Lineare Algebra II

N. Perrin

Düsseldorf
Sommersemester 2013

Inhaltsverzeichnis

1 Wiederholung

In diesem Semester werden wir weiter mit linearen Abbildungen arbeiten. Wir nehmen an, dass alles, was im Skript LA1 steht, bekannt ist. Wir werden aber mit einigen Wiederholungen anfangen.

1.1 Äquivalenzrelationen

Definition 1.1.1 1. Sei M eine Menge. Eine **Relation** auf M ist eine Teilmenge R von $M \times M$. Seien x, y zwei Elemente in M, für $(x, y) \in R$ schreibt man $x \sim_R y$.

2. R heißt **reflexiv**, wenn $x \sim_R x$ für alle $x \in M$.

3. R heißt **symmetrisch**, wenn $x \sim_R y \Rightarrow y \sim_R x$.

4. R heißt **transitiv**, wenn $(x \sim_R y$ und $y \sim_R z) \Rightarrow x \sim_R z$.

Definition 1.1.2 Eine Relation R heißt **Äquivalenzrelation**, wenn R reflexiv, symmetrisch und transitiv ist.

Definition 1.1.3 Sei R eine Äquivalenzrelation auf M.

1. Die **Äquivalenzklasse** $[x]$ ist

$$[x] = \{y \in M \mid x \sim_R y\} \subset M.$$

2. Die **Quotientenmenge** M/R ist die Gesamtheit der Äquivalenzklassen:

$$M/R \ = \{[x] \in \mathfrak{P}(M) \mid x \in M\}.$$

Satz 1.1.4 Sei R eine Äquivalenzrelation auf M. Dann sind alle Elemente aus M in genau einer Äquivalenzklasse. $\square$

Für eine Äquivalenzrelation sind die folgenden Fragen wichtig:

Frage 1.1.5

1. Wann sind zwei Elemente $x, y \in M$ äquivalent?

2. Suche ein Element in jede Äquivalenzklasse.

1.2 Lineare Abbildungen, Matrizen, Basiswechsel

Für die Definitionen von Abbildungen, Körpern, Vektorräumen und Basen verweisen wir auf das Skript LA1 (Definition 2.2.1, Definition 3.1.1 und Definition 5.1.1). Sei K ein Körper und seien V und W zwei K-Vektorräume.

Definition 1.2.1 Eine Abbildung $f : V \to W$ heißt **linear**, wenn für alle $x, y \in K$ und alle $v, v' \in V$ gilt

$$f(xv + yv') = xf(v) + yf(v').$$

Sei $\mathcal{B} = (v_1, \cdots, v_n)$ eine Basis von V und $\mathcal{B}' = (w_1, \cdots, w_m)$ eine Basis von W. Da $\mathcal{B}'$ eine Basis ist, gibt es, für alle $j \in [1, n]$, Skalare $(a_{i,j})_{i\in[1,m]}$ aus K mit

$$f(v_j) = \sum_{i=1}^{m} a_{i,j} w_i.$$

Für Definition und Eigenschaften von Matrizen verweisen wir auf das Skript LA1.

Definition 1.2.2 Die Matrix $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ **von** f **mit den Basen** $\mathcal{B}$**,** $\mathcal{B}'$ ist

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) = (a_{i,j})_{i\in[1,m],\ j\in[1,n]} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}.$$

Sei $f : V \to W$ eine lineare Abbildung. Wenn wir die Basen $\mathcal{B}, \mathcal{B}'$ wechseln, wird sich die Matrix $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ verändern. Der Basiswelchelsatz erklärt, wie sich die Matrix verändert.

Satz 1.2.3 Sei $f : V \to W$ eine lineare Abbildung. Seien $\mathcal{B}, \mathcal{C}$ Basen von V und seien $\mathcal{B}', \mathcal{C}'$ Basen von W. Sei $A = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ und $B = \mathrm{Mat}_{\mathcal{C},\mathcal{C}'}(f)$. Dann gilt

$$B = QAP$$

wobei $P = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\mathrm{Id}_V)$ und $Q = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(\mathrm{Id}_W)$. □

Wir werden zwei Beispiele von Äquivalenzrelationen für Matrizen einführen.

1.3 Äquivalenz von Matrizen

Definition 1.3.1 1. Seien $A, B \in M_{m,n}(K)$. A und B sind **äquivalent**, falls es $P \in \mathrm{GL}_n(K)$ und $Q \in \mathrm{GL}_m(K)$ gibt mit

$$B = QAP.$$

In diesem Fall schreiben wir $A \sim B$.

2. Sei R die Relation $R = \{(A, B) \in M_{m,n}(K) \mid A \sim B\}$.

Lemma 1.3.2 Die Relation R ist eine Äquivalenzrelation. □

Satz 1.3.3 Seien $A, B \in M_{m,n}(K)$.

$$A \sim B \Leftrightarrow \mathrm{Rg}(A) = \mathrm{Rg}(B).$$

Wir können also die Frage: *wann sind zwei Elemente $A, B \in M$ äquivalent?* antworten: Zwei Matrizen A, B sind äquivalent genau dann, wenn $\mathrm{Rg}(A) = \mathrm{Rg}(B)$.

Um die zweite Frage: *suche ein Element aus jeder Äquivalenzklasse* zu beantworten brauchen wir die folgende Definition.

Definition 1.3.4 Sei $A \in M_{m,n}(K)$ mit $\mathrm{Rg}(A) = r$ Dann heißt

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(K)$$

die **Normalform** von A bzg. Äquivalenz von Matrizen.

Wir haben gesehen, dass die Äquivalenzklasse einer Matrix A mir $\mathrm{Rg}(A) = r$ die folgende Menge ist:

$$[A]_\sim = \{B \in M_{n,m}(K) \mid \mathrm{Rg}(B) = \mathrm{Rg}(A) = r\}.$$

Wir haben in $[A]$ ein sehr einfaches Element: die **Normalform** von A.

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in [A]_\sim.$$

1.4 Basiswechsel für Endomorphismen, Ähnlichkeit

Satz 1.4.1 Sei V ein n-dimensionaler Vektorraum. Seien $\mathcal{B}$ und $\mathcal{C}$ Basen von V und sei $f : V \to V$ linear. Sei $A = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f)$ und $B = \mathrm{Mat}_{\mathcal{C},\mathcal{C}}(f)$. Dann gilt

$$B = P^{-1}AP,$$

wobei $P = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\mathrm{Id}_V)$. □

Definition 1.4.2 1. Seien $A, B \in M_n(K)$. Dann sind A und B **ähnlich**, falls es ein $P \in \mathrm{GL}_n(K)$ gibt mit

$$B = P^{-1}AP.$$

In diesem Fall schreiben wir $A \approx B$.

2. Sei R' die Relation $R' = \{(A, B) \in M_n(K) \mid A \approx B\}$.

Lemma 1.4.3 Die Relation R' ist eine Äquivalenzrelation. □

Die zwei wichtigen Fragen für die Ähnlichkeitrelation sind:

Frage 1.4.4

1. Wann sind zwei Matrizen $A, B \in M_n(K)$ ähnlich?

2. Suche eine Normalform bzg. Ähnlichkeit von Matrizen.

Wir werden dieses Semester diese Fragen beantworten.

1.5 Erste Invarianten für die Ähnlichkeitsrelation

Lemma 1.5.1 Seien $A, B \in M_n(K)$. Es gilt

$$A \approx B \Rightarrow A \sim B.$$

Beweis. Seien $A, B \in M_n(K)$ mit $A \approx B$. Nach der Definition gibt es ein $P \in \mathrm{GL}_n(K)$ mit $B = P^{-1}AP$. Sei $Q = P^{-1} \in \mathrm{GL}_n(K)$, dann gilt $B = QAP$ und $A \sim B$. ■

Korollar 1.5.2 Seien $A, B \in M_n(K)$ mit $A \approx B$. Dann gilt $\mathrm{Rg}(A) = \mathrm{Rg}(B)$.

Beweis. Folgt aus Satz 1.3.3. ■

Beispiel 1.5.3 In Korollar 1.5.2 haben wir nicht $\mathrm{Rg}(A) = \mathrm{Rg}(B) \Rightarrow A \approx B$. Seien

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Für $C \approx A$ gilt: es gibt $P \in \mathrm{GL}_2(K)$ mit

$$C = P^{-1}AP = P^{-1}I_2P = P^{-1}P = I_2 = A.$$

Es gilt also

$$[A]_{\approx} = \{A\}.$$

Die einzige Matrix die ähnlich zu A ist, ist die Matrix A. Also gilt $\mathrm{Rg}(A) = 2 = \mathrm{Rg}(B)$ (z.B. beide Determinanten sind ungleich 0) aber $A \not\approx B$.

Nächstes Semester haben wir den folgende Satz bewiesen.

Satz 1.5.4 Seien $A, B \in M_n(K)$ mit $A \approx B$. Dann gilt $\chi_A = \chi_B$. □

Korollar 1.5.5 Seien $A, B \in M_n(K)$ mit $A \approx B$. Dann sind die Eigenwerte von A und B gleich.

Beispiel 1.5.6 In Satz 1.5.4 haben wir nicht $\chi_A = \chi_A \Rightarrow A \approx B$. Seien

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Es gilt

$$\chi_A = (X-1)^2 = \chi_B.$$

Die Eigenwerte von A und B sind gleich (der einzige Eigenwert ist 1). Aber, wie in Beispiel 1.5.3, gilt $A \not\approx B$.

Wir geben hier eine hinreichende Bedingung für die Ähnlichkeit von Matrizen.

Satz 1.5.7 Seien $A \in M_n(K)$ mit n paarweise verschiedenen Eigenwerten $\lambda_1, \cdots, \lambda_n$ und sei $B \in M_n(K)$ mit $\lambda_1, \cdots, \lambda_n$ als Eigenwerten. Dann gilt $A \approx B$. □

Beweis. Wir wissen (siehe Satz 1.7.5), dass die Matrix A und auch die Matrix B diagonalisierbar mit den Eigenwerten $\lambda_1, \cdots, \lambda_n$ sind. Es gibt also Matrizen $P, Q \in \mathrm{GL}_n(K)$ mit

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = Q^{-1}BQ.$$

Es gilt also $A \approx D \approx B$. ■

Beispiel 1.5.8 Im Satz 1.5.7 haben wir nicht

$(A \approx B) \Rightarrow$ (A und B haben die gleichen n paarweise verschiedenen Eigenwerte).

Seien

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = B.$$

Dann gilt $A \approx B$ und A und B haben die gleichen Eigenwerte, aber A und B haben nur einen Eigenwert und nicht 2 paarweise verschiedene Eigenwerte.

Diese Beispiele und erste Invarianten zeigen, dass Diagonalisierbarkeit einen starken Zusammenhang mit Ähnlichkeit hat. Wir werden aber mehr brauchen. Wir wiederholen jetzt die Eigenschaften von diagonalisierbaren Matrizen.

1.6 Eigenwerte und Eigenvektoren

Definition 1.6.1 1. Sei $f : V \to V$ ein Endomorphismus von V. Ein Vektor $v \in V \setminus \{0\}$ heißt **Eigenvektor mit Eigenwert** $\lambda \in K$ falls gilt

$$f(v) = \lambda v.$$

2. Sei $A \in M_n(K)$ eine Matrix. Ein Vektor $v \in K^n \setminus \{0\}$ heißt **Eigenvektor mit Eigenwert** $\lambda \in K$ falls gilt

$$Av = \lambda v.$$

Definition 1.6.2 Sei $\lambda \in K$ und $f : V \to V$ ein Endomorphismus. Der **Eigenraum** $E(f, \lambda)$ **zu** f **und** λ ist der Unterraum

$$E(f, \lambda) = \mathrm{Ker}(\lambda \mathrm{Id}_V - f) = \{v \in V \mid f(v) = \lambda v\}.$$

Satz 1.6.3 Die Eigenwerte von f sind die Nullstelen von χ_f. □

Satz 1.6.4 Sei $f \in \mathrm{End}(V)$.

1. Für $\lambda \neq \mu$ gilt $E(f, \lambda) \cap E(f, \mu) = 0$.

2. Systeme von Eigenvektoren mit paarweise verschiedenen Eigenwerten von f sind linear unabhängig. □

Sei $n = \dim V$

Korollar 1.6.5 Sei $f \in \mathrm{End}(V)$. Dann hat f höchstens n Eigenwerte.

Korollar 1.6.6 Sei $f \in \mathrm{End}(V)$. Dann gilt

$$\sum_{\lambda \in K} E(f, \lambda) = \bigoplus_{\lambda \in K} E(f, \lambda).$$

1.7 Diagonalisierbare Matrizen

Definition 1.7.1 Eine Matrix $A = (a_{i,j}) \in M_n(K)$ heißt diagonal wenn gilt: $a_{i,j} = 0$ für alle $i \neq j$.

Definition 1.7.2 Eine Matrix $A \in M_n(K)$ ist **diagonalisierbar** falls sie ähnlich zu einer Diagonalmatrix ist, *i.e.* falls es $P \in \mathrm{GL}_n(K)$ gibt so dass PAP^{-1} eine Diagonalmatrix ist.

Bemerkung 1.7.3 Eine Matrix A ist diagonalisierbar genau dann, wenn es in der Ähnlichkeitsklasse von A eine Diagonalmatrix D gibt. Für diagonalisierbare Matrizen gibt es ein sehr einfaches Element: die Diagonalmatrix D. Diese Diagonalmatrix D wird die **(jordansche) Normalform** von A sein.

Satz 1.7.4 Sei $A \in M_n(K)$. Dann sind folgende Aussagen äquivalent:

1. A ist diagonalisierbar.
2. Es gibt eine Basis $\mathcal{B}$ von K^n, welche aus Eigenvektoren von A besteht.
3. $\sum_{\lambda \in K} \dim E(A, \lambda) = n$.
4. $\oplus_{\lambda \in K} E(A, \lambda) = K^n$. $\square$

Satz 1.7.5 Sei $n = \dim V$ und $f \in \mathrm{End}(V)$. Hat f genau n verschiedene Eigenwerte, dann ist f diagonalisierbar. $\square$

1.8 Eigenwerte und das charakteristische Polynom

Satz 1.8.1 Sei $A \in M_n(K)$ und sei $f \in \mathrm{End}(V)$. Es gilt

$$\{\text{Eigenwerte von } A\} = \{\text{Nullstellen von } \chi_A\}$$
$$\{\text{Eigenwerte von } f\} = \{\text{Nullstellen von } \chi_f\}.$$

Satz 1.8.2 Sei $n = \dim V$ und sei $f \in \mathrm{End}(V)$. Für jedes $\lambda \in K$ gilt dann

$$\dim E(f, \lambda) \leq m(\chi_f, \lambda),$$

wobei $m(\chi_f, \lambda)$ die Vielfachkeit von λ in χ_f ist. $\square$

Korollar 1.8.3 Sei $n = \dim V$ und sei $f \in \mathrm{End}(V)$. Der Endomorphismus f ist diagonalisierbar genau dann, wenn χ_f vollständig in Linearfaktoren zerfällt und für jedes $\lambda \in K$, gilt $\dim E(f, \lambda) = m(\chi_f, \lambda)$.

1.9 Trigonalisierbarkeit

Definition 1.9.1 1. Eine Matrix $A = (a_{i,j}) \in M_n(K)$ ist eine obere Dreieckmatrix wenn $a_{i,j} = 0$ für $i > j$.

2. Sei $n = \dim V$ und $f \in \mathrm{End}(V)$. Der Endomorphismus f heißt **trigonalisierbar** falls es eine Basis $\mathcal{B}$ gibt mit $\mathrm{Mat}_{\mathcal{B}}(f)$ eine obere Dreieckmatrix.

Bemerkung 1.9.2 Eine Matrix A is diagonalisierbar genau dann, wenn es in der Ähnlichkeitsklasse von A eine obere Dreieckmatrix D gibt.

Satz 1.9.3 Sei $f \in \mathrm{End}(V)$. Die folgende Aussagen sind äquivalent:

1. f ist trigonalisierbar.
2. χ_f zerfällt über K vollstandig in Linearfaktoren. □

Korollar 1.9.4 Falls K algebraisch abgeschlossen ist, falls also jedes Polynom in $K[X]\backslash\{0\}$ über K in Linearfaktoren zerfällt, dann ist jedes $f \in \mathrm{End}(V)$ mit $\dim V < \infty$ trigonalisierbar.

Bemerkung 1.9.5 Für K algebraisch abgeschlossen, gibt es immer in der Ähnlichkeitsklasse $[A]_\approx$ von A eine obere Dreieckmatrix. Wir können also als einfaches Element in der Ähnlichkeitsklasse eine obere Dreieckmatrix wählen. Wir werden sehen, dass man eine noch einfachere Matrix wählen kann: die **(jordansche) Normalform** von A.

1.10 Minimal Polynom

Sei V mit $\dim V = n$ und sei $f \in \mathrm{End}(V)$.

Satz 1.10.1 ann existiert genau ein normiertes Polynom $\mu_f \in K[X]$, das **Minimalpolynom von** f mit

1. $\mu_f(f) = 0$
2. Ist $P \in K[X]$ mit $P(f) = 0$, so ist μ_f ein Teiler von P. □

Satz 1.10.2 Dann sind folgende Aussagen äquivalent:

1. f ist diagonalisierbar.
2. μ_f zerfällt vollständig in Linearfaktoren und besitzt nur einfache Nullstellen. □

Satz 1.10.3 (Satz von Cayley-Hamilton) Es gilt $\chi_f(f) = 0$. □

Korollar 1.10.4 Es gilt: μ_f ist ein Teiler von χ_f.

Korollar 1.10.5 μ_f und χ_f haben die gleichen Nullstellen (die Eigenwerte). Sei λ eine solche Nullstelle, es gilt

$$m(\mu_f, \lambda) \leq m(\chi_f, \lambda).$$

Satz 1.10.6 Seien $A, B \in M_n(K)$ mit $A \approx B$. Dann gilt $\mu_A = \mu_B$. □

Beweis. Sei $P \in \mathrm{GL}_n(K)$ mit $B = P^{-1}AP$. Es gilt also auch $A = PBP^{-1}$. Eine einfache Induktion gibt für alle $i \in \mathbb{N}$:

$$B^i = P^{-1}A^iP.$$

Sei $\mu_A = \sum_{i=0}^{k} a_i X_i \in K[X]$. Es gilt $\mu_A(A) = 0$. Wir zeigen, dass $\mu_A(B) = 0$. Es gilt

$$\mu_A(B) = \sum_{i=0}^{k} a_i B^k = \sum_{i=0}^{k} a_i P^{-1}A^kP = P^{-1}\left(\sum_{i=0}^{k} a_i A^k\right)P = P^{-1}\mu_A(A)P = 0.$$

Es gilt also: $\mu_A(B) = 0$ und μ_B ist ein Teiler von μ_A.

Wir können A und B vertauchen und so gilt auch $\mu_B(A) = 0$. Daraus folgt, dass μ_A ein Teiler von μ_B ist. Es folgt, dass $\mu_A = \lambda\mu_B$ mit $\lambda \in K$, und weil μ_A und μ_B beide normiert sind, folgt $\mu_A = \mu_B$. ■

Beispiel 1.10.7 Im Satz 1.10.6 haben wir nicht $\mu_A = \mu_B \Rightarrow A \approx B$. Seien

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Nach Korollar 1.10.5 hat μ_A (bzg. μ_B) die Eigenwerte von A (bzg. B) als Nullstellen. Also haben μ_A und μ_B die Zahlen 1 und 2 als Nullstellen. Die beiden Matrizen A und B sind Diagonalmatrizen, also diagonalisierbar. Nach Satz 1.10.2 folgt, dass μ_A und μ_B einfache Nullstellen haben. Es folgt

$$\mu_A = (X-1)(X-2) = \mu_B.$$

Wir zeigen, dass $A \not\approx B$. Hätten wir $A \approx B$, dann folgt nach Satz 1.5.4 $\chi_A = \chi_B$. Aber es gilt

$$\chi_A = (X-1)^2(X-2) \neq (X-1)(X-2)^2 = \chi_B.$$

Also $A \not\approx B$.

Beispiel 1.10.8 Es gibt Matrizen A und B mit

$$\mathrm{Rg}(A) = \mathrm{Rg}(B),\ \chi_A = \chi_B \text{ und } \mu_A = \mu_B$$

, aber mit $A \not\approx B$.

Seien

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Es gilt

$$\mathrm{Rg}(A) = 4 = \mathrm{Rg}(B),\ \chi_A = (X-1)^4 = \chi_B \text{ und } \mu_A = (X-1)^2 = \mu_B.$$

Aber es gilt $A \not\approx B$.

Übung 1.10.9 Seien A und B wie im Beispiel 1.10.8.

1. Zeigen Sie, dass $\mathrm{Rg}(A) = 4 = \mathrm{Rg}(B)$, $\chi_A = (X-1)^4 = \chi_B$ und $\mu_A = (X-1)^2 = \mu_B$.
2. Zeigen Sie, dass $A \not\approx B$.

2 Jordansche Normalform

Sei V ein K-Vektorraum der Dimension n und sei $f \in \mathrm{End}(V)$ ein Endomorphismus.

2.1 Invariante Unterräume

Definition 2.1.1 Ein Unterraum U von V heißt **invariant** für f (oder f-**invariant**) falls $f(U) \subset U$.

Lemma 2.1.2 Sei U ein Unterraum von V.

(I) Wenn U f-invariant ist, dann ist U auch $P(f)$-invariant für alle $P \in K[X]$.

(II) Sei $\lambda \in K$. Dann ist U genau dann f-invariant, wenn U $(f - \lambda \mathrm{Id}_V)$-invariant ist.$\square$

Beweis. (I) Sei $P \in K[X]$ und $u \in U$. Dann ist $f(u) \in U$ und per Induktion gilt $f^k(u) \in U$ für alle $k \in \mathbb{N}$. Daraus folgt $P(f)(u) \in U$.

(II) Angenommen U sei f-invariant. Dann gilt $f(u) \in U$ für alle $u \in U$. Es folgt $(f - \lambda \mathrm{Id}_V)(u) = f(u) - \lambda u \in U$ und U ist $(f - \lambda \mathrm{Id}_V)$-invariant. Umgekehrt, sei $u \in U$, dann gilt $(f - \lambda \mathrm{Id}_V)(u) \in U$ also $f(u) - \lambda u \in U$. daraus folgt $f(u) \in U$ und U ist f-invariant. ∎

Lemma 2.1.3 Seien $U_1, \cdots, U_r$ f-invariante Unterräume so dass, $V = U_1 \oplus \cdots \oplus U_r$.

(I) Seien $\mathcal{B}_1, \cdots, \mathcal{B}_r$ Basen von $U_1, \cdots, U_r$. Dann ist $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ eine Basis von V.

(II) Sei $A_i = \mathrm{Mat}_{\mathcal{B}_i}(f|_{U_i})$ für $i \in [1, r]$, dann gilt

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_r \end{pmatrix}.$$

Beweis. Übung. ∎

Lemma 2.1.4 Umgekehrt, sei $\mathcal{B} = (v_1, \cdots, v_n)$ eine Basis mit

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_r \end{pmatrix}$$

wobei $A_i \in M_{n_i}(K)$. Dann sind die Unterräume

$$U_i = \langle v_{n_1+\cdots+n_{i-1}+1}, \cdots, v_{n_1+\cdots+n_{i-1}+n_i} \rangle$$

f-invariant und es gilt $V = U_1 \oplus \cdots \oplus U_r$. □

Beweis. Übung. ■

2.2 Verallgemeinerte Eigenräume

Definition 2.2.1 Seien $k \in \mathbb{N}$ und $\lambda \in K$. **Der k-te verallgemeinerte Eigenraum zum Eigenwert** λ ist $E_k(f, \lambda) = \mathrm{Ker}(f - \lambda \mathrm{Id}_V)^k$.

Bemerkung 2.2.2 Es gilt $E_1(f, \lambda) = E(f, \lambda)$ also ist der erste verallgemeinerte Eigenraum zum Eigenwert λ der Eigenraum zum Eigenwert λ.

Lemma 2.2.3 Sei $\lambda \in K$.

(i) Für jedes $k \in \mathbb{N}$ ist $E_k(f, \lambda)$ f-invariant.

(ii) Es gilt $E_k(f, \lambda) \subset E_l(f, \lambda)$ für $k \leq l$.

(iii) Es gibt ein $k \in \mathbb{N}$ mit $E_k(f, \lambda) = E_{k+1}(f, \lambda)$.

(iv) Sei k mit $E_k(f, \lambda) = E_{k+1}(f, \lambda)$, dann gilt $E_k(f, \lambda) = E_l(f, \lambda)$ für alle $l \geq k$. □

Beweis. (i) Sei $v \in E_k(f, \lambda)$. Dann gilt $(f - \lambda \mathrm{Id}_V)^k(v) = 0$. Wir zeigen, dass $f(v) \in E_k(f, \lambda)$ also $(f - \lambda \mathrm{Id}_V)^k(f(v)) = 0$. Es gilt

$$(f - \lambda \mathrm{Id}_V)^k(f(v)) = ((f - \lambda \mathrm{Id}_V)^k \circ f)(v) = f \circ (f - \lambda \mathrm{Id}_V)^k(v) = f((f - \lambda \mathrm{Id}_V)^k(v)) = 0.$$

(ii) Sei $v \in E_k(f, \lambda)$ und $l \geq k$. Dann gilt $(f - \lambda \mathrm{Id}_V)^k(v) = 0$, Also gilt $(f - \lambda \mathrm{Id}_V)^l(v) = (f - \lambda \mathrm{Id}_V)^{l-k}((f - \lambda \mathrm{Id}_V)^k(v)) = (f - \lambda \mathrm{Id}_V)^{l-k}(0) = 0$. Es gilt also $v \in E_l(f, \lambda)$.

(iii) Wir betrachten $d_k = \dim E_k(f, \lambda$. Die Folge $(d_k)_{k \in \mathbb{N}}$ ist steigend und $d_k \leq n$. Es gibt also ein k mit $d_k = d_{k+1}$ also $\dim E_k(f, \lambda) = \dim E_{k+1}(f, \lambda)$. Daraus folgt $E_k(f, \lambda) = E_{k+1}(f, \lambda)$.

(iv) Sei k mit $E_k(f,\lambda) = E_{k+1}(f,\lambda)$ und sei $l \geq k$. Es gilt $E_k(f,\lambda) \subset E_l(f,\lambda)$. Umgekehrt zeigen wir per Induktion über $l \geq k$, dass $E_l(f,\lambda) \subset E_k(f,\lambda)$. Für $l = k$ ist dies wahr.

Angenommen $E_l(f,\lambda) \subset E_k(f,\lambda)$. Wir zeigen $E_{l+1}(f,\lambda) \subset E_k(f,\lambda)$. Sei $v \in E_{l+1}(f,\lambda)$. Es gilt $(f - \lambda\mathrm{Id}_V)^{l+1}(v) = 0$, also $(f - \lambda\mathrm{Id}_V)^{k+1}((f - \lambda\mathrm{Id}_V)^{l-k}(v)) = 0$. Es folgt $(f - \lambda\mathrm{Id}_V)^{l-k}(v) \in E_{k+1}(f,\lambda) = E_k(f,\lambda)$. Es gilt also $0 = (f - \lambda\mathrm{Id}_V)^k((f - \lambda\mathrm{Id}_V)^{l-k}(v)) = (f - \lambda\mathrm{Id}_V)^l(v) = 0$ und $v \in E_l(f,\lambda) \subset E_k(f,\lambda)$. ∎

Korollar 2.2.4 Sei $\lambda \in K$. Dann gibt es ein $M_\lambda \in \mathbb{N}$ mit $E_k(f,\lambda) \subsetneq E_{k+1}(f,\lambda)$ für $k < M_\lambda$ und $E_k(f,\lambda) = E_{k+1}(f,\lambda)$ für $k \geq M_\lambda$.

2.3 Haupträume

Definition 2.3.1 Der Hauptraum zum Eigenwert λ ist $H(f,\lambda) = E_{M_\lambda}(f,\lambda)$.

Lemma 2.3.2 (i) Ist λ ein Eigenwert von f, so gilt $H(f,\lambda) \neq 0$.

(ii) Sonst gilt $H(f,\lambda) = 0$. □

Beweis. (i) Sei v ein Eigenvektor zu λ. Es gilt $v \neq 0$ und $(f - \lambda\mathrm{Id}_V)(v) = 0$. Es gilt also $0 \neq E(f,\lambda) \subset H(f,\lambda)$.

(ii) Angenommen $H(f,\lambda) \neq 0$. Dann gibt es ein $v \in H(f,\lambda)$ mit $v \neq 0$. Es gilt $(f-\lambda)^{M_\lambda}(v) = 0$ und daraus folgt $(f-\lambda)^l(v) = 0$ für alle $l \geq M_\lambda$. Sei k maximal mit der Eigenschaft $(f-\lambda)^k(v) \neq 0$ (z.B. hat $k = 0$ diese Eigenschaft, aber alle $k \geq M_\lambda$ haben diese Eigenschaft nicht mehr). Es gilt also $(f-\lambda)^k(v) \neq 0$ und $(f-\lambda)^{k+1}(v) = 0$. Daraus folgt

$$0 = (f-\lambda)^{k+1}(v) = (f - \lambda\mathrm{Id}_V)((f-\lambda)^k(v)).$$

Also ist $(f-\lambda)^k(v)$ ein Eigenvektor von f mit dem Eigenwert λ. Widerspruch. ∎

Wir werden die Haupträume dank dem Minimalpolynom studieren. Zuerst brauchen wir ein Lemma.

Definition 2.3.3 Seien $P_1, \cdots, P_r \in K[X]$. Die Polynome $P_1, \cdots, P_r$ sind **teilerfremd**, falls es kein $Q \in K[X]$ mit $\deg(Q) > 0$ und $Q|P_i$ für alle $i \in [1,r]$ gibt.

Beispiel 2.3.4 (i) X und $X - 1$ sind teilerfremd.

(ii) Für $\lambda_1, \cdots, \lambda_r$ paarweise verschieden sind $P_1 = (X-\lambda_1)^{m_1}, \cdots, P_r = (X-\lambda_r)^{m_r}$ teilerfremd.

(iii) Für $\lambda_1, \cdots, \lambda_r$ paarweise verschieden sei

$$P_i = \prod_{j \neq i} (X - \lambda_j)^{m_j}.$$

Dann sind $P_1, \cdots, P_r$ teilerfremd.

(iv) Für $P_1 = \cdots = P_r = 0$ sind $P_1, \cdots, P_r$ nicht teilerfremd. Jedes Polynom P teilt $P_1, \cdots, P_r$: $P_i = 0 = 0 \cdot P$.

Lemma 2.3.5 Seien $P_1, \cdots, P_r \in K[X]$ teilerfremd. Dann gibt es Polynome $Q_1, \cdots, Q_r \in K[X]$ mit

$$Q_1 P_1 + \cdots + Q_r P_r = 1.$$

Beweis. Nach Induktion über $N = \deg(P_1) + \cdots + \deg(P_r)$.

Für $N = 0$ gilt $\deg(P_1) = \deg(P_r) = 0$. Es gibt also Skalare $\lambda_1 \cdots, \lambda_r \in K$ mit $P_i = \lambda_i$ für alle $i \in [1, r]$. Es gibt ein i mit $\lambda_i \neq 0$. (Wenn nicht gilt $\lambda_1 = \cdots = \lambda_r = 0$, also $P_1 = \cdots = P_r = 0$ und $P_1, \cdots, P_r$ sind nicht teilerfremd.) Sei $Q_i = \frac{1}{\lambda_i}$ und $Q_j = 0$ für $j \neq i$, also gilt $Q_1 P_1 + \cdots + Q_r P_r = 1$.

Wir nehmen an, dass es für alle teilerfremden Polynome $R_1, \cdots, R_r$ mit $N \geq \deg(R_1) + \cdots + \deg(R_r)$ Polynome $S_1, \cdots, S_r$ gibt mit $S_1 R_1 + \cdots + S_r R_r = 1$. Seien $P_1, \cdots, P_r$ teilerfremde Polynome mit $\deg(P_1) + \cdots + \deg(P_r) = N + 1$. Ohne Beschränkung können wir annehmen, dass $\deg(P_1) \geq \cdots \deg(P_r)$. Wir wissen, dass es für alle $i \in [1, r-1]$ Polynome U_i, R_i mit $P_i = T_i P_r + R_i$ und $\deg(R_i) < \deg(P_r) \leq \deg(P_i)$ gibt. Es gilt also $\deg(R_1) + \cdots + \deg(R_{r-1}) + \deg(P_r) < \deg(P_1) + \cdots + \deg(P_{r-1}) + \deg(P_r)$.

Wir zeigen, dass $R_1, \cdots, R_{r-1}, R_r = P_r$ teilerfremd sind. Sei $P \in K[X]$ mit $P | R_i$ für alle $i \in [1, r]$. Es gilt $P | R_i$ und $P | R_r = P_r$. Also teilt P alle Polynome $T_i P_r + R_i = P_i$. Da $P_1, \cdots, P_r$ teilerfremd sind gilt $\deg(P) = 0$ und $R_1, \cdots, R_r$ sind teilerfremd. Nach Induktion gibt es Polynome $S_1, \cdots, S_r$ mit $S_1 R_1 + \cdots + S_r R_r = 1$. Wir setzen $R_i = P_i - T_i P_r$ für $i \in [1, r-1]$ und $R_r = P_r$. Es gilt

$$1 = S_1 R_1 + \cdots + S_r R_r = S_1 (P_1 - T_1 P_r) + \cdots + S_{r-1} (P_{r-1} - T_{r-1} P_r) + S_r P_r.$$

Wir setzen $Q_i = S_i$ für $i \in [1, r-1]$ und $Q_r = S_r - (S_1 T_1 + \cdots + S_{r-1} T_{r-1})$. Die Gleichung $Q_1 P_1 + \cdots + Q_r P_r = 1$ folgt. ∎

Beispiel 2.3.6 Sei $P_1 = X$ und $P_2 = X - 1$. Dann sind P_1 und P_2 teilerfremd und für $Q_1 = 1$, $Q_2 = -1$ gilt $Q_1 P_1 + Q_2 P_2 = 1$.

Sei μ_f das Minimalpolynom von f. Wir nehmen an, dass μ_f in Linearfaktoren zerfällt:

$$\mu_f = (X - \lambda_1)^{m_1} \cdots (X - \lambda^r)^{m_r},$$

wobei $\lambda_1, \cdots, \lambda_r$ paarweise verschieden sind.

Satz 2.3.7 Sei $H_i = \mathrm{Ker}(f - \lambda_i)^{m_i}$ für $i \in [1, r]$. Es gilt

$$V = H_1 \oplus \cdots \oplus H_r.$$

Beweis. Wir zeigen $V = H_1 + \cdots + H_r$. Sei $v \in V$. Wir zeigen, dass es Vektoren $v_i \in H_i$ für $i \in [1, r]$ gibt mit $v = v_1 + \cdots + v_r$. Sei $P_i = \prod_{j \neq i}(X - \lambda_j)^{m_j}$ für $i \in [1, r]$. Dann sind $P_1, \cdots, P_r$ teilerfremd. Nach dem obigen Lemma gibt es Polynome $Q_1, \cdots, Q_r$ mit $P_1Q_1 + \cdots + P_rQ_r = 1$. Es gilt also

$$v = \mathrm{Id}_V(v) = (P_1(f)Q_1(f) + \cdots + P_r(f)Q_r(f))(v).$$

Sei $v_i = P_i(f)Q_i(f)(v)$. Es gilt $v = v_1 + \cdots + v_r$. Wir zeigen $v_i \in H_i$. Es gilt

$$(f - \lambda_i)_i^m(v_i) = (f - \lambda_i)_i^m P_i(f)Q_i(f)(v) = \mu_f(f)Q_i(f)(v) = 0(f)Q_i(f)(v) = 0.$$

Daraus folgt $v_i \in H_i$.

Wir zeigen jetzt, dass die Summe $H_1 + \cdots + H_r$ eine direkte Summe ist. Seien also $v_i \in H_i$ mit $v_1 + \cdots + v_r = 0$. Wir zeigen $v_i = 0$ für alle $i \in [1, r]$. Es gilt

$$0 = P_i(f)(v_1) + \cdots + P_i(f)(v_r) = P_i(f)(v_i)$$

da $(X - \lambda_j)^{m_j}$ P_i für alle $j \neq i$ teilt. Sei $R = (X - \lambda_i)^{m_i}$. Es gilt $R(f)(v_i)$. Die Polynome P_i und $R = (X - \lambda_i)^{m_i}$ sind teilerfremd. Es gibt also Polynome Q und S mit $QP_i + SR = 1$. Daraus folgt

$$v_i = Q(f)P_i(f)(v_i) + S(f)R(f)(v_i) = 0.$$

Da der obige Beweis für alle $i \in [1, r]$ gilt, gilt also $v_i = 0$ für alle $i \in [1, r]$. ∎

Korollar 2.3.8 Für alle $i \in [1, r]$ gilt $H(f, \lambda_i) = H_i$ und $M_{\lambda_i} = m_i$.

Beweis. Für alle $i \in [1, r]$ und $k \leq M_{\lambda_i} \leq l$ gilt

$$\mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^k \subset \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{M_{\lambda_i}} \subset \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^l.$$

Es gilt also $H_i \subset H(f, \lambda_i)$.

Umgekehrt, sei $v \in H(f, \lambda_i)$. Wir zeigen, dass $v \in H_i$. Nach dem obigen Satz gilt $v = v_1 + \cdots + v_r$ mit $v_j \in H_j$ für alle $j \in [1, r]$. Sei $P_i = \prod_{j \neq i}(X - \lambda_j)^{m_j}$ und $R = (X - \lambda_i)^{M_{\lambda_i}}$. Es gilt $P_i(f)(v_j) = 0$ für alle $j \in [1, r]$ und $R(f)(v) = 0$. Die Polynome P_i und R sind teilerfremd. Es gibt also $Q, S \in K[X]$ mit $1 = QP_i + SR$. Daraus folgt

$$v = Q(f)P_i(f)(v_1 + \cdots + v_r) + S(f)R(f)(v) = Q(f)P_i(f)(v_i).$$

Da $v_i \in H_i$ und H_i f-invariant, gilt $v = Q(f)P_i(f)(v_i) \in H_i$.

Wir zeigen $m_i = M_{\lambda_i}$. Es gilt

$\mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i+1} \subset \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{M_{\lambda_i}} = \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i} \subset \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i+1}$.

Alle Enthaltungen sind Gleichungen und es folgt $\mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i} = \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i+1}$. Nach der Definition von M_{λ_i} gilt $M_{\lambda_i} \leq m_i$. Sei

$$P = (X - \lambda_i)^{M_{\lambda_i}} \cdots (X - \lambda_r)^{M_{\lambda_r}}.$$

Wir zeigen, dass $P(f) = 0$. Sei also $v \in V$. Wir zeigen $P(f)(v) = 0$. Nach dem obigen Satz gilt $v = v_1 + \cdots + v_r$ mit $v_i \in H_i$. Es gilt also $P(f)(v_i) = 0$ für alle $i \in [1, r]$. Daraus folgt $P(f)(v) = 0$. Aus der Definition von μ_f folgt, dass μ_f ein Teiler von P ist. Daraus folgt $m_i \leq M_{\lambda_i}$. Es folgt $M_{\lambda_i} = m_i$. ■

Korollar 2.3.9 Sei U ein f-invarianter Unterraum. Dann gilt

$$U = (U \cap H_1) \oplus \cdots \oplus (U \cap H_r).$$

Beweis. Da wir eine direkte Summe $H_1 \oplus \cdots \oplus H_r$ haben ist die Summe $(U \cap H_1) + \cdots + (U \cap H_r)$ auch eine direkte Summe. Wir haben eine Enthaltung $(U \cap H_1) \oplus \cdots (U \cap H_r) \subset U$. Umgekehrt, sei $v \in U$, und wie oben sei $P_i = \prod_{j \neq i}(X - \lambda_j)^{m_j}$ für $i \in [1, r]$. Dann sind $P_1, \cdots, P_r$ teilerfremd und es gibt Polynome $Q_1, \cdots, Q_r$ mit $P_1 Q_1 + \cdots + P_r Q_r = 1$. Es gilt also

$$v = v_1 + \cdots + v_r$$

wobei $v_i = P_i(f) Q_i(f)(v) \in H_i$. Da U ein f-invarianter Unterraum ist und $v \in U$ ist, gilt $v_i = P_i(f) Q_i(f)(v) \in U$. Es folgt $v_i \in U \cap H_i$ und $U = (U \cap H_1) \oplus \cdots \oplus (U \cap H_r)$. ■

Korollar 2.3.10 Sei $i \in [1, r]$. Dann gibt es ein $v \in V$ mit

$$(f - \lambda_i \mathrm{Id}_V)^{m_i - 1}(v) \neq 0 \text{ und } (f - \lambda_i \mathrm{Id}_V)^{m_i}(v) \neq 0.$$

Beweis. Es gilt $m_i = M_{\lambda_i}$. Also gilt $\mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i - 1} = E_{m_i - 1}(f, \lambda_i) \subsetneq E_{m_i}(f, \lambda_i) = \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i}$. Sei $v \in \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i} \setminus \mathrm{Ker}(f - \lambda_i \mathrm{Id}_V)^{m_i - 1}$. Dann erfüllt v die obige Eigenschaft. ■

2.4 Jordan-Kette

Definition 2.4.1 Ein System $(v_1, \cdots, v_t)$ von Vektoren heißt **Jordan-Kette (für f zum Eigenwert λ)**, falls für alle $k \in [1, t-1]$ gilt

- $v_1 \neq 0$,
- $(f - \lambda \mathrm{Id}_V)(v_1) = 0$

- $(f - \lambda \mathrm{Id}_V)(v_{k+1}) = v_k$.

Lemma 2.4.2 (i) Es gibt einen Vektor $v \in V$ mit $(f - \lambda_i \mathrm{Id}_V)^{m_i - 1}(v) \neq 0$ und $(f - \lambda_i \mathrm{Id}_V)^{m_i}(v) = 0$.

(ii) Sei $v_k = (f - \lambda_i \mathrm{Id}_V)^{m_i - k}(v)$. Das System $(v_1, \cdots, v_{m_i})$ ist eine Jordan-Kette für f zum Eigenwert λ_i. □

Beweis. (i) Siehe Korollar 2.3.10

(ii) Folgt aus den Definitionen von v und der Jordan-Kette. ■

Lemma 2.4.3 Sei $(v_1, \cdots, v_t)$ eine Jordan-Kette für f zum Eigenwert λ.

(i) Dann ist $((f - \lambda_i \mathrm{Id}_V)(v_2), \cdots, (f - \lambda_i \mathrm{Id}_V)(v_t)) = (v_1, \cdots, v_{t-1})$ eine Jordan-Kette für f zum Eigenwert λ.

(ii) Dann ist $\langle v_1, \cdots, v_t \rangle$ f-invariant und $(v_1, \cdots, v_t)$ linear unabhängig. □

Beweis. (i) Folgt aus der Definition.

(ii) Nach (i) folgt, dass $f - \lambda \mathrm{Id}_V$ die Jordan-Kette auf $(0, v_1, \cdots, v_{t-1})$ schickt. Daraus folgt, dass $\langle v_1, \cdots, v_t \rangle$ $(f - \lambda \mathrm{Id}_V)$-invariant, also f-invariant ist.

Nach Induktion über t. Seien $x_1, \cdots, x_t$ Skalare mit $\sum_i x_i v_i = 0$. Es folgt $0 = \sum_i x_i (f - \lambda \mathrm{Id}_V)(v_i) = \sum_{i \leq r-1} x_{i+1} v_i$. Da $(v_1, \cdots, v_{t-1})$ eine Jordan-Kette ist, ist das System linear unabhängig. Es folgt $x_2 = \cdots = x_r = 0$. Es gilt dann auch $x_1 v_1 = 0$. Da $v_1 \neq 0$ folgt $x_1 = 0$. Das System $(v_1, \cdots, v_t)$ ist linear unabhängig. ■

Korollar 2.4.4 Sei $(v_1, \cdots, v_t)$ eine Jordan-Kette für f zum Eigenwert λ. Sei $U = \langle v_1, \cdots, v_t \rangle$ und sei $\mathcal{B} = (v_1, \cdots, v_t)$.

(i) Das System $\mathcal{B}$ ist eine Basis von U.

(ii) Es gilt

$$\mathrm{Mat}_{\mathcal{B}}(f|_U) = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{pmatrix} := J(\lambda, t).$$

Beweis. (i) Folgt aus dem obigen Lemma.

(ii) Es gilt $(f - \lambda \mathrm{Id}_V)(v_{k+1}) = v_k$ für $k \in [1, t-1]$. Daraus folgt $f(v_{k+1}) = \lambda v_{k+1} + v_k$. Es gilt auch $(f - \lambda \mathrm{Id}_V)(v_1) = 0$, also $f(v_1) = \lambda v_1$. Das Lemma ist bewiesen. ■

Definition 2.4.5 Die Matrix $J(\lambda, t)$ heißt **Jordan-Block der Größe t zum Eigenwert λ**.

2.5 Endomorphismus mit einem Eigenwert

Sei $f \in \mathrm{End}(V)$. In diesem Kapitel nehmen wir an, dass $\chi_f = (X-\lambda)^n$ und $\mu_f = (X-\lambda)^m$. Wie schreiben $E_i = E_i(f, \lambda)$. Es gilt

$$0 \subsetneq E_1 \subsetneq \cdots \subsetneq E_m = V.$$

Wir schreiben $g = f - \lambda \mathrm{Id}_V$.

Lemma 2.5.1 Sei U ein Unterraum mit $E_1 \cap U = 0$. Dann ist $g|_U : U \to V$ injektiv.

Insbesondere gilt: Sei $\mathcal{B}$ eine Basis von U, dann ist $g(\mathcal{B})$ eine Basis von $g(U)$. □

Beweis. Es gilt $\mathrm{Ker}(g|_U) = U \cap \mathrm{Ker} g = U \cap E_1 = 0$. ■

Wir bauen jetzt eine Zerlegung von $V = E_m$ in eine direkte Summe.

Satz 2.5.2 Für alle $i \in [1, m]$, gibt es Unterräume $U_i \subset E_i$ so dass

$$E_i = E_{i-1} \oplus \bigoplus_{j=0}^{m-i} g^j(U_{i+j}).$$

Beweis. Nach absteigender Induktion über $i \in [1, m]$.

Für $i = m$, wählen wir U_m ein Komplement von E_{m-1} in E_m. Es gilt $E_m = E_{m-1} \oplus U_m$.

Induktionsannahme: für $k \in [i+1, m]$ gibt es Unterräume $U_k \subset E_k$ mit:

$$E_{i+1} = E_i \oplus \bigoplus_{j=0}^{m-i-1} g^j(U_{i+j+1}).$$

Lemma 2.5.3 Es gilt $E_{i-1} + \sum_{j=1}^{m-i} g^j(U_{i+j}) = E_{i-1} \oplus \bigoplus_{i=1}^{m_i} g^j(U_{i+j})$. □

Beweis. Sei $v \in E_{i-1}$ und $u_{i,j} \in U_{i+j}$ mit

$$v + \sum_{j=1}^{m-i} g^j(u_{i,j}) = 0.$$

Wir zeigen, dass $v = g^j(u_{i,j}) = 0$ für alle $j \in [1, m-i]$. Es gilt

$$\sum_{j=0}^{m-i-1} g^{i+j}(u_{i,j+1}) = \sum_{j=1}^{m-i} g^{i+j-1}(u_{i,j}) = -g^{i-1}(v) = 0.$$

Daraus folgt

$$\sum_{j=0}^{m-i-1} g^j(u_{i,j+1}) \in E_i.$$

Nach Induktionsannahme gilt $g^j(u_{i,j+1}) = 0$ für alle $j \in [0, m-i-1]$. Es folgt $g^{j+1}(u_{i,j+1}) = 0$ für alle $j \in [0, m-i-1]$ und $g^j(u_{i,j}) = 0$ für alle $j \in [1, m-i]$. Es folgt auch $v = 0$. ∎

Sei U_i ein Komplement von $E_{i-1} \oplus \bigoplus_{i=1}^{m_i} g^j(U_{i+j})$ in E_i. Es gilt

$$E_i = E_{i-1} \oplus \bigoplus_{j=0}^{m-i} g^j(U_{i+j}).$$

Korollar 2.5.4 Seien U_i für $i \in [1, m]$ wie im Satz 2.5.2. Sei $\mathcal{B}_i$ eine Basis von U_i, dann ist $g^j(\mathcal{B}_i)$ eine Basis von $g^j(U_i)$ für alle $j \in [0, i-1]$.

Beweis. Nach Lemma 2.5.1, genügt es zu zeigen, dass $g^{j-1}(U_i) \cap E_1 = 0$ für alle $j \in [0, i-1]$. Sei $v \in g^{j-1}(U_i) \cap E_1$ und sei $u \in U_i$ mit $g^{j-1}(u) = v$. Es gilt $g^j(u) = g(v) = 0$. Daraus folgt $u \in U_i \cap E_j \subset U_i \cap E_{i-1} = 0$. Es folgt $u = 0$ und $v = 0$. ∎

Korollar 2.5.5 Für alle $i \in [1, m]$ gilt

$$\dim E_i = \dim E_{i-1} + \sum_{k=i}^{m} \dim U_k \text{ und } \dim U_i = 2\dim E_i - \dim E_{i+1} - \dim E - i - 1,$$

wobei $E_{m+1} = E_m = V$.

Beweis. Die erste Dimensionsformel folgt aus dem Satz 2.5.2 und dem Korollar 2.5.4. Die zweite Dimensionsformel folgt aus der ersten nach absteigender Induktion über i.

Für $i = m$ gilt $\dim E_m = \dim E_{m-1} + \dim U_m$. Daraus folgt die Dimensionformel. Induktionsannahme: Für $k \in [i+1, m]$ gilt $dim U_k = \dim E_k - \dim E_{k+1} - \dim E_{k-1}$.

Es gilt $\dim E_i = \dim E_{i-1} + \sum_{k=i}^m \dim U_k$. Daraus folgt

$$\begin{aligned}
\dim U_i &= \dim E_i - \dim E_{i-1} - \textstyle\sum_{k=i+1}^m \dim U_k \\
&= \dim E_i - \dim E_{i-1} - 2\textstyle\sum_{k=i+1}^m \dim E_k + \sum_{k=i+1}^m \dim E_{k+1} + \sum_{k=i+1}^m \dim E_{k-1} \\
&= \dim E_i - \dim E_{i-1} - 2\textstyle\sum_{k=i+1}^m \dim E_k + \sum_{k=i+2}^{m+1} \dim E_k + \sum_{k=i}^{m-1} \dim E_k \\
&= \dim E_i - \dim E_{i-1} - \dim E_{i+1} - \dim E_m + \dim E_{m+1} + \dim E_i \\
&= 2\dim E_i - \dim E_{i-1} - \dim E_{i+1}.
\end{aligned}$$

Korollar 2.5.6 Es gilt

$$V = \bigoplus_{i=1}^{m} \left(\bigoplus_{j=0}^{i-1} g^j(U_i) \right).$$

Korollar 2.5.7 Seien U_i für $i \in [1, m]$ wie im Satz 2.5.2 und seien $\mathcal{B}_i$ Basen von U_i. Dann ist

$$\mathcal{B} = \bigcup_{i=1}^{m} \left(\bigcup_{j=1}^{i-1} g^j(\mathcal{B}_i) \right)$$

eine Basis von V.

Definition 2.5.8 Sei $m \in \mathbb{N}$. Für $i \in [1, m]$, sei $n_i, d_i \in \mathbb{N}$ und $A_i \in M_{n_i}(K)$. Wir schreiben $\mathrm{diag}(d_1 A_1, \cdots, d_m A_m)$ für die blockdiagonale Matrix mit d_1-Mal A_1, $\cdots$, d_m-Mal A_m auf der Diagonale.

Korollar 2.5.9 Es gilt $\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{diag}(d_1 J(\lambda, 1), \cdots, d_m J(\lambda, m))$ wobei $d_i = \dim U_i$.

Beweis. Für $v \in U_i$, ist $(g^{i-1}(v), \cdots, g(v), v)$ eine Jordan-Kette für f zum Eigenwert λ. Die Basis $\mathcal{B}$ ist also eine Vereinigung von Jordan-Ketten und die obere Diagonalform der Matrix folgt daraus. ■

2.6 Jordansche Normalform

Satz 2.6.1 (Jordansche Normalform) Sei $f \in \mathrm{End}(V)$, so dass χ_f (oder μ_f) in Linearfaktoren zerfällt. Dann gibt es eine Basis $\mathcal{B}$ von V so, dass

$$\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{diag}(J(\zeta_1, n_1), \cdots, J(\zeta_s, n_s))$$

, wobei die Matrizen $J(\lambda_i, n_i)$ sind, bis auf Vertauschen, eindeutig bestimmt sind. Diese Matrix heißt **Jordan-Normalform** von f. Die $\zeta_1, \cdots, \zeta_s$ sind nicht notwendig paarweise verschieden. □

Beweis. Ist $\mathcal{B}$ eine Basis mit $\mathrm{Mat}_{\mathcal{B}}(f)$ wie oben, so sagen wir, dass $\mathrm{Mat}_{\mathcal{B}}(f)$ in Jordan-Normalform ist. Wir zeigen zuerst, dass es eine solche Basis gibt.

Für die Einschränkung f_i von f auf $H(f, \lambda_i)$ gilt $(f_i - \lambda_i \mathrm{Id})^{m_i} = 0$. Nach Korollar 2.5.9 gibt es eine Basis $\mathcal{B}_i$ von $H(f, \lambda_i)$ so, dass $\mathrm{Mat}_{\mathcal{B}_i}(f_i)$ in Jordan-Normalform ist. Sei $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$. Nach dem Satz 2.3.7, ist $\mathcal{B}$ eine Basis von V und es gilt

$$\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{diag}(\mathrm{Mat}_{\mathcal{B}_1}(f_1), \cdots, \mathrm{Mat}_{\mathcal{B}_r}(f_r)).$$

Es folgt, dass $\mathrm{Mat}_{\mathcal{B}}(f)$ in Jordan-Normalform ist.

Wir zeigen jetzt, dass die Jordan-Blöcke, bis auf Vertauschen, eindeutig bestimmt sind. Sei $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_s$ eine Basis mit $\mathcal{B}_i = (v_{i,1}, \cdots, v_{i,n_i})$ so, dass $\mathrm{Mat}_{\mathcal{B}}(f)$ in Jordan-Normalform ist.

Sei $J(\zeta_i, n_i)$ ein Jordan-Block von $\mathrm{Mat}_{\mathcal{B}}(f)$. Dann ist $v_{i,1}$ ein Eigenvektor für den Eigenwert ζ_i. Es folgt, dass die Skalare $\zeta_1, \cdots, \zeta_s$ die Eigenwerte von f sind und also eindeutig bestimmt. Außerdem, gilt $v_{i,j} \in E_j(f, \zeta_i) \setminus E_{j-1}(f, \zeta_i)$ für $j \in [1, n_i]$. Sei $e_j(\lambda) = \dim E_j(f, \lambda)$. Es folgt

$$d_j(\lambda) = e_j(\lambda) - e_{j-1}(\lambda) = \text{Anzahl der Elemente von } \{i \mid \zeta_i = \lambda \text{ und } n_i \leq j\}.$$

Sei $j_t(\lambda)$ die Anzahl von Jordan-Blöcken der Gestalt $J(\lambda, t)$ in alle Jordan-Blöcken $J(\zeta_1, n_1), \cdots, J(\zeta_s, n_s)$. Es folgt

$$d_j(\lambda) = \sum_{t \geq j} j_t(\lambda).$$

Daraus folgt

$$j_t(\lambda) = d_t(\lambda) - d_{t-1}(\lambda) = e_t(\lambda) + e_{t-2}(\lambda) - 2e_{t-1}(\lambda).$$

Es folgt, dass $j_t(\lambda)$ nur von f abhängt und dass die Jordan-Blöcke, bis auf Vertauschen, eindeutig bestimmt sind. ∎

Sei $e_t(f, \lambda) = \dim E_t(f, \lambda)$ und sei $j_t(f, \lambda)$ die Anzahl von Jordan-Blöcke der Größe t zum Eingenwert λ.

Korollar 2.6.2 Sei $f \in \mathrm{End}(V)$ und J eine jordansche Normalform für f.

(i) J hat $\dim E(f, \lambda) = e_1(f, \lambda)$ Jordan-Blöcke zum Eigenwert λ.

(ii) J hat $j_t(f, \lambda) = 2e_t(f, \lambda) - e_{t+1}(f, \lambda) - e_{t-1}(f, \lambda)$ Jordan-Blöcke $J(\lambda, t)$.

(iii) Es gilt
$$e_i(\lambda) = \sum_{t \geq 1} \min(i, t) j_t(f, \lambda).$$

Beweis. (i) folgt aus (iii) für $i = 1$.

Wir haben im Beweis des obigen Satzes gezeigt, dass

$$e_i(\lambda) - e_{i-1}(\lambda) = d_i(\lambda) = \sum_{t \geq i} j_t(\lambda).$$

Es folgt

$$j_t(f, \lambda) = d_t(\lambda) - d_{t+1}(\lambda) = e_t(f, \lambda) - e_{t-1}(f, \lambda) - (e_{t+1}(f, \lambda) - e_t(f, \lambda))$$

und (ii) folgt.

Es gilt auch

$$e_i(f,\lambda) = \sum_{k=1}^{i} d_k(\lambda) = \sum_{k=1}^{i}\sum_{t\geq k} j_t(\lambda) = \sum_{t\geq 1} j_t(\lambda) \sum_{k\leq i,t} 1 = \sum_{t\geq 1} \min(i,t) j_t(f,\lambda).$$

Definition 2.6.3 Das Spektrum $\Sigma(f)$ eines Endomorphismus f (bzw. $\Sigma(A)$ einer Matrix A) ist die Menge aller Eigenwerte von f (bzw. von A).

Korollar 2.6.4 Zwei Matrizen $A, B \in M_n(K)$, so dass χ_A und χ_B in Linearfaktoren zerfallen, sind genau dann ähnlich, wenn $\dim E_i(A,\lambda) = \dim E_i(B,\lambda)$ für alle $\lambda \in K$.

Beweis. Nach dem Satz sind A und B genau dann ähnlich, wenn A und B die selben Jordan-Blöcke haben. Nach dem obigen Korollar ist dies äquivalent zu $j_t(A,\lambda) = j_t(B,\lambda)$ für alle $\lambda \in K$ und alle $t \in \mathbb{N}$. Nach dem obigen Korollar gilt

$$j_t(f,\lambda) = 2\dim E_t(f,\lambda) - (\dim E_{t-1}(f,\lambda) + \dim E_{t+1}(f,\lambda)$$

und (nach Induktion) gilt auch

$$\dim E_i(f,\lambda) = \sum_{t\geq 1} \min(i,t) j_t(f,\lambda).$$

Es folgt, dass A und B genau dann ähnlich sind, wenn $\dim E_i(A,\lambda) = \dim E_i(B,\lambda)$ für alle $\lambda \in K$. ∎

Beispiel 2.6.5 Seien

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Diese Matrizen sind in Jordan-Normalform. Die Jordan-Blöcke für A sind

$$J(1,1); J(1,1); J(1,2).$$

Die Jordan-Blöcke für B sind

$$J(1,2); J(1,2).$$

Es folgt, dass $A \not\approx B$ (dies ist eine Lösung für die Übung nach dem Beispiel 1.10.8). Wir können die Dimension aller erweiterten Eigenräume Bestimmen. Für $\lambda \neq 1$ gilt $\dim E_i(A,\lambda) = \dim E_i(B,\lambda) = 0$ für alle i. Es gilt

$$\dim E_0(A,1) = 0, \dim E_1(A,1) = 3 \text{ und } \dim E_i(A,\lambda) = 4 \text{ für alle } i \geq 2.$$

$$\dim E_0(B,1) = 0, \dim E_1(B,1) = 2 \text{ und } \dim E_i(B,\lambda) = 4 \text{ für alle } i \geq 2.$$

3 Symmetrische Gruppe

3.1 Definition

Sei $n \in \mathbb{N}$, mit $n \geq 1$ und sei $I_n = [1, n]$.

Definition 3.1.1 Die symmetrische Gruppe S_n ist die Gruppe $(\mathrm{Bij}(I_n), \circ)$, wobei $\mathrm{Bij}(I_n)$ die Menge aller bijektiven Abbildungen $I_n \to I_n$ ist und die Verknupfung $\circ$ die Komposition ist. Ein Element von S_n heißt **Permutation**.

Notation 3.1.2 Für $\sigma : I_n \to I_n$ ein Element in S_n, wir schreiben

$$\sigma = (\sigma(1), \cdots, \sigma(n)).$$

Beispiel 3.1.3 (i) Die Gruppe S_1. Die Menge I_1 hat nur ein Element: 1. Es gibt also nur eine Abbildung $I_1 \to I_1$: die Identität-Abbildung. Es ist eine Bijektive Abbildung. Es gilt

$$S_1 = \{\mathrm{Id}_{I_1}\}.$$

(ii) Die Gruppe S_2. Die Menge I_2 hat zwei Elemente: 1 und 2. Es gibt zwie Bijektion $I_2 \to I_2$: die Identität-Abbildung und $\tau_{1,2}$ die Abbildung definiert durch $\tau_{1,2}(1) = 2$ und $\tau_{1,2}(2) = 1$. Es gilt

$$S_2 = \{\mathrm{Id}_{I_2}, \tau_{1,2}\} = \{(1,2), (2,1)\}.$$

(iii) Die Gruppe S_3 hat 6 Elemente:

$$S_3 = \{(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)\}.$$

Lemma 3.1.4 Die Gruppe S_3 is nicht abelsch. □

Beweis. Es gilt $(2,1,3) \circ (2,3,1) = (1,3,2)$ und $(2,3,1) \circ (2,1,3) = (3,2,1)$. Es folgt $(2,1,3) \circ (2,3,1) \neq (2,3,1) \circ (2,1,3)$ und S_3 ist nicht abelsch. ■

Lemma 3.1.5 Die Abbildung $\iota_{n+1} : S_n \to S_{n+1}$ definiert durch

$$\iota_{n+1}(\sigma)(i) = \begin{cases} \sigma(i) & \text{für } i \in [1, n] \\ n+1 & \text{für } i = n+1, \end{cases}$$

ist ein injektiv Gruppenhomomorphismus. Das bild ist die Untergruppe

$$\iota_{n+1}(S_n) = \{\sigma \in S_{n+1} \mid \sigma(n+1) = n+1\} = S_{n+1}(n+1).$$

Beweis. Die Abbildung ist injektiv: seien $\sigma, \tau \in S_n$ mit $\iota_{n+1}(\sigma) = \iota_{n+1}(\tau)$. Dann gilt für alle $i \in [1, n]$, dass $\sigma(i) = \iota_{n+1}(\sigma)(i) = \iota_{n+1}(\tau)(i) = \tau(i)$. Es folgt $\sigma = \tau$.

Seien $\sigma, \tau \in S_n$. Es gilt

$$\iota_{n+1}(\sigma \circ \tau)(i) = \begin{cases} \sigma \circ \tau(i) & \text{für } i \in [1, n] \\ n+1 & \text{für } i = n+1, \end{cases}$$

Es gilt auch

$$\iota_{n+1}(\sigma) \circ \iota_{n+1}(\tau)(i) = \begin{cases} \iota_{n+1}(\sigma)(\tau(i)) & \text{für } i \in [1, n] \\ \iota_{n+1}(\sigma)(n+1) & \text{für } i = n+1, \end{cases} = \begin{cases} \sigma(\tau(i)) & \text{für } i \in [1, n] \\ n+1 & \text{für } i = n+1, \end{cases}$$

Daraus folgt $\iota_{n+1}(\sigma \circ \tau) = \iota_{n+1}(\sigma) \circ \iota_{n+1}(\tau)$ und ι_{n+1} is ein Gruppenhomomorphismus.

Das bild ist enthalten in $S_{n+1}(n)$. Sei $\sigma \in S_{n+1}(n+1)$. Dann gilt $\sigma(I_n) \subset I_n$ und $\sigma|_{I_n} \in S_n$. Es gilt $\iota_{n+1}(\sigma|_{I_n}) = \sigma$. ■

Korollar 3.1.6 Die Gruppe S_n für $n \geq 3$ ist nicht abelsch.

Beweis. Nach Induktion nach n. Für $n = 3$, es ist Lemma 3.1.4. Angenommen S_n ist nicht abelsch. Es gibt also Elemente $\sigma, \tau \in S_n$ mit $\sigma \circ \tau \neq \tau \circ \sigma$. Wir betrachten $\iota_{n+1}(\sigma), \iota_{n+1}(\tau \in S_{n+1}$. Da ι_{n+1} injektiv und ein Gruppenhomomorphismus ist, gilt $\iota_{n+1}(\sigma) \circ \iota_{n+1}(\tau) \neq \iota_{n+1}(\tau) \circ \iota_{n+1}(\sigma)$. Es folgt S_{n+1} ist nicht abelsch. ■

3.2 Transpositionen

Definition 3.2.1 Seien $i, j \in [1, n]$ mit $i \neq j$. **Die Transposition** $\tau_{i,j}$ ist die Permutation definiert durch

$$\tau_{i,j}(k) = \begin{cases} j & \text{für } k = i \\ i & \text{für } k = j, \\ k & \text{sonst.} \end{cases}$$

Bemerkung 3.2.2 Es gilt $\tau_{i,j}^2 = \mathrm{Id}_{i_n}$ oder $\tau_{k.n}^{-1} = \tau_{k,n}$.

Lemma 3.2.3 Jedes $\sigma \in S_n$ ist ein Produkt von $r \leq n-1$ Transpositionen. □

Beweis. Nach Induktion nach n. Klar für $n = 1$ und $n = 2$. Sei $\sigma \in S_{n+1}$ und sei $i = \sigma(n+1)$. Sei $\tau = \tau_{i,n+1} \circ \sigma$. Es gilt $\tau(n+1) = \tau_{i,n+1}(i) = n+1$. Es gilt $\tau \in S_n(n+1)$ und nach Induktion ist τ ein Produkt von $r \leq n-1$ Transpositionen. Es gilt $\sigma = \tau_{i,n+1} \circ \tau$. Daraus folgt, dass σ ein Produkt von $r+1 \leq n$ Transpositionen. ■

Lemma 3.2.4 Sei G eine Gruppe und sei $g \in G$. Sei $\mathrm{Int}_g : G \to G$ definiert durch $\mathrm{Int}_g(h) = ghg^{-1}$ für alle $h \in G$. Dann ist Int_g ein Gruppenautomorphismus von G und es gilt $\mathrm{Int}_g^{-1} = \mathrm{Int}_{g^{-1}}$. □

Beweis. Es gilt $\mathrm{Int}_g(h)\mathrm{Int}_g(k) = ghg^{-1}gkg^{-1} = ghkg^{-1} = \mathrm{Int}_g(hk)$. Daraus folgt, dass Int_g ein Gruppenhomomorphismus ist. Es gilt $\mathrm{Int}_g(\mathrm{Int}_{g^{-1}}(h) = g(g^{-1}hg)g^{-1} = h$ und $\mathrm{Int}_{g^{-1}}(\mathrm{Int}_g(h) = g^{-1}(ghg^{-1})g = h$. Daraus folgt, dass $\mathrm{Int}_g^{-1} = \mathrm{Int}_{g^{-1}}$ und Int_g ist ein Gruppenautomorphismus. ∎

Korollar 3.2.5 Sei $k \in [1, n+1]$ und sei $\iota_k : S_n \to S_{n+1}$ definiert durch

$$\iota_k(\sigma) = \tau_{k,n+1} \circ \iota_{n+1}(\sigma) \circ \tau_{i,n+1}^{-1}.$$

Dann ist ι_k injektiv und ein Gruppenhomomorphismus. Das bild von ι_k ist die Untergruppe

$$\iota_k(S_n) = \{\sigma \in S_{n+1} \mid \sigma(k) = k\} = S_{n+1}(k).$$

Beweis. Es gilt $\iota_k = \mathrm{Int}_{\tau_{k,n+1}} \circ \iota_{n+1}$. Es folgt, dass ι_k injektiv und ein Gruppenhomomorphismus ist. Das bild von ι_k ist

$$\iota_k(S_n) = \mathrm{Int}_{\tau_{k,n+1}}(\iota_{n+1}(S_n)) = \mathrm{Int}_{\tau_{k,n+1}}(\{\sigma \in S_{n+1} \mid \sigma(n+1) = n+1\}).$$

Wir zeigen $\mathrm{Int}_{\tau_{k,n+1}}(\{\sigma \in S_{n+1} \mid \sigma(n+1) = n+1\}) = \{\sigma \in S_{n+1} \mid \sigma(k) = k\}$. Sei $\sigma \in S_{n+1}$ mit $\sigma(n+1) = n+1$. Es gilt $\mathrm{Int}_{\tau_{k,n+1}}(\sigma)(k) = \tau_{k,n+1}\sigma\tau_{k,n+1}(k) = \tau_{k,n+1}\sigma(n+1) = \tau_{k,n+1}(n+1) = k$. Es folgt $\mathrm{Int}_{\tau_{k,n+1}}(\sigma) \in S_{n+1}(k)$. Sei $\sigma \in S_{n+1}$ mit $\sigma(k) = k$. Es gilt $\sigma = \mathrm{Int}_{\tau_{k,n+1}}\mathrm{Int}_{\tau_{k,n+1}^{-1}}(\sigma)$. Sei $\tau = \mathrm{Int}_{\tau_{k,n+1}^{-1}}(\sigma) = \mathrm{Int}_{\tau_{k,n+1}}(\sigma)$. Es gilt $\tau(n+1) = \tau_{k,n+1}\sigma\tau_{k,n+1}(n+1) = \tau_{k,n+1}\sigma(k) = \tau_{k,n+1}(k) = n+1$. Es folgt $\tau = \mathrm{Int}_{\tau_{k,n+1}}(\sigma) \in S_{n+1}(n+1)$ und $\sigma = \mathrm{Int}_{\tau_{k,n+1}}(\tau)$. ∎

Lemma 3.2.6 Sei $S_{n+1}^i = \{\sigma \in S_{n+1} \mid \sigma(n+1) = i\}$. Dann ist die Abbildung $S_{n+1}^i \to S_{n+1}(n+1)$ definiert durch $\sigma \mapsto \tau_{i,n+1} \circ \sigma$ eine Bijektion. □

Beweis. Wir zeigen, dass diese Abbilfung wohl definiert ist *i.e.*, dass für $\sigma \in S_{n+1}^i$ gilt $\tau_{i,n+1} \circ \sigma \in S_{n+1}(n+1)$. Es gilt

$$\tau_{i,n+1}\sigma(n+1) = \tau_{,n+1}(i) = n+1.$$

Umgekehrt, wir betrachten die Abbildung $S_{n+1}(n+1) \to S_{n+1}^i$ definiert durch $\sigma \mapsto \tau_{i,n+1} \circ \sigma$. Diese Abbildung ist wohl definiert: für $\sigma \in S_{n+1}(n+1)$ gilt

$$\tau_{i,n+1} \circ \sigma(n+1) = \tau_{i,n+1}(n+1) = i.$$

Diese Abbildungen sind Inverse Abbildungen. Es folgt das Lemma. ∎

Korollar 3.2.7 Die Gruppe S_n hat $n!$ Elemente.

Beweis. Nach Induktion nach n. Für $n = 1$, es ist wahr. Angenommen S_n hat $n!$ Elemente. Die Gruppe S_{n+1} ist die disjunkte Vereinnigung

$$S_{n+1} = \coprod_{i=1}^{n+1} S_{n+1}^i.$$

Nach dem Lemma, folgt, dass S_{n+1}^i genau so viele Elemente als S_n hat. Nach Induktion, hat S_{n+1}, für alle $i \in [1, n+1]$, genau $n!$ Elemente. Es folgt, dass S_{n+1} genau $n \cdot n! = (n+1)!$ Elemente hat. ∎

3.3 Support

Definition 3.3.1 **Der Support** einer Permutation $\sigma \in S_n$ ist die Teilmenge $\mathrm{Supp}(\sigma) \subset I_n$ definiert durch

$$\mathrm{Supp}(\sigma) = \{i \in I_n \mid \sigma(i) \neq i\}.$$

Lemma 3.3.2 Seien $\sigma, \tau \in S_n$ mit $\mathrm{Supp}(\sigma) \cap \mathrm{Supp}(\tau) = \emptyset$. Dann gilt $\sigma \circ \tau = \tau \circ \sigma$. □

Beweis. Übung. ■

3.4 Permutationsmatrix

Definition 3.4.1 Sei $\sigma \in S_n$. Das zugehörige **Permutationsendomoprhismus** $f_\sigma \in \mathrm{End}(k^n)$ und die zugehörige **Permutationsmatrix** P_σ sind definiert durch

$$f_\sigma(e_i) = e_{\sigma(i)} \text{ und } P_\sigma = \mathrm{Mat}_{\mathcal{B}}(f_\sigma),$$

wobei $\mathcal{B} = (e_1, \cdots, e_n)$ die kanonische basis des k^n ist.

Beispiel 3.4.2 Sei $\sigma = (2, 3, 1) \in S_3$. Dann ist

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Lemma 3.4.3 Seien $\sigma, \tau \in S_n$. Dann gilt $f_\sigma \circ f_\tau = f_{\sigma\circ\tau}$ und $P_\sigma \circ P_\tau = P_{\sigma\circ\tau}$. □

Beweis. Die Gleichung $P_\sigma \circ P_\tau = P_{\sigma\circ\tau}$ folgt aus $f_\sigma \circ f_\tau = f_{\sigma\circ\tau}$. Da $\mathcal{B}$ eine Basis ist, genügt es zu zeigen, dass $f_\sigma \circ f_\tau(e_i) = f_{\sigma\circ\tau}(e_i)$. Es gilt

$$f_\sigma \circ f_\tau(e_i) = f_\sigma(e_{\tau(i)}) = e_{\sigma(\tau(i))} = f_{\sigma\circ\tau}(e_i).$$

Korollar 3.4.4 Sei $\sigma \in S_n$.

(i) Die Matrix P_σ ist invertierbar.

(ii) Die Abbildung $S_n \to \mathrm{GL}_n(K)$ ist ein Gruppenhomomorphismus.

Beweis. (i) Es gilt $P_\sigma \circ P_{\sigma^{-1}} = P_{\mathrm{Id}} = I_n = P_{\sigma^{-1}} \circ P_\sigma$. Daraus folgt, dass P_σ invertierbar mit Inverse $P_{\sigma^{-1}}$ ist.

(ii) Folgt aus dem Lemma. ■

Korollar 3.4.5 Die Abbildung $\varepsilon : S_n \to K \setminus \{0\}$ definiert durch $\sigma \mapsto \det(P_\sigma)$ ist ein Gruppenhomomorphismus.

Beweis. Wir wissen, dass det : $\mathrm{GL}_n(K) \to K \setminus \{0\}$ eine Gruppenhomomorphismus ist. Daraus folgt, dass die obige Komposition auch ein Gruppenhomomorphismus ist. ∎

Lemma 3.4.6 Sei $\tau_{i.j}$ eine Transposition, es gilt $\varepsilon(\tau_{i,j}) = -1$. □

Beweis. Man sieht einfach, dass $P_{\tau_{i,j}} = E_{i,j}^{(n)}$ wobei $E_{i,j}^{(n)}$ die zugehörige Elementarmatrix ist. Daraus folgt das Lemma. ∎

Korollar 3.4.7 Die Abbilgung $\varepsilon : S_n \to \{1, -1\}$ definiert durch $\varepsilon(\sigma) = \det(P_\sigma)$ ist ein Gruppenhomomorphismus.

Beweis. Es bleibt nur zu zeigen, dass $det(P_\sigma) \in \{-1, -1\}$. Nach Lemma 3.2.3, gilt, dass σ ein Produkt $\tau_1 \cdots \tau_k$ von Transpositionen ist. Es folgt $det(P_\sigma) = (-1)^k \in \{-1, -1\}$. ∎

Definition 3.4.8 Eine Permutation $\sigma \in S_n$ heißt **gerade** falls $\varepsilon(\sigma) = 1$ und **ungerade** falls $\varepsilon(\sigma) = -1$.

3.5 Elementare Transpositionen

Definition 3.5.1 Sei $i \in [1, n-1]$. **Die elementare Transposition** s_i ist die Transposition $\tau_{i,i+1}$.

Lemma 3.5.2 Seien $i, j \in [1, n]$ mit $i < j$. Es gilt

$$\tau_{i,j} = s_i \cdots s_{j-2} s_{j-1} s_{j-2} \cdots s_i.$$

Insbesondere ist $\tau_{i,j}$ ein Produkt von $2(j-i) - 1$ elementare Transpositionen. □

Beweis. Nach Induktion nach $j-i$. Für $j-i = 1$ gilt $j = i+1$ und $\tau_{i,j} = s_i$. Angenommen $\tau_{i,j} = s_i \cdots s_{j-2} s_{j-1} s_{j-2} \cdots s_i$, wir zeigen $\tau_{i-1,j} = s_{i-1} \cdots s_{j-2} s_{j-1} s_{j-2} \cdots s_{i-1}$. Es gilt

$$s_{i-1} \tau_{i,j} s_{i-1} = \tau_{i-1,j}.$$

Das Lemma folgt nach Induktionsannahme. ∎

Satz 3.5.3 Jedes $\sigma \in S_n$ ist ein Produkt von $R \leq \frac{n(n-1)}{2}$ elementare Transpositionen. □

Beweis. Nach Induktion nach n. Für $n = 1$ oder $n = 2$ ist es wahr. Wir nehmen an, dass das Lemma für S_n wahr ist. Sei $\sigma \in S_{n+1}$ und sei $i = \sigma(n+1)$. Sei $\tau = s_n \cdots s_i \sigma$. Es gilt $\tau(n+1) = n+1$. Es folgt, dass $\tau \in S_n$ und nach Induktion, gibt es $R \leq \frac{n(n-1)}{2}$ elementare Tranpositionen $s_{i_1}, \cdots, s_{i_R}$ mit $\tau = s_{i_1} \cdots s_{i_R}$. Es folgt, dass σ ein Produkt von weniger als

$$\frac{n(n-1)}{2} + n - i + 1 \leq \frac{n(n-1)}{2} + n = \frac{n(n+1)}{2}$$

elementare Transpositionen ist. ∎

Lemma 3.5.4 Es gilt

(i) $s_i^2 = \mathrm{Id}$, für alle $i \in [1, n-1]$.

(ii) $(s_i s_{i+1} s_i)^2 = \mathrm{Id}$, für alle $i \in [1, n-2]$.

(iii) $(s_i s_j)^2 = \mathrm{Id}$, für alle $i, j \in [1, n-1]$ mit $|i - j| > 1$. □

Beweis. Übung. ■

Satz 3.5.5 Sei $\sigma \in S_n$ und sei

$$I(\sigma) = \{(i,j) \in [1,n] \times [1,n] \mid i < j \text{ und } \sigma(i) > \sigma(j)\}$$

und $\ell(\sigma) = |I(\sigma)|$. Dann gilt $\varepsilon(\sigma) = (-1)^{\ell(\sigma)}$. □

Beweis. Siehe Übungsblatt 5. ■

3.6 Determinante

Satz 3.6.1 Sei $A \in M_n(K)$ eine Matrix. Dann gilt

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}.$$

Beweis. Die zweite Formel folgt aus der erste und $\det(A^T) = \det(A)$.

Für die erste Formel zeigen wir, dass die Abbildung

$$A \mapsto D(A) = \sum_{\sigma \in S_n} \varepsilon \prod_{i=1}^{n} a_{i,\sigma(i)}$$

linear in den Zeilen ist, dass $D(A) = 0$ für $\mathrm{Rg}(A) < n$ und, dass $D(I_n) = 1$.

Wir schreiben $I_n = (\delta_{i,j})$. Es gilt

$$D(I_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} \delta_{i,\sigma(i)}.$$

Es gilt $\delta_{i,\sigma(i)} \neq 0$ genau dann, wenn $i = \sigma(i)$. Daraus folgt, dass gilt $\prod_{i=1}^{n} \delta_{i,\sigma(i)} \neq 0$ genau dann, wenn $i = \sigma(i)$ für alle $i \in [1, n]$ *i.e.* $\prod_{i=1}^{n} \delta_{i,\sigma(i)} \neq 0$ genau dann, wenn $\sigma = \mathrm{Id}$. Daraus folgt

$$D(I_n) = 1.$$

Sei $A \in M_n(K)$ und seien $Z_1, \cdots, Z_k, \cdots, Z_n$ die Zeilen von A. Sei B mit Zeilen $Z_1, \cdots, Z_k + Z'_k, \cdots, Z_n$ und C mit Zeilen $Z_1, \cdots, Z'_k, \cdots, Z_n$. Wir schreiben $Z_i =$

$(a_{i,1}, \cdots, a_{i,n})$ und $Z'_k = (a'_{k,1}, \cdots, a'_{i,n})$. Es gilt also $A = (a_{i,j})$, $B = (b_{i,j})$ und $c_{i,j})$ wobei

$$b_{i,j} = \begin{cases} a_{i,j} & \text{für } i \neq k \\ a_{k,j} + a'_{k,j} & \text{für } i = k. \end{cases} \quad \text{und } c_{i,j} = \begin{cases} a_{i,j} & \text{für } i \neq k \\ a'_{k,j} & \text{für } i = k. \end{cases}$$

Es gilt

$$\prod_{i=1}^{n} b_{i,\sigma(i)} = \prod_{i=1}^{n} a_{i,\sigma(i)} + a'_{k,\sigma(k)} \prod_{i=1,\ i\neq k}^{n} a_{i,\sigma(i)} = \prod_{i=1}^{n} a_{i,\sigma(i)} + \prod_{i=1}^{n} c_{i,\sigma(i)}.$$

Daraus folgt $D(B) = D(A) + D(C)$ und D inst linear in den Zeilen.

Sei A mit $\mathrm{Rg}(A) < n$. Es gibt eine Zeile Z_k mit $Z_k = \sum_{t=1,\ t\neq k}^{n} x_t Z_t$. Sei A_t die Matrix mit Zeilen $(Z_1, \cdots, Z_{k-1}, Z_t, Z_{k+1}, \cdots, Z_n)$. Nach Linearität gilt

$$D(A) = \sum_{t=1,\ t\neq k} x_t D(A_t).$$

Es genügt zu zeigen, dass $D(A_t) = 0$. Wir schreiben $A_t = (b_{i,j})$. Es gilt

$$D(A_t) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} b_{i,\sigma(i)}.$$

Da die t-te und die k-te Zeilen von A_t gleich sind gilt $b_{t,j} = b_{k,j}$ für alle j. Sei $\tau = \tau_{j,k}$, es gilt

$$\prod_{i=1}^{n} b_{i,\sigma\tau(i)} = b_{t,\sigma(k)} b_{k,\sigma(t)} \prod_{i\neq t,k}^{n} b_{i,\sigma(i)} = b_{k,\sigma(k)} b_{t,\sigma(t)} \prod_{i\neq t,k}^{n} b_{i,\sigma(i)} = \prod_{i=1}^{n} b_{i,\sigma(i)}.$$

Es folgt

$$D(A_t) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} b_{i,\sigma(i)} = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} b_{i,\sigma\tau(i)}.$$

Wir setzen $\theta = \sigma\tau$ *i.e.* $\sigma = \theta\tau^{-1} = \theta\tau$. Es gilt

$$D(A_t) = \sum_{\theta\in S_n} \varepsilon(\theta\tau) \prod_{i=1}^{n} b_{i,\theta(i)} = -\sum_{\theta\in S_n} \varepsilon(\theta) \prod_{i=1}^{n} b_{i,\theta(i)} = -D(A_t).$$

Es folgt $D(A_t) = 0$. ∎