11.A Eigenvalues, Characteristic polynomial, Minimal polynomial

We begin our investigation of a linear operator f on a vector space over a field K. Let $f: V \longrightarrow V$ be a such operator. The basic idea of the description of f is the following: We decompose V into a direct sum $V = \sum_{i \in I}^{\oplus} V_i$ into subspaces V_i, $i \in I$, which are mapped into themselves by f and the restrictions $f|_{V_i} : V_i \to V_i$ are "simplest possible". Then if $\mathfrak{A}_i$ is the matrix of f with respect to a basis of V_i, $i \in I$, then the matrix of f with respect to the basis of V obtained by the bases of V_i, $i \in I$, is the diagonal block matrix $\mathrm{Diag}(\mathfrak{A}_i : i \in I)$

To procecute this programme we are interested in subspaces U of V which are mapped into themselves under f, i.e. $f(U) \subseteq U$:

11.A.1 Definition Let f be an operator on the K-vector space V. A subspace U of V is called f-invariant if $f(U) \subseteq U$.

Otherthan the zero subspace the most simplest invariant subspaces are of dimension one. A line Kx, $x \neq 0$, is f-invariant if and only if $f(x) \in Kx$, i.e. if $f(x) = \lambda x$ for some $\lambda \in K$.

11.A.2 Definition Let V be a K-vector space and $f: V \longrightarrow V$ be a linear operator on V. Further, let $\lambda \in K$.

(1) λ is called an eigen-value of f if there exists $x \in V$, $x \neq 0$ such that $f(x) = \lambda x$. Such a vector $x \in V$ is called an eigen-vector of f corresponding to the eigen-value λ. The set of eigen values of f is denoted by $E(f) := \mathrm{Eig}(f) \subseteq K$.

(2) The subspace
$$V_f(\lambda) := \{x \in V \mid f(x) = \lambda x\} = \mathrm{Ker}(\lambda\, \mathrm{id}_V - f)$$
of V is called the eigen-space of f corresponding to λ.

The operator $f_{\mathfrak{a}}: K^{(I)} \longrightarrow K^{(I)}$ is described by using the matrix $\mathfrak{a} \in M_I(K)$ (f maps the standard basis element $e_j \longmapsto \sum_{i \in I} a_{ij} e_i$)

The eigen-values of $f_{\mathfrak{a}}$ (resp. eigen-vectors of $f_{\mathfrak{a}}$) are also called the eigen-values (resp. eigen-vectors) of $\mathfrak{a}$.

Let $f: V \longrightarrow V$ be an arbitrary linear operator. The eigen-space $V_f(\lambda)$ contains, other than the zero vector, exactly eigen-vectors of f corresponding to the eigen-value λ. Further, $\lambda \in K$ is an eigen-value of f if and only if the eigen-space

$V_f(\lambda)$ of f corresponding to λ is $\neq 0$, equivalently, $\lambda\, id_V - f$ is not injective. If V is finite-dimensional, this is exactly the case when $\lambda\, id_V - f$ is not bijective. In general, the last condition is weaker. With this we define:

11.A.3 Definition Let V be a K-vector space and let $f: V \longrightarrow V$ be a linear operator on V. An element $\lambda \in K$ is called a spectral-value of f if the operator $\lambda\, id_V - f$ is not an automorphism of V. The set of spectral-values of f is called the spectrum of f and is denoted by $\mathrm{Spec}(f)$.

As we have noted above: for a finite dimensional vector space the set of eigen-values of f is equal to the spectrum of f. In arbitrary case the set of eigen-values of f is a subset of the spectrum of f, i.e. $\mathrm{Eig}(f) \subseteq \mathrm{Spec}(f)$. and equality holds if V is finite dimensional.

Suppose that V is finite dimensional. Then $\lambda \in K$ is an eigen-value of $f: V \longrightarrow V$ if and only if $\lambda\, id_V - f \notin \mathrm{Aut}_K V$. By the theory of determinants, this is exactly the case when $\mathrm{Det}(\lambda\, id_V - f) = 0$.

Now, let $\underline{v} = \{v_i \mid i \in I\}$ be a basis of V and let $\mathfrak{A} = (a_{ij}) = \mathfrak{M}^{\underline{v}}_{\underline{v}}(f) \in M_I(K)$ be the matrix of f with respect to $\underline{v}$. Then $\lambda \mathcal{E}_I - \mathfrak{A} = (\lambda \delta_{ij} - a_{ij}) \in M_I(K)$ is the matrix of $\lambda \mathrm{id}_V - f$ with respect to $\underline{v}$ and $\mathrm{Det}(\lambda \mathrm{id}_V - f) = \mathrm{Det}(\lambda \mathcal{E}_I - \mathfrak{A})$. Therefore eigen-values of f are the zeros of the polynomial

$$\chi_f = \chi_{\mathfrak{A}} := \mathrm{Det}(X \mathcal{E}_I - \mathfrak{A}) \in K[X].$$

This polynomial is naturally independent of the choice of the indexed set I. If $I = \{1, 2, \cdots, n\}$, then

$$\chi_{\mathfrak{A}} = \mathrm{Det} \begin{pmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X - a_{nn} \end{pmatrix}$$

The explicit formula for calculation of a determinants shows that the polynomial $\chi_{\mathfrak{A}}$ is a monic polynomial of degree $n = \mathrm{Dim}_K V$ in $K[X]$. Moreover, $\chi_{\mathfrak{A}}$ does not change if we change basis of V.

For, if $\underline{v}' = \{v_j' \mid \in \}$ is another basis of V and if $\mathfrak{B} \in GL_I(V)$ is the matrix of the change of basis from $\underline{v}$ to $\underline{v}'$, $\mathfrak{A}' = \mathfrak{M}_{\underline{v}'}^{\underline{v}'}(f)$ is the matrix of f with respect to $\underline{v}'$, then $\mathfrak{A}' = \mathfrak{B}\mathfrak{A}\mathfrak{B}^{-1}$ and $X\mathcal{E}_I - \mathfrak{A}' = \mathfrak{B}(X\mathcal{E}_I - \mathfrak{A})\mathfrak{B}^{-1}$ and so (by the product formula for Determinants)

$$\chi_{\mathfrak{A}'} = \mathrm{Det}(X\mathcal{E}_I - \mathfrak{A}' = \mathrm{Det}(X\mathcal{E}_I - \mathfrak{A}) = \chi_{\mathfrak{A}}.$$

We now define

11.A.4 Definition Let $f : V \longrightarrow V$ be an operator on the finite dimensional K-vector space V with basis $\underline{v} = \{v_i \mid i \in I\}$ and $\mathfrak{v} := \mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ be the matrix of f with respect to $\underline{v}$. Then the polynomial

$$\chi_f = \chi_{\mathfrak{v}} = \mathrm{Det}\,(X\,\mathfrak{E}_I - \mathfrak{v})$$

is called the <u>characteristic polynomial</u> of the matrix $\mathfrak{v} \in M_I(K)$ resp. the operator $f \in \mathrm{End}_K V$. We note that $\chi_f = \chi_{\mathfrak{v}}$ is a monic polynomial in $K[X]$ of degree $|I| = \mathrm{Dim}_K V$ and we already have proved:

11.A.5 Theorem *The spectrum $\mathrm{Spec}\, f$ of an operator f on a finite dimensional K-vector space V is the set of zeros of the characteristic polynomial χ_f of f. -- In particular, f has at most $\mathrm{Dim}_K V$ eigen-values.*

Let $f : V \longrightarrow V$ be a linear operator on the K-vector space V. We consider the substitution homomorphism $K[X] \xrightarrow{\Phi_f} \mathrm{End}_K V$ with $X \mapsto f$, the polynomial $F = a_0 + a_1 X + \cdots + a_m X^m \in K[X]$, $\Phi_f(F) = F(f) = a_0 \mathrm{id}_V + a_1 f + \cdots + a_m f^m \in \mathrm{End}_K V$. The operators $F(f)$, $F \in K[X]$ are called polynomials in f. They form the smallest K-subalgebra of

$\mathrm{End}_K V$ which contain f:

$K[f] = \{F(f) \mid F \in K[X]\} = \mathrm{Im}\, \Phi_f \subseteq \mathrm{End}_K V$.

We say that f is <u>algebraic</u> (over K) if there exists a <u>non-zero</u> polynomial $F \in K[X]$, $F \neq 0$, such that $F(f) = 0$, i.e. $F \in \mathrm{Ker}\, \Phi_f$; equivalently, $\mathrm{Ker}\, \Phi_f \neq 0$. In this case there is a <u>unique</u> monic polynomial $\mu_f \in \mathrm{Ker}\, \Phi_f$ which generates the ideal $\mathrm{Ker}\, \Phi_f$. This unique polynomial μ_f is called the <u>minimal polynomial of f over K</u>.

Therefore $\mu_f(f) = 0$ and if $F(f) = 0$, i.e. $F \in \mathrm{Ker}\, \Phi_f$, then μ_f divides F in $K[X]$. Further,

$\mathrm{Im}\, \Phi_f = K[f] \cong K[X]/K[X]\mu_f$. We note this:

11.A.6 Theorem *Let $f : V \to V$ be a K-linear operator of the K-vector space V with the minimal polynomial μ_f. For a polynomial $F \in K[X]$ $F(f) = 0$ if and only if F is a multiple of μ_f in $K[X]$. If V is finite dimensional, then $\mu_f \neq 0$.*

For the last assertion note that

$$K[X]/(\mu_f) \cong K[f] \subseteq \mathrm{End}_K V$$

and that $\mathrm{End}_K V$ is finite dimensional.

For a matrix $\mathfrak{A} \in M_I(K)$, I <u>finite</u> set. The substitution homomorphism $K[X] \longrightarrow M_I(K)$, $X \longmapsto \mathfrak{A}$. If V is finite dimensional K-vector space with basis $\underline{v} = \{v_i \mid i \in I\}$ and let $f_{\mathfrak{A}} : V \to V$ be a linear operator with $\mathfrak{A} = \mathfrak{M}^{\underline{v}}_{\underline{v}}(f)$ the matrix of $f_{\mathfrak{A}}$ with respect to the basis $\underline{v}$. Then the diagram

$$\begin{tikzcd}
 & K[X] \arrow[dl, "\Phi_{f_{\mathfrak{A}}}"'] \arrow[dr, "\Phi_{\mathfrak{A}}"] & \\
\mathrm{End}_K V \arrow[rr, "\mathfrak{M}^{\underline{v}}_{\underline{v}}(\ldots)", "\approx"'] & & M_I(K)
\end{tikzcd}$$

is commutative, where $\Phi_{f_{\mathfrak{A}}}$ (resp. $\Phi_{\mathfrak{A}}$) denote the substitution homomorphism $X \mapsto f_{\mathfrak{A}}$ (resp. $X \mapsto \mathfrak{A}$). In particular, $F(f_{\mathfrak{A}}) = 0$ if and only $F(\mathfrak{A}) = 0$. Further, f and $\mathfrak{A}$ are algebraic with the equal minimal polynomial $\mu_f = \mu_{\mathfrak{A}}$.

Moreover, we have the following fundamental theorem:

11.A.7 Cayley-Hamilton Theorem Let V be a finite dimensional K-vector space with the basis $\underline{v} = \{v_i \mid i \in I\}$ and $f : V \to V$ be a K-linear operator with $\mathfrak{M}^{\underline{v}}_{\underline{v}}(f) = \mathfrak{A} = (a_{ij}) \in M_I(K)$ $\left(j \in I,\ f(v_j) = \sum_{i \in I} a_{ij} v_i\right)$. Then $\chi_{\mathfrak{A}}(\mathfrak{A}) = 0$ and $\chi_f(f) = 0$, i.e. the minimal polynomial $\mu_{\mathfrak{A}} = \mu_f$ divides the characteristic polynomial $\chi_{\mathfrak{A}} = \chi_f$.

Proof By earlier preparation it is enough to prove the assertion for $\mathfrak{A}$. We may assume that $I = \{1, 2, \cdots, n\}$.

Write

$$\chi_{\mathfrak{a}} = \mathrm{Det}(X\mathfrak{E}_n - \mathfrak{a}) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$$

and consider the adjoint matrix $\mathrm{Adj}(X\mathfrak{E}_n - \mathfrak{a})$ of the matrix $X\mathfrak{E}_n - \mathfrak{a} \in M_n(K(X))$. The coefficients of $\mathrm{Adj}(X\mathfrak{E}_n - \mathfrak{a})$ are determinants of certain $(n-1)\times(n-1)$ submatrices of $X\mathfrak{E}_n - \mathfrak{a}$ and hence

$\mathrm{Adj}(X\mathfrak{E}_n - \mathfrak{a}) = X^{n-1}\mathfrak{b}_{n-1} + \cdots + X\mathfrak{b}_1 + \mathfrak{b}_0$ with matrices $\mathfrak{b}_0, \ldots, \mathfrak{b}_{n-1} \in M_n(K)$. Now, by <u>adjoint-determinant formula</u>, we have:

$$\chi_{\mathfrak{a}}\mathfrak{E}_n = \mathrm{Det}(X\mathfrak{E}_n - \mathfrak{a})\mathfrak{E}_n = (X\mathfrak{E}_n - \mathfrak{a})\,\mathrm{Adj}(X\mathfrak{E}_n - \mathfrak{a}) =$$
$$(X\mathfrak{E}_n - \mathfrak{a})(X^{n-1}\mathfrak{b}_{n-1} + \cdots + X\mathfrak{b}_1 + \mathfrak{b}_0) =$$
$$X^n\mathfrak{b}_{n-1} + X^{n-1}(\mathfrak{b}_{n-2} - \mathfrak{a}\mathfrak{b}_{n-1}) + \cdots + X(\mathfrak{b}_0 - \mathfrak{a}\mathfrak{b}_1) - \mathfrak{a}\mathfrak{b}_0$$

and hence (Comparing the coefficients of powers of X) we get the equations:

$$\begin{array}{rcl}
\mathfrak{b}_{n-1} & = & \mathfrak{E}_n \\
\mathfrak{b}_{n-2} - \mathfrak{a}\mathfrak{b}_{n-1} & = & a_{n-1}\mathfrak{E}_n \\
\cdots\cdots & & \cdots\cdots \\
\mathfrak{b}_0 - \mathfrak{a}\mathfrak{b}_1 & = & a_1\mathfrak{E}_n \\
- \mathfrak{a}\mathfrak{b}_0 & = & a_0\mathfrak{E}_n
\end{array}$$

Multiplying these equations by (from left) $\mathfrak{a}^n, \mathfrak{a}^{n-1}, \ldots, \mathfrak{a}^1 = \mathfrak{a}, \mathfrak{a}^0 = \mathfrak{E}_n$, we get:

$$
\begin{aligned}
\mathfrak{a}^n \mathfrak{b}_{n-1} & & &= E_n \\
\mathfrak{a}^{n-1}\mathfrak{b}_{n-2} &- \mathfrak{a}^n \mathfrak{b}_{n-1} & &= a_{n-1}\mathfrak{a}^{n-1} \\
\cdots & & &\cdots \\
\mathfrak{a}\mathfrak{b}_0 &- \mathfrak{a}^2\mathfrak{b}_1 & &= a_1 \mathfrak{a} \\
&- \mathfrak{a}\mathfrak{b}_0 & &= a_0 E_n
\end{aligned}
$$

Adding all these equations, we get: $0 = \chi_{\mathfrak{a}}(\mathfrak{a})$.

The above proof shows that: The equation $\chi_{\mathfrak{a}}(\mathfrak{a}) = 0$ for all matrices $\mathfrak{a} \in M_n(A)$ with coefficients in an arbitrary commutative ring A.

We would like to connect minimal polynomial μ_f and the characteristic polynomial χ_f more closely. For this study, first we prove the following two useful lemmas:

11.A.8 Lemma Let $f: V \longrightarrow V$ be a K-linear operator on the K-vector space V and let $U \subseteq V$ be an f-invariant subspace, $\overline{f}: \overline{V} \longrightarrow \overline{V}$ denote the K-linear operator on the quotient space $\overline{V} := V/U$ induced by f ($\overline{f}(\overline{x}) := \overline{f(x)}$, $x \in V$).

(1) If V is finite dimensional, then $\chi_f = \chi_{f|U} \cdot \chi_{\overline{f}}$

In particular, $\chi_{f|U}$ and $\chi_{\overline{f}}$ are divisors of χ_f in $K[X]$.

(2) *f is algebraic over K if and only if $f|U$ and $\bar{f}$ are algebraic over K. Moreover, in this case $\mu_{f|U}$ divides μ_f, $\mu_{\bar{f}}$ divides μ_f and μ_f divides $\mu_{f|U} \cdot \mu_{\bar{f}}$.*

Proof Let $u_1, \dots, u_r$ be a K-basis of U. Extend this to a basis $u_1, \dots, u_r, w_1, \dots, w_s$ of V. Then the matrix of f with respect to this basis is

$$\mathfrak{f} = \begin{pmatrix} \mathfrak{a} & \tau \\ 0 & \mathfrak{b} \end{pmatrix}$$

$\mathfrak{a} \in M_r(K)$, $\mathfrak{b} \in M_s(K)$, $\tau \in M_{r,s}(K)$,

where $\mathfrak{a} = \mathfrak{M}_{\underline{u}}^{\underline{u}}(f|U)$ and $\mathfrak{b} = \mathfrak{M}_{\underline{\bar{w}}}^{\underline{\bar{w}}}(\bar{f})$.

It follows that

$$X\mathfrak{E}_n - \mathfrak{f} = \begin{pmatrix} X\mathfrak{E}_r - \mathfrak{a} & -\tau \\ 0 & X\mathfrak{E}_s - \mathfrak{b} \end{pmatrix}, \quad n = r+s$$

and hence

$$\chi_f = \mathrm{Det}(X\mathfrak{E}_n - \mathfrak{f}) = \mathrm{Det}(X\mathfrak{E}_r - \mathfrak{a}) \cdot \mathrm{Det}(X\mathfrak{E}_s - \mathfrak{b}) =$$

$\chi_{f|U} \cdot \chi_{\bar{f}}$. This proves (1). For the proof of (2), first suppose that f is algebraic with minimal polynomial μ_f $(\neq 0)$. From $\mu_f(f) = 0$, it follows that $\mu_f(f|U) = 0$ and also $\mu_f(\bar{f}) = 0$ and hence $f|U$ and $\bar{f}$ are algebraic; Moreover, $\mu_{\bar{f}} \mid \mu_f$ and $\mu_{f|U} \mid \mu_f$.

Now, it is enough to show that

$$(\mu_{f|U}\cdot\mu_{\bar f})(f) = \mu_{f|U}(f)\cdot\mu_{\bar f}(f) = 0 .$$

For this let $x\in V$ be arbitrary. Then $\mu_{\bar f}(\bar f)(\bar x)=0$, i.e. $\mu_{\bar f}(f)(x)\in U$ and hence $\mu_{f|U}(f)\left(\mu_{\bar f}(f)(x)\right)=0$. Therefore $\mu_{f|U}(f)\cdot\mu_{\bar f}(f)=0$ as required.

11.A.9 Lemma Let $f: V\longrightarrow V$ be an algebraic operator on the K-vector space V with the minimal polynomial $\mu_f\,(\neq 0)$. Then for every non-constant divisor F of μ_f, $\operatorname{Ker} F(f)\neq 0$.

Proof Suppose that $\mu_f = FG$, $F, G\in K[X]$. Then $\deg G<\deg\mu_f$, since $\deg F\geqslant 1$. If $F(f)$ is injective, then since $0=\mu_f(f) = F(f)\,G(f)$, we must have $G(f)=0$ which contradicts the fact that μ_f is a $\neq 0$ polynomial of least degree with $\mu_f(f)=0$.

11.A.10 Theorem Let $f: V\longrightarrow V$ be an algebraic operator on the K-vector space V. Then the eigenvalues of f are precisely the spectral values of f and these are precisely the zeros of the minimal polynomial μ_f of f, i.e.

$$\operatorname{Eig}(f) = \operatorname{Spec} f = N(\mu_f) = \{\lambda\in K \mid \mu_f(\lambda)=0\}.$$

Proof We shall prove the inclusions:

$$\operatorname{Eig}(f)\subseteq \operatorname{Spec} f \subseteq N(\mu_f)\subseteq \operatorname{Eig}(f)$$

$\mathrm{Eig}(f) \subseteq \mathrm{Spec}\, f$ remarked above. For the proof of $\mathrm{Spec}\, f \subseteq N(\mu_f)$. Suppose that $\lambda \in K$ and $\lambda \notin N(\mu_f)$. Then the polynomials $X-\lambda$ and μ_f are relatively prime in $K[X]$ and hence there exist polynomials $S, T \in K[X]$ such that $1 = S\cdot(X-\lambda) + T\mu_f$. Now, by substituting $X = f$, we get
$\mathrm{id}_V = S(f)\circ(f-\lambda\,\mathrm{id}_V) + T(f)\circ\mu_f(f) =$

$S(f)\circ(f-\lambda\,\mathrm{id}_V) = (f-\lambda\,\mathrm{id}_V)\circ S(f)$. This means that $S(f) \in \mathrm{End}_K V$ is the inverse of $f-\lambda\,\mathrm{id}_V$ and hence λ is <u>not</u> a spectral value of f.
Finally, suppose $\lambda \in N(\mu_f)$. Then $X-\lambda$ is a divisor of μ_f and hence $\mathrm{Ker}(f-\lambda\,\mathrm{id}_V) \neq 0$ by 11.A.9, i.e. λ is an eigen-value of f.

From 11.A.5 and 11.A.10 it follows that:

11.A.11 Corollary The minimal polynomial μ_f and the characteristic polynomial χ_f of an operator $f : V \to V$ on a finite dimensional K-vector space V have the same zero-set, i.e.

$$N(\mu_f) = N(\chi_f)$$

Moreover, this common zero-set of μ_f and χ_f is the $\mathrm{Spec}\, f$ spectrum of f which is the set of eigen-values $\mathrm{Eig}(f)$ of f.

We now generalize 11.A.11 : The polynomials μ_f and χ_f have the same prime factors (in $K[X]$). Since by the Cayley-Hamilton Theorem 11.A.7 μ_f divides χ_f, clearly every prime factor of μ_f is also a prime factor of χ_f. For the proof of converse we first prove the following theorem which is of interest independently.

11.A.12 Theorem Let $f : V \longrightarrow V$ be an algebraic operator on the K-vector space V. For every monic prime factor P of the minimal polynomial μ_f of f, there exists an f-invariant subspace U of V such that $\mathrm{Dim}_K U = \deg P$ and $\mu_{f|U} = \chi_{f|U} = P$.

Proof By 11.A.9 $\mathrm{Ker}\, P(f) \neq 0$. Suppose that $x \in \mathrm{Ker}\, P(f)$, $x \neq 0$, i.e $P(f)(x) = 0$, $x \neq 0$. Consider the smallest f-invariant subspace U of V containing x. This subspace U is precisely the subspace $\{F(f)(x) \mid F \in K[X]\}$. Note that if we write $F = QP + R$ with $\deg R < \deg P$, then $F(f) = Q(f)P(f) + R(f)$ and hence $F(f)(x) = R(f)(x)$, since $P(f)(x) = 0$. This proves that

$$U = \{R(f)(x) \mid R \in K[X],\ \deg R < \deg P\}$$

$$= Kx + K \quad + \cdots + K f^{m-1}(x), \quad m = \deg$$

In particular, $\mathrm{Dim}_K U \leq \deg P$. Note that for all

$y \in U$, $P(f)(y)=0$, since $y = F(f)(x)$ with $F \in K[X]$ and hence $P(f)(y) = P(f)\big(F(f)(x)\big) = \big(P(f)\cdot F(f)\big)(x) = \big(F(f)\cdot P(f)\big)(x) = F(f)\big(P(f)(x)\big) = F(f)(0) = 0$. Therefore we have proved that $P(f|U)=0$. But then $\mu_{f|U}$ divides P and hence (since P is prime) either $\mu_{f|U} = 1$ or $\mu_{f|U} = P$. Now, since $U \neq 0$ (as $0 \neq x \in U$) $\mu_{f|U} \neq 1$ and so $\mu_{f|U} = P$. By 11.A.7 we have $\mu_{f|U}$ $\chi_{f|U}$. This proves that

$$\deg \mu_{f|U} \leq \deg \chi_{f|U} = \mathrm{Dim}_K\, U \leq \deg P = \deg \mu_{f|U}$$

and hence $\chi_{f|U} = P = \mu_{f|U}$ and $\mathrm{Dim}_K U = \deg P$.

Proof of 11.A.13 shows how can we find an f-invariant subspace $U \subseteq V$: Choose $x \in V$ such that $P(f)(x)=0$, and $x \neq 0$. Then $x, f(x), \ldots, f^{m-1}(x)$, where $m := \deg P$, is a K-basis of U.

Since all prime polynomials in $\mathbb{R}[X]$ are of degree ≤ 2, we have the following special case:

11.A.13 Corollary Let $V \neq 0$ be an $\mathbb{R}$-vector space and let $f: V \longrightarrow V$ be an algebraic operator on V. Then there exists an f-invariant subspace $U \neq 0$ of dimension ≤ 2.

Now, we can prove

11.A.14 Theorem Let V be a finite dimensional K-vector space and let $f: V \longrightarrow V$ be a linear operator on V. Then the characteristic polynomial χ_f and the minimal polynomial μ_f have the same prime factors.

Proof By 11.A.7 μ_f divides χ_f and hence every prime factor of μ_f is also a prime factor of χ_f. We shall prove the converse by induction on $\mathrm{Dim}_K V$. Suppose that $\mathrm{Dim}_K V > 0$ and that P is a prime factor of μ_f. By 11.A.12 there exists an f-invariant subspace $U \subseteq V$ such that $\chi_{f|U} = \mu_{f|U} = P$. Let $\bar{f}: \bar{V} \longrightarrow \bar{V}$ be the linear operator on the quotient space $\bar{V} = V/U$ induced by f. Now by 11.A.8 we have $\chi_f = \chi_{f|U}\,\chi_{\bar{f}} = P\chi_{\bar{f}}$ and $\mu_{\bar{f}} \mid \mu_f$. Further, since $\mathrm{Dim}_K \bar{V} < \mathrm{Dim}_K V$ ($U \neq 0$), by induction hypothesis, every prime divisor of $\chi_{\bar{f}}$ is also a prime divisor of $\mu_{\bar{f}}$ and hence also of μ_f. From this it follows that all prime divisors of χ_f are also prime divisors of μ_f, since P is also a divisor of μ_f.

11.A.15 Remark 11.A.14 also follows directly from 11.A.11: It is enough to prove 11.A.14 for matrices $\mathfrak{A} \in M_n(K)$. Suppose that π is a prime divisor of $\chi_{\mathfrak{A}}$. Let $L|K$ be a field extension such that π has a zero in L, for example, $L := K[X]/(\pi)$ and consider $\mathfrak{A}$ as a matrix in $M_n(L)$. Then by Exercise 10,

$$\chi_{\mathfrak{A},K} = \chi_{\mathfrak{A},L} \text{ and } \mu_{\mathfrak{A},K} = \mu_{\mathfrak{A},L}.$$

If π is not a divisor of $\mu_{\mathfrak{A},K} = \mu_{\mathfrak{A},L}$, then π and $\mu_{\mathfrak{A},K}$ are relatively prime in $K[X]$ and hence also in $L[X]$. But λ is a zero of π and hence of $\chi_{\mathfrak{A},L} = \chi_{\mathfrak{A},K}$; therefore λ is a zero $\mu_{\mathfrak{A},L} = \mu_{\mathfrak{A},K}$. This contradicts $\gcd_{L[X]}(\pi, \mu_{\mathfrak{A}}) = 1$, since $X-\lambda \mid \pi$ and $X-\lambda \mid \mu_{\mathfrak{A}}$ in $L[X]$.

The method used above is known as the <u>method of base field extension</u>. An important special case is the passage from $\mathbb{R}$ to $\mathbb{C}$.

Finally we prove:

<u>**11.A.16 Theorem**</u> *Let $f : V \longrightarrow V$ be a linear operator on the finite dimensional K-vector space V. Then f and the dual operator $f^* : V^* \longrightarrow V^*$ have the same characteristic polynomial and the same minimal polynomial, i.e. $\chi_f = \chi_{f^*}$ and $\mu_f = \mu_{f^*}$.*

<u>Proof</u> If $\mathfrak{A} = \mathfrak{M}^{\underline{v}}_{\underline{v}}(f)$ is the matrix of f with respect to the basis $\underline{v}$ of V, then ${}^t\mathfrak{A} = \mathfrak{M}^{\underline{v}^*}_{\underline{v}^*}(f^*)$
the transpose

is the matrix of f^* with respect to the dual basis $\underline{v}^*$ of $\underline{v}$ and hence it follows that

$$\chi_f = \operatorname{Det}(X\mathfrak{E} - \mathfrak{A}) = \operatorname{Det}({}^t(X\mathfrak{E} - \mathfrak{A})) =$$
$$\operatorname{Det}(X\mathfrak{E} - {}^t\mathfrak{A}) = \chi_{f^*}.$$

For arbitrary polynomial $F \in K[X]$, clearly we have $F({}^t\mathfrak{A}) = {}^t(F(\mathfrak{A}))$ and hence $\mathfrak{A}$ and ${}^t\mathfrak{A}$ have the same minimal polynomial. Therefore $\mu_f = \mu_{f^*}$.

Let $f: V \longrightarrow V$ be a K-linear operator on the finite dimensional K-vector space V. Then $\chi_f(\lambda) = \operatorname{Det}(\lambda\, \mathrm{id}_V - f)$, $\lambda \in K$, and in particular,
$$\chi_f(0) = \operatorname{Det}(-f) = (-1)^{\operatorname{Dim}_K V} \operatorname{Det} f.$$

The constant term of the characteristic polynomial of f is, therefore, up to a sign, the determinant of f. Analogously, for a matrix $\mathfrak{A} \in M_n(K)$, we have $\chi_{\mathfrak{A}}(0) = (-1)^n \operatorname{Det} \mathfrak{A}$.

Further, the other coefficients of χ_f (resp. $\chi_{\mathfrak{A}}$) are invariants of f (resp. $\mathfrak{A}$). The coefficient of X^{n-1} in χ_f is used more often, $n = \operatorname{Dim}_K V$.

11.A.17 Definition Let $f: V \longrightarrow V$ be an operator on the finite dimensional K-vector space V (resp. $\mathfrak{A} \in M_I(K)$ be a matrix, I finite set.

characteristic polynomial of f (resp. $\mathfrak{A}$) is $\chi_f = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in K[X]$. Then $-a_{n-1} \in K$ is called the trace of f (resp. $\mathfrak{A}$) and is denoted by $\operatorname{Tr} f$ (resp. $\operatorname{Tr}\mathfrak{A}$). If $\mathfrak{A}$ is the matrix of f with respect to basis of V, then $\operatorname{Tr} f = \operatorname{Tr}\mathfrak{A}$. For $\mathfrak{A} = (a_{ij}) \in M_n(K)$

$$\operatorname{Tr}\mathfrak{A} = \sum_{i \in I} a_{ii}$$

the sum of the diagonal elements of $\mathfrak{A}$ as the explicit formula for $\chi_{\mathfrak{A}} = \operatorname{Det}(X\mathfrak{E}_I - \mathfrak{A}) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$. More generally, $(-1)^{n-r} a_r$ is the sum of diagonal-minors of the order $n-r$ of $\mathfrak{A}$. See Exercise

11.A.18 Theorem Let V be a finite dimensional K-vector space. The trace-map $\operatorname{Tr} : \operatorname{End}_K V \to K$ is K-linear with $\operatorname{Tr}(\mathrm{id}_V) = \operatorname{Dim}_K V$. Moreover, if $f, g \in \operatorname{End}_K V$, then $\operatorname{Tr}(fg) = \operatorname{Tr}(gf)$.
For the trace-map $\operatorname{Tr} : M_I(K) \longrightarrow K$ on the matrix algebra $M_I(K)$, I finite set, the analogous assertions hold.

Proof It is enough to prove these assertions for the matrix algebra. The K-linearity of Tr follows from the above formula for $\operatorname{Tr}\mathfrak{A}$. Further, $\operatorname{Tr}(\mathfrak{E}_I) = n \cdot 1_K$, $n := |I|$. The equality $\operatorname{Tr}(\mathfrak{A}\mathfrak{B}) = \operatorname{Tr}(\mathfrak{B}\mathfrak{A})$ follows from the equality $\chi_{\mathfrak{A}\mathfrak{B}} = \chi_{\mathfrak{B}\mathfrak{A}}$ of characteristic polynomials, See Exercise It also directly follows from the formula: $\operatorname{Tr}(\mathfrak{A}\mathfrak{B}) = \sum_{(i,j) \in I \times I} a_{ij} b_{ji} = \sum_{(j,i) \in I \times I} b_{ji} a_{ij} = \operatorname{Tr}(\mathfrak{B}\mathfrak{A})$

11.A.19 Examples

(1) An operator $f: V \longrightarrow V$ on a K-vector space V has an eigenvalue 0 if and only if f is not injective and has an eigenvalue 1 if and only if f has non-zero fixed point x, i.e. the elements of a line Kx, $x \neq 0$, are pointwise fixed.
The eigen space $V_f(0) = \operatorname{Ker} f$; the eigen space $V_f(1)$ is the space of fixed points of f. The 0 is a spectral value of $f: V \to V$ if and only if f is not an automorphism of V.
If $f: V \longrightarrow V$ is an automorphism of V, then the eigen-vectors of exactly those lines which are fixed points of the projective collineation

$$\langle f \rangle : \mathbb{P}(V) \longrightarrow \mathbb{P}(V)$$

If λ_i, $i \in I$, are pairwise distinct eigenvalues of f with the eigen-spaces $V_f(\lambda_i)$, $i \in I$, then $\biguplus_{i \in I} \mathbb{P}(V_f(\lambda_i)) \subseteq \mathbb{P}(V)$ is the set of fixed points of $\langle f \rangle$. What are the possibilities for $V = K^2$, i.e. for $\mathbb{P}(V)$ and how can one characterize this, e.g. if $K = \mathbb{R}$ or $K = \mathbb{C}$.

(2) The homothecy $a\,\mathrm{id}_V$ on a finite dimensional K-vector space V has the matrix $a\mathfrak{E}_I$ with respect to any basis $\underline{v} = \{v_i \mid i \in I\}$. Therefore the characteristic polynomial $\chi_{a\,\mathrm{id}_V} = \operatorname{Det}(X\mathfrak{E}_I - a\mathfrak{E}_I) = \operatorname{Det}((X-a)\mathfrak{E}_I) = (X-a)^{\operatorname{Dim}_K V}$. The minimal polynomial $\mu_{a\,\mathrm{id}_V} = X - a$ if $V \neq 0$.

(3) The operator $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ defined with respect to the standard basis by the matrix

$$\mathfrak{A} = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

Then $\chi_f = X^3 - 12X - 16 = (X+2)^2(X-4)$. The eigen-values of f (and $\mathfrak{A}$) are therefore -2 and 4. The eigenspaces are

$$V_f(-2) = \operatorname{Ker}(f + 2\,\mathrm{id}) = \mathbb{R}\begin{pmatrix}1\\1\\0\end{pmatrix} + \mathbb{R}\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

$$V_f(4) = \operatorname{Ker}(f - 4\,\mathrm{id}) = \mathbb{R}\begin{pmatrix}1\\1\\2\end{pmatrix}$$

Further, $V = V_f(-2) \oplus V_f(4)$, since the vectors $\begin{pmatrix}1\\1\\0\end{pmatrix}$, $\begin{pmatrix}-1\\0\\1\end{pmatrix}$ and $\begin{pmatrix}1\\1\\2\end{pmatrix}$ are linearly independent.

The possibilities for minimal polynomial of f are $\mu_f = (X+2)(X-4)$ and $(X+2)^2(X-4)$. Already $g := (f + 2\,\mathrm{id})(f - 4\,\mathrm{id}) = 0$, since $g \,|\, V_f(-2) = 0$ and $g \,|\, V_f(4) = 0$ $(\because V = V_f(-2) \oplus V_f(4))$. Therefore $\mu_f = (X+2)(X-4)$.

11.A.20 Example Let $F = a_0 + a_1 X + \cdots + a_m X^m \in K[X]$ be a polynomial which annihilates the operator $f : V \longrightarrow V$, i.e. $0 = F(f) = a_0\,\mathrm{id} + a_1 f + \cdots + a_m f^m$. If the constant term $a_0 = F(0)$ of F is non-zero, i.e. $a_0 \neq 0$. Then f is an automorphism of V with

$$f^{-1} = -\frac{1}{a_0}\left(a_1 + \cdots + a_m f^{m-1}\right).$$

If f is an automorphism and if V is finite dimensional, then for F we may take either the characteristic polynomial χ_f or even the minimal polynomial μ_f.

11.A.21 Example (Operators and Matrices of rank one) Let $g: V \longrightarrow V$ be an operator of rank 1 and $\mathrm{Dim}_K V \geq 2$. Then g operates on $\mathrm{Im}\, g$ like a homothecy, say by multiplication by $b \in K$, i.e. $g: \mathrm{Im}\, g \longrightarrow \mathrm{Im}\, g$, $x \longmapsto bx$. It follows that $g^2 = bg$ and hence g annihilates the polynomial $X^2 - bX = X(X-b)$. Since g itself cannot be homothecy ($\mathrm{Dim}_K V \geq 2$), the minimal polynomial of g is $X(X-b)$.
For an arbitrary $a \in K$, the minimal polynomial $\mu_{g+\mathrm{id}_V}$ is $(X-a)(X-(a+b)) = X^2 - (2a+b)X + a(a+b)$.

Therefore $g + a\,\mathrm{id}_V$ is an automorphism if and only if $a(a+b) \neq 0$. Moreover, in this case

$$(g + a\,\mathrm{id}_V)^{-1} = \frac{-1}{a(a+b)}\left(g - (a+b)\,\mathrm{id}_V\right).$$

If V is finite dimensional, then $\chi_g = X^{n-1}(X-b)$, $n = \mathrm{Dim}_K V$ and $b = \mathrm{Tr}\, g$, since $g = 0$ on the $(n-1)$ dimensional subspace $\mathrm{Ker}\, g$, see 11.A.8 (1).
Therefore we have:

If $n \geq 2$ and $\mathfrak{B} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \in M_n(K)$,

is a matrix of rank 1 with the trace $b := b_{11} + \cdots + b_{nn}$, then $\mu_{\mathfrak{B}} = X(X-b)$, $\chi_{\mathfrak{B}} = X^{n-1}(X-1)$ and $\mathfrak{B} + a\mathfrak{E}_n$ is invertible if and only if $a(a+b) \neq 0$. Moreover, in this case

$$(\mathfrak{B} + a\mathfrak{E}_n)^{-1} = \frac{-1}{a(a+b)} \begin{pmatrix} b_{11}-(a+b) & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22}-(a+b) & \cdots & b_{2n} \\ \vdots & \vdots & & \\ b_{n1} & b_{n2} & \cdots & b_{nn}-(a+b) \end{pmatrix}$$

11.A.22 Example The characteristic polynomial of an upper triangular matrix

$$\mathfrak{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \in M_n(K)$$

is $\chi_{\mathfrak{A}} = \operatorname{Det} \begin{pmatrix} X-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a & X-a_{22} & \cdots & -a_{2n} \\ & & \ddots & \\ 0 & 0 & \cdots & X-a_{nn} \end{pmatrix}$

$$= (X-a_{11})(X-a_{22}) \cdots (X-a_{nn}).$$

$\chi_{\mathfrak{A}}$ factors into linear factors and its zeros are the diagonal elements of $\mathfrak{A}$ (counted with multiplicities)

Analogous assertion also holds for lower-triangular matrices.

11.A.23 Example (Companion matrices) For every monic polynomial
$$F = a_0 + a_1 X + \cdots + a_{n-1} X^{n-1} + X^n \in K[X]$$
is the characteristic polynomial of a matrix $\mathfrak{A} \in M_n(K)$ and hence of an operator $f \in \mathrm{End}_K V$, $n = \mathrm{Dim}_K V$. For example, F is the characteristic polynomial of the well-known Companion matrix:
$$\mathfrak{A}_F = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in M_n(K) \text{ of } F.$$

In the matrix $(XE_n - \mathfrak{A}_F)$ adding successively X-times $(n-i)$-th row to the $(n-i-1)$-th row, $i = 0, \ldots, n-2$, we get the matrix:
$$\begin{pmatrix} 0 & 0 & \cdots & 0 & X^n + a_{n-1}X^{n-1} + \cdots + a_0 \\ -1 & 0 & \cdots & 0 & X^{n-1} + a_{n-1}X^{n-2} + \cdots + a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & X^2 + a_{n-1}X + a_{n-2} \\ 0 & 0 & \cdots & -1 & X + a_{n-1} \end{pmatrix}$$
Now, expanding by using the first row the assertion follows. Further, the minimal polynomial of $\mathfrak{A}_F$ is also F: $\mu_{\mathfrak{A}_F} = \chi_{\mathfrak{A}_F} = F$.

If $f : K^n \longrightarrow K^n$ is the operator with the matrix $\mathfrak{A}_F$ with respect to the standard basis $e_1, \ldots, e_n$ of K^n and if $G = b_0 + b_1 X + \cdots + b_{n-1} X^{n-1}$ is a polynomial of degree $< n$ in $K[X]$, then $G(f)(e_1) = b_0 e_1 + b_1 e_2 + \cdots + b_{n-1} e_{n-1} \neq 0$ if $G \neq 0$ and hence $G(f) \neq 0$ ($f(e_i) = e_{i+1}$ for $i = 1, \ldots, n-1$ and $f(e_n) = -a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n$)

11.A.24 Example Let $f : V \longrightarrow V$ be a K-linear operator on a K-vector space V and let $H \subseteq V$ be an hyperplane. Then $H = \operatorname{Ker} h$ for a linear form $h \in V^*$ (which is unique a factor $a \in K^*$). Then: H is f-invariant if and only if h is an eigen-vector of a dual operator $f^* : V^* \rightarrow V^*$.

Proof Suppose that H is f-invariant. Then $f^*(h) = h \circ f$ is 0 on H and $f^*(h)$ is a multiple λh of h. Conversely, $f^*(h) = h \circ f = \lambda h$ with some $\lambda \in K$. Then $h(f(H)) = \lambda h(H) = 0$ and hence $f(H) \subseteq H$. Therefore we have:

11.A.25 Lemma For an operator f on the K-vector space V, there exists f-invariant hyperplane in V if and only if the dual operator f^* has an eigen-value.

By 11.A.16 for an finite dimensional K-vector space V the eigenvalues of f and f^* are same ($\operatorname{Eig} f = N(\chi_f) = N(\chi_{f^*}) = \operatorname{Eig} f^*$). Therefore:

11.A.26 Corollary *For an operator f on the finite dimensional K-vector space V, there exists an f-invariant hyperplane in V if and only if f has an eigenvalue in K.*

One can also prove 11.A.26 without going to the dual operator f^*: If $\lambda \in K$ is an eigenvalue of f, then $f - \lambda\, id_V$ is not injective and hence also not surjective and since $f(x) = \lambda x + (f - \lambda\, id_V)(x)$ every subspace, particularly, every hyperplane in V which contain $\mathrm{Im}(f - \lambda\, id_V)$ is f-invariant. Conversely, if $H \subseteq V$ is an f-invariant hyperplane, and $x \in V \setminus H$, then $f(x) = \lambda x + y$ with $\lambda \in K$ and $y \in H$. Therefore $\mathrm{Im}(f - \lambda\, id_V) \subseteq H$ and also $\mathrm{Ker}(f - \lambda\, id_V) \neq 0$, i.e. λ is an eigen-value of f.

11.A.27 Example (Characteristic polynomials in Algebras) Let A be a finite dimensional algebra over a field K. For an element $x \in A$, the characteristic polynomial χ_{λ_x} of the left-multiplication $\lambda_x : A \longrightarrow A$, $y \longmapsto xy$ is also called the characteristic polynomial of x and is denoted by $\chi_x = \chi^A_{K,x} \in K[X]$.

Since $\chi_x(\lambda_x) : A \longrightarrow A$ is $\lambda_{\chi_x(x)}$,

$(\chi_x(\lambda_x)(y) = \chi_x(\lambda_x(y)) =$

the Cayley-Hamilton Theorem 11.A.7 follows:

11.A.28 Theorem *Let x be an element of the finite dimensional algebra A over the field K. Then $\chi_x(x) = 0$, i.e. the minimal polynomial μ_x of x is a divisor of the characteristic polynomial χ_x of x (in $K[X]$).*

By 11.A.28 the characteristic polynomial χ_x is a canonical algebraic equation for $x \in A$. The determinant of λ_x is the *norm* $Nx = N_K^A x$ of x and the trace of λ_x is called the *trace* of x. *The norm is multiplicative*: $Nxy = Nx\,Ny$; *the trace is K-linear*: $\mathrm{Tr}(ax+by) = a\,\mathrm{Tr}\,x + b\,\mathrm{Tr}\,y$, $a, b \in K$, $x, y \in A$.

The characteristic polynomial $\chi_{K,\mathfrak{a}}^{M_n(K)}$ of $\mathfrak{a} \in M_n(K)$ as an element in the K-algebra $M_n(K)$ is the n-th power of the characteristic polynomial $\chi_{\mathfrak{a}}$ of $\mathfrak{a}$ as the matrix.

Analogous remark for operators on finite dimensional K-vector spaces.