11.B Diagonalisable and Trigonalisable Operators

11.B.1 Definition An operator $f: V \longrightarrow V$ on the K-vector space V is called *diagonalisable* if V has a basis consisting of eigenvectors of f.

If $v_i, i \in I$, is a basis of V consisting of eigenvectors of f, then $f(v_i) = \lambda_i v_i$ with $\lambda_i \in K$, $i \in I$, and the matrix of f with respect to $\underline{v} = \{v_i \mid i \in I\}$ is a diagonal matrix which has the eigenvalues $\lambda_i, i \in I$, on the diagonal. Further, $V = \bigoplus_{\lambda \in K} V_f(\lambda)$ is the sum of eigenspaces $V_f(\lambda), \lambda \in K$, of f and this sum is direct, more generally,

11.B.2 Lemma *Let $f: V \longrightarrow V$ be an operator on the K-vector space V. Then the sum of the eigenspaces $V_f(\lambda)$, $\lambda \in K$, of f is direct.*

Proof This assertion is equivalent with: any family of eigen-vectors of f corresponding to distinct eigenvalues are linearly independent over K: Let $x_1, \ldots, x_n \in V$ be eigen-vectors of f with pairwise distinct $\lambda_1, \ldots, \lambda_n \in K$. We need to show that the subspace $U = Kx_1 + \cdots + Kx_n$ generated by $x_1, \dots, x_n$ (in V) has dimension n. But U is f-invariant and $f|_U$ has n distinct eigen-values $\lambda_1, \dots, \lambda_n$. Then $\mathrm{Dim}_K U = \deg \chi_{f|U} \geq n$ by 11.A.5.

It follows that:

11.B.3 Theorem For an operator $f: V \longrightarrow V$, the following statements are equivalent:

(1) f is diagonalisable.

(2) $V = \sum_{\lambda \in K} V_f(\lambda)$

(3) $V = \sum_{\lambda \in K}^{\oplus} V_f(\lambda)$

Proof The equivalence of (2) and (3) is immediate from 11.B.2. (1) $\Rightarrow$ (2): Clear by definition 11.B.1. (3) $\Rightarrow$ (1): Any basis of $V_f(\lambda)$ contains only eigen-vectors (corresponding to eigenvalue λ) and hence by putting together bases of the eigen-spaces $V_f(\lambda)$, $\lambda \in K$, we get a basis[1] of V consisting of eigenvectors of f.

Let $f: V \longrightarrow V$ be a diagonalisable operator

---

[1] For the implication (3) $\Rightarrow$ (1) we have used the fact that: every vector space has a basis. Therefore one can define the diagonalisibility of an operator by one of the equivalent conditions: (2) or (3) in 11.B.3.

on V, therefore $V = \sum\limits_{\lambda \in K}^{\oplus} V_f(\lambda)$. Let

$p_f(\lambda) = p(\lambda)$, $\lambda \in K$, be the corresponding family of projections of V. Therefore $p(\lambda)$ is the projection onto $V_f(\lambda)$ along the sum of the other eigenspaces $\sum\limits_{\lambda' \in K, \lambda' \neq \lambda} V_f(\lambda')$.

On every eigen-space $V_f(\lambda)$, f induces the homothecy by λ, i.e. $f|_{V_f(\lambda)} : V_f(\lambda) \longrightarrow V_f(\lambda)$, $x \longmapsto \lambda x$

Therefore it follows that $f = \sum\limits_{\lambda \in K} \lambda\, p_f(\lambda)$. This <u>representation</u> or <u>decomposition</u> of f is called the <u>spectral-representation</u> or <u>spectral-decomposition of</u> f.

The operator f commutes with all projections $p(\lambda)$, $\lambda \in K$, i.e. $f\, p(\lambda) = p(\lambda)\, f = \lambda\, p(\lambda)$.
The family of these projections $p(\lambda)$, $\lambda \in K$, is called the <u>spectral-characters</u> of f.

We now would like to give some criterions for the diagonalisibility of operators on the finite dimensional vector spaces. Let V be a finite dimensional K-vector space and $f : V \longrightarrow V$ be an operator on V with characteristic polynomial χ_f.
For $\lambda \in K$, we denote $\alpha(\lambda) = \alpha_f(\lambda)$ the multiplicity of the zero λ of χ_f; this is the exponent of the prime factor $X - \lambda$ in the prime decomposition of

of χ_f : $\chi_f = \prod_{\lambda \in K} (X-\lambda)^{\alpha(\lambda)} \cdot G$ with $G \in K[X]$, $N(G) = \emptyset$.

The integer $\alpha_f(\lambda) \in \mathbb{N}$ is called the algebraic multiplicity of λ for the operator f.
Note that $\alpha_f(\lambda) > 0$ if and only if λ is an eigenvalue and hence a spectral-value of f, see 11.A.5.

The dimension $\mathrm{Dim}_K V_f(\lambda)$ of the eigen-space $V_f(\lambda)$ is denoted by $\gamma(f) = \gamma_f(\lambda)$, $\lambda \in K$ and is called the geometric multiplicity of λ for the operator f.
Note that $\gamma_f(\lambda) > 0$ if and only if λ is an eigenvalue of f. Further, we have:

11.B.4 Lemma Let $f: V \longrightarrow V$ be an operator on the finite dimensional K-vector space V. Then $\gamma_f(\lambda) \leq \alpha_f(\lambda)$ for all $\lambda \in K$.

Proof The restriction $f|_{V_f(\lambda)} : V_f(\lambda) \longrightarrow V_f(\lambda)$ is the homothecy : $x \longmapsto \lambda x$. Therefore the characteristic polynomial of $\chi_f|_{V_f(\lambda)}$ is $(X-\lambda)^{\gamma_f(\lambda)}$ and by 11.A.8(1) this polynomial divides $\chi_f =$ and hence $\gamma_f(\lambda) \leq \alpha_f(\lambda)$.

11.B.5 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 diagonalisable.

(2) The characteristic polynomial χ_f of f splits into linear factors in $K[X]$ and $\gamma_f(\lambda) = \alpha_f(\lambda)$ for all $\lambda \in K$.

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

Proof (1) $\Rightarrow$ (2) and (3): Let $\lambda_1, \cdots, \lambda_r$ be the distinct eigenvalues of f, i.e.

$$N(\chi_f) = N(\mu_f) = \{\lambda_1, \cdots, \lambda_r\}$$

Then $V = V_f(\lambda_1) \oplus \cdots \oplus V_f(\lambda_r)$ by (1)

$$\chi_f = (X-\lambda_1)^{\gamma_f(\lambda_1)} \cdots (X-\lambda_r)^{\gamma_f(\lambda_r)}, \text{ since}$$

$$\deg \chi_f = \mathrm{Dim}_K V = \sum_{i=1}^{r} \mathrm{Dim}_K V_f(\lambda_i) = \sum_{i=1}^{r} \gamma_f(\lambda_i)$$

Further, $\chi_{f|V_f(\lambda_i)} = (X-\lambda_i)^{\gamma_f(\lambda_i)}$ for all $i = 1, \cdots, r$,

since $\mathrm{Dim}_K V_f(\lambda_i) = \gamma_f(\lambda_i)$ and $f|_{V_f(\lambda_i)} : V_f(\lambda_i) \to V_f(\lambda_i)$

is the homothecy $x \longmapsto \lambda_i x$, $i = 1, \cdots r$.

Finally, note that $\mu_f = (X-\lambda_1) \cdots (X-\lambda_r)$, since $f - \lambda_i \, \mathrm{id}_V \big|_{V_f(\lambda_i)} = 0$ for all $i = 1, \cdots, r$. This proves (2) and (3).

(2) $\Rightarrow$ (1): Let $N(\chi_f) = \{\lambda_1, \cdots, \lambda_r\}$. Then $\chi_f = (X-\lambda_1)^{\alpha(\lambda_1)} \cdots (X-\lambda_r)^{\alpha(\lambda_r)}$. We need to show that $V = \sum_{i=1}^{r} V_f(\lambda_i)$, since we already know (see 11.B.2) that this sum is direct. This is immediate from:

$$\mathrm{Dim}_K \sum_{i=1}^{r} V_f(\lambda_i) = \sum_{i=1}^{r} \mathrm{Dim}_K V_f(\lambda_i) = \sum_{i=1}^{r} \gamma_f(\alpha_i)$$
$$= \sum_{i=1}^{r} \alpha_f(\lambda_i) = \mathrm{Dim}_K V.$$

(3) $\Rightarrow$ (1): (Simple and constructive) Suppose that $\mu_f = (X-\lambda_1)\cdots(X-\lambda_r)$, $\mathrm{Eig}(f) = \{\lambda_1, \cdots, \lambda_r\}$. Then $1 = a_1 F_1 + \cdots + a_r F_r$, where $F_\rho = \prod_{\substack{i \neq \rho \\ i=1,\cdots,r}} (X-\lambda_i)$ and $a_\rho := \frac{1}{F_\rho(\lambda_\rho)}$, $\rho = 1, \cdots, r$. (Lagrange's interpolation formula). Therefore

$$id_V = a_1 F_1(f) + \cdots + a_r F_r(f) \text{ and so } V = \sum_{\rho=1}^{r} \mathrm{Im}\, F_\rho(f)$$

Since $\mu_f = (X-\lambda_\rho) F_\rho$, $0 = \mu_f(f) = (f - \lambda_\rho id_V) F_\rho(f)$ and so $\mathrm{Im}\, F_\rho(f) \subseteq \mathrm{Ker}(f - \lambda_\rho id_V) = V_f(\lambda_\rho)$, $\rho = 1, \cdots, r$. Therefore $V = \sum_{\rho=1}^{r} V_f(\lambda_\rho)$.

11.B.6 Corollary Let $f: V \longrightarrow V$ be an operator on the finite dimensional K-vector space V. Suppose that the characteristic polynomial χ_f of f factors into simple linear factors in $K[X]$. Then f is diagonalisable.

Proof Let $n := \mathrm{Dim}_K V = \deg \chi_f$ and let $\lambda_1, \ldots, \lambda_n$ be the n distinct eigen-values of f. Let $x_1, \ldots, x_n$ be eigenvectors corresponding to eigen-values $\lambda_1, \ldots, \lambda_n$. Then $x_1, \ldots, x_n$ are linearly independent over K and hence is a basis of V. Now, f is diagonalisable by definition 11.B.1.

A matrix $\mathfrak{A} \in M_I(K)$ is called diagonalisable if the corresponding operator $f_{\mathfrak{A}} : K^{(I)} \longrightarrow K^{(I)}$ is diagonalisable; this is precisely the case if there exists an invertible matrix $\mathfrak{B} \in GL_I(K)$ such that $\mathfrak{B}\mathfrak{A}\mathfrak{B}^{-1}$ is a diagonal matrix, i.e. $\mathfrak{A}$ is similar to diagonal matrix. Using 11.B.5 and 11.B.6 we can give the corresponding criterion for the diagonalisability of a matrix $\mathfrak{A} \in M_I(K)$ with the help of the characteristic polynomial $\chi_{\mathfrak{A}}$ and the minimal polynomial μ of $\mathfrak{A}$.

Remark Since $\chi_f = \prod_{i=1}^{r} (X - \lambda_i)^{\alpha(\lambda_i)}$ and $\mu_f = \prod_{i=1}^{r} (X - \lambda_i)$ for a diagonalisable operator f, $\mathrm{Eig}\, f = \{\lambda_1, \ldots, \lambda_r\}$, on the finite dimensional K-vector space V. For such an operator f (and the corresponding matrix $\mathfrak{A} \in M_n(K)$; in particular, for matrices in $M_n(K)$ with n distinct eigen values) the

Cayley-Hamilton Theorem (see 11.A.7) $\chi_f(f)=0$ is trivial. With the Kronecker's method of indeterminates (see 10.B.3) we prove this for arbitrary matrices $\mathfrak{A} \in M_n(K)$:
The equation $\chi_{\mathfrak{A}}(\mathfrak{A})=0$ is a system of n^2 polynomial equations in n^2 variables X_{ij}, $1 \leq i, j \leq n$ with coefficients in $\mathbb{Z}$. It is enough to show that these polynomials $f_1, \dots, f_{n^2}$ are identically zero. But these polynomial equations vanish for all $\mathfrak{A} = (x_{ij}) \in M_n(\mathbb{C})$ with n distinct eigenvalues (by above observation) and since the set $\{\mathfrak{A} \in M_n(\mathbb{C}) \mid \mathfrak{A} \text{ has } n \text{ distinct eigenvalues}\} \subseteq \mathbb{C}^{n^2}$ is an non-empty open subset of $\mathbb{C}^{n^2}$, from the identity theorem it follows that $f_1, \dots, f_{n^2}$ are identically zero.
This proves that $f_1=0, \dots, f_{n^2}=0$, i.e. $\chi_{\mathfrak{A}}(\mathfrak{A})=0$.

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 $U_j = \mathrm{Ker}((f-\lambda_1 \mathrm{id}) \cdots (f-\lambda_j \mathrm{id}))$, $j = 0, \dots, n$. 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

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

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 $n := \operatorname{Dim}_K V$. 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 $U_j = \operatorname{Ker}((f-\lambda_1 id)\cdots(f-\lambda_j id))$, $j = 0, \dots, n$. 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 $x_1, \dots, x_{n_j}$, $n_j = \operatorname{Dim}_K U_j$ 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 consequitively 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 $W := \operatorname{Im}(f - \lambda_1 id) \cdots (f - \lambda_j id)$, $j = 0, \dots, n$, 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 $\underline{v} = \{v_1, \dots, v_n\}$, $n = \mathrm{Dim}_K V$, 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

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

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 $s_i := S_i(\lambda_1, \dots, \lambda_n)$ are the i-th elementary symmetric function in elements $\lambda_1, \dots, \lambda_n$. 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 $\operatorname{Tr} f^m$, $m \in \mathbb{N}^*$, 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 $\mathfrak{B} \in M_n(K)$ is the matrix of f with respect to an arbitrary basis $\underline{v}$ of V, then $\mathfrak{B}$ is triagonalisable over L, since $\chi_{\mathfrak{B}} = \chi_f$.

We now go on to prove some results on simultaneous 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 <u>simultaneously diagonalisable</u> 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 simultaneously diagonalisable if and only $V = \sum_{\lambda, \lambda' \in K} V(\lambda, \lambda')$. This sum is obviously direct. Simultaneous 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 $V(\lambda_1, \dots, \lambda_r) := V_{f_1}(\lambda_1) \cap \cdots \cap V_{f_r}(\lambda_r)$

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

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 $\sum_{\lambda_1, \dots, \lambda_r \in K} p_{f_1}(\lambda_1) \cdots p_{f_r}(\lambda_r) = \mathrm{id}_V$ directly follows from the equations $\sum_{\lambda_i \in K} p_{f_i}(\lambda_i) = \mathrm{id}_V$, $i = 1, \dots, r$, 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 $f_1, \dots, f_r$ are pairwise commuting diagonalisable operators on the K-vector space V, then every polynomial in the operators $f_1, \dots, f_r$ 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 $\mathfrak{L} \in GL_I(K)$ such that all the matrices $\mathfrak{L}\mathfrak{A}_j\mathfrak{L}^{-1}$, $j \in J$, are diagonal matrices.

Pairwise commuting triagonalisable operators are also simultaneously triagonalisable:

11.B.17 Theorem Let $f_1, \dots, f_r$ be triagonalisable operators on the n-dimensional K-vector space V. If $f_1, \dots, f_r$ 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) $\Rightarrow$ (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.