$, then there exists a $c \in K \setminus \{0\}$ with $p_1 = cp_2$.

15.3. Examine whether the following polynomials are irreducible:

$$p_1 = t^3 - t^2 + t - 1 \in \mathbb{Q}[t], \qquad p_4 = t^3 - t^2 + t - 1 \in \mathbb{R}[t],$$
$$p_2 = t^3 - t^2 + t - 1 \in \mathbb{C}[t], \qquad p_5 = 4t^3 - 4t^2 - t + 1 \in \mathbb{Q}[t],$$
$$p_3 = 4t^3 - 4t^2 - t + 1 \in \mathbb{R}[t], \qquad p_6 = t^3 - 4t^2 - t + 1 \in \mathbb{C}[t].$$

Determine the decompositions into irreducible factors.

15.4. Decompose the polynomials $p_1 = t^2 - 2$, $p_2 = t^2 + 2$, $p_3 = t^4 - 1$ and $p_4 = t^2 + t + 1$ into irreducible factors over the fields $K = \mathbb{Q}$, $K = \mathbb{R}$ and $K = \mathbb{C}$.

15.5. Show the following assertions for $p \in K[t]$:

(a) If $\deg(p) = 1$, then p is irreducible.
(b) If $\deg(p) \geq 2$ and p has a root, then p is not irreducible.
(c) If $\deg(p) \in \{2, 3\}$, then p is irreducible if and only if p does not have a root.

15.6. Let $A \in GL_n(\mathbb{C})$, $n \geq 2$, and let $\mathrm{adj}(A) \in \mathbb{C}^{n,n}$ be the adjunct of A. Show that there exist $n - 1$ matrices $A_j \in \mathbb{C}^{n,n}$ with $\det(-A_j) = \det(A)$, $j = 1, \ldots, n-1$, and

$$\mathrm{adj}(A) = \prod_{j=1}^{n-1} A_j.$$

(*Hint:* Use P_A to construct a polynomial $p \in \mathbb{C}[t]_{\leq n-1}$ with $\mathrm{adj}(A) = p(A)$ and express p as product of linear factors.)

15.7. Show that two polynomials $p, q \in \mathbb{C}[t] \setminus \{0\}$ have a common root if and only if there exist polynomials $r_1, r_2 \in \mathbb{C}[t]$ with $0 \leq \deg(r_1) < \deg(p)$ such that $0 \leq \deg(r_2) < \deg(q)$ and $p \cdot r_2 + q \cdot r_1 = 0$.

15.8. Let $\mathcal{V}$ be a finite dimensional unitary vector space, $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$, $\mathcal{H} = \{g \in \mathcal{L}(\mathcal{V}, \mathcal{V}) \mid g = g^{ad}\}$ and let

$$h_1 : \mathcal{H} \to \mathcal{L}(\mathcal{V}, \mathcal{V}), \quad g \mapsto \frac{1}{2}(f \circ g + g \circ f^{ad}),$$

$$h_2 : \mathcal{H} \to \mathcal{L}(\mathcal{V}, \mathcal{V}), \quad g \mapsto \frac{1}{2\mathbf{i}}(f \circ g - g \circ f^{ad}).$$

Show that $h_1, h_2 \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ and $h_1 \circ h_2 = h_2 \circ h_1$.

15.9. Let $A \in \mathbb{C}^{n,n}$, $\mathcal{S} = \{B \in \mathbb{C}^{n,n} \mid B = B^T\}$ and let

$$h_1 : \mathcal{S} \to \mathbb{C}^{n,n}, \quad B \mapsto AB + BA^T,$$
$$h_2 : \mathcal{S} \to \mathbb{C}^{n,n}, \quad B \mapsto ABA^T.$$

Show that $h_1, h_2 \in \mathcal{L}(\mathcal{S}, \mathcal{S})$ and $h_1 \circ h_2 = h_2 \circ h_1$.

15.10. Let $\mathcal{V}$ be a $\mathbb{C}$-vector space, $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ and let $\mathcal{U} \neq \{0\}$ be a finite dimensional f-invariant subspace of $\mathcal{V}$. Show that $\mathcal{U}$ contains at least one eigenvector of f.

15.11. Let $\mathcal{V} \neq \{0\}$ be a K-vector space and let $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$. Show the following statements:

(a) If $K = \mathbb{C}$, then there exists an f-invariant subspace $\mathcal{U}$ of $\mathcal{V}$ with $\dim(\mathcal{U}) = 1$.
(b) If $K = \mathbb{R}$, then there exists an f-invariant subspace $\mathcal{U}$ of $\mathcal{V}$ with $\dim(\mathcal{U}) \in \{1, 2\}$.

15.12. Prove Theorem 15.17.

15.13. Construct an example showing that the condition $f \circ g = g \circ f$ in Theorem 15.17 is sufficient but not necessary for the simultaneous unitary triangulation of f and g.

15.14. Let $A \in K^{n,n}$ be a diagonal matrix with pairwise distinct diagonal entries and $B \in K^{n,n}$ with $AB = BA$. Show that in this case B is a diagonal matrix. What can you say about B, when the diagonal entries of A are not all pairwise distinct?

15.15. Show that the matrices

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

commute and determine a unitary matrix Q such that $Q^H AQ$ and $Q^H BQ$ are upper triangular.

15.16. Show the following statements for $p \in K[t]$:

(a) For all $A \in K^{n,n}$ and $S \in GL_n(K)$ we have $p(SAS^{-1}) = Sp(A)S^{-1}$.
(b) For all $A, B, C \in K^{n,n}$ with $AB = CA$ we have $Ap(B) = p(C)A$.
(c) If $K = \mathbb{C}$ and $A \in \mathbb{C}^{n,n}$, then there exists a unitary matrix Q, such that $Q^H AQ$ and $Q^H p(A)Q$ are upper triangular.

15.17. Let $\mathcal{V}$ be a finite dimensional unitary vector space. Let $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ be *normal*, i.e., f satisfies $f \circ f^{ad} = f^{ad} \circ f$.

(a) Show that if $\lambda \in \mathbb{C}$ is an eigenvalue of f, then $\mathcal{V}_f(\lambda)^\perp$ is an f-invariant subspace.
(b) Show (using (a)) that f is diagonalizable. (*Hint:* Show by induction on $\dim(\mathcal{V})$, that $\mathcal{V}$ is the direct sum of the eigenspaces of f.)
(c) Show (using(a) or (b)), that f is even *unitarily diagonalizable*, i.e., there exists an orthonormal basis B of $\mathcal{V}$ such that $[f]_{B,B}$ is a diagonal matrix.
(d) Let $g \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ be unitarily diagonalizable. Show that g is normal. (This shows that an endomorphism on a finite dimensional unitary vector space is normal if and only if it is unitarily diagonalizable. We will give a different proof of this result in Theorem 18.2.)

15.18. Let $\mathcal{V}$ be a finite dimensional K-vector space, $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ and $\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2$, where $\mathcal{U}_1, \mathcal{U}_2$ are f-invariant subspaces of $\mathcal{V}$. Let, furthermore, $f_j := f|_{\mathcal{U}_j} \in \mathcal{L}(\mathcal{U}_j, \mathcal{U}_j)$, $j = 1, 2$.

(a) For every $v \in \mathcal{V}$ there exist unique $u_1 \in \mathcal{U}_1$ and $u_2 \in \mathcal{U}_2$ with $v = u_1 + u_2$. Show that then also $f(v) = f(u_1) + f(u_2) = f_1(u_1) + f_2(u_2)$. (We write this as $f = f_1 \oplus f_2$ and call f the *direct sum* of f_1 and f_2 with respect to the decomposition $\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2$.)
(b) Show that $\operatorname{rank}(f) = \operatorname{rank}(f_1) + \operatorname{rank}(f_2)$ and $P_f = P_{f_1} \cdot P_{f_2}$.
(c) Show that $a(\lambda, f) = a(\lambda, f_1) + a(\lambda, f_2)$ for all $\lambda \in K$. (Here we set $a(\lambda, h) = 0$, if λ is not an eigenvalue of $h \in \mathcal{L}(\mathcal{V}, \mathcal{V})$.)
(d) Show that $g(\lambda, f) = g(\lambda, f_1) + g(\lambda, f_2)$ for all $\lambda \in K$. (Here we set $g(\lambda, h) = \dim(\ker(\lambda \mathrm{Id}_{\mathcal{V}} - h))$ even if λ is not an eigenvalue of $h \in \mathcal{L}(\mathcal{V}, \mathcal{V})$.)
(e) Show that $p(f) = p(f_1) \oplus p(f_2)$ for all $p \in K[t]$.