10.B Polynomials in several variables

Let K be a field. A polynomial function on K^I, where I is an arbitrary set, is a function $K^I \longrightarrow K$ with $t \longmapsto \sum_{\nu \in \mathbb{N}^{(I)}} a_\nu t^\nu$, where the (constant) coefficients $a_\nu \in K$, $\nu \in \mathbb{N}^{(I)}$, and are 0 for almost all $\nu = (\nu_i)_{i \in I} \in \mathbb{N}^{(I)}$ and t^ν denote the power-product $t^\nu := \prod_{i \in I} t_i^{\nu_i}$ for an I-tuple $t = (t_i) \in K^{(I)}$ (in this product almost all factors are 1).

Like in the case of one variable (see 10.A) also in the case of several variables, it is often comfortable to consider <u>polynomials</u> $P = \sum_{\nu \in \mathbb{N}^{(I)}} a_\nu X^\nu$ instead of polynomial functions which can be identified with its coefficient-tuples $(a_\nu)_{\nu \in \mathbb{N}^{(I)}}$. Note that in every polynomial almost all coefficients $a_\nu \in K$ are 0 and X^ν denote the <u>monomial</u> $X^\nu = \prod_{i \in I} X_i^{\nu_i}$, in the <u>indeterminates</u> or <u>variables</u> X_i, $i \in I$. Sometimes we also choose another notations like Y_i, Z_i or similar for the indeterminates. In the case of finitely many variables, i.e. in the case of a <u>finite</u> indexed set I, often we also denote the variables by isolated letters, e.g. by X, Y, Z in the case $\# I = 3$. The set of all polynomials over K in the variables X_i, $i \in I$, is denoted by $K[X_i]_{i \in I}$ or $K[X_i \mid i \in I]$.

If $I=\{1,2,\cdots,n\}$, then we also use the notation $K[X_1,\cdots,X_n]$. In any case every polynomial P only finitely many indeterminates X_i appear. Therefore often one can reduce the investigation of polynomials in finitely many variables.

Polynomials $P=\sum_\nu a_\nu X^\nu$ and $Q=\sum_\nu b_\nu X^\nu$ from $K[X_i]_{i\in I}$ can be formally added and multiplied:

$$P+Q:=\sum_\nu (a_\nu+b_\nu)X^\nu \;;\; PQ=\sum_\nu c_\nu X^\nu,\; c_\nu:=\sum_{\mu+\lambda=\nu} a_\mu b_\lambda .$$

Also the multiplication with constant $a\in K$ is defined coefficientwise. With these binary operations $K[X_i]_{i\in I}$ is a commutative K-algebra; the monomials X^ν, $\nu\in\mathbb{N}^{(I)}$, form its K-basis. We define the degree of a monomial X^ν as $|\nu|:=\sum_{i\in I}\nu_i$. The degree of a polynomial $P=\sum_\nu a_\nu X^\nu\neq 0$ is by definition the maximum of the degrees of the monomials X^ν whose coefficients a_ν are non-zero. We denote it by $\deg P$. The degree of the zero-polynomial is by definition $-\infty$. For $m\in\mathbb{N}$, the sum $P_m:=\sum_{\nu\in\mathbb{N}^{(I)},\,|\nu|=m} a_\nu X^\nu$ is called the m-th homogeneous component of P. Then $P=\sum_{m\in\mathbb{N}} P_m$ is the canonical decomposition of P into its homogeneous components. If $P=P_m$, then P is called homogeneous of degree m; accordingly the zero-polynomial is homogeneous of every degree.

The set of homogeneous polynomials of degree m form a K-subspace A_m of $A := K[X_i]_{i\in I}$ and A is the direct sum of A_m, $m \in \mathbb{N}$:

$$A = \sum_{m\in\mathbb{N}}^{\oplus} A_m .$$

Clearly $A_m \cdot A_n \subseteq A_{m+n}$, i.e. the product of a homogeneous polynomial of degree m and a homogeneous polynomial of degree $m+n$. Let $j \in I$ be fixed. Then every polynomial $P \in K[X_i]_{i\in I}$ can be written in the form $P = \sum_{k\in\mathbb{N}} P_k X_j^k$ with uniquely determined polynomials P_k in the indeterminates X_i, $i \in I \setminus \{j\}$; the indeterminate X_j does not appear in them explicitly. If $P \neq 0$ and $P_n \neq 0$, but $P_k = 0$ for all $k > n$, then n is called the <u>partial degree of P with respect to the indeterminate X_j</u> and is denoted by $\deg_{X_j} P$.

<u>The product PQ of two non-zero polynomials $P, Q \in K[X_i]_{i\in I}$ is again non-zero.</u> Since (in) P and Q only finitely indeterminate appear, for the proof we may assume that I is finite, say $I = \{1, 2, \cdots, n\}$. We prove the assertion by induction on n. The cases $n=0$ and $n=1$ are trivial, see 10.A.1. For the inductive step from n to $n+1$, we write P and Q in the form: $P = \sum_{i=0}^{r} P_i X_{n+1}^i$ and $Q = \sum_{j=0}^{s} Q_j X_{n+1}^j$ with

polynomials $P_i, Q_j \in K[X_1, \cdots, X_n]$ and $P_r \neq 0, Q_s \neq 0$.

Then $PQ = \sum_{k=0}^{r+s} \left(\sum_{i+j=k} P_i Q_j \right) X_{n+1}^k = P_r Q_s X_{n+1}^{r+s} + \cdots$

By induction hypothesis $P_r \cdot Q_s \neq 0$ and hence $PQ \neq 0$. From this the <u>degree-formula</u> analogous to 10.A.1 also follows: If $P \neq 0$ and $Q \neq 0$ have the decompositions $P = P_0 + \cdots + P_m$ and $Q = Q_0 + \cdots + Q_t$ into homogeneous components, $P_m, Q_t \neq 0$, then $PQ = P_0 Q_0 + \cdots + P_m Q_t$ is the corresponding decomposition for the product PQ with $P_m Q_t \neq 0$. Therefore $\deg PQ = m + t = \deg P + \deg Q$.

The definition of a prime polynomial or irreducible polynomial is analogous to 10.A.7. Once again we have ~~the~~ the theorem on the unique prime factorisation, see 10.A.9. But for polynomials in several variable a division with remainder hold only in a restricted form and its proof is more complicated

We now discuss the substitution homomorphisms. Let A be a K-algebra and let x_i, $i \in I$, be a family of pairwise commuting elements in A. For $\nu \in \mathbb{N}^{(I)}$, the monomial $x^\nu := \prod_{i\in I} x_i^{\nu_i} \in A$ in x_i, $i \in A$,

is well-defined and in the same way, for every polynomial $F = \sum_{\nu \in \mathbb{N}^{(I)}} a_\nu X^\nu$, the <u>value</u> $F(x) =$

$F(x_i \mid i \in I) = \sum_{\nu \in \mathbb{N}^{(I)}} a_\nu x^\nu$ <u>of F at the place</u> $x = (x_i)_{i \in I}$.

The assignment $F \longmapsto F(x)$ is a K-algebra homomorphism from $K[X_i]_{i\in I}$ into A; it is called the <u>substitution homomorphism</u> $X_i \longmapsto x_i$, $i \in I$ and its image is the smallest K-subalgebra $K[x_i]_{i\in I} = K[x_i \mid i \in I]$ of A which contains all x_i, $i \in I$. Its elements are also called <u>polynomials in the x_i, $i \in I$</u>. If the substitution homomorphism $X_i \longmapsto x_i$, $i \in I$, is injective, then we say that the family x_i, $i \in I$, is <u>algebraically independent</u>; otherwise the x_i, $i \in I$, are <u>algebraically dependent</u> (<u>over K</u>).

For example, if $(a_i) \in K^{(I)}$ is arbitrary, then the substitution homomorphism $X_i \longmapsto X_i - a_i$, $i \in I$, defined from $K[X_i]_{i \in I}$ into itself is bijective. It inverse is again the substitution homomorphism $X_i \longmapsto X_i + a_i$, $i \in I$. In particular, the polynomials

$$(X-a)^\nu = \prod_{i \in I} (X_i - a_i)^{\nu_i}, \quad \nu = (\nu_i) \in \mathbb{N}^{(I)},$$

form a K-basis of $K[X_i]_{i\in I}$, since it is the image of the K-basis X^ν, $\nu \in \mathbb{N}^{(I)}$ of $K[X_i]_{i\in I}$. The representation

$$P = \sum_{\nu \in \mathbb{N}^{(I)}} b_\nu (X-a)^\nu$$

of a polynomial $P \in K[X_i]_{i\in I}$ with (uniquely determined) coefficients $b_\nu \in K$, $\nu \in \mathbb{N}^{(I)}$, is called the Taylor-expansion of P at $a=(a_i)_{i\in I}$. The substitution homomorphism $X_i \longmapsto p_i$, $i \in I$, where $p_i : K^I \longrightarrow K$ is the i-th projection $(t_j)_{j\in I} \longmapsto t_i$, associates a polynomial $F = \sum_\nu a_\nu X^\nu$ to the polynomial function $K^I \longrightarrow K$,

$$t=(t_i)_{i\in I} \longmapsto F(t) = F(t_i \mid i\in I) = \sum_{\nu\in \mathbb{N}^{(I)}} a_\nu t^\nu .$$

If K is infinite then this substitution homomorphism is injective, therefore the polynomial F is uniquely determined by the corresponding polynomial function. This follows from the following more general theorem:

10.B.1 Identity Theorem for polynomials Let $F \in K[X_1,\dots,X_n]$ be a polynomial with the partial degree with respect to the indeterminate X_i is atmost d_i, $i=1,\dots,n$. Suppose that $N_i \subseteq K$ are subsets with $\#N_i > d_i$, $i=1,\dots,n$ and F vanish on all $t=(t_1,\dots,t_n) \in N_1\times\cdots\times N_n$. Then F is the zero polynomial, i.e. $F=0$.

Proof By induction on n. Suppose $n\geq 1$ and $F = \sum_{k=0}^{d_n} F_k(X_1,\dots,X_{n-1}) X_n^k$. Let $(a_1,\dots,a_{n-1}) \in N_1\times\cdots\times N_{n-1}$.

For all $a_n \in N_n$, the polynomial $F(a_1,\dots,a_{n-1},X_n) = \sum_{k=0}^{d_n} F_k(a_1,\dots,a_{n-1})\, X_n^k \in K[X_n]$ vanish at a_n. Since $\# N_n > d_n$, the polynomial $F(a_1,\dots,a_{n-1},X_n)$ is the zero polynomial and hence $F_k(a_1,\dots,a_{n-1}) = 0$ for all $(a_1,\dots,a_{n-1}) \in N_1 \times \dots \times N_{n-1}$ and for all $k = 0,\dots,d_n$.
Therefore by induction hypothesis $F_k = 0$ for all $k=0,\dots,d_n$, and hence $F=0$.

The K-algebra of fractions F/G with $F, G \in K[X_i]_{i\in I}$, $G \neq 0$, constructed as in the case of one variable, is a field. It is called the <u>field of rational functions in the indeterminates X_i, $i \in I$</u>. We denote this field by $K(X_i)_{i\in I} = K(X_i \mid i \in I)$.

10.B.2 Remark The polynomial rings $A[X_i]_{i\in I}$ for arbitrary rings A can be constructed analogously as in the case when A is a field. In particular, for commutative base rings A play a fundamental role in the entire algebra. We shall use this sometimes, see for example the next example.
The polynomial rings $A[X_1,\dots,X_n]$, $n \in \mathbb{N}$, can be

---

The proof of 10.B.1 can also be formulated as follows: The functions $t_i^{\nu_i}$, $\nu_i = 0,\dots,d_i$, are linearly independent K-valued functions on N_i, $i = 1,\dots,n$, by 10.A.12. Then the functions $t_1^{\nu_1} \otimes \dots \otimes t_n^{\nu_n}$, $0 \le \nu_i \le d_i$, $i=1,\dots,n$, are also linearly independent on $N_1 \times \dots \times N_n$ and this is the assertion.

recursively constructed by adjoining only one indeterminate at a time:

A, $A[X_1]$, $A[X_1,X_2]=A[X_1][X_2]$, $A[X_1,X_2,X_3]=A[X_1,X_2][X_3]$, ...

For induction-proofs this view-point may be very useful.

10.B.3 Example (Method of indeterminates -- Determinants over arbitrary commutative rings)

The polynomial rings $A[X_i \mid i\in I]$ are used for the proofs of general known identities. For this basic ingredients are the substitution homomorphisms:

Let $\varphi: A \longrightarrow B$ be a homomorphism of commutative rings and x_i, $i\in I$, be elements in B. Then there exists a unique ring homomorphism

$A[X_i \mid i\in I] \longrightarrow B$, $X_i \longmapsto b_i$, $i\in I$, i.e.

$$\sum_{\nu\in\mathbb{N}^{(I)}} a_\nu X^\nu \longmapsto \sum_{\nu\in\mathbb{N}^{(I)}} \varphi(a_\nu)\, x^\nu$$ which extends φ

to the polynomial rings $A[X_i \mid i\in I]$ uniquely.

Whenever one would like to prove an (algebraic) equation for the elements $x_i\in B$, $i\in I$, most often, it is enough to prove the corresponding identity for the indeterminates X_i, $i\in I$, (over A) and then transport this identity to B by using the substitution homomorphism. This process is called the Kronecker's method of indeterminates.

For example, the binomial equation

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

11.B.7 Definition An operator $f: V \longrightarrow V$ on a finite dimensional K-vector space V is called trigonalisable if there exists a basis $\underline{v} = \{v_i \mid i \in I\}$ of V such that the matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ of f is an upper-triangular matrix, i.e. there exists a flag

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V, \quad n = \mathrm{Dim}_K V$$

of f-invariant subspaces V_i, $i = 1, \ldots, n$.

11.B.8 Theorem Let $f: V \longrightarrow V$ be an operator on the finite dimensional K-vector space V. Then the following statements are equivalent:

(1) f is trigonalisable.

(2) The characteristic polynomial χ_f of f splits into linear factors in $K[X]$.

(3) The minimal polynomial μ_f of f splits into linear factors in $K[X]$.

Proof The equivalence of (2) and (3) follows from the fact that χ_f and μ_f have the same prime factors, see 11.A.14. The implication (1) $\Rightarrow$ (2) is trivial.

We shall prove the implication (2) $\Rightarrow$ (1) by induction on $n := \mathrm{Dim}_K V$. The cases $n \leq 1$ are trivial. Now, assume $n \geq 1$. By 11.A.26 there exists an f-invariant hyperplane V_{n-1} in V. Since the characteristic polynomial $\chi_{f|V_{n-1}}$ divides χ_f by 11.A.8 (1), $\chi_{f|V_{n-1}}$ also splits into linear factors in $K[X]$ and hence by induction hypothesis, there exists an f-inv.

ariant flag $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1}$ of V_{n-1}. Then $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1} \subsetneq V_n = V$ is the required f-invariant flag of V.

One can also prove: Let $x \in V$ be an eigenvector of f (exists by (1)) and $V_1 := Kx$. Further, let $\bar{f} : V/V_1 \longrightarrow V/V_1$ be the induced operator by f on $\bar{V} := V/V_1$. Since $\chi_{\bar{f}}$ divides χ_f by 11.A.8(1), it follows that $\chi_{\bar{f}}$ also splits into linear factors in $K[X]$, therefore by induction hypothesis there exists an $\bar{f}$-invariant flag $0 = V_1/V_1 \subsetneq V_2/V_1 \subsetneq \cdots \subsetneq V_n/V_1 = V/V_1 = \bar{V}$ of $\bar{V}$. Then $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ is an f-invariant flag of V.

11.B.9 Corollary Let K be an algebraically closed field. Then every operator on an finite dimensional K-vector space is triagonalisable. In particular, every operator on a finite dimensional $\mathbb{C}$-vector space is triagonalisable.

A matrix $\mathfrak{A} \in M_n(K)$, K an arbitrary field, is called *triagonalisable* if the corresponding operator $f_{\mathfrak{A}} : K^n \longrightarrow K^n$ is triagonalisable, i.e. if it is similar to an upper triangular matrix. Theorem 11.B.8 also hold for matrices.

11.B.10 Example We modify the second proof in 11.B.8 with the following formulation: Suppose that $\chi_f = (X-\lambda_1)\cdots(X-\lambda_n)$ where

$\lambda_1, \dots, \lambda_n$ are the zeros of χ_f counted with their multiplicities. Let $U_j = \mathrm{Ker}\big((f-\lambda_1 \mathrm{id}) \cdots (f-\lambda_j \mathrm{id})\big)$, $j = 0, \dots, n$. Then $0 = U_0 \subseteq U_1 \subseteq \cdots \subseteq U_n = V$ and $(f - \lambda_j \mathrm{id})\, U_j \subseteq U_{j-1}$, $j = 1, \dots, n$. But in general, U_j do not form a flag. Now, we choose a basis $x_1, \dots, x_n$ of V such that the part $x_1, \dots, x_{n_j}$, $n_j = \mathrm{Dim}_K U_j$ form a basis of U_j and the matrix of f with respect to this basis is an upper triangular matrix. Numbering the zeros $\lambda_1, \dots, \lambda_n$ such that the equal zeros are consequtively numbered so that $\lambda_1, \dots, \lambda_n$ is also the sequence on the main diagonal of this upper triangular matrix.

The first implication $(2) \Rightarrow (1)$ in 11.B.8 lead to the following process for constructing an f-invariant flag of V. Consider the sequence of images $W := \mathrm{Im}\,(f-\lambda_1 \mathrm{id}) \cdots (f - \lambda_j \mathrm{id})$, $j = 0, \dots, n$, where $\chi_f = (X-\lambda_1) \cdots (X-\lambda_n)$. Then $0 = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_n = V$ and every flag of V containing these subspaces W_j is invariant under f.

11.B.11 Example Nilpotent operators and matrices are triagonalisable and hence 11.B.8 holds for them.

11.B.12 Example Let $f : V \to V$ be a triagonalisable operator on the finite dimensional K-vector space V.

11.B.7 Definition An operator $f: V \longrightarrow V$ on a finite dimensional K-vector space V is called trigonalisable if there exists a basis $\underline{v} = \{v_i \mid i \in I\}$ of V such that the matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ of f is an upper triangular matrix, i.e. there exists a flag

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V, \quad n = \mathrm{Dim}_K V$$

of f-invariant subspaces V_i, $i = 1, \ldots, n$.

11.B.8 Theorem Let $f: V \longrightarrow V$ be an operator on the finite dimensional K-vector space V. Then the following statements are equivalent:

(1) f is trigonalisable.

(2) The characteristic polynomial χ_f of f splits into linear factors in $K[X]$.

(3) The minimal polynomial μ_f of f splits into linear factors in $K[X]$.

Proof The equivalence of (2) and (3) follows from the fact that χ_f and μ_f have the same prime factors, see 11.A.14. The implication (1) $\Rightarrow$ (2) is trivial.

We shall prove the implication (2) $\Rightarrow$ (1) by induction on $n := \mathrm{Dim}_K V$. The cases $n \leq 1$ are trivial. Now, assume $n \geq 1$. By 11.A.26 there exists an f-invariant hyperplane V_{n-1} in V. Since the characteristic polynomial $\chi_{f|V_{n-1}}$ divides χ_f by 11.A.8 (1), $\chi_{f|V_{n-1}}$ also splits into linear factors in $K[X]$ and hence by induction hypothesis, there exists an f-inv.

ariant flag $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1}$ of V_{n-1}. Then $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1} \subsetneq V_n = V$ is the required f-invariant flag of V.

One can also prove: Let $x \in V$ be an eigenvector of f (exists by (1)) and $V_1 := Kx$. Further, let $\overline{f} : V/V_1 \longrightarrow V/V_1$ be the induced operator by f on $\overline{V} := V/V_1$. Since $\chi_{\overline{f}}$ divides χ_f by 11.A.8(1), it follows that $\chi_{\overline{f}}$ also splits into linear factors in $K[X]$, therefore by induction hypothesis there exists an $\overline{f}$-invariant flag $0 = V_1/V_1 \subsetneq V_2/V_1 \subsetneq \cdots \subsetneq V_n/V_1 = V/V_1 = \overline{V}$ of $\overline{V}$. Then $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ is an f-invariant flag of V.

11.B.9 Corollary *Let K be an algebraically closed field. Then every operator on an finite dimensional K-vector space is triagonalisable. In particular, every operator on a finite dimensional $\mathbb{C}$-vector space is triagonalisable.*

A matrix $\mathfrak{A} \in M_n(K)$, K an arbitrary field, is called *triagonalisable* if the corresponding operator $f_{\mathfrak{A}} : K^n \longrightarrow K^n$ is triagonalisable, i.e. if it is similar to an upper triangular matrix. Theorem 11.B.8 also hold for matrices.

11.B.10 Example We modify the second proof in 11.B.8 with the following formulation: Suppose that $\chi_f = (X-\lambda_1)\cdots(X-\lambda_n)$ where

$\lambda_1, \dots, \lambda_n$ are the zeros of χ_f counted with their multiplicities. Let $U_j = \mathrm{Ker}((f-\lambda_1 id)\cdots(f-\lambda_j id))$, $j = 0, \dots, n$. Then $0 = U_0 \subseteq U_1 \subseteq \cdots \subseteq U_n = V$ and $(f-\lambda_j id)U_j \subseteq U_{j-1}$, $j = 1, \dots, n$. But in general, U_j do not form a flag. Now, we choose a basis $x_1, \dots, x_n$ of V such that the part $x_1, \dots, x_{n_j}$, $n_j = \mathrm{Dim}_K U_j$ form a basis of U_j and the matrix of f with respect to this basis is an upper triangular matrix. Numbering the zeros $\lambda_1, \dots, \lambda_n$ such that the equal zeros are consequtively numbered so that $\lambda_1, \dots, \lambda_n$ is also the sequence on the main diagonal of this upper triangular matrix.

The first implication $(2) \Rightarrow (1)$ in 11.B.8 lead to the following process for constructing an f-invariant flag of V. Consider the sequence of images $W := \mathrm{Im}\,(f-\lambda_1 id)\cdots(f-\lambda_j id)$, $j = 0, \dots, n$, where $\chi_f = (X-\lambda_1)\cdots(X-\lambda_n)$. Then $0 = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_n = V$ and every flag of V containing these subspaces W_j is invariant under f.

11.B.11 Example Nilpotent operators and matrices are triagonalisable and hence 11.B.8 holds for them.

11.B.12 Example Let $f : V \to V$ be a triagonalisable operator on the finite dimensional K-vector space V.

Then there is a basis $\underline{v} = \{v_1, \cdots, v_n\}$, $n = \mathrm{Dim}_K V$, of V such that the matrix $\mathfrak{M}^{\underline{v}}_{\underline{v}}(f)$ of f with respect to $\underline{v}$ is an upper triangular matrix

$$\mathfrak{A} := \mathfrak{M}^{\underline{v}}_{\underline{v}}(f) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Therefore the characteristic polynomial of f is $\chi_f = (X - \lambda_1) \cdots (X - \lambda_n)$. Let $F \in K[X]$ be an arbitrary polynomial, then the matrix of $F(f)$ with respect to the same basis $\underline{v}$ is:

$$\mathfrak{M}^{\underline{v}}_{\underline{v}}(F(f)) = F(\mathfrak{A}) = \begin{pmatrix} F(\lambda_1) & * & \cdots & * \\ 0 & F(\lambda_2) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F(\lambda_n) \end{pmatrix}$$

In particular, the operator $F(f)$ is also triagonalisable and its characteristic polynomial $\chi_{F(f)}$ is $\chi_{F(f)} = (X - F(\lambda_1)) \cdots (X - F(\lambda_n))$. In particular,

$$\mathrm{Det}\, F(f) = \prod_{i=1}^{n} F(\lambda_i), \quad \mathrm{Tr}\, F(f) = \sum_{i=1}^{n} F(\lambda_i).$$

Note that $\chi_f = X^n - s_1 X^{n-1} + s_2 X^{n-2} - \cdots + (-1)^n s_n$, where $s_i := S_i(\lambda_1, \cdots, \lambda_n)$ are the i-th elementary symmetric function in elements $\lambda_1, \cdots, \lambda_n$. The sums of powers $p_m = \sum_{i=1}^{n} \lambda_i^m$ are the $\mathrm{Tr}\, f^m$, $m \in \mathbb{N}^*$.

The Newton's formulas:

$$\mathrm{Tr}\, f^{m+1} + \sum_{k=1}^{m} (-1)^k s_k \,\mathrm{Tr}\, f^{m+1-k} + (m+1)(-1)^{m+1} s_{m+1} = 0,$$

$m \geq 0$ and $s_m = 0$ for $m > n$.

These equations between the coefficients of χ_f and the traces $\mathrm{Tr} f^m$, $m \in \mathbb{N}^*$, also hold for not neccessarily triagonalisable operators f. By 10.A.30 there exists a (finite) field extension $L|K$ such that χ_f splits into linear factors in $L[X]$. Now, if $\mathfrak{F} \in M_n(K)$ is the matrix of f with respect to an arbitrary basis $\underline{v}$ of V, then $\mathfrak{F}$ is triagonalisable over L, since $\chi_{\mathfrak{F}} = \chi_f$.

We now go on to prove some results on simutaneous diagonalisability and triagonalisability

Let f and g be diagonalisable operators on the vector space V with the spectral-decompositions

$$f = \sum_\lambda \lambda\, p_f(\lambda) \quad \text{and} \quad g = \sum_\lambda \lambda\, p_g(\lambda).$$

Then f and g are said to be simutaneously diagonalisable if there is a basis of V consisting of common eigen-vectors for f and g.

We put

$$V(\lambda, \lambda') := V_f(\lambda) \cap V_g(\lambda'), \quad \lambda, \lambda' \in K.$$

Then f and g are simuntaneously diagonalisable if and only $V = \sum_{\lambda, \lambda' \in K} V(\lambda, \lambda')$. This sum is obviously direct. Simutaneous diagonalisable operators f, g are always comuting, since for $x \in V(\lambda, \lambda')$, we have $f \circ g(x) = f(g(x)) = f(\lambda' x) = \lambda' f(x) = \lambda' \lambda x = \lambda \lambda' x = g(\lambda x) = g \circ f(x)$. The converse is also true, for this

We need the following lemma:

11.B.14 Lemma Let f and g be diagonalisable operators on the K-vector space V with the spectral-decomposition $f = \sum_\lambda \lambda\, p_f(\lambda)$ and $g = \sum_{\lambda'} \lambda' p_g(\lambda')$.

Then f and g commute if and only if for all $\lambda, \lambda' \in K$, the projections $p_f(\lambda)$ and $p_g(\lambda')$ commute.

Proof Since the projections of the spectral-decompositions of f and g commute, we have
$$fg = \sum_{\lambda,\lambda'} \lambda\lambda' p_f(\lambda)\, p_g(\lambda') = \sum_{\lambda,\lambda'} \lambda'\lambda\, p_g(\lambda')\, p_f(\lambda) = gf.$$

Conversely, suppose that f and g commute. Then the eigen-spaces $V_f(\lambda)$ of f are invariant under g and the eigen-spaces $V_g(\lambda')$ of g are invariant under f, since if $x \in V_f(\lambda)$, then $f(x) = \lambda x$ and hence $f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x)$, i.e. $g(x) \in V_f(\lambda)$. Now, it follows that the eigen-spaces $V_f(\lambda)$ of f are also invariant under all projections $p_g(\lambda')$, $\lambda' \in K$. For, if $x = \sum_{\lambda'} x_{\lambda'}$, $x_{\lambda'} \in V_g(\lambda')$ and $f(x) = \lambda x$, then $f(x) = \sum_{\lambda'} \lambda x_{\lambda'} = \sum_{\lambda'} f(x_{\lambda'})$ and hence $f(x_{\lambda'}) = \lambda x_{\lambda'}$, i.e. $f\big(p_g(\lambda')(x)\big) = \lambda\, p_g(\lambda')(x)$.

Finally, if $\lambda, \lambda' \in K$ and $x = \sum_\alpha x_\alpha \in V$, $x_\alpha \in V_f(\alpha)$. Then $p_f(\lambda)\, p_g(\lambda')(x) = \sum_\alpha p_f(\lambda)\, p_g(\lambda')(x_\alpha) = p_g(\lambda')(x_\lambda)$
$= p_g(\lambda')\, p_f(\lambda)(x)$ and hence $p_f(\lambda)\, p_g(\lambda') = p_g(\lambda')\, p_f(\lambda)$.

Now we can easily prove:

11.B.15 Theorem Let $f_1, \dots, f_r$ be diagonalisable operators on the K-vector space V. Then $f_1, \dots, f_r$ are simultaneously diagonalisable if and only if $f_1, \dots, f_r$ are pairwise commuting, i.e. $f_i f_j = f_j f_i$ for all $1 \le i, j \le r$. Moreover, in this case we have

$$V = \sum_{\lambda_1, \dots, \lambda_r \in K}^{\oplus} V(\lambda_1, \dots, \lambda_r)$$

with $V(\lambda_1, \dots, \lambda_r) := V_{f_1}(\lambda_1) \cap \dots \cap V_{f_r}(\lambda_r)$

$$= \operatorname{Im} p_{f_1}(\lambda_1) \cdots p_{f_r}(\lambda_r)$$

Proof Suppose that $f_1, \dots, f_r$ are pairwise commuting. For $\lambda_1, \dots, \lambda_r \in K$, the projections $p_{f_1}(\lambda_1), \dots, p_{f_r}(\lambda_r)$ are pairwise commuting by 11.B.14. Therefore $p_{f_1}(\lambda_1) \cdots p_{f_r}(\lambda_r)$, $\lambda_1, \dots, \lambda_r \in K$, is a family of projections and this family satisfy the conditions of 5.F.9. The equation $\sum_{\lambda_1, \dots, \lambda_r \in K} p_{f_1}(\lambda_1) \cdots p_{f_r}(\lambda_r) = id_V$ directly follows from the equations $\sum_{\lambda_i \in K} p_{f_i}(\lambda_i) = id_V$, $i = 1, \dots, r$, by multiplying them. Therefore by 5.F.9 it follows that $\operatorname{Im} p_{f_1}(\lambda_1) \cdots p_{f_r}(\lambda_r) = \bigcap_{i=1}^{r} \operatorname{Im} p_{f_i}(\lambda_i) = \bigcap_{i=1}^{r} V_{f_i}(\lambda_i)$.

Conversely, if $f_1, \dots, f_r$ are simultaneously diagonalisable, then $f_1, \dots, f_r$ are pairwise commutative.

11.B.16 Example One can prove 11.B.15 in some what more general situation: If f_j, $j \in J$ is a family of diagonalisable operators, then it is simultaneously diagonalisable if the family f_j, $j \in J$, is pairwise commuting and they generate finite dimensional subspace. For, if x_i, $i \in I$, is a basis consisting of common eigen-vectors for a family f_j, $j \in J$ of operators, then x_i, $i \in I$ is also a basis consisting of eigen-vectors for every linear combination of these operators, moreover, even for every polynomial in f_j, $j \in J$. In particular, if $f_1, \ldots, f_r$ are pairwise commuting diagonalisable operators on the K-vector space V, then every polynomial in the operators $f_1, \ldots, f_r$ is also diagonalisable.

For matrices the formulation of 11.B.15 is: If $\mathfrak{A}_j$, $j \in J$, is a family of diagonalisable matrices in $M_I(K)$, I finite set and if $\mathfrak{A}_j$, $j \in J$, are pairwise commuting, then $\mathfrak{A}_j$, $j \in J$, are simultaneously diagonalisable, i.e. there exists a $\mathfrak{B} \in GL_I(K)$ such that all the matrices $\mathfrak{B}\mathfrak{A}_j\mathfrak{B}^{-1}$, $j \in J$, are diagonal matrices.

Pairwise commuting triagonalisable operators are also simultaneously triagonalisable:

11.B.17 Theorem Let $f_1, \ldots, f_r$ be triagonalisable operators on the n-dimensional K-vector space V. If $f_1, \ldots, f_r$ are pairwise commuting, then they are simultaneously triagonalisable, i.e. there exists a flag $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ of V such that each V_i, $i = 1, 2, \ldots, n$, is invariant under all $f_1, \ldots, f_r$.

Proof The proof is completely analogous to the proof of the implication (2) ⇒ (1) in 11.B.9. For $V \neq 0$, it is enough to construct a common eigen-vector for all $f_1, \ldots, f_r$. This is done by induction on r. For the inductive step from r to $r+1$, let x be an eigen-vector of $f_1, \ldots, f_r$ corresponding to eigen-values $\lambda_1, \ldots, \lambda_r \in K$. Then $x \in U := V_{f_1}(\lambda_1) \cap \cdots \cap V_{f_r}(\lambda_r)$. Since f_{r+1} commutes with all $f_1, \ldots, f_r$, U is f_{r+1}-invariant. The restriction $f_{r+1}|U$ is also triagonalisable since the characteristic polynomial of $f_{r+1}|U$ is a divisor of the characteristic polynomial of f_{r+1}. In particular, U contains an eigen-vector y of f_{r+1}. Then y is an eigen-vector for all $f_1, \ldots, f_r, f_{r+1}$.

11.B.17 also hold for an arbitrary family f_j, $j \in J$, of pairwise commuting triagonalisable operators on the finite dimensional K-vector space, see the first part of the Example 11.B.16. Further (it follows that): If $f_1, \ldots, f_r$ are pairwise commuting triagonalisable operators on the (finite dimensional) K-vector space V, then every polynomial in $f_1, \ldots, f_r$ is also triagonalisable, see Example 11.B.13. One can also give a matrix-theoretic formulation of 11.B.17 which we leave it to the reader.