LINEARE ALGEBRA FÜR INFORMATIKER

OLAF M. SCHNÜRER

Inhaltsverzeichnis

Datum: 30. November 2020.

Skript zur Vorlesung *Lineare Algebra für Informatiker* im Sommersemester 2020 an der Universität Paderborn. Ich danke Jürgen Klüners für die Vorlage. Fabian Schneider, Tobias Teumert, Paul Schübeler, Vincent Griebel, Heike Herr, Jessica Müller, Philip Coutinho de Sousa, Christian Günther, Steffen Sassalla, Linus Jungemann und Kevin Andrew Baier haben mich dankenswerterweise auf Fehler oder Unklarheiten im Skript aufmerksam gemacht. (Falls jemand hier nicht namentlich erscheinen möchte, bitte melden.)

1. Gruppen

1.1. Grundlegendes zu Gruppen.

1.1.1. Wir wiederholen einiges zu Gruppen aus der Analysis-Vorlesung [Sch20].

Definition 1.1.2. Eine **Gruppe** $(G, *)$ ist eine Menge G zusammen mit einer Verknüpfung $*$, also einer Abbildung

$$*\colon G \times G \to G,$$

$$(g, h) \mapsto g * h,$$

so dass die folgenden Bedingungen erfüllt sind:

(G1) (Assoziativität) Für alle $g, h, l \in G$ gilt

$$g * (h * l) = (g * h) * l.$$

(G2) (Existenz eines neutralen Elements) Es existiert ein Element $e = e_G \in G$ mit

$$e * g = g = g * e$$

für alle $g \in G$. Ein solches Element ist eindeutig[1] und heißt das **neutrale Element von** G.

(G3) (Existenz von Inversen) Zu jedem $g \in G$ existiert ein $h \in G$ mit

$$h * g = e = g * h.$$

Ein solches Element ist eindeutig[2], wird als g^{-1} notiert und heißt das **Inverse von** g.

Eine Gruppe $(G, *)$ heißt genau dann **kommutativ** (oder **abelsch**), wenn

$$g * h = h * g$$

für alle $g, h \in G$ gilt.

Beispiel 1.1.3. (a) $(\mathbb{Z}, +)$ ist eine abelsche Gruppe.

(b) $(\mathbb{N}, +)$ ist keine Gruppe, da kein Inverses zu n für $n > 0$ existiert. (Wir erinnern an unsere Konvention $\mathbb{N} = \{0, 1, 2, 3, \dots\}$; manche Leute schreiben $\mathbb{N}_0$ dafür.)

(c) Sei K ein Körper (wie in der Analysis definiert, etwa $\mathbb{Q}$ oder $\mathbb{R}$ oder $\mathbb{C}$) und $K^\times := K \setminus \{0\}$. Dann sind $(K, +)$ und $(K^\times, \cdot)$ kommutative Gruppen.

(d) Sei M eine Menge. Betrachte auf der Menge

$$\mathrm{Sym}(M) := \{\sigma : M \to M \mid \sigma \text{ bijektiv}\}$$

die Verknüpfung Komposition (= Hintereinanderausführung) $\circ$ von Abbildungen: Gegeben $\sigma, \tau \in \mathrm{Sym}(M)$ ist $\sigma \circ \tau$ durch

$$(\sigma \circ \tau)(m) := \sigma(\tau(m))$$

für beliebiges $m \in M$ definiert. Dann ist $(\mathrm{Sym}(M), \circ)$ eine Gruppe mit neutralem Element id_M (warum?). Sie heißt **symmetrische Gruppe von** M. Für $|M| \geq 3$ ist sie nicht kommutativ (warum?).

(e) Die Menge der Drehungen der Ebene $\mathbb{R}^2$ um den Ursprung $0 \in \mathbb{R}^2$ (im Uhrzeigersinn) mit der Komposition als Verknüpfung ist eine kommutative Gruppe.

Lemma 1.1.4. *Sei $(G, *)$ eine Gruppe. Dann gelten*

$$(g * h)^{-1} = h^{-1} * g^{-1} \qquad \text{für alle } g, h \in G,$$
$$(g^{-1})^{-1} = g \qquad \text{für alle } g \in G,$$
$$e^{-1} = e.$$

[1] Sei $e' \in G$ ein weiteres Element mit $e' * g = g * e' = g$ für alle $g \in G$, so liefert dies für $g = e$ die Gleichheiten $e = e * e' = e'$.

[2] Sei h' ein weiteres Element mit $h' * g = g * h' = e$ für alle $g \in G$, so erhalten wir $h = h * e = h * (g * h') = (h * g) * h' = e * h' = h'$.

Beweis. Es gilt

$$(g*h)*(h^{-1}*g^{-1}) \overset{(G1)}{=} g*(h*(h^{-1}*g^{-1})) \overset{(G1)}{=} g*((h*h^{-1})*g^{-1}) \overset{(G3)}{=} g*(e*g^{-1}) \overset{(G2)}{=} g*g^{-1} \overset{(G3)}{=} e.$$

Analog zeigt man $(h^{-1} * g^{-1}) * (g * h) = e$. Wegen der Eindeutigkeit des Inversen von $g * h$ folgt die erste Behauptung. Wegen $g^{-1} * g = e = g^{-1} * g$ ist g invers zu g^{-1}. Da das Inverse von g^{-1} eindeutig bestimmt ist, folgt die zweite Behauptung. Wegen $e * e = e = e * e$ ist e Inverses von e, und die Eindeutigkeit von Inversen zeigt die dritte Behauptung. □

1.1.5 (Kürzen in Gruppen)**.** Gilt in einer Gruppe $g * a = g * b$, so folgt $a = b$ (multipliziere von links mit g^{-1}). Analog folgt aus $a * g = b * g$ bereits $a = b$.

1.1.6. In einer beliebigen Gruppe darf man wegen der Assoziativität Klammern weglassen und kurz $g * h * l$ statt $(g * h) * l = g * (h * l)$ und $a * b * c * d * e$ statt $a * (b * (c * (d * e))) = ((a * b) * (c * d)) * e = \dots$ schreiben (genaue Begründung siehe [Sch20]).

1.1.7. In abelschen Gruppen ist die Reihenfolge der verknüpften Gruppenelemente egal, beispielsweise gilt $a * b * c = c * a * b$. Die Reihenfolge der Faktoren ist in nicht abelschen Gruppen aber äußerst wichtig.

Bemerkung 1.1.8. Wird als Verknüpfungssymbol $\cdot$ verwendet, so spricht man von einer *multiplikativ* geschriebenen Gruppe, nennt die Verknüpfung Multiplikation, das neutrale Element **Einselement** (und notiert es oft als 1). Meist schreibt man dann gh statt $g \cdot h$ und nennt dies das *Produkt* von g und h.

Erwähnt man die Verknüpfung nicht explizit, so geht man in der Regel von einer multiplikativ geschriebenen Gruppe aus.

Das Verknüpfungssymbol $+$ wird nur dann verwendet, wenn die betrachtete Gruppe kommutativ ist. Man spricht dann von *additiver* Notation, schreibt das neutrale Element meist als 0, nennt es **Nullelement**, nennt $g + h$ die *Summe* und schreibt $-g$ für das Inverse eines Elements $g \in G$, und $h - g$ statt $h + (-g)$.

1.1.9. Sei $(G, *)$ eine Gruppe und seien $g \in G$ und $n \in \mathbb{Z}$. Definiere

$$g^n := \begin{cases} g * \dots * g & \text{falls } n \geq 1 \text{ (dabei } n \text{ Faktoren im Produkt)}, \\ e & \text{falls } n = 0, \\ g^{-1} * \dots * g^{-1} & \text{falls } n \leq -1 \text{ (dabei } -n \text{ Faktoren im Produkt)}. \end{cases}$$

Mit dieser Definition gilt $g^n * g^m = g^{n+m}$ für alle $n, m \in \mathbb{Z}$. Ebenso gilt $g^{nm} = (g^n)^m$.

In einer additiv geschriebenen Gruppe G schreibt man ng statt g^n, es gilt also

$$ng = \begin{cases} g + \dots + g & \text{falls } n \geq 1 \text{ (n Summanden)}, \\ 0 & \text{falls } n = 0, \\ -g - \dots - g & \text{falls } n \leq -1 \text{ ($-n$ Summanden)}. \end{cases}$$

1.2. **Untergruppen.**

Definition 1.2.1. Sei G eine Gruppe. Eine Teilmenge[3] $H \subset G$ heißt genau dann **Untergruppe**, wenn die folgenden Bedingungen gelten:

(a) Für alle $g, h \in H$ gilt $gh \in H$.

[3]Wir verwenden das Zeichen $\subset$ für beliebige Teilmengen (es kann also auch Gleichheit gelten).

(b) Für alle $g \in H$ gilt $g^{-1} \in H$.
(c) $e \in H$.

Dann ist H zusammen mit der induzierten Verknüpfung $H \times H \to H$ ebenfalls eine Gruppe.

Beispiele 1.2.2. (a) $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{Q}, +)$.
(b) Die Menge

$$\{\text{Drehung um } \alpha \text{ Grad} \mid \alpha \in \{0, 60, 120, 180, 240, 300\}\}$$

der Drehungen um ein Vielfaches des Winkels 60° ist eine Untergruppe der Gruppe aller Drehungen der Ebene aus Beispiel 1.1.3.(e).
(c) Jede Gruppe G hat $\{e\}$ und G als Untergruppen.
(d) Der Durchschnitt von (endlich vielen oder auch unendlich vielen) Untergruppen einer Gruppe ist eine Untergruppe. Beispielsweise ist der Schnitt auf der rechten Seite in (1.2.1) in Lemma 1.2.4 eine Untergruppe.

Beispiel 1.2.3 (Symmetriegruppen)**.** (a) Wir verwenden den Betrag $|\cdot|$ auf $\mathbb{R}$. Eine Abbildung $g\colon \mathbb{R} \to \mathbb{R}$ heißt **abstandserhaltend**, wenn $|x - y| = |g(x) - g(y)|$ für alle $x, y \in \mathbb{R}$ gilt. Man zeigt leicht, dass es für jede solche Abbildung eindeutige reelle Zahlen $a \in \{+1, -1\}$ und $b \in \mathbb{R}$ mit $g(x) = ax + b$ für alle $x \in \mathbb{R}$ gibt. Die Menge aller abstandserhaltenden Abbildungen ist eine Untergruppe der symmetrischen Gruppe $\mathrm{Sym}(\mathbb{R})$.
(b) Ist $A \subset \mathbb{R}$ eine Teilmenge, etwa die im folgenden Bild durch dicke Punkte angedeutete, so kann man die **Symmetriegruppe von** A betrachten (bitte nicht mit der symmetrischen Gruppe verwechseln). Sie besteht per Definition aus allen abstandserhaltenden Abbildungen $g\colon \mathbb{R} \to \mathbb{R}$ mit $g(A) = A$. Für die skizzierte Teilmenge

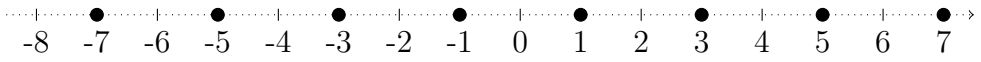


Abbildung 1. Zur Symmetriegruppe der Teilmenge der ungeraden Zahlen

$A = \{2n + 1 \mid n \in \mathbb{Z}\}$ aller ungeraden Zahlen sind die folgenden Abbildungen in der Symmetriegruppe:
- alle Verschiebungen $x \mapsto x + 2n$ um gerade Zahlen, für $n \in \mathbb{Z}$;
- allen Spiegelungen $x \mapsto -(x - n) + n$ an Punkten $n \in \mathbb{Z}$.

Man überzeugt sich leicht, dass die Symmetriegruppe genau aus diesen Elementen besteht.
(c) Analog definiert man abstandserhaltende Abbildungen $g\colon \mathbb{R}^n \to \mathbb{R}^n$ mit Hilfe der Norm $\|v\| = \sqrt{\sum v_i^2}$ und Symmetriegruppen von Teilmengen von $\mathbb{R}^n$. Elemente der Symmetriegruppe mag man als *Symmetrien* der gegebenen Teilmenge bezeichnen.
(d) Beispielsweise besteht die Symmetriegruppe des Quadrats in Abbildung 2 aus acht Elementen (vier Drehungen, vier Spiegelungen an Geraden), die des regulären Fünfecks aus 10 Elementen (fünf Drehungen, fünf Spiegelungen an Geraden); diese Aussage ist anschaulich klar (vgl. Aufgabe 8.4.5). Analoges gilt für reguläre n-Ecke für $n \geq 3$.
(e) Slogan: Die Größe der Symmetriegruppe ist ein Maß für die Anzahl der Symmetrien eines Objekts.

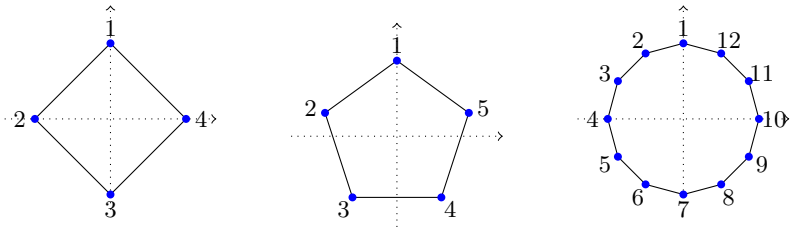


ABBILDUNG 2. Zu Symmetriegruppen von regulären n-Ecken

Symmetrisch erscheinende Objekte haben große Symmetriegruppen. Eine zufällig gewählte Teilmenge von $\mathbb{R}^2$ hat in der Regel keine Symmetrie: Ihre Symmetriegruppe besteht nur aus der Identitätsabbildung.

Lemma 1.2.4. *Sei G eine Gruppe und $M \subset G$ eine Teilmenge. Dann ist die Teilmenge*[4]

$$\langle M\rangle := \left\{ g_1^{\varepsilon_1} \dots g_r^{\varepsilon_r} \mid r \geq 0,\ g_1, \dots, g_r \in M,\ \varepsilon_1, \dots, \varepsilon_r \in \{\pm 1\} \right\} \subset G$$

*eine Untergruppe von G; sie heißt die **von M erzeugte Untergruppe** und ist die kleinste Untergruppe von G, die M enthält. Weiter gilt*

$$\langle M\rangle = \bigcap_{H \text{ Untergruppe von } G \text{ mit } M \subset H} H. \tag{1.2.1}$$

Beweis. Es ist klar, dass das Produkt zweier Elemente von $\langle M\rangle$ wieder in $\langle M\rangle$ liegt, und dass $e \in \langle M\rangle$ gilt (da e das Produkt von $r = 0$ Elementen ist). Sei $g = g_1^{\varepsilon_1} \dots g_r^{\varepsilon_r} \in \langle M\rangle$. Wegen $g^{-1} = g_r^{-\varepsilon_r} \dots g_1^{-\varepsilon_1}$ gilt $g^{-1} \in \langle M\rangle$. Dies zeigt, dass $\langle M\rangle$ eine Untergruppe von G ist. Offensichtlich gilt $M \subset \langle M\rangle$.

Ist $H \subset G$ eine Untergruppe, die alle Elemente von M enthält, so enthält sie auch alle Inversen von Elementen von M und somit alle Produkte $g_1^{\varepsilon_1} \dots g_r^{\varepsilon_r}$ mit $g_i \in M$ und $\varepsilon_i \in \{\pm 1\}$. Es folgt $\langle M\rangle \subset H$. Also ist $\langle M\rangle$ die kleinste Untergruppe von G, die M enthält.

Die Gleichheit (1.2.1) ist damit klar: Die Inklusion $\subset$ gilt wegen $\langle M\rangle \subset H$ für alle Untergruppen $H \subset G$ mit $M \subset H$. Die Inklusion $\supset$ gilt, da $\langle M\rangle$ eine Untergruppe von G ist, die M enthält, also ein mögliches H im Schnitt auf der rechten Seite ist. □

1.2.5. Gegeben Elemente $m_1, \dots, m_s \in G$ schreiben wir $\langle m_1, \dots, m_s\rangle$ statt $\langle\{m_1, \dots, m_s\}\rangle$. Gilt $G = \langle M\rangle$, so sagt man, dass G **von** M **erzeugt** ist und nennt die Elemente von M **Erzeuger von** G.

Beispiele 1.2.6.
(a) Es gilt $(\mathbb{Z}, +) = \langle 1\rangle$, denn jedes Element $n \in \mathbb{Z}$ lässt sich als $1 + \dots + 1$ oder als $-1 - 1 \cdots - 1$ schreiben. Analog gilt $(\mathbb{Z}, +) = \langle -1\rangle$.
(b) Die Untergruppe $\langle 6, 10\rangle \subset \mathbb{Z}$ besteht genau aus allen geraden Zahlen.
(c) Die Symmetriegruppe in Beispiel 1.2.3.(b) wird von der Translation (= Verschiebung) $x \mapsto x + 2$ und der Reflektion (= Spiegelung) $x \mapsto -x$ in Null erzeugt.

1.3. **Gruppenhomomorphismen.**

Definition 1.3.1. Seien $(G, \circ)$ und $(H, *)$ Gruppen. Ein **Gruppenhomomorphismus** oder **Homomorphismus von Gruppen** $\varphi : G \to H$ ist eine Abbildung $\varphi : G \to H$ von Mengen

[4] Im Fall $r = 0$ ist das Produkt $g_1^{\varepsilon_1} \dots g_r^{\varepsilon_r}$ als neutrales Element $e \in G$ zu verstehen.

mit

$$\varphi(x \circ y) = \varphi(x) * \varphi(y)$$

für alle $x, y \in G$.

Beispiele 1.3.2. (a) Sei $\lambda \in \mathbb{R}$ eine reelle Zahl. Die Abbildung $\varphi\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto \lambda x$, also die Streckung um den Faktor λ, ist ein Gruppenmorphismus, denn $\varphi(x+y) = \lambda(x+y) = \lambda x + \lambda y = \varphi(x) + \varphi(y)$ für alle $x, y \in \mathbb{R}$.

(b) Die Exponentialfunktion $\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$, $x \mapsto \exp(x)$, ist ein Gruppenhomomorphismus, da $\exp(x+y) = \exp(x)\exp(y)$ für alle $x, y \in \mathbb{R}_{>0}$ gilt (Funktionalgleichung der Exponentialfunktion).

(c) Ihre Umkehrabbildung, die Logarithmusfunktion $\log : (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +)$, ist ein Gruppenhomomorphismus.

Lemma 1.3.3. *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann gelten*

$$\varphi(e_G) = e_H,$$
$$\varphi(g^{-1}) = (\varphi(g))^{-1} \qquad \text{für alle } g \in G.$$

Beweis. Aus $\varphi(e_G) = \varphi(e_G e_G) = \varphi(e_G)\varphi(e_G)$ folgt durch Linksmulitplikation mit $\varphi(e_G)^{-1}$ die erste Behauptung. Es gilt

$$\varphi(g)\varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(e_G) = e_H = \varphi(g)\varphi(g)^{-1}.$$

Kürzen von $\varphi(g)$ liefert die zweite Behauptung. □

Beispiel 1.3.4. Sei $\lambda \in \mathbb{R}$ eine reelle Zahl mit $\lambda \neq 0$. Die Verschiebe-Abbildung $\varphi\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x + \lambda$, ist kein Gruppenhomomorphismus, da $\varphi(0) = \lambda \neq 0$.

Lemma 1.3.5. *Bilder und Urbilder von Untergruppen unter Gruppenhomomorphismen sind Untergruppen. Explizit: Ist $\varphi : G \to H$ ein Gruppenhomomorphismus, so gelten:*

(a) Ist $U \subset G$ eine Untergruppe, so ist $\varphi(U) = \{\varphi(u) \mid u \in U\}$ eine Untergruppe von H.

(b) Ist $V \subset H$ eine Untergruppe, so ist $\varphi^{-1}(V) = \{g \in G \mid \varphi(g) \in V\}$ eine Untergruppe von G.

Beweis. Übungsaufgabe, Blatt 1. □

Definition 1.3.6. Sei $\varphi\colon G \to H$ ein Gruppenhomomorphismus.

(a) $\ker(\varphi) := \varphi^{-1}(\{e_H\})$ heißt der **Kern** von φ.

(b) $\mathrm{im}(\varphi) := \varphi(G) \subset H$ heißt das **Bild** von φ. (Bezeichnung wegen englisch *image* für *Bild*.)

Kern und Bild sind Untergruppen von G bzw. H wegen Lemma 1.3.5

Lemma 1.3.7. *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann gilt*

$$\varphi \text{ injektiv} \iff \ker(\varphi) = \{e_G\}.$$

Beweis. $\Rightarrow$: Nach Lemma 1.3.3 gilt stets $\varphi(e_G) = e_H$, also $\{e_G\} \subset \ker(\varphi)$. Gilt umgekehrt $x \in \ker(\varphi)$, also $\varphi(x) = e_H = \varphi(e_G)$, so folgt aus der angenommen Injektivität $x = e_G$, also $\ker(\varphi) \subset \{e_G\}$. Dies zeigt die Gleichheit $\ker(\varphi) = \{e_G\}$.

$\Leftarrow$: Seien $x, y \in G$ mit $\varphi(x) = \varphi(y)$. Multiplizieren wir diese Gleichung von links mit $\varphi(x)^{-1} = \varphi(x^{-1})$ (Lemma 1.3.3), so ergibt sich $e = \varphi(x^{-1})\varphi(x) = \varphi(x^{-1})\varphi(y) = \varphi(x^{-1}y)$, also $x^{-1}y \in \ker(\varphi) = \{e_G\}$, also $x^{-1}y = e_G$, also $y = x$ per Linksmultiplikation mit x. Somit ist φ injektiv. □

Definition 1.3.8. Ein **Gruppenisomorphismus** oder **Isomorphismus von Gruppen** ist ein bijektiver Gruppenhomomorphismus $\varphi : G \to H$ und wird als $\varphi\colon G \xrightarrow{\sim} H$ notiert. Zwei Gruppen G und H heißen **isomorph**, falls es einen Gruppenisomorphismus $\varphi\colon G \xrightarrow{\sim} H$ gibt.

Bemerkung 1.3.9. Ist $\varphi : G \xrightarrow{\sim} H$ ein Gruppenisomorphismus, so ist die mengentheoretische Umkehrabbildung $\varphi^{-1}\colon H \to G$ ebenfalls ein Gruppenisomorphismus: Bijektivität ist klar. Für beliebige $x, y \in H$ bleibt $\varphi^{-1}(x)\varphi^{-1}(y) = \varphi^{-1}(xy)$ zu zeigen. Da φ injektiv ist, genügt es, die Gleichheit $\varphi(\varphi^{-1}(x)\varphi^{-1}(y)) = \varphi(\varphi^{-1}(xy))$ zu zeigen. Diese gilt aber wegen $\varphi(\varphi^{-1}(x)\varphi^{-1}(y)) = \varphi(\varphi^{-1}(x))\varphi(\varphi^{-1}(y)) = xy = \varphi(\varphi^{-1}(xy))$.

Ende 1. Vorlesung am 21.04.2020

1.4. Permutationen.

Definition 1.4.1. Sei $n \in \mathbb{N}$. Die Gruppe

$$S_n := \mathrm{Sym}(\{1, \dots, n\}) = \{\pi\colon \{1, \dots, n\} \to \{1, \dots, n\} \mid \pi \text{ bijektiv}\}$$

aller bijektiven Selbstabbildungen der Menge $\{1, \dots, n\}$ (mit Komposition als Verknüpfung) heißt **n-te symmetrische Gruppe** (Spezialfall von Beispiel 1.1.3.(d)). Elemente von S_n heißen **Permutationen**.

1.4.2 (Permutationsschreibweise)**.** Um ein Element $\pi \in S_n$ anzugeben, verwendet man oft die **Permutationsschreibweise**

$$\pi = \begin{bmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{bmatrix}.$$

Da π bijektiv ist, kommt jede der Zahlen $1, \dots, n$ in der zweiten Zeile genau einmal vor.

Satz 1.4.3. *Die n-te symmetrische Gruppe hat $n! = n(n-1)\cdots 2 \cdot 1$ Elemente, d. h.*

$$|S_n| = n!.$$

Beweis. Ein beliebiges Element π von S_n ist eindeutig bestimmt und definiert durch die Folge $\pi(1), \pi(2), \dots, \pi(n)$ von paarweise verschiedenen Elementen von $\{1, \dots, n\}$. Für $\pi(1)$ hat man n Möglichkeiten, für $\pi(2)$ hat man dann noch $n-1$ Möglichkeiten, etc., und für $\pi(n)$ hat man noch genau eine Möglichkeit. □

Definition 1.4.4. Sind $p_1, p_2, \dots, p_l$ verschiedene Elemente von $\{1, \dots, n\}$, so bezeichne

$$(p_1\ p_2\ \dots\ p_l)$$

die Permutation, die die gegebenen Elemente zyklisch vertauscht, also wie folgt abbildet,

$$p_1 \mapsto p_2, \quad p_2 \mapsto p_3, \quad \dots, \quad p_l \mapsto p_1,$$

und alle anderen Elemente fixiert. Elemente dieser Art heißen **l-Zykel**.

Beispiel 1.4.5. Es gilt $(3\ 4\ 7\ 9) = (4\ 7\ 9\ 3) = (7\ 9\ 3\ 4) = (9\ 3\ 4\ 7)$ in S_9. Oft ist man etwas nachlässig und schreibt (3479) statt $(3\ 4\ 7\ 9)$. Kommen mindestens zweistellige Zahlen vor, sind klare Abstände in der Notation wichtig: $(24) \neq (2\ 4)$ in S_{24}. 1-Zykel wie (3) oder (24) sind nur eine komplizierte Schreibweise für das neutrale Element $e = \mathrm{id}_{\{1,\dots,n\}}$.

1.4.6. Wegen $(12)(23) = (123) \neq (132) = (23)(12)$ ist die symmetrische Gruppe S_n für $n \geq 3$ nicht kommutativ.

Faustregel: Beim Verknüpfen von Zykeln lies von rechts nach links!

Beispiel 1.4.7. In der folgenden Auflistung sind alle Elemente von S_3 in Permutationsschreibweise angegeben. Unter jeder Permutation steht dieselbe Permutation als Verknüpfung von Zykeln.

$$\begin{aligned} S_3 &= \left\{ \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \right\} \\ &= \quad \{(1)(2)(3), \quad (1)(23), \quad (12)(3), \quad (123), \quad (132), \quad (13)(2)\}. \end{aligned}$$

In der zweiten Zeile kann man alle 1-Zykel (1), (2) und (3) weglassen. Beachte $(123) = (231) = (312)$ und $(132) = (321) = (213)$.

Beispiel 1.4.8. Gruppentafel der S_3. Wählt man in der linken Spalte eine Permutation π und in der oberen Zeile eine Permutation σ, so ist $\pi \circ \sigma$ im zugehörigen Kästchen eingetragen.

	id	(23)	(12)	(123)	(132)	(13)
id	id	(23)	(12)	(123)	(132)	(13)
(23)	(23)	id	(132)	(13)	(12)	(123)
(12)	(12)	(123)	id	(23)	(13)	(132)
(123)	(123)	(12)	(13)	(132)	id	(23)
(132)	(132)	...	...	...	...	...
(13)	(13)	...	...	...	...	...

Lemma 1.4.9 (Zykelschreibweise oder genauer Schreibweise als Produkt von Zykeln)**.** *Jedes Element $\pi \in S_n$ kann als Produkt von Zykeln geschrieben werden, d. h. es gibt eine natürliche Zahl $r \geq 0$ und Elemente $a_1, \ldots, a_r \in \{1, \ldots, n\}$, so dass*

$$\pi = \Big(a_1\ \pi(a_1)\ \pi(\pi(a_1)) \cdots \pi^{-1}(a_1)\Big)\Big(a_2\ \pi(a_2)\ \cdots \pi^{-1}(a_2)\Big) \cdots \Big(a_r\ \pi(a_r) \cdots \pi^{-1}(a_r)\Big).$$

Man kann dabei genauer annehmen, dass kein Element von $\{1, \ldots, n\}$ in mehr als einem Zykel vorkommt.

Beweis. Betrachte die Folge $1, \pi(1), \pi^2(1), \ldots$. Da die Menge $\{1, \ldots, n\}$ endlich ist, gibt es natürliche Zahlen $i < j$ mit $\pi^i(1) = \pi^j(1)$. Gilt $i > 0$, so gilt wegen der Injektivität von π auch $\pi^{i-1}(1) = \pi^{j-1}(1)$. Deswegen können wir ohne Einschränkung annehmen, dass $i = 0$ gilt, d. h. $1 = \pi^j(1)$. Weiter können wir annehmen, dass $j > 0$ minimal gewählt ist mit $1 = \pi^j(1)$. Der Zykel

$$\Big(1\ \pi(1)\ \ldots\ \pi^{j-1}(1)\Big)$$

und π stimmen offensichtlich auf der j-elementigen Teilmenge $\{1, \pi(1), \pi^2(1), \ldots, \pi^{j-1}(1)\}$ von $\{1, \ldots, n\}$ überein. Falls diese Teilmenge schon die ganze Menge $\{1, \ldots, n\}$ ist, sind wir fertig. Sonst gibt es ein b in ihrem Komplement, für das man nun die Folge $b, \pi(b), \pi^2(b), \ldots$ betrachtet und ein minimales $k > 0$ mit $b = \pi^k(b)$ findet. Dann stimmen

$$\Big(1\ \pi(1)\ \pi(\pi(1)) \cdots \pi^{j-1}(1)\Big)\Big(b\ \pi(b)\ \pi^2(b) \cdots \pi^{k-1}(b)\Big)$$

und π auf der $(j+k)$-elementigen (warum?) Teilmenge $\{1, \ldots, \pi^{j-1}(1), b, \ldots, \pi^{k-1}(b)\}$ überein, und wir können analog fortfahren. Da $\{1, \ldots, n\}$ endlich ist, terminiert unser Verfahren nach endlich vielen Schritten. Setze $a_1 = 1, a_2 = b, \ldots$. □

Beispiel 1.4.10. Wir geben eine Permutation in Permutationsschreibweise (diese ist offensichtlich eindeutig) und in zwei möglichen Zykelschreibweisen an.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 9 & 3 & 5 & 10 & 2 & 8 & 7 & 6 & 1 \end{bmatrix} = (1\ 4\ 5\ 10)\ (2\ 9\ 6)\ (3)\ (7\ 8) = (8\ 7)\ (5\ 10\ 1\ 4\ )\ (9\ 6\ 2)$$

In der Zykelschreibweise kann man innerhalb von Zykeln zyklisch vertauschen, Zykel vertauschen und 1-Zykel weglassen, ohne das Element zu verändern. Man könnte hinten auch noch etwa $(1\ 5\ 6\ 9)(9\ 6\ 5\ 1) = \mathrm{id}_{1,\dots,10}$ dranschreiben, aber meist nimmt man in der Zykelschreibweise implizit an, dass kein Element in zwei Zykeln vorkommt.

Beispiel 1.4.11. Sei G die Symmetriegruppe des Quadrats in Beispiel 1.2.3.(d). Die vier Eckpunkte des Quadrats bezeichnen wir wie dort mit 1,2,3,4 (formal genauer wäre p_1, p_2, p_3, p_4). Sei $g \in G$ eine Symmetrie des Quadrats. Da $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ jeden Eckpunkt auf genau einen Eckpunkt abbildet, induziert g durch Einschränkung eine Permutation der Menge $\{1,2,3,4\}$ der Eckpunkte. Wir nennen diese Permutation $\varphi(g)$. Die Abbildung

$$\begin{aligned} \varphi\colon G &\to S_4, \\ g &\mapsto \varphi(g) = g|_{\{1,2,3,4\}}, \end{aligned}$$

ist ein Gruppenhomomorphismus, denn Einschränken vertauscht mit Komposition von Abbildungen: $(g \circ h)|_{\{1,2,3,4\}} = g|_{\{1,2,3,4\}} \circ h|_{\{1,2,3,4\}}$. Die acht Elemente von G werden wir folgt abgebildet:

(1) Die Drehungen um 0°, 90°, 180° und 270° werden in dieser Reihenfolge auf die folgenden Elemente abgebildet:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix} = (1234), \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{bmatrix} = (13)(24), \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix} = (1432).$$

(2) Die Spiegelungen an vertikaler und horizontaler Achse und den beiden diagonalen Achsen SW-NO und SO-NW werden in dieser Reihenfolge auf die folgenden Elemente abgebildet:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{bmatrix} = (24), \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix} = (13), \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix} = (14)(23), \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{bmatrix} = (12)(34).$$

Wir sehen spätestens jetzt, dass φ injektiv ist (direkt oder per Lemma 1.3.7). Also induziert φ einen Gruppenisomorphismus $G \xrightarrow{\sim} \mathrm{im}(\varphi)$. Das Bild $\mathrm{im}(\varphi)$ besteht aus den angegebenen acht Elementen und kann durch Erzeuger beispielsweise als $\mathrm{im}(\varphi) = \langle (1234), (24) \rangle$ beschrieben werden. Die beiden Erzeuger (1234), (24) kommutieren nicht: $(1234)(24) = (12)(34) \neq (14)(23) = (24)(1234)$. Geometrisch bedeutet dies, dass die Spiegelung an der vertikalen Achse nicht mit der Drehung um 90° kommutiert.

Definition 1.4.12. Eine **Transposition** ist ein 2-Zykel, hat also die Form $\tau_{i,j} := (ij) \in S_n$ für geeignete Zahlen $i, j \in \{1, \dots, n\}$ mit $i \neq j$. Eine Transposition der Form $\tau_{i,i+1} = (i\ (i+1))$ heißt **einfache Transposition**.

Satz 1.4.13. *Jedes Element von S_n kann als Produkt einfacher Transpositionen geschrieben werden.*

Insbesondere ist die symmetrische Gruppe S_n von der Menge der einfachen Transpositionen erzeugt, in Formeln $S_n = \langle \{\text{einfache Transpositionen}\} \rangle$.

Beweis. Wir führen Induktion über n.

Induktionsanfang $n = 1$ (oder auch $n = 0$): Offensichtlich, da $S_1 = \{\mathrm{id}_{\{1\}}\}$ (und $S_0 = \{\mathrm{id}_\emptyset\}$).

Induktionsschritt: Sei $n \geq 1$ und sei die Aussage für $n-1$ per Induktionsannahme bereits gezeigt. Sei $\pi \in S_n$ beliebig. Setze $i := \pi(n)$ und definiere $\pi' := \tau_{n-1,n} \circ \cdots \circ \tau_{i,i+1} \circ \pi$. Dann gilt $\pi'(n) = n$, d. h. wir können π' als Element von S_{n-1} auffassen (indem wir π' auf die Menge $\{1, \ldots, n-1\}$ einschränken). Nach Induktionvoraussetzung ist π' ein Produkt von einfachen Transpositionen. Also ist auch $\pi = \tau_{i,i+1} \circ \ldots \circ \tau_{n-1,n} \circ \pi'$ ein solches Produkt. □

Beispiel 1.4.14. In S_3 gelten $(13) = (12)(23)(12) = (23)(12)(23)$ und $(123) = (12)(23)$ und $(132) = (23)(12)$.

Definition 1.4.15. Sei $\pi \in S_n$.

(a) Ein **Fehlstand von** π ist ein Paar $(i, j) \in \{1, \ldots, n\}^2$ mit $i < j$ und $\pi(i) > \pi(j)$.

(b) Die **Länge** $\ell(\pi)$ **von** π ist die Anzahl der Fehlstände von π, in Formeln[5]

$$\ell(\pi) := \#\Big\{(i, j) \in \{1, \ldots, n\}^2 \mid i < j \text{ und } \pi(i) > \pi(j)\Big\}.$$

(c) Die Zahl

$$\mathrm{sgn}(\pi) := (-1)^{\ell(\pi)}$$

heißt **Signum** (oder **Signatur** oder **Vorzeichen**) **von** π.

Beispiele 1.4.16. Erläutere per Bild: Kreuzungen entsprechen Fehlständen.

(a) Die Permutation $(123) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$ hat zwei Fehlstände, nämlich $(1, 3)$ und $(2, 3)$, somit Länge 2 und Signatur 1.

(b) Die Permutation $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 9 & 3 & 5 & 10 & 2 & 8 & 7 & 6 & 1 \end{bmatrix}$ in Beispiel 1.4.10 hat 26 Fehlstände, also Länge 26 und Signatur 1.

Satz 1.4.17. *Sei $\pi \in S_n$ als Produkt $\pi = \tau_1 \circ \ldots \circ \tau_r$ einfacher Transpositionen $\tau_1, \ldots, \tau_r$ geschrieben (eine solche Darstellung existiert nach Satz 1.4.13). Dann gilt*

$$\mathrm{sgn}(\pi) = (-1)^r.$$

Insbesondere ist $\mathrm{sgn} : S_n \to (\{\pm 1\}, \cdot)$ *ein Gruppenhomomorphismus.*

Beweis. Sei $\pi \in S_n$ und sei $\tau = \tau_{a,a+1}$ eine einfache Transposition. Wir behaupten $\ell(\tau \circ \pi) = \ell(\pi) \pm 1$. Sei ein Paar (i, j) mit $1 \leq i < j \leq n$ gegeben. Die „Reihenfolge“ von $\pi(i)$ und $\pi(j)$ unterscheidet sich genau dann von der „Reihenfolge“ von $\tau(\pi(i))$ und $\tau(\pi(j))$, wenn $\{\pi(i), \pi(j)\} = \{a, a+1\}$ gilt. Da π bijektiv ist, erfüllt genau ein Paar (i, j) die Bedingung $\{\pi(i), \pi(j)\} = \{a, a+1\}$. Also unterscheiden sich die Längen von π und $\tau \circ \pi$ um eins, in Formeln $\ell(\tau \circ \pi) = \ell(\pi) \pm 1$. Es folgt

$$\mathrm{sgn}(\tau \circ \pi) = (-1)^{\ell(\tau \circ \pi)} = (-1)^{\ell(\pi)} \cdot (-1)^{\pm 1} = -\mathrm{sgn}(\pi).$$

[5] In Zukunft will ich das besser **Fehlstandsanzahl** nennen (damit es keine Verwechslungen mit der „Länge eines Zykels“ gibt). Gute Abkürzung dafür? Vielleicht $\mathrm{FS}(\pi)$ als Menge der Fehlstände definieren und dann einfach $\#\,\mathrm{FS}(\pi)$ schreiben? Dann muss auch gar keinen Namen vergeben.

Ist nun $\pi = \tau_1 \circ \ldots \circ \tau_r$ ein Produkt einfacher Transpositionen $\tau_1, \ldots, \tau_r$, so folgt daraus die erste Behauptung:

$$\operatorname{sgn}(\pi) = \operatorname{sgn}(\tau_1 \circ \ldots \circ \tau_r) = -\operatorname{sgn}(\tau_2 \circ \ldots \circ \tau_r) = \cdots = (-1)^r \operatorname{sgn}(\mathrm{id}) = (-1)^r$$

Sei nun $\sigma \in S_n$ gegeben. Schreibe $\sigma = \tau_1' \circ \ldots \circ \tau_s'$ als Produkt einfacher Transpositionen. Dann gilt $\sigma \circ \pi = \tau_1' \circ \ldots \circ \tau_s' \circ \tau_1 \circ \ldots \circ \tau_r$ und somit nach obigem wie gewünscht

$$\operatorname{sgn}(\sigma \circ \pi) = (-1)^{r+s} = (-1)^r (-1)^s = \operatorname{sgn}(\sigma) \operatorname{sgn}(\pi). \qquad \square$$

Beispiele 1.4.18. (a) Der m-Zykel $(1\ 2\ 3\ \ldots m)$ hat Länge $m-1$, also Signatur $(-1)^{m-1}$.

(b) Allgemeiner hat jeder m-Zykel $(p_1\ p_2\ \ldots\ p_m)$ Signatur $(-1)^{m-1}$ (aber nicht Länge $m-1$: die Länge des 2-Zykels (13) ist 3 und nicht 1).

Beweis. Für beliebige Permutationen $\sigma, \tau \in S_n$ gilt nach Satz 1.4.17

$$\operatorname{sgn}(\tau\sigma\tau^{-1}) = \operatorname{sgn}(\tau) \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau^{-1}) = \operatorname{sgn}(\tau) \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)^{-1} = \operatorname{sgn}(\sigma). \tag{1.4.1}$$

Für

$$\tau := \begin{bmatrix} 1 & 2 & 3 & \ldots & m-1 & m & m+1 & \ldots & n \\ p_1 & p_2 & p_3 & \ldots & p_{m-1} & p_m & m+1 & \ldots & n \end{bmatrix}$$

und $\pi := (p_1\ p_2\ \ldots\ p_m)$ berechnen wir

$$\tau^{-1}\pi\tau = (1\ 2\ \ldots\ m).$$

Äquivalent gilt $\pi = \tau(1\ 2\ \ldots\ m)\tau^{-1}$ (multipliziere von links mit τ und von rechts mit τ^{-1}). Aus (1.4.1) und (a) folgt damit wie gewünscht

$$\operatorname{sgn}(\pi) = \operatorname{sgn}(\tau(1\ 2\ \ldots\ m)\tau^{-1}) = \operatorname{sgn}\big((1\ 2\ \ldots\ m)\big) = (-1)^{m-1}. \qquad \square$$

(c) Mit unserem neuen Wissen berechnen wir noch einmal die Signatur der Permutation in Beispiel 1.4.10:

$$\operatorname{sgn}\big((1\ 4\ 5\ 10)\ (2\ 9\ 6)\ (7\ 8)\big) = \operatorname{sgn}\big((1\ 4\ 5\ 10)\big) \cdot \operatorname{sgn}\big((2\ 9\ 6)\big) \cdot \operatorname{sgn}\big((7\ 8)\big) = (-1) \cdot 1 \cdot (-1) = 1.$$

Ende 2. Vorlesung am 23.04.2020

1.5. Äquivalenzrelationen.

1.5.1. Sei A eine Menge. Wir erinnern daran, dass eine Teilmenge $R \subset A \times A$ auch als **Relation auf** A bezeichnet wird. Statt $(x, y) \in R$ sagen wir auch, dass $R(x, y)$ gilt.

Definition 1.5.2. Eine Relation $R \subset A \times A$ auf einer Menge A heißt **Äquivalenzrelation**, wenn sie reflexiv, symmetrisch und transitiv ist, wenn also die folgenden Bedingungen gelten:

(a) Reflexivität: Für alle $x \in A$ gilt $R(x, x)$.
(b) Symmetrie: Für alle $x, y \in A$ gilt: Aus $R(x, y)$ folgt $R(y, x)$.
(c) Transitivität: Für alle $x, y, z \in A$ gilt: Aus $R(x, y)$ und $R(y, z)$ folgt $R(x, z)$.

Oft verwendet man das Symbol $\sim$ für Äquivalenzrelationen und schreibt $x \sim y$ statt $R(x, y)$ und sagt „x ist äquivalent zu y“.

Definition 1.5.3. Sei $\sim$ eine Äquivalenzrelation auf einer Menge A. Für $x \in A$ heißt die Teilmenge

$$[x] := \{y \in A \colon y \sim x\}$$

die **Äquivalenzklasse von x bezüglich $\sim$**. Eine Teilmenge $K \subset A$ heißt **Äquivalenzklasse bezüglich $\sim$**, wenn es ein $x \in A$ mit $K = [x]$ gibt. Jedes Element einer Äquivalenzklasse K heißt **Repräsentant** von K. Die Menge aller Äquivalenzklassen wird als

$$A/\sim := \{[x] \colon x \in A\}$$

bezeichnet. Formal ist $A/\sim$ eine Teilmenge der Potenzmenge $\mathcal{P}(A)$.

1.5.4. Die Abbildung $A \to A/\sim$, $a \mapsto [a]$, ist surjektiv.

1.5.5. Ist $\sim$ eine Äquivalenzrelation auf einer Menge A, so sind für beliebige Elemente $x, y \in A$ die folgenden Aussagen äquivalent:

- $x \sim y$,
- $[x] = [y]$,
- $[x] \cap [y] \neq \emptyset$.

Zwei Äquivalenzklassen sind also entweder gleich oder disjunkt. Da $x \in [x]$ für jedes $x \in A$ gilt, ist A die disjunkte Vereinigung seiner Äquivalenzklassen, in Formeln

$$A = \dot{\bigcup}_{K \in A/\sim} K.$$

Beispiel 1.5.6. Wir nehmen an, dass jeder Mensch in genau einer Stadt wohnt. Dann ist die Relation „x wohnt in der gleichen Stadt wie y“ eine Äquivalenzrelation auf der Menge aller Menschen. Die Äquivalenzklassen bestehen jeweils aus allen Bürgern von Paderborn, Mainz, ...

Beispiel 1.5.7. Die Relation $\geq$ auf $\mathbb{R}$ ist keine Äquivalenzrelation, da sie nicht symmetrisch ist.

Beispiel 1.5.8. Auf der Menge $A := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definieren wir eine Äquivalenzrelation $\sim$ durch die Festlegung, dass $(a, b) \sim (c, d)$ genau dann gilt, falls $ad = bc$ gilt. (Der Leser prüfe, dass dies „wohldefiniert“ ist, also wirklich eine Äquivalenzrelation auf A definiert.) Die rationalen Zahlen werden formal als Menge

$$\mathbb{Q} := A/\sim$$

der Äquivalenzklassen bezüglich dieser Äquivalenzrelation definiert. Die Äquivalenzklasse von (a, b) wird in der Regel als Bruch geschrieben, d. h. man definiert $\frac{a}{b} := [(a, b)]$. Wer mag, kann sich nun überlegen, dass die in der Schule definierte Addition und Multiplikation wohldefiniert ist, also nicht von der Wahl von Repräsentanten abhängt.

Schon aus der Schulmathematik sind Sie also mit Äquivalenzklassen vertraut.

Proposition 1.5.9. *Seien G eine Gruppe und $H \subset G$ eine Untergruppe. Für $x, y \in G$ schreibe $x \sim_H y$ genau dann, wenn $x^{-1}y \in H$ gilt. Dann ist $\sim_H$ ein Äquivalenzrelation auf G.*

Beweis. Seien $x, y, z \in G$ beliebig. Wir schreiben $\sim$ statt $\sim_H$.

Reflexivität: $x \sim x$ gilt wegen $x^{-1}x = e \in H$.

Symmetrie: Gelte $x \sim y$. Es folgt $x^{-1}y \in H$, also $H \ni (x^{-1}y)^{-1} = y^{-1}(x^{-1})^{-1} = y^{-1}x$, also $y \sim x$.

Transitivität: Gelten $x \sim y$ und $y \sim z$. Es folgen $x^{-1}y \in H$ und $y^{-1}z \in H$, also $H \ni x^{-1}yy^{-1}z = x^{-1}z$, also $x \sim z$. □

1.5.10. Im Setting von Proposition 1.5.9 ist die Äquivalenzklasse eines beliebigen Elements $g \in G$ durch $[g] = \{gh \mid h \in H\} =: gH$ gegeben. Man nennt diese Menge gH die **Rechtsnebenklasse von** g. Meist schreibt man G/H statt $G/\sim_H$ für die Menge dieser Klassen.

Beispiel 1.5.11. Für $G = S_3 \supset H = \langle(12)\rangle = \{\mathrm{id}, (12)\}$ gilt

$$S_3/H = \Big\{\big\{\mathrm{id}, (12)\big\}, \big\{(23), (132)\big\}, \big\{(13), (123)\big\}\Big\}.$$

Satz 1.5.12 (Satz von Lagrange)**.** *Sei G eine endliche Gruppe und sei H eine Untergruppe von G. Dann gilt*

$$\frac{|G|}{|H|} = \big|G/H\big| \in \mathbb{N}$$

Insbesondere ist $|H|$ ein Teiler von $|G|$.

Beweis. Betrachte auf G die Äquivalenzrelation $x \sim y \iff x^{-1}y \in H$ aus Proposition 1.5.9. Dann ist G die disjunkte Vereinigung seiner Äquivalenzklassen (siehe 1.5.5), d. h.

$$G = \dot{\bigcup}_{K \in G/H} K.$$

Wir behaupten, dass jede Äquivalenzklasse K aus genau $|H|$ Elementen besteht. In der Tat, für beliebiges $x \in G$ ist die Abbildung

$$\begin{aligned} [e] = eH = H &\to [x] = xH, \\ h &\mapsto xh. \end{aligned}$$

bijektiv, denn $k \mapsto x^{-1}k$ ist ihre Umkehrabbildung.

Weil G endlich ist, liefert die obige disjunkte Zerlegung von G somit

$$|G| = \sum_{K \in G/H} |K| = \sum_{K \in G/H} |H| = |G/H| \cdot |H|.$$

Als Kardinalität einer endlichen Menge ist $|G/H|$ eine natürliche Zahl. □

Beispiele 1.5.13. Der Satz von Lagrange 1.5.12 zeigt:

(a) Ist G eine endliche Gruppe, deren Kardinalität $|G|$ eine Primzahl ist, so hat G genau zwei Untergruppen, nämlich G und $\{e\}$.
(b) Die Gruppe S_3 hat $3! = 6$ Elemente. Ist H eine Untergruppe von S_3, so folgt $|H| \in \{1, 2, 3, 6\}$.

1.6. **Rechnen mit Nebenklassen / Modulo-Rechnung / Rechnen mit Uhrzeiten (Clock arithmetic).**

1.6.1. Seien $(A, +)$ eine additiv geschriebene abelsche Gruppe, $U \subset A$ eine Untergruppe und $\sim_U$ die Äquivalenzrelation aus Proposition 1.5.9. Dann gilt

$$x \sim_U y \iff y - x \in H \iff x - y \in H.$$

Die Äquivalenzklasse von $x \in A$ ist durch $x + U = \{x + u \mid u \in U\}$ gegeben und wird als **Nebenklasse von** x bezeichnet.

Beispiel 1.6.2. Betrachte die additive Gruppe $A = \mathbb{Z}$ und ihre Untergruppe $U = n\mathbb{Z} := \{nx \mid x \in \mathbb{Z}\}$ der Vielfachen von n. Die Äquivalenzklasse von $x \in \mathbb{Z}$ ist

$$[x]_n := [x] = x + n\mathbb{Z} = \{\dots, x-2n, x-n, x, x+n, x+2n, \dots\}.$$

Beispielsweise gilt $[2] = [2+n] = [2+42n] = [2-13n]$. Wir folgern, dass es n verschiedene Äquivalenzklassen gibt, nämlich

$$[0], [1], [2], \dots, [n-1].$$

Die Menge der Äquivalenzklassen ist also

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\sim = \{[0], [1], [2], \dots, [n-1]\}.$$

Der folgende Satz 1.6.3 zeigt, dass die Menge $\mathbb{Z}/n\mathbb{Z}$ in naheliegender Weise eine abelsche Gruppe wird.

Satz 1.6.3. *Sei $(A, +)$ eine abelsche Gruppe und sei $U \subset A$ eine Untergruppe. Dann wird die Menge A/U durch die (wohldefinierte) Verknüpfung*

$$[x] + [y] := [x + y] \qquad \text{für } x, y \in A$$

zu einer abelschen Gruppe.

Beweis. Seien $x, y, x', y' \in A$ mit $[x] = [x']$ und $[y] = [y']$. Dies bedeutet $x - x' \in U$ und $y - y' \in U$. Es folgt

$$(x + y) - (x' + y') = (x - x') + (y - y') \in U,$$

d. h. $[x + y] = [x' + y']$. Also ist die Abbilung $+\colon A/U \times A/U \to A/U$ wohldefiniert.

Alle Eigenschaften, die in der Definition 1.1.2 einer Gruppe gefordert werden, vererben sich von A auf A/U: Assoziativität gilt wegen $([x]+[y])+[z] = [x+y]+[z] = [(x+y)+z] = [x+(y+z)] = [x]+[y+z] = [x]+([y]+[z])$. Wegen $[0]+[x] = [0+x] = [x] = [x+0] = [x]+[0]$ ist $[0]$ das neutrale Element. Zu $[x]$ ist $[-x]$ invers. Also ist A/U eine Gruppe. Sie ist wegen $[x] + [y] = [x + y] = [y + x] = [y] + [x]$ abelsch. □

Beispiel 1.6.4. Nach Satz 1.6.3 ist $\mathbb{Z}/n\mathbb{Z}$ in naheliegender Weise eine abelsche Gruppe.

In $\mathbb{Z}/15\mathbb{Z} = \{[0], \dots, [14]\}$ gelten beispielsweise $[10] + [7] = [17] = [2] = [-13]$ und $[1] + [1] + [1] = [3]$.

1.6.5. Gegeben $x \in \mathbb{Z}$ schreibt man oft nachlässig x statt $[x]$, wenn klar ist, dass es sich um ein Element von $\mathbb{Z}/n\mathbb{Z}$ handelt.

1.6.6 (Clock arithmetic)**.** Man kann sich die Gruppe $\mathbb{Z}/n\mathbb{Z}$ als Uhr mit n Stunden vorstellen. Auf der 13-Stunden-Uhr $\mathbb{Z}/13\mathbb{Z}$ in Abbildung 3 gelten etwa $3 + 5 = 8$ und $10 + 7 = 4$ oder genauer $[3]+[5] = [8]$ und $[10]+[7] = [4]$. Schon seit Kindertagen rechnen Sie also in Gruppen der Form $\mathbb{Z}/12\mathbb{Z}$ oder $\mathbb{Z}/24\mathbb{Z}$ oder $\mathbb{Z}/365\mathbb{Z}$.

1.7. **Zyklische Gruppen.**

Definition 1.7.1. Eine Gruppe G heißt **zyklisch**, wenn sie von einem Element erzeugt wird, wenn es also ein $g \in G$ mit $G = \langle g \rangle$ gibt. Beachte $\langle g \rangle = \{\dots, g^{-2}, g^{-1}, e, g, g^2, \dots\}$.

1.7.2. Jede zyklische Gruppe ist abelsch, denn es gilt $g^a g^b = g^{a+b} = g^{b+a} = g^b g^a$ für alle $a, b \in \mathbb{Z}$.

Beispiel 1.7.3. (a) $(\mathbb{Z}, +) = \langle 1 \rangle$ ist zyklisch.

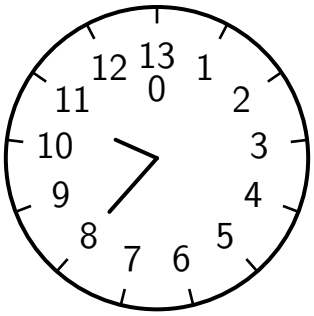


ABBILDUNG 3. Uhr mit 13 Stunden

(b) $(\mathbb{Z}/n\mathbb{Z}, +) = \langle [1] \rangle$ ist zyklisch.
(c) $\mathbb{Z}/6\mathbb{Z}$ ist zyklisch und nicht nur von $[1]$, sondern auch von $[5]$ erzeugt. Das Element $[4]$ ist aber kein Erzeuger, denn $\langle [4] \rangle = \{[4], [2], [0]\}$. Ebenso sind $[2]$ und $[0]$ keine Erzeuger.
(d) Sei G eine Gruppe und $g \in G$. Dann ist die von g erzeugte Untergruppe $\langle g \rangle$ zyklisch.
(e) Die symmetrische Gruppe S_3 ist nicht zyklisch. Dies folgt aus 1.7.2, aber man kann es sich auch explizit überlegen: Die von $e = \mathrm{id}$ erzeugte Untergruppe ist $\{e\}$, jedes der Elemente (12), (23) und (13) erzeugt eine Untergruppe, die aus genau zwei Elementen besteht, und jedes der Elemente (123) und (132) erzeugt die Untergruppe $\{\mathrm{id}, (123), (132)\}$.

1.7.4. Motiviere den folgenden Satz durch Bilder.

Satz 1.7.5 (Klassifikation zyklischer Gruppen bis auf Isomorphie)**.** *Jede zyklische Gruppe ist zu genau einer Gruppe aus der folgenden Liste zyklischer Gruppen isomorph:*

$$\mathbb{Z}, \quad \mathbb{Z}/1\mathbb{Z} = \{0\}, \quad \mathbb{Z}/2\mathbb{Z}, \quad \mathbb{Z}/3\mathbb{Z}, \quad \mathbb{Z}/4\mathbb{Z}, \quad \mathbb{Z}/5\mathbb{Z}, \quad \dots.$$

Genauer gilt: Sei G eine zyklische Gruppe mit Erzeuger $g \in G$.

(a) Ist G endlich, so ist die Abbildung

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} &\xrightarrow{\sim} G, \\ [a] &\mapsto g^a, \end{aligned} \tag{1.7.1}$$

ein (wohldefinierter) Isomorphismus von Gruppen, wobei $n := |G|$. Insbesondere gilt $g^{|G|} = e$.

(b) Ist G unendlich, so ist die Abbildung

$$\begin{aligned} \mathbb{Z} &\xrightarrow{\sim} G, \\ a &\mapsto g^a, \end{aligned} \tag{1.7.2}$$

ein (wohldefinierter) Isomorphismus von Gruppen.

Beweis. Wegen $G = \langle g \rangle$ besteht G aus den Elementen $\dots, g^{-2}, g^{-1}, g^0 = e, g^1 = g, g^2, \dots$. Wir unterscheiden zwei Fälle:

(i) Falls all diese Elemente verschieden sind, ist G unendlich, und wegen $g^{a+b} = g^a g^b$ (siehe 1.1.9) ist klar, dass (1.7.2) ein Gruppenisomorphismus ist.

(ii) Nun nehmen wir an, dass $g^i = g^j$ für zwei ganze Zahlen $i < j$ gilt. Dann folgt $e = g^{j-i}$. Wegen $j - i > 0$ können wir $n > 0$ minimal mit $g^n = e$ wählen.

Multipliziert man ein beliebiges Element der Menge

$$\{e = g^0 = g^n, g^1, g^2, \dots, g^{n-1}\}$$

von links oder rechts mit g oder g^{-1}, so ist das Resultat wieder in dieser Menge. Dies zeigt, dass diese Menge mit $\langle g \rangle$ übereinstimmt, dass also

$$G = \langle g \rangle = \{e = g^0 = g^n, g^1, g^2, \dots, g^{n-1}\}$$

gilt. Die Elemente $e = g^0, g^1, g^2, g^{n-1}$ sind alle verschieden, denn sonst wäre $g^s = g^t$ für $0 \leq s < t \leq n-1$ und damit $e = g^{t-s}$ mit $0 < t - s \leq n - 1$ im Widerspruch zur Minimalität von n.

Also besteht G genau aus $|G| = n$ Elementen und es gilt $g^{|G|} = g^n = e$.

Die Abbildung (1.7.1) ist wohldefiniert, denn für beliebige $a, x \in \mathbb{Z}$ gilt $g^{a+xn} = g^a g^{xn} = g^a (g^n)^x = g^a e^x = g^a$. Sie ist bijektiv, denn die linke Menge besteht genau aus den Elementen $[0], [1], \dots, [n-1]$, und diese gehen auf die Elemente $e, g, \dots, g^{n-1}$. Wegen $g^{a+b} = g^a g^b$ ist sie ein Gruppenhomomorphismus.

Es ist nun klar, dass der Fall (ii) genau dann eintritt, wenn G endlich ist.

Offensichtlich sind alle Gruppen in unserer Liste zyklisch, und wir haben soeben bewiesen, dass jede zyklische Gruppe isomorph zu einer Gruppe aus dieser Liste ist. Da alle Gruppen in der Liste unterschiedlich viele Elemente haben, sind sie paarweise nicht isomorph. □

Korollar 1.7.6. *Sei G eine endliche Gruppe und $g \in G$. Dann gilt $g^{|G|} = e$.*

Beweis. Sei $H = \langle g \rangle \subset G$ die von g erzeugte Untergruppe. Sie ist zyklisch und endlich. Nach Satz 1.7.5 gilt $g^{|H|} = e$. Nach dem Satz von Lagrange 1.5.12 ist $|H|$ ein Teiler von $|G|$, es gilt also $|G| = |H| \cdot a$ für ein $a \in \mathbb{N}$ (nämlich für $a = |G/H|$). Wir erhalten $g^{|G|} = g^{|H| \cdot a} = (g^{|H|})^a = e^a = e$. □

Ende 3. Vorlesung am 28.04.2020

2. Ringe

2.1. Definitionen und Beispiele.

Definition 2.1.1. Ein **Ring** ist eine Menge R mit zwei Abbildungen

$$+ \colon R \times R \to R, \quad (a, b) \mapsto a + b, \qquad (\textbf{Addition})$$
$$\cdot \colon R \times R \to R, \quad (a, b) \mapsto a \cdot b, \qquad (\textbf{Multiplikation})$$

so dass die folgenden Bedingungen erfüllt sind:

(a) $(R, +)$ ist eine abelsche Gruppe. Das neutrale Element wird mit 0 bezeichnet. Das inverse Element von $a \in R$ wird mit $-a$ bezeichnet.
(b) Für alle $a, b, c \in R$ gilt $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität der Multiplikation).
(c) Es existiert ein Element $1 \in R$ mit $1 \cdot a = a = a \cdot 1$ für alle $a \in R$ (Existenz eines neutrale Element bzgl. der Multiplikation). Ein solches Element ist eindeutig[6].

[6] Ist $1' \in R$ ein weiteres Element mit $1' \cdot a = a = a \cdot 1'$ für alle $a \in R$, so folgt $1 = 1 \cdot 1' = 1'$.

(d) Für alle $a, b, c \in R$ gelten

$$a \cdot (b + c) = a \cdot b + a \cdot c,$$
$$(a + b) \cdot c = a \cdot c + b \cdot c$$

(Distributivität).

Ein Ring $(R, +, \cdot)$ heißt genau dann **kommutativ**, wenn $a \cdot b = b \cdot a$ für alle $a, b \in R$ gilt.

2.1.2. In diesem Kapitel 2 betrachten wir meist kommutative Ringe.

2.1.3. Ist a ein Element eines Ringes und $n \in \mathbb{N}$, so schreiben wir abkürzend $a^n = a \cdot \dots \cdot a$ (n Faktoren).

2.1.4. Wir schreiben meist ab statt $a \cdot b$. In dieser Notation schreiben sich die Forderungen in der Distributivität als $a(b + c) = ab + ac$ $(a + b)c = ac + bc$.

Beispiele 2.1.5.
(a) Jeder Körper (wie in der Analysis definiert) ist ein kommutativer Ring. Zum Beispiel sind die Körper $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ und $\mathbb{F}_2 = \{0, 1\}$ kommutative Ringe.
(b) $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring (aber kein Körper).
(c) Ist M eine beliebige Menge, etwa $\mathbb{R}$, so ist die Menge aller Funktionen $M \to \mathbb{R}$ ein Ring, indem wir $f + g$ und $f \cdot g$ „punktweise“ durch $(f + g)(m) = f(m) + g(m)$ und $(f \cdot g)(m) = f(m) \cdot g(m)$ definieren. Nullelement und Einselement sind die konstanten Funktionen $0\colon M \to \mathbb{R}$, $m \mapsto 0$ und $1\colon M \to \mathbb{R}$, $m \mapsto 1$.
(d) $R = \{0\}$ ist ein Ring, genannt *Nullring* (für Addition und Multiplikation gibt es nur eine Möglichkeit). Im Nullring gilt $1 = 0$.

Lemma 2.1.6. *Seien R ein Ring und $a, b \in R$ beliebige Elemente. Dann gelten*

- $0a = a0 = 0$*;*
- $(-1)a = -a = a(-1)$*;*
- $(-1)(-1) = 1$ *(„Minus mal Minus ist Plus“);*
- $(-a)(-b) = ab$*.*

Beweis. Aus $0a = (0 + 0)a = 0a + 0a$ folgt per Addition von $-(0a)$, dass $0 = 0a$. Analog sieht man $0 = a0$.

Aus $(-1)a + a = (-1)a + 1a = (-1 + 1)a = 0a = 0$ folgt, dass $(-1)a = -a$. Aus $a(-1) + a = a(-1) + a1 = a((-1) + 1) = a0 = 0$ folgt analog $a(-1) = -a$. Für $a = -1$ liefert dies $(-1)(-1) = -(-1) = 1$ (zweite Gleichheit nach Lemma 1.1.4).

Die vorigen beiden Punkte liefern $(-a)(-b) = (-1)a(-1)b = (-1)(-1)ab = 1ab = ab$. □

Satz 2.1.7. *Sei $n \in \mathbb{N}$. Die abelsche Gruppe $(\mathbb{Z}/n\mathbb{Z}, +)$ aus Beispiel 1.6.4 wird mit der (wohldefinierten) Multiplikation*

$$[x] \cdot [y] := [x \cdot y] \qquad \text{für } x, y \in \mathbb{Z}$$

ein kommutativer Ring.

Beweis. Seien $x, y, x', y' \in \mathbb{Z}$ mit $[x] = [x']$ und $[y] = [y']$. Dies bedeutet $x - x' \in n\mathbb{Z}$ und $y - y' \in n\mathbb{Z}$. Es folgt

$$x \cdot y - x' \cdot y' = x \cdot y - x \cdot y' + x \cdot y' - x' \cdot y' = x \cdot (y - y') + (x - x') \cdot y' \in n\mathbb{Z}.$$

Also ist die Abbildung $\cdot\colon \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ wohldefiniert. Alle Eigenschaften, die in der Definition 2.1.1 eines kommutativen Rings gefordert werden, rechnet man leicht nach. Das Einselement von $\mathbb{Z}/n\mathbb{Z}$ ist $[1]$. □

2.1.8. Uhrzeiten der n-Stunden-Uhr 1.6.6 kann man also auch multiplizieren.

Beispiel 2.1.9. In $\mathbb{Z}/15\mathbb{Z} = \{[0], \ldots, [14]\}$ gelten $[6] \cdot [10] = [60] = [0]$ und $[13] = [-2]$, also $[13]^{10} = [-2]^{10} = [(-2)^{10}] = [1024] = [124] = [34] = [4]$.

Beispiel 2.1.10. Die Additions- und Multiplikationstafel von $\mathbb{Z}/4\mathbb{Z} := \{[0], [1], [2], [3]\}$. sind wie folgt gegeben.

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Wir schreiben hier abkürzend 0 statt [0] usw. In dieser Schreibweise addiert und multipliziert man also wie in $\mathbb{Z}$ und nimmt dann den Rest bei Division durch 4.

Der Ring $\mathbb{Z}/4\mathbb{Z}$ ist kein Körper, denn 2 hat kein multiplikatives Inverses.

2.1.11. Zahlentheoretische Schreibweise: Für $x, y \in \mathbb{Z}$ schreibe

$$x \equiv y \pmod{n}$$

und sage „x kongruent (zu) y modulo n“ genau dann, wenn $[x] = [y]$ gilt. Diese Sprechweise kommt daher, dass $[x] = [y]$ äquivalent zu der Aussage ist, dass x und y bei Division durch n denselben Rest haben[7]. Division mit Rest wird in Satz 2.5.7 genau erklärt.

Beispiel 2.1.12. In dieser Schreibweise gelten $10+7 \equiv 2 \pmod{15}$ und $6 \cdot 10 \equiv 0 \pmod{15}$ und $13 \equiv -2 \pmod{15}$, also $13^3 \equiv -8 \equiv 7 \pmod{15}$.

2.2. Einheiten, Teiler, Nullteiler, Integritätsbereiche.

Definition 2.2.1. Sei R ein Ring. Ein Element $a \in R$ heißt genau dann **Einheit** oder **invertierbar**, wenn ein $b \in R$ existiert mit $ab = 1 = ba$. Ein solches Element b ist eindeutig (Beweis wie Eindeutigkeit des Inversen bei Gruppen) und wird als a^{-1} oder $1/a$ oder $\frac{1}{a}$ notiert. Wir schreiben

$$R^\times := \{a \in R \mid a \text{ ist Einheit}\}$$

für die Menge der Einheiten von R.

2.2.2. Die Menge $R^\times$ mit Multiplikation als Verknüpfung ist eine Gruppe: Für $a, b \in R^\times$ ist ab wegen $(ab)(b^{-1}a^{-1}) = 1 = (b^{-1}a^{-1})(ab)$ invertierbar mit Inversem $(ab)^{-1} = b^{-1}a^{-1}$. Assoziativität ist klar, 1 ist neutrales Element. Existenz von Inversen: Sei $a \in R^\times$ beliebig. Dann gibt es nach Definition einer Einheit ein Element $a^{-1} \in R$ mit $aa^{-1} = 1 = a^{-1}a$. Also gilt $a^{-1} \in R^\times$ und dieses Element ist das gesuchte Inverse.

Ist R kommutativ, so ist $R^\times$ eine kommutative Gruppe.

Beispiele 2.2.3.
(a) $(\mathbb{Z}/4\mathbb{Z})^\times = \{1, 3\}$.
(b) $\mathbb{Z}^\times = \{-1, +1\}$.
(c) Ist K ein Körper, so gilt $K^\times = K \setminus \{0\}$.
Genauer gilt: Ein Ring R ist genau dann ein Körper, wenn er kommutativ ist und $R^\times = R \setminus \{0\}$ gilt.

[7] Begründung: Gelte $x = an + r$ und $y = bn + s$ mit $0 \le r < n$ und $0 \le s < n$ und $a, b \in \mathbb{Z}$. Wegen $x - y = (a - b)n + r - s$ ist $x - y$ genau dann durch n teilbar, wenn $r - s$ durch n teilbar ist. Wegen $-n < r - s < n$ ist dies aber genau dann der Fall, wenn $r - s = 0$ gilt.

Definition 2.2.4. Sei R ein kommutativer Ring.

(a) Seien $a, b \in R$ Elemente. Man sagt a **teilt** b oder a **ist Teiler von** b oder b **ist Vielfaches von** a und schreibt $a \mid b$, wenn ein $c \in R$ mit $ac = b$ existiert. Die Negation dieser Aussage wird als $a \nmid b$ notiert.

(b) Ein Element $a \in R$ heißt genau dann **Nullteiler**, wenn es ein Element $0 \neq b \in R$ mit $ab = 0$ gibt.

(c) Ein kommutativer Ring R heißt genau dann **nullteilerfrei**, wenn er außer der Null keinen Nullteiler enthält: Für alle $a, b \in R$ gilt: Aus $ab = 0$ folgt $a = 0$ oder $b = 0$. (Äquivalent: Aus $a \neq 0$ und $b \neq 0$ folgt $ab \neq 0$.)

(d) Ein **Integritätsbereich** ist ein nullteilerfreier kommutativer Ring mit[8] $1 \neq 0$.

Beispiel 2.2.5. (a) Jeder Körper ist ein Integritätsbereich.

Beweis. Sei K ein Körper. Dann ist K sicher ein kommutativer Ring mit $1 \neq 0$. Gelte $ab = 0$ für zwei Elemente $a, b \in K$. Ist $a = 0$, so sind wir fertig, sonst ist a multiplikativ invertierbar und es folgt $0 = a^{-1}0 = a^{-1}ab = 1b = b$. □

(b) Der Ring $\mathbb{Z}/4\mathbb{Z}$ ist kein Integritätsbereich: Das Element $2 \neq 0$ ist wegen $2 \cdot 2 = 0$ ein Nullteiler.

2.2.6 (Kürzen in Integritätsbereichen)**.** Es gilt

$$ab = ac \text{ und } a \neq 0 \quad \Longrightarrow \quad b = c$$

für beliebige Elemente a, b, c eines Integritätsbereichs. Aus $a(b - c) = 0$ folgt nämlich wegen der Nullteilerfreiheit $b - c = 0$.

2.2.7. Sei a, b Elemente eines Integritätsbereichs R. Ist a ein Teiler von b, so gibt es genau ein $c \in R$ mit $ac = b$; dies folgt aus 2.2.6.

2.3. Unterringe und Ringhomomorphismen.

Definition 2.3.1. Sei R ein Ring. Eine Teilmenge $S \subset R$ heißt **Unterring**, wenn sie unter Addition und Multiplikation abgeschlossen ist (aus $s, t \in S$ folgen also $s + t, st \in S$), mit jedem Element sein additives Inverses enthält (aus $s \in S$ folgt $-s \in S$) und die Eins von R enthält (also $1_R \in S$).

Dann ist S in offensichtlicher Weise selbst ein Ring, was die Terminologie rechtfertigt.

Beispiele 2.3.2. Jede der Inklusionen $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ ist die Inklusion eines Unterrings in einen Ring.

Beispiel 2.3.3. Die Teilmenge $\mathbb{Z}[i] := \{a + bi \in \mathbb{C} \mid a, b \in \mathbb{Z}\} \subset \mathbb{C}$ ist ein Unterring von $\mathbb{C}$.

2.3.4. Unterringe von Integritätsbereichen sind Integritätsbereiche. Dies zeigt, dass die Ringe $\mathbb{Z}$ und $\mathbb{Z}[i]$ in den obigen Beispielen Integritätsbereiche sind.

Definition 2.3.5. Seien R und S Ringe. Ein **Ringhomomorphismus** oder **Homomorphismus von Ringen** von R nach S ist eine Abbildung $\varphi\colon R \to S$ mit $\varphi(x+y) = \varphi(x)+\varphi(y)$ und $\varphi(xy) = \varphi(x)\varphi(y)$ für alle $x, y \in R$ und $\varphi(1) = 1$. Ein **Ringisomorphismus** ist ein bijektiver Ringhomomorpismus (oder äquivalent, ein Ringhomomorpismus $\varphi\colon R \to S$, für den es einen Ringhomomorpismus $\psi\colon S \to R$ mit $\psi \circ \varphi = \mathrm{id}_R$ und $\varphi \circ \psi = \mathrm{id}_S$ gibt, vgl.

[8] Gilt in einem Ring $1 = 0$, so besteht dieser Ring aus genau einem Element (ist also der Nullring), denn für jedes beliebige Element x gilt dann $x = x \cdot 1 = x \cdot 0 = 0$.

Bemerkung 1.3.9). Wir notieren einen Ringisomorphismus als $\varphi\colon R \xrightarrow{\sim} S$. Zwei Ringe heißen **isomorph**, wenn es einen Ringisomorphismus zwischen ihnen gibt.

Beispiel 2.3.6. Die Abbildung $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $x \mapsto [x]$, ist ein Ringhomomorphismus.

Beispiel 2.3.7. Ist $S \subset R$ ein Unterring eines Ringes, so ist die injektive Inklusionsabbildung $S \to R$, $s \mapsto s$ ein Ringhomomorphismus.

2.4. Polynomringe in einer Variablen über Körpern.

2.4.1. Im gesamten Abschnitt 2.4 sei K ein Körper, etwa $\mathbb{R}$ oder $\mathbb{C}$.[9]

Definition 2.4.2. Ein **Polynom in einer Variablen X mit Koeffizienten in K** (oder **Polynom in X über K**) ist ein formaler Ausdruck der Form

$$f = f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_2 X^2 + a_1 X + a_0$$

für $n \in \mathbb{N}$ und $a_0, a_1, \ldots, a_n \in K$.[10] [11] Die $a_0, \ldots, a_n$ heißen **Koeffizienten** des Polynoms, genauer ist a_i der Koeffizient von X^i.

Die Menge $K[X]$ aller Polynome in X über K wird mit der (koeffizientenweisen) Addition

$$\sum_{i=0}^{m} a_i X^i + \sum_{j=0}^{n} b_j X^j := \sum_{k=0}^{\max(n,m)} (a_k + b_k) X^k,$$

wobei „nicht vorhandene“ Koeffizienten als Null verstanden werden (d. h. $a_k := 0$ für $k > m$ und $b_k := 0$ für $k > n$), und der Multiplikation

$$\Big(\sum_{i=0}^{m} a_i X^i\Big) \cdot \Big(\sum_{j=0}^{n} b_j X^j\Big) := \sum_{i=0}^{m} \sum_{j=0}^{n} a_i b_j X^{i+j} = \sum_{k=0}^{m+n} \Big(\sum_{\substack{i,j \in \mathbb{N}, \\ i+j=k}} a_i b_j \Big) X^k$$

ein kommutativer Ring[12], der als **Polynomring über K in der Variablen X** bezeichnet wird.

2.4.3. Oft werden Elemente $f \in K[X]$ wie oben als $f(X)$ notiert, um anzudeuten, dass sie von der Variablen X abhängen. Das Nullelement 0 in $K[X]$ wird auch als **Nullpolynom** bezeichnet.

Ein Polynom der Form $f = a_0$, für $a_0 \in K$, heißt **konstantes** Polynom. Wir fassen K als Unterring $K \subset K[X]$ auf, indem wir Elemente von K als konstante Polynome auffassen.

Beispiel 2.4.4. Im Polynomring $\mathbb{R}[X]$ gelten

$$(X^2 + 2X + 1) + (X^3 - X) = X^3 + X^2 + X + 1,$$
$$(X^2 + X)(X - 1) = X^3 - X.$$

[9] Man kann allgemeiner Polynomringe über kommutativen Ringen betrachten, jedoch ist dies für diese Vorlesung nicht von Bedeutung.

[10] In einem solchen Ausdruck lassen wir oft Unterausdrücke der Form $+0X^i$ weg und schreiben X^i statt $1X^i$ und X statt X^1, beispielsweise sind $2X^3 + X + 1$ und $0X^4 + 2X^3 + 0X^2 + 1X^1 + 1$ verschiedene Schreibweisen für dasselbe Polynom.

[11] Korrekter, aber vollkommen unintuitiv, ist ein Polynom eine Folge $(a_0, a_1, a_2, \ldots)$ reeller Zahlen, für die es ein $N \in \mathbb{N}$ mit $a_i = 0$ für alle $i \geq N$ gibt.

[12] Wir überlassen dem Leser den Beweis.

2.4.5 (Polynome als polynomiale Abbildungen)**.** Sei $f = f(X) = a_n X^n + \cdots + a_1 X + a_0 \in K[X]$ ein Polynom. Dann liefert f eine Abbildung[13]

$$\begin{aligned} K &\to K, \\ z &\mapsto f(z) := a_n z^n + \cdots + a_1 z + a_0. \end{aligned}$$

Jede Abbildung $K \to K$ von dieser Form heißt **polynomiale Abbildung**. Ist $f = a_0$ ein konstantes Polynom, so ist die zugeordnete polynomiale Abbildung konstant, denn sie nimmt überall den Wert a_0 an.

Definition 2.4.6. Sei $f(X) = \sum_{i=0}^{n} a_i X^i \in K[X]$ ein Polynom. Eine **Nullstelle von f (in K)** ist ein Element $z \in K$ mit $f(z) = 0$ gilt.

Beispiele 2.4.7.
(a) Das Polynom $X^2 \in \mathbb{R}[X]$ hat genau eine reelle Nullstelle.
(b) Das Polynom $X^2 + 1 \in \mathbb{R}[X]$ hat keine reelle Nullstelle.
(c) Das Polynom $X^2 + 1 \in \mathbb{C}[X]$ hat zwei komplexe Nullstellen, nämlich $\pm i$.
(d) Das Polynom $X^3 + X + 1 \in \mathbb{F}_2[X]$ hat keine Nullstelle in $\mathbb{F}_2$.

Definition 2.4.8. Sei $f(X) = a_n X^n + \cdots + a_1 X + a_0 \in K[X]$ ein Polynom.
(a) Der **Grad von** f (englisch *degree*) ist definiert durch

$$\deg(f) := \begin{cases} \max\{i \mid a_i \neq 0\} & \text{falls } f \neq 0, \\ -\infty & \text{falls } f = 0. \end{cases}$$

Es gilt also $\deg(f) \in \mathbb{N} \cup \{-\infty\}$.
(b) Gilt $f \neq 0$, so ist der **Leitkoeffizient von** f das Element $a_{\deg(f)} \neq 0$.
(c) Ein Polynom f heißt **normiert**, falls $f \neq 0$ gilt und sein Leitkoeffizient 1 ist.

Beispiel 2.4.9. Das Polynom $2X^3 + 7X + 1$ hat Grad 3 und ist nicht normiert. Das Polynom $X^3 + X$ ist normiert und hat Grad $\deg(X^3 + X) = 3$.

Ende 4. Vorlesung am 30.04.2020

2.4.10. Für das Symbol $-\infty$ sind im Hinblick auf das folgende Lemma die folgenden Konventionen sinnvoll, wobei $n \in \mathbb{N}$:

$$\begin{aligned} -\infty < n, \qquad & -\infty \leq -\infty \leq n, \\ \max(-\infty, n) = n = \max(n, -\infty), \qquad & \max(-\infty, -\infty) = -\infty, \\ (-\infty) + n = -\infty = n + (-\infty), \qquad & (-\infty) + (-\infty) = -\infty. \end{aligned}$$

Proposition 2.4.11. *Sei K ein Körper. Für beliebige Polynome $f, g \in K[X]$ gelten:*
(a) $\deg(f + g) \leq \max(\deg(f), \deg(g))$*; im Fall* $\deg(f) \neq \deg(g)$ *gilt* $\deg(f + g) = \max(\deg(f), \deg(g))$.
(b) $\deg(f \cdot g) = \deg(f) + \deg(g)$.
Ist insbesondere f ein Teiler[14] *von $q \neq 0$ in $K[X]$, so gilt* $\deg(f) \leq \deg(q)$.

Beweis. Ist f oder g das Nullpolynom, so sind alle Behauptungen trivial. Gelte $f \neq 0 \neq g$. Schreibe $f = a_m X^m + \ldots$ und $g = b_n X^n + \ldots$ wobei $m = \deg(f)$ und $n = \deg(g)$. Die erste Behauptung ist dann klar nach Definition der Addition.

[13] Diese Abbildung ist kein Homomorphismus von Ringen, sie ist aber „K-linear“.
[14] Definition 2.2.4 nachträglich ergänzt.

Nach Definition der Multiplikation hat fg Grad $\leq n+m$. Der Koeffizient von fg bei X^{n+m} ist $a_n b_m$, und dieses Produkt verschwindet nicht, da K als Körper insbesondere ein Integritätsbereich ist. Also hat fg Grad $n+m$. □

Beispiele 2.4.12. (a) Sei $g(X) = X^3 + X$ in $\mathbb{R}[X]$.
- Für $f(X) = 2X^3$ hat hat $f+g = 3X^3 + X$ Grad $3 = \max(1,3) = 3$.
- Für $f(X) = -X^3$ hat $f+g = X$ Grad $1 < \max(1,3) = 3$.
- Für $f(X) = X$ hat $f+g = X^3 + 2X$ Grad $3 = \max(1,3)$.

(b) Für $f(X) = X^3 + 1$ und $g(X) = X^3 + X$ in $\mathbb{F}_2[X]$ haben $f+g = X+1$ Grad $1 < \max(3,3) = 3$ und $fg = X^6 + X^4 + X^3 + X$ Grad $6 = 3+3$.

Satz 2.4.13. *Sei K ein Körper. Dann ist $K[X]$ ein Integritätsbereich.*

Beweis. Wir wissen, dass $K[X]$ kommutativ ist; sicherlich gilt $1 \neq 0$. Seien $f, g \in K[X]$ mit $fg = 0$. Dann gilt $-\infty = \deg(0) = \deg(fg) = \deg(f) + \deg(g)$ nach Proposition 2.4.11.(b), also $\deg(f) = -\infty$ oder $\deg(g) = -\infty$. Wir folgern $f = 0$ oder $g = 0$, denn nur das Nullpolynom hat Grad $-\infty$. □

Proposition 2.4.14. *Sei K ein Körper. Dann sind die Einheiten von $K[X]$ genau die Einheiten von K, in Formeln $K[X]^\times = K^\times = K \setminus \{0\}$.*

Beweis. Ist $a \in K^\times$ eine Einheit, so ist klar, dass a als konstantes Polynom invertierbar in $K[X]$ ist. Sei umgekehrt $f \in K[X]$ eine Einheit. Dann existiert ein $g \in K[X]$ mit $fg = 1$. Es folgt $0 = \deg(1) = \deg(fg) = \deg(f) + \deg(g)$ und somit $\deg(f) = 0 = \deg(g)$, also $f, g \in K$. Dann gilt aber schon $fg = 1$ im Unterring K und f ist eine Einheit in K. □

Lemma 2.4.15 (Auswerten von Polynomen)**.** *Seien $z \in K$. Dann ist **Auswerten bei** z (oder **Evaluation bei** z)*

$$\begin{aligned} \mathrm{ev}_z \colon K[X] &\to K, \\ f(X) = \sum_{i=0}^{m} a_i X^i &\mapsto f(z) = \sum_{i=0}^{m} a_i z^i, \end{aligned}$$

ein Homomorphismus von Ringen, es gelten also $f(z) + g(z) = (f+g)(z)$ und $(fg)(z) = f(z)g(z)$ für alle $f, g \in K[X]$ und $1(z) = 1$.

Beweis. Einfaches Nachrechnen: Seien $f(X) = \sum_{i=0}^{m} a_i X^i$ und $g(X) = \sum_{i=0}^{n} b_i X^i$ Elemente von $K[X]$. Dann gilt

$$\mathrm{ev}_z(f)\mathrm{ev}_z(g) = f(z)g(z) = \Big(\sum_{i=0}^{m} a_i z^i\Big) \cdot \Big(\sum_{j=0}^{n} b_j z^j\Big) = \sum_{k=0}^{m+n} \Big(\sum_{\substack{i,j\in\mathbb{N},\\ i+j=k}} a_i b_j\Big) z^k = (fg)(z) = \mathrm{ev}_z(fg).$$

Dies besagt, dass Auswerten mit Multiplikation verträglich ist. Analog zeigt man, dass es mit Addition verträglich ist. Außerdem gilt $\mathrm{ev}_z(1) = 1(z) = 1$. □

Satz 2.4.16 (Polynomdivision - Division mit Rest in $K[X]$)**.** *Sei K ein Körper. Seien $f, g \in K[X]$ Polynome mit $g \neq 0$. Dann existieren eindeutig bestimmte Polynome $q, r \in K[X]$ mit*

$$f = qg + r \text{ und } \deg(r) < \deg(g).$$

Beweis. Seien $0 \neq g \in K[X]$ fixiert und $n := \deg(g) \geq 0$.

Existenz der Darstellung: Wir zeigen die Aussage per Induktion über den Grad von f.

Induktionsanfang: Die Aussage gilt für alle Polynome $f \in K[X]$ vom Grad $\deg(f) < n$. In der Tat, für jedes solche Polynom ist $f = 0g + f$ wegen $\deg(f) < n = \deg(g)$ eine gesuchte Darstellung.

Sei nun $m := \deg(f) \geq n$ und sei die Aussage bereits für alle Polynome vom Grad $< m$ bewiesen. Seien a_m der Leitkoeffizient von f und b_n der Leitkoeffizient von g. Wegen $m \geq n$ gilt $X^{m-n} \in K[X]$ und wir können das Polynom

$$f' := f - \frac{a_m}{b_n} X^{m-n} g$$

betrachten.[15] Da $\frac{a_m}{b_n} X^{m-n} g$ ein Polynom vom Grad m mit Leitkoeffizient a_m ist, was auch für f zutrifft, gilt $\deg(f') < m = \deg(f)$. Per Induktion finden wir $q', r \in K[X]$ mit

$$f' = q'g + r \text{ und } \deg(r) < \deg(g).$$

Wir erhalten wie gewünscht

$$f = f' + \frac{a_m}{b_n} X^{m-n} g = \underbrace{(q' + \frac{a_m}{b_n} X^{m-n})}_{=:q} g + r \text{ mit } \deg(r) < \deg(g).$$

Eindeutigkeit der Darstellung: Gelte $f = qg + r = q'g + r'$ mit $\deg(r) < \deg(g)$ und $\deg(r') < \deg(g)$. Es folgt

$$r' - r = (q - q')g.$$

Gilt $q = q'$, so folgt $r = r'$ und wir sind fertig. Gelte $q \neq q'$, also insbesondere $\deg(q-q') \geq 0$. Wir erhalten

$$\deg(r'-r) \leq \max(\deg(r), \deg(r')) < \underbrace{\deg(g)}_{\geq 0} \leq \deg(g)+\deg(q-q') = \deg(g(q-q')) = \deg(r'-r).$$

Dies zeigt $\deg(r' - r) \geq 0$ und $\deg(r' - r) < \deg(r' - r)$, was nicht sein kann. □

Algorithmus 2.4.17 (Polynomdivison)**.** Der Beweis von Satz 2.4.16 ist konstruktiv und liefert das möglicherweise aus der Schule bekannten Verfahren zur Polynomdivision. Wir illustrieren dies in Beispiel 2.4.18.

Beispiel 2.4.18 (Polynomdivision)**.** Betrachte die Polynome $f = f(X) = X^5 + X^4 + 2X^3 - X^2 - X - 2$ und $g = g(X) = X^3 + 4X^2 + 5X + 6$ in $\mathbb{R}[X]$. Das folgende Rechenschema (je nach Geschmack eventuell mit dem Summanden -2 in der dritten Zeile ergänzt) illustriert die Polynomdivision von f durch g.

$$\begin{array}{rrrrrr}
X^5 & +X^4 & +2X^3 & -X^2 & -X & -2 = \left(X^3 + 4X^2 + 5X + 6\right)\left(X^2 - 3X + 9\right) - 28X^2 - 28X - 56 \\
\cline{1-4}
-X^5 & -4X^4 & -5X^3 & -6X^2 & & \\
\cline{1-4}
 & -3X^4 & -3X^3 & -7X^2 & -X & \\
 & 3X^4 & +12X^3 & +15X^2 & +18X & \\
\cline{2-5}
 & & 9X^3 & +8X^2 & +17X & -2 \\
 & & -9X^3 & -36X^2 & -45X & -54 \\
\cline{2-5}
 & & & -28X^2 & -28X & -56
\end{array}$$

[15] Hier ist f' nur ein Name für ein Polynom. Bitte nicht an die Ableitung im Sinne der Analysis denken.

Proposition 2.4.19 (Abspalten von Nullstellen)**.** *Seien K ein Körper, $f \in K[X]$ ein Polynom und $z \in K$. Dann sind äquivalent:*

- *z ist eine Nullstelle von f, d. h. $f(z) = 0$;*
- *$X - z$ ist Teiler von f, d. h. es gibt ein (eindeutiges) Polynom $q \in K[X]$ mit $f = (X - z)q$.*

Beweis. $\Rightarrow$: Sei z eine Nullstelle von f. Per Division mit Rest (Satz 2.4.16) gilt $f = (X - z)q + r$ für geeignete $q, r \in K[X]$ mit $\deg(r) < \deg(X - z) = 1$. Also ist $r = a_0 \in K$ ein konstantes Polynom. Werten wir $f = f(X) = (X - z)q(X) + r$ bei z aus, so erhalten wir

$$0 = f(z) = (z - z)q(z) + r(z) = 0 + a_0$$

in K, wobei die zweite Gleichheit genaugenommen verwendet, dass Auswerten bei z ein Ringhomomorphismus ist (Lemma 2.4.15). Also gilt $r = a_0 = 0$ und damit $f = (X - z)q$.

$\Leftarrow$: Wertet man $f = (X - z)q$ bei z aus, so folgt $f(z) = (z - z)q(z) = 0$. □

Definition 2.4.20. Seien K ein Körper, $f \in K[X] \setminus \{0\}$ ein Polynom und $z \in K$. Die **Nullstellenordnung** $\mathrm{ord}_z(f)$ **von** f **in** z ist die größte natürliche Zahl n, so dass $(X - z)^n$ ein Teiler von f ist. Dies ist wohldefiniert, denn jeder Teiler von f hat Grad $\leq \deg(f) \in \mathbb{N}$ und $1 = (X - z)^0$ ist ein Teiler von f.

Man nennt z eine **n-fache Nullstelle von** f, wenn $\mathrm{ord}_z(f) = n$ gilt.

2.4.21 (Berechnung der Nullstellenordnung)**.** Man berechnet die Nullstellenordnung in z, indem man solange den Faktor $X - z$ abspaltet, wie z eine Nullstelle ist.

Beispiele 2.4.22.

(a) Das Polynom $f = X^3 - X^2 - 2X + 2 \in \mathbb{R}[X]$ hat 1 als reelle Nullstelle. Durch Abspalten dieser Nullstelle erhält man $X^3 - X^2 - 2X + 2 = (X - 1)(X^2 - 2)$. Da $X^2 - 2$ bei 1 nicht verschwindet, gilt $\mathrm{ord}_1(f) = 1$; in Worten ist 1 eine *einfache Nullstelle* von f. Ebenso sind $\sqrt{2}$ und $-\sqrt{2}$ einfache Nullstellen von f. Alle Elemente $z \in \mathbb{R} \setminus \{1, \pm\sqrt{2}\}$ sind keine Nullstellen (= nullfache Nullstellen) von f.

(b) Das Polynom $X(X+1)^2(X-5)^4$ hat die einfache Nullstelle 0, die doppelte Nullstelle -1 und die 4-fache Nullstelle 5 (da etwa $X(X + 1)^2$ nicht bei 5 verschwindet).

Die folgenden beiden Beispiele verwenden den Körper $\mathbb{F}_3 := \mathbb{Z}/3\mathbb{Z}$ (der Leser überzeuge sich, dass dieser Ring ein Körper ist, vgl. Satz 2.9.1).

(a) Das Polynom $f = X^4+X^3+2X^2+X+1 \in \mathbb{F}_3[X]$ hat in $\mathbb{F}_3$ nur die Nullstelle $1 = -2$. Wir können also $X-1 = X+2$ abspalten und erhalten $f = (X-1)(X^3+2X^2+X+2)$. Der zweite Faktor hat nur 1 als Nullstelle und wir erhalten $f = (X - 1)^2(X^2 + 1)$. Der Faktor (X^2+1) hat keine Nullstelle in $\mathbb{F}_3$. Also gilt $\mathrm{ord}_1(f) = 2$ und $\mathrm{ord}_z(f) = 0$ für alle $z \in \mathbb{F}_3 \setminus \{0\} = \{1, 2\}$.

(b) Das Polynom $X^3 - X \in \mathbb{F}_3[X]$ hat $0, 1$ und 2 als Nullstellen. Durch Abspalten dieser Nullstellen erhält man $X^3 - X = X(X^2 - 1) = X(X - 1)(X - 2)$.

Korollar 2.4.23. *Ein Polynom $f \in K[X] \setminus \{0\}$ vom Grad $n \in \mathbb{N}$ hat höchstens n Nullstellen in K.*

Beweis. Wir beweisen dies per Induktion über n.

Induktionsanfang: Im Fall $n = 0$ gilt $f \in K \setminus \{0\}$ und f hat keine Nullstelle.

Induktionsschritt. Gelte $n \geq 1$. Sei $z \in K$ eine Nullstelle von f. Nach Proposition 2.4.19 gilt $f = (X - z)q$ für ein $q \in K[X]$. Wegen $n = \deg(f) = \deg(X - z) + \deg(q) = 1 + \deg(q)$ gilt $\deg(q) = n - 1$. Nach Induktionsannahme hat q höchstens $n - 1$ Nullstellen. Ist $z' \in K$

eine Nullstelle von f mit $z' \neq z$, so gilt $0 = f(z') = \underbrace{(z'-z)}_{\neq 0} q(z')$, also $q(z') = 0$, und z' ist eine Nullstelle von q. Dies zeigt, dass f höchstens n Nullstellen hat. □

Beispiel 2.4.24. Die Polynome $X(X-1)$, X^2, X^2+1 in $\mathbb{R}[X]$ vom Grad zwei haben zwei, eine und keine reellen Nullstellen.

Definition 2.4.25. Ein Polynom $f \in K[X] \setminus \{0\}$ **zerfällt vollständig in Linearfaktoren**, wenn

$$f(X) = a(X-\lambda_1)\cdots(X-\lambda_n)$$

für geeignete Elemente $\lambda_1, \ldots, \lambda_n \in K$ und $a \in K^\times$ gilt. Offensichtlich sind dann a der Leitkoeffizient und n der Grad von f.

2.4.26. In diesem Fall ist $\{\lambda_1, \ldots, \lambda_n\}$ die Menge der Nullstellen von f. Beachte, dass die Elemente $\lambda_1, \ldots, \lambda_n$ nicht verschieden sein müssen.

2.4.27. Der Beweis von Korollar 2.4.23 zeigt: Hat ein Polynom $f \in K[X]$ vom Grad $n \in \mathbb{N}$ genau n verschiedene Nullstellen, so zerfällt f vollständig in Linearfaktoren.

Beispiel 2.4.28.
(a) Jedes lineare (und auch jedes konstante Polynom $\neq 0$) zerfällt vollständig in Linearfaktoren.
(b) Sei $p \in K[X]$ ein Polynom mit $\deg(p) = 2$. Dann zerfällt p genau dann vollständig in Linearfaktoren, wenn p eine Nullstelle hat (vgl. Proposition 2.4.19).
(c) Hat ein Polynom $f \in K[X]$ vom Grad $\deg(f) \geq 2$ keine Nullstelle in K, so kann es nicht vollständig in Linearfaktoren zerfallen.
(d) Über $\mathbb{R}$ hat X^2+1 keine Nullstelle, es zerfällt also nicht vollständig in Linearfaktoren.
(e) Über $\mathbb{C}$ hat X^2+1 die beiden verschiedenen Nullstellen i und $-i$ und zerfällt als $X^2+1 = (X-i)(X+i)$ vollständig in Linearfaktoren.

Satz 2.4.29 (Fundamentalsatz/Hauptsatz der Algebra - im Rahmen dieser Vorlesung ohne Beweis)**.** *Jedes Polynom $f \in \mathbb{C}[X]$ vom Grad ≥ 1 hat eine komplexe Nullstelle.*
*(Dies bedeutet, dass $\mathbb{C}$ ein **algebraisch abgeschlossener** Körper ist.)*

Bemerkung 2.4.30. Der Beweis des Hauptsatzes der Algebra benötigt Mittel aus der Analysis. Der kürzeste mir bekannte elementare Beweis steht in [dO11] und verwendet nur Vorkenntnisse aus der Analysisvorlesung für Informatiker (noch elementarer, aber länger ist [dO12]).

Korollar 2.4.31. *Jedes Polynom $f \in \mathbb{C}[X] \setminus \{0\}$ zerfällt vollständig in Linearfaktoren.*

Beweis. Wir zeigen die Aussage per Induktion über den Grad von f. Sie ist klar für $\deg(f) = 0$. Gelte $\deg(f) \geq 1$. Nach dem Fundamentalsatz der Algebra (siehe Satz 2.4.29) hat f eine Nullstelle $z \in \mathbb{C}$. Nach Proposition 2.4.19 können wir diese Nullstelle abspalten und erhalten $f = (X-z)g$ für ein $g \in \mathbb{C}[X]$. Wegen $\deg(f) = \deg(X-z) + \deg(g)$ gilt $\deg(g) = \deg(f) - 1$. Nach Induktionsannahme zerfällt g vollständig in Linearfaktoren. Dasselbe gilt dann auch für $f = (X-z)g$. □

2.5. Größter gemeinsamer Teiler und euklidischer Algorithmus.

Definition 2.5.1. Seien a und b Elemente eines Integritätsbereichs R. Ein Element $d \in R$ heißt genau dann **größter gemeinsamer Teiler (ggT) von a und b**, wenn gelten:

(a) $d \mid a$ und $d \mid b$.
(b) Für alle $t \in R$ gilt: Aus $t \mid a$ und $t \mid b$ folgt $t \mid d$.

In Worten ist d also ein gemeinsamer Teiler von a und b und jeder gemeinsame Teiler von a und b teilt d.

Beispiel 2.5.2. In $\mathbb{Z}$ sind -8 und 8 die beiden ggT von 56 und 16. Man kann dies entweder direkt sehen (die Teiler von 56 sind $\pm1, \pm2, \pm4, \pm7, \pm8, \pm14, \pm28, \pm56$, die von 16 sind $\pm1, \pm2, \pm4, \pm8, \pm16$; da alle Teiler einer ganzen Zahl n Betrag $\leq |n|$ haben, muss man nur endlich viele Zahlen testen; vgl. Hasse-Diagramm [Wik19, Hasse-Diagramm]) oder formal aus Beispiel 2.5.11 unten folgern. Obwohl der ggT nicht eindeutig ist, schreiben wir $\mathrm{ggT}(56, 16) = 8$.

2.5.3 (Eindeutigkeit von ggT)**.** Ein ggT ist eindeutig bis auf Multiplikation mit einer Einheit: Sind d und d' zwei ggT von a und b, so existiert eine Einheit $r \in R^\times$ mit $d = rd'$. Ist d ein ggT von a und b, so ist für jede Einheit $r \in R^\times$ auch rd ein ggT von a und b.

Beweis. Die zweite Aussage ist klar. Wir beweisen die erste. Es gilt $d \mid d'$ und $d' \mid d$, also $sd = d'$ und $rd' = d$ für geeignete $r, s \in R$. Somit $d = rd' = rsd$. Ist $d \neq 0$, so folgt $1 = rs$ nach 2.2.6, also ist r eine Einheit und es gilt $d = rd'$. Ist $d = 0$, so folgt $d' = 0$ und es gilt $d = 0 = 1 \cdot 0 = 1 \cdot d'$. □

Bemerkung 2.5.4 (Existenz von ggT)**.** Ein ggT existiert im Allgemeinen nicht. Beispielsweise haben im Integritätsring $\mathbb{Z}[i\sqrt{3}] = \{a + bi\sqrt{3} \mid a, b \in \mathbb{Z}\}$ die beiden Elemente $a = 4 = 2 \cdot 2 = (1 + \sqrt{3}i)(1 - \sqrt{3}i)$ und $b = 2 + 2\sqrt{3}i = 2 \cdot (1 + \sqrt{3}i)$ keinen ggT. Die Begründung ist dem interessierten Leser überlassen.

Definition 2.5.5. Ein **euklidischer Ring** ist ein Integritätsbereich R zusammen mit einer Abbildung $\delta\colon R \setminus \{0\} \to \mathbb{N}$, so dass für alle $a, b \in R$ mit $b \neq 0$ Elemente $q, r \in R$ mit

$$a = qb + r \text{ und } \delta(r) < \delta(b)$$

existieren; um hier auch den Fall $r = 0$ abzudecken, setzen wir hier und später stets $\delta(0) := -\infty$. Wir nennen eine solche Darstellung $a = qb + r$ eine *Division von a durch b mit Rest r*.

Beispiel 2.5.6. Der Polynomring $K[X]$ über einem Körper K zusammen mit der Abbildung $\deg\colon K[X] \setminus \{0\} \to \mathbb{N}$ ist nach Satz 2.4.16 ein euklidischer Ring.

Satz 2.5.7. *Der Ring $\mathbb{Z}$ der ganzen Zahlen mit der Betragsfunktion $|\cdot|\colon \mathbb{Z} \setminus \{0\} \to \mathbb{N}$ ist ein euklidischer Ring.*

Genauer gilt (Division mit Rest in den ganzen Zahlen): Für alle $a, b \in \mathbb{Z}$ mit $b \neq 0$ existieren $q, r \in \mathbb{Z}$ mit $a = qb + r$ und $0 \leq r < |b|$.

Beweis. Die Menge $\{a - q'b \mid q' \in \mathbb{Z} \text{ und } a - q'b \in \mathbb{N}\}$ ist nicht leer (im Fall $b > 0$ ist jedes $q' \leq \frac{a}{b}$ in der Menge, im Fall $b < 0$ jedes $q' \geq \frac{a}{b}$) und hat somit ein kleinstes Element $r \geq 0$. Sei $q \in \mathbb{Z}$ mit $a - qb = r$. Wir behaupten $r < |b|$. Sonst gilt $r \geq |b|$. Im Fall $b > 0$ gilt $r > r - b = a - (q + 1)b \geq 0$ im Widerspruch zur Wahl von r. Im Fall $b < 0$ gilt $r > r + b = a - (q - 1)b \geq 0$ im Widerspruch zur Wahl von r. Die Behauptung folgt. □

Beispiel 2.5.8. Neben der Division $31 = 4 \cdot 7 + 3$ mit nicht-negativem Rest 3 (wie in Satz 2.5.7) ist auch $31 = 5 \cdot 7 - 4$ eine Division mit Rest im Sinne von Definition 2.5.5.

Ende 5. Vorlesung am 05.05.2020

Satz 2.5.9. *Sei R ein euklidischer Ring und seien $a, b \in R$. Dann existiert ein ggT von a und b. Ist d ein ggT von a und b, so existieren $x, y \in R$ mit $d = xa + yb$ (Berechnung durch den im Beweis erklärten* ***euklidischen Algorithmus****).*

Beispiel 2.5.10. Der ggT -8 von 56 und 16 hat die Darstellung $-8 = 1 \cdot 56 - 4 \cdot 16$. Der ggT 8 von 56 und 16 hat die Darstellungen $8 = 1 \cdot 56 - 3 \cdot 16 = (-1) \cdot 56 + 4 \cdot 16$.

Beweis. Ohne Einschränkung gelte $a \neq 0 \neq b$ (denn im Fall $a = b = 0$ ist $0 = 0 \cdot 0 + 0 \cdot 0$ ein ggT, im Fall $a = 0 \neq b$ ist $b = 0 \cdot 0 + 1 \cdot b$ ein ggT, und der Fall $a \neq 0 = b$ geht analog).

Bestimmung des ggT mit dem **euklidischen Algorithmus**: Sei $\delta\colon R \setminus \{0\} \to \mathbb{N}$ die Abbildung, die R zu einem euklidischen Ring macht. Gelte ohne Einschränkung $\delta(b) \leq \delta(a)$ (sonst vertausche a und b). Wir werden induktiv Elemente $a_0, a_1, a_2, \ldots$ in R definieren. Setze $a_0 := a$ und $a_1 := b$.

Sei $i \geq 2$ und seien a_{i-2} und a_{i-1} bereits definiert und gelte $a_{i-1} \neq 0$. Wir definieren a_i als „Rest einer Divison von a_{i-2} durch a_{i-1}“. Genauer gibt es nach Definition eines euklidischen Rings Elemente $q_i, a_i \in R$ mit

$$a_{i-2} = q_i a_{i-1} + a_i \text{ und } \delta(a_i) < \delta(a_{i-1}).$$

Solange $a_i \neq 0$ gilt, iterieren wir dies. Wegen $\delta(a_0) \geq \delta(a_1) > \delta(a_2) > \ldots$ bricht unser Verfahren nach endlich vielen Schritten ab (da es keine unendliche echt absteigende Folge natürlicher Zahlen gibt), d. h. es gibt ein $n \geq 1$ mit $a_{n+1} = 0$.

Wir behaupten, dass $d := a_n$ ein ggT von a und b ist.

Per Konstruktion haben wir Gleichungen

$$\begin{aligned}
a_0 &= q_2 a_1 + a_2, \\
a_1 &= q_3 a_2 + a_3, \\
&\vdots \\
a_{n-3} &= q_{n-1} a_{n-2} + a_{n-1}, \\
a_{n-2} &= q_n a_{n-1} + a_n, \\
a_{n-1} &= q_{n+1} a_n.
\end{aligned}$$

Trivialerweise ist a_n ein Vielfaches von $d = a_n$. Liest man unsere Gleichungen von unten nach oben, so sieht man, dass auch $a_{n-1}, a_{n-2}, \ldots, a_2, a_1 = b, a_0 = a$ Vielfache von d sind. Also ist d Teiler von a und b.

Sei x ein Teiler von $a = a_0$ und $b = a_1$. Nach der ersten Gleichung ist dann x auch ein Teiler von a_2. Nach der zweiten Gleichung ist dann x auch ein Teiler von a_3, ..., nach der vorletzten Gleichung ist x auch ein Teiler von $a_n = d$. Dies zeigt, dass d ein ggT von a und b ist.

Um eine Darstellung $d = ax + by$ zu erhalten, beginnen wir mit der vorletzten Gleichung. Sie liefert die Darstellung

$$d = a_n = a_{n-2} - q_n a_{n-1}.$$

Die vorvorletzte Gleichung erlaubt, darin a_{n-1} durch $a_{n-3} - q_{n-1} a_{n-2}$ zu ersetzen. ... Die erste Gleichung erlaubt, a_2 durch $a_0 - q_2 a_1$ zu ersetzen. Zusammengefasst liefert das eine Darstellung $d = xa + yb$.

Ist d' ein beliebiger ggT von a und b, so gilt $d' = ud$ für eine Einheit $u \in R$ (siehe 2.5.3), und wir erhalten die Darstellung $d' = ud = uxa + uyb$. □

Beispiel 2.5.11. Wir bestimmen einen ggT von 150 und 42 im euklidischen Ring $\mathbb{Z}$ mit dem euklidischen Algorithmus.

$$\begin{aligned} 150 &= 3 \cdot 42 + 24 \\ 42 &= 1 \cdot 24 + 18 \\ 24 &= 1 \cdot 18 + 6 \\ 18 &= 3 \cdot 6. \end{aligned}$$

Nach dem Beweis von Satz 2.5.9 gelten also $6 = \mathrm{ggT}(150, 42)$ und

$$6 = 24 - 18 = 24 - (42 - 24) = -42 + 2 \cdot 24 = -42 + 2 \cdot (150 - 3 \cdot 42) = 2 \cdot 150 - 7 \cdot 42.$$

Aufgabe 2.5.12. Seien $n \geq 1$ und $a \in \mathbb{Z}$. Dann gilt $\mathrm{ggT}(n, a) = \pm 1$ genau dann, wenn $\mathbb{Z}/n\mathbb{Z} = \langle a \rangle$.

Satz 2.5.13. *Sei K ein Körper. Dann besitzen je zwei Polynome $f, g \in K[X]$, die nicht beide Null sind, genau einen normierten ggT (Berechnung mit euklidischem Algorithmus).*

Beweis. Existenz: Nach Satz 2.4.16 ist $K[X]$ ein euklidischer Ring. Also besitzen f und g nach Satz 2.5.9 einen ggT d, der mit dem euklidischen Algorithmus berechnet werden kann. Es gilt $d \neq 0$, da nicht $f = 0 = g$ gilt. Sei $a \neq 0$ der Leitkoeffizient von d. Dann ist $\frac{1}{a}d$ ein normierter ggT von f und g (siehe 2.5.3).

Eindeutigkeit: Ist d' ein weiterer normierter ggT von f und g, so gilt $d = ud'$ für eine Einheit $u \in K[X]^\times = K \setminus \{0\}$ nach 2.5.3 und Proposition 2.4.14. Da d und d' normiert sind und $u \in K$ gilt, folgt $u = 1$ per Vergleich der Leitkoeffizienten, also $d = d'$. □

Beispiel 2.5.14. Wir berechnen den normierten ggT von $f = f(X) = X^2 + X$ und $g = g(X) = X^2 + 1$ in $\mathbb{R}[X]$ mit Hilfe des euklidischen Algorithmus.

$$\begin{array}{rl} (X^2 + X) : (X^2 + 1) = 1 & \text{Rest } X - 1 \\ \underline{-(X^2 + 1)} & \\ X - 1 & \end{array} \qquad \Longrightarrow\ X^2 + X = 1 \cdot (X^2 + 1) + \boxed{(X - 1)}$$

$$\begin{array}{rl} (X^2 + 1) : (X - 1) = X + 1 & \text{Rest } 2 \\ \underline{-(X^2 - X)} & \\ X + 1 & \\ -(X - 1) & \\ 2 & \end{array} \qquad \Longrightarrow\ X^2 + 1 = (X + 1) \cdot \boxed{(X - 1)} + 2$$

$$(X - 1) : 2 = \frac{1}{2}X - \frac{1}{2} \qquad \text{Rest } 0 \qquad \Longrightarrow\ (X - 1) = (\frac{1}{2}X - \frac{1}{2}) \cdot 2 + 0$$

Dies bedeutet, dass 2 ein ggT von f und g in $\mathbb{R}[X]$ ist. Der normierte ggT von f und g ist also 1.

Lesen unserer Gleichungen von unten und Ersetzen liefert (wir starten mit der vorletzten Gleichung und ersetzen den umrahmten Ausdruck durch den aus der vorigen Gleichung erhaltenen Ausdruck)

$$\begin{aligned} 2 &= (X^2+1) - (X+1)(X-1) \\ &= (X^2+1) - (X+1)[(X^2+X) - (X^2+1)] \\ &= (X+2)(X^2+1) - (X+1)(X^2+X) \end{aligned}$$

und somit die folgende Darstellung des normierten ggT.

$$1 = \frac{X+2}{2}(X^2+1) - \frac{X+1}{2}(X^2+X)$$

Aufgabe 2.5.15. Bestimmen Sie den normierten ggT von $f = f(X) = X^2 + X$ und $g = g(X) = X^2 + 1$ in $\mathbb{F}_2[X]$ (statt in $\mathbb{R}[X]$ wie im vorigen Beispiel) mit Hilfe des euklidischen Algorithmus und stellen Sie ihn in der Form $pf + qg$ dar.

Beispiel 2.5.16 (Fortsetzung von Beispiel 2.4.18)**.** Wir bestimmen den normierten ggT von $f = f(X) = X^5 + X^4 + 2X^3 - X^2 - X - 2$ und $g = g(X) = X^3 + 4X^2 + 5X + 6$ in $\mathbb{R}[X]$ mit Hilfe des euklidischen Algorithmus. Erste Polynomdivision:

$$\begin{array}{rrrrrrl} X^5 & +X^4 & +2X^3 & -X^2 & -X & -2 & = (X^3+4X^2+5X+6)(X^2-3X+9) - 28X^2 - 28X - 56 \\ -X^5 & -4X^4 & -5X^3 & -6X^2 & & & \\ \hline & -3X^4 & -3X^3 & -7X^2 & -X & & \\ & 3X^4 & +12X^3 & +15X^2 & +18X & & \\ \hline & & 9X^3 & +8X^2 & +17X & -2 & \\ & & -9X^3 & -36X^2 & -45X & -54 & \\ \hline & & & -28X^2 & -28X & -56 & \end{array}$$

Zweite Polynomdivision:

$$\begin{array}{rrrl} X^3 & +4X^2 & +5X+6 & = (-28X^2-28X-56)(-\frac{1}{28}X - \frac{3}{28}) \\ -X^3 & -X^2 & -2X & \\ \hline & 3X^2 & +3X+6 & \\ & -3X^2 & -3X-6 & \\ \hline & & 0 & \end{array}$$

Dies bedeutet, dass $-28X^2 - 28X - 56$ ein ggT bzw. $X^2 + X + 2$ der normierte ggT von f und g sind. Die Darstellung als Summe von Vielfachen von f und g liest man direkt aus der ersten Polynomdivision ab.

Definition 2.5.17. Seien a und b Elemente eines Integritätsbereichs R. Ein Element $v \in R$ heißt genau dann **kleinstes gemeinsames Vielfaches (kgV) von a und b**, wenn gelten:

(a) $a \mid v$ und $b \mid v$.

(b) Für alle $w \in R$ gilt: Aus $a \mid w$ und $b \mid w$ folgt $v \mid w$.

Ein kgV ist eindeutig bis auf Multiplikation mit einer Einheit und wird als $\mathrm{kgV}(a, b)$ notiert.

Satz 2.5.18. *Seien a und b zwei Elemente eines euklidischen Rings R. Dann existiert ein kgV von a und b und es gilt*

$$\mathrm{ggT}(a, b) \cdot \mathrm{kgV}(a, b) = ab,$$

wenn man ggT und kgV geeignet wählt.

Beweis. Sei $d = \mathrm{ggT}(a,b)$ ein ggT von a und b und gelte $d = xa + yb$ für geeignete $x, y \in R$ (Satz 2.5.9). Da d Teiler von a und b ist, gibt es $\alpha, \beta \in R$ mit $a = \alpha d$ und $b = \beta d$. Wir behaupten, dass

$$v := \alpha\beta d$$

ein kgV von a und b ist. Wegen $dv = d\alpha\beta d = ab$ gilt dann die behauptete Formel.

Wegen $v = \alpha\beta d = a\beta = \alpha b$ ist v ein gemeinsames Vielfaches von a und b.

Im Fall $d = 0$ ist $v = \alpha\beta d = 0$ offensichtlich ein kgV von $a = \alpha d = 0$ und $b = \beta d = 0$.

Gelte also $d \neq 0$. Aus $d1 = d = xa + yb = d(x\alpha + y\beta)$ folgt durch Kürzen $1 = x\alpha + y\beta$. Sei $w \in R$ ein Vielfaches von a und b, d. h. es gibt $s, t \in R$ mit $w = sa = tb$. Dann ist

$$w = w1 = w(x\alpha + y\beta) = tbx\alpha + say\beta = t\beta dx\alpha + s\alpha dy\beta = \alpha\beta d(tx + sy) = v(tx + sy)$$

ein Vielfaches von v wie gewünscht. □

Aufgabe 2.5.19. Seien $f, g \in K[X]$ Polynome, die nicht beide Null sind. Dann ist der normierte ggT von f und g das eindeutige Polynom größten Grades in der Menge

$$\{h \in K[X] \mid h \text{ ist normiert und teilt } f \text{ und } g\}.$$

In Worten ist der normierte größte gemeinsame Teiler der normierte gemeinsame Teiler größten Grades.

2.6. **Chinesischer Restsatz.**

2.6.1. Sind R und S Ringe, so ist $R \times S$ mit komponentenweiser Addition und Multiplikation, d. h. $(r,s) + (r',s') := (r + r', s + s')$ und $(r,s) \cdot (r',s') := (r \cdot r', s \cdot s')$, ebenfalls ein Ring.

Satz 2.6.2 (Chinesischer Restsatz (Version für $\mathbb{Z}$))**.** *Seien $m, n \geq 1$ ganze Zahlen mit $\mathrm{ggT}(m,n) = 1$. Dann ist die Abbildung*

$$\varphi\colon \mathbb{Z}/(mn)\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z},$$
$$[a]_{mn} \mapsto ([a]_m, [a]_n),$$

für $a \in \mathbb{Z}$, ein Isomorphismus von Ringen.

Beweis. Sei $\varphi\colon \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ die angegebene Abbildung. Sie ist wohldefiniert, denn jedes Vielfache von mn ist auch ein Vielfaches von m und von n. Sie ist offensichtlich ein Ringhomomorpismus. Da Start- und Zielbereich von φ je mn Elemente haben, genügt es zu zeigen, dass φ surjektiv ist[16] [17].

Wegen $\mathrm{ggT}(m,n) = 1$ gibt es ganze Zahlen $x, y \in \mathbb{Z}$ mit $1 = xm + yn$ (siehe Satz 2.5.9). Für das Element $a := [yn]_{mn}$ gilt dann

$$\varphi(a) = ([yn]_m, [yn]_n) = ([1 - xm]_m, [0]_n) = ([1]_m, [0]_n).$$

[16] Für eine Abbildung zwischen endlichen Mengen sind die Eigenschaften injektiv, surjektiv und bijektiv äquivalent.

[17] Alternativ reicht es auch zu zeigen, dass φ injektiv ist, was ähnlich kompliziert ist:

Nach Lemma 1.3.7 reicht es zu zeigen, dass der Kern von φ (als Gruppenhomomorphismus) nur aus dem Element $[0]_{mn}$ besteht. Sei $a \in \mathbb{Z}$ mit $\varphi([a]_{nm}) = ([a]_m, [a]_n) = 0 = ([0]_m, [0]_n)$. Dies bedeutet, dass a sowohl durch m als auch durch n teilbar ist, dass es also $u, v \in \mathbb{Z}$ mit $a = mu = nv$ gibt.

Wegen $\mathrm{ggT}(m,n) = 1$ finden wir mit dem euklidischen Algorithmus ganze Zahlen $x, y \in \mathbb{Z}$ mit $xm + yn = 1$. Multiplizieren wir diese Gleichung mit a, so erhalten wir

$$a = axm + ayn = nvxm + muyn = mn(vx + uy).$$

Dies bedeutet $mn \mid a$, also $[a]_{mn} = 0$. Also ist φ injektiv.

Für das Element $b := [xm]_{mn}$ gilt analog

$$\varphi(b) = ([xm]_m, [xm]_n) = ([0]_m, [1 - yn]_n) = ([0]_m, [1]_n).$$

Für beliebige $s, t \in \mathbb{Z}$ erhalten wir damit

$$\begin{aligned}\varphi(sa + tb) = \varphi(sa) + \varphi(tb) = s\varphi(a) + t\varphi(b) = s([1]_m, [0]_n) + t([0]_m, [1]_n) \\ = ([s]_m, [0]_n) + ([0]_m, [t]_n) = ([s]_m, [t]_n).\end{aligned}$$

Also ist φ surjektiv. □

2.7. **Primfaktorzerlegung - Zerlegung in irreduzible Faktoren.**

Definition 2.7.1. Sei R ein kommutativer Ring. Ein Element $p \in R$ heißt genau dann **irreduzibel**, wenn gelten:

(a) Das Element p ist keine Einheit, d. h. $p \notin R^\times$.
(b) Ist $p = xy$ das Produkt zweier Elemente $x, y \in R$, so ist x oder y eine Einheit in R.

2.7.2. Sei $u \in R^\times$ eine Einheit. Dann ist $p \in R$ genau dann irreduzibel, wenn up irreduzibel ist.

Bemerkung 2.7.3. Das Nullelement 0 eines beliebigen Ringes ist nie irreduzibel: Im Nullring ist $0 = 1$ eine Einheit. Sonst ist $0 = 0 \cdot 0$ das Produkt zweier Nichteinheiten.

Beispiel 2.7.4 (Irreduzible Elemente in $\mathbb{Z}$)**.** Wir erinnern daran, dass eine ganze Zahl $p \in \mathbb{Z}$ genau dann **Primzahl** heißt, wenn $p > 1$ gilt und für alle ganzen Zahlen $x \geq 1$ aus $x \mid p$ entweder $x = 1$ oder $x = p$ folgt.

Die Primzahlen sind also genau die positiven irreduziblen Elemente von $\mathbb{Z}$.

Die irreduziblen Elemente von $\mathbb{Z}$ sind also die Primzahlen und ihre negativen.

Beispiel 2.7.5 (Irreduzible Polynome)**.** Im Falle des Polynomrings $K[X]$ in einer Variablen über einem Körper K gilt offensichtlich: Ein Polynom $f \in K[X]$ ist genau dann irreduzibel, wenn die beiden folgenden Bedingungen erfüllt sind:

(a) Es gilt $\deg(f) \geq 1$.[18]
(b) Alle Teiler von f in $K[X]$ haben Grad 0 oder $\deg(f)$.[19]

Wir geben einige Beispiele:

(a) Lineare Polynome, d. h. Polynome der Form $aX + b$, für $a, b \in K$ mit $a \neq 0$, sind irreduzibel.
(b) Das Polynom $X^3 - 1$ ist nicht irreduzibel wegen $(X^3 - 1) = (X - 1)(X^2 + X + 1)$.
(c) Hat ein Polynom in $K[X]$ vom Grad ≥ 2 eine Nullstelle, so ist es nicht irreduzibel (Abspalten von Nullstellen, Proposition 2.4.19).

[18] Nach Proposition 2.4.14 sind die Einheiten in $K[X]$ genau die Elemente vom Grad Null. Die Bedingung $\deg(f) \geq 1$ bedeutet also, dass f keine Einheit ist und auch nicht das Nullelement, das niemals irreduzibel ist, siehe Bemerkung 2.7.3.

[19] Diese Bedingung, nennen wir sie $(\star)$ ist äquivalent zur Bedingung (b) in Definition 2.7.1:

$(\star) \implies$ (b): Gelte $f = pq$ für Elemente $p, q \in K[X]$, also insbesondere $\deg(f) = \deg(p) + \deg(q)$. Aus $(\star)$ folgt $\deg(p) = 0$ oder $\deg(p) = \deg(f)$. Im Fall $\deg(p) = 0$ ist p eine Einheit in $K[X]$. Im Fall $\deg(p) = \deg(f)$ folgt $\deg(q) = 0$ und q ist eine Einheit in $K[X]$.

(b) $\implies (\star)$: Sei p ein Teiler von f, also $f = pq$ für ein $q \in K[X]$. Wegen (b) ist p oder q eine Einheit in $K[X]$. Nach Proposition 2.4.14 gilt also $p \in K^\times$ oder $q \in K^\times$. Also hat p Grad Null oder Grad $\deg(f)$.

(d) Ein Polynom vom Grad 2 oder 3 in $K[X]$ ist genau dann irreduzibel, wenn es keine Nullstelle in K hat.

Beispiel 2.7.6. Die normierten irreduziblen Polynome in $\mathbb{C}[X]$ sind gegeben durch die linearen Polynome $X - z$ für $z \in \mathbb{C}$. Dies folgt sofort aus Korollar 2.4.31.

Aufgabe 2.7.7. Jedes irreduzible normierte Polynom $p \in \mathbb{R}[X]$ ist entweder linear (d. h. $\deg(p) = 1$) oder quadratisch (d. h. $\deg(p) = 2$). Im linearen Fall hat es die Gestalt $X - a$ für ein $a \in \mathbb{R}$, im quadratischen Fall hat es die Gestalt $(X - a)^2 + b$ mit $a, b \in \mathbb{R}$ und $b > 0$. Umgekehrt sind alle Polynome der angegebenen Gestalten irreduzibel.

Hinweis: Verwenden Sie den Fundamentalsatz der Algebra. Hat p eine komplexe, nicht reelle Nullstelle $z \in \mathbb{C} \setminus \mathbb{R}$, so ist auch $\overline{z}$ eine Nullstelle von p und das Polynom $(X-z)(X-\overline{z})$ (mit reellen Koeffizienten) ist ein Teiler von p sowohl in $\mathbb{C}[X]$ als auch in $\mathbb{R}[X]$ (nutze Polynomdivision in beiden Ringen).

Ende 6. Vorlesung am 07.05.2020

Lemma 2.7.8. *Sei R ein Integritätsbereich. Seien $p, b \in R$ mit p irreduzibel und $p \nmid b$. Dann ist 1 ein ggT von p und b, d. h.* $\mathrm{ggT}(p, b) = 1$.

Beweis. Sicherlich ist 1 ein gemeinsamer Teiler von p und b.

Sei x ein gemeinsamer Teiler von p und b, d. h. $x \mid p$ und $x \mid b$. Dann gibt es ein $y \in R$ mit $p = xy$. Da p irreduzibel ist, ist x oder y eine Einheit.

Ist y eine Einheit, so folgt aus $x \mid b$ sofort $p = xy \mid b$ (denn $b = xc = (xy)y^{-1}c$ für ein $c \in R$) im Widerspruch zur Annahme.

Also ist x eine Einheit, d. h. $x \mid 1$, und 1 ist ein größter gemeinsamer Teiler von p und b. □

Lemma 2.7.9. *Sei R ein euklidischer Ring und sei $p \in R$ irreduzibel. Dann gilt: Teilt p ein Produkt, so teilt es einen der Faktoren. Explizit: Für alle $a, b \in R$ folgt aus $p \mid ab$ bereits $p \mid a$ oder $p \mid b$.*

Beweis. Gelte $p \mid ab$. Gilt $p \mid b$, so sind wir fertig. Sonst gilt $\mathrm{ggT}(p, b) = 1$ nach Lemma 2.7.8. Da R euklidisch ist, liefert Satz 2.5.9 Elemente $x, y \in R$ mit $1 = xp + yb$. Es folgt $a = axp + yab$. Wegen $p \mid ab$ impliziert dies $p \mid a$. □

2.7.10. Der Beweis des Satzes 2.7.12 benötigt als Vorbereitung ein weiteres Lemma. Für $x \in R$ bezeichne $Rx := \{rx \mid r \in R\}$ die Menge aller Vielfachen von x.

Lemma 2.7.11. *Sei (R, δ) ein euklidischer Ring und sei $x \in R \setminus \{0\}$. Sei y ein Element von $Rx \setminus \{0\}$ mit minimalem δ-Wert (d. h. $\delta(y) \leq \delta(z)$ für alle $z \in Rx \setminus \{0\}$). Dann gilt $Rx = Ry$.*

Beweis. Aus $y \in Rx$ folgt offensichtlich $Ry \subset Rx$.

Zu zeigen bleibt $Rx \subset Ry$. Sei $a \in Rx$. Per „Division mit Rest“ schreibe $a = qy + r$ mit $\delta(r) < \delta(y)$. Da die Summe zweier Elemente aus Rx wieder in Rx liegt, folgt $r = a - qy \in Rx$. Ist $r \neq 0$, so liefert die Minimalitätsannahme $\delta(y) \leq \delta(r)$ im Widerspruch zu $\delta(r) < \delta(y)$. Also muss $r = 0$ gelten, d. h. $a = qy \in Ry$. □

Satz 2.7.12 („Primfaktorzerlegung“ - Zerlegung in irreduzible Faktoren)**.** *Sei R ein euklidischer Ring.*

(a) *Jedes Element $a \in R \setminus \{0\}$ lässt sich als Produkt einer Einheit und irreduzibler Elemente darstellen, in Formeln*

$$a = up_1 \cdots p_n,$$

für eine Einheit $u \in R^\times$, eine natürliche Zahl $n \in \mathbb{N}$ und irreduzible Elemente $p_1, \cdots, p_n \in R$.

(b) *Eine solche Darstellung ist eindeutig bis auf Einheiten und Reihenfolge der Faktoren: Sind $a = up_1 \cdots p_n = vq_1 \dots q_m$ zwei Darstellungen wie oben, so gilt $n = m$ und es gibt eine Permutation $\sigma \in S_n$ und Einheiten $\varepsilon_i \in R^\times$ mit $q_i = \varepsilon_i p_{\sigma(i)}$ für alle $i = 1, \dots, n$.*

Beweis. (a) Sei $\delta : R \setminus \{0\} \to \mathbb{N}$ die Abbildung, die R zu einem euklidischen Ring macht.

Angenommen, es existiert ein Element in $R \setminus \{0\}$, das keine solche Darstellung hat. Sei $a \in R \setminus \{0\}$ unter all diesen Elementen eines mit minimalem δ-Wert. Da a dann weder irreduzibel noch eine Einheit ist (sonst hätte es ja eine Darstellung wie gefordert), gilt $a = bc$ für geeignete Nicht-Einheiten $b, c \in R \setminus \{0\}$.

Offensichtlich gilt $Ra \subset Rb$. Wir behaupten, dass die Inklusion echt ist, d. h. $Ra \subsetneq Rb$.

Nehmen wir an, dass Gleichheit $Ra = Rb$ gilt, also $b = 1b = xa$ für ein $x \in R$ und somit $a = bc = xac = acx$. Kürzen (mit $a \neq 0$) liefert $1 = cx$ im Widerspruch zu $c \notin R^\times$. Dies zeigt $Ra \subsetneq Rb$.

Wähle y in $Rb \setminus \{0\}$ mit minimalem δ-Wert (es existiert wegen $b \neq 0$). Lemma 2.7.11 liefert $Ry = Rb$.

Wegen $0 \neq a = cb \in Rb$ gilt also $\delta(y) \leq \delta(a)$. Gilt Gleichheit, also $\delta(y) = \delta(a)$, so zeigt Lemma 2.7.11 die Gleichheit $Ra = Rb$ im Widerspruch zu $Ra \subsetneq Rb$. Es gilt also $\delta(y) < \delta(a)$. Nach Wahl von a bedeutet dies, dass y ein Produkt einer Einheit und endlich vieler irreduzibler Elemente ist.

Aus $Ry = Rb$ erhalten wir $y = sb$ und $b = ty$ für geeignete Elemente $s, t \in R$, also $y = sb = sty$ und damit $1 = st$, also $t \in R^\times$. Also ist auch $b = ty$ ein Produkt einer Einheit und endlich vieler irreduzibler Elemente.

Dasselbe Argument zeigt, dass c ebenfalls ein Produkt einer Einheit und endlich vieler irreduzibler Elemente ist. Dies folgt dann auch für $a = bc$ und liefert den gesuchten Widerspruch zur Annahme vom Anfang des Beweises. Dies zeigt die erste Behauptung.

(b) Ohne Einschränkung gelte $n \leq m$. Aus $p_1 \mid a = vq_1 \dots q_m$ folgt mit Lemma 2.7.9, dass $p_1 \mid q_j$ für ein j gilt (der Fall $p_1 \mid v$ kann nicht eintreten, da p_1 sonst als Teiler einer Einheit eine Einheit wäre).

Per Umnummerieren können wir ohne Einschränkung $j = 1$ annehmen, d. h. $p_1 \mid q_1$, also $\varepsilon_1 p_1 = q_1$ für ein $\varepsilon_1 \in R$. Da q_1 irreduzibel ist und p_1 keine Einheit ist, ist ε_1 eine Einheit.

Aus $up_1 \cdots p_n = vq_1 q_2 \cdots q_m = v\varepsilon_1 p_1 q_2 \dots q_m$ erhalten wir per Kürzen von $p_1 \neq 0$ die Gleichung $up_2 \cdots p_n = v\varepsilon_1 q_2 \dots q_m$.

Wir fahren induktiv fort. Dies liefert $q_i = \varepsilon_i p_i$ mit $\varepsilon_i \in R^\times$ für alle $i = 1, \dots, n$ und $u = v\varepsilon_1 \cdots \varepsilon_n q_{n+1} \cdots q_m$. Falls $n < m$ gilt, so ist q_{n+1} als Teiler der Einheit u ebenfalls eine Einheit in R, was der Irreduzibilität von q_{n+1} widerspricht. Damit gilt $n = m$ und wir sind fertig. (Die Permutation σ ist im „Umnummerieren“ versteckt.)

□

Korollar 2.7.13 (Primfaktorzerlegung)**.** *Jede ganze Zahl $n \in \mathbb{Z} \setminus \{0\}$ lässt sich als Produkt*

$$n = up_1p_2 \cdots p_n$$

einer Einheit $u \in \{\pm 1\}$ und von Primzahlen $p_1, \ldots, p_n$ schreiben, die bis auf die Reihenfolge eindeutig bestimmt sind.

Beweis. Da $\mathbb{Z}$ ein euklidischer Ring ist (Satz 2.5.7), folgt dies sofort aus Satz 2.7.12 und Beispiel 2.7.4. □

Beispiel 2.7.14. $-60984 = (-1) \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 7 \cdot 11 \cdot 11 = (-1) \cdot 2^3 \cdot 3^2 \cdot 7 \cdot 11^2$

Bemerkung 2.7.15. Man kann Korollar 2.7.13 auch so formulieren: Sei $\mathbf{P} = \{2, 3, 5, \ldots\}$ die Menge aller Primzahlen. Dann existieren für jede ganze Zahl $n \in \mathbb{Z}$ eindeutig bestimmte natürliche Zahlen $\nu_p \in \mathbb{N}$, für alle $p \in \mathbf{P}$, so dass bis auf endlich viele Ausnahmen alle ν_p Null sind, und eine eindeutig bestimmte Einheit $u \in \{\pm 1\}$ mit

$$n = u \prod_{p \in \mathbf{P}} p^{\nu_p}.$$

Daraus folgert man leicht: Ist $m = v \prod_{p \in \mathbf{P}} p^{\mu_p}$ eine analoge Darstellung für $m \in \mathbb{Z}$, so gelten $\mathrm{ggT}(m, n) = \prod_{p \in \mathbf{P}} p^{\min(\nu_p, \mu_p)}$ und $\mathrm{kgV}(m, n) = \prod_{p \in \mathbf{P}} p^{\max(\nu_p, \mu_p)}$.

Satz 2.7.16. *Sei K ein Körper. Sei $f \in K[X] \setminus \{0\}$ ein Polynom mit Leitkoeffizient $a \in K$. Dann existieren irreduzible normierte Polynome $f_1, \ldots, f_r$ mit*

$$f = af_1 \cdots f_r.$$

Die Polynome $f_1, \ldots, f_r$ sind bis auf ihre Reihenfolge eindeutig bestimmt.

Beweis. Da $K[X]$ ein euklidischer Ring ist (Beispiel 2.5.6), folgt dies sofort aus Satz 2.7.12; um ein irreduzibles Polynom in eine normiertes irreduzibles Polynom zu verwandeln, multipliziere man es mit dem Inversen seines Leitkoeffizienten. □

2.8. **Ganzheit von Nullstellen.**

Satz 2.8.1 (Nullstellen von normierten Polynomen mit ganzen Koeffizienten)**.** *Rationale Nullstellen von normierten Polynomen mit ganzen Koeffizienten sind ganzzahlig und Teiler des konstanten Koeffizienten:*

Sei $f(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0 \in \mathbb{Q}[X]$ ein normiertes Polynom vom Grad $n \in \mathbb{N}$ mit ganzen Koeffizienten $a_{n-1}, \ldots, a_0 \in \mathbb{Z}$. Dann ist jede rationale Nullstelle $z \in \mathbb{Q}$ von f ganzzahlig, d. h. $z \in \mathbb{Z}$, und ein Teiler von a_0 in $\mathbb{Z}$.

Beweis. Für $z = 0$ ist die Aussage trivial, denn dann gilt $0 = f(0) = a_0$.

Gelte $z \neq 0$. Schreibe $z = \frac{r}{s}$ mit $r, s \in \mathbb{Z} \setminus \{0\}$.

Wir können annehmen, dass dieser Bruch „vollständig gekürzt“ ist, dass es also keine Primzahl gibt, die sowohl r als auch s teilt (Korollar 2.7.13).

Wir zeigen zuerst $z \in \mathbb{Z}$. Zu zeigen ist also $s = \pm 1$. Sonst gibt es eine Primzahl p, die s teilt (vgl. Korollar 2.7.13). Schreibe $s = ps'$ für ein $s' \in \mathbb{Z}$.

Da $z = \frac{r}{ps'}$ eine Nullstelle von f ist, gilt

$$\left(\frac{r}{ps'}\right)^n + a_{n-1}\left(\frac{r}{ps'}\right)^{n-1} + \ldots + a_1\frac{r}{ps'} + a_0 = 0 \qquad \text{in } \mathbb{Q}.$$

Multiplikation mit $(ps')^n$ ergibt:

$$r^n + \underbrace{ps'a_{n-1}r^{n-1} + \ldots + (ps')^{n-1}a_1 r + (ps')^n a_0}_{\text{Vielfaches von } p} = 0 \qquad \text{in } \mathbb{Z}.$$

Also ist p ein Teiler von r^n und damit von r (Lemma 2.7.9). Somit ist p ein gemeinsamer Teiler von s und r, im Widerspruch zu unserer Annahme. Dies zeigt $s = \pm 1$ und damit $z = \pm r \in \mathbb{Z}$.

Nun zeigen wir, dass z ein Teiler von a_0 in $\mathbb{Z}$ ist. Durch Abspalten der Nullstelle z erhalten wir $f(X) = (X - z)g(X)$ für ein notwendigerweise normiertes Polynom $g \in \mathbb{Q}[X]$ vom Grad $n - 1$ (Proposition 2.4.19), sagen wir $g(X) = X^{n-1} + b_{n-2}X^{n-2} + \cdots + b_0$ für geeignete Koeffizienten $b_i \in \mathbb{Q}$, die eindeutig durch die Gleichungen

$$\begin{aligned}
a_{n-1} &= b_{n-2} - z, \\
a_{n-2} &= b_{n-3} - zb_{n-2}, \\
&\vdots \\
a_i &= b_{i-1} - zb_i, \\
&\vdots \\
a_1 &= b_0 - zb_1, \\
a_0 &= -zb_0
\end{aligned}$$

bestimmt sind. Liest man diese Gleichungen von oben und verwendet, dass z und alle a_i ganze Zahlen sind, so sieht man sukzessive, dass alle b_i ganze Zahlen sind. Dafür braucht man die letzte Gleichung nicht; diese zeigt aber, dass z ein Teiler von a_0 in $\mathbb{Z}$ ist. □

Beispiel 2.8.2. Satz 2.8.1 kann zum Raten von Nullstellen (bzw. zum Bestimmen der rationalen Nullstellen) von normierten Polynomen mit ganzzahligen Koeffizienten verwendet werden.

Betrachten wir beispielsweise das normierte Polynom $X^3 + 5X^2 + X + 5$ mit ganzzahligen Koeffizienten. Hat es eine rationale Nullstelle, so ist diese ein Teiler von 5 in $\mathbb{Z}$, also ± 1 oder ± 5. Es ist klar, dass 1 und 5 keine Nullstellen sind. Es ist -1 keine Nullstelle. Aber -5 ist eine Nullstelle. In der Tat ist unser Polynom gleich $(X + 5)(X^2 + 1)$.

Beispiel 2.8.3. Das Polynom $X^3 + X + 1 \in \mathbb{Q}[X]$ hat nach Satz 2.8.1 keine rationale Nullstelle, da weder 1 noch -1 Nullstellen sind. Also ist es irreduzibel in $\mathbb{Q}[X]$.

2.9. **Endliche Körper von Primzahlordnung.**

Satz 2.9.1. *Sei $n \geq 1$ eine ganze Zahl. Dann ist $\mathbb{Z}/n\mathbb{Z}$ genau dann ein Körper, wenn n eine Primzahl ist.*

Beweis. $\Rightarrow$: Wir zeigen äquivalent: Ist n keine Primzahl, so ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper.

Sei also n keine Primzahl. Ist $n > 1$, so existieren ganze Zahlen $1 < r < n$ und $1 < s < n$ mit $n = rs$. Dann gelten $[r] \cdot [s] = [n] = [0]$ aber $[r], [s] \neq [0]$. Also ist $\mathbb{Z}/n\mathbb{Z}$ (kein Integritätsbereich und erst recht) kein Körper. Ist $n = 1$, so gilt $\mathbb{Z}/n\mathbb{Z} = \{[0]\}$, was wegen $[1] = [0]$ kein Körper ist.

$\Leftarrow$: Sei n eine Primzahl und sei $x \in \mathbb{Z}$ mit $[x] \neq [0]$ in $\mathbb{Z}/n\mathbb{Z}$, d. h. $n \nmid x$. Da n eine Primzahl ist, hat sie nur die Teiler $\pm 1, \pm n$, und es folgt $1 = \mathrm{ggT}(x, n)$ (dies folgt auch aus

Lemma 2.7.8). Nach Satz 2.5.9 (und da $\mathbb{Z}$ mit dem Absolutbetrat ein euklidischer Ring ist, siehe Satz 2.5.7) gibt es $y, s \in \mathbb{Z}$ mit $yx + sn = 1$. Wegen $[1] = [yx + sn] = [yx] + [sn] = [yx] + [0] = [y][x]$ ist $[x]$ invertierbar in $\mathbb{Z}/n\mathbb{Z}$. Weiter gilt $[1] \neq [0]$. Also ist $\mathbb{Z}/n\mathbb{Z}$ ein Körper. □

Definition 2.9.2. Sei p eine Primzahl. Der Körper $\mathbb{Z}/p\mathbb{Z}$ wird meist mit $\mathbb{F}_p$ bezeichnet (dies verwendet Satz 2.9.1).

(Der Buchstabe $\mathbb{F}$ wird vermutlich verwendet, weil *field* das englische Wort für Körper ist. Überdies ist $\mathbb{F}_p$ endlich, also *finite*.)

Beispiel 2.9.3. Wir geben Additions- und Multiplikationstafeln von $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ an (wobei wir $0 = [0]$ und $1 = [1]$ abkürzen).

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Es gilt also $-1 = 1$.

Beispiel 2.9.4. Wir geben Additions- und Multiplikationstafeln von $\mathbb{F}_5 = \mathbb{Z}/5\mathbb{Z}$ an.

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Es gelten beispielsweise $-2 = 3$ und $2^{-1} = 3$ und $4^{-1} = 4$.

Aufgabe 2.9.5. Welche Elemente von $\mathbb{F}_5$ sind Quadrate?

Finden Sie alle Lösungen der Gleichung $X^4 - 1 = 0$ in $\mathbb{F}_5$.

Finden Sie alle Lösungen der Gleichung $X^3 + X + 1 = 0$ in $\mathbb{F}_5$.

Ende 7. Vorlesung am 12.05.2020

Satz 2.9.6 (Kleiner Satz von Fermat)**.** *Seien p eine Primzahl und $a \in \mathbb{Z}$ eine ganze Zahl mit $p \nmid a$. Dann gilt*

$$a^{p-1} \equiv 1 \pmod{p}$$

oder äquivalent

$$a^{p-1} = 1 \text{ in } \mathbb{F}_p.$$

2.9.7. Durch Multiplikation mit a erhält man $a^p \equiv a \pmod{p}$ oder äquivalent $a^p = a$ in $\mathbb{F}_p$. Dies bedeutet, dass alle Elemente von $\mathbb{F}_p$ die Gleichung $X^p - X = 0$ erfüllen.

Beweis. Da $\mathbb{F}_p$ ein Körper ist, sind all seine Elemente $\neq 0$ invertierbar, d. h. die Gruppe $\mathbb{F}_p^\times$ der Einheiten hat $p - 1$ Elemente. Wegen $p \nmid a$ gilt $[a] \in \mathbb{F}_p^\times$. Nach Korollar 1.7.6 folgt $[a]^{p-1} = [1]$ in $\mathbb{F}_p$, d. h. $a^{p-1} \equiv 1 \pmod{p}$ in zahlentheoretischer Schreibweise. □

Beispiel 2.9.8. Für $p = 13$ und $a = 2$ gilt

$$2^{12} = 4096 \equiv 196 \equiv 66 \equiv 1 \pmod{13}.$$

Für $p = 5$ und $a = 3$ gilt

$$3^4 = 81 \equiv 1 \pmod{5}.$$

Beispiel 2.9.9. Dieser Satz kann verwendet werden um zu zeigen, dass eine Zahl keine Primzahl ist: $2^{14} = (2^4)^3 \cdot 2^2 = 16^3 \cdot 4 \equiv 1^3 \cdot 4 \equiv 4 \not\equiv 1 \pmod{15}$. Also ist 15 nach dem Kleinen Fermatschen Satz 2.9.6 keine Primzahl.

Zeigen Sie analog, dass 91 keine Primzahl ist (Hinweis: $a = 2$).

Satz 2.9.10 (hier ohne Beweis)**.** *Sei p eine Primzahl. Die Gruppe $\mathbb{F}_p^\times$ der Einheiten ist eine zyklische Gruppe (mit $p-1$ Elementen), d. h. es gibt ein $w \in \mathbb{F}_p^\times$ mit $\mathbb{F}_p^\times = \langle w \rangle$.*

Beispiel 2.9.11. (a) Ein Erzeuger von $\mathbb{F}_5^\times$ ist 2, d. h. $\mathbb{F}_5^\times = \langle 2 \rangle$, denn es gilt $2^0 = 1$, $2^1 = 2, 2^2 = 4, 2^3 = 8 = 3$. Ebenso gilt $\mathbb{F}_5^\times = \langle 3 \rangle$.

Die Elemente 1 und 4 sind keine Erzeuger von $\mathbb{F}_5^\times$.

(b) Es gilt $\mathbb{F}_7^\times \neq \langle 2 \rangle$, denn $2^1 = 2, 2^2 = 4, 2^3 = 1$.

(c) Es gilt $\mathbb{F}_7^\times = \langle 3 \rangle = \{3, 2, 6, 4, 5, 1\}$.

Bemerkung 2.9.12 (Diskretes Logarithmenproblem)**.** Sei eine zyklische Gruppe G zusammen mit einem Erzeuger g gegeben. Dann ist die Abbildung

$$\begin{aligned} \mathbb{Z} &\to G, \\ a &\mapsto g^a, \end{aligned}$$

offensichtlich surjektiv (und ein Gruppenhomomorphismus) (vgl. Satz 1.7.5 für eine präzisere Aussage). Das diskrete Logarithmenproblem besteht darin, für ein Element $x \in G$ ein $a \in \mathbb{Z}$ mit $g^a = x$ zu finden. Dann ist nämlich „a ein Logarithmus von x zur Basis g“.

Für die zyklische Gruppe $G = \mathbb{F}_p^\times$, wobei p ein große Primzahl ist, gilt dieses Problem als schwer (in dem Sinne, dass heutzutage kein effizientes Lösungsverfahren bekannt ist).

Es kommt aber sehr auf die Wahl der Gruppe an. Für die zyklische Gruppe $G = \mathbb{Z}/n\mathbb{Z}$ mit Erzeuger $g = [1]$ ist dieses Problem trivial.

2.10. **Public-Key-Kryptographie.**

Algorithmus 2.10.1 (Schlüsselvereinbarung mit dem Verfahren nach Diffie–Hellman–Merkle von 1976)**.** [20] Alice und Bob wollen sich auf einen nur ihnen bekannten Schlüssel K einigen, haben aber nur eine Kommunikationsverbindung (z. B. das Internet), die abgehört werden kann.

(1) Sie wählen öffentlich eine sehr große Primzahl $p > 2$ und eine ganze Zahl $2 \leq g \leq p-1$ mit[21] $\mathbb{F}_p^\times = \langle g \rangle$. Damit gilt (siehe Satz 1.7.5)

$$\begin{aligned} \{0, 1, 2, \ldots, p-2\} &\xrightarrow{\sim} \mathbb{F}_p^\times, \\ n &\mapsto g^n. \end{aligned}$$

(2) Alice wählt geheim ein $a \in \{0, \ldots, p-2\}$, berechnet $A := g^a \pmod{p}$ und schickt A an Bob. (Die Zahl a is Alices *private key*, A ist ihr *public key*.)

(3) Bob wählt geheim ein $b \in \{0, \ldots, p-2\}$, berechnet $B := g^b \pmod{p}$ und schickt B an Alice. (Die Zahl b is Bobs *private key*, B ist sein *public key*.)

(4) Alice berechnet $B^a \equiv (g^b)^a \equiv g^{ba} \pmod{p}$.

(5) Bob berechnet $A^b \equiv (g^a)^b \equiv g^{ab} \pmod{p}$.

[20] Der Begriff *Schlüsselvereinbarung* erscheint mir sinnvoller als der oft verwendete Begrif Schlüsselaustausch (key exchange).

[21] Es gibt wohl keine effizienten Algorithmen, ein solches g zu bestimmen, jedoch gibt es bessere Methoden, als einfach alle möglichen Zahlen naiv zu testen.

(6) Nun kennnen beide $K := g^{ab} \pmod p$. Dies ist ihr gemeinsamer Schlüssel (*shared secret key*).

öffentlich bekannt sind: $p, g, A = g^a \pmod p$, $B = g^b \pmod p$, aber nicht a und b. Wenn die Zahlen groß sind, dann dauert es nach heutigem Wissen sehr lange, aus diesen Informationen $g^{ab} = (g^a)^b = (g^b)^a$ zu berechnen.

Das Problem, a aus A (bzw. b aus B) zu bestimmen, ist das diskrete Logarithmenproblem.

2.10.2. Dieses Verfahren funktioniert analog für alle zyklischen Gruppen, in denen das diskrete Logarithmenproblem schwer ist.

Algorithmus 2.10.3 (RSA-Verschlüsselungsverfahren (Rivest–Shamir–Adleman 1977))**.**

(1) Alice wählt zufällig zwei verschiedene Primzahlen $p > 2$ und $q > 2$, sowie eine natürliche Zahl $1 < e < (p-1)(q-1)$, mit
$$\mathrm{ggT}(e, (p-1)(q-1)) = 1.$$
(2) Weiter berechnet Alice mit dem euklidischen Algorithmus ganze Zahlen d und t mit $1 = de + t(p-1)(q-1)$. Wir können ohne Einschränkung zusätzlich annehmen (um Rechenarbeit zu sparen), dass $1 \leq d < (p-1)(q-1)$ gilt (dividiere d mit Rest durch $(p-1)(q-1)$; dabei tritt der Fall $d = 0$ nicht ein, denn sonst wäre $(p-1)(q-1)$ eine Einheit in $\mathbb{Z}$, also ± 1). Es gilt dann
$$de \equiv 1 \pmod{(p-1)(q-1)}.$$
(3) Alice berechnet $n := pq$.
(4) Alice veröffentlicht das Paar (n, e) (ihren *public key*) und hält die Zahl d (ihren *private key*) geheim. (Sie kann p und q vergessen, darf diese aber nicht veröffentlichen!)
(5) Will eine beliebige Person, sagen wir Charly, Alice eine Nachricht m mit $0 \leq m < n$ schicken, so berechnet er $c := m^e \pmod n$ und schickt c an Alice.
(Jeder kann dies tun, da (n, e) öffentlich ist!).
(6) Alice berechnet $c^d \pmod n$. Fassen wir dies als ganze Zahl zwischen 0 und $n-1$ auf, so ist das die ursprüngliche Nachricht m, denn es gilt $c^d = (m^e)^d = m^{ed} = m \pmod n$ wegen Satz 2.10.5.

2.10.4. Da kein effizientes Verfahren bekannt ist, die RSA-Verschlüsselung zu knacken (also ein mögliches d aus n und e zu berechnen), gilt diese als sicher. Ein möglicher Angriff wäre, die beiden Primfaktoren p und q von n zu bestimmen. Das Faktorisierungsproblem gilt aber für sehr große Zahlen als schwer (im Sinne, dass kein effizientes Verfahren bekannt ist).

Satz 2.10.5 (Korrektheit des RSA-Verfahrens 2.10.3)**.** *Unter den obigen Voraussetzungen gilt*
$$m^{de} \equiv m \pmod n.$$

Beweis. Da $p \neq q$ verschiedene Primzahlen sind, besagt der Chinesische Restsatz 2.6.2, dass die offensichtliche Abbildung ein Isomorphismus
$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/(pq)\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} = \mathbb{F}_p \times \mathbb{F}_q$$
von Ringen ist. Die behauptete Gleichheit ist also äquivalent zu den beiden Gleichheiten
$$m^{de} = m \quad \text{in } \mathbb{F}_p,$$
$$m^{de} = m \quad \text{in } \mathbb{F}_q.$$

Da die Rollen von p und q im folgenden Argument austauschbar sind, genügt es, die erste Gleichheit zu zeigen.

Sie gilt sicher, falls $m = 0$ in $\mathbb{F}_p$ gilt. Gelte nun $m \neq 0$ in $\mathbb{F}_p$. Wähle $t \in \mathbb{Z}$ mit $1 = de + t(p-1)(q-1)$ wie in Schritt (2) des RSA-Verfahrens. In $\mathbb{F}_p$ erhalten wir dann unter Verwendung des Kleinen Fermatschen Satzes 2.9.6 wie gewünscht

$$m^{de} = m^{1-t(p-1)(q-1)} = m(m^{p-1})^{-t(q-1)} = m \cdot 1 = m. \qquad \square$$

3. Lineare Gleichungssysteme

3.0.1. Sei K stets ein Körper (z. B. $K = \mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ oder $K = \mathbb{F}_p$ für eine Primzahl p). Wir empfehlen dem Leser, zunächst anzunehmen, dass $K = \mathbb{Q}$ oder $K = \mathbb{R}$ gilt. Dadurch motiviert bezeichnen wir im folgenden Elemente von K als Zahlen.

3.1. Lineare Gleichungssysteme und Matrizen.

Definition 3.1.1. Seien $m, n \in \mathbb{N}$ natürliche Zahlen. Ein **lineares Gleichungssystem** (**LGS**) über K mit m Gleichungen in n Variablen (oder Unbekannten) $X_1, X_2, \ldots, X_n$ ist ein System von Gleichungen der Form

$$\begin{array}{ccccccccc} a_{11}X_1 & + & a_{12}X_2 & + & \cdots & + & a_{1n}X_n & = & b_1, \\ a_{21}X_1 & + & a_{22}X_2 & + & \cdots & + & a_{2n}X_n & = & b_2, \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots & \vdots \\ a_{m1}X_1 & + & a_{m2}X_2 & + & \cdots & + & a_{mn}X_n & = & b_m, \end{array} \tag{3.1.1}$$

wobei $a_{ij} \in K$, für $1 \leq i \leq m$ und $1 \leq j \leq n$, und $b_i \in K$, für $1 \leq i \leq m$. Die a_{ij} heißen **Koeffizienten** des LGS. Ein LGS der Form (3.1.1) heißt genau dann **homogen**, wenn $b_1 = b_2 = \ldots = b_m = 0$ gilt. Sonst heißt es **inhomogen**.

Ein n-Tupel $(x_1, \ldots, x_n) \in K^n$ von Zahlen heißt genau dann **Lösung** des LGS (3.1.1), wenn sämtliche Gleichungen des LGS erfüllt sind, wenn wir die Zahl x_i für die Variable X_i einsetzen, für alle $1 \leq i \leq n$.

Es ist üblich und sinnvoll (wie wir später sehen werden), Lösungen als sogenannte **Spaltenvektoren** $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ zu notieren.

Die **Lösungsmenge** des LGS (3.1.1) ist die Menge aller Lösungen, also

$$\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n \mid x_1, \ldots, x_n \text{ erfüllen (3.1.1)} \right\}.$$

Beispiel 3.1.2. (a) Betrachte das folgende LGS in 3 Variablen mit reellen Koeffizienten:

$$\begin{array}{ccccccccc} 2X_1 & + & X_2 & + & 4X_3 & = & -1, \\ X_1 & & & & -X_3 & = & 0 \end{array} \tag{3.1.2}$$

Die Lösungsmenge ist:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in K^3 \mid \begin{array}{ccccc} 2x_1 & +x_2 & +4x_3 & = -1, \\ x_1 & & -x_3 & = 0 \end{array} \right\} = \left\{ \begin{pmatrix} t \\ -1-6t \\ t \end{pmatrix} \mid t \in K \right\}.$$

(b) Ergänzen wir das LGS (3.1.2) um die Gleichung $X_1 = 1$, so hat das so erhaltene LGS genau eine Lösung, nämlich $\begin{pmatrix} 1 \\ -7 \\ 1 \end{pmatrix}$.

(Diese Aussage bleibt auch richtig, wenn wir zusätzlich etwa die Gleichung $7X_1 + X_2 = 0$ fordern.)

(c) Fordern wir zusätzlich zu $X_1 = 1$ auch noch die Gleichung $X_1 = 0$, so hat das so erhaltene LGS keine Lösung mehr, seine Lösungsmenge ist leer.

Definition 3.1.3. Seien $m, n \in \mathbb{N}$. Eine $(m \times n)$**-Matrix mit Einträgen in** K ist eine Abbildung $A\colon \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to K$. Wir schreiben meist a_{ij} (oder A_{ij}) statt $A(i, j)$ und stellen A wie folgt graphisch dar:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{3.1.3}$$

Man schreibt meist $A = (a_{ij})_{i=1,j=1}^{m,n}$ oder $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ oder abkürzend $A = (a_{ij})$.

Die **Zeilen** von A sind die n-Tupel $(a_{i1}, a_{i2}, \ldots, a_{in}) \in K^n$, für $1 \leq i \leq m$. Die **Spalten** von A sind die Spaltenvektoren $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$, für $1 \leq j \leq n$.

Bemerkung 3.1.4. Ein Spaltenvektor mit m Einträgen aus K ist dasselbe wie eine $(m \times 1)$-Matrix mit Einträgen aus K.

Definition 3.1.5. Gegeben ein LGS der Form (3.1.1), heißt die $(m \times n)$-Matrix $A = (a_{ij})_{i=1,j=1}^{m,n}$ die **Koeffizientenmatrix** des LGS. Ergänzt man diese Matrix rechts um den Spaltenvektor $\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$, so erhält man die **erweiterte Koeffizientenmatrix**, bei der man zusätzlich meist einen senkrechten Strich ergänzt:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Formal ist sie eine $(m \times (n + 1))$-Matrix.

3.1.6. Die erweiterte Koeffizientenmatrix ist eine effiziente Schreibweise für ein LGS.

Beispiel 3.1.7. Die erweiterte Koeffizientenmatrix des LGS (3.1.2) ist

$$\left(\begin{array}{ccc|c} 2 & 1 & 4 & -1 \\ 1 & 0 & -1 & 0 \end{array}\right).$$

3.2. **Das Gauß-Verfahren.**

Definition 3.2.1. Eine **elementare Zeilenumformung** einer $(m \times n)$-Matrix oder eines LGS ist eine der folgenden Operationen:

(a) **Zeilenumformung vom Typ I**, Addieren: Seien $\lambda \in K$ und $1 \leq i \neq j \leq m$. Addiere das λ-fache der i-ten Zeile zur j-ten Zeile, kurz: $Z_j \rightsquigarrow Z_j + \lambda \cdot Z_i$.

(b) **Zeilenumformung vom Typ II**, Vertauschen: Seien $1 \leq i \neq j \leq m$. Vertausche die i-te und die j-te Zeile, kurz: $Z_i \leftrightsquigarrow Z_j$.

(c) **Zeilenumformung vom Typ III**, Skalieren: Sei $\lambda \in K^\times = K \setminus \{0\}$ und $1 \leq i \leq m$. Multipliziere die i-te Zeile mit λ, kurz: $Z_i \rightsquigarrow \lambda \cdot Z_i$.

Satz 3.2.2. *Elementare Zeilenumformungen ändern die Lösungsmenge eines LGS nicht.*

Beweis. Es ist offensichtlich, dass jede Lösung eines LGS auch eine Lösung des LGS ist, das man durch eine elementare Zeilenumformung erhält. Für Typ I (Addieren, $Z_j \rightsquigarrow Z_j + \lambda Z_i$) sieht man das beispielsweise so: Gilt $a_{s1}x_1 + \ldots a_{sn}x_n = b_n$ für alle $s = 1, \ldots, m$, so erhalten wir

$$\begin{aligned}(a_{j1} + \lambda a_{i1})x_1 + \ldots (a_{jn} + \lambda a_{in})x_n &= a_{j1}x_1 + \ldots a_{jn}x_n + \lambda(a_{i1}x_1 + a_{in}x_n) \\ &= b_j + \lambda b_i.\end{aligned}$$

Die Menge der Lösungen bleibt dabei gleich, denn man kann elementare Zeilenumformungen durch elementare Zeilenumformungen rückgängig machen:

Typ	Operation	Umkehroperation
I	Addiere λ-fache i-te Zeile zu j-ter Zeile	Addiere $(-\lambda)$-fache i-te Zeile zu j-ter Zeile
II	Vertausche i-te und j-te Zeile	dieselbe Operation
III	Multipliziere i-te Zeile mit $\lambda \in K^\times$	Multipliziere i-te Zeile mit $\lambda^{-1} \in K^\times$

Die Behauptung folgt. □

3.2.3. Das **Gauß-Verfahren** (oder **Gaußsches Eliminationsverfahren**), das wir zunächst an einem Beispiel erklären, bringt ein beliebiges LGS durch elementare Zeilenumformungen auf eine so einfache Gestalt, dass man die Lösungen sofort ablesen kann. Wir hoffen, dass der Leser dem Beispiel bereits das allgemeine Verfahren entnehmen kann, beschreiben es aber später präzise und erklären, warum es die korrekte Lösungsmenge liefert.

Beispiel 3.2.4. Gegeben sei das LGS

$$\begin{array}{rcrcrcrcr} 5X_1 & - & 5X_2 & - & 5X_3 & - & 25X_4 & = & -60, \\ -3X_1 & + & 3X_2 & + & 3X_3 & + & 17X_4 & = & 42, \\ 2X_1 & + & X_2 & + & 7X_3 & - & 4X_4 & = & -9, \end{array}$$

oder äquivalent seine erweiterte Koeffizientenmatrix

$$\left(\begin{array}{cccc|c} 5 & -5 & -5 & -25 & -60 \\ -3 & 3 & 3 & 17 & 42 \\ 2 & 1 & 7 & -4 & -9 \end{array}\right).$$

Multiplizere die erste Zeile mit $\frac{1}{5}$, symbolisch $Z_1 \rightsquigarrow \frac{1}{5}Z_1$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ -3 & 3 & 3 & 17 & 42 \\ 2 & 1 & 7 & -4 & -9 \end{array}\right)$$

Addiere das 3-fache der 1. Zeile zur 2. Zeile, symbolisch $Z_2 \rightsquigarrow Z_2 + 3Z_1$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ 0 & 0 & 0 & 2 & 6 \\ 2 & 1 & 7 & -4 & -9 \end{array}\right)$$

Addiere das (-2)-fache der 1. Zeile zur 3. Zeile, d. h. subtrahiere das 2-fache der 1. Zeile von der 3. Zeile, symbolisch $Z_3 \rightsquigarrow Z_3 - 2Z_1$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ 0 & 0 & 0 & 2 & 6 \\ 0 & 3 & 9 & 6 & 15 \end{array}\right)$$

Vertausche 2. und 3. Zeile, symbolisch $Z_2 \leftrightsquigarrow Z_3$.

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ 0 & 3 & 9 & 6 & 15 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right)$$

Multipliziere die 2. Zeile mit $\frac{1}{3}$, symbolisch $Z_2 \rightsquigarrow \frac{1}{3}Z_2$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ 0 & 1 & 3 & 2 & 5 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right)$$

Multipliziere die 3. Zeile mit $\frac{1}{2}$, symbolisch $Z_3 \rightsquigarrow \frac{1}{2}Z_3$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ 0 & 1 & 3 & 2 & 5 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

$Z_2 \rightsquigarrow Z_2 - 2Z_3$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & -5 & -12 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

$Z_1 \rightsquigarrow Z_1 + 5Z_3$:

$$\left(\begin{array}{cccc|c} 1 & -1 & -1 & 0 & 3 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

$Z_1 \rightsquigarrow Z_1 + Z_2$:

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

Als LGS geschrieben:

$$\begin{array}{rcrcrcr} X_1 & & & + & 2X_3 & = & 2 \\ & & X_2 & + & 3X_3 & = & -1 \\ & & & & X_4 & = & 3 \end{array}$$

Nun kann man die Lösungsmenge ablesen (man geht die Gleichungen von unten her durch und sucht geeignete Werte für die Variablen in der Reihenfolge $X_4, X_3, \ldots, X_1$). Sie ist gegeben durch

$$\left\{ \begin{pmatrix} 2-2x_3 \\ -1-3x_3 \\ x_3 \\ 3 \end{pmatrix} \mid x_3 \in K \right\}.$$

Ende 8. Vorlesung am 14.05.2020

Definition 3.2.5. Sei A eine $(m \times n)$-Matrix. Dann ist A genau dann in **Zeilenstufenform** (abgekürzt ZSF), wenn gelten:

(a) Falls es Nullzeilen (also Zeilen, die nur Nullen enthalten) gibt, stehen diese nur am Ende.
(b) Der erste Eintrag $\neq 0$ in jeder Zeile ist 1 (dieser Eintrag mitsamt seiner Position heißt **Pivot**).
(c) Für alle $i = 1, \ldots, m-1$ gilt: Der Pivot der $(i+1)$-ten Zeile (falls er existiert) steht echt rechts vom Pivot der i-ten Zeile.
(d) Alle Einträge oberhalb eines Pivots sind 0.

A hat dann also die folgende Form, wobei jedes Symbol $*$ für ein beliebiges Element von K steht:

$$\begin{pmatrix}
0\cdots 0 & 1*\cdots* & 0*\cdots* & 0*\cdots & \cdots & 0*\cdots* \\
0\cdots & \cdots\cdots 0 & 1*\cdots* & 0*\cdots & \cdots & 0*\cdots* \\
0\cdots & \cdots\cdots & \cdots\cdots 0 & 1*\cdots & \cdots & 0*\cdots* \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0\cdots & \cdots & \cdots & \cdots & \cdots 0 & 1*\cdots* \\
0\cdots & \cdots & \cdots & \cdots & \cdots & \cdots 0 \\
\vdots & & & & & \vdots \\
0\cdots & \cdots & \cdots & \cdots & \cdots & \cdots 0
\end{pmatrix}$$

Satz 3.2.6. *Jede Matrix kann durch elementare Zeilenumformungen (unter denen sich die Lösungsmenge nicht ändert, siehe Satz 3.2.2) auf Zeilenstufenform gebracht werden. Das im Beweis erklärte Verfahren ist das* ***Gauß-Verfahren****.*

Beweis. Sei A eine $(m \times n)$-Matrix. Wir verwenden Induktion über die Anzahl m der Zeilen. (Im Fall $m = 0$ ist A bereits in ZSF.)

Induktionsanfang $m = 1$**:** Falls A nur aus Nullen besteht, ist A bereits in Zeilenstufenform. Sonst hat A die Gestalt $A = (0 \ldots 0 c * \cdots *)$ für ein $c \neq 0$. Multipliziere die erste Zeile mit c^{-1} (Operation vom Typ III). Dann ist A in Zeilenstufenform.

Induktionsschritt $m \geq 2$**:** Falls A nur aus Nullen besteht, so ist A bereits in ZSF. Sonst suche von links beginnend die erste Spalte, die einen Eintrag $c \neq 0$ enthält.

Transportiere diesen Eintrag durch Zeilenvertauschung (Operation vom Typ II) in die 1. Zeile.

Multipliziere die 1. Zeile mit c^{-1} (Typ III).

Ziehe geeignete Vielfache der 1. Zeile von den anderen Zeilen ab, so dass unter dem Eintrag nur Nullen stehen (Typ I).

Wir erhalten eine Matrix der Form (die vertikalen Striche haben nichts mit dem vertikalen Strich in der erweiterten Koeffizientenmatrix zu tun)

$$\left(\begin{array}{ccc|c|ccc} 0 & \cdots & 0 & 1 & * & \cdots & * \\ \hline 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{array}\right) = \left(\begin{array}{ccc|c|c} 0 & \cdots & 0 & 1 & z \\ \hline 0 & \cdots & 0 & 0 & \\ \vdots & \ddots & \vdots & \vdots & B \\ 0 & \cdots & 0 & 0 & \end{array}\right). \tag{3.2.1}$$

Hier ist B eine Matrix mit $m-1$ Zeilen und z ist ein **Zeilenvektor** (= eine Matrix mit genau einer Zeile = ein Tupel von Zahlen).

Nach Induktionsvoraussetzung können wir B durch elementare Zeilenumformungen auf ZSF B' bringen. Diese Zeilenumformungen können wir in offensichtlicher Weise als Zeilenumformungen der gesamten Matrix (3.2.1) auffassen. Wir erhalten so die Matrix

$$\left(\begin{array}{ccc|c|c} 0 & \cdots & 0 & 1 & z \\ \hline 0 & \cdots & 0 & 0 & \\ \vdots & \ddots & \vdots & \vdots & B' \\ 0 & \cdots & 0 & 0 & \end{array}\right).$$

Durch Operationen vom Typ I machen wir alle Einträge in z oberhalb von Pivots von B' zu Nullen. Die entstehende Matrix ist in ZSF. □

Bemerkung 3.2.7. Man kann zeigen, dass die ZSF einer Matrix eindeutig bestimmt ist (Beweis per Induktion mit Hilfe der späteren Bemerkung 5.9.4). Wir sprechen daher von *der* ZSF statt von *einer* ZSF.

Bemerkung 3.2.8 (Ablesen der Lösungen bei ZSF)**.** Sei A die erweiterte Koeffizienten-Matrix eines linearen Gleichungssystems. Wir nehmen an, dass A bereits in ZSF ist.

1. Fall: In der letzten Spalte steht ein Pivot, d. h. (hier ist der vertikale Strich wieder der Markierungsstrich in der erweiterten Koeffizientenmatrix)

$$A = \left(\begin{array}{ccc|c} & & & 0 \\ & A' & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{array}\right)$$

wobei A' eine geeignete Matrix ist. Dann besitzt das LGS keine Lösung.

2. Fall: In der letzten Spalte steht kein Pivot.

$$A = \left(\begin{array}{ccc|c} \ddots & \ddots & \ddots & \vdots \\ \ddots & 1 * \cdots * & 0 * \cdots * & * \\ \cdots & \cdots\cdots 0 & 1 * \cdots * & * \\ 0 & \cdots & \cdots 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots 0 & 0 \end{array}\right).$$

Alle Lösungen des zugehörigen LGS kann man dann wie folgt bestimmen: Für jede Unbekannte X_i, deren zugehörige Spalte i kein Pivot enthält, wähle man einen beliebigen Wert $x_i \in K$. Die Werte der verbleibenden Variablen sind dann in offensichtlicher Weise eindeutig bestimmt. Man erhält so eine Lösung, und jede Lösung muss diese Gestalt haben.

Ist r die Anzahl der Pivots (= die Anzahl der Nicht-Nullzeilen), so gibt es also $n - r$ Variablen, die wir frei wählen können. Die Menge der Lösungen ist also bijektiv zu der Menge K^{n-r}.[22]

Man beachte, dass der erste Fall nur im Falle eines inhomogenen LGS eintreten kann, denn jedes homogene LGS hat die triviale Lösung, bei der alle Variablen Null sind.

Beispiel 3.2.9. Wir betrachten das LGS

$$\begin{array}{rrrl} X_1 + 6X_2 & & +X_4 & = 1 \\ & X_3 & +2X_4 & = 3 \\ & & 0 & = 0 \end{array}$$

bzw. seine erweiterte Koeffizientenmatrix

$$\left(\begin{array}{cccc|c} \boxed{1} & 6 & 0 & 1 & 1 \\ 0 & 0 & \boxed{1} & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Diese ist in ZSF und hat $r = 2$ Pivots eingerahmt dargestellte Pivots.

Die Lösungsmenge des LGS ist

$$\left\{ \begin{pmatrix} 1 - 6x_2 - x_4 \\ x_2 \\ 3 - 2x_4 \\ x_4 \end{pmatrix} \mid x_2, x_4 \in K \right\}.$$

Ist nämlich $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ eine Lösung, so liefert die erste Gleichung $x_1 = 1 - 6x_2 - x_4$ und die zweite $x_3 = 3 - 2x_4$. Also hat jede Lösung die angegebene Gestalt. Umgekehrt sind die Elemente der obigen Menge offensichtlich Lösungen.

[22] Gilt $K = \mathbb{F}_p$, so gibt es also p^{n-r} Lösungen.

3.2.10. Wir erinnern daran, dass ein LGS der Form (3.1.1) genau dann homogen heißt, wenn $b_1 = b_2 = \ldots = b_m = 0$ gilt. Ein homogenes LGS hat stets eine „triviale“ Lösung, nämlich den „Nullvektor“ $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

Satz 3.2.11. *Sei ein homogenes LGS*

$$\begin{array}{ccccccccc} a_{11}X_1 & + & a_{12}X_2 & + & \cdots & + & a_{1n}X_n & = & 0, \\ a_{21}X_1 & + & a_{22}X_2 & + & \cdots & + & a_{2n}X_n & = & 0, \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots & \vdots \\ a_{m1}X_1 & + & a_{m2}X_2 & + & \cdots & + & a_{mn}X_n & = & 0 \end{array}$$

mit m Gleichungen in n Unbekannten gegeben. Gilt $m < n$ („mehr Variablen als Gleichungen“), so hat das LGS eine nicht-triviale Lösung, also eine Lösung $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ mit $x_i \neq 0$ für ein $1 \leq i \leq n$.

Beweis. Wir bringen unser LGS mit dem Gauß-Verfahren durch elementare Zeilenumformungen auf Zeilenstufenform (siehe Satz 3.2.6). Bei diesen Umformungen bleibt das LGS homogen und die Lösungsmenge ändert sich nicht (Satz 3.2.2). Sei r die Anzahl der Pivots der zugehörigen Koeffizientenmatrix. Dann gilt $r \leq m$, denn in jeder der m Zeilen steht maximal ein Pivot. Nach Bemerkung 3.2.8 können wir die Werte von $n - r$ Variablen frei wählen (der erste Fall in dieser Bemerkung tritt nicht ein, denn unser LGS ist homogen hat mindestens den Nullvektor als Lösung). Wegen $n - r \geq n - m > 0$ können wir mindestens eine Variable $\neq 0$ wählen. □

Aufgabe 3.2.12. Finden Sie eine nicht-triviale Lösung des homogenen LGS

$$\begin{array}{ccccccc} x_1 & + & 2x_2 & + & x_3 & = & 0, \\ 2x_1 & + & x_2 & & & = & 0. \end{array}$$

Aufgabe 3.2.13. Bestimmen Sie ein Polynom $f \in \mathbb{Q}[X]$ vom Grad 3 mit $f(-1) = 2$, $f(0) = 4$, $f(1) = 10$, $f(2) = 26$.

3.3. Zur Struktur der Lösungsmenge eines homogenen LGS.

3.3.1. Der folgende Satz ist als Motivation für den Begriff eines Vektorraums (siehe Definition 4.1.1 im nächsten Kapitel) gedacht. Wir werden ihn und auch die Aussage der nachfolgenden Aufgabe 3.3.5 in 5.7.2 und Satz 5.7.4 in kürzerer Schreibweise zeigen.

Satz 3.3.2. *Sei ein homogenes LGS in n Variablen über einem Körper K gegeben. Dann gelten:*

(a) Das LGS besitzt eine „triviale“ Lösung, nämlich den „Nullvektor“ $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

(b) *Jedes „skalare Vielfache“ einer Lösung ist eine Lösung: Seien* $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *eine Lösung und* $\lambda \in K$*. Dann ist* $\begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$ *ebenfalls eine Lösung.*

(c) *Die „Summe“ zweier Lösungen ist eine Lösung: Seien* $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *und* $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ *Lösungen. Dann ist auch* $\begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$ *eine Lösung.*

Bemerkung 3.3.3. Satz 3.3.2 besagt, dass die Lösungsmenge eines homogenen LGS ein K-Vektorraum im Sinne der Definition 4.1.1 im nächsten Kapitel ist.

Beweis. Wir dürfen annehmen, dass unser LGS die Gestalt (3.1.1) hat, wobei $b_1 = \cdots = b_m = 0$ gilt.

(a) Dies ist klar, denn

$$a_{i1}0 + \cdots + a_{in}0 = 0 \qquad \text{für alle } 1 \leq i \leq m.$$

(b) Aus

$$a_{i1}x_1 + \cdots + a_{in}x_n = 0 \qquad \text{für alle } 1 \leq i \leq m$$

folgt per Multiplikation mit λ

$$a_{i1}(\lambda x_1) + \cdots + a_{in}(\lambda x_n) = 0 \qquad \text{für alle } 1 \leq i \leq m.$$

(c) Aus

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= 0 \qquad \text{für alle } 1 \leq i \leq m \text{ und} \\ a_{i1}y_1 + \cdots + a_{in}y_n &= 0 \qquad \text{für alle } 1 \leq i \leq m \end{aligned}$$

folgt per Addition

$$a_{i1}(x_1 + y_1) + \cdots + a_{in}(x_n + y_n) = 0 \qquad \text{für alle } 1 \leq i \leq m.$$

□

3.3.4. Jedem LGS der Form (3.1.1) kann man ein homogenes LGS zuordnen, indem man alle b_i durch 0 ersetzt.

Aufgabe 3.3.5. Sei ein LGS (L) (etwa der Gestalt (3.1.1)) gegeben. Sei (HL) das zugeordnete homogene LGS (siehe 3.3.4). Zeigen Sie:

(a) Ist $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ eine Lösung von (L) und $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ eine Lösung von (HL), so ist $\begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$ eine Lösung von (L).

(b) Sind $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ Lösungen von (L), so ist $\begin{pmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{pmatrix}$ eine Lösung von (HL).

Fazit: Hat (L) eine Lösung x, so erhält man sämtlich Lösungen von (L) durch Addition von Lösungen von (HL) zu x.

3.4. **Beispiele.**

3.4.1. Hier einige Beispiele, die in der Vorlesung vermutlich nicht erklärt werden.

Beispiel 3.4.2 (Zum Gauß-Verfahren)**.** Betrachte das LGS:

$$\begin{array}{rcrcrcl} X_1 & + & 2X_2 & + & X_3 & = & 1 \\ 2X_1 & + & X_2 & & & = & 0 \\ -X_1 & & & & & = & 1 \end{array}$$

Der aufmerksame Leser kann die Lösungsmenge natürlich sofort ablesen. Wir führen trotzdem das Gauß-Verfahren durch.

Die erweiterte Koeffizienten-Matrix ist

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right)$$

$Z_2 \rightsquigarrow Z_2 - 2Z_1$ und $Z_3 \rightsquigarrow Z_3 + Z_1$:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & -3 & -2 & -2 \\ 0 & 2 & 1 & 2 \end{array}\right)$$

$Z_2 \rightsquigarrow -\frac{1}{3}Z_2$:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} \\ 0 & 2 & 1 & 2 \end{array}\right)$$

$Z_3 \rightsquigarrow Z_3 - 2Z_2$ und $Z_1 \rightsquigarrow Z_1 - 2Z_2$ (diese zweite Umformung wird im Gauß-Verfahren, so wie wir es erklärt haben, erst später durchgeführt):

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} \\ 0 & 0 & -\frac{1}{3} & \frac{2}{3} \end{array}\right)$$

$Z_3 \rightsquigarrow -3Z_3$:

$$\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{2}{3} & \frac{2}{3} \\ 0 & 0 & 1 & -2 \end{array}\right)$$

$Z_1 \rightsquigarrow Z_1 + \frac{1}{3}Z_3$ und $Z_2 \rightsquigarrow Z_2 - \frac{2}{3}Z_3$:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Lösungsmenge:

$$\left\{ \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix} \right\}.$$

Manchmal ist es cleverer, elementare Zeilenumformungen in einer anderen Reihenfolge durchzuführen. So auch hier:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right) \overset{\substack{Z_1 \leftrightarrow Z_3 \\ Z_1 \rightsquigarrow -1 \cdot Z_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 \end{array}\right) \overset{\substack{Z_2 \rightsquigarrow Z_2 - 2Z_1 \\ Z_3 \rightsquigarrow Z_3 - Z_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 2 & 1 & 2 \end{array}\right)$$

$$\overset{Z_3 \rightsquigarrow Z_3 - 2Z_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -2 \end{array}\right)$$

Beispiel 3.4.3 (Zum Ablesen der Lösungen eines LGS in ZSF)**.** Ein LGS sei durch die folgende erweiterte Koeffizientenmatrix gegeben:

$$\left(\begin{array}{ccccccccc|c} \underline{1} & 2 & 0 & 3 & 3 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & \underline{1} & 2 & 4 & 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & \underline{1} & 0 & 3 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{1} & 4 & 1 & 4 \end{array}\right)$$

Das LGS ist also in ZSF. Ablesen der Lösung:

$$\begin{aligned} x_1 &= 1 - 2x_2 - 3x_4 - 3x_5 - x_8, \\ x_3 &= 2 - 2x_4 - 4x_5 - 2x_8 - x_9, \\ x_6 &= 3 - 3x_8 - 2x_9, \\ x_7 &= 4 - 4x_8 - x_9. \end{aligned}$$

Die Lösungsmenge ist also

$$\left\{ \begin{pmatrix} -2x_2 - 3x_4 - 3x_5 - x_8 + 1 \\ x_2 \\ -2x_4 - 4x_5 - 2x_8 - x_9 + 2 \\ x_4 \\ x_5 \\ -3x_8 - 2x_9 + 3 \\ -4x_8 - x_9 + 4 \\ x_8 \\ x_9 \end{pmatrix} \mid x_2, x_4, x_5, x_8, x_9 \in K \right\}.$$

4. Vektorräume

4.0.1. Im gesamten Kapitel 4 bezeichnet K einen beliebigen, fest gewählten Körper.

4.1. **Vektorräume.**

Definition 4.1.1. Ein **Vektorraum (über K)** oder ein **K-Vektorraum** ist eine Menge V mit zwei Abbildungen:

$$\begin{aligned} +&: V \times V \to V, && (v,w) \mapsto v+w, && \textbf{(Addition)} \\ .&: K \times V \to V, && (a,v) \mapsto a.v, && \textbf{(Skalarmultiplikation)} \end{aligned}$$

so dass gelten:

(VR1) $(V,+)$ ist eine abelsche (= kommutative) Gruppe.

Wie in abelschen Gruppen üblich wird das neutrale Element als 0 und das zu $v \in V$ inverse Element als $-v$ notiert.

(VR2) $a.(v+w) = a.v + a.w \quad \forall a \in K,\ \forall v,w \in V.$
(VR3) $(a+b).v = a.v + b.v \quad \forall a,b \in K,\ \forall v \in V.$
(VR4) $(a \cdot b).v = a.(b.v) \quad \forall a,b \in K,\ \forall v \in V.$
(VR5) $1.v = v \quad \forall v \in V.$

Meist schreibt man abkürzend $av := a.v$. Elemente eines Vektorraums nennt man **Vektoren**. Elemente des Körpers nennt man in diesem Zusammenhang **Skalare**. Statt $\mathbb{R}$-Vektorraum bzw. $\mathbb{C}$-Vektorraum sagt man auch reeller bzw. komplexer Vektorraum.

Beispiel 4.1.2. (a) Sei $n \geq 1$ eine natürliche Zahl. Die Menge

$$K^n = \{(x_1,\dots,x_n) \mid x_i \in K\}$$

mit komponentenweiser Addition $(x_1,\dots,x_n)+(y_1,\dots,y_n) := (x_1+y_1,\dots,x_n+y_n)$ und komponentenweiser Skalarmultiplikation $a.(x_1,\dots,x_n) := (ax_1,\dots,ax_n)$ ist ein K-Vektorraum.

Dies ist offensichtlich: (VR1) gilt, da $(K,+)$ eine abelsche Gruppe ist. (VR2) gilt, da $a(b+c) = ab+ac$ für alle $a,b,c \in K$ gilt. (VR3) gilt, da $(a+b)c = ac+bc$ für alle $a,b,c \in K$ gilt. (VR4) gilt, da $(ab)c = a(bc)$ für alle $a,b,c \in K$ gilt. (VR5) gilt, da $1a = a$ für alle $a \in K$ gilt.

(b) Für $n=1$ erhält man den Spezialfall, dass $K = K^1$ selbst ein K-Vektorraum ist.

(c) $\{0\}$ ist ein K-Vektorraum, der **Nullvektorraum**. Dabei sind Addition und Skalarmultiplikation die einzig möglichen Abbildungen.

(d) Der Polynomring $K[X]$ wird ein K-Vektorraum mit der Skalarmultiplikation $\lambda.f = \lambda f$ für $\lambda \in K$ und $f \in K[X]$, wobei wir λ rechts als konstantes Polynom auffassen (jeder Koeffizient von f wird also mit λ multipliziert).

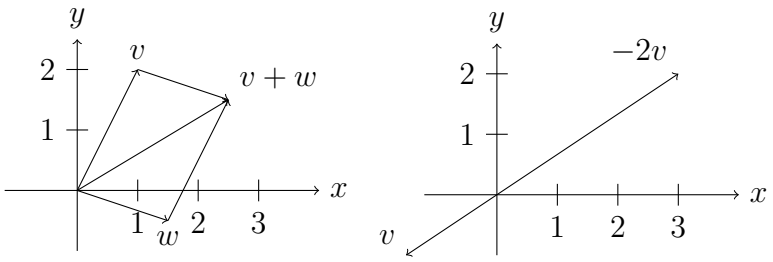
4.1.3 (Vorstellung)**.** (a) Man kann sich $\mathbb{R} = \mathbb{R}^1$ anschaulich als reelle Gerade vorstellen. Ebenso kann man sich $\mathbb{C} = \mathbb{C}^1$ als komplexe Ebene vorstellen.

(b) Man kann sich $\mathbb{R}^2 = \{(x,y) \mid x,y \in \mathbb{R}\}$ anschaulich als Zeichenebene mit gewählter x- und y-Achse (samt Achsenbeschriftung) vorstellen.

Geometrische Interpretation der Addition und Skalarmultiplikation:

- Addition $\widehat{=}$ Aneinandersetzung von Vektoren
- Skalarmultiplikation mit $\alpha \in \mathbb{R}$ $\widehat{=}$ Streckung (oder Skalierung) um den Faktor α. Daher kommt der Name *Skalarmultiplikation*.

(c) Analog kann man sich $\mathbb{R}^3$ vorstellen und vielleicht auch $\mathbb{R}^n$, für $n \geq 4$.



Beispiel 4.1.4. Sei I eine Menge und V ein K-Vektorraum. Dann ist die Menge

$$V^I := \{x\colon I \to V\} = \{(x_i)_{i\in I} \mid x_i \in V \text{ für alle } i \in I\}$$

aller Abbildungen von I nach V mit „punktweiser“ Addition

$(x+y)(i) := x(i)+y(i)$ oder in alternativer Schreibweise $(x_i)_{i\in I}+(y_i)_{i\in I} := (x_i+y_i)_{i\in I}$

und „punktweiser“ Skalarmultiplikation

$$(a.x)(i) := a.(x(i)) \qquad \text{oder} \qquad a.(x_i)_{i\in I} := (ax_i)_{i\in I}$$

ebenfalls ein K-Vektorraum.

- Im Spezialfall $V = K$ und $I = \{1, 2, \dots, n\}$ erhalten wir den Vektorraum $K^n = K^{\{1,\dots,n\}}$.[23]
- Im Spezialfall $V = K$ und $I = \mathbb{N}$ erhalten wir den K-Vektorraum $K^{\mathbb{N}}$ aller Folgen in K. Beispielsweise ist $\mathbb{R}^{\mathbb{N}} = \{(x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{R}\}$ der reelle Vektorraum aller reellen Folgen.

Definition 4.1.5. Seien $m, n \in \mathbb{N}$. Wir schreiben $K^{m\times n}$ für die Menge der $(m\times n)$-Matrizen mit Einträgen in K.

Beispiel 4.1.6. Wir erinnern daran, dass eine $(m \times n)$-Matrix mit Einträgen in K formal eine Abbildung $\{1, \dots, m\} \times \{1, \dots, n\} \to K$ ist (siehe Definition 3.1.3). Nach Beispiel 4.1.4 ist deshalb $K^{m\times n}$ ein K-Vektorraum. Addition und Skalarmultiplikation sind also komponentenweise definiert (ähnlich wie im K^n, siehe Beispiel 4.1.2.(a)): Gegeben Matrizen $A = (a_{ij})_{i=1,j=1}^{m,n} \in K^{m\times n}$ und $B = (b_{ij})_{i=1,j=1}^{m,n} \in K^{m\times n}$ und $\lambda \in K$ gelten also

$$A + B := \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \qquad \text{und} \qquad \lambda.A := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Die **Nullmatrix** $0 := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$ ist das Nullelement des Vektorraums $K^{m\times n}$.

Lemma 4.1.7. *Sei V ein K-Vektorraum. Dann gelten für alle $a, b \in K$ und $v, w \in V$:*

(a) $0v = 0$ *(oder präziser* $0_K.v = 0_V$*),*
(b) $a0 = 0$ *(oder präziser* $a.0_V = 0_V$*),*

[23] Für $n = 0$ gilt $I = \{1, 2, \dots, 0\} = \emptyset$ und es ist sinnvoll, $K^0 := K^\emptyset$ zu definieren. Da es genau eine Abbildung von der leeren Menge nach K gibt (man kann solch eine Abbildung auch als ein 0-Tupel in K auffassen), besteht K^0 nur aus einem Element, dem Nullelement, und K^0 ist der Nullvektorraum.

(c) $(-a)v = -(av) = a(-v)$ *und speziell* $(-1)v = -v$,
(d) $a(v - w) = av - aw$,
(e) $(a - b)v = av - bv$,
(f) $av = 0$ *implizert* $a = 0$ *oder* $v = 0$.

Beweis. (a) Folgt aus $0v = (0 + 0)v = 0v + 0v$ durch Addition von $-(0v)$.
(b) Folgt aus $a0 = a(0 + 0) = a0 + a0$ durch Addition von $-(a0)$.
(c) Folgt aus $av+(-a)v = (a+(-a))v = (a-a)v = 0v = 0$ und $av+a(-v) = a(v+(-v)) = a(v - v) = a0 = 0$ und der Eindeutigkeit des additiven Inversen in der abelschen Gruppe $(V, +)$.
(d) $a(v - w) = a(v + (-w)) = av + a(-w) = av + (-aw) = av - aw$.
(e) $(a - b)v = (a + (-b))v = av + (-b)v = av + (-bv) = av - bv$.
(f) Gelte $av = 0$. Ist $a \neq 0$, so folgt $v = 1v = (a^{-1}a)v = a^{-1}(av) = a^{-1}0 = 0$. □

Ende 9. Vorlesung am 19.05.2020

Definition 4.1.8. Sei V ein K-Vektorraum. Eine Teilmenge $U \subset V$ heißt genau dann **Untervektorraum** oder **Unterraum** oder **Teilraum** von V, wenn gelten:
(a) $0_V \in U$.
(b) Für alle $x, y \in U$ gilt $x + y \in U$.
(c) Für alle $a \in K$ und $x \in U$ gilt $ax \in U$.

4.1.9. Jeder Untervektorraum ist (mit der induzierten Addition und Skalarmultiplikation) selbst ein K-Vektorraum. Der Beweis, dass U mit der induzierten Addition eine abelsche (Unter-)Gruppe ist, verwendet $-u = (-1)u$, siehe Lemma 4.1.7.(c); die restlichen Bedingungen sind offensichtlich.

Bemerkung 4.1.10 (Alternative Definition eines Untervektorraums)**.** Eine Teilmenge $U \subset V$ eines Vektorraums ist genau dann ein Untervektorraum, wenn sie nicht leer ist und wenn für alle $x, y \in U$ und alle $a \in K$ auch $ax + y \in U$ gilt. Wir überlassen den Beweis dem Leser als einfache Übung.

Beispiel 4.1.11.
(a) Jeder K-Vektorraum V hat $\{0\}$ und V als Untervektorräume.
(b) Sei ein *homogenes* LGS in n Variablen gegeben, etwa in der Form (3.1.1) mit $b_1 = \cdots = b_m = 0$. Dann bildet seine Lösungsmenge einen Untervektorraum von K^n (wobei man Elemente von K^n wie im vorigen Kapitel als Spaltenvektoren schreibe). Dies ist der Inhalt von Satz 3.3.2.
Slogan: „Homogene lineare Gleichungssysteme liefern Untervektorräume."
Beispielsweise sind
$$\{(x, y, z) \in \mathbb{R}^3 \mid y = 0\}$$
oder
$$\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$$
Untervektorräume von $\mathbb{R}^3$.
(c) Die Menge $U := \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$ ist kein Untervektorraum von $\mathbb{R}^2$. Beispielsweise gelten $(1, 1) \in U$ und $(2, 4) \in U$, aber $(1, 1) + (2, 4) = (3, 5) \notin U$.
Faustregel (aber nicht immer korrekt): „Gleichungen, die nicht homogen linear sind, liefern keine Untervektorräume."
Beispiele, wo diese Faustregel versagt:
- $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 0\} = \{(0, 0)\}$ ist ein Untervektorraum von $\mathbb{R}^2$.

- $\{(x,y) \in \mathbb{F}_2^2 \mid x = y^2\}$ ist ein Untervektorraum von $\mathbb{F}_2^2$ da $y^2 = y$ für alle $y \in \mathbb{F}_2$ gilt. Dasselbe Beispiel funktioniert in $\mathbb{F}_p^2$ mit der Gleichung $x = y^p$.

(d) Sei I eine Menge. Dann ist

$$K^{(I)} := \{(x_i)_{i\in I} \in K^I \mid \text{nur endlich viele } x_i \text{ sind } \neq 0\}$$

ein Untervektorraum von K^I.

(e) Betrachte den Vektorraum $V = \mathbb{R}^{\mathbb{N}}$ aller reellen Folgen. Seine Teilmenge aller konvergenten Folgen ist ein Untervektorraum. Seine Teilmenge aller Folgen, die ab einem gewissen Folgeglied Null sind, ist ein Untervektorraum.

(f) Im reellen Vektorraum $\mathbb{R}^{\mathbb{R}}$ aller Abbildungen $\mathbb{R} \to \mathbb{R}$ bilden die stetigen (differenzierbaren, polynomialen, konstanten) Funktionen (mit Nullstelle bei 42, falls differenzierbar: mit Ableitung Null bei 3) einen Untervektorraum.

Lemma 4.1.12. *Sei V ein K-Vektorraum und seien U_1 und U_2 Untervektorräume. Dann sind*

*(a) der **Schnitt** $U_1 \cap U_2$ und*

*(b) die **Summe** $U_1 + U_2 := \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$*

Untervektorräume von V.[24] *Es gilt $U_1 \cap U_2 \subset U_i \subset U_1 + U_2$ für $i = 1, 2$.*

4.1.13. Achtung: $U_1 \cup U_2$ ist im Allgemeinen kein Untervektorraum!

Beweis. (a): Dem Leser als Übung überlassen.

(b):

- Wegen $0 \in U_1$ und $0 \in U_2$ gilt $0 = 0 + 0 \in U_1 + U_2$.
- Seien $x, x' \in U_1 + U_2$ beliebig. Dann existieren $u_1, u_1' \in U_1$ und $u_2, u_2' \in U_2$ mit $x = u_1 + u_2$ und $x' = u_1' + u_2'$. Es folgt $x + x' = (u_1 + u_2) + (u_1' + u_2') = \underbrace{(u_1 + u_1')}_{\in U_1} + \underbrace{(u_2 + u_2')}_{\in U_2} \in U_1 + U_2$.
- Sei $x \in U_1 + U_2$. Dann existieren $u_1 \in U_1$ und $u_2 \in U_2$ mit $x = u_1 + u_2$. Für jedes $a \in K$ folgt $ax = a(u_1 + u_2) = \underbrace{au_1}_{\in U_1} + \underbrace{au_2}_{\in U_2} \in U_1 + U_2$.

Die Inklusionen sind offensichtlich. Beispielsweise gilt $U_2 \subset U_1 + U_2$ wegen $u = u + 0 \in U_1 + U_2$ für $u \in U_2$. □

Definition 4.1.14. Sei V ein K-Vektorraum und $M \subset V$ eine Teilmenge. Eine **Linearkombination** (oder genauer **K-Linearkombination** oder **K-lineare Kombination**) von Elementen aus M ist ein Ausdruck der Gestalt

$$a_1 v_1 + \cdots + a_n v_n$$

für geeignete $n \in \mathbb{N}$, $v_1, \ldots, v_n \in M$ und $a_1, \ldots, a_n \in K$ (für $n = 0$ interpretieren wir einen solchen Ausdruck als 0). Die Menge

$$\langle M \rangle := \langle M \rangle_K := \{v \in V \mid v \text{ lässt sich als Linearkombination von Elementen von } M \text{ schreiben}\}$$

heißt **Spann** von M oder **lineare Hülle** von M.

[24] Sinngemäß gilt Lemma 4.1.12 auch für Untervektorräume U_i von V, für $i \in I$, wobei I eine beliebige Indexmenge ist: Der Schnitt $\bigcap_{i\in I} U_i$ und die Summe $\sum_{i\in I} U_i := \{\sum_{f\in F} u_f \mid F \subset I \text{ endliche Teilmenge}, u_f \in U_f \text{ für alle } f \in F\}$ sind Untervektorräume von V.

Gegeben Vektoren $v_1, \ldots, v_n$ sagt man abkürzend „Linearkombination der $v_1, \ldots, v_n$“ statt „Linearkombination von Elementen aus $\{v_1, \ldots, v_n\}$“. Jede solche Linearkombination hat die Form $a_1v_1 + \cdots + a_nv_n$ für geeignete $a_1, \ldots, a_n \in K$. Man schreibt abkürzend $\langle v_1, \ldots, v_m\rangle := \langle\{v_1, \ldots, v_m\}\rangle$.

Beispiel 4.1.15. (a) $\langle\emptyset\rangle = \{0\}$.
(b) Sei $v = (1, 1, 0) \in \mathbb{R}^3$. Dann ist $\langle v\rangle = \{(t, t, 0) \mid t \in \mathbb{R}\}$.

Lemma 4.1.16. *Sei M eine Teilmenge eines Vektorraums V. Dann ist $\langle M\rangle$ der kleinste Untervektorraum von V, der M enthält.*[25] *Er wird deswegen auch als der **von** M **erzeugte Untervektorraum** oder als der **von** M **aufgespannte Untervektorraum** bezeichnet.*

Beweis. Es ist leicht zu sehen, dass $\langle M\rangle$ ein Untervektorraum von V ist, der M enthält. Da jeder Untervektorraum von V, der M enthält, jede Linearkombination von Elementen von M enthält, und $\langle M\rangle$ genau aus diesen Elementen besteht, ist $\langle M\rangle$ der kleinste solche Untervektorraum. □

4.1.17. Auch wenn wir die folgende Aufgabe später leicht mit etwas Theorie lösen können, empfehlen wir dem Leser, sie von Hand zu lösen.

Aufgabe 4.1.18. Sei U ein Untervektorraum von $\mathbb{R}^2$. Dann tritt genau einer der folgenden drei Fälle ein:

(a) $U = \{0\}$.
(b) $U = \langle v\rangle = \{av \mid a \in \mathbb{R}\}$ für ein $v \in \mathbb{R}^2$ mit $v \neq 0$, d. h. U ist die Gerade durch v und den Nullpunkt.
(c) $U = \mathbb{R}^2$.

Aufgabe 4.1.19 (Produkt von Vektorräumen)**.** Seien V und W Vektorräume. Dann ist auch $V \times W$ mit komponentenweiser Addition $(v, w) + (v', w') := (v + v', w + w')$ und Skalarmultiplikation $\lambda.(v, w) := (\lambda.v, \lambda.w)$ ein Vektorraum.

4.2. Lineare Unabhängigkeit, Erzeugendensysteme, Basen.

4.2.1. Gegeben Vektoren $v_1, \ldots, v_n$ eines Vektorraums V mag man sich fragen, ob man jeden Vektor $v \in V$ als Linearkombination $v = a_1v_1 + \cdots + a_nv_n$ schreiben kann, und wenn ja, ob dies auf genau eine Weise so möglich ist. Diese Fragestellungen führen zu den folgenden Definitionen. Bei der Definition der linearen Unabhängigkeit ist dies auf den ersten Blick vielleicht nicht so offensichtlich, vergleiche aber 4.2.6.

Definition 4.2.2. Sei V ein K-Vektorraum. Ein Tupel $(v_1, v_2, \ldots, v_n)$ von Vektoren $v_i \in V$ heißt genau dann

(a) **Erzeugendensystem (ES)**, wenn sich jedes Element $v \in V$ als Linearkombination $v = a_1v_1 + \cdots + a_nv_n$ der Vektoren $v_1, \ldots, v_n$ schreiben lässt, für geeignete Elemente $a_1, \ldots, a_n \in K$, wenn in Formeln also $V = \langle v_1, v_2, \ldots, v_n\rangle$ gilt.
(b) **linear unabhängig (über K)**, wenn für alle $a_1, \ldots, a_n \in K$ gilt: Aus $a_1v_1 + \ldots + a_nv_n = 0$ folgt $a_1 = a_2 = \ldots = a_n = 0$.

[25] Es ist auch der Schnitt über alle Untervektorräume von V, die M enthalten.

(c) **Basis**, wenn es für alle $v \in V$ eindeutig bestimmte $a_1, \ldots, a_n \in K$ mit $v = a_1v_1 + \ldots + a_nv_n$ gibt. Diese Elemente $a_1, \ldots, a_n$ heißen dann die **Koordinaten von** v (bezüglich der Basis $(v_1, \ldots, v_n)$).[26]

Statt „nicht linear unabhängig“ sagen wir meist **linear abhängig**. Das Tupel $(v_1, \ldots, v_n)$ ist also genau dann linear abhängig ist, wenn der Nullvektor eine Darstellung $0 = a_1v_1 + \cdots + a_nv_n$ hat, in der nicht alle a_i Null sind.

Wir sagen auch lax: $v_1, \ldots, v_n$ sind linear unabhängig / linear abhängig / ein Erzeugendensystem / eine Basis anstatt $(v_1, \ldots, v_n)$ sind linear unabhängig / linear abhängig / ein Erzeugendensystem / eine Basis.

4.2.3. Ein Tupel $(v_1, \ldots, v_n)$ hat genau dann eine der vier Eigenschaften aus Definition 4.2.2, wenn jede Umsortierung $(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)})$ diese Eigenschaft hat, für $\sigma \in S_n$.

4.2.4. Ist ein Tupel $(v_1, \ldots, v_n)$ linear unabhängig, so ist auch jedes Teiltupel linear unabhängig: Seien $1 \leq i_1 < \cdots < i_r \leq n$. Dann ist $v_{i_1}, \ldots, v_{i_r}$ linear unabhängig.

Lemma 4.2.5. *Sei $(v_1, v_2, \ldots, v_n)$ ein Tupel von Vektoren in V. Dann sind die beiden folgenden Aussagen äquivalent:*

(a) $(v_1, v_2, \ldots, v_n)$ ist linear unabhängig.
*(b) (**Koeffizientenvergleich**) Für alle Elemente $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$ folgt aus*

$$a_1v_1 + \cdots + a_nv_n = b_1v_1 + \cdots + b_nv_n,$$

dass $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_n = b_n$ gilt.

Beweis. (a) $\Rightarrow$ (b): Aus der Gleichung $a_1v_1 + \cdots + a_nv_n = b_1v_1 + \cdots + b_nv_n$ folgt $(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0$, woraus per linearer Unabhängigkeit $a_i = b_i$ für alle i folgt.
(b) $\Leftarrow$ (a): Gelte $a_1v_1 + \cdots + a_nv_n = 0$. Wegen $0v_1 + \cdots + 0v_n = 0$ folgt dann $a_1 = 0$, $a_2 = 0$, $\ldots$, $a_n = 0$. Dies zeigt die lineare Unabhängigkeit. □

4.2.6. Vektoren $v_1, \ldots, v_n \in V$ sind also genau dann

(a) eine Erzeugendensystem von V, wenn sich jeder Vektor $v \in V$ als $v = a_1v_1 + \cdots + a_nv_n$ schreiben lässt („Existenz der Schreibweise als Linearkombination“);
(b) linear unabhängig, wenn für jeden Vektor v, der sich als $v = a_1v_1 + \cdots + a_nv_n$ schreiben lässt, die Koeffizienten $a_1, \ldots, a_n$ eindeutig bestimmt sind („Eindeutigkeit der Schreibweise als Linearkombination“);
(c) eine Basis von V, wenn sich jeder Vektor $v \in V$ eindeutig als $v = a_1v_1 + \cdots + a_nv_n$ schreiben („Existenz und Eindeutigkeit der Schreibweise als Linearkombination“).

Dies folgt aus den Definitionen und Lemma 4.2.5.

Proposition 4.2.7. *Genau dann ist $(v_1, \ldots, v_n)$ eine Basis von V, wenn $(v_1, \ldots, v_n)$ ein linear unabhängiges Erzeugendensystem von V ist.*

Beweis. Dies folgt sofort aus 4.2.6. □

[26] Insbesondere erlaubt eine Basis, jedes Element von V eindeutig durch seine Koordinaten zu beschreiben. In diesem Sinne ist eine Basis ein Koordinatensystem von V. (Der Begriff eines Koordinatensystems wird aber selten formal definiert.)

Beispiele 4.2.8. (a) Sei $V = K^n$. Die Vektoren $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\ldots$, $e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ bilden offensichtlich eine Basis des K^n. Man nennt $e_1, \ldots, e_n$ die **Standardbasis** des K^n und den Spaltenvektor e_i den **i-ten Standardbasisvektor**. Er hat als i-ten Eintrag eine Eins und besteht sonst aus Nullen.

(b) Das (leere 0-)Tupel () ist eine Basis des Nullvektorraums $\{0\}$.

(c) Seien $m, n \in \mathbb{N}$. Für $1 \leq i \leq m$, $1 \leq j \leq n$ sei

$$E_{ij} := \begin{array}{c} \begin{matrix} 1 & \cdots & j-1 & j & j+1 & \cdots & n \end{matrix} \\ \begin{pmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{matrix} 1 \\ \vdots \\ i-1 \\ i \\ i+1 \\ \vdots \\ m \end{matrix} \end{array} \tag{4.2.1}$$

die Matrix mit einer Eins an Position (i, j) und Nullen an allen anderen Positionen. Dann ist $(E_{ij})_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$ offensichtlich eine Basis von $K^{m \times n}$, wobei wir diese mn Elemente in beliebiger Reihenfolge als mn-Tupel auffassen können.

(d) Im K-Vektorraum $K[X]$ sind für jedes $n \in \mathbb{N}$ die **Monome** $1, X, X^2, \ldots, X^n$ linear unabhängig. Sie bilden eine Basis des Untervektorraums der Polynome vom Grad $\leq n$.

Beispiele 4.2.9. (a) Sei $v \in V$. Dann ist (v) genau dann linear unabhängig, wenn $v \neq 0$.

Beweis: Fall $v \neq 0$: Gilt $av = 0$ für $a \in K$, so folgt $a = 0$ nach Lemma 4.1.7. Also ist (v) linear unabhängig.

Fall $v = 0$. Setzt man $a = 1$, so gilt $av = 1v = v = 0$ aber $a = 1 \neq 0$. Also ist (v) linear abhängig.

(b) Seien $u, v \in V$. Dann ist (u, v) genau dann linear abhängig, wenn u ein skalares Vielfaches von v ist oder v ein skalares Vielfaches von u ist, es also ein $a \in K$ mit $u = av$ oder $v = au$ gibt:

Beweis: $\Rightarrow$: Seien (u, v) linear abhängig. Dann existieren $a, b \in K$, nicht beide 0, mit $au + bv = 0$. Im Fall $a \neq 0$ gilt $u + a^{-1}bv = 0$, also $u = -a^{-1}bv = (-a^{-1}b)v = (-\frac{b}{a})v$. Im Fall $b \neq 0$ gilt analog $v = (-b^{-1}a)u$.

$\Leftarrow$: Gelte $u = av$ oder $v = au$ für ein $a \in K$. Dann folgt $1u + (-a)v = 0$ oder $(-a)u + 1v = 0$. Also ist (u, v) linear abhängig.

Beispiel 4.2.10. Sei $a \in \mathbb{R}$. In $V = \mathbb{R}^3$ sind die beiden Vektoren $v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} a \\ 3 \\ 3 \end{pmatrix}$ genau dann linear unabhängig, wenn $a \neq 6$ gilt.

Beispiel 4.2.11. Wir kennen bereits die Standardbasis (e_1, e_2) von $\mathbb{R}^2$. Wir behaupten, dass die beiden Vektoren $v = \begin{pmatrix}1\\2\end{pmatrix}$, $w = \begin{pmatrix}1\\3\end{pmatrix}$ ebenfalls eine Basis des $\mathbb{R}^2$ bilden. Insbesondere zeigt dies, dass ein Vektorraum im Allgemeinen mehrere Basen hat. Daher sprechen wir immer von *einer* Basis und nicht von *der* Basis.

Um zu zeigen, dass v, w eine Basis von $\mathbb{R}^2$ ist, genügt es zu zeigen, dass sich jeder Vektor $x = \begin{pmatrix}x_1\\x_2\end{pmatrix} \in \mathbb{R}^2$ als Linearkombination $x = av + bw$ für eindeutige $a, b \in \mathbb{R}$ schreiben lässt.

Wir interessieren uns also für Paare (a, b) reeller Zahlen mit

$$\begin{array}{rcrcll} 1a & + & 1b & = & x_1 & \text{und}\\ 2a & + & 3b & = & x_2, & \end{array}$$

sind also an der Lösungsmenge des LGS

$$\begin{array}{rcrcl} A & + & B & = & x_1,\\ 2A & + & 3B & = & x_2 \end{array}$$

in den zwei Variablen A und B interessiert. Bringen wir dieses LGS auf Zeilenstufenform, so erhalten wir

$$\begin{array}{rcl} A & = & 3x_1 - x_2,\\ B & = & x_2 - 2x_1. \end{array}$$

Seine Lösungsmenge besteht also genau aus dem Vektor $\begin{pmatrix}3x_1 - x_2\\x_2 - 2x_1\end{pmatrix} \in \mathbb{R}^2$. Dies zeigt, dass (v, w) eine Basis ist.

Hier ist die Testrechnung, dass dies wirklich eine Lösung ist:

$$\begin{aligned}(3x_1 - x_2)v + (x_2 - 2x_1)w &= (3x_1 - x_2)\begin{pmatrix}1\\2\end{pmatrix} + (x_2 - 2x_1)\begin{pmatrix}1\\3\end{pmatrix}\\ &= \begin{pmatrix}3x_1 - x_2\\6x_1 - 2x_2\end{pmatrix} + \begin{pmatrix}x_2 - 2x_1\\3x_2 - 6x_1\end{pmatrix} = \begin{pmatrix}x_1\\x_2\end{pmatrix} = x.\end{aligned}$$

Ende 10. Vorlesung am 26.05.2020

Satz 4.2.12. *Sei V ein K-Vektorraum. Für Vektoren $v_1, \ldots, v_n \in V$ sind die folgenden Aussagen äquivalent:*

(a) $(v_1, \ldots, v_n)$ ist eine Basis von V.

(b) $(v_1, \ldots, v_n)$ ist ein minimales Erzeugendensystem, d. h. es ist ein Erzeugendensystem und hat die folgende Eigenschaft: Entfernt man mindestens einen Vektor aus dem Tupel, so ist das verbleibende Tupel kein Erzeugendensystem mehr.

(c) $(v_1, \ldots, v_n)$ ist ein maximal linear unabhängiges Tupel, d. h. es ist linear unabhängig und hat die folgende Eigenschaft: Ergänzt man das Tupel um mindestens einen Vektor aus V, so erhält man ein linear abhängiges Tupel.

Beweis. Wir zeigen die Implikationen (a) $\Rightarrow$ (b) $\Rightarrow$ (a) $\Rightarrow$ (c) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b): Es ist klar, dass $(v_1, \ldots, v_n)$ ein Erzeugendensystem ist. Ist $(v_1, \ldots, v_n)$ kein minimales Erzeugendensystem, so existiert ein Index i, so dass $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ ein Erzeugendensystem ist. Also kann man v_i als $v_i = \sum_{\substack{j=1\\j\neq i}}^{n} a_j v_j$ darstellen, für geeignete $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in K$. Dies liefert $0 = (-1)v_i + \sum_{\substack{j=1\\j\neq i}}^{n} a_j v_j$. Also ist $(v_1, \ldots, v_n)$ linear abhängig und somit keine Basis (nach Proposition 4.2.7).

(b) $\Rightarrow$ (a): Laut Annahme ist $(v_1, \ldots, v_n)$ ein Erzeugendensystem. Nehmen wir an, dass es keine Basis ist. Dann ist $(v_1, \ldots, v_n)$ linear abhängig (nach Proposition 4.2.7). Also existieren Skalare $a_1, \ldots, a_n \in K$, die nicht alle Null sind, mit

$$a_1 v_1 + \ldots + a_n v_n = 0.$$

Sei i ein Index mit $a_i \neq 0$. Dann gilt

$$v_i = -\frac{1}{a_i} \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j v_j.$$

Also ist bereits $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ ein Erzeugendensystem, d. h. $(v_1, \ldots, v_n)$ ist kein minimales Erzeugendensystem. Dieser Widerspruch zeigt, dass $(v_1, \ldots, v_n)$ eine Basis ist.

(a) $\Rightarrow$ (c): Es reicht zu zeigen, dass für jeden Vektor $w \in V$ das Tupel $(v_1, \ldots, v_n, w)$ linear abhängig ist. Da $(v_1, \ldots, v_n)$ ein Erzeugendensystem ist, gilt $w = a_1 v_1 + \ldots + a_n v_n$ für geeignete $a_1, \ldots, a_n \in K$. Also ist $a_1 v_1 + \ldots + a_n v_n + (-1)w = 0$ eine Linearkombination der Null, in der nicht alle Koeffizienten $= 0$ sind. Also ist $(v_1, \ldots, v_n, w)$ linear abhängig.

(c) $\Rightarrow$ (a): Da $(v_1, \ldots, v_n)$ linear unabhängig ist, bleibt zu zeigen, dass $(v_1, \ldots, v_n)$ ein Erzeugendensystem ist (nach Proposition 4.2.7), dass also $\langle v_1, \ldots, v_n \rangle = V$ gilt.

Sei $v \in V$. Dann ist $(v_1, \ldots, v_n, v)$ linear abhängig, es gibt also Skalare $a_1, \ldots, a_n, b \in K$, die nicht alle Null sind, mit

$$a_1 v_1 + \cdots + a_n v_n + bv = 0.$$

Es muss $b \neq 0$ gelten, denn sonst zeigt die lineare Unabhängigkeit von $(v_1, \ldots, v_n)$, dass alle $a_i = 0$ sind. Indem wir die obige Gleichung mit b^{-1} multiplizieren, sehen wir $v \in \langle v_1, \ldots, v_n \rangle$. Dies zeigt $\langle v_1, \ldots, v_n \rangle = V$ wie gewünscht. □

Definition 4.2.13. Ein Vektorraum V über einem Körper K heißt **endlich erzeugt**, wenn es eine endliche Teilmenge $M \subset V$ mit $V = \langle M \rangle_K$ gibt. Eine äquivalente Bedingung ist, dass es Vektoren $v_1, \ldots, v_n \in V$ gibt, die ein Erzeugendensystem bilden, für die also $V = \langle v_1, \ldots, v_n \rangle$ gilt.

Satz 4.2.14. *Sei V ein endlich erzeugter K-Vektorraum. Dann besitzt V eine Basis.*

Beweis. Sei $(v_1, \ldots, v_n)$ ein Erzeugendensystem von V. Ist es nicht minimal, so können wir einige der v_i weglassen und haben immer noch ein Erzeugendensystem von V. Dies iterieren wir so lange, bis wir ein minimales Erzeugendensystem erhalten (und dieser Prozess terminiert, da wir es mit endlich vielen Vektoren zu tun haben). Dieses ist dann eine Basis von V, gemäß der Äquivalenz der Aussagen (a) und (b) in Satz 4.2.12. □

Bemerkung 4.2.15. Wenn man den Begriff der Basis etwas allgemeiner definiert, kann man zeigen, dass jeder Vektorraum eine Basis besitzt. Dieser Beweis verwendet das Zornsche Lemma und ist nicht konstruktiv.

Lemma 4.2.16. *Sei V ein K-Vektorraum, seien die Vektoren $v_1, \ldots, v_n \in V$ linear unabhängig, und sei $v \in V$. Dann gilt*

$$(v_1, \ldots, v_n, v) \text{ ist linear unabhängig} \iff v \notin \langle v_1, \ldots, v_n \rangle.$$

Beweis. Wir zeigen äquivalent: $(v_1, \ldots, v_n, v)$ linear abhängig $\iff$ $v \in \langle v_1, \ldots, v_n \rangle$.

$\Rightarrow$: Sei $(v_1, \ldots, v_n, v)$ linear abhängig. Dann gibt es Skalare $a_1, \ldots, a_n, b \in K$, die nicht alle 0 sind, mit $a_1v_1 + \ldots + a_nv_n + bv = 0$. Da $v_1, \ldots, v_n$ linear unabhängig sind, muss $b \neq 0$ gelten. Wir folgern

$$v = -\frac{1}{b}(a_1v_1 + \ldots + a_nv_n) \in \langle v_1, \ldots, v_n \rangle.$$

$\Leftarrow$: Gelte $v \in \langle v_1, \ldots, v_n \rangle$. Dann existieren Skalare $a_1, \ldots, a_n \in K$ mit $v = a_1v_1 + \ldots + a_nv_n$. Also ist

$$-v + a_1v_1 + \ldots + a_nv_n = 0$$

eine Linearkombination der Null, in der nicht alle Koeffizienten Null sind. Also ist $(v_1, \ldots, v_n, v)$ linear abhängig. □

Satz 4.2.17 (Basisergänzungssatz)**.** *Sei V ein endlich erzeugter K-Vektorraum, und seien $v_1, \ldots, v_r$ linear unabhängige Vektoren. Dann existieren Vektoren $v_{r+1}, \ldots, v_n \in V$, so dass*

$$(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$$

eine Basis ist.

Beweis. Sei $(w_1, \ldots, w_m)$ ein Erzeugendensystem von V.

Falls möglich, wählen wir ein $i_1 \in \{1, \ldots, m\}$ mit $w_{i_1} \notin \langle v_1, \ldots, v_r \rangle$ und setzen $v_{r+1} := w_{i_1}$; nach Lemma 4.2.16 sind die Vektoren $v_1, \ldots, v_r, v_{r+1}$ linear unabhängig.

Falls es ein solches i_1 gab, iterieren wir dieses Vorgehen: Falls möglich, wählen wir ein $i_2 \in \{1, \ldots, m\}$ mit $w_{i_2} \notin \langle v_1, \ldots, v_r, v_{r+1} \rangle$ und setzen $v_{r+2} := w_{i_2}$. Nach Lemma 4.2.16 sind die Vektoren $v_1, \ldots, v_r, v_{r+1}, v_{r+2}$ linear unabhängig.

Wir iterieren diese Vorgehensweise solange wie möglich. Alle Indizes $i_1, i_2, \ldots$, die wir so produzieren, sind paarweise verschiedene Elemente der endlichen Menge $\{1, \ldots, m\}$. Also terminiert unser Verfahren nach endlich vielen Schritten.

Wir erhalten linear unabhängige Vektoren $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$, für ein geeignetes $n \geq r$, so dass $w_i \in \langle v_1, \ldots, v_n \rangle$ für alle $i \in \{1, \ldots, m\}$ gilt. Daraus folgt $V = \langle w_1, \ldots, w_m \rangle \subset \langle v_1, \ldots, v_n \rangle \subset V$, also $V = \langle v_1, \ldots, v_n \rangle$. Also ist $v_1, \ldots, v_n$ ein Erzeugendensystem und damit eine Basis von V. □

Satz 4.2.18. *Sei V ein K-Vektorraum, sei $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V und sei $(w_1, \ldots, w_n)$ linear unabhängig in V. Dann gilt $n \leq m$.*

In Worten hat jedes Erzeugendensystem mehr Elemente als oder gleich viele Elemente wie jedes linear unabhängige Tupel.

Beweis. Wir nehmen an, dass $n > m$ gilt. Es genügt zu zeigen, dass $(w_1, \ldots, w_n)$ linear abhängig ist. Es gilt, Skalare $c_1, \ldots, c_n \in K$ zu finden, die nicht alle Null sind, so dass

$$0 = c_1w_1 + \ldots + c_nw_n$$

gilt. Da $(v_1, \ldots, v_m)$ ein Erzeugendensystem ist, lässt sich jedes w_j als Linearkombination

$$w_j = \sum_{i=1}^{m} a_{ij}v_i$$

für geeignete $a_{ij} \in K$ darstellen. Wir möchten, dass der Ausdruck

$$\sum_{j=1}^{n} c_jw_j = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{m} a_{ij}v_i \right) = \sum_{j=1,i=1}^{n,m} c_ja_{ij}v_i = \sum_{i=1}^{m} \boxed{\left(\sum_{j=1}^{n} c_ja_{ij} \right)} v_i$$

für geeignete Elemente $c_1, \ldots, c_n \in K$ Null ist. Dies ist sicherlich der Fall, wenn alle eingerahmten Summanden Null sind, wenn also

$$\sum_{j=1}^{n} a_{ij} c_j = 0 \qquad \text{für alle } 1 \leq i \leq m \text{ gilt.}$$

Wir können dies als ein homogenes LGS mit m Gleichungen in den n Unbekannten $c_1, \ldots, c_n$ auffassen. Gesucht ist eine Lösung $(c_1, \ldots, c_n)$ dieses homogenen LGS, in der nicht alle c_j Null sind. Eine solche nicht-triviale Lösung existiert aber nach Satz 3.2.11 wegen $m < n$.

Dies zeigt, dass $(w_1, \ldots, w_n)$ linear abhängig ist. Die Behauptung des Satzes folgt. □

4.3. **Dimension.**

Satz 4.3.1. *Je zwei Basen eines endlich erzeugten K-Vektorraums V haben gleich viele Elemente.*

Beweis. Seien $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ zwei Basen von V. Satz 4.2.18 liefert $n \geq m$ und $m \geq n$, also $n = m$. □

Definition 4.3.2. Sei V ein endlich erzeugter K-Vektorraum. Die **Dimension** von V ist die Anzahl der Elemente in einer Basis von V. Sie wird als

$$\dim V = \dim(V) \in \mathbb{N}$$

notiert und ist eine natürliche Zahl. Hierbei beachte man einerseits, dass V eine Basis hat (Satz 4.2.14), und andererseits, dass jede Basis aus gleich vielen Elementen besteht (Satz 4.3.1), $\dim V$ also nicht von der Wahl der Basis abhängt.

Beispiel 4.3.3. Aus den beiden Beispielen (a) und (c) in 4.2.8 folgern wir

(a) $\dim K^n = n$,
(b) $\dim K^{m \times n} = mn$.

Bemerkung 4.3.4. Man kann allgemeiner für jeden Vektorraum seine Dimension definieren (als Kardinalität einer oder alternativ jeder Basis, vgl. Bemerkung 4.2.15). Man sagt dann, dass ein Vektorraum *endlichdimensional* ist, falls diese Dimension eine natürliche Zahl ist. Es ist dann offensichtlich, dass die endlichdimensionalen Vektorräume in diesem Sinne genau unsere endlich erzeugten Vektorräume sind. Da der Begriff *endlichdimensional* der in der Literatur über Vektorräume üblicherweise verwendete Begriff ist und er mir auch griffiger erscheint, machen wir diese Erkenntnis nun einfach zu einer Definition.

Definition 4.3.5. Ein K-Vektorraum heißt genau dann **endlichdimensional**, wenn er endlich erzeugt ist (vgl. Bemerkung 4.3.4). Sei $n \in \mathbb{N}$. Ein Vektorraum V heißt genau dann **n-dimensional**, wenn er endlichdimensional ist und $\dim(V) = n$ gilt.

Satz 4.3.6. *Sei V ein endlichdimensionaler K-Vektorraum.*

(a) Sei $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V. Dann gelten:
  (i) $m \geq \dim V$.
  (ii) $m = \dim V \Leftrightarrow (v_1, \ldots, v_m)$ Basis.
(b) Sei $(w_1, \ldots, w_n)$ linear unabhängig in V. Dann gelten:
  (i) $n \leq \dim V$.
  (ii) $n = \dim V \Leftrightarrow (w_1, \ldots, w_n)$ Basis.

Beweis. (a): Die erste Behauptung folgt aus Satz 4.2.18. In der zweiten Behauptung ist die Implikation $\Leftarrow$ klar. Wir zeigen die Implikation $\Rightarrow$. Gelte also $m = \dim V$. Indem wir aus dem Erzeugendensystem $(v_1, \dots, v_m)$ geeignete Einträge streichen, erhalten wir ein minimales Erzeugendensystem, also eine Basis (siehe Satz 4.2.12). Da alle Basen $\dim V = m$ Elemente haben (Satz 4.3.1), haben wir gar keinen Eintrag gestrichen, und $(v_1, \dots, v_m)$ ist eine Basis.

(b): Die erste Behauptung folgt aus Satz 4.2.18. In der zweiten Behauptung ist die Implikation $\Leftarrow$ klar. Gelte $n = \dim V$. Nach dem Basisergänzungssatz 4.2.17 können wir $(w_1, \dots, w_n)$ zu einer Basis von V ergänzen. Da alle Basen $\dim V = n$ Elemente haben (Satz 4.3.1), haben wir gar keinen Vektor ergänzt. Somit ist $(w_1, \dots, w_n)$ bereits eine Basis von V. □

Beispiel 4.3.7. (a) Es ist $v_1 = \begin{pmatrix}1\\1\end{pmatrix}, v_2 = \begin{pmatrix}1\\4\end{pmatrix} \in \mathbb{R}^2$ eine Basis von $\mathbb{R}^2$ (denn (v_1, v_2) ist linear unabhängig (Beispiel 4.2.9) und $\dim \mathbb{R}^2 = 2$).

(b) Betrachte die Vektoren $v_1 = \begin{pmatrix}1\\2\\1\end{pmatrix}$, $v_2 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $v_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}$, $v_4 = \begin{pmatrix}1\\1\\1\end{pmatrix} \in \mathbb{R}^3$. Dann gelten:

- „zu wenig Elemente für ein Erzeugendensystem“: (v_1, v_2) ist kein Erzeugendensystem (und erst recht keine Basis): Nach Satz 4.3.6.(a) muss jedes Erzeugendensystem mindestens $\dim \mathbb{R}^3 = 3$ Elemente haben.
- „zu viele Elemente für linear unabhängig“: (v_1, v_2, v_3, v_4) ist linear abhängig (und erst recht keine Basis): Nach Satz 4.3.6.(b) sind maximal $\dim \mathbb{R}^3 = 3$ Vektoren linear unabhängig. Explizit ist $v_2 + v_3 - v_4 = 0$ eine nicht-triviale lineare Abhängigkeitsrelation.
- „richtige Anzahl an Elementen für Basis (= Erzeugendensystem und linear unabhängig)“:
  - (v_2, v_3, v_4) ist keine Basis, da wegen $v_2+v_3-v_4 = 0$ nicht linear unabhängig. (Alternativ sieht man auch leicht, dass es sich nicht um ein Erzeugendensystem handelt.)
  - (v_1, v_2, v_3) ist Basis: Wegen Satz 4.3.6.(b) reicht es zu zeigen, dass (v_1, v_2, v_3) linear unabhängig ist. Dies bedeutet, dass das LGS $a_1v_1 + a_2v_2 + a_3v_3 = 0$ in den Unbestimmten a_1, a_2, a_3 nur die triviale Lösung $\begin{pmatrix}0\\0\\0\end{pmatrix}$ besitzt. Dies stimmt, wie eine einfache Rechnung zeigt.

Satz 4.3.8. *Sei V ein endlichdimensionaler K-Vektorraum, und sei $U \subset V$ ein Untervektorraum. Dann ist U endlich erzeugt (und somit endlichdimensional, vgl. Definition 4.3.5) und es gilt*

$$\dim U \leq \dim V.$$

Ferner gilt Gleichheit der Dimensionen genau dann, wenn der Untervektorraum der ganze Raum ist, in Formeln

$$\dim U = \dim V \quad \Longleftrightarrow \quad U = V.$$

Beweis. Wir nehmen an, dass U nicht endlich erzeugt ist. Dann gilt $U \neq \{0\}$. Es existiert also ein Vektor $v_1 \in U \setminus \{0\}$. Da $\langle v_1 \rangle$ ein endlich erzeugter Unterraum von U ist, ist er echt in U enthalten. Es existiert als ein Vektor $v_2 \in U \setminus \langle v_1 \rangle$. Da $\langle v_1, v_2 \rangle$ ein endlich erzeugter

Unterraum von U ist, ist er echt in U enthalten. Es existiert also ein Vektor $v_3 \in U \setminus \langle v_1, v_2\rangle$, und so weiter.

Per Induktion konstruieren wir so eine Folge $v_1, v_2, v_3, \ldots$ von Vektoren in U mit der Eigenschaft, dass $v_n \notin \langle v_1, \ldots, v_{n-1}\rangle$ für alle $n \geq 1$. Induktiv sehen wir mit Lemma 4.2.16, dass für jedes $n \in \mathbb{N}$ das Tupel $(v_1, \ldots, v_n)$ linear unabhängig ist. Dies kann aber nicht sein, da $\dim V$ nach Satz 4.3.6.(b) die maximale Länge eines linear unabhängigen Tupels in V ist. Dieser Widerspruch zeigt, dass U endlich erzeugt ist. Insbesondere ist $\dim U$ definiert.

Sei $(v_1, \ldots, v_m)$ eine Basis von U. Nach dem Basisergänzungssatz 4.2.17 können wir dieses Tupel zu einer Basis $(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ von V ergänzen. Dies zeigt $\dim U = m \leq n = \dim V$.

Gilt $\dim U = \dim V$, also $m = n$, so haben wir nichts ergänzt, und $(v_1, \ldots, v_m)$ ist bereits eine Basis von V. Es folgt $U = \langle v_1, \ldots, v_m\rangle = V$. Die Implikation $U = V \Rightarrow \dim U = \dim V$ ist trivial. □

Beispiele 4.3.9. Im Video erklärt: Untervektorräume von $\mathbb{R}^2$; Dimension des Lösungsraums eines konkreten homogenen LGS mit zwei Gleichungen in vier Variablen.

Ende 11. Vorlesung am 28.05.2020

Definition 4.3.10. Sei X ein Untervektorraum eines K-Vektorraums V. Ein Untervektorraum $Y \subset V$ heißt genau dann **Komplementärraum** (oder **Komplement**) **zu** X **in** V, wenn

$$X \cap Y = \{0\} \qquad \text{und} \qquad X + Y = V$$

gelten. In diesem Fall schreibt man $V = X \oplus Y$.

Dann ist offensichtlich auch X ein Komplementärraum zu Y und man nennt X und Y **komplementäre** Untervektorräume von V.

Lemma 4.3.11. *Seien X und Y Untervektorräume eines Vektorraums V. Dann sind X und Y genau dann komplementär in V, wenn es für jeden Vektor $v \in V$ eindeutige Vektoren $x \in X$ und $y \in Y$ mit $v = x + y$ gibt.*

Beweis. Nach Definition von $X+Y$ ist die Aussage $V = X+Y$ äquivalent zur Aussage, dass sich jedes $v \in V$ als $v = x + y$ mit $x \in X$ und $y \in Y$ schreiben lässt.

$\Rightarrow$: Sei $v \in V$. Es bleibt die Eindeutigkeit der Darstellung $v = x + y$ mit $x \in X$ und $y \in Y$ zu zeigen. Gelte $v = x' + y'$ für $x' \in X$ und $y' \in Y$. Aus $x + y = v = x' + y'$ folgt $x - x' = y' - y \in X \cap Y = \{0\}$, also $x = x'$ und $y = y'$.

$\Leftarrow$: Es bleibt $X \cap Y = \{0\}$ zu zeigen. Sei $v \in X \cap Y$. Dann sind $v = v + 0 = 0 + v$ zwei Darstellungen von v als Summe eines Elements von X und eines Elements von Y. Aus der Eindeutigkeit folgt $v = 0$. □

Beispiel 4.3.12. Sei $V := \mathbb{R}^2$ und sei $X := \langle e_1\rangle$ die „erste Koordinatenachse“. Dann ist jeder Untervektorraum der Form $Y = \langle y\rangle$ mit $y \in \mathbb{R}^2 \setminus X$ ein Komplementärraum zu X.

Satz 4.3.13. *Sei X ein Untervektorraum eines endlichdimensionalen K-Vektorraums V. Dann existiert ein Komplementärraum Y zu X in V.*

Beweis. Sei $(v_1, \ldots, v_r)$ eine Basis von X. Nach dem Basisergänzugssatz 4.2.17 können wir sie zu einer Basis $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ von V ergänzen. Wir behaupten, dass $Y := \langle v_{r+1}, \ldots, v_n\rangle$ ein Komplementärraum zu X ist.

(a) $\boxed{X \cap Y = \{0\}}$: Sei $v \in X \cap Y$. Wegen $v \in X$ und $v \in Y$ gibt es Darstellungen

$$v = \sum_{i=1}^{r} a_i v_i \qquad \text{für geeignete } a_1, \ldots, a_r \in K \text{ und}$$

$$v = \sum_{j=r+1}^{n} a_j v_j \qquad \text{für geeignete } a_{r+1}, \ldots, a_n \in K.$$

Ziehen wir die beiden Gleichungen voneinander ab, so erhalten wir

$$0 = v - v = \sum_{i=1}^{r} a_i v_i + \sum_{j=r+1}^{n} (-a_j) v_j.$$

Da $(v_1, \ldots, v_n)$ linear unabhängig ist, folgt $a_1 = a_2 = \cdots = a_r = a_{r+1} = \cdots = a_n = 0$, und damit $v = \sum_{i=1}^{r} a_i v_i = 0$. Dies zeigt $X \cap Y \subset \{0\}$. Die Inklusion $\{0\} \subset X \cap Y$ ist klar, da X und Y Untervektorräume sind.

(b) $\boxed{X + Y = V}$: Aus der Definition des Spanns folgt sofort $X + Y = \langle v_1, \ldots, v_r \rangle + \langle v_{r+1}, \ldots, v_n \rangle = \langle v_1, \ldots, v_r, v_{r+1}, \ldots, v_n \rangle = V$.

□

Satz 4.3.14 (Dimensionsformel für zwei Untervektorräume)**.** *Seien X und Y Untervektorräume eines endlichdimensionalen K-Vektorraums V. Dann gilt*

$$\dim(X) + \dim(Y) = \dim(X \cap Y) + \dim(X + Y).$$

Beweis. Beachte, dass nach Satz 4.3.8 alle Untervektorräume $X, Y, X \cap Y, X + Y$ von V endlichdimensional sind und somit ihre Dimensionen wohldefinierte natürliche Zahlen sind.

Sei $(z_1, \ldots, z_r)$ eine Basis von $X \cap Y$. Nach dem Basisergänzungssatz 4.2.17 ist sie sowohl zu einer Basis $(z_1, \ldots, z_r, x_1, \ldots, x_s)$ von X als auch zu einer Basis $(z_1, \ldots, z_r, y_1, \ldots, y_t)$ von Y ergänzbar.

Wir behaupten, dass $(z_1, \ldots, z_r, x_1, \ldots, x_s, y_1, \ldots, y_t)$ eine Basis von $X + Y$ ist.

Sicherlich handelt es sich um ein Erzeugendensystem von $X + Y$. Um die lineare Unabhängigkeit zu beweisen, gelte für Skalare $a_1, \ldots, a_r, b_1, \ldots, b_s, c_1, \ldots c_t \in K$ die Gleichung

$$0 = \underbrace{\sum_{i=1}^{r} a_i z_i}_{=:z \in X \cap Y} + \underbrace{\sum_{j=1}^{s} b_j x_j}_{=:x \in X} + \underbrace{\sum_{k=1}^{t} c_k y_k}_{=:y \in Y}.$$

Aus $z \in X \cap Y \subset Y$ und $y \in Y$ folgt $z + y \in Y$. Somit gilt $x = -z - y = -(z + y) \in Y$ und wegen $x \in X$ folgt $x \in X \cap Y$. Da $(z_1, \ldots, z_r)$ eine Basis von $X \cap Y$ ist, finden wir eindeutige Skalare $d_1, \ldots, d_r$ mit

$$x = \sum_{i=1}^{r} d_i z_i.$$

Nach Definition von x gilt

$$x = \sum_{j=1}^{s} b_j x_j.$$

Da $(z_1, \dots, z_r, x_1, \dots, x_s)$ eine Basis von X ist, folgt $d_1 = \dots = d_r = 0$ und $b_1 = \dots = b_s = 0$, also $x = 0$.

Analog zeigt man $c_1 = \dots = c_t = 0$ und $y = 0$.

Es folgt $0 = z + x + y = z = \sum_{i=1}^{r} a_i z_i$. Da $(z_1, \dots, z_r)$ eine Basis von $X \cap Y$ ist, folgt $a_1 = \dots = a_r = 0$. Insgesamt sind also alle Skalare a_i, b_j, c_k Null. Dies zeigt unsere Behauptung, dass $(z_1, \dots, z_r, x_1, \dots, x_s, y_1, \dots, y_t)$ eine Basis von $X + Y$ ist.

Wir erhalten

$$\dim X + \dim Y = (r+s) + (r+t) = r + (r+s+t) = \dim(X \cap Y) + \dim(X+Y). \qquad \square$$

Korollar 4.3.15. *Sind X und Y komplementäre Untervektorräume eines endlichdimensionalen Vektorraums V, so gilt $\dim(X) + \dim(Y) = \dim(V)$. Sind $(x_1, \dots, x_s)$ eine Basis von X und $(y_1, \dots, y_t)$ eine Basis von Y, so ist $(x_1, \dots, x_s, y_1, \dots, y_t)$ eine Basis von V.*

Beweis. Aus $V = X + Y$ und $X \cap Y = \{0\}$ und $\dim\{0\} = 0$ folgt die erste Behauptung sofort aus Satz 4.3.14. Die zweite Behauptung folgt sofort aus dem Beweis von Satz 4.3.14, oder auch direkt aus der gerade bewiesenen Gleichung $\dim(X) + \dim(Y) = \dim(V)$: Es ist klar, dass $(x_1, \dots, x_s, y_1, \dots, y_t)$ ein Erzeugendensystem von $X + Y = V$ ist, und dann ist es schon eine Basis nach Satz 4.3.6, denn $s + t = \dim(X) + \dim(Y) = \dim(V)$. $\square$

Aufgabe 4.3.16 (Dimension des Produkts)**.** Seien V und W endlichdimensionale Vektorräume. Dann ist auch $V \times W$ endlichdimensional und hat Dimension $\dim(V \times W) = \dim(V) + \dim(W)$ (die Vektorraumstruktur auf $V \times W$ ist in Aufgabe 4.1.19 definiert).

5. Lineare Abbildungen

Der Buchstabe K bezeichne stets einen Körper.

5.1. Erste Eigenschaften linearer Abbildungen.

Definition 5.1.1. Seien V und W K-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt genau dann **K-linear** oder kurz **linear** oder **Homomorphismus von K-Vektorräumen**, wenn

$$\begin{aligned} f(v+v') &= f(v) + f(v') && \text{für alle } v, v' \in V \text{ und} \\ f(av) &= af(v) && \text{für alle } a \in K \text{ und } v \in V \text{ gelten.} \end{aligned}$$

Wir notieren die Menge aller linearen Abbildungen $f\colon V \to W$ als

$$\mathrm{Hom}_K(V, W) := \{f\colon V \to W \mid f \text{ ist } K\text{-linear}\}.$$

Bemerkung 5.1.2 (Alternative Definitionen einer linearen Abbildung)**.** Seien V und W Vektorräume und sei $f\colon V \to W$ eine Abbildung. Dann sind äquivalent:

(a) f ist linear.
(b) $f(a_1v_1 + \dots + a_nv_n) = a_1f(v_1) + \dots + a_nf(v_n)$ für alle $n \in \mathbb{N}$, alle $a_1, \dots, a_n \in K$ und alle $v_1, \dots, v_n \in V$.
(c) $f(av + v') = af(v) + f(v')$ für alle $a \in K$ und alle $v, v' \in V$.

Die Äquivalenz dieser Aussagen ist eine einfache Übung, die wir dem Leser überlassen.

5.1.3. Ist $f\colon V \to W$ eine lineare Abbildung, so gelten

$$\begin{aligned} f(0) &= 0 && \text{und} \\ f(v - v') &= f(v) - f(v') && \text{für alle } v, v' \in V, \end{aligned}$$

denn f ist insbesondere ein Homomorphismus $(V,+) \to (W,+)$ der additiven Gruppen (man kann dies auch mit Hilfe der Skalarmultiplikation und Lemma 4.1.7 sehen: $f(0_V) = f(0_K 0_V) = 0_K f(0_V) = 0_W$ und $f(v-v') = f(v+(-v')) = f(v+(-1)v') = f(v)+(-1)f(v') = f(v) - f(v')$).

Beispiele 5.1.4. (a) Die drei Abbildungen

$$p\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ -y \end{pmatrix}, \qquad s\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}, \qquad d\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -y \\ x \end{pmatrix},$$

sind linear. Fassen wir $\mathbb{R}^2$ als Zeichebene mit dem üblichen rechtwinkligen Koordinatensystem auf, so sind

- p die Punktspiegelung im Ursprung,
- s die Spiegelung an der ersten Winkelhalbierenden und
- d die Drehung um 90° (im Gegenuhrzeigersinn).

(b) Die Abbildung

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x + 42y - 3z,$$

ist linear. Würden wir rechts $x^2 + 42y$ schreiben, wäre die Abbildung nicht linear.

(c) Sind $f, g\colon \mathbb{R} \to \mathbb{R}$ differenzierbare Funktionen und ist $a \in K$ ein Skalar, so gilt $(af+g)' = af' + g'$. Dies bedeutet, dass Ableiten eine lineare Abbildung vom Vektorraum der differenzierbaren reellwertigen Funktionen auf $\mathbb{R}$ in den Vektorraum der reellwertigen Funktionen auf $\mathbb{R}$ ist.

(d) Seien $n \in \mathbb{N}$ und $1 \leq i \leq n$. Die **Projektion** auf die i-te Komponente

$$K^n \to K, \quad (x_1, \ldots, x_n) \mapsto x_i,$$

ist eine lineare Abbildung. Ebenso ist die i-te **Inklusion**

$$K \to K^n, \quad t \mapsto (0, \ldots, 0, t, 0, \ldots, 0),$$

linear (hier steht das t rechts an i-ter Stelle).

(e) Ist $U \subset V$ ein Untervektorraum, so ist die Inklusionsabbildung $U \to V$, $u \mapsto u$, linear.

Definition 5.1.5. Ein **Isomorphismus** von Vektorräumen ist ein Homomorphismus $f\colon V \to W$ von Vektorräumen, für den es einen Homomorphismus $g\colon W \to V$ mit $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$ gibt (für eine alternative Definition siehe Lemma 5.1.7.(b)). Ein solches g ist eindeutig[27], ist ein Isomorphismus, und wird als f^{-1} notiert. Ein Isomorphismus wird als $f\colon V \xrightarrow{\sim} W$ notiert. Zwei Vektorräume V und W heißen genau dann **isomorph**, notiert $V \cong W$, wenn es einen Isomorphismus $f\colon V \xrightarrow{\sim} W$ gibt.

[27] Ist $g'\colon W \to V$ ein weiterer Homomorphismus mit $g' \circ f = \mathrm{id}_V$ und $f \circ g' = \mathrm{id}_W$, so folgt $g = g \circ \mathrm{id}_W = g \circ (f \circ g') = (g \circ f) \circ g' = \mathrm{id}_V \circ g' = g'$.

Beispiel 5.1.6. Die Abbildung „Schreibe die Zeilen einer Matrix hintereinander“

$$K^{m\times n} \xrightarrow{\sim} K^{mn},$$
$$A = (a_{ij}) \mapsto (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{m1}, \dots, a_{mn}),$$

ist ein Isomorphismus von K-Vektorräumen.

Lemma 5.1.7. *(a) Die Verknüpfung linearer Abbildungen ist linear: Sind $g\colon U \to V$ und $f\colon V \to W$ K-lineare Abbildungen von K-Vektorräumen, so ist auch ihre Verknüpfung $f \circ g\colon U \to W$ eine K-lineare Abbildung.*

(b) Ein Homomorphismus $f\colon V \to W$ von Vektorräumen ist genau dann ein Isomorphismus, wenn er bijektiv ist.

Beweis. Wir verwenden die äquivalente Charakterisierung von Linearität einer Abbildung aus Bemerkung 5.1.2.

(a) Seien $a \in K$ und $u, u' \in U$. Aus der Linearität von f und g erhalten wir

$$\begin{aligned}(f \circ g)(au + u') = f(g(au + u')) = f(ag(u) + g(u')) = af(g(u)) + f(g(u')) \\ = a(f \circ g)(u) + (f \circ g)(u').\end{aligned}$$

Also ist $f \circ g$ linear.

(b): Sei f ein Isomorphismus. Dann gibt es nach Definition einen Homomorphismus $g\colon W \to V$ mit $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$. Die erste (bzw. zweite) dieser Gleichungen zeigt, dass f injektiv (bzw. surjektiv) ist. Also ist f bijektiv.

Sei f ein bijektiver Homomorphismus. Sei $g\colon W \to V$ die mengentheoretisch inverse Abbildung. Es genügt zu zeigen, dass g linear ist.

Seien $a \in K$ und $w, w' \in W$. Da f surjektiv ist, gibt es $v, v' \in V$ mit $f(v) = w$ und $f(v') = w'$. Da g invers zu f ist, folgt $g(w) = v$ und $g(w') = v'$. Mit Hilfe der Linearität von f erhalten wir daraus

$$g(aw + w') = g(af(v) + f(v')) = g(f(av + v')) = av + v' = ag(w) + g(w').$$

Dies zeigt die Linearität von g. □

Definition 5.1.8. Ein **Epimorphismus** von Vektorräumen ist ein surjektiver Homomorphismus von Vektorräumen. Ein **Monomorphismus** von Vektorräumen ist ein injektiver Homomorphismus von Vektorräumen.

Definition 5.1.9. Ein Homomorphismus eines Vektorraums in sich selbst wird als **Endomorphismus** bezeichnet. Wir schreiben

$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V)$$

für die Menge aller Endomorphismen eines Vektorraums V. Ein Isomorphismus eines Vektorraums in sich selbst wird als **Automorphismus** bezeichnet. Wir schreiben

$$\mathrm{Aut}_K(V) := \mathrm{GL}(V) := \{f \in \mathrm{End}_K(V) \mid f \text{ ist Isomorphismus}\}$$

für die Menge aller Automorphismen eines Vektorraums V. Diese Menge bildet mit Komposition als Verknüpfung eine Gruppe (diese Verknüpfung ist wegen Lemma 5.1.7.(a) wohldefiniert; der Leser prüfe, dass es sich um eine Gruppe handelt). Sie heißt **allgemeine lineare Gruppe** von V. Die Notation GL kommt von der englischen Bezeichnung *general linear group*.

Beispiel 5.1.10. (a) Die drei Abbildungen (Punktspiegelung im Ursprung, Spiegelung an der ersten Winkelhalbierenden, Drehung) in Beispiel 5.1.4.(a) sind Automorphismen des $\mathbb{R}^2$.

(b) Die *Identitätsabbildung* (oder kurz *Identität*) $\mathrm{id}_V\colon V \to V$, $v \mapsto v$, ist eine lineare Abbildung. Genauer ist sie ein Automorphismus von V.

(c) Die *Nullabbildung* $0\colon V \to W$, $v \mapsto 0$, ist linear.

(d) Sei V ein Vektorraum über K. Skalarmultiplikation $K \times V \to V$ liefert zwei lineare Abbildungen:

(i) Fixieren wir $a \in K$, so ist *Skalarmultiplikation mit a*

$$\begin{aligned} a.\colon V &\to V, \\ v &\mapsto av = a.v, \end{aligned}$$

eine lineare Abbildung. Diese Abbildung ist ein Endomorphismus von V und genau dann ein Automorphismus, wenn $a \neq 0$ oder $V = \{0\}$ gilt.

(ii) Fixieren wir $v \in V$, so ist das *Multiplizieren eines Skalars mit v*

$$\begin{aligned} .v\colon K &\to V, \\ a &\mapsto av = a.v, \end{aligned}$$

eine lineare Abbildung.

Ende 12. Vorlesung am 02.06.2020

Lemma 5.1.11. *Seien V und W K-Vektorräume. Dann ist $\mathrm{Hom}_K(V,W)$ ein K-Vektorraum, wenn wir Addition und Skalarmultiplikation wie folgt definieren: Für $f, g \in \mathrm{Hom}_K(V,W)$ und $a \in K$ definiere $f+g$ und af durch*

$$(f+g)(v) := f(v) + g(v) \quad \textit{und} \quad (af)(v) := af(v) = f(av) \quad \textit{für } v \in V.$$

Beweis. Wir zeigen genauer, dass die Teilmenge $\mathrm{Hom}_K(V,W)$ des Vektorraums W^V aller Abbildungen von V nach W (siehe Beispiel 4.1.4, wobei V schlicht als Menge betrachtet wird) ein Untervektorraum ist[28]. Die im Lemma angegebene Addition und Skalarmultiplikation kommt offensichtlich von der Addition und Skalarmultiplikation auf W^V her. Offensichtlich liegt die Nullabbildung $0\colon V \to W$ in $\mathrm{Hom}_K(V,W)$. Zu zeigen bleibt also, dass die so definierten Abbildungen $f+g$ und af von V nach W linear sind, also in $\mathrm{Hom}_K(V,W)$ liegen.

Seien $v, v' \in V$ und $\lambda \in K$. Dann gelten

$$\begin{aligned} (f+g)(\lambda v + v') &= f(\lambda v + v') + g(\lambda v + v') = \lambda f(v) + f(v') + \lambda g(v) + g(v') \\ &= \lambda(f(v) + g(v)) + f(v') + g(v') = \lambda((f+g)(v)) + (f+g)(v'). \end{aligned}$$

und

$$\begin{aligned} (af)(\lambda v + v') = af(\lambda v + v') = a(\lambda f(v) + f(v')) = (a\lambda)f(v) + af(v') = (\lambda a)f(v) + af(v') \\ = \lambda(af(v)) + af(v') = \lambda((af)(v)) + (af)(v'). \end{aligned}$$

Nach Bemerkung 5.1.2 sind also $f+g$ und af K-linear. □

[28] Alternativ zeigt man zuerst, dass Addition und Skalarmultiplikation wohldefiniert sind, dass also $f+g$ und af K-lineare Abbildungen sind (siehe unten), und überlegt sich dann direkt, dass alle Bedingungen in der Definition eines Vektorraums erfüllt sind.

Lemma 5.1.12. *Seien $h\colon T \to U$ und $g, g'\colon U \to V$ und $f, f'\colon V \to W$ Homomorphismen von K-Vektorräumen und sei $\lambda \in K$. Dann gelten die folgenden Formeln:*

$$\begin{aligned}(f \circ g) \circ h &= f \circ (g \circ h),\\ f \circ \mathrm{id}_V &= f = \mathrm{id}_W \circ f,\\ f \circ (g + g') &= f \circ g + f \circ g',\\ (f + f') \circ g &= fg + f'g,\\ f \circ (\lambda g) &= \lambda(f \circ g) = (\lambda f) \circ g.\end{aligned}$$

Insbesondere ist $\mathrm{End}_K(V)$ mit Addition $+$ und Multiplikation $\circ$ ein Ring (der im Allgemeinen nicht kommutativ ist, wie wir in Beispiel 5.4.12 sehen werden). Das Nullelement ist die Nullabbildung $0\colon V \to V$, das Einselement ist die Identität $\mathrm{id}_V\colon V \to V$.

Beweis. Die ersten beiden Formeln sind trivial, die dritte Formel folgt aus $f((g+g')(u)) = f(g(u)+g'(u)) = f(g(u)) + f(g'(u))$ für $u \in U$, die vierte Formel folgt analog, und die letzte Formel folgt aus $f((\lambda g)(u)) = f(\lambda g(u)) = \lambda f(g(u)) = (\lambda(f \circ g))(u)$ und $(\lambda(f \circ g))(u) = \lambda f(g(u)) = (\lambda f)(g(u))$. □

5.1.13. Die Gruppe der Einheiten des Ringes $\mathrm{End}_K(V)$ ist $\mathrm{Aut}_K(V) = \mathrm{GL}(V) = \mathrm{End}_K(V)^\times$.

5.2. Lineare Abbildungen und Untervektorräume.

Lemma 5.2.1. *Bilder und Urbilder von Untervektorräumen unter linearen Abbildungen sind Untervektorräume: Sei $f\colon V \to W$ eine lineare Abbildung.*

(a) Sei V' ein Untervektorraum von V. Dann ist sein mengentheoretisches Bild

$$f(V') := \{f(v') \mid v' \in V'\}$$

*ein Untervektorraum von W. Insbesondere ist das **Bild** von f,*

$$\mathrm{im}(f) := f(V),$$

ein Untervektorraum von W.

(b) Sei W' ein Untervektorraum von W. Dann ist sein mengentheoretisches Urbild

$$f^{-1}(W') = \{v \in V \mid f(v) \in W'\}$$

*ein Untervektorraum von V. Insbesondere ist der **Kern** von f,*

$$\ker(f) := f^{-1}(\{0\}),$$

ein Untervektorraum von V.

Beweis. Der einfache Beweis ist dem Leser überlassen. □

Proposition 5.2.2. *Sei $f\colon V \to W$ eine lineare Abbildung. Dann gelten:*

(a) f injektiv $\iff$ $\ker(f) = \{0\}$;
(b) f surjektiv $\iff$ $\mathrm{im}(f) = W$.

Beweis. Die Behauptung zur Surjektivität ist trivial, und die Behauptung zur Injektivität folgt aus Lemma 1.3.7, denn f ist insbesondere ein Homomorphismus der abelschen Gruppen $(V,+) \to (W,+)$. Wir geben den Beweis noch einmal:

Ist f injektiv, so gilt sicherlich $\ker(f) = \{0\}$. Gilt $\ker(f) = \{0\}$, so folgt aus $f(v) = f(v')$ für beliebige Vektoren $v, v' \in V$, dass $f(v - v') = f(v) - f(v') = 0$. Dies bedeutet $v - v' \in \ker(f) = \{0\}$, also $v = v'$. □

5.3. **Lineare Abbildungen und Basen.**

Satz 5.3.1 (Lineare Abbildungen durch Werte auf Basis)**.** *Sei $(v_1, \dots, v_n)$ eine Basis eines K-Vektorraums V und sei W ein weiterer K-Vektorraum. Dann ist die Abbildung*

$$\begin{aligned} \mathrm{Hom}_K(V, W) &\xrightarrow{\sim} W^n, \\ f &\mapsto (f(v_1), \dots, f(v_n)), \end{aligned} \tag{5.3.1}$$

bijektiv. Genauer ist sie ein Isomorphismus von Vektorräumen, wenn wir $\mathrm{Hom}_K(V, W)$ wie in Lemma 5.1.11 und W^n wie in Beispiel 4.1.4 als Vektorräume auffassen.

Jede lineare Abbildung ist also eindeutig festgelegt und eindeutig festlegbar durch ihre Werte auf einer Basis.

Beweis. Seien Vektoren $w_1, \dots, w_n \in W$ gegeben. Für die Bijektivität unserer Abbildung müssen wir zeigen, dass es genau eine lineare Abbildung $f\colon V \to W$ mit

$$f(v_i) = w_i \quad \text{für alle } i = 1, \dots, n$$

gibt.

Existenz von f: Da $(v_1, \dots, v_n)$ eine Basis von V ist, existieren für jeden Vektor $v \in V$ eindeutige Skalare $a_1, \dots, a_n$ mit $v = a_1v_1 + \ldots + a_nv_n$. Wir definieren $f(v) := a_1w_1 + \ldots + a_nw_n$ und erhalten so eine wohldefinierte Abbildung $f\colon V \to W$ mit $f(v_i) = w_i$.

Wir behaupten, dass diese Abbildung linear ist. Sei $v' \in V$. Dann gibt es eindeutige $b_1, \dots, b_n$ mit $v' = b_1v_1 + \cdots + b_1v_n$. Sei $c \in K$. Aus $cv + v' = (ca_1 + b_1)v_1 + \cdots + (ca_n + b_n)v_n$ erhalten wir

$$\begin{aligned} f(cv + v') = (ca_1 + b_1)w_1 + \cdots + (ca_n + b_n)w_n &= ca_1w_1 + \cdots + ca_nw_n + b_1w_1 + \cdots + b_nw_n \\ &= c(a_1w_1 + \cdots + a_nw_n) + b_1w_1 + \cdots + b_nw_n = cf(v) + f(v'). \end{aligned}$$

Also ist f linear.

Eindeutigkeit von f: Seien $f, f'\colon V \to W$ zwei lineare Abbildungen mit $f'(v_i) = f(v_i)$ für alle $i = 1, \dots n$. Ein beliebiger Vektor $v \in V$ kann (eindeutig) als Linearkombination $v = a_1v_1 + \ldots + a_nv_n$ geschrieben werden. Auf Grund der Linearität von f' und f gilt

$$f'(v) = a_1f'(v_1) + \ldots + a_nf'(v_n) = a_1f(v_1) + \ldots + a_nf(v_n) = f(v),$$

also $f' = f$.

Dies zeigt, dass unsere Abbildung bijektiv ist. Es ist offensichtlich, dass sie K-linear und damit ein Isomorphismus von K-Vektorräumen ist (siehe Lemma 5.1.7.(b)). □

Proposition 5.3.2. *Sei $f\colon K^n \to V$ eine lineare Abbildung. Setze $v_i := f(e_i)$, wobei $(e_1, \dots, e_n)$ die Standardbasis von K^n ist. Dann gelten:*

(a) f ist injektiv (= ein Monomorphismus) $\iff$ $(v_1, \dots, v_n)$ ist linear unabhängig;
(b) f ist surjektiv (= ein Epimorphismus) $\iff$ $(v_1, \dots, v_n)$ ist ein Erzeugendensystem von V;
(c) f ist bijektiv (= ein Isomorphismus) $\iff$ $(v_1, \dots, v_n)$ ist eine Basis von V.

Beweis. Beachte, dass

$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = f\left(\sum_{i=1}^n x_ie_i\right) = \sum_{i=1}^n x_if(e_i) = x_1v_1 + \cdots + x_nv_n$$

für alle $x_1, \dots, x_n \in K$ gilt.

(a): Lineare Unabhängigkeit von $(v_1, \dots, v_n)$ ist offensichtlich äquivalent zu $\ker(f) = \{0\}$, was nach Proposition 5.2.2 zur Injektivität von f äquivalent ist.

(b): Dies ist offensichtlich.

(c): Folgt sofort aus (a) und (b), da eine Basis dasselbe ist wie ein linear unabhängiges Erzeugendensystem (siehe Proposition 4.2.7). □

Korollar 5.3.3. *Sei $f\colon V \xrightarrow{\sim} W$ ein Isomorphismus von Vektorräumen. Dann sind gegebene Vektoren $v_1, \dots, v_n \in V$ genau dann linear unabhängig bzw. bilden ein Erzeugendensystem von V bzw. eine Basis von V, wenn ihre Bilder $f(v_1), \dots, f(v_n)$ linear unabhängig sind bzw. ein Erzeugendensystem bzw. eine Basis von W bilden.*

Beweis. Sei $g\colon K^n \to V$ die nach Satz 5.3.1 eindeutige lineare Abbildung mit $g(e_i) = v_i$ für alle $i = 1, \dots, n$. Dann ist die Injektivität bzw. Surjektivität bzw. Bijektivität von g äquivalent zur Injektivität bzw. Surjektivität bzw. Bijektivität von $f \circ g$, da f bijektiv ist. Wegen $f(v_i) = f(g(e_i))$ folgt die Behauptung damit aus Proposition 5.3.2, wenn wir diese sowohl auf g als auch auf $f \circ gK^n \to V$ anwenden. □

Korollar 5.3.4. *Sei V ein Vektorraum. Die Wahl einer Basis von V ist gleichbedeutend mit der Wahl eines Isomorphismus $K^n \xrightarrow{\sim} V$. Genauer ist die Abbildung*[29]

$$\{f\colon K^n \xrightarrow{\sim} V \text{ Isomorphismus}\} \xrightarrow{\sim} \{(v_1, \dots, v_n) \text{ Basis von } V\},$$
$$f \mapsto (f(e_1), \dots, f(e_n)),$$

bijektiv.[30]

Beweis. Satz 5.3.1 liefert die Bijektion $\mathrm{Hom}_K(K^n, V) \xrightarrow{\sim} V^n$, $f \mapsto (f(e_1), \dots, f(e_n))$. Nach Proposition 5.3.2 entsprechen unter dieser Bijektion die Isomorphismen genau den Basen. □

Satz 5.3.5 (Klassifikation endlichdimensionaler Vektorräume bis auf Isomorphie)**.** *Zwei endlichdimensionale K-Vektorräume V und W sind genau dann isomorph, wenn $\dim V = \dim W$ gilt.*

Insbesondere ist jeder endlichdimensionale K-Vektorraum der Dimension n zu K^n isomorph.

Beweis. $\Rightarrow$: Sei $g\colon V \xrightarrow{\sim} W$ ein Isomorphismus. Da g dann jede Basis von V auf eine Basis von W abbildet (siehe Korollar 5.3.3), folgt $\dim V = \dim W$.

$\Leftarrow$: Sei $n = \dim V = \dim W$. Seien $(v_1, \dots, v_n)$ eine Basis von V und $(w_1, \dots, w_n)$ eine Basis von W. Nach Korollar 5.3.4 entsprechen diesen Basen Isomorphismen $f\colon K^n \xrightarrow{\sim} V$ und $g\colon K^n \xrightarrow{\sim} W$. Dann ist auch $g \circ f^{-1}\colon V \xrightarrow{\sim} W$ ein Isomorphismus. □

5.3.6. Isomorphie $V \cong W$ definiert eine Äquivalenzrelation auf der Menge aller Vektorräume und auch auf der Menge aller endlichdimensionalen Vektorräume; dies ist klar nach Definition 5.1.5, Lemma 5.1.7.(a) und Korollar 5.3.3.

Satz 5.3.5 und $\dim(K^n) = n$ zeigen, dass $V \mapsto \dim V$ eine wohldefinierte bijektive Abbildung

$$\{K\text{-Vektorräume}\}/\cong \xrightarrow{\sim} \mathbb{N}$$

[29] Wir notieren hier und im Folgenden auch Bijektionen von Mengen mit dem Zeichen $\xrightarrow{\sim}$.

[30] Falls V nicht n-dimensional ist, so sind beide Mengen leer.

definiert. Bis auf Isomorphie werden endlichdimensionale Vektorräume also durch ihre Dimension klassifiziert. Eine vollständige Liste aller endlichdimensionalen K-Vektorräume bis auf Isomorphie ist durch $K^0 = \{0\}, K^1 = K, K^2, K^3, \dots$ gegeben.

Aufgabe 5.3.7. Sei $f\colon V \to W$ eine lineare Abbildung und seien Vektoren $v_1, \dots, v_n \in V$ gegeben. Dann gelten:

(a) $(f(v_1), \dots, f(v_n))$ linear unabhängig $\Rightarrow$ $(v_1, \dots, v_n)$ linear unabhängig.
(b) $(v_1, \dots, v_n)$ linear unabhängig und f injektiv $\Rightarrow$ $(f(v_1), \dots, f(v_n))$ linear unabhängig.
(c) $(v_1, \dots, v_n)$ Erzeugendensystem von V und f surjektiv $\Rightarrow$ $(f(v_1), \dots, f(v_n))$ Erzeugendensystem von W.

Aufgabe 5.3.8. Sei $f\colon V \to W$ eine lineare Abbildungen von K-Vektorräumen, und sei $(v_1, \dots, v_n)$ eine Basis von V. Dann gelten:

(a) f injektiv $\iff$ $(f(v_1), \dots, f(v_n))$ linear unabhängig.
(b) f surjektiv $\iff$ $(f(v_1), \dots, f(v_n))$ Erzeugendensystem von W.
(c) f Isomorphismus $\iff$ $(f(v_1), \dots, f(v_n))$ Basis von W.

Aufgabe 5.3.9. Sei V ein Vektorraum mit einer Basis $(v_1, \dots, v_n)$. Unter der Bijektion (5.3.1) aus Satz 5.3.1 entsprechen die Isomorphismen $V \xrightarrow{\sim} W$ von Vektorräumen genau den Basen von W, in Formeln ist also

$$\{f\colon V \xrightarrow{\sim} W \text{ Isomorphismus}\} \to \{(w_1, \dots, w_n) \text{ Basis von } W\},$$
$$f \mapsto (f(v_1), \dots, f(v_n)),$$

bijektiv.

5.4. Lineare Abbildungen $K^n \to K^m$ und Matrizen.

Definition 5.4.1. Seien $m, n \in \mathbb{N}$ und sei $f\colon K^n \to K^m$ eine lineare Abbildung. Die **darstellende Matrix von** f ist die $(m \times n)$-Matrix $[f] \in K^{m\times n}$ mit Einträgen aus K, deren j-te Spalte das Bild $f(e_j)$ des j-ten Standardbasisvektors e_j ist; in Formeln gilt also

$$[f] = \big(f(e_1)|f(e_2)|\dots|f(e_n)\big),$$

wobei wir der besseren Lesbarkeit halber die Spaltenvektoren $f(e_i) \in K^m$ durch senkrechte Striche voneinander getrennt haben.

Satz 5.4.2 (Lineare Abbildungen $K^n \to K^m$ und Matrizen)**.** *Seien $m, n \in \mathbb{N}$. Dann ist die Abbildung*

$$\operatorname{Hom}_K(K^n, K^m) \xrightarrow{\sim} K^{m\times n}, \tag{5.4.1}$$
$$f \mapsto [f],$$

die eine lineare Abbildung $f\colon K^n \to K^m$ auf ihre darstellende Matrix $[f]$ abbildet, bijektiv. Genauer ist sie ein Isomorphismus von Vektorräumen, wenn wir $\operatorname{Hom}_K(K^n, K^m)$ wie in Lemma 5.1.11 und $K^{m\times n}$ wie in Beispiel 4.1.6 als Vektorräume auffassen.

Erster Beweis. Nach Satz 5.3.1, angewandt auf $V = K^n$ mit der Standardbasis $(e_1, \dots, e_n)$ und $W = K^m$, ist eine lineare Abbildung $f\colon K^n \to K^m$ eindeutig festgelegt und eindeutig festlegbar durch das n-Tupel $(f(e_1), f(e_2), \dots, f(e_n))$ von Spaltenvektoren in K^m. Die darstellende Matrix $[f]$ ist aber nur eine äquivalente Art, dieses n-Tupel darzustellen. Es ist offensichtlich, dass die so erhaltene Bijektion K-linear ist, dass also die Formeln $[f+f'] = [f]+[f']$ und $[\lambda f] = \lambda[f]$ gelten. □

Zweiter, sehr expliziter Beweis. (Dieser Beweis ist nur angegeben, um den Leser hoffentlich davon zu überzeugen, dass die Aussage sehr einfach ist) Jedes Element von K^n lässt sich als Linearkombination der Standardbasisvektoren schreiben:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

Sei $f\colon K^n \to K^m$ eine K-lineare Abbildung. Definiere Elemente $a_{ij} \in K$ durch

$$f(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, f(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix}, \ldots, f(e_n) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Dies definiert eine Matrix $A := (a_{ij}) \in K^{m\times n}$, und dies ist genau die darstellende Matrix $[f]$ von f. Weil f linear ist, gilt

$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = f(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) = x_1 f(e_1) + \cdots + x_n f(e_n)$$

$$= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + \cdots + x_n a_{1n} \\ x_1 a_{21} + \cdots + x_n a_{2n} \\ \vdots \\ x_1 a_{m1} + \cdots + x_n a_{mn} \end{pmatrix}.$$

Dies bedeutet, dass die Matrix A die K-lineare Abbildung f eindeutig festlegt. Dies zeigt die Injektivität der im Satz angegebenen Abbildung.

Ist umgekehrt eine Matrix $A = (a_{ij}) \in K^{m\times n}$ gegeben, so sieht man leicht, dass die durch die obige Formel definierte Abbildung $K^n \to K^m$ K-linear ist und A als darstellende Matrix hat. Dies zeigt die Surjektivität der angegebenen Abbildung.

Da unsere Abbildung offensichtlich K-linear ist, ist sie als bijektive lineare Abbildung ein Isomorphismus. □

5.4.3. Die Umkehrabbildung zu (5.4.1) ist die Abbildung, die einer Matrix $A = (a_{ij}) \in K^{m\times n}$ die Abbildung

$$K^n \to K^m,$$
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + \cdots + x_n a_{1n} \\ \vdots \\ x_1 a_{m1} + \cdots + x_n a_{mn} \end{pmatrix}.$$

zuordnet: Dies folgt aus dem Beweis von Satz 5.3.1; alternativ rechnet man leicht nach, dass diese Abbildung linear ist[31]) und offensichtlich den j-ten Standardbasisvektor auf die j-te Spalte von A abbildet. Nach Einführung der Matrizenmultiplikation werden wir eine kurze Schreibweise (nämlich $x \mapsto Ax$) für diese Abbildung kennenlernen (siehe 5.4.13).

Ende 13. Vorlesung am 04.06.2020

Beispiel 5.4.4. (a) Die drei Abbildungen in Beispiel 5.1.4.(a) haben als darstellende Matrizen

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

(b) Die darstellende Matrix der Identität $\mathrm{id}_{K^n} \colon K^n \to K^n$ ist die sogenannte **Einheitsmatrix** oder genauer $(n \times n)$**-Einheitmatrix**

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{5.4.2}$$

deren Diagonaleinträge Eins und deren restliche Einträge Null sind. Es gilt also $[\mathrm{id}_{K^n}] = I_n$.

(c) Die darstellende Matrix der Nullabbildung $0 \colon K^n \to K^m$ ist die Nullmatrix, in Formeln $[0] = 0$.

Definition 5.4.5 (Matrizenmultiplikation)**.** Seien $m, n, r \in \mathbb{N}$. Seien $A = (a_{ij}) \in K^{m \times n}$ und $B = (b_{jk}) \in K^{n \times r}$ Matrizen. Wir betonen, dass die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt. Das **Matrixprodukt** oder kurz **Produkt** von A und B ist definiert als die $(m \times r)$-Matrix $A \cdot B$ mit

$$(A \cdot B)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + \ldots + a_{in} b_{nk}.$$

Abkürzend schreiben wir meist $AB := A \cdot B$. Formal ist Matrizenmultiplikation eine Abbildng

$$\begin{aligned} K^{m \times n} \times K^{n \times r} &\to K^{m \times r}, \\ (A, B) &\mapsto AB. \end{aligned} \tag{5.4.3}$$

Bemerkung 5.4.6 (Rechenschema)**.** Setzen wir $C = (c_{ik})_{i=1,k=1}^{m,r} := AB$, so ist der Eintrag $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$ das „Produkt der i-ten Zeile von A mit der k-ten Spalte von B“. Genauer multipliziert man für jedes $1 \leq j \leq n$ den j-ten Eintrag der i-ten Zeile von A mit dem j-ten Eintrag der k-ten Spalte von B, und bildet die Summe dieser n Produkte. Dieses

[31] oder folgert dies aus unseren abstrakten Resultaten: Sie ist die Summe, für $j = 1, \ldots, n$, der Verknüpfungen der linearen Abbildungen „Projektion auf die j-te Koordinate“ (Beispiel 5.1.4.(d)) mit „Multiplizieren eines Skalars mit der j-ten Spalte der Matrix A“ (Beispiel 5.1.10.(d)). Diese Summe von Verknüpfungen ist linear nach Lemma 5.1.7.(a) und Lemma 5.1.11.

Rechenschema merke ich mir graphisch wie folgt:

$$\begin{pmatrix} \cdots & b_{1k} & \cdots & \cdots & \cdots \\ \ddots & \vdots & \ddots & \ddots & \ddots \\ \ddots & \vdots & \ddots & \ddots & \ddots \\ \cdots & b_{nk} & \cdots & \cdots & \cdots \end{pmatrix}$$

$$\begin{pmatrix} \vdots & \ddots & \ddots & \vdots \\ a_{i1} & \cdots & \cdots & a_{in} \\ \vdots & \ddots & \ddots & \vdots \end{pmatrix} \cdot \begin{pmatrix} \ddots & \vdots & \ddots & \ddots & \ddots \\ \cdots & c_{ik} & \cdots & \cdots & \cdots \\ \ddots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

Beispiel 5.4.7. Für $A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \\ 3 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ gilt

$$A \cdot B = \begin{pmatrix} 3 & 1 \\ -1 & -1 \\ 4 & 1 \end{pmatrix}.$$

Beachte, dass $B \cdot A$ nicht definiert ist.

Satz 5.4.8 (Darstellende Matrizen: Komposition und Matrizenmultiplikation)**.** *Sind $g\colon K^r \to K^n$ und $f\colon K^n \to K^m$ lineare Abbildungen von K-Vektorräumen, so ist die darstellende Matrix der Verknüpfung $f \circ g\colon K^r \to K^m$ das Matrixprodukt der darstellenden Matrizen von g und f, in Formeln*

$$[f \circ g] = [f] \cdot [g].$$

Beweis. Seien $(e_1, \ldots, e_r)$ bzw. $(e'_1, \ldots, e'_n)$ bzw. $(e''_1, \ldots, e''_m)$ die Standardbasen von K^r bzw. K^n bzw. K^m. Nach Definition der darstellenden Matrizen gilt

$$g(e_k) = \sum_{j=1}^{n} [g]_{jk} e'_j \quad \text{für alle } k = 1, \ldots, r,$$

$$f(e'_j) = \sum_{i=1}^{m} [f]_{ij} e''_i \quad \text{für alle } j = 1, \ldots, n, \text{ und}$$

$$(f \circ g)(e_k) = \sum_{i=1}^{m} [f \circ g]_{ik} e''_i \quad \text{für alle } k = 1, \ldots, r. \tag{5.4.4}$$

Aus den ersten beiden Gleichungen erhalten wir für beliebiges $k \in \{1, \ldots, r\}$

$$\begin{aligned}(f \circ g)(e_k) = f(g(e_k)) = f\Big(\sum_{j=1}^{n} [g]_{jk} e'_j\Big) = \sum_{j=1}^{n} [g]_{jk} f(e'_j) = \sum_{j=1}^{n} [g]_{jk} \Big(\sum_{i=1}^{m} [f]_{ij} e''_i\Big) \\ = \sum_{j=1}^{n} \sum_{i=1}^{m} [f]_{ij} [g]_{jk} e''_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} [f]_{ij} [g]_{jk}\Big) e''_i = \sum_{i=1}^{m} ([f] \cdot [g])_{ik} e''_i.\end{aligned}$$

Der Ausdruck rechts stimmt also mit der rechten Seite von (5.4.4) überein. Per Koeffizientenvergleich (siehe Lemma 4.2.5.(b), dies verwendet die lineare Unabhängigkeit der $(e''_1, \ldots, e''_m)$)

erhalten wir $([f]\cdot[g])_{ik} = [f\circ g]_{ik}$ für alle $i \in \{1,\dots,m\}$ und alle $k \in \{1,\dots,r\}$. Dies bedeutet $[f]\cdot[g] = [f\circ g]$. □

Definition 5.4.9. Eine Matrix mit gleich vielen Zeilen wie Spalten heißt **quadratische** Matrix. Eine Matrix $A \in K^{m\times n}$ ist also genau dann quadratisch, wenn $m = n$ gilt.

Lemma 5.4.10 (Rechenregeln für Matrizen)**.** *Für alle $A, A' \in K^{m\times n}$, $B, B' \in K^{n\times r}$, $C \in K^{r\times s}$ und $\lambda \in K$ gelten:*

$$\begin{aligned}(A\cdot B)\cdot C &= A\cdot(B\cdot C),\\ A\cdot I_n &= A = I_m\cdot A,\\ A\cdot(B+B') &= A\cdot B + A\cdot B',\\ (A+A')\cdot B &= AB + A'B,\\ A\cdot(\lambda B) &= \lambda(A\cdot B) = (\lambda A)\cdot B.\end{aligned}$$

*Insbesondere ist die Menge $K^{n\times n}$ der quadratischen Matrizen ein Ring, der sogenannte **Matrizenring** oder genauer **Ring der** $(n\times n)$**-Matrizen**. Sein Nullelement ist die Nullmatrix und sein Einselement die Einheitsmatrix.*

Erster Beweis. Diese Formeln entsprechen genau den Formeln in Lemma 5.1.12, wenn man sie mit Hilfe der Sätze 5.4.2 und 5.4.8 in Aussagen über lineare Abbildungen übersetzt. Wir zeigen dies exemplarisch für die erste Formel.

Nach Satz 5.4.2 gibt es eindeutige Homomorphismen $h\colon K^s \to K^r$ und $g\colon K^r \to K^n$ und $f\colon K^n \to K^m$ von Vektorräumen mit $[f] = A$ und $[g] = B$ und $[h] = C$. Unter Verwendung von Satz 5.4.8 und der ersten Formel in Lemma 5.1.12 erhalten wir

$$\begin{aligned}(A\cdot B)\cdot C = ([f]\cdot[g])\cdot[h] = [f\circ g]\cdot[h] = [(f\circ g)\circ h] = [f\circ(g\circ h)]\\ = [f]\cdot[g\circ h] = [f]\cdot([g]\cdot[h]) = A\cdot(B\cdot C).\end{aligned}$$

Für die zweite Formel benötigt man die offensichtliche Gleichheit $[\mathrm{id}_{K^n}] = I_n$, und für die anderen Formeln, dass $[\lambda f] = \lambda[f]$ und $[f+f'] = [f]+[f']$ gelten (dies folgt aus Satz 5.4.2, denn die Abbildung $f \mapsto [f]$ ist K-linear. □

Zweiter Beweis durch stupides Nachrechnen. Alle Aussagen sind offensichtliche Indexschlachten, wenn man die Definitionen ausschreibt. Wir zeigen die erste Aussage zur Demonstration:

$$\begin{aligned}((AB)C)_{ij} = \sum_{x=1}^{r}(AB)_{ix}C_{xj} = \sum_{x=1}^{r}\Big(\sum_{y=1}^{n}A_{iy}B_{yx}\Big)C_{xj} = \sum_{x=1}^{r}\sum_{y=1}^{n}A_{iy}B_{yx}C_{xj}\\ = \sum_{y=1}^{n}A_{iy}\Big(\sum_{x=1}^{r}B_{yx}C_{xj}\Big) = \sum_{y=1}^{n}A_{iy}(BC)_{yj} = (A(BC))_{ij}.\end{aligned}$$

Um $A = A\cdot I_n = I_m\cdot A$ zu zeigen, ist es hilfreich, das sogenannte **Kronecker-**δ zu verwenden: Es ist definiert durch

$$\delta_{ij} := \begin{cases}1 & \text{falls } i = j,\\ 0 & \text{sonst.}\end{cases}$$

Dabei sind i und j meist ganze Zahlen (allgemeiner können i und j Elemente einer beliebigen Menge sein). Damit gilt $(I_n)_{ij} = \delta_{ij}$ und man kann nun $A = A\cdot I_n$ direkt nachrechnen. □

Bemerkung 5.4.11. Die Abbildung

$$\begin{aligned} \mathrm{End}_K(K^n) &\xrightarrow{\sim} K^{n\times n}, \\ f &\mapsto [f], \end{aligned}$$

ist ein Ringisomorphismus. Dies folgt aus den Sätzen 5.4.2 und 5.4.8 und aus $[\mathrm{id}_{K^n}] = I_n$.

Beispiel 5.4.12. (a) Der Ring $K^{n\times n}$ ist für $n \geq 2$ nicht kommutativ: Für $n = 2$ und die Matrizen $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ gilt beispielsweise

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

Für $n > 2$ findet man leicht ähnliche Beispiele; zum Beispiel kann man A und B rechts und unten durch Nullen ergänzen.

Nach Bemerkung 5.4.11 zeigt dies auch, dass $\mathrm{End}_K(K^n)$ für $n \geq 2$ nicht kommutativ ist.

(b) Der Ring $K^{1\times 1}$ ist der Körper K.

(c) Für alle $n \geq 2$ gibt es eine Matrix $A \in K^{n\times n}$ mit $A^2 = 0$ aber $A \neq 0$.

Im Fall $n = 2$ nehme man etwa die Matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Für $n > 2$ kann man analog eine Matrix nehmen, die sich nur im rechten oberen Eintrag von der Nullmatrix unterscheidet.

5.4.13 (Matrix-Vektor-Multiplikation)**.** Wir erinnern daran, dass ein Spaltenvektor mit n Einträgen dasselbe ist wie eine $(n \times 1)$-Matrix, wir also $K^n = K^{n\times 1}$ identifizieren können. Damit liefert Matrizenmultiplikation (5.4.3) eine Abbildung

$$\begin{aligned} \cdot\colon K^{m\times n} \times K^n &\to K^m, \\ (A, x) &\mapsto A \cdot x = Ax, \end{aligned}$$

die man als **Matrix-Vektor-Multiplikation** bezeichnet. Fixieren wir $A \in K^{m\times n}$, so erhalten wir die Abbildung *Linksmultiplikation mit A*

$$\begin{aligned} A\cdot\colon K^n &\to K^m, \\ x &\mapsto A \cdot x = Ax. \end{aligned} \tag{5.4.5}$$

Diese Abbildung ist K-linear, denn nach unseren Rechenregeln in Lemma 5.4.10 gelten $A(x+y) = Ax + Ay$ für alle $x, y \in K^n$ und $A(\lambda x) = \lambda Ax$ für alle $\lambda \in K$ und $x \in K^n$. Manchmal wird sie die **zu** A **assoziierte lineare Abbildung** genannt.

Die Abbildung (5.4.5) ist in weniger kompakter Schreibweise explizit durch

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto Ax = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \qquad \ddots \qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \tag{5.4.6}$$

gegeben und uns bereits in 5.4.3 begegnet. Es ist also Ax die angegebene Linearkombination der Spalten von A, und speziell ist Ae_j die j-te Spalte von A, wobei e_j der j-te Standardbasisvektor von K^n ist.

Spätestens nun ist klar, dass die Abbildung $K^{m\times n} \to \mathrm{Hom}_K(K^n, K^m)$, die eine Matrix A auf $A\cdot$ abbildet, invers zu der Bijektion (5.4.1) in Satz 5.4.2 ist. Mit anderen Worten ist die darstellende Matrix von $A\cdot$ gerade A ist, in Formeln $[A\cdot] = A$.

5.4.14. Die Identifikation (5.4.1) von $(m\times n)$-Matrizen und linearen Abbildungen $K^n \to K^m$ sollte dem Leser in Fleisch und Blut übergehen. Gegeben eine Matrix $A \in K^{m\times n}$ notieren wir die zugehörige lineare Abbildung $A\cdot\colon K^n \to K^m$ oft kurz als $A\colon K^n \to K^m$ und fassen somit die Matrix als lineare Abbildung auf (deren darstellende Matrix A ist). Ist umgekehrt $A\colon K^n \to K^m$ eine lineare Abbildung, so sollte es für den Leser vollkommen natürlich sein, A als Matrix aufzufassen.

Beispiel 5.4.15 (Beispiele assoziierter Abbildungen)**.** Nun ist es einfach, Beispiele linearer Abbildungen anzugeben.

(a) Die zur $(1\times n)$-Matrix $A = (1\ 1\ \ldots\ 1)$ assoziierte lineare Abbildung ist die Abbildung „Summiere die Einträge“

$$A\colon K^n \to K,$$
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1 + \ldots + x_n.$$

(b) Für $\varphi \in \mathbb{R}$ liefert die **Drehmatrix** $A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$ die lineare Abbildung

$$A\colon \mathbb{R}^2 \to \mathbb{R}^2,$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos(\varphi) - y\sin(\varphi) \\ x\sin(\varphi) + y\cos(\varphi) \end{pmatrix}.$$

Als Abbildung der Zeichenebene (mit den üblichen, aufeinander senkrecht stehenden Koordinatenachsen) ist dies die „Drehung um den Winkel φ (im Bogenmaß)“, denn sie bildet den Vektor $\begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$ nach den Additionstheoremen auf $\begin{pmatrix} \cos(\alpha+\varphi) \\ \sin(\alpha+\varphi) \end{pmatrix}$ ab. Speziell für $\varphi = \frac{\pi}{2}$ erhalten wir die Drehmatrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, die eine Drehung um 90° beschreibt.

Aufgabe 5.4.16. Unter welchen Bedingungen an $s, t, u \in K$ ist die lineare Abbildung

$$\begin{pmatrix} s & u \\ 0 & t \end{pmatrix} : K^2 \to K^2$$

ein Automorphismus von Vektorräumen?

Ende 14. Vorlesung am 09.06.2020

5.5. Dimensionsformeln für lineare Abbildungen.

Definition 5.5.1. Ist $f\colon V \to W$ eine lineare Abbildung endlichdimensionaler Vektorräume, so heißt

$$\mathrm{rk}(f) := \dim(\mathrm{im}(f)) \in \mathbb{N}$$

der **Rang** von f. Ist $A \in K^{m\times n}$ eine Matrix, so ist der **Rang** $\mathrm{rk}(A)$ von A der Rang der linearen Abbildung $A = A\cdot\colon K^n \to K^m$, in Formeln $\mathrm{rk}(A) := \mathrm{rk}(A\cdot) = \dim(\mathrm{im}(A\cdot))$.

Bemerkung 5.5.2. Das Bild der Abbildung $A\colon K^n \to K^m$ ist der von den Vektoren Ae_j, für $j = 1, \ldots, n$, aufgespannte Untervektorraum von K^m. Diese Vektoren sind gerade die Spalten von A (vgl. (5.4.6)). Deswegen nennt man $\mathrm{rk}(A)$ auch den **Spaltenrang** von A.

Satz 5.5.3 (Dimensionssatz für lineare Abbildungen)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann gilt die* ***Dimensionsformel (für lineare Abbildungen)***

$$\dim V = \dim(\ker(f)) + \dim(\mathrm{im}(f)). \tag{5.5.1}$$

5.5.4 (Rangsatz)**.** Die Dimensionsformel (5.5.1) lautet umgeschrieben

$$\dim V = \dim(\ker(f)) + \mathrm{rk}(f) \tag{5.5.2}$$

und wird auch als **Rangformel** bezeichnet.

5.5.5. Man beachte, dass $\ker(f)$ und $\mathrm{im}(f)$ als Untervektorräume endlichdimensionaler Vektorräume automatisch endlichdimensional sind (Satz 4.3.8), so dass also $\dim(\ker(f))$ und $\dim(\mathrm{im}(f))$ wohldefinierte natürliche Zahlen sind.

Beweis. Sei $U \subset V$ ein Komplementärraum zu $\ker(f)$ (er existiert nach Satz 4.3.13). Es gilt

$$\dim V = \dim(\ker(f)) + \dim U \tag{5.5.3}$$

nach Korollar 4.3.15.

Sei $g\colon U \to W$ die Einschränkung von f auf U. Sie ist offensichtlich linear (als Komposition der Inklusionsabbildung $U \to V$ mit f) und es gilt $\mathrm{im}(g) = g(U) = f(U) \subset f(V) = \mathrm{im}(f)$.

Wir behaupten, dass auch die Inklusion $\mathrm{im}(f) \subset \mathrm{im}(g)$ gilt. Jedes $v \in V$ lässt sich (eindeutig) als $v = u + w$ mit $u \in U$ und $w \in \ker(f)$ darstellen (siehe Lemma 4.3.11). Es gilt $f(v) = f(u) + f(w) = f(u) + 0 = f(u) = g(u)$. Daraus folgt die behauptete Inklusion und insgesamt $\mathrm{im}(f) = \mathrm{im}(g)$.

Wir können g also als surjektive lineare Abbildung $g\colon U \to \mathrm{im}(f)$ auffassen. Diese Abbildung ist aber auch injektiv: Für jedes $u \in U$ mit $0 = g(u) = f(u)$ gilt $u \in \ker(f) \cap U = \{0\}$, also $u = 0$. Somit gilt $\ker(g) = \{0\}$ und g ist injektiv (Proposition 5.2.2).

Also ist $g\colon U \xrightarrow{\sim} \mathrm{im}(f)$ bijektiv und damit ein Isomorphismus. Satz 5.3.5 liefert $\dim(U) = \dim(\mathrm{im}(f))$. Kombiniert mit (5.5.3) zeigt das die Dimensionsformel. □

5.5.6. Das folgende Korollar zeigt, dass man jeden Homomorphismus zwischen endlichdimensionalen Vektorräumen durch geeignete Wahl von Basen auf eine sehr einfache Gestalt bringen kann.

Korollar 5.5.7 (im Wesentlichen Smith-Normalform)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen; sei $r := \mathrm{rk}(f)$ der Rang von f. Dann gibt es eine Basis $(v_1, \ldots, v_n)$ von V und eine Basis $(w_1, \ldots, w_m)$ von W, so dass gelten:*

$$f(v_i) = w_i \qquad \text{für alle } 1 \le i \le r,$$
$$f(v_i) = 0 \qquad \text{für alle } r+1 \le i \le n.$$

Beweis. Wir verwenden den Beweis von Satz 5.5.3. Es gilt $r = \mathrm{rk}(f) = \dim(\mathrm{im}(f)) = \dim(U)$. Sei $(b_1, \ldots, b_r)$ eine Basis von U. Da die Einschränkung $g\colon U \xrightarrow{\sim} \mathrm{im}(f)$ von f ein Isomorphismus ist, ist $(g(b_1), \ldots, g(b_r))$ eine Basis von $\mathrm{im}(f)$ (Korollar 5.3.3). Nach dem Basisergänzungssatz 4.2.17 können wir die linear unabhängigen Vektoren $w_1 := g(b_1) = f(b_1), \ldots, w_r := g(b_r) = f(b_r)$ zu einer Basis $(w_1, \ldots, w_r, w_{r+1}, \ldots, w_m)$ von W ergänzen.

Sei $(b_{r+1}, \ldots, b_n)$ eine Basis von $\ker(f)$. Weil U und $\ker(f)$ komplementäre Untervektorräume von V sind, ist $(b_1, \ldots, b_r, b_{r+1}, \ldots, b_n)$ eine Basis von V (Korollar 4.3.15). Nach Konstruktion hat f auf den Elementen dieser Basis die gewünschten Werte. □

Korollar 5.5.8. *Sei $f\colon V \to W$ eine lineare Abbildung endlichdimensionaler Vektorräume. Dann gelten die folgenden Implikationen:*

(a) f ist injektiv (= ein Monomorphismus) $\iff \dim(\ker(f)) = 0 \iff \operatorname{rk}(f) = \dim V \implies \dim V \leq \dim W$.
(b) f ist surjektiv (= ein Epimorphismus) $\iff \dim W = \operatorname{rk}(f) \implies \dim V \geq \dim W$.
(c) f ist bijektiv (= ein Isomorphismus) $\iff \dim V = \operatorname{rk}(f) = \dim W$.

Beweis. (a): Injektivität von f ist äquivalent zu $\ker(f) = \{0\}$ (siehe Proposition 5.2.2), was offensichtlich zu $\dim(\ker(f)) = 0$ äquivalent ist. Nach der Rangformel (5.5.2) ist dies zur Bedingung $\dim V = \operatorname{rk}(f)$ äquivalent. Diese Gleichheit impliziert $\dim V \leq \dim W$, da stets $\operatorname{rk}(f) = \dim(\operatorname{im}(f)) \leq \dim(W)$ gilt (siehe Satz 4.3.8).

(b): Surjektivität von f ist äquivalent zu $\operatorname{im}(f) = W$ (siehe Proposition 5.2.2), was nach Satz 4.3.8 äquivalent zu $\operatorname{rk}(f) = \dim(\operatorname{im}(f)) = \dim W$ ist. Diese Gleichheit impliziert $\dim V \geq \dim W$, da nach der Rangformel (5.5.2) stets $\dim V \geq \operatorname{rk}(f)$ gilt.

(c): Dies folgt sofort aus (a) und (b). □

Korollar 5.5.9. *Sei $f\colon V \to W$ eine lineare Abbildung endlichdimensionaler Vektorräume derselben Dimension $n := \dim V = \dim W$. Dann sind äquivalent:*

(a) f ist injektiv (= ein Monomorphismus);
(b) f ist surjektiv (= ein Epimorphismus);
(c) f ist bijektiv (= ein Isomorphismus);
(d) $\operatorname{rk}(f) = n$.

Insbesondere ist ein Endomorphismus eines endlichdimensionalen Vektorraums genau dann injektiv, wenn er surjektiv ist, was genau dann der Fall ist, wenn er bijektiv ist.

Beweis. Laut Annahme gilt $n = \dim V = \dim W$. Nach Korollar 5.5.8 sind sowohl Injektivität als auch Surjektivität von f äquivalent zu der Aussage $\operatorname{rk}(f) = n$. □

Korollar 5.5.10. *Für $A \in K^{m\times n}$ gilt* $\operatorname{rk}(A) \in \{0, 1, \ldots, \min(m, n)\}$.

Beweis. Da $\operatorname{rk}(A)$ die Dimension des von den Spalten von A aufgespannten Untervektorraums des K^m ist, ist $\operatorname{rk}(A) \leq m$ klar (siehe Satz 4.3.8). Nach der Dimensionsformel (siehe Satz 5.5.3) für die lineare Abbildung $A\colon K^n \to K^m$ gilt $n = \dim K^n = \dim(\ker(A)) + \dim(\operatorname{im}(A)) \geq \dim(\operatorname{im}(A)) = \operatorname{rk}(A)$. □

Aufgabe 5.5.11. Folgere die Dimensionsformel für zwei Untervektorräume 4.3.14 aus dem Dimensionssatz für lineare Abbildungen 5.5.3.

Hinweis: Betrachte die (lineare) Abbildung $X \times Y \to V$, $(x, y) \mapsto x + y$, deren Kern isomorph zu $X \cap Y$ ist.

5.6. **Invertierbare Matrizen.**

Definition 5.6.1. Eine quadratische Matrix $A \in K^{n\times n}$ heißt genau dann **invertierbar** (oder **regulär**), wenn eine quadratische Matrix $B \in K^{n\times n}$ mit

$$AB = BA = I_n$$

existiert. Mit anderen Worten ist eine Matrix genau dann invertierbar, wenn sie eine Einheit im Matrizenring $K^{n\times n}$ ist. Da die Menge aller Einheiten in einem Ring eine Gruppe bildet (siehe 2.2.2), ist

$$\mathrm{GL}_n(K) := (K^{n\times n})^\times = \{A \in K^{n\times n} \mid A \text{ ist invertierbar}\}$$

mit Matrizenmultiplikation eine Gruppe. Sie wird die **allgemeine lineare Gruppe** der $(n \times n)$-Matrizen genannt. Die Notation GL kommt von der englischen Bezeichnung *general linear group*.

5.6.2. Unter unserer Identifikation $\mathrm{End}_K(K^n) \xrightarrow{\sim} K^{n\times n}$ (Bemerkung 5.4.11) von Endomorphismen $K^n \to K^n$ mit $(n \times n)$-Matrizen gilt $\mathrm{GL}(K^n) \xrightarrow{\sim} \mathrm{GL}_n(K)$ (siehe Definition 5.1.9).

5.6.3. Die Gruppe $\mathrm{GL}_1(K) = K^\times = K \setminus \{0\}$ ist kommutativ. Für $n \geq 2$ ist die Gruppe $\mathrm{GL}_n(K)$ nicht kommutativ. In $\mathrm{GL}_2(K)$ gilt beispielsweise

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Dieses Beispiel verallgemeinert sich leicht für $n \geq 2$: Man nehme die Einheitsmatrix und ändere die Null an Position $(1, 2)$ bzw. an Position $(2, 1)$ in eine Eins.

5.6.4. Wie in Gruppen üblich, wird das Inverse eines Elements $A \in \mathrm{GL}_n(K)$ mit A^{-1} bezeichnet. Das neutrale Element von $\mathrm{GL}_n(K)$ ist die Einheitsmatrix I_n. Nach Lemma 1.1.4 gilt $(A^{-1})^{-1} = A$ und $I_n^{-1} = I_n$; sind weiter $A, B \in K^{n\times n}$ invertierbar, so ist auch AB invertierbar mit $(AB)^{-1} = B^{-1}A^{-1}$.

Satz 5.6.5. *Für eine quadratische $(n \times n)$-Matrix $A \in K^{n\times n}$ sind die folgenden Aussagen äquivalent:*

(a) A ist invertierbar, d. h. $A \in \mathrm{GL}_n(K)$.
(b) $A\colon K^n \to K^n$ ist ein Isomorphismus (äquivalent: die Spalten von A bilden eine Basis von K^n).
(c) $A\colon K^n \to K^n$ ist ein Monomorphismus (äquivalent: die Spalten von A sind linear unabhängig).
(d) $A\colon K^n \to K^n$ ist ein Epimorphismus (äquivalent: die Spalten von A sind ein Erzeugendensystem von K^n).
(e) $\mathrm{rk}(A) = n$.
(f) (Existenz eines Rechtsinversen) Es gibt ein $B \in K^{n\times n}$ mit $AB = I_n$.
(g) (Existenz eines Linksinversen) Es gibt ein $B \in K^{n\times n}$ mit $BA = I_n$.

Ist B wie in (f) oder (g), so gilt $B = A^{-1}$.

Beweis. Die Aussagen (a) und (b) sind offensichtlich äquivalent.

Für einen Endomorphismus eines endlichdimensionalen Vektorraums sind die vier Bedingungen surjektiv, injektiv, bijektiv, voller Rang äquivalent (siehe Korollar 5.5.9). Dies zeigt die Äquivalenz der vier Aussagen (b), (c), (d) und (e), und die Äquivalenz zu den drei in Klammern angegebenen Aussagen folgt aus Proposition 5.3.2.

(a) $\Rightarrow$ (f) und (a) $\Rightarrow$ (g): Trivial.

(f) $\Rightarrow$ (d): Fassen wir A und B als Endomorphismen von K^n auf, so folgt aus $AB = I_n = \mathrm{id}_{K^n}$, dass A surjektiv ist. Da dann nach dem bereits Bewiesenen A invertierbar ist, können wir $AB = I_n$ von links mit A^{-1} multiplizieren und erhalten $B = A^{-1}$.

(f) $\Rightarrow$ (c): Fassen wir A und B als Endomorphismen von K^n auf, so folgt aus $BA = I_n = \mathrm{id}_{K^n}$, dass A injektiv ist. Da dann nach dem bereits Bewiesenen A invertierbar ist, können wir $BA = I_n$ von rechts mit A^{-1} multiplizieren und erhalten $B = A^{-1}$. □

Definition 5.6.6. Eine quadratische Matrix $N \in K^{n\times n}$ heißt **nilpotent** (also „nichtskönnend"), falls es ein $m \in \mathbb{N}$ mit $N^m = 0$ gibt.

Lemma 5.6.7. *Sei $N \in \mathrm{K}^{n\times n}$ eine nilpotente Matrix. Dann ist die Matrix $I_n - N$ invertierbar (und ebenso ist $I_n + N$ invertierbar).*

Beweis. Gelte $N^m = 0$. Wir behaupten, dass $I_n + N + N^2 + \cdots + N^{m-2} + N^{m-1}$ invers zu $I_n - N$ ist. In der Tat, es gilt

$$\begin{aligned}(I_n - N)(I_n + N + N^2 + \cdots + N^{m-2} + N^{m-1}) &= I_n + N + N^2 + \cdots + N^{m-2} + N^{m-1} \\ &\quad - N - N^2 - \cdots - N^{m-2} - N^{m-1} - N^m \\ &= I_n - 0 = I_n.\end{aligned}$$

Also ist unsere Matrix rechtsinvers zu $I_n - N$ und Satz 5.6.5 zeigt, dass $I_n - N$ invertierbar ist (man kann auch analog nachrechnen, dass unsere Matrix auch linksinvers ist).

Da mit N auch $-N$ nilpotent ist, ist dann auch $I_n + N = I_n - (-N)$ invertierbar. □

Beispiel 5.6.8. Die Matrix $\begin{pmatrix} 1 & 17 \\ 0 & 1 \end{pmatrix}$ ist invertierbar als Summe der Einheitsmatrix mit der nilpotenten Matrix $\begin{pmatrix} 0 & 17 \\ 0 & 0 \end{pmatrix}$. Ihr Inverses ist $\begin{pmatrix} 1 & -17 \\ 0 & 1 \end{pmatrix}$.

Aufgabe 5.6.9. Ist $A \in K^{n\times n}$ nilpotent, so gilt bereits $A^n = 0$.

Hinweis: Betrachte die aufsteigende Folge der Untervektorräume $\{0\} = \ker(A^0) \subset \ker(A) \subset \ker(A^2) \subset \ldots$ des K^n und zeige: Gilt an einer Stelle Gleichheit $\ker(A^s) = \ker(A^{s+1})$, so auch an jeder nachfolgenden Stelle.

Lemma 5.6.10 (Rang einer Verknüpfung)**.** *Seien $g\colon U \to V$ und $f\colon V \to W$ Homomorphismen zwischen endlichdimensionalen Vektorräumen. Dann gilt*

$$\mathrm{rk}(f \circ g) \leq \min(\mathrm{rk}(f), \mathrm{rk}(g)). \tag{5.6.1}$$

Insbesondere gelten:

(a) Ist f ein Isomorphismus, so gilt $\mathrm{rk}(f \circ g) = \mathrm{rk}(g)$.
(b) Ist g ein Isomorphismus, so gilt $\mathrm{rk}(f \circ g) = \mathrm{rk}(f)$.
(c) Gegeben eine Matrix $A \in K^{m\times n}$ und invertierbare Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$, so gilt

$$\mathrm{rk}(A) = \mathrm{rk}(SAT).$$

Beweis. Wegen $g(U) \subset V$ folgt $\mathrm{im}(f \circ g) = f(g(U)) \subset f(V) = \mathrm{im}(f)$ und somit $\mathrm{rk}(f \circ g) \leq \mathrm{rk}(f)$ nach Satz 4.3.8. Wir wenden den Dimensionssatz 5.5.3 auf die durch Einschränkung erhaltene lineare Abbildung $f|_{g(U)}\colon g(U) \to W$ an und erhalten

$$\begin{aligned}\mathrm{rk}(g) = \dim(g(U)) &\overset{(5.5.1)}{=} \dim(\ker(f|_{g(U)})) + \dim(\mathrm{im}(f|_{g(U)})) \\ &\geq \dim(\mathrm{im}(f|_{g(U)})) = \dim(f(g(U))) = \mathrm{rk}(f \circ g).\end{aligned}$$

Insgesamt zeigt dies die behauptete Abschätzung (5.6.1).

(a) Ist f ein Isomorphismus, so folgt die behauptete Gleichheit aus

$$\operatorname{rk}(f \circ g) \overset{(5.6.1)}{\leq} \operatorname{rk}(g) = \operatorname{rk}(f^{-1} \circ f \circ g) \overset{(5.6.1)}{\leq} \operatorname{rk}(f \circ g).$$

(b) Ist g ein Isomorphismus, so folgt die behauptete Gleichheit aus

$$\operatorname{rk}(f \circ g) \overset{(5.6.1)}{\leq} \operatorname{rk}(f) = \operatorname{rk}(f \circ g \circ g^{-1}) \overset{(5.6.1)}{\leq} \operatorname{rk}(f \circ g).$$

(c) Damit folgt sofort $\operatorname{rk}(A) \overset{(a)}{=} \operatorname{rk}(SA) \overset{(b)}{=} \operatorname{rk}(SAT)$. □

Aufgabe 5.6.11. Für lineare Abbildungen $f, g\colon V \to W$ gilt $\operatorname{rk}(f+g) \leq \operatorname{rk}(f) + \operatorname{rk}(g)$.

Hinweis: Dimensionsformel für zwei Unterverktorräume 4.3.14 oder Lemma 5.6.10 samt der folgenden Beobachtung: Die lineare Abbildung $f+g\colon V \to W$ kann als Verknüpfung $V \xrightarrow{\Delta} V \times V \xrightarrow{f \times g} W \times W \xrightarrow{+} V$ linearer Abbildungen geschrieben werden, wobei $\Delta(v) := (v, v)$ und $(f \times g)(v, w) := (f(v), g(w))$.

Ende 15. Vorlesung am 16.06.2020

5.7. Lineare Abbildungen und lineare Gleichungssysteme.

5.7.1. Sei ein LGS der Form (3.1.1) gegeben. Sei $A = (a_{ij})_{i=1,j=1}^{m,n}$ seine Koeffizientenmatrix und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$ die „rechte Seite“ des LGS. Wir erkennen aus (5.4.6): Ein Spaltenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ ist genau dann eine Lösung des LGS, wenn $Ax = b$ gilt.

Man schreibt deshalb kurz $AX = b$ für das LGS (3.1.1) (oder auch $Ax = b$, wobei man sich $x \in K^n$ variabel denkt).

5.7.2. Ein LGS $AX = b$ ist also dasselbe wie eine lineare Abbildung $A\colon K^n \to K^m$ zusammen mit einem Vektor $b \in K^m$. Seine Lösungsmenge ist das Urbild von b unter dieser linearen Abbildung.

Genau dann ist $AX = b$ für alle $b \in K^m$ lösbar, wenn $A\colon K^n \to K^m$ surjektiv ist, wenn also $\operatorname{rk}(A) = m$ gilt (siehe Korollar 5.5.8.(b)).

Die Lösungsmenge eines homogenes LGS $AX = 0$ ist genau der Kern der linearen Abbildung $A\colon K^n \to K^m$ und somit ein Untervektorraum des K^n (siehe Satz 3.3.2) der Dimension

$$\dim(\ker(A)) = n - \operatorname{rk}(A)$$

(siehe Satz 5.5.3 bzw. (5.5.2)).

- Offensichtlich gilt $\operatorname{rk}(A) \leq m$. Im Fall $n > m$ („mehr Variablen als Gleichungen“) folgt also $\dim(\ker(A)) = n - \operatorname{rk}(A) \geq n - m > 0$ und somit hat das homogene LGS $AX = 0$ mindestens eine nicht-triviale Lösung. Diese Aussage ist genau der Inhalt von Satz 3.2.11.[32]
- Ist A in Zeilenstufenform, so ist $\operatorname{rk}(A)$ offensichtlich die Anzahl der Pivotelemente von A (vergleiche der im homogenen Fall stets eintretende zweite Fall in Bemerkung 3.2.8, wo man die Werte von $n - \#\{\text{Pivotelemente}\} = n - \operatorname{rk}(A)$ Variablen frei wählen kann).

[32] Das hier gegebene Argument ist aber kein unabhängiger Beweis, denn er verwendet Resultate, die wir aus Satz 3.2.11 gefolgert haben.

5.7.3. Die mittlerweile eingeführten Begriffe (Vektorraum, Matrix-Vektor-Multiplikation) und die damit verbundene effiziente Sprech- und Schreibweise gestatten uns nun, Aufgabe 3.3.5 kompakt zu formulieren und zu beweisen.

Satz 5.7.4. *Sei $AX = b$ ein LGS mit $A \in K^{m\times n}$ und $b \in K^m$. Hat das LGS $AX = b$ eine Lösung $y \in K^n$ (eine solche muss aber nicht existieren, falls $b \neq 0$), so ist seine Lösungsmenge $\{x \in K^n \mid Ax = b\}$ durch $y + \ker(A)$ gegeben. Sie entsteht also durch Addition von y aus der Lösungsmenge des homogenen LGS $AX = 0$.*

Beweis. Sei $S := \{x \in K^n \mid Ax = b\}$ die Lösungsmenge und sei $y \in S$ eine Lösung. Zu zeigen ist $S = y + \ker(A)$.

$\subset$: Sei $s \in S$. Wegen $A(s - y) = As - Ay = b - b = 0$ gilt $s - y \in \ker(A)$ und damit $s = y + (s - y) \in y + \ker(A)$.

$\supset$: Sei $x \in y + \ker(A)$, d. h. es gibt ein $l \in \ker(A)$ mit $x = y + l$. Dann gilt $Ax = Ay + Al = b + 0 = b$, also $x \in S$. □

Aufgabe 5.7.5. Sei $f\colon V \to W$ ein Homomorphismus von Vektorräumen. Zeigen Sie für jeden Vektor $w \in W$: Für jedes $v \in f^{-1}(w)$ gilt die Gleichheit $f^{-1}(w) = v + \ker(f)$.

Folgern Sie daraus Satz 5.7.4.

Aufgabe 5.7.6. Sei ein LGS $AX = b$ mit $A \in K^{m\times n}$ und $b \in K^m$ gegeben. Überlegen Sie sich, dass das Gauß-Verfahren (oder eine beliebige andere Abfolge elementarer Zeilenoperationen, die das LGS auf ZSF bringt) einen Isomorphismus von K^{n-p} in den Vektorraum $\ker(A)$ der Lösungen des LGS liefert, wobei p die Menge der Pivot-Elemente in der Zeilenstufenform ist.

5.8. **Algorithmus zur Berechnung der inversen Matrix.**

Algorithmus 5.8.1 (zur Berechnung der inversen Matrix (modifiziertes Gauß-Verfahren))**.** Sei $A \in \mathrm{GL}_n(K)$ gegeben (den Fall, dass $A \in K^{n\times n}$ nicht als invertierbar vorausgesetzt ist, besprechen wir in Algorithmus 5.8.2). Wir möchten die inverse Matrix A^{-1} ausrechnen. Gesucht ist also die eindeutige Matrix B mit $AB = I_n$. Bezeichnen wir die Spalten von B mit $s_1, \ldots, s_n$ (d. h. $s_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}$), so ist die Gleichung $AB = I_n$ in $K^{n\times n}$ äquivalent zu den n Gleichungen

$$As_j = e_j \qquad \text{in } K^n, \text{ für } j = 1, \ldots, n,$$

wobei e_j der j-te Standardbasisvektor ist.

Für jedes j können wir $As_j = e_j$ als LGS (mit n Gleichungen) in den n Einträgen von s_j als Variablen auffassen, und wir können dieses Gleichungssystem bzw. die zugehörige erweiterte Koeffizientenmatrix $(A \mid e_j)$ mit dem Gauß-Verfahren auf Zeilenstufenform bringen und dann die Lösungsmenge ablesen. Da A als invertierbar angenommen ist, wissen wir bereits abstrakt, dass die Lösungsmenge aus genau einem Element besteht, nämlich aus dem Vektor $s_j = A^{-1}e_j$. Also muss die Zeilenstufenform die Gestalt $(I_n \mid s_j)$ haben (siehe Bemerkung 3.2.8 zum Ablesen der Lösung bei ZSF).

Wir können die erweiterten Koeffizientenmatrizen $(A \mid e_j)$, für $j = 1, \ldots, n$, wie folgt zusammenfassen und sie dann simultan auf Zeilenstufenform bringen: Schreibe die Matrix A

links eines vertikalen Strichs und rechts davon die Einheitsmatrix I_n:

$$(A \mid I_n) \tag{5.8.1}$$

Formal ist das eine $(n \times 2n)$-Matrix. Wir bringen diese Matrix mit dem Gauß-Verfahren (oder anderen geschickten elementaren Zeilenumformungen) auf Zeilenstufenform. Diese muss nach der obigen Erklärung die Gestalt $(I_n \mid s_1\, s_2\, \ldots\, s_n)$ haben. Da die s_j die Spalten von B sind, gilt

$$(I_n \mid s_1\, s_2\, \ldots\, s_n) = (I_n \mid B). \tag{5.8.2}$$

Mit anderen Worten haben wir das Inverse $B = A^{-1}$ berechnet.

Algorithmus 5.8.2 (um herauszufinden, ob eine Matrix invertierbar ist, und um gegebenenfalls die inverse Matrix zu berechnen (modifiziertes Gauß-Verfahren))**.** Bei der Vorstellung des Algorithmus 5.8.1 haben wir angenommen, dass $A \in K^{n\times n}$ invertierbar ist. Falls nur $A \in K^{n\times n}$ gegeben ist, können wir auch mit dem „simultanen“ Lösungsschema $(A \mid I_n)$ starten und dieses auf ZSF bringen.

Wir begründen unten, dass es in der ZSF genau n Pivotelemente geben muss. Nehmen wir das einstweilen an, so gibt es zwei Fälle für die ZSF.

(a) Alle Pivots befinden sich links des senkrechten Strichs: Dann hat die ZSF notwendig die Form $(I_n \mid B)$. Dies bedeutet, dass A invertierbar ist und $B = A^{-1}$ gilt (vgl. Bemerkung 3.2.8 zum Ablesen der Lösung).

(b) Ein Pivot befindet sich rechts des senkrechten Strichs: Dann ist A nicht invertierbar. Befindet sich nämlich ein Pivot in der j-ten Spalte rechts des Strichs (wobei $j \in \{1, \ldots, n\}$), so ist das Gleichungssystem $AX = e_j$ nicht lösbar (vgl. Bemerkung 3.2.8 zum Ablesen der Lösung).

Zu begründen ist noch, dass es genau n Pivotelemente gibt, dass es also keine Nullzeilen am Ende gibt. Dazu überlegt man sich (was nicht schwer ist), dass bei elementaren Zeilenumformungen der **Zeilenrang**, also die Dimension des von den Zeilen einer Matrix aufgespannten Untervektorraums konstant ist (in unserem Fall handelt es sich um einen Unterraum des K^{2n}). Der Zeilenrang von $(A \mid I_n)$ ist offensichtlich n, denn die n Zeilen sind linear unabhängig (was an der Matrix I_n liegt) und bilden somit eine Basis ihres linearen Spanns. Der Zeilenrang der ZSF muss dann auch n sein, was nicht möglich ist, wenn es am Ende Nullzeilen gibt.

Beispiel 5.8.3. Gegeben sei die Matrix $A = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Wir erhalten

$$(A \mid I_n) = \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right).$$

Wir erhalten per $Z_1 \rightsquigarrow Z_1 - Z_3$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & -1 \\ 3 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right),$$

per $Z_2 \rightsquigarrow Z_2 - 3Z_1$ und $Z_3 \rightsquigarrow Z_3 - Z_1$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 4 & 1 & -3 & 1 & 3 \\ 0 & 2 & 1 & -1 & 0 & 2 \end{array}\right),$$

per $Z_2 \rightsquigarrow Z_2 - Z_3$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 2 & 0 & -2 & 1 & 1 \\ 0 & 2 & 1 & -1 & 0 & 2 \end{array}\right),$$

per $Z_3 \rightsquigarrow Z_3 - Z_2$ und $Z_2 \rightsquigarrow \frac{1}{2}Z_2$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right),$$

per $Z_1 \rightsquigarrow Z_1 + Z_2$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & -1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right).$$

Dies bedeutet

$$A^{-1} = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ -1 & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 1 \end{pmatrix}.$$

Zur Probe berechnen wir

$$\begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ -1 & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = I_3.$$

Beispiel 5.8.4. In der Vorlesung wird der Algorithmus für $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$ vorgeführt. Ersetzt man die 10 durch eine 9, so ist die Matrix nicht mehr invertierbar (eine lineare Abhängigkeit der Spalten findet man schnell, wenn man die Differenzen aufeinanderfolgender Spalten betrachtet).

5.9. **Elementarmatrizen und elementare Umformungen.**

5.9.1. Sei $n \in \mathbb{N}$ eine natürliche Zahl. Wir erinnern an die in Beispiel 4.2.8.(c) definierten Matrizen E_{ij}, für $i, j \in \{1, \ldots, n\}$, die eine Basis des $K^{n\times n}$ bilden.

Definition 5.9.2. Die im Folgenden definierten quadratischen Matrizen heißen **Elementarmatrizen**, denn sie gehen jeweils aus der Einheitsmatrix durch die naheliegende elementare Zeilenumformung hervor.

(a) Für $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ und $\lambda \in K$ definiere

$$T_{ij}(\lambda) := I_n + \lambda E_{ij} \in K^{n\times n}.$$

Diese Matrix hat also Einsen auf der Diagonalen, den Eintrag λ an der Position (i, j) und besteht sonst aus Nullen.

Matrizen dieser Gestalt heißen **Elementarmatrizen vom Typ I**.

(b) Für $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ definiere

$$P_{ij} := I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji} \in K^{n \times n}.$$

Diese Matrix hat also Einsen auf der Diagonalen außer an den Positionen (i, i) und (j, j), wo Nullen stehen, hat Einsen an den Positionen (i, j) und (j, i) und besteht sonst aus Nullen.

Matrizen dieser Gestalt heißen **Elementarmatrizen vom Typ II**.

(c) Für $i \in \{1, \ldots, n\}$ und $\mu \in K^\times = K \setminus \{0\}$ definiere

$$D_i(\mu) := I_n - E_{ii} + \mu E_{ii}.$$

Diese Matrix hat also Einsen auf der Diagonalen außer an Position (i, i), wo μ steht, und besteht sonst aus Nullen.

Matrizen dieser Gestalt heißen **Elementarmatrizen vom Typ III**.

5.9.3. Alle Elementarmatrizen haben offensichtlich Rang n und sind deswegen invertierbar (siehe Satz 5.6.5), in Formeln gilt also

$$T_{ij}(\lambda), P_{ij}, D_i(\mu) \in \mathrm{GL}_n(K).$$

Alternativ kann man auch einfach die Inversen angeben und sieht zusätzlich, dass diese ebenfalls Elementarmatrizen sind: Es gelten

$$\begin{aligned} T_{ij}(\lambda)T_{ij}(-\lambda) = (I_n + \lambda E_{ij})(I_n - \lambda E_{ij}) &= I_n + \lambda E_{ij} - \lambda E_{ij} - \lambda^2 E_{ij}E_{ij} = I_n, \\ P_{ij}P_{ji} &= I_n, \\ D_i(\mu)D_i(\mu^{-1}) &= I_n. \end{aligned}$$

Hierbei haben wir verwendet, dass $E_{ij}^2 = E_{ij}E_{ij} = 0$ gilt. Dies muss man nachrechnen: Die Abbildung $E_{ij} \colon K^n \to K^n$ schickt den Standardbasisvektor e_i auf e_j, und alle anderen Standardbasisvektoren auf Null. Wegen $i \neq j$ schickt also E_{ij}^2 alle Standardbasisvektoren auf Null und ist somit die Nullabbildung. Man kann dies auch brutal durchrechnen, etwa mit Hilfe von $(E_{ij})_{st} = \delta_{is}\delta_{jt}$ unter Verwendung des Kronecker-δ.

Bemerkung 5.9.4 (Elementare Zeilenumformungen als Linksmultiplikation mit Elementarmatrizen)**.** Sei $A \in K^{m \times n}$ eine $(m \times n)$-Matrix. Seien $i, j \in \{1, \ldots, m\}$ mit $i \neq j$ und $\lambda \in K$ und $\mu \in K \setminus \{0\}$.

(I)$_\mathrm{Z}$ Die Matrix, die aus A entsteht, indem man das λ-fache der j-ten Zeile zur i-ten Zeile addiert, ist $T_{ij}(\lambda)A$.

(II)$_\mathrm{Z}$ Die Matrix, die aus A entsteht, indem man i-te und j-te Zeile miteinander vertauscht, ist $P_{ij}A$.

(III)$_\mathrm{Z}$ Die Matrix, die aus A entsteht, indem man i-te Zeile mit μ multipliziert, ist $D_i(\mu)A$.

Definition 5.9.5. Elementare Spaltenumformungen einer Matrix sind analog zu den elementaren Zeilenumformungen definiert.

Bemerkung 5.9.6 (Elementare Spaltenumformungen als Rechtsmultiplikation mit Elementarmatrizen)**.** Sei $A \in K^{m \times n}$ eine $(m \times n)$-Matrix. Seien $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ und $\lambda \in K$ und $\mu \in K \setminus \{0\}$.

(I)$_\mathrm{S}$ Die Matrix, die aus A entsteht, indem man das λ-fache der i-ten Spalte zur j-ten Spalte addiert, ist $AT_{ij}(\lambda)$.

(II)$_S$ Die Matrix, die aus A entsteht, indem man i-te und j-te Spalte miteinander vertauscht, ist AP_{ij}.

(III)$_S$ Die Matrix, die aus A entsteht, indem man die i-te Spalte von A mit μ multipliziert, ist $AD_i(\mu)$.

Bemerkung 5.9.7. Da elementare Zeilen- bzw. Spaltenumformungen Links- bzw. Rechtsmultiplikationen mit invertierbaren Matrizen entsprechen (siehe 5.9.3 und Bemerkungen 5.9.4 und 5.9.6), erhalten sie den Rang: Die Matrix und die umgeformte Matrix haben denselben Rang. Dies folgt aus Lemma 5.6.10.

Es sei darauf hingewiesen, dass elementare Spaltenumformungen offensichtlich den Spaltenrang erhalten; für den Zeilenrang ist dies a priori nicht klar. Ebenso offensichtlich erhalten elementare Zeilenumformungen den Zeilenrang; für den Spaltenrang ist dies a priori nicht klar. Jedoch stimmen Zeilen- und Spaltenrang überein (siehe Satz 5.11.7).

Satz 5.9.8. *Jedes Element von* $\mathrm{GL}_n(K)$ *kann als Produkt von Elementarmatrizen der Typen I, II und III geschrieben werden. Insbesondere wird die Gruppe* $\mathrm{GL}_n(K)$ *von diesen Elementarmatrizen erzeugt.*

Beweis. Dies folgt sofort aus dem Algorithmus 5.8.1 zum Berechnen der inversen Matrix. Wir haben dort gesehen, dass für jedes Element $A \in \mathrm{GL}_n$ die Matrix (5.8.1) durch elementare Zeilenumformungen auf die Gestalt (5.8.2) gebracht werden kann. Insbesondere kann die Matrix A durch elementare Zeilenumformungen auf die Gestalt I_n gebracht werden. Da elementare Zeilenumformungen schlicht Linksmultiplikationen mit Elementarmatrizen sind (Bemerkung 5.9.4), gibt es Elementarmatrizen $C_1, \dots, C_m$, so dass $C_m C_{m-1} \dots C_2 C_1 A = I_n$ gilt. Nach 5.9.3 sind Elementarmatrizen invertierbar mit Elementarmatrizen als Inverse. Unsere Gleichung is somit äquivalent zu der gesuchten Darstellung $A = C_1^{-1} C_2^{-1} \dots C_m^{-1}$. □

Ende 16. Vorlesung am 18.06.2020

5.10. Darstellende Matrizen bezüglich Basen.

Definition 5.10.1. Sei $f\colon V \to W$ ein Homomorphismus von einem n-dimensionalen Vektorraum V in einen m-dimensionalen Vektorraum W. Sei $\mathcal{B} = (b_1, \dots, b_n)$ eine Basis von V und sei $\mathcal{C} = (c_1, \dots, c_m)$ eine Basis von W. Die **darstellende Matrix** ${}_{\mathcal{C}}[f]_{\mathcal{B}} \in K^{m\times n}$ **von** f **in Bezug auf die Basen** $\mathcal{B}$ **und** $\mathcal{C}$ ist diejenige $(m \times n)$-Matrix, deren j-te Spalte als Einträge die Koordinaten von $f(b_j)$ bezüglich der Basis $\mathcal{C}$ hat, für alle $j = 1, \dots, n$; in Formeln

$$f(b_j) = \sum_{i=1}^{m} ({}_{\mathcal{C}}[f]_{\mathcal{B}})_{ij} c_i \qquad \text{für alle } j = 1, \dots, n. \tag{5.10.1}$$

5.10.2. Setze $A := {}_{\mathcal{C}}[f]_{\mathcal{B}}$. Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ das Koordinatentupel (als Spaltenvektor) eines beliebigen Vektors $v \in V$ bezüglich der Basis $\mathcal{B}$, d. h. es gilt $v = \sum_{j=1}^{n} x_j b_j$. Wir erhalten

$$f(v) = f\left(\sum_{j=1}^{n} x_j b_j\right) = \sum_{j=1}^{n} x_j f(b_j) \overset{(5.10.1)}{=} \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} c_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) c_i$$

$$= \sum_{i=1}^{m} (Ax)_i c_i = \sum_{i=1}^{m} ({}_{\mathcal{C}}[f]_{\mathcal{B}} x)_i c_i.$$

Also ist $Ax = {}_{\mathcal{C}}[f]_{\mathcal{B}} x \in K^m$ das Koordinatentupel von $f(v)$ bezüglich der Basis $\mathcal{C}$.

Das Koordinatentupel von $f(v)$ (bezüglich $\mathcal{C}$) erhält man also durch Multiplikation der darstellenden Matrix $A = {}_{\mathcal{C}}[f]_{\mathcal{B}}$ mit dem Koordinatentupel von v (bezüglich $\mathcal{B}$).

Dies erklärt den Begriff *darstellende Matrix*.

Bemerkung 5.10.3. Seien $\mathcal{S}_n = (e_1, \ldots, e_n)$ die Standardbasis von K^n und $\mathcal{S}_m$ die Standardbasis von K^m. Dann ist die darstellende Matrix $[f]$ einer linearen Abbildung $f\colon K^n \to K^m$, wie in Definition 5.4.1 erklärt, die darstellende Matrix ${}_{\mathcal{S}_m}[f]_{\mathcal{S}_n}$ bezüglich der Standardbasen im Sinne der Definition 5.10.1, in Formeln $[f] = {}_{\mathcal{S}_m}[f]_{\mathcal{S}_n}$. Dies ist klar, denn die j-te Spalte von $[f]$ ist $f(e_j)$ und die Einträge von $f(e_j)$ sind nun einmal die Koordinaten von $f(e_j)$ bezüglich $\mathcal{S}_m$.

5.10.4. Für jede Basis $\mathcal{B} = (b_1, \ldots, b_n)$ eines Vektorraums V gibt es nach Korollar 5.3.4 genau einen Isomorphismus

$$\varphi_{\mathcal{B}}\colon K^n \xrightarrow{\sim} V$$

mit $\varphi_{\mathcal{B}}(e_i) = b_i$ für alle $i = 1, \ldots, n$.

Proposition 5.10.5. *Sei $f\colon V \to W$ ein Homomorphismus eines n-dimensionalen Vektorraums V in einen m-dimensionalen Vektorraum W. Seien $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V und $\mathcal{C} = (c_1, \ldots, c_m)$ eine Basis von W. Fassen wir die Matrix ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ als lineare Abbildung $K^n \to K^m$ auf, so gilt*

$$ {}_{\mathcal{C}}[f]_{\mathcal{B}} = \varphi_{\mathcal{C}}^{-1} \circ f \circ \varphi_{\mathcal{B}}\colon K^n \to K^m. \tag{5.10.2}$$

Genauer ist ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ durch diese Gleichung eindeutig charakterisiert: Genau dann erfüllt eine Matrix $A \in K^{m\times n}$ die Gleichung

$$A = \varphi_{\mathcal{C}}^{-1} \circ f \circ \varphi_{\mathcal{B}}\colon K^n \to K^m, \tag{5.10.3}$$

wenn $A = {}_{\mathcal{C}}[f]_{\mathcal{B}}$ gilt.

5.10.6. Ich merke mir den Inhalt von Proposition 5.10.5 mit Hilfe von Diagrammen. Die Gleichheit (5.10.2) ist äquivalent zu $\varphi_{\mathcal{C}} \circ {}_{\mathcal{C}}[f]_{\mathcal{B}} = f \circ \varphi_{\mathcal{B}}$ und besagt, dass es im Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle\varphi_{\mathcal{B}}}\uparrow{\scriptstyle\sim} & & {\scriptstyle\varphi_{\mathcal{C}}}\uparrow{\scriptstyle\sim} \\ K^n & \xrightarrow{{}_{\mathcal{C}}[f]_{\mathcal{B}}} & K^m \end{array} \tag{5.10.4}$$

egal ist, ob man von K^n nach W mit Hilfe der Abbildungen über die rechte untere Ecke K^m oder über die linke obere Ecke V geht. Analog kann man sich die Gleichheit (5.10.3) graphisch vorstellen, indem man die Beschriftung ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ in diesem Diagramm durch A ersetzt.

Beweis. Für $A \in K^{m\times n}$ sind die folgenden Aussagen äquivalent:

- Die Gleichheit (5.10.3) gilt.
- Die Gleichheit $\varphi_{\mathcal{C}} \circ A = f \circ \varphi_{\mathcal{B}}$ von linearen Abbildungen $K^n \to W$ gilt.

- Für alle $j = 1, \dots, n$ gilt
$$\sum_{i=1}^{m} a_{ij} c_i = f(b_j).$$
(Da zwei lineare Abbildungen nach Satz 5.3.1 genau dann gleich sind, wenn sie auf einer Basis des Ausgangsraums übereinstimmen.)
- $A = {}_{\mathcal{C}}[f]_{\mathcal{B}}$. □

5.10.7. Da die beiden Abbildungen $\varphi_{\mathcal{B}}$ und $\varphi_{\mathcal{C}}$ in (5.10.2) (also die beiden vertikalen Pfeile in Diagramm (5.10.4)) bijektiv sind, sehen wir: Der Homomorphismus f ist genau dann injektiv (bzw. surjektiv, bijektiv), wenn ${}_{\mathcal{C}}[f]_{\mathcal{B}} \colon K^n \to K^m$ injektiv (bzw. surjektiv, bijektiv) ist.

Insbesondere ist f genau dann ein Isomorphismus, wenn $n = m$ und ${}_{\mathcal{C}}[f]_{\mathcal{B}} \in \mathrm{GL}_n(K)$ gelten (dies verwendet Satz 5.3.5).

5.10.8 (Rang einer Abbildung als Rang jeder darstellenden Matrix)**.** Der Rang von f stimmt mit dem Rang der darstellenden Matrix von f bezüglich beliebig gewählter Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W überein, in Formeln
$$\mathrm{rk}(f) = \mathrm{rk}({}_{\mathcal{C}}[f]_{\mathcal{B}}).$$
Dies folgt aus der Gleichheit (5.10.2) und der folgenden Rechnung, in der wir Lemma 5.6.10 zweimal anwenden:
$$\mathrm{rk}(f) = \mathrm{rk}(f \circ \varphi_{\mathcal{B}}) = \mathrm{rk}(\varphi_{\mathcal{C}} \circ {}_{\mathcal{C}}[f]_{\mathcal{B}}) = \mathrm{rk}({}_{\mathcal{C}}[f]_{\mathcal{B}}).$$

5.10.9. Fixieren wir wie oben Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W, so ist die Abbildung
$$\begin{aligned} \mathrm{Hom}_K(V, W) &\xrightarrow{\sim} K^{m \times n}, \\ f &\mapsto {}_{\mathcal{C}}[f]_{\mathcal{B}}, \end{aligned} \tag{5.10.5}$$
bijektiv[33]. Dies folgt einerseits direkt aus der Definition 5.10.1 der darstellenden Matrix bezüglich gewählter Basen und Satz 5.3.1 oder andererseits aus der Gleichung (5.10.2) und der Bijektivität von $\varphi_{\mathcal{B}}$ und $\varphi_{\mathcal{C}}$.

Beispiele 5.10.10. (a) Sei V ein Vektorraum mit Basis $\mathcal{B} = (b_1, \dots, b_n)$. Die darstellende Matrix der Identität $\mathrm{id}_V \colon V \to V$ bezüglich der Basen $\mathcal{B}$ und $\mathcal{B}$ ist die Einheitsmatrix, in Formeln ${}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}} = I_n$. Beachte, dass diese Aussage für jede Wahl einer Basis von V gilt.

(b) Sei $\mathcal{S} = (e_1, e_2)$ die Standardbasis von $\mathbb{R}^2$. Sei $\mathcal{B} = \left(x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right)$ eine beliebige Basis von $\mathbb{R}^2$. Betrachte die eindeutige lineare Abbildung $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(e_1) = x$ und $f(e_2) = y$ (siehe Satz 5.3.1). Dann gilt
$${}_{\mathcal{B}}[f]_{\mathcal{S}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$
(obwohl f nicht die Identität ist, falls $\mathcal{B} \neq \mathcal{S}$) aber
$${}_{\mathcal{S}}[f]_{\mathcal{S}} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$

[33] Sie ist offensichtlich sogar ein Isomorphismus von Vektorräumen, wenn wir $\mathrm{Hom}_K(V, W)$ wie in Lemma 5.1.11 und $K^{m \times n}$ wie in Beispiel 4.1.6 als Vektorraum auffassen.

(c) Betrachte die lineare Abbildung

$$f := \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} : \mathbb{R}^3 \to \mathbb{R}^2.$$

Sei $\mathcal{S} = (e_1, e_2, e_3)$ die Standardbasis von $\mathbb{R}^3$ und $\mathcal{C} := \left(c_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, c_2 := \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$, was eine Basis von $\mathbb{R}^2$ ist. Wegen $f(e_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot c_1 + 0 \cdot c_2$ und $f(e_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2 \cdot c_1 - 1 \cdot c_2$ und $f(e_3) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 3 \cdot c_1 - 2 \cdot c_2$ erhalten wir

$${}_{\mathcal{C}}[f]_{\mathcal{S}} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix}.$$

Satz 5.10.11. *Seien $g\colon U \to V$ und $f\colon V \to W$ lineare Abbildungen zwischen endlichdimensionalen Vektorräumen. Seien $\mathcal{A}$ eine Basis von U, $\mathcal{B}$ eine Basis von V und $\mathcal{C}$ Basis von W. Dann gilt*

$${}_{\mathcal{C}}[f]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[g]_{\mathcal{A}} = {}_{\mathcal{C}}[f \circ g]_{\mathcal{A}}$$

Beweis. Fassen wir die drei Matrizen als lineare Abbildungen auf, so ist nach Satz 5.4.8 die Gleichheit ${}_{\mathcal{C}}[f]_{\mathcal{B}} \circ {}_{\mathcal{B}}[g]_{\mathcal{A}} = {}_{\mathcal{C}}[f \circ g]_{\mathcal{A}}$ zu zeigen. Nach Proposition 5.10.5 gilt aber

$$\begin{aligned} {}_{\mathcal{C}}[f]_{\mathcal{B}} \circ {}_{\mathcal{B}}[g]_{\mathcal{A}} &= (\varphi_{\mathcal{C}}^{-1} \circ f \circ \varphi_{\mathcal{B}}) \circ (\varphi_{\mathcal{B}}^{-1} \circ g \circ \varphi_{\mathcal{A}}) \\ &= \varphi_{\mathcal{C}}^{-1} \circ f \circ \varphi_{\mathcal{B}} \circ \varphi_{\mathcal{B}}^{-1} \circ g \circ \varphi_{\mathcal{A}} \\ &= \varphi_{\mathcal{C}}^{-1} \circ (f \circ g) \circ \varphi_{\mathcal{A}} \\ &= {}_{\mathcal{C}}[f \circ g]_{\mathcal{A}}. \end{aligned}$$

Hier noch ein Diagramm, das den Beweis veranschaulicht (vgl. 5.10.6):

$$\begin{array}{ccccc} U & \xrightarrow{g} & V & \xrightarrow{f} & W \\ {\scriptstyle \varphi_{\mathcal{A}}}\uparrow{\scriptstyle \sim} & & {\scriptstyle \varphi_{\mathcal{B}}}\uparrow{\scriptstyle \sim} & & {\scriptstyle \varphi_{\mathcal{C}}}\uparrow{\scriptstyle \sim} \\ K^r & \xrightarrow[{}_{\mathcal{B}}[g]_{\mathcal{A}}]{} & K^n & \xrightarrow[{}_{\mathcal{C}}[f]_{\mathcal{B}}]{} & K^m \end{array}$$

□

Korollar 5.10.12. *Seien $\mathcal{B}$ und $\mathcal{C}$ Basen eines n-dimensionalen Vektorraums V. Dann gilt*

$${}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{C}} = {}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{C}} = I_n.$$

Somit ist ${}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}} \in \mathrm{GL}_n(K)$ invertierbar mit Inversem ${}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{C}}$, in Formeln

$$({}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}})^{-1} = {}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{C}}.$$

Beweis. Die erste Gleichheit ist Satz 5.10.11 spezialisiert zu $U = V = W$ und $f = g = \mathrm{id}_V$ und $\mathcal{A} = \mathcal{C}$. Die zweite Gleicheit ist offensichtlich und wurde bereits in Beispiel 5.10.10.(a) beobachtet. Aus Symmetriegründen folgt die letzte Behauptung des Korollars; alternativ ist die Existenz einer rechtsinversen Matrix schon ausreichend für Invertierbarkeit nach Satz 5.6.5. □

Satz 5.10.13 (Smith-Normalform)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen vom Rang $r := \operatorname{rk}(f)$. Dann existieren Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W, so dass*

$$ {}_{\mathcal{C}}[f]_{\mathcal{B}} = \left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{array}\right) = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \tag{5.10.6} $$

gilt. Die angedeutete Matrix hat nur auf der Diagonalen[34] *von Null verschiedene Einträge, und zwar zuerst r Einsen und dann lauter Nullen. Man beachte, dass diese Matrix im Allgemeinen nicht quadratisch ist: Sie hat $\dim W$ Zeilen und $\dim V$ Spalten.*

Beweis. Dies ist schlicht eine Umformulierung von Korollar 5.5.7. □

Satz 5.10.14 (Smith-Normalform für Matrizen)**.** *Sei $A \in K^{m\times n}$ eine Matrix. Dann gibt es Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit*

$$ SAT = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) $$

wobei $r := \operatorname{rk}(A)$ und die angedeutete Matrix die aus (5.10.6) ist.

Beweis. Wenden wir Satz 5.10.13 auf $f = A\colon V = K^n \to W = K^m$ an, so erhalten wir Basen $\mathcal{B}$ von K^n und $\mathcal{C}$ von K^m, so dass die darstellende Matrix $M := {}_{\mathcal{C}}[A]_{\mathcal{B}}$ die gesuchte Gestalt hat. Nach Proposition 5.10.2 gilt also in Formeln

$$ M = \varphi_{\mathcal{C}}^{-1} \circ A \circ \varphi_{\mathcal{B}}. $$

Für $T := \varphi_{\mathcal{B}} \in \mathrm{GL}_n(K)$ und $S := \varphi_{\mathcal{C}}^{-1} \in \mathrm{GL}_m(K)$ gilt damit die gewünschte Behauptung. Hier ein Diagramm zu Veranschaulichung:

$$ \begin{array}{ccc} K^n & \xrightarrow{A} & K^m \\ {\scriptstyle T=\varphi_{\mathcal{B}}}\uparrow{\scriptstyle \sim} & & {\scriptstyle \sim}\uparrow{\scriptstyle S^{-1}=\varphi_{\mathcal{C}}} \\ K^n & \xrightarrow{M} & K^m \end{array} $$

□

Aufgabe 5.10.15. Geben Sie einen Alternativbeweis von Satz 5.10.14:

- Bringen Sie A zuerst auf ZSF, indem sie A von links mit geeigneten Elementarmatrizen multiplizieren. Hinweis: Gauß-Verfahren.
- Bringen Sie das Resultat dann auf die Smith-Normalform, indem sie von rechts mit geeigneten Elementarmatrizen multiplizieren, also geeignete elementare Spaltenumformumgen vornehmen.

[34] Auch bei nicht notwendig quadratischen Matrizen A sprechen wir von der Diagonalen und meinen damit alle Einträge a_{ii} an Koordinaten mit demselben Zeilen- und Spaltenindex.

5.11. Spaltenrang = Zeilenrang.

Definition 5.11.1. Sei $A = (a_{ij}) \in K^{m\times n}$ eine $(m \times n)$-Matrix. Die **transponierte Matrix** von A ist die durch

$$(A^{\mathrm{t}})_{ij} := a_{ji}$$

definierte $(n \times m)$-Matrix $A^{\mathrm{t}} \in K^{n\times m}$.

5.11.2. Anschaulich entsteht die transponierte Matrix A^{t} aus A durch Spiegeln an der Diagonalen, denn der Eintrag von A^{t} an der Position (i, j) ist der Eintrag von A an der Position (j, i). Mit anderen Worten ist die i-te Zeile von A^{t} die i-te Spalte von A (als Zeilenvektor), und die j-te Spalte von A^{t} ist die j-te Zeile von A (als Spaltenvektor).

Beispiel 5.11.3. Die Transponierte von $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ ist $A^{\mathrm{t}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Lemma 5.11.4. *Transponieren $(-)^{\mathrm{t}}\colon K^{m\times n} \to K^{n\times m}$, $A \mapsto A^{\mathrm{t}}$, ist eine lineare Abbildung. Es gelten*

$$(A^{\mathrm{t}})^{\mathrm{t}} = A \qquad \text{für alle } A \in K^{m\times n} \text{ und}$$
$$(AB)^{\mathrm{t}} = B^{\mathrm{t}}A^{\mathrm{t}} \qquad \text{für alle } A \in K^{m\times n} \text{ und alle } B \in K^{n\times r}.$$

Insbesondere gilt: Eine quadratische Matrix A ist genau dann invertierbar, wenn A^{t} invertierbar ist; in diesem Fall gilt $(A^{\mathrm{t}})^{-1} = (A^{-1})^{\mathrm{t}}$.

Beweis. Offensichtlich ist Transponieren linear. Die Gleichheit $(A^{\mathrm{t}})^{\mathrm{t}} = A$ ist ebenfalls klar. Die Gleichheit $(AB)^{\mathrm{t}} = B^{\mathrm{t}}A^{\mathrm{t}}$ rechnet man wie folgt komponenten weise nach:

$$((AB)^{\mathrm{t}})_{ij} = (AB)_{ji} = \sum_{s=1}^{n} A_{js}B_{si} = \sum_{s=1}^{n} (B^{\mathrm{t}})_{is}(A^{\mathrm{t}})_{sj} = (B^{\mathrm{t}}A^{\mathrm{t}})_{ij}$$

Ist $A \in \mathrm{GL}_n(K)$, so gilt $AA^{-1} = I_n$, woraus $(A^{-1})^{\mathrm{t}}A^{\mathrm{t}} = (AA^{-1})^{\mathrm{t}} = (I_n)^{\mathrm{t}} = I_n$ folgt, also ist A^{t} invertierbar mit Inversem $(A^{\mathrm{t}})^{-1} = (A^{-1})^{\mathrm{t}}$. Ist A^{t} invertierbar, so ist also $A = (A^{\mathrm{t}})^{\mathrm{t}}$ invertierbar. □

5.11.5. Sei $A \in K^{m\times n}$. Wir erinnern daran, dass der Rang $\mathrm{rk}(A)$ von A (siehe Definition 5.5.1) die Dimension des von den Spalten von A in K^m aufgespannten Untervektorraums ist (siehe Bemerkung 5.5.2) und deswegen auch *Spaltenrang* heißt.

Man nennt $\mathrm{rk}(A^{\mathrm{t}})$ den **Zeilenrang** von A, denn $\mathrm{rk}(A^{\mathrm{t}})$ ist die Dimension des von den Spalten von A^{t}, also den Zeilen von A, in K^n aufgespannten Untervektorraums.

Korollar 5.11.6. *Eine quadratische Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn ihre Zeilen linear unabhängig sind (bzw. ein Erzeugendensystem bzw. eine Basis des K^n bilden bzw. wenn $\mathrm{rk}(A^{\mathrm{t}}) = n$ gilt).*

Beweis. Nach Lemma 5.11.4 ist die Invertierbarkeit von A äquivalent zur Invertierbarkeit von A^{t}. Letztere ist nach Satz 5.6.5 äquivalent zu den angegebenen Bedingungen, da die Spalten von A^{t} die Zeilen von A sind. □

Satz 5.11.7 (Spaltenrang=Zeilenrang)**.** *Sei $A \in K^{m\times n}$. Dann gilt* $\mathrm{rk}(A) = \mathrm{rk}(A^{\mathrm{t}})$.

Beweis. Sei $r := \operatorname{rk}(A)$. Nach Satz 5.10.14 gibt es Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit

$$SAT = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

wobei die angedeutete Matrix die aus (5.10.6) ist. Für diese Matrix stimmen offenbar Spalten- und Zeilenrang überein, es gilt also $\operatorname{rk}(SAT) = \operatorname{rk}((SAT)^{\mathrm{t}})$. Aus dieser Gleichheit, Lemma 5.6.10 und Lemma 5.11.4 folgern wir

$$\operatorname{rk}(A) = \operatorname{rk}(SAT) = \operatorname{rk}((SAT)^{\mathrm{t}}) = \operatorname{rk}(T^{\mathrm{t}}A^{\mathrm{t}}S^{\mathrm{t}}) = \operatorname{rk}(A^{\mathrm{t}}). \qquad \square$$

Ende 17. Vorlesung am 23.06.2020

5.12. **Basiswechsel.**

Definition 5.12.1 (Basiswechselmatrix)**.** Sei V ein endlichdimensionaler Vektorraum der Dimension $n = \dim V$. Sind $\mathcal{B}$ und $\mathcal{C}$ zwei Basen von V, so heißt die darstellende Matrix

$${}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}} \in \mathrm{GL}_n(K)$$

der Identität von V bezüglich dieser Basen die **Basiswechselmatrix von $\mathcal{B}$ nach $\mathcal{C}$** (sie ist nach Korollar 5.10.12 invertierbar).

5.12.2. Gelte $\mathcal{B} = (b_1, \ldots, b_n)$ und $\mathcal{C} := (c_1, \ldots, c_n)$. Die Gleichungen (5.10.1) spezialisieren für $f = \mathrm{id}_V$ zu

$$b_j = \sum_{i=1}^{n} ({}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}})_{ij} c_i \quad \text{für alle } j = 1, \ldots, n. \tag{5.12.1}$$

Die j-te Spalte der Basiswechselmatrix von $\mathcal{B}$ nach $\mathcal{C}$ besteht also gerade aus den Koordinaten von b_j bezüglich der Basis $\mathcal{C}$.

Spezialisiert man 5.10.2 auf den Fall $f = \mathrm{id}_V$, so erhält man: Das Koordinatentupel von $v \in V$ (bezüglich $\mathcal{C}$) erhält man durch Multiplikation der Basiswechselmatrix $A = {}_{\mathcal{C}}[\mathrm{id}_V]_{\mathcal{B}}$ mit dem Koordinatentupel von v (bezüglich $\mathcal{B}$). Dies erklärt den Begriff *Basiswechselmatrix*.

Satz 5.12.3 (Unabhängiger Basiswechsel im Start- und Zielvektorraum eines Homomorphismus)**.** *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Seien $\mathcal{B}$ und $\mathcal{B}'$ Basen von V, und seien $\mathcal{C}$ und $\mathcal{C}'$ Basen von W. Dann gilt*

$${}_{\mathcal{C}'}[f]_{\mathcal{B}'} = {}_{\mathcal{C}'}[\mathrm{id}_W]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[f]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}'} = {}_{\mathcal{C}'}[\mathrm{id}_W]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[f]_{\mathcal{B}} \cdot ({}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}})^{-1},$$

die darstellende Matrix ${}_{\mathcal{C}'}[f]_{\mathcal{B}'}$ kann also aus der darstellenden Matrix ${}_{\mathcal{C}}[f]_{\mathcal{B}}$ von f und den Basiswechselmatrizen berechnet werden.

Setzen wir $A := {}_{\mathcal{C}}[f]_{\mathcal{B}}$, $A' := {}_{\mathcal{C}'}[f]_{\mathcal{B}'}$, $S := {}_{\mathcal{C}'}[\mathrm{id}_W]_{\mathcal{C}}$ und $T := {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}}$, so gilt also

$$A' = SAT^{-1}.$$

Beweis. Dies folgt sofort aus Satz 5.10.11 und ${}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}'} = ({}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}})^{-1}$ (siehe Korollar 5.10.12). $\square$

Definition 5.12.4. Gegeben Matrizen $A, B \in K^{m\times n}$ schreiben wir genau dann $A \sim B$ und sagen, dass A zu B **äquivalent** ist, wenn es invertierbare Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $B = SAT^{-1}$ gibt.

5.12.5. Dies definiert eine Äquivalenzrelation $\sim$ auf $K^{m\times n}$ (und rechtfertigt den Namen *äquivalent*):

- reflexiv: $A \sim A$: Nimm $T = I_n$ und $S = I_m$.
- symmetrisch: $A \sim B \iff B = SAT^{-1}$ für geeignete $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ $\iff S^{-1}BT = A$ für geeignete $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K) \iff B \sim A$.
- transitiv: $A \sim B$ und $B \sim C \iff B = SAT^{-1}$ und $C = S'BT'^{-1}$ für geeignete $S, S' \in \mathrm{GL}_m(K)$ und $T, T' \in \mathrm{GL}_n(K) \implies C = S'BT'^{-1} = S'SAT^{-1}T'^{-1} = (S'S)A(T'T)^{-1} \iff A \sim C$.

Korollar 5.12.6. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Setze $n := \dim(V)$ und $m := \dim(W)$. Sei $\mathcal{B}$ eine Basis von V, sei $\mathcal{C}$ eine Basis von W, und sei $A := {}_{\mathcal{C}}[f]_{\mathcal{B}}$. Dann lässt sich die Menge aller darstellenden Matrizen von f bezüglich aller Basen von V und W wie folgt beschreiben:*

$$\{{}_{\mathcal{C}'}[f]_{\mathcal{B}'} \mid \mathcal{B}' \text{ Basis von } V \text{ und } \mathcal{C}' \text{ Basis von } W\} = \{SAT^{-1} \mid S \in \mathrm{GL}_m(K), T \in \mathrm{GL}_n(K)\}.$$

Die Menge auf der rechten Seite ist die Äquivalenzklasse von A bezüglich der Äquivalenzrelation $\sim$, also die Menge aller zu A äquivalenten Matrizen.

Beweis. Die Inklusion $\subset$ folgt sofort aus Satz 5.12.3.

Die Inklusion $\supset$ benötigt die folgende Hilfsaussage: Jedes Element $T \in \mathrm{GL}_n(K)$ ist eine Basiswechselmatrix von $\mathcal{B}$ zu einer geeigneten (sogar eindeutigen) Basis $\mathcal{B}'$ von V, in Formeln $T = {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}}$.

Betrachte die Verknüpfung

$$\varphi_{\mathcal{B}} \circ T^{-1}\colon K^n \xrightarrow[\sim]{T^{-1}} K^n \xrightarrow[\sim]{\varphi_{\mathcal{B}}} W$$

von Isomorphismen. Sie bildet die Elemente der Standardbasis von K^n nach Korollar 5.3.4 auf eine Basis von W ab, die wir $\mathcal{B}'$ nennen. Es gilt also

$$\varphi_{\mathcal{B}} \circ T^{-1} = \varphi_{\mathcal{B}'}$$

oder äquivalent

$$T = \varphi_{\mathcal{B}'}^{-1} \circ \varphi_{\mathcal{B}} = \varphi_{\mathcal{B}'}^{-1} \circ \mathrm{id}_V \circ \varphi_{\mathcal{B}}$$

Nach Proposition 5.10.5 folgt $T = {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}}$. Dies zeigt die Hilfsaussage.

Seien nun $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$. Nach der Hilfsaussage existieren Basen $\mathcal{B}'$ von V und $\mathcal{C}'$ von W mit $T = {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}}$ und $S = {}_{\mathcal{C}'}[\mathrm{id}_W]_{\mathcal{C}}$. Satz 5.12.3 liefert dann

$$SAT^{-1} = {}_{\mathcal{C}'}[\mathrm{id}_W]_{\mathcal{C}} \cdot {}_{\mathcal{C}}[f]_{\mathcal{B}} \cdot ({}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}})^{-1} = {}_{\mathcal{C}'}[f]_{\mathcal{B}'}.$$

Dies zeigt die Inklusion $\supset$. Dass rechts die Äquivalenzklasse von A steht, ist offensichtlich. □

Satz 5.12.7 (Klassifikation von Matrizen bis auf Äquivalenz)**.** *Seien $m, n \in \mathbb{N}$ fixiert. Zwei Matrizen $A, B \in K^{m\times n}$ sind genau dann äquivalent, wenn $\mathrm{rk}(A) = \mathrm{rk}(B)$ gilt. Insbesondere ist jede $(m \times n)$-Matrix vom Rang r äquivalent zur Matrix $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \in K^{m\times n}$.*

Beweis. Seien A und B äquivalent, d. h. es gibt $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $B = SAT^{-1}$. Daraus folgt $\mathrm{rk}(A) = \mathrm{rk}(SAT^{-1}) = \mathrm{rk}(B)$ nach Lemma 5.6.10.

Gelte umgekehrt $\mathrm{rk}(A) = \mathrm{rk}(B)$. Dann sind nach Satz 5.10.14 sowohl A als auch B zur Matrix $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ in Smith-Normalform äquivalent. Weil Äquivalenz von Matrizen eine Äquivalenzrelation ist, sind A und B äquivalent. □

5.12.8. Für jedes $r \in \{0, 1, \ldots, \min(m,n)\}$ hat die Matrix $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \in K^{m\times n}$ offenbar Rang r. Diese Matrizen bilden eine vollständige Liste aller $(m \times n)$-Matrizen bis auf Äquivalenz, nach Satz 5.12.7. In Formeln liefert die Rankabbildung $A \mapsto \mathrm{rk}(A)$ also eine Bijektion

$$K^{m\times n}/\sim \xrightarrow{\sim} \{0, 1, 2, \ldots, \min(m,n)\},$$

wobei links die Menge der Äquivalenzklassen bezüglich $\sim$ steht.

5.12.9. Wir haben gesehen, dass jeder Homomorphismus $f\colon V \to W$ zwischen endlichdimensionalen Vektorräumen bei geschickter Wahl von Basen im Start- und Zielraum durch eine Matrix in Smith-Normalform $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ dargestellt wird (siehe Satz 5.10.13).

Insbesondere kann man dies auf Endomorphismen $f\colon V \to V$ anwenden. Jedoch ist es in diesem Fall naheliegend zu fragen, ob es eine einzige Basis $\mathcal{B}$ von V gibt (statt zweier Basen $\mathcal{B}$ und $\mathcal{C}$ wie bei der Smith-Normalform), so dass die darstellende Matrix ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ von f bezüglich diese Basis eine möglichst einfache Gestalt hat. Die Jordansche Normalform, die wir später kennenlernen werden (siehe Satz 7.5.5), liefert für $K = \mathbb{C}$ (und jeden anderen algebraisch abgeschlossenen Körper) eine befriedigende Antwort auf diese Frage.

Satz 5.12.10 (Basiswechsel für Endomorphismen)**.** *Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums V mit Basen $\mathcal{B}$ und $\mathcal{B}'$. Dann gilt*

$${}_{\mathcal{B}'}[f]_{\mathcal{B}'} = {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[f]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}'} = {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[f]_{\mathcal{B}} \cdot ({}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}})^{-1},$$

die darstellende Matrix ${}_{\mathcal{B}'}[f]_{\mathcal{B}'}$ kann also aus der darstellenden Matrix ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ von f und der Basiswechselmatrix ${}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}}$ berechnet werden.

Setzen wir $A := {}_{\mathcal{B}}[f]_{\mathcal{B}}$, $A' := {}_{\mathcal{B}'}[f]_{\mathcal{B}'}$ und $T := {}_{\mathcal{B}'}[\mathrm{id}_V]_{\mathcal{B}}$, so gilt also

$$A' = TAT^{-1}.$$

Beweis. Dies wird wie Satz 5.12.3 bewiesen bzw. ist auch ein Spezialfall dieses Satzes. □

Definition 5.12.11. Gegeben quadratische Matrizen $A, B \in K^{n\times n}$ schreiben wir genau dann $A \approx B$ und sagen, dass A und B **ähnlich** oder **konjugiert** sind, wenn es eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$ mit $B = TAT^{-1}$ gibt.

5.12.12. Dies definiert eine Äquivalenzrelation $\approx$ auf $K^{n\times n}$. Der Beweis ist vollkommen analog zu dem Beweis in 5.12.5.

Korollar 5.12.13. *Sei $f\colon V \to V$ ein Endomorphismus eines n-dimensionalen Vektorraums. Seien $\mathcal{B}$ eine Basis von V und $A := {}_{\mathcal{B}}[f]_{\mathcal{B}}$. Dann gilt:*

$$\{{}_{\mathcal{B}'}[f]_{\mathcal{B}'} \mid \mathcal{B}' \text{ Basis von } V\} = \{TAT^{-1} \mid T \in \mathrm{GL}_n(K)\}. \tag{5.12.2}$$

Die Menge auf der rechten Seite ist die Äquivalenzklasse von A bezüglich der Äquivalenzrelation $\approx$, also die Menge aller zu A ähnlichen Matrizen.

Beweis. Der Beweis verwendet Satz 5.12.10 und ist ansonsten vollkommen analog zu dem Beweis von Korollar 5.12.3. □

5.13. **Diagonalmatrizen und obere Dreiecksmatrizen.**

Definition 5.13.1. Eine quadratische Matrix $A = (a_{ij}) \in K^{n\times n}$ heißt genau dann **Diagonalmatrix**, wenn all ihre Einträge außerhalb der Diagonalen Nullen sind, wenn also für alle $i, j \in \{1, \dots, n\}$ gilt: Aus $a_{ij} \neq 0$ folgt $i = j$.

5.13.2. Eine Diagonalmatrix A ist genau dann invertierbar, wenn all ihre Diagonaleinträge $a_{11}, a_{22}, \dots, a_{nn}$ von Null verschieden sind. In diesem Fall ist ihr Inverses die Diagonalmatrix mit Diagonaleinträgen $a_{11}^{-1}, a_{22}^{-1}, \dots, a_{nn}^{-1}$. Wir folgern sofort, dass die Menge aller invertierbaren Diagonalmatrizen eine Untergruppe der $\mathrm{GL}_n(K)$ ist.

Definition 5.13.3. Eine **obere Dreiecksmatrix** ist eine quadratische Matrix $A = (a_{ij}) \in K^{n\times n}$, deren Einträge echt unterhalb der Diagonalen Null sind (d. h. für alle $i, j \in \{1, \dots, n\}$ folgt aus $a_{ij} \neq 0$ bereits $i \leq j$), also eine Matrix der Form

$$\begin{pmatrix} a_{11} & * & \dots & \dots & * \\ 0 & a_{22} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n-1} & * \\ 0 & \dots & \dots & 0 & a_{nn} \end{pmatrix}.$$

Eine **untere Dreiecksmatrix** ist eine quadratische Matrix $A \in K^{n\times n}$, deren Einträge echt oberhalb der Diagonalen Null sind, für die also A^{t} eine obere Dreiecksmatrix ist.

5.13.4. Eine obere Dreiecksmatrix ist genau dann invertierbar, wenn all ihre Diagonaleinträge $a_{11}, a_{22}, \dots, a_{nn}$ in $K^\times = K \setminus \{0\}$ liegen: Sind alle Diagonaleinträge in $K^\times$, so ist leicht zu sehen, dass die Spalten linear unabhängig sind; ist ein Diagonaleintrag Null, sagen wir $a_{ii} = 0$ mit i minimal, so ist die i-te Spalte eine Linearkombination der ersten $i - 1$ Spalten, und somit sind die Spalten linear abhängig.

Aufgabe 5.13.5. Die Menge

$$\begin{aligned} \mathcal{B} := \mathcal{B}_n &:= \{\text{invertierbare obere Dreiecksmatrizen in } K^{n\times n}\} \\ &= \left\{ \begin{pmatrix} d_1 & * & \dots & \dots & * \\ 0 & d_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & d_{n-1} & * \\ 0 & \dots & \dots & 0 & d_n \end{pmatrix} \in K^{n\times n} \mid d_1, \dots, d_n \in K^\times \right\} \end{aligned}$$

ist eine Untergruppe von $\mathrm{GL}_n(K)$.[35] Die Teilmenge $\mathcal{U} = \mathcal{U}_n$ aller oberen Dreiecksmatrizen mit Einsen auf der Diagonalen ist eine Untergruppe von $\mathcal{B}$.[36] Analoge Aussagen gelten für untere Dreiecksmatrizen.

[35] Sie ist ein Beispiel einer **Borel-Untergruppe** (zumindest falls K ein algebraisch abgeschlossener Körper ist), was die Wahl des Buchstabens $\mathcal{B}$ erklärt.

[36] Die Wahl des Buchstabens $\mathcal{U}$ kommt daher, dass alle Elemente von $\mathcal{U}$ *unipotent* sind.

Hinweis: Verwende entweder Lemma 5.6.7 oder gehe wie folgt vor: Sei $F_r := K^r \times \{0\}^{n-r} \subset K^n$ der von den ersten r Standardbasisvektoren aufgespannte Untervektorraum. Dann gilt $\mathcal{B}_n = \{g \in \mathrm{GL}_n(K) \mid g(F_r) = F_r \text{ für alle } r = 0, 1, \ldots, n\}$.

5.14. **Permutationsmatrizen.**

Definition 5.14.1. Sei $\pi \in S_n$ eine Permutation. Die **Permutationsmatrix** zu π ist die Matrix $P_\pi \in K^{n\times n}$, deren j-te Spalte der Standardbasisvektor $e_{\pi(j)}$ ist, für $j = 1, \ldots, n$, d. h. symbolisch

$$P_\pi = (e_{\pi(1)} \mid \cdots \mid e_{\pi(n)})$$

oder als Formel

$$(P_\pi)_{ij} = \delta_{i,\pi(j)} = \begin{cases} 1 & \text{falls } i = \pi(j), \\ 0 & \text{sonst.} \end{cases}$$

Die auf diese Weise erhaltenen Matrizen heißen **Permutationsmatrizen**.

Beispiel 5.14.2. (a) Elementarmatrizen vom Typ II sind Permutationsmatrizen zu Transpositionen: Es gilt $P_{ij} = P_{(ij)}$.

(b) Die Permutationsmatrix zu $\pi = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{bmatrix} \in S_4$ ist

$$P_\pi = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

5.14.3. Eine Matrix ist genau dann eine Permutationsmatrix (zu einer eindeutig bestimmten Permuation), wenn sie in jeder Zeile und in jeder Spalte genau eine Eins enthält, und sonst aus Nullen besteht.

Lemma 5.14.4. *Für jedes $\pi \in S_n$ gilt $P_\pi \in \mathrm{GL}_n(K)$, und die Abbildung*

$$S_n \to \mathrm{GL}_n(K),$$
$$\pi \mapsto P_\pi,$$

ist ein injektiver Gruppenhomomorphismus. Insbesondere gilt $(P_\pi)^{-1} = P_{\pi^{-1}}$.

Beweis. Die Spalten von P_π bilden eine Basis des K^n, denn bis auf ihre Reihenfolge stimmen sie mit den Elementen der Standardbasis $e_1, \ldots, e_n$ überein. Deshalb ist P_π invertierbar, also in $\mathrm{GL}_n(K)$ (siehe Satz 5.6.5).

Per Definition ist $P_\pi \colon K^n \to K^n$ eindeutig dadurch festgelegt, dass $P_\pi e_j = e_{\pi(j)}$ für alle $j \in \{1, \ldots, n\}$ gilt. Für $\pi, \sigma \in S_n$ und beliebiges $j \in \{1, \ldots, n\}$ berechnen wir

$$P_\pi(P_\sigma(e_j)) = P_\pi(e_{\sigma(j)}) = e_{\pi(\sigma(j))} = P_{\pi\circ\sigma}(e_j).$$

Somit gilt $P_\pi \circ P_\sigma = P_{\pi\circ\sigma}$. Also ist die angegebene Abbildung ein Gruppenhomomorphismus.

Sie ist injektiv, denn es gilt $P_\pi(e_j) = e_{\pi(j)}$ für alle $j \in \{1, \ldots, n\}$ bzw. ausführlich: Gelte $P_\pi = P_\sigma$ für beliebige $\pi, \sigma \in S_n$. Dann folgt $e_{\pi(j)} = P_\pi(e_j) = P_\sigma(e_j) = e_{\sigma(j)}$ und damit $\pi(j) = \sigma(j)$ für alle $j \in \{1, \ldots, n\}$, also $\pi = \sigma$. □

6. Determinanten

Seien K ein Körper und $n \in \mathbb{N}$ eine natürliche Zahl.

6.1. **Definition der Determinante.**

Definition 6.1.1. Die **Determinante** ist die Abbildung

$$\det\colon K^{n\times n} \to K,$$

die eine quadratische Matrix $A = (a_{ij}) \in K^{n\times n}$ auf

$$\det(A) := \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \tag{6.1.1}$$

abbildet. Dieser Ausdruck für die Determinante heißt **Leibniz-Formel**.

Beispiel 6.1.2. Wir schreiben die Leibniz-Formel in den Fällen $n = 1, 2, 3$ explizit aus:

$$\det(a) = a$$

$$\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ad - bc$$

$$\begin{aligned}\det\left(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\right) &= +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}\end{aligned}$$

Die letzte Formel heißt **Regel von Sarrus** oder weidmannssprachlich **Jägerzaunformel**, siehe Illustration in der Vorlesung. Die naive Verallgemeinerung dieser Regel für $n \geq 4$ ist falsch (was man schon allein deshalb vermuten sollte, dass in ihr nur $2n$ statt $n!$ Summanden auftauchen).

6.1.3. In der Praxis ist die Leibniz-Formel im Allgemeinen (also für große Matrizen) keine sinnvolle Methode, um die Determinante auszurechnen, jedoch ist sie theoretisch nützlich.

Anschauung 6.1.4. Eine gute geometrische Anschauung für die Determinante $\det(A)$ einer quadratischen Matrix A mit Einträgen in einem beliebigen Körper habe ich nicht. Hat A aber reelle Einträge, so gibt es eine gute Anschauung für deren Betrag $|\det(A)|$: Diese reelle Zahl ist das Volumen des von den Spalten von A aufgespannten Parallelepipeds (später wird das zur Definition erklärt, siehe Definition 8.4.12). Wir erklären dies für reelle (2×2)-Matrizen.

Sind $v, w \in \mathbb{R}^2$ zwei Vektoren, so bezeichen wir mit $\operatorname{area}(v, w)$ die Fläche (= das zweidimensionale Volumen) des Parallelogramms in der „Zeichenebene“ $\mathbb{R}^2$, das von den Vektoren v und w aufgespannt wird. Elementargeometrisch sind die folgenden Eigenschaften der Fläche plausibel (siehe Zeichnungen in der Vorlesung):

- Invarianz unter Scherungen: $\operatorname{area}(v + \lambda w, w) = \operatorname{area}(v, w) = \operatorname{area}(v, w + \lambda v)$ für alle $v, w \in \mathbb{R}^2$ und alle $\lambda \in \mathbb{R}$.
- Verhalten unter Skalierungen: $\operatorname{area}(\lambda v, w) = |\lambda| \operatorname{area}(v, w) = \operatorname{area}(v, \lambda w)$ für alle $v, w \in \mathbb{R}^2$ und alle $\lambda \in \mathbb{R}$.
- Normierung: $\operatorname{area}(e_1, e_2) = 1$.

Seien Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \in \mathbb{R}^2$ gegeben. Wir wollen die Gleichung

$$\operatorname{area}\left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) = \left|\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\right|$$

plausibel machen. Mit Hilfe der obigen Eigenschaften berechnen wir unter der Annahme $d \neq 0$:

$$\begin{aligned}
\operatorname{area}\left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) &= \operatorname{area}\left(\begin{pmatrix} a \\ c \end{pmatrix} - \frac{c}{d}\begin{pmatrix} b \\ d \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) \\
&= \operatorname{area}\left(\begin{pmatrix} \frac{ad-bc}{d} \\ 0 \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) \\
&= \frac{|ad-bc|}{|d|} \operatorname{area}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) \\
&= \frac{|ad-bc|}{|d|} \operatorname{area}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} - b\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \\
&= \frac{|ad-bc|}{|d|} \operatorname{area}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ d \end{pmatrix}\right) \\
&= |ad-bc| \operatorname{area}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
&= |ad-bc| \\
&= \left|\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\right|.
\end{aligned}$$

Den Fall $d = 0$ erledigt man in ähnlicher Weise (mit einer weiteren Fallunterscheidung $b = 0$ bzw. $b \neq 0$) unter Beachtung von $\operatorname{area}(v, 0) = 0$ (wegen des Skalierungsverhaltens) und $\operatorname{area}(v, w) = \operatorname{area}(w, v)$ (wegen $\operatorname{area}(v, w) = \operatorname{area}(v, w - v) = \operatorname{area}(v + (w - v), w - v) = \operatorname{area}(w, w - v) = \operatorname{area}(w, w - v - w) = \operatorname{area}(w, -v) = \operatorname{area}(w, v)$).

Ende 18. Vorlesung am 25.06.2020

Lemma 6.1.5. *Ist $A \in K^{n\times n}$ eine obere oder untere Dreiecksmatrix, so ist $\det(A)$ das Produkt der Diagonaleinträge, in Formeln*

$$\det(A) = a_{11}a_{22}\cdots a_{nn}.$$

Insbesondere gilt dies für Diagonalmatrizen und noch spezieller gilt $\det(I_n) = 1$.

Beweis. Sei $A = (a_{ij})$ eine obere Dreiecksmatrix. Es genügt zu zeigen, dass in der Leibniz-Formel (6.1.1) alle Summanden für $\sigma \neq \mathrm{id}_{\{1,\dots,n\}}$ verschwinden.

Sei $\sigma \in S_n$ mit $a_{1\sigma(1)}\cdots a_{n\sigma(n)} \neq 0$. Dann gilt $a_{i\sigma(i)} \neq 0$ für alle Faktoren dieses Produkts, also $i \leq \sigma(i)$ für alle $i = 1, \dots, n$. Der Reihe nach folgert man $\sigma(n) = n$, $\sigma(n-1) = n-1$, $\dots$, $\sigma(2) = 2$, $\sigma(1) = 1$; es gilt also $\sigma = \mathrm{id}$.

Für untere Dreiecksmatrizen argumentiert man analog (aus $\sigma(i) \leq i$ für alle $i = 1, \dots, n$ folgert man der Reihe nach $\sigma(1) = 1$, $\sigma(2) = 2$, $\dots$, $\sigma(n) = n$) oder verwendet Lemma 6.1.8 unten. □

Beispiel 6.1.6. Lemma 6.1.5 erlaubt die schnelle Berechnung der Determinante aller Elementarmatrizen vom Typ I oder III (siehe Definition 5.9.2):

$$\det(T_{ij}(\lambda)) = 1,$$
$$\det(D_i(\mu)) = \mu.$$

Beispiel 6.1.7 (Determinante von Permutationsmatrizen)**.** Die Determinante der Permutationsmatix P_π, für $\pi \in S_n$, ist das Vorzeichen von π, denn

$$\begin{aligned}
\det(P_\pi) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (P_\pi)_{i\sigma(i)} \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \underbrace{\prod_{i=1}^{n} \delta_{i,\pi(\sigma(i))}}_{\neq 0 \iff \forall i\colon\, i=\pi(\sigma(i)) \iff \mathrm{id}=\pi\circ\sigma \iff \pi^{-1}=\sigma} \\
&= \operatorname{sgn}(\pi^{-1}) \\
&= \operatorname{sgn}(\pi)^{-1} \\
&= \operatorname{sgn}(\pi),
\end{aligned}$$

wobei die vorletzte Gleichheit verwendet, dass $\operatorname{sgn}\colon S_n \to \{\pm 1\}$ ein Gruppenhomomorphismus ist (Satz 1.4.17), und die letzte, dass $(-1)^{-1} = -1$ und $1^{-1} = 1$.

Insbesondere hat jede Elementarmatrix vom Typ II Determinante -1, in Formeln

$$\det(P_{ij}) = \det(P_{(ij)}) = \operatorname{sgn}((ij)) = -1,$$

denn die Anzahl der Fehlstände der Permutation (ij) ist ungerade.

Lemma 6.1.8. *Die Determinante ändert sich nicht beim Transponieren einer Matrix, in Formeln*

$$\det(A) = \det(A^{\mathrm{t}}).$$

Beweis. Wir erinnern daran, dass $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ für alle $\sigma \in S_n$ gilt (siehe Beispiel 6.1.7). Damit berechnen wir

$$\det(A^{\mathrm{t}}) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (A^{\mathrm{t}})_{i\sigma(i)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{\substack{i=1,\dots,n,\\ j=1,\dots,n,\\ \sigma^{-1}(j)=i}} a_{ji}$$

$$= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{j=1}^{n} a_{j\sigma^{-1}(j)} = \sum_{\tau\in S_n} \operatorname{sgn}(\tau) \prod_{j=1}^{n} a_{j\tau(j)} = \det(A). \qquad \square$$

6.1.9. Erinnert man sich nicht, ob in der Leibniz-Formel (6.1.1) $a_{i\sigma(i)}$ oder $a_{\sigma(i)i}$ steht, so ändert dies nichts am Ergebnis: Es gilt

$$\det(A) \overset{(6.1.1)}{=} \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

nach (dem Beweis von) Lemma 6.1.8.

6.2. **Charakterisierung der Determinante.**

Definition 6.2.1. Eine Abbildung

$$D\colon \underbrace{K^n \times \cdots \times K^n}_{n \text{ Faktoren}} \to K$$

heißt genau dann **Determinantenfunktion**, wenn sie die folgenden beiden Bedingungen erfüllt:

(a) Sie ist **multilinear**, d. h. linear in jedem Argument, wenn die anderen Argumente fixiert sind: Für alle $i \in \{1, \dots, n\}$ und alle $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n \in K^n$ ist die Abbildung

$$D(v_1, \dots, v_{i-1}, ?, v_{i+1}, \dots, v_n)\colon K^n \to K, \quad x \mapsto D(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_n)$$

K-linear.

(b) Sie ist **alternierend**, d. h. sie ist Null, wenn zwei Argumente übereinstimmen: Gegeben beliebige Vektoren $v_1, \dots, v_n \in K^n$, für die $v_i = v_j$ für zwei Indizes $i, j \in \{1, \dots, n\}$ mit $i \neq j$ gilt, so gilt

$$D(v_1, \dots, v_n) = 0.$$

Eine solche Determinantenfunktion heisst **normiert**, falls sie auf der Standardbasis den Wert Eins annimmt, in Formeln

$$D(e_1, e_2, \dots, e_n) = 1.$$

Lemma 6.2.2. *Sei $D\colon K^n \times \cdots \times K^n \to K$ eine Determinantenfunktion. Vertauscht man zwei Argumente von D, so ändert sich das Vorzeichen*[37]*: Seien $v_1, \dots, v_n \in K^n$ und $1 \leq i < j \leq n$. Dann gilt*

$$D(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n) = -D(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n). \tag{6.2.1}$$

Insbesondere gilt

$$D(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \operatorname{sgn}(\sigma) D(v_1, \dots, v_n) \tag{6.2.2}$$

für alle $\sigma \in S_n$.

Beweis. Wir definieren $D'(x, y) := D(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_{j-1}, y, v_{j+1}, \dots, v_n)$. Für (6.2.1) ist $D'(x, y) = -D'(y, x)$ zu zeigen. Da D alternierend und multilinear ist, gilt

$$\begin{aligned} 0 = D'(x+y, x+y) &= D'(x, x+y) + D'(y, x+y) \\ &= D'(x, x) + D'(x, y) + D'(y, x) + D'(y, y) = D'(x, y) + D'(y, x), \end{aligned}$$

also $D'(x, y) = -D'(y, x)$. Dies zeigt (6.2.1).

Sei $\sigma \in S_n$. Schreibe $\sigma = \tau_1 \circ \ldots \circ \tau_r$ als Produkt einfacher Transpositionen $\tau_1, \dots, \tau_r$ (eine solche Darstellung existiert nach Satz 1.4.13), und es gilt $\operatorname{sgn}(\sigma) = (-1)^r$ nach Satz 1.4.17. Indem wir (6.2.1) r-mal anwenden, erhalten wir

$$D(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = D(v_{\tau_1(\dots(\tau_r(1))\dots)}, \dots, v_{\tau_1(\dots(\tau_r(n))\dots)})$$

[37] Dies erklärt den Begriff *alternierend*.

$$
\begin{aligned}
&= -D(v_{\tau_2(\dots(\tau_r(1))\dots)}, \dots, v_{\tau_2(\dots(\tau_r(n))\dots)}) \\
&= \dots \\
&= (-1)^{r-1} D(v_{\tau_r(1)}, \dots, v_{\tau_r(n)}) \\
&= (-1)^r D(v_1, \dots, v_n) \\
&= \operatorname{sgn}(\sigma) D(v_1, \dots, v_n). \qquad \square
\end{aligned}
$$

6.2.3. Da eine $(n \times n)$-Matrix nichts anderes ist als ein n-Tupel von Spaltenvektoren, können wir die in Definition 6.1.1 per Leibniz-Formel definierte Determinante $\det\colon K^{n\times n} \to K$ auch als Abbildung

$$
\begin{aligned}
\det\colon K^n \times \dots \times K^n &\to K, \\
(v_1, \dots, v_n) &\mapsto \det(v_1, v_2, \dots, v_n) := \det\big((v_1 \mid v_2 \mid \dots \mid v_n)\big),
\end{aligned} \tag{6.2.3}
$$

auffassen. Hier bezeichnet $(v_1 \mid \dots \mid v_n)$ die Matrix, deren Spalten die Vektoren $v_1, \dots, v_n$ sind.

Satz 6.2.4. *Es gibt genau eine normierte Determinantenfunktion $K^n \times \dots \times K^n \to K$, nämlich die durch die Leibniz-Formel definierte Determinante* $\det\colon K^n \times \dots \times K^n \to K$ *(siehe Definition 6.1.1).*

Ist genauer $D\colon K^n \times \dots \times K^n \to K$ eine (nicht notwendig normierte) Determinantenfunktion, so stimmt sie bis auf den Faktor $D(e_1, \dots, e_n) \in K$ mit der Determinante det *überein, es gilt also*

$$
D(v_1, \dots, v_n) = D(e_1, \dots, e_n) \det(v_1, \dots, v_n) \tag{6.2.4}
$$

für alle $v_1, \dots, v_n \in K^n$.

6.2.5. In anderen Worten ist die Abbildung $\det\colon K^{n\times n} \to K$ eindeutig durch die folgenden Eigenschaften bestimmt:

- det ist linear in jeder Spalte, wenn die anderen Spalten fixiert sind,
- $\det(A) = 0$, falls zwei Spalten von A übereinstimmen,
- $\det(I_n) = 1$.

Beweis. Sei $D\colon K^n \times \dots \times K^n \to K$ eine Determinantenfunktion. Seien $v_1, \dots, v_n \in K^n$ und $A = (a_{ij}) := (v_1 \mid \dots \mid v_n)$, also $v_j = \sum_{i=1}^n a_{ij} e_i$ für alle $j \in \{1, \dots, n\}$. Wir berechnen

$$
\begin{aligned}
D(v_1, \dots, v_n) &= D\left(\sum_{i_1=1}^n a_{i_1 1} e_{i_1}, \dots, \sum_{i_j=1}^n a_{i_j j} e_{i_j}, \dots, \sum_{i_n=1}^n a_{i_n n} e_{i_n}\right) \\
&= \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} \dots a_{i_n n} D(e_{i_1}, \dots, e_{i_n}) && \text{(da } D \text{ multilinear)} \\
&= \sum_{\sigma \in S_n} D(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \prod_{j=1}^n a_{\sigma(j) j} && \text{(da } D \text{ alternierend)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) D(e_1, \dots, e_n) \prod_{j=1}^n a_{\sigma(j) j} && \text{(nach (6.2.2) in Lemma 6.2.2)}
\end{aligned}
$$

$$
\begin{aligned}
&= D(e_1,\dots,e_n)\Big(\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{j=1}^{n} a_{\sigma(j)j}\Big)\\
&= D(e_1,\dots,e_n)\det(A^{\mathrm{t}})\\
&= D(e_1,\dots,e_n)\det(A) && \text{(nach Lemma 6.1.8)}\\
&= D(e_1,\dots,e_n)\det(v_1,\dots,v_n).
\end{aligned}
$$

Dies zeigt (6.2.4).

Wir folgern die Eindeutigkeit einer normierten Determinantenfunktion: Ist D eine normierte Determinantenfunktion, so gilt $D = \det$.

Es bleibt die Existenz einer normierten Determinantenfunktion zu zeigen: Nach unseren bisherigen Überlegungen ist det der einzige Kandidat, und wir müssen nachrechnen, dass det eine normierte Determinantenfunktion ist. Seien $v_1,\dots,v_n$ und $A = (v_1 \mid \cdots \mid v_n)$ wie oben. Schreiben wir $v_j = \begin{pmatrix}(v_j)_1\\ \vdots \\ (v_j)_n\end{pmatrix}$, so gilt $(v_j)_i = a_{ij}$ und wir erhalten mit Lemma 6.1.8

$$
\det(v_1,\dots,v_n) = \det(A) = \det(A^{\mathrm{t}}) = \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{i=1}^{n}(v_i)_{\sigma(i)}.
$$

Für fixiertes $i' \in \{1,\dots,n\}$ können wir diesen Ausdruck zu

$$
\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)(v_{i'})_{\sigma(i')}\Big(\prod_{\substack{i=1\\ i\neq i'}}^{n}(v_i)_{\sigma(i)}\Big)
$$

umschreiben. Dieser ist offensichtlich linear in $v_{i'} \in K^n$, wenn wir die anderen Argumente v_i mit $i \neq i'$ fixiert lassen. Dies zeigt, dass det multilinear ist.

Seien nun $1 \leq i' < i'' \leq n$ gegeben mit $v_{i'} = v_{i''}$ Dann gilt

$$
\det(v_1,\dots,v_n) = \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)(v_{i'})_{\sigma(i')}(v_{i''})_{\sigma(i'')}\Big(\prod_{\substack{i=1\\ i\neq i'\\ i\neq i''}}^{n}(v_i)_{\sigma(i)}\Big). \tag{6.2.5}
$$

Sei $\tau := (i'i'') \in S_n$ die Transposition. Es gilt $\operatorname{sgn}(\tau) = -1$.

Sei $A_n := \ker(\operatorname{sgn}) = \operatorname{sgn}^{-1}(\{1\})$. Für jedes Element $\sigma \in S_n$ mit $\operatorname{sgn}(\sigma) = -1$ gibt es genau ein Element $\pi \in A_n$ mit $\sigma = \pi\tau$, nämlich $\pi := \sigma\tau^{-1} = \sigma\tau$.

Wir setzen die Rechnung (6.2.5) fort zu

$$
\begin{aligned}
\det(v_1,\dots,v_n) = &\sum_{\pi\in A_n}\operatorname{sgn}(\pi)(v_{i'})_{\pi(i')}(v_{i''})_{\pi(i'')}\Big(\prod_{\substack{i=1\\ i\neq i'\\ i\neq i''}}^{n}(v_i)_{\pi(i)}\Big)\\
&+\sum_{\pi\in A_n}\operatorname{sgn}(\pi\tau)(v_{i'})_{\pi(\tau(i'))}(v_{i''})_{\pi(\tau(i''))}\Big(\prod_{\substack{i=1\\ i\neq i'\\ i\neq i''}}^{n}(v_i)_{\pi(\tau(i))}\Big).
\end{aligned}
$$

Diese Summe ist aber Null: Betrachte den zweiten Summanden und verwende

- $\operatorname{sgn}(\pi\tau) = -\operatorname{sgn}(\pi)$,
- $\pi(\tau(i')) = \pi(i'')$ und $\pi(\tau(i'')) = \pi(i')$,
- $v_{i'} = v_{i''}$ nach Annahme,
- für alle $i \in \{1, \dots, n\}$ mit $i \neq i'$ und $i \neq i''$ gilt $\pi(\tau(i)) = \pi(i)$.

Dies zeigt, dass det alternierend ist.

Wegen $\det(e_1, \dots, e_n) = \det(I_n) = 1$ (siehe Lemma 6.1.5) ist det normiert. □

Satz 6.2.6 (Determinante verschwindet bei linearer Abhängigkeit)**.** *Sei $A \in K^{n\times n}$ eine quadratische Matrix mit linear abhängigen Spalten oder linear abhängigen Zeilen. Dann gilt*[38] $\det(A) = 0$.

Beweis. Wegen $\det(A) = \det(A^{\mathrm{t}})$ (Lemma 6.1.8) genügt es, die Aussage im Fall linear abhängiger Spalten zu beweisen.

Seien $v_1, \dots, v_n$ die linear abhängigen Spalten von A. Dies bedeutet, dass es Elemente $c_1, \dots, c_n \in K$ mit $c_i \neq 0$ für ein i gibt, so dass

$$0 = \sum_{j=1}^{n} c_j v_j$$

gilt. Indem wir $c_i v_i$ auf die linke Seite bringen und mit $-c_i^{-1}$ multiplizieren, erhalten wir die Darstellung

$$v_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} b_j v_j$$

für geeignete $b_j \in K$. Da det multilinear und alternierend ist erhalten wir

$$\begin{aligned}
\det(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n) &= \det(v_1, \dots, v_{i-1}, \sum_{\substack{j=1 \\ j\neq i}}^{n} b_j v_j, v_{i+1}, \dots, v_n) \\
&= \sum_{\substack{j=1 \\ j\neq i}}^{n} b_j \det(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_n) \\
&= 0.
\end{aligned}$$

□

Satz 6.2.7 (Rechenregeln für die Determinante)**.** *Es gelten die folgenden Rechenregeln:*

(a) Die Determinante ändert sich nicht, wenn man ein Vielfaches einer Spalte zu einer anderen Spalte addiert (elementare Spaltenoperation vom Typ I):[39]

$$\det(\dots, v_i, \dots, v_j + \lambda v_i, \dots) = \det(\dots, v_i, \dots, v_j, \dots)$$

(b) Die Determinante ändert sich um den Faktor -1, wenn man zwei Spalten miteinander vertauscht (elementare Spaltenoperation vom Typ II):

$$\det(\dots, v_i, \dots, v_j, \dots) = -\det(\dots, v_j, \dots, v_i, \dots)$$

[38] Der Beweis zeigt die entsprechende Aussage allgemeiner für jede Determinantenfunktion.

[39] Nicht aufgeführte Argumente sind fixiert.

(c) Die Determinante ändert sich um den Faktor λ, wenn man eine Spalte mit $\lambda \in K$ multipliziert (elementare Spaltenoperation vom Typ III, falls $\lambda \neq 0$):

$$\det(\dots, \lambda v_i, \dots) = \lambda \det(\dots, v_i, \dots)$$

Dieselben Regeln gelten auch, wenn man „Spalte“ durch „Zeile“ ersetzt.

Beweis. Wegen $\det(A) = \det(A^{\mathrm{t}})$ (Lemma 6.1.8) gelten die Regeln genau dann für Spaltenoperationen, wenn sie für Zeilenoperationen gelten. Wir beweisen sie für Spaltenoperationen.

Die erste Regel folgt daraus, dass die Determinante multilinear und alternierend ist (alle nicht angegebenen Argumente sind fixiert; gelte ohne Einschränkung $i < j$):

$$\begin{aligned}\det(\dots, v_i, \dots, v_j + \lambda v_i, \dots) &= \det(\dots, v_i, \dots, v_j, \dots) + \lambda \det(\dots, v_i, \dots, v_i, \dots) \\ &= \det(\dots, v_i, \dots, v_j, \dots)\end{aligned}$$

Die zweite Regel haben wir bereits in Lemma 6.2.2 für beliebige Determinantenfunktionen bewiesen, und nach Satz 6.2.4 ist det eine solche. Die dritte Regel folgt direkt aus der Multilinearität. □

Bemerkung 6.2.8 (Berechnung der Determinante)**.** Um die Determinante einer (hinreichend komplizierten) quadratischen Matrix A zu berechnen, ist es sinnvoll, sie per geeigneter elementarer Zeilen- und Spaltenoperationen auf obere oder untere Dreiecksform zu bringen. Die Rechenregeln aus Satz 6.2.7 sagen uns, wie sich die Determinante dabei verändert; wenn man sich dabei auf elementare Operationen vom Typ I und II beschränkt, ändert sie sich um den Faktor $(-1)^{\text{Anzahl Operationen vom Typ II}}$. Die Determinante einer oberen (oder unteren) Dreiecksmatrix ist einfach zu berechnen: Sie ist das Produkt der Diagonaleinträge (siehe Lemma 6.1.5). Stellt man unterdessen fest, dass die Zeilen oder Spalten linear abhängig sind, so ist die Determinante Null (Satz 6.2.6). Hilfreich ist auch Korollar 6.5.2 weiter hinten im Text, das erklärt, wie man die Determinante von Oberen-Dreiecks-Blockmatrizen berechnet.

Ende 19. Vorlesung am 30.06.2020

6.3. Multiplikativität der Determinante.

Satz 6.3.1. *Für alle Matrizen $A, B \in K^{n\times n}$ gilt*

$$\det(AB) = \det(A)\det(B).$$

6.3.2. Offensichtlich impliziert dies $\det(AB) = \det(BA)$.

Beweis. Fixiere $A \in K^{n\times n}$ und betrachte die Abbildung

$$\begin{aligned}D\colon K^n \times \cdots \times K^n &\to K, \\ (v_1, \dots, v_n) &\mapsto \det(Av_1, \dots, Av_n).\end{aligned}$$

Diese Abbildung ist offensichtlich multilinear und alternierend, also eine Determinantenfunktion. Nach Satz 6.2.4 gilt

$$\begin{aligned}D(v_1, \dots, v_n) &= D(e_1, \dots, e_n)\det(v_1, \dots, v_n) \\ &= \det(Ae_1, \dots, Ae_n)\det(v_1, \dots, v_n) = \det(A)\det(v_1, \dots, v_n)\end{aligned}$$

für alle $v_1, \dots, v_n \in K^n$. Nehmen wir speziell $v_j := Be_j$ die j-te Spalte von B, so gelten

$$\det(v_1, \dots, v_n) = \det(Be_1, \dots, Be_n) = \det(B) \qquad \text{und}$$
$$D(v_1, \dots, v_n) = \det(Av_1, \dots, Av_n) = \det(ABe_1, \dots, ABe_n) = \det(AB).$$

Setzen wir dies oben ein, erhalten wir die Behauptung. □

Korollar 6.3.3. *Ähnliche Matrizen haben dieselbe Determinente: Seien $A, B \in K^{n\times n}$ mit $B = TAT^{-1}$ für ein $T \in \mathrm{GL}_n(K)$. Dann gilt*

$$\det(A) = \det(B).$$

Beweis. Die Multiplikativität der Determinante (siehe Satz 6.3.1) liefert

$$\begin{aligned}\det(B) = \det(TAT^{-1}) = \det(T)\det(A)\det(T^{-1}) = \det(T)\det(T^{-1})\det(A)\\ = \det(TT^{-1})\det(A) = \det(I_n)\det(A) = \det(A).\end{aligned}$$ □

Bemerkung 6.3.4. Aus Satz 6.3.1 kann man alle Rechenregeln in Satz 6.2.7 noch einmal herleiten:

(a) $\det(T_{ij}(\lambda)A) = \det(A) = \det(AT_{ij}(\lambda))$;
(b) $\det(P_{ij}A) = -\det(A) = \det(AP_{ij})$.
(c) $\det(D_j(\mu)A) = \mu\det(A) = \det(AD_j(\mu))$ (dies gilt auch für $\mu = 0$, wenn man $D_i(0)$ in offensichtlichter Weise definiert);

6.4. Determinantenkriterium für Invertierbarkeit.

Satz 6.4.1 (Determinantenkriterium für Invertierbarkeit einer Matrix)**.** *Eine quadratische Matrix $A \in K^{n\times n}$ ist genau dann invertierbar (also ein Element von $\mathrm{GL}_n(K)$), wenn $\det(A) \neq 0$ gilt.*

Anschauung 6.4.2. Im Fall $K = \mathbb{R}$ ist dieser Satz recht anschaulich, wenn man glaubt, dass $|\det(A)|$ das Volumen des von den Spalten von A aufgespannten Parallelepipeds ist (vgl. Anschauung 6.1.4). Sind die Spalten von A linear unabhängig, d. h. gilt $A \in \mathrm{GL}_n(\mathbb{R})$, so hat das aufgespannte Parallelepiped (als „durch Scherungen verformter“ Quader positiven Volumens) positives Volumen, und somit gilt $\det(A) \neq 0$. Sind die Spalten von A linear abhängig, d. h. gilt $A \notin \mathrm{GL}_n(\mathbb{R})$, so liegt das aufgespannte Parallelepiped in einem echten Unterraum des $\mathbb{R}^n$, und somit ist sein n-dimensionales Volumen Null, es gilt also $\det(A) = 0$.

Beweis. Aus $A \in \mathrm{GL}_n(K)$ folgt $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. Also ist $\det(A)$ eine Einheit in K, d. h. es gilt $\det(A) \neq 0$.

Sei $A \notin \mathrm{GL}_n(K)$. Dann sind die Spalten von A linear abhängig (siehe Satz 5.6.5), so dass $\det(A) = 0$ nach Satz 6.2.6 gilt. □

Definition 6.4.3. Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums. Dann ist die **Determinante von** f durch

$$\det(f) := \det({}_{\mathcal{B}}[f]_{\mathcal{B}})$$

definiert, wobei $\mathcal{B}$ eine beliebige Basis von V und ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ die darstellende Matrix von f bezüglich dieser Basis sind. Nach Satz 5.12.10 und Korollar 6.3.3 hängt $\det(f)$ nicht von der Wahl der Basis ab.

Korollar 6.4.4 (Determinantenkriterium für Invertierbarkeit eines Endomorphismus)**.** *Ein Endomorphismus $f\colon V \to V$ eines endlichdimensionalen Vektorraums ist genau dann invertierbar, wenn $\det(f) \neq 0$ gilt.*

Beweis. Sei $\mathcal{B}$ eine Basis von V. Nach 5.10.7 ist f genau dann invertierbar, wenn ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ invertierbar ist. Nach Satz 6.4.1 ist das genau dann der Fall, wenn $\det(f) = \det({}_{\mathcal{B}}[f]_{\mathcal{B}}) \neq 0$ gilt. □

Satz 6.4.5. *Die Determinante*

$$\det\colon \mathrm{GL}_n(K) \to K^\times = K \setminus \{0\}$$

ist ein Gruppenhomomorphismus. Insbesondere gilt $\det(A^{-1}) = \det(A)^{-1}$ *für alle* $A \in \mathrm{GL}_n(K)$.

Beweis. Nach Satz 6.4.1 ist die Abbildung wohldefiniert, und nach Satz 6.3.1 ist sie verträglich mit den Gruppenverknüpfungen Matrizenmultiplikation bzw. Multiplikation. □

6.5. Determinante von Oberen-Dreiecks-Blockmatrizen.

Proposition 6.5.1. *Sei*

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

eine ***Obere-Dreiecks-Blockmatrix****, d. h.* $B \in K^{r\times r}$ *und* $D \in K^{s\times s}$ *sind quadratische Matrizen, es gilt* $C \in K^{r\times s}$ *und* A *ist die angedeutete quadratische* $(r+s)\times(r+s)$*-Matrix. Dann gilt*

$$\det(A) = \det(B)\det(D).$$

Erster Beweis. Wir betrachten zunächst den Fall, dass B nicht invertierbar ist. Dies bedeutet, dass die Spalten von B linear abhängig sind. Dann sind die Spalten von A erst recht linear abhängig, es ist also A nicht invertierbar. Satz 6.4.1 liefert $\det(B) = 0 = \det(A)$, und die behauptete Gleichheit gilt.

Sei im Folgenden B invertierbar. Es gilt (in hoffentlich verständlicher Notation)

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} B & 0 \\ \hline 0 & I_s \end{array}\right) \cdot \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & D \end{array}\right) \cdot \left(\begin{array}{c|c} I_r & B^{-1}C \\ \hline 0 & I_s \end{array}\right).$$

Multiplikativität der Determinante (siehe Satz 6.3.1) zeigt, dass $\det(A)$ das Produkt der Determinanten der drei Faktoren ist. Der dritte Faktor ist eine obere Dreiecksmatrix mit Einsen auf der Diagonalen und hat deswegen Determinante Eins. Aus der Leibniz-Formel sieht man sofort, dass $\det(B)$ die Determinante des ersten Faktors ist (denn in der Summe in der Leibniz-Formel steuern nur die Permutationen $\sigma \in S_{r+s}$ etwas bei, die die Zahlen $r+1,\dots,r+s$ fixieren, die man also als Elemente von S_r auffassen kann). Analog ist $\det(D)$ die Determinante des zweiten Faktors. □

Zweiter Beweis. [40] Nach der Leibniz-Formel (6.1.1) gilt

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_{r+s}} \operatorname{sgn}(\sigma) \prod_{i=1}^{r+s} a_{i\sigma(i)} \\
&= \sum_{\substack{\sigma\in S_{r+s},\\ \sigma(\{r+1,\dots,r+s\})\subset\{r+1,\dots,r+s\}}} \operatorname{sgn}(\sigma) \prod_{i=1}^{r+s} a_{i\sigma(i)} \\
&= \sum_{\tau\in S_r,\rho\in S_s} \operatorname{sgn}(\tau)\operatorname{sgn}(\rho) \prod_{i=1}^{r} a_{i\tau(i)} \prod_{i=1}^{s} a_{r+i,r+\rho(i)}
\end{aligned}$$

[40] Dieser Beweis hat den Vorteil, dass er auch für Matrizen mit Einträgen in einem kommutativen Ring funktioniert, vgl. Abschitt 7.2.

$$= \sum_{\tau\in S_r}\sum_{\rho\in S_s} \operatorname{sgn}(\tau)\operatorname{sgn}(\rho)\prod_{i=1}^{r} b_{i\tau(i)}\prod_{i=1}^{s} d_{i,\rho(i)}$$
$$= \Big(\sum_{\tau\in S_r} \operatorname{sgn}(\tau)\prod_{i=1}^{r} b_{i\tau(i)}\Big)\Big(\sum_{\rho\in S_s} \operatorname{sgn}(\rho)\prod_{i=1}^{s} d_{i,\rho(i)}\Big)$$
$$= \det(B)\det(D).$$

□

Korollar 6.5.2. *Sei*

$$A = \left(\begin{array}{c|c|c|c} A_1 & B_{11} & \cdots & B_{1n} \\ \hline 0 & A_2 & \ddots & B_{2n} \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline 0 & \cdots & 0 & A_n \end{array}\right)$$

eine Obere-Dreiecks-Blockmatrix mit quadratischen Matrizen $A_i \in K^{r_i\times r_i}$ auf der „Blockdiagonalen", Nullmatrizen unterhalb und beliebigen Matrizen oberhalb der Blockdiagonalen. Dann gilt

$$\det(A) = \det(A_1)\det(A_2)\cdots\det(A_n).$$

6.5.3. Ein Spezialfall dieses Korollars ist, dass die Determinante einer oberen Dreicksmatrix das Produkt der Diagonaleinträge ist. Wir haben dies bereits in Lemma 6.1.5 gesehen.

Beweis. Dies folgt per Induktion sofort aus Proposition 6.5.1, indem man A in vier Blöcke einteilt. □

6.6. **Determinante der Vandermonde-Matrix.**

Satz 6.6.1. *Seien $x_1,\ldots,x_n \in K$ beliebig. Dann hat die sogenannte* ***Vandermonde-Matrix***

$$V = V(x_1,\ldots,x_n) := \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix} = (x_i^{j-1})_{i,j=1,\ldots,n} \in K^{n\times n}$$

die Determinante

$$\det(V) = \prod_{\substack{i,j\in\{1,\ldots,n\}\\ i<j}} (x_j - x_i).$$

Beispiel 6.6.2. Im Fall $n = 3$ gilt also $\det(V) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$. Der Leser überzeuge sich von Hand in den Fällen $n = 1, 2, 3$, dass die Formel für $\det(V)$ stimmt.

Beweis. Die Formel stimmt offensichtlich im Fall $n = 1$. Der allgemeine Beweis ist leicht ersichtlich aus dem im Folgenden angegebenen Beweis für $n = 5$. Zu berechnen ist also die

Determinante der Matrix

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 & x_1^4 \\ 1 & x_2 & x_2^2 & x_2^3 & x_2^4 \\ 1 & x_3 & x_3^2 & x_3^3 & x_3^4 \\ 1 & x_4 & x_4^2 & x_4^3 & x_4^4 \\ 1 & x_5 & x_5^2 & x_5^3 & x_5^4 \end{pmatrix}.$$

Wir ziehen die erste Zeile von allen anderen Zeilen ab und erhalten die Matrix

$$V' := \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 & x_1^4 \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & x_2^3 - x_1^3 & x_2^4 - x_1^4 \\ 0 & x_3 - x_1 & x_3^2 - x_1^2 & x_3^3 - x_1^3 & x_3^4 - x_1^4 \\ 0 & x_4 - x_1 & x_4^2 - x_1^2 & x_4^3 - x_1^3 & x_4^4 - x_1^4 \\ 0 & x_5 - x_1 & x_5^2 - x_1^2 & x_5^3 - x_1^3 & x_5^4 - x_1^4 \end{pmatrix}.$$

Mit Hilfe der allgemeinen Formel

$$a^N - b^N = (a-b)(a^{N-1} + a^{N-2}b + \cdots + ab^{N-2} + b^{N-1})$$

sehen wir, dass die Matrix V' aus der Matrix V''

$$V'' := \left(\begin{array}{c|cccc} 1 & x_1 & x_1^2 & x_1^3 & x_1^4 \\ \hline 0 & 1 & x_2 + x_1 & x_2^2 + x_2x_1 + x_1^2 & x_2^3 + x_2^2x_1 + x_2x_1^2 + x_1^3 \\ 0 & 1 & x_3 + x_1 & x_3^2 + x_3x_1 + x_1^2 & x_3^3 + x_3^2x_1 + x_3x_1^2 + x_1^3 \\ 0 & 1 & x_4 + x_1 & x_4^2 + x_4x_1 + x_1^2 & x_4^3 + x_4^2x_1 + x_4x_1^2 + x_1^3 \\ 0 & 1 & x_5 + x_1 & x_5^2 + x_5x_1 + x_1^2 & x_5^3 + x_5^2x_1 + x_5x_1^2 + x_1^3 \end{array}\right)$$

hervorgeht, indem wir, für jedes $i \geq 2$, die i-te Zeile mit $x_i - x_1$ multiplizieren. In der angedeuteten Weise ist V'' eine Obere-Dreicks-Blockmatrix. Sei W der 4×4-Block rechts unten von V''. Dann gilt

$$W = \begin{pmatrix} 1 & x_2 + x_1 \cdot 1 & x_2^2 + x_1(x_2 + x_1) & x_2^3 + x_1(x_2^2 + x_2x_1 + x_1^2) \\ 1 & x_3 + x_1 \cdot 1 & x_3^2 + x_1(x_3 + x_1) & x_3^3 + x_1(x_3^2 + x_3x_1 + x_1^2) \\ 1 & x_4 + x_1 \cdot 1 & x_4^2 + x_1(x_4 + x_1) & x_4^3 + x_1(x_4^2 + x_4x_1 + x_1^2) \\ 1 & x_5 + x_1 \cdot 1 & x_5^2 + x_1(x_5 + x_1) & x_5^3 + x_1(x_5^2 + x_5x_1 + x_1^2) \end{pmatrix}.$$

Wir ziehen nun von der vierten Spalte von W das x_1-fache der dritten Spalte ab, dann von der dritten Spalte das x_1-fache der zweiten Spalte, und schließlich von der zweiten Spalte das x_1-fache der ersten Spalte. Das Ergebnis ist die Matrix

$$W' = \begin{pmatrix} 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \\ 1 & x_5 & x_5^2 & x_5^3 \end{pmatrix}.$$

Dies ist wiederum eine Vandermonde-Matrix der Größe 4×4. Per Induktion können wir annehmen, dass ihre Determinante bereits als

$$\det(W') = \prod_{\substack{i,j \in \{2,\ldots,5\} \\ i<j}} (x_j - x_i)$$

bekannt ist. Auf Grund der Rechenregeln in Satz 6.2.7 und Proposition 6.5.1 folgt wie gewünscht

$$\begin{aligned}\det(V) = \det(V') &= (x_2-x_1)(x_3-x_1)(x_4-x_1)(x_5-x_1)\det(V'')\\ &= (x_2-x_1)(x_3-x_1)(x_4-x_1)(x_5-x_1)\det(W)\\ &= (x_2-x_1)(x_3-x_1)(x_4-x_1)(x_5-x_1)\det(W') = \prod_{\substack{i,j\in\{1,\dots,5\}\\ i<j}} (x_j-x_i)\end{aligned}$$ □

Korollar 6.6.3 (Polynom zu vorgeschriebenen Werten)**.** *Sei $n \in \mathbb{N}$. Seien $x_1, \dots, x_n \in K$ paarweise verschiedene Elemente und $b_1, \dots, b_n \in K$ beliebige Elemente. Dann gibt es genau ein Polynom $p \in K[X]$ vom Grad $\leq n-1$ mit $p(x_i) = b_i$ für alle $i = 1, \dots, n$.*

Beweis. Der Ansatz $p = C_0 + C_1X + \dots C_{n-1}X^{n-1} \in K[X]$ führt zu einem linearen Gleichungssystem mit n Gleichungen in den n Variablen $C_0, \dots, C_{n-1}$. Dieses hat die Form $VC = b$, wobei $V = V(x_1, \dots, x_n)$ die Vandermonde-Matrix, C der Spaltenvektor mit den Variablen $C_0, \dots, C_{n-1}$ als Einträgen und b der Spaltenvektor mit den Einträgen $b_0, \dots, b_{n-1}$ ist. Nach Satz 6.6.1 und unserer Annahme $x_i \neq x_j$ für alle $i \neq j$ gilt

$$\det(V) = \prod_{\substack{i,j\in\{1,\dots,n\}\\ i<j}} \underbrace{(x_j - x_i)}_{\neq 0} \neq 0.$$

Nach dem Determinantenkriterium 6.4.1 ist V invertierbar. Die eindeutige Lösung unseres linearen Gleichungssystems $VC = b$ ist somit der Spaltenvektor $c = V^{-1}b$. Dies zeigt die behauptete Existenz und Eindeutigkeit von p. □

6.6.4. Man kann das Polynom p aus Korollar 6.6.3 auch direkt hinschreiben. Für jedes $i \in \{1, \dots, n\}$ hat das Polynom

$$q_i := \frac{\prod_{j\neq i}(X - x_j)}{\prod_{j\neq i}(x_i - x_j)} \in K[X]$$

(mit Nenner $\neq 0$) die Eigenschaft $q_i(x_s) = \delta_{is}$. Setzen wir $p := \sum_i b_i q_i$ so gilt $p(x_s) = b_s$ für alle s. Wegen $\deg(q_i) = n-1$ hat p höchstens Grad $n-1$. Dies zeigt die Existenz des gesuchten Polynoms p.

In der Notation vom Anfang des Beweises von Korollar 6.6.3 zeigt dies, dass es für jeden Vektor $b = (b_1, \dots, b_n) \in K^n$ einen Vektor $c = (c_0, c_1, \dots, c_{n-1}) \in K^n$ mit $Vc = b$ gibt. Die lineare Abbildung $V\colon K^n \to K^n$ ist also surjektiv. Nach Korollar 5.5.9 ist sie dann auch injektiv. Dies zeigt die Eindeutigkeit von p. Insgesamt liefert dies einen Alternativbeweis von Korollar 6.6.3.

6.6.5. Sei $\zeta \in \mathbb{C}$ eine primitive n-te Einheitswurzel[41], etwa $\zeta = e^{2\pi i/n}$, wobei $n \in \mathbb{N}_{>0}$. Die Vandermonde-Matrix zur Folge der n-ten Einheitswurzeln $1, \zeta, \zeta^2, \dots, \zeta^{n-1}$, genannt **Fourier-Matrix**, ist nützlich beim Multiplizieren mit diskreter Fourier-Transformation, siehe etwa [Sed90, Chapter 41, The Fast Fourier Transform]; vgl. [Wik19, Diskrete Fourier-Transformation].

[41] Ein komplexe Zahl $z \in \mathbb{C}$ mit $z^n = 1$ heißt n-te komplexe Einheitswurzel; eine solche heißt primitiv, falls $z^m \neq 1$ für alle $m \in \mathbb{N}$ mit $0 < m < n$ gilt.

Aufgabe 6.6.6. Sei $\zeta \in \mathbb{C}$ eine primitive n-te Einheitswurzel mit $\zeta \neq 1$, etwa $\zeta = e^{2\pi i/n}$ für $n \in \mathbb{N}_{>0}$. Dann ist die Vandermonde-Matrix zu $1, \zeta, \zeta^2, \ldots, \zeta^{n-1}$ invertierbar und hat als Inverse das $\frac{1}{n}$-fache der Vandermonde-Matrix zu $1, \zeta^{-1}, \zeta^{-2}, \ldots, \zeta^{-n+1}$. Beispielsweise ist für $n = 3$ das Produkt der beiden Fourier-Matrizen

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta^4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta \end{pmatrix} \qquad \text{und} \qquad \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta^{-1} & \zeta^{-2} \\ 1 & \zeta^{-2} & \zeta^{-4} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta^2 & \zeta \\ 1 & \zeta & \zeta^2 \end{pmatrix}$$

das 3-fache der Einheitsmatrix.

Hinweis: Es gilt $1 + \xi + \xi^2 + \cdots + \xi^{n-1} = 0$ für jede komplexe n-te Einheitswurzel $\xi \neq 1$ (das Polynom $X^n - 1 \in \mathbb{C}[X]$ hat die Nullstelle 1 und somit ist der Faktor $X - 1$ abspaltbar).

Ende 20. Vorlesung am 02.07.2020

6.7. Laplacescher Entwicklungssatz.

Satz 6.7.1 (Laplacescher Entwicklungssatz)**.** *Für eine quadratische Matrix $A \in K^{n\times n}$ und $s, t \in \{1, \ldots, n\}$ bezeichne $A\langle s, t\rangle$ die sogenannte **Streichmatrix**, die man aus A erhält, indem man die s-te Zeile und die t-te Spalte streicht.*

*Dann gilt für jedes fest gewählte $i \in \{1, \ldots, n\}$ die **Entwicklung der Determinanten nach der i-ten Zeile***

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A\langle i, j\rangle).$$

*Analog gilt für jedes fest gewählte $j \in \{1, \ldots, n\}$ die **Entwicklung der Determinanten nach der j-ten Spalte***

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A\langle i, j\rangle).$$

Beweis. Wir beweisen die Entwicklung nach der j-ten Spalte. Die Entwicklung nach der i-ten Zeile folgt dann sofort mit Hilfe der Formeln $\det(A) = \det(A^{\mathrm{t}})$ (siehe Lemma 6.1.8) und $A^{\mathrm{t}}\langle s, t\rangle = (A\langle t, s\rangle)^{\mathrm{t}}$ (oder kann analog bewiesen werden).

Es gilt $Ae_j = \sum_{i=1}^{n} a_{ij} e_i$. Damit berechnen wir

$$\begin{aligned} \det(A) &= \det(Ae_1 \mid \cdots \mid Ae_j \mid \cdots \mid Ae_n) \\ &= \sum_{i=1}^{n} a_{ij} \det(Ae_1 \mid \cdots \mid Ae_{j-1} \mid e_i \mid Ae_{j+1} \mid \cdots \mid Ae_n). \end{aligned}$$

Es bleibt also

$$\det(Ae_1 \mid \cdots \mid Ae_{j-1} \mid e_i \mid Ae_{j+1} \mid \cdots \mid Ae_n) = (-1)^{i+j} \det(A\langle i, j\rangle) \tag{6.7.1}$$

für alle $i \in \{1, \ldots, n\}$ zu zeigen.

Wir verwenden dazu die Rechenregeln aus Satz 6.2.7 und fixieren i. Wenn man in der Matrix $(Ae_1 \mid \cdots \mid Ae_{j-1} \mid e_i \mid Ae_{j+1} \mid \cdots \mid Ae_n)$ geeignete Vielfache der j-ten Spalte e_i zu den anderen Spalten dazuaddiert, kann man erreichen, dass der i-te Eintrag all dieser Spalten Null wird. Dabei ändert sich die Determinante nicht.[42] Nun lässt man die j-te Spalte bis ganz nach rechts wandern, indem man sie sukzessive mit der nächsten Spalte vertauscht.

[42] Auf diesen ersten Schritt kann auch verzichtet werden, wenn man am Ende des Beweises die Formel für die Determinante einer Unteren-Dreiecks-Blockmatrix verwendet.

Dabei ändert sich das Vorzeichen der Determinante um den Faktor $(-1)^{n-j}$, denn es finden $n-j$ Vertauschungen statt. Nun lässt man die i-te Zeile bis ganz nach unten wandern, indem man sie sukzessive mit der nächsten Zeile vertauscht. Dabei ändert sich das Vorzeichen der Determinante um den Faktor $(-1)^{n-i}$, denn es finden $n-i$ Vertauschungen statt. Wir erhalten so die Matrix

$$\left(\begin{array}{c|c} A\langle i,j\rangle & 0 \\ \hline 0 & 1 \end{array}\right).$$

Diese hat offensichtlich dieselbe Determinante wie die Matrix $A\langle i,j\rangle$. Wegen $(-1)^{n-j+n-i} = (-1)^{i+j}$ zeigt das die Gleichheit (6.7.1). □

6.8. **Adjunkte Matrix und Cramersche Regel.** Abschnitt im Sommersemester 2020 weggelassen.

Bemerkung 6.8.1 (Herleitung der Cramerschen Formeln)**.** Seien $A \in K^{n\times n}$ und $b \in K^n$ und betrachte das LGS $AX = b$. Wir nehmen an, dass es eine Lösung $x \in K^n$ hat. Dies bedeutet, dass b wie folgt als Linearkombination der Spalten von A geschrieben werden kann:

$$b = x_1 Ae_1 + \cdots + x_n Ae_n.$$

Wir berechnen für beliebiges $i \in \{1,\dots,n\}$

(6.8.1)

$$\begin{aligned}\det(Ae_1,\dots,Ae_{i-1},b,Ae_{i+1},\dots,Ae_n) &= \sum_{j=1}^{n} x_j \det(Ae_1,\dots,Ae_{i-1},Ae_j,Ae_{i+1},\dots,Ae_n)\\ &= x_i \det(Ae_1,\dots,Ae_{i-1},Ae_i,Ae_{i+1},\dots,Ae_n)\\ &= x_i \det(A).\end{aligned}$$

Ist $\det(A) \neq 0$ (also $A \in \mathrm{GL}_n(K)$, siehe Satz 6.4.1), so erhalten wir

$$x_i = \frac{\det(Ae_1,\dots,Ae_{i-1},b,Ae_{i+1},\dots,Ae_n)}{\det(A)}. \tag{6.8.2}$$

Weiss man von Anfang an, dass $A \in \mathrm{GL}_n(K)$ gilt, so hat das LGS $AX = b$ eine eindeutige Lösung x (nämlich $x = A^{-1}b$), und die obige Begründung zeigt, dass die Koordinaten des Lösungsvektors durch die **Cramerschen Formeln** (6.8.2) gegeben sind. In Worten steht im Nenner die Determinante von A, und im Zähler steht die Determinante der Matrix, die man aus A erhält, indem man die i-te Spalte durch b ersetzt.

Diese Lösungsformeln sind leider in der Praxis meist unbrauchbar.

Satz 6.8.2 (Cramersche Lösungsformeln)**.** *Sei $A \in \mathrm{GL}_n(K)$ und $b \in K^n$. Sei $x \in K^n$ der Spaltenvektor, dessen Einträge x_i durch die Cramerschen Formeln (6.8.2) definiert sind, für $i = 1,\dots,n$. Dann ist x die eindeutige Lösung des LGS $AX = b$, d. h. es gilt $x = A^{-1}b$.*

Beweis. Dies ist klar nach der Herleitung in Bemerkung 6.8.1. □

Bemerkung 6.8.3 (Motivation der adjunkten Matrix und der Cramerschen Regel)**.** Sei $A \in \mathrm{GL}_n(K)$. Sei $j \in \{1,\dots,n\}$ beliebig und sei $x^j \in K^n$ die Lösung von $Ax^j = e_j$. Wir erhalten aus (6.8.1) (für $b = e_j$) die Gleichung

$$x_i^j \det(A) = \det(Ae_1,\dots,Ae_{i-1},e_j,Ae_{i+1},\dots,Ae_n).$$

Wir erhalten per Laplace-Entwicklung der Determinanten auf der rechten Seite nach der i-ten Spalte

$$x_i^j \det(A) = (-1)^{i+j} \det(A\langle j, i\rangle).$$

Sei $B \in K^{n\times n}$ durch

$$b_{ij} := (-1)^{i+j} \det(A\langle j, i\rangle) \tag{6.8.3}$$

definiert. Die j-te Spalte von B ist also gerade $\det(A)x^j$. Aus $Ax^j = e_j$ folgt $A\det(A)x^j = \det(A)e_j$ und somit $AB = \det(A)I_n$.

Zusammengefasst definiert also für jedes $A \in \mathrm{GL}_n(K)$ die Formel (6.8.3) (die nur von A abhängt) eine Matrix B mit $AB = \det(A)I_n$.

Es ist B die zu A *adjunkte Matrix* im Sinne der folgenden Definition 6.8.4, in der A nicht als invertierbar vorausgesetzt ist. Die Cramersche Regel (siehe Satz 6.8.5) verallgemeinert unsere Formel $AB = \det(A)I_n$ zu nicht notwendig invertierbarem A.

Definition 6.8.4. Sei $A \in K^{n\times n}$ eine quadratische Matrix. Die **adjunkte Matrix** $A^\# \in K^{n\times n}$ ist definiert durch

$$(A^\#)_{ij} = (-1)^{i+j} \det(A\langle j, i\rangle),$$

wobei $A\langle j, i\rangle$ die in Satz 6.7.1 erklärte Streichmatrix ist (beachte die Reihenfolge von i und j).

Satz 6.8.5 (Cramersche Regel)**.** *Für alle quadratischen Matrizen $A \in K^{n\times n}$ gilt*

$$A \cdot A^\# = \det(A) \cdot I_n = A^\# \cdot A.$$

In Worten ist das Matrixprodukt von A mit ihrer adjunkten $A^\#$ das $\det(A)$-fache der Einheitsmatrix.

Beweis. Wir zeigen nur die erste Gleichheit. Der Beweis der zweiten Gleichheit geht vollkommen analog.

Zu zeigen ist $(A \cdot A^\#)_{ij} = \delta_{ij} \det(A)$ für alle $i, j \in \{1, \ldots, n\}$, wobei δ_{ij} das Kronecker-δ ist.

Für $i = j$ gilt

$$(A \cdot A^\#)_{ii} = \sum_{s=1}^{n} a_{is}(A^\#)_{si} = \sum_{s=1}^{n} a_{is}(-1)^{s+i} \det(A\langle i, s\rangle) = \det(A),$$

wobei die letzte Gleichheit die Laplace-Entwicklung der Determinante nach der i-ten Zeile ist (siehe Satz 6.7.1).

Für $i \neq j$ gilt

$$(A \cdot A^\#)_{ij} = \sum_{s=1}^{n} a_{is}(A^\#)_{sj} = \sum_{s=1}^{n} a_{is}(-1)^{s+j} \det(A\langle j, s\rangle). \tag{6.8.4}$$

Sei B die Matrix, die aus A entsteht, indem man in die j-te Zeile eine Kopie der i-ten Zeile schreibt. Dann hat B zwei gleiche Zeilen, und wir erhalten mit der Laplace-Entwicklung der Determinante nach der j-ten Zeile (siehe Satz 6.7.1))

$$0 = \det(B) = \sum_{s=1}^{n} (-1)^{j+s} b_{js} \det(B\langle j, s\rangle) = \sum_{s=1}^{n} (-1)^{j+s} a_{is} \det(A\langle j, s\rangle).$$

Vergleich mit der rechten Seite von (6.8.4) liefert wie gewünscht $(A \cdot A^\#)_{ij} = 0$. □

Korollar 6.8.6. *Für $A \in \mathrm{GL}_n(K)$ gilt*

$$A^{-1} = \frac{1}{\det(A)} A^{\#}.$$

Beweis. Da wir bereits wissen, dass $\det(A)$ in K invertierbar ist (siehe Satz 6.4.1), folgt dies aus Satz 6.8.5. □

6.8.7. Somit ist die eindeutige Lösung eines „quadratischen" LGS $AX = b$ mit $\det(A) \neq 0$ durch $A^{-1}b = \frac{1}{\det(A)} A^{\#} b$ gegeben.

Beispiel 6.8.8. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist die adjunkte Matrix $A^{\#} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Es gilt

$$A \cdot A^{\#} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = \det(A) I_2.$$

Ist $\det(A) = ad - bc \neq 0$, so gilt $A \in \mathrm{GL}_2(K)$ und

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{6.8.5}$$

Bemerkung 6.8.9. Die Definition der Determinanten (6.1.1) verallgemeinert in offensichtlicher Weise zu quadratischen Matrizen mit Einträgen in einem beliebigen kommutativen Ring R. Das in offensichtlicher Weise definierte Matrix-Produkt macht die Menge $R^{n\times n}$ all solcher Matrizen zu einem Ring. Multiplikativität der Determinanten (Satz 6.3.1) und die Cramersche Regel (Satz 6.8.5) gelten auch in diesem Kontext (wie auch viele andere Resultate über Determinanten(funktionen)).

Daraus kann man das folgende Determinantenkriterium für Invertierbarkeit folgern (vgl. Satz 6.4.1): Eine $(n \times n)$-Matrix $A \in R^{n\times n}$ ist genau dann invertierbar, wenn $\det(A)$ eine Einheit in R ist, wenn also in Formeln $\det(A) \in R^{\times}$ gilt.

Die Implikation $\Rightarrow$ folgt aus der Multiplikativität der Determinanten, die Implikation $\Leftarrow$ aus der Cramerschen Regel.

Beispielsweise ist die 2×2-Matrix $\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$ mit ganzzahligen Einträgen invertierbar im Ring $\mathbb{Z}^{2\times 2}$ der ganzzahligen 2×2-Matrizen, denn ihre Determinante $14 - 15 = -1$ ist eine Einheit in $\mathbb{Z}$. Ihre Inverse ist die Matrix $\begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix}$, vgl. (6.8.5).

Die Matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ hat Determinante $4 - 6 = -2$, was keine Einheit in $\mathbb{Z}$ ist. Also ist sie im Ring $\mathbb{Z}^{2\times 2}$ nicht invertierbar. Im Ring $\mathbb{Q}^{2\times 2}$ ist sie aber invertierbar, mit Inversem $\begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.

7. Endomorphismen

7.1. Eigenwerte und Eigenvektoren.

7.1.1. Wir möchten Endomorphismen endlichdimensionaler Vektorräume möglichst gut verstehen. Es ist naheliegend, alle Vektoren $v \in V$ zu betrachten, die von einem gegebenen Endomorphismus $f\colon V \to V$ auf Null abgebildet werden (also $f(v) = 0$ erfüllen) oder von f

fixiert werden (also $f(v) = v$ erfüllen). Etwas allgemeiner betrachtet man alle Vektoren v, die von f auf ein Vielfaches von v abgebildet werden. Dies führt zum Begriff des Eigenvektors.

Definition 7.1.2. Sei $f\colon V \to V$ ein Endomorphismus eines K-Vektorraums.

- Ein Vektor $v \in V$ heißt genau dann **Eigenvektor (EV) von** f, wenn $v \neq 0$ gilt und $f(v)$ ein skalares Vielfaches von v ist, es also ein $\lambda \in K$ mit $f(v) = \lambda v$ gibt.
- Ein Skalar $\lambda \in K$ heißt genau dann **Eigenwert (EW) von** f, wenn es einen Eigenvektor $v \in V$ mit $f(v) = \lambda v$ gibt.
- Ein **Eigenvektor von** f **zum Eigenwert** λ ist ein Eigenvektor v von f mit $f(v) = \lambda v$.
- Für beliebiges $\lambda \in K$ nennen wir den Untervektorraum
$$\operatorname{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda v\} = \ker(\lambda \mathrm{id}_V - f)$$
von V den **Eigenraum von** f **zu** λ.

7.1.3. Skalare Vielfache $\neq 0$ von Eigenvektoren sind Eigenvektoren: Ist v ein Eigenvektor von f zum Eigenwert λ und $a \in K \setminus \{0\}$ ein Skalar ungleich Null, so ist auch $av \neq 0$ ein Eigenvektor von f zum selben Eigenwert λ.

7.1.4. Ist λ kein Eigenwert, so gilt $\operatorname{Eig}(f, \lambda) = \{0\}$. Ist λ ein Eigenwert, so besteht $\operatorname{Eig}(f, \lambda)$ genau aus allen Eigenvektoren von f zum Eigenwert λ und aus der Null (denn $(\lambda \mathrm{id}_V - f)(v) = 0$ ist äquivalent zu $f(v) = \lambda v$).

7.1.5. Ist $v \in V$ ein Eigenvektor von f, so gibt es genau ein $\lambda \in K$, so dass v Eigenvektor von f zum Eigenwert λ ist: Aus $f(v) = \lambda v$ und $f(v) = \mu v$ für $\lambda, \mu \in K$ folgt $(\lambda - \mu)v = 0$ und somit $\lambda = \mu$ wegen $v \neq 0$ und Lemma 4.1.7.(f).

7.1.6. Lemma 4.1.7.(f) zeigt auch
$$\operatorname{Eig}(f, \lambda) \cap \operatorname{Eig}(f, \mu) = \{0\}$$
für alle Skalare $\lambda \neq \mu$: Ist v im Schnitt, so gilt $\lambda v = f(v) = \mu v$, also $(\lambda - \mu)v = 0$ und wegen $\lambda - \mu \neq 0$ folgt $v = 0$.

Beispiel 7.1.7. Betrachte die drei linearen Abbildungen (siehe 5.1.4.(a))
$$P = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2.$$

(a) Punktspiegelung P im Ursprung: Jeder Vektor $v \in \mathbb{R}^2$ mit $v \neq 0$ ist Eigenvektor von P zum Eigenwert -1.

(b) Spiegelung S an der ersten Winkelhalbierenden: Der Vektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (und jedes Vielfache $\neq 0$) ist ein Eigenvektor von S zum Eigenwert 1. Der Vektor $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ (und jedes Vielfache $\neq 0$) ist ein Eigenvektor von S zum Eigenwert -1. Kein anderer Vektor ist ein Eigenvektor.

(c) Drehung D um 90°: Es gibt keine Eigenvektoren. Dies ist hoffentlich anschaulich klar, die formale Rechnung geht wie folgt: Die Annahme, dass $0 \neq \begin{pmatrix} x \\ y \end{pmatrix}$ ein Eigenvektor zum Eigenwert $\lambda \in K$ ist, liefert die Gleichungen $x = \lambda y$ und $-y = \lambda x$. Dies

liefert $x = \lambda y = -\lambda^2 x$, also $(1+\lambda^2)x = 0$. Wegen $1+\lambda^2 > 0$ folgt $x = 0$ und dann $y = -\lambda x = 0$. Dies widerspricht der Annahme, denn der Nullvektor ist kein Eigenvektor.

Beispiel 7.1.8. In Beispiel 7.1.7 haben wir die Drehung D betrachtet, die keinen Eigenwert hatte. Wir verändern dieses Beispiel, indem wir $\mathbb{R}^2$ durch $\mathbb{C}^2$ ersetzen und die Abbildung

$$D' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : \mathbb{C}^2 \to \mathbb{C}^2$$

betrachten. Sie hat die Eigenvektoren $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ zum Eigenwert i und $\begin{pmatrix} 1 \\ i \end{pmatrix}$ zum Eigenwert $-i$. Diese beiden Vektoren bilden eine Basis von $\mathbb{C}^2$ aus Eigenvektoren.

Beispiel 7.1.9. Sei $A \in K^{n\times n}$ eine Matrix. Ist A eine Diagonalmatrix, so ist der i-te Standardbasisvektor e_i ein Eigenvektor von $A\colon K^n \to K^n$ zum Eigenwert a_{ii}, für all i; außerdem ist die Menge der Eigenwerte genau die Menge $\{a_{11}, \dots, a_{nn}\}$ der Diagonaleinträge, denn für jedes andere λ gilt $\ker(\lambda I_n - A) = \{0\}$.

Sind umgekehrt alle e_i Eigenvektoren von A, so hat A notwendig Diagonalgestalt.

Beispiel 7.1.10. Sei $\lambda \in \mathbb{R}$. Die Matrix $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ hat λ als einzigen Eigenwert: Zugehörige Eigenvektoren sind alle Vielfachen von e_1 bis auf den Nullvektor.

Beispiel 7.1.11. Sei $x \in K$. Hat eine Matrix $A \in K^{n\times n}$ den Eigenwert λ (mit zugehörigem Eigenvektor v), so hat $A + xI_n$ den Eigenwert $\lambda + x$ (mit demselben Eigenvektor v).

Im vorigen Beispiel 7.1.10 hätte es also genügt, denn Fall $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ zu betrachten.

Beispiel 7.1.12. Die Abbildung

$$A = \begin{pmatrix} -13 & 14 \\ -7 & 8 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$$

hat die Eigenvektoren $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ zum Eigenwert 1 und $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ zum Eigenwert -6. Diese beiden Vektoren bilden eine Basis $\mathcal{B}$ von $\mathbb{R}^2$, für die

$${}_{\mathcal{B}}[A]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -6 \end{pmatrix}$$

gilt (im Sinne der späteren Definition 7.4.4 ist A also diagonalisierbar).

Satz 7.1.13 (Lineare Unabhängigkeit von Eigenvektoren)**.** *Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. Genauer sei $f\colon V \to V$ ein Endomorphismus eines Vektorraums V und v_i sei ein Eigenvektor von f zum Eigenwert λ_i, für $i = 1, \dots, r$. Sind die Eigenwerte $\lambda_1, \dots, \lambda_r \in K$ paarweise verschieden, so sind die Vektoren $v_1, \dots, v_r$ linear unabhängig.*[43]

[43] Oft ist die folgende äquivalente Aussage nützlich (diese wird etwa im Beweis von Satz 7.4.7 verwendet): Sei $f\colon V \to V$ ein Endomorphismus eines Vektorraums V. Seien $\lambda_1, \dots, \lambda_r \in K$ paarweise verschiedene Elemente des Körpers und gelte

$$0 = v_1 + \cdots + v_r$$

für Vektoren $v_i \in \mathrm{Eig}(f, \lambda_i)$, wobei $i = 1, \dots, r$. Dann sind alle v_i Null, in Formeln $0 = v_1 = \cdots = v_r$.

Beispiel 7.1.14. Beispielsweise sind in Beispiel 7.1.8 die beiden dort angegebenen Eigenvektoren linear unabhängig.

Erster Beweis. Wir beweisen dies per Induktion nach r. Für $r = 1$ ist die Aussage richtig, denn $v_1 \neq 0$.

Gelte nun $r \geq 2$. Gelte

$$a_1 v_1 + \cdots + a_r v_r = 0 \tag{7.1.1}$$

für Skalare $a_1, \ldots, a_r \in K$. Wir wenden einerseits f auf diese Gleichung an und erhalten

$$a_1 \lambda_1 v_1 + \cdots + a_r \lambda_r v_r = 0.$$

Andererseits multiplizieren wir sie mit λ_r und erhalten

$$\lambda_r a_1 v_1 + \cdots + \lambda_r a_r v_r = 0.$$

Die Differenz der beiden Gleichungen ist

$$(\lambda_1 - \lambda_r) a_1 v_1 + (\lambda_2 - \lambda_r) a_2 v_2 + \cdots + (\lambda_{r-1} - \lambda_r) a_{r-1} v_{r-1} = 0.$$

Die Induktionsvoraussetzung besagt, dass die Vektoren $v_1, \ldots, v_{r-1}$ linear unabhängig sind. Es folgt $0 = (\lambda_1 - \lambda_r) a_1 = \cdots = (\lambda_{r-1} - \lambda_r) a_{r-1}$. Da die λ_i paarweise verschieden sind, folgt $0 = a_1 = \cdots = a_{r-1}$. Setzen wir diese Erkenntnis in (7.1.1) ein, so erhalten wir $a_r v_r = 0$. Wegen $v_r \neq 0$ folgt $a_r = 0$. Dies zeigt die lineare Unabhängigkeit von $v_1, \ldots, v_r$. □

Zweiter Beweis. Gelte

$$a_1 v_1 + \cdots + a_r v_r = 0 \tag{7.1.2}$$

für Skalare $a_1, \ldots, a_r \in K$.

Die Abbildung $f - \lambda_i := f - \lambda_i \mathrm{id}_V \colon V \to V$ bildet v_i auf Null ab und jedes v_j, für $j \neq i$, auf $(\lambda_j - \lambda_i) v_j \neq 0$. Wenden wir die Verknüpfung

$$(f - \lambda_2)(f - \lambda_3) \cdots (f - \lambda_r) = (f - \lambda_2 \mathrm{id}_V) \circ (f - \lambda_3 \mathrm{id}_V) \circ \cdots \circ (f - \lambda_r \mathrm{id}_V)$$

auf die Gleichung (7.1.2) an (je zwei der verknüpften Abbildungen kommutieren), so erhalten wir also

$$(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \cdots (\lambda_1 - \lambda_r) a_1 v_1 = 0.$$

Da v_1 und alle Skalare $\lambda_1 - \lambda_2, \ldots \lambda_1 - \lambda_r$ von Nullverschieden sind, folgt $a_1 = 0$. Analog erhält man $0 = a_2 = \cdots = a_r$. □

Korollar 7.1.15. *Ist $f \colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums V, so hat f höchstens $\dim(V)$ verschiedene Eigenwerte.*

Beweis. Seien $\lambda_1, \ldots, \lambda_r$ verschiedene Eigenwerte von f. Sei v_i ein Eigenvektor mit Eigenwert λ_i. Nach Satz 7.1.13 sind $v_1, \ldots, v_r$ linear unabhängig. Somit gilt $r \leq \dim(V)$ nach Satz 4.3.6.(b). □

Lemma 7.1.16. *Sei $f \colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums und sei $\lambda \in K$. Dann besitzt f genau dann einen Eigenvektor zum Eigenwert λ, falls es eine Basis $\mathcal{B}$ von V mit*

$${}_{\mathcal{B}}[f]_{\mathcal{B}} = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$

gibt.

Beweis. Besitze f einen Eigenvektor v zum Eigenwert λ. Nach dem Basisergänzungssatz 4.2.17 (und da v wegen $v \neq 0$ linear unabhängig ist) können wir v zu einer Basis $\mathcal{B} = (v, b_2, b_3, \ldots, b_n)$ ergänzen. Wegen $f(v) = \lambda v$ hat ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ die gewünschte Gestalt.

Ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V, so dass ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ die angegebene Gestalt hat, so gilt $f(b_1) = \lambda b_1$. Wegen $b_1 \neq 0$ ist b_1 ein Eigenvektor zum Eigenwert λ. □

7.2. **Charakteristisches Polynom.**

7.2.1. Sei R ein kommutativer Ring.

7.2.2. Wir verallgemeinern einige Begriffsbildungen von Körpern zu kommutativen Ringen.

Eine $(m \times n)$**-Matrix mit Einträgen in** R ist eine Abbildung

$$A\colon \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to R.$$

Sie wird meist als

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{7.2.1}$$

notiert. Wir schreiben $R^{m\times n}$ für die Menge aller $(m \times n)$-Matrizen mit Einträgen in R.

Addition zweier Matrizen $A, B \in R^{n\times m}$, Skalarmultiplikation eines Skalars $\lambda \in R$ mit einer Matrix $A \in R^{n\times m}$ und Multiplikation zweier Matrizen $A \in R^{m\times n}$ und $B \in R^{n\times r}$ werden durch die üblichen Formeln definiert (siehe Beispiel 4.1.6 und Definition 5.4.5, wo man K durch R ersetze). Alle Rechenregeln für Matrizen in Lemma 5.4.10 gelten analog in diesem verallgemeinerten Rahmen. Viele Begriffsbildungen (Zeilen, Spalten, quadratische Matrizen, obere Dreiecksmatrizen, elementare Zeilen- und Spaltenoperationen, etc.) verallgemeinern in offensichtlicher Weise (vgl. Bemerkung 6.8.9).

Insbesondere definieren wir die Determinante $\det(A) \in R$ einer quadratischen Matrix $A \in R^{n\times n}$, indem wir in Definition 6.1.1 den Körper K durch den kommutativen Ring R ersetzen:

$$\det(A) := \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \tag{7.2.2}$$

Die Determinante $\det\colon R^{n\times n} \to R$ ist durch die folgenden Eigenschaften eindeutig bestimmt (vgl. 6.2.5; der Beweis ist vollkommen analog zum Beweis von Satz 6.2.4)

- $\det$ ist R-linear in jeder Spalte, wenn die anderen Spalten fixiert sind (wobei R-Linearität in offensichtlicher Weise definiert ist),
- $\det(A) = 0$, falls zwei Spalten von A übereinstimmen,
- $\det(I_n) = 1$.

7.2.3. Die meisten Resultate aus Kapitel 6 gelten auch[44] für Matrizen mit Einträgen in R, zum Beispiel die Berechnung der Determinante per Entwicklung nach einer Spalte oder Zeile

[44] Im Wesentlichen mit den angegebenen Beweisen. Da wir eigentlich nur am Polynomring $K[X]$ interessiert sind, kann man alternativ auch seinen Quotientenkörper $K(X)$ (nicht definiert) betrachten. Alle Resultate über Determinanten gelten für Matrizen über diesem Körper und implizieren die entsprechenden Resultate für Matrizen mit Einträgen in $K[X]$, da $K[X]$ ein Unterring von $K(X)$ ist.

(Laplacescher Entwicklungssatz 6.7.1), die Berechnung von oberen Dreiecksmatrizen (siehe Lemma 6.1.5) oder Oberen-Dreiecks-Blockmatrizen (siehe Korollar 6.5.2), die Multiplikativität der Determinante $\det(AB) = \det(A)\det(B)$ (siehe Satz 6.3.1), $\det(A^{\mathrm{t}}) = \det(A)$ (siehe Lemma 6.1.8), die Rechenregeln aus 6.2.7.

Definition 7.2.4. Das **charakteristische Polynom** einer quadratischen Matrix $A \in K^{n\times n}$ alias einer linearen Abbildung $A\colon K^n \to K^n$ ist das Polynom

$$\chi_A = \chi_A(X) := \det(XI_n - A) \in K[X].$$

7.2.5. Ausgeschrieben ist das charakteristische Polynom von $A = (a_{ij}) \in K^{n\times n}$ also die Determinante der Matrix

$$XI_n - A = \begin{pmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & & -a_{2n} \\ \vdots & & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & X - a_{nn} \end{pmatrix} \in K[X]^{n\times n},$$

d. h. nach der Leibniz-Formel gilt

$$\chi_A(X) = \det(XI_n - A) := \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (X\delta_{i\sigma(i)} - a_{i\sigma(i)})$$

unter Verwendung des Kronecker-δ.

Ende 21. Vorlesung am 07.07.2020

Satz 7.2.6. *Sei $A \in K^{n\times n}$ eine beliebige quadratische Matrix. Dann sind die Nullstellen von χ_A in K genau die Eigenwerte des Endomorphismus $A\colon K^n \to K^n$, in Formeln*

$$\{\textit{Eigenwerte von } A\colon K^n \to K^n\} = \{\textit{Nullstellen von } \chi_A \textit{ in } K\}.$$

Beweis. Die Auswertung des charakteristische Polynoms $\chi_A(X)$ bei $\lambda \in K$ ist

$$\chi_A(\lambda) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (\lambda\delta_{i\sigma(i)} - a_{i\sigma(i)}) = \det(\lambda I_n - A).$$

Für jedes $\lambda \in K$ sind die folgenden Aussagen äquivalent:

- λ ist Nullstelle von χ_A;
- $\det(\lambda I_n - A) = 0$;
- $\lambda I_n - A\colon K^n \to K^n$ ist nicht injektiv (Äquivalenz zum vorigen Punkt nach den Sätzen 6.4.1 und 5.6.5);
- es gibt ein $x \in K^n \setminus \{0\}$ mit $(\lambda I_n - A)x = 0$;
- es gibt ein $x \in K^n \setminus \{0\}$ mit $\lambda x = Ax$;
- λ ist Eigenwert von $A\colon K^n \to K^n$.

Dies zeigt die behauptete Gleichheit. □

Beispiel 7.2.7. Betrachte den Endomorphismus

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2.$$

Nach Satz 7.2.6 sind die Eigenwerte von A genau die Nullstellen des charakteristischen Polynoms

$$\chi_A = \det\left(\begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}\right) = \det\left(\begin{pmatrix} X-2 & -1 \\ -2 & X-3 \end{pmatrix}\right)$$
$$= (X-2)(X-3) - 2 = X^2 - 5X + 4 \in \mathbb{R}[X].$$

Nach einer der üblichen Lösungsformeln für quadratische Gleichungen (p-q-Formel oder Mitternachtsformel) sind 1 und 4 die Nullstellen. Die zugehörigen Eigenräume berechnen sich zu

$$\mathrm{Eig}(A,1) = \ker(I_2 - A) = \ker\begin{pmatrix} -1 & -1 \\ -2 & -2 \end{pmatrix} = \left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle,$$
$$\mathrm{Eig}(A,4) = \ker(4I_2 - A) = \ker\begin{pmatrix} 2 & -1 \\ -2 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle,$$

und die angegebenen Erzeuger dieser Eigenräume bilden eine Basis des $\mathbb{R}^2$ aus Eigenvektoren von A.

Beispiel 7.2.8. Betrachten wir die Drehung $D\colon \mathbb{R}^2 \to \mathbb{R}^2$ aus Beispiel 7.1.7, so ist ihr charakteristisches Polynom

$$\chi_D = \det\left(\begin{pmatrix} X & 1 \\ -1 & X \end{pmatrix}\right) = X^2 + 1.$$

Dieses Polynom hat keine *reelle* Nullstelle, denn jedes Quadrat in $\mathbb{R}$ ist nicht-negativ. Also hat D keine Eigenwerte.

Jedoch hat dieses Polybom die beiden *komplexen* (aber nicht reellen) Nullstellen $\pm i$. Diese sind genau die beiden Eigenwerte der Abbildung D' in Beispiel 7.1.8.

Definition 7.2.9. Sei $A = (a_{ij}) \in K^{n\times n}$ eine quadratische Matrix. Dann heißt

$$\mathrm{tr}(A) := a_{11} + a_{22} + \cdots + a_{nn}$$

die **Spur** von A (im Englischen *trace*).

Beispiel 7.2.10. $\mathrm{tr}\begin{pmatrix} 1 & 0 & 3 \\ 1 & 2 & 3 \\ 4 & 6 & -11 \end{pmatrix} = 1 + 2 - 11 = -8.$

Lemma 7.2.11. *Die Spur* $\mathrm{tr}\colon K^{n\times n} \to K$ *ist eine lineare Abbildung. Für alle Matrizen* $A, B \in K^{n\times n}$ *gilt*

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Beweis. Die erste Aussage ist offensichtlich, die zweite rechnet man sofort nach: $\mathrm{tr}(AB) = \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{s=1}^n a_{is}b_{si}$ und analog für $\mathrm{tr}(BA)$. □

Beispiele 7.2.12. (a) Für eine (2×2)-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ gilt

$$\chi_A(X) = \det\left(\begin{pmatrix} X-a & -b \\ -c & X-d \end{pmatrix}\right) = (X-a)(X-d) - bc = X^2 - \underbrace{(a+d)}_{=\mathrm{tr}(A)} X + \underbrace{(ad-bc)}_{=\det(A)}.$$

(b) Für eine (3×3)-Matrix $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ gilt (vgl. Beispiel 6.1.2)

$$\begin{aligned}\chi_A(X) &= +(X-a_{11})(X-a_{22})(X-a_{33}) + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{13}(X-a_{22})a_{31} - a_{12}a_{21}(X-a_{33}) - (X-a_{11})a_{23}a_{32} \\ &= \quad X^3 \\ &\quad -(\underbrace{a_{11}+a_{22}+a_{33}}_{=\mathrm{tr}(A)})X^2 \\ &\quad +(a_{11}a_{22}+a_{11}a_{33}+a_{22}a_{33}-a_{13}a_{31}-a_{12}a_{21}-a_{23}a_{32})X \\ &\quad -\det(A).\end{aligned}$$

(c) Ist $A \in K^{n\times n}$ eine obere Dreiecksmatrix mit Diagonaleinträgen $\lambda_1, \lambda_2, \ldots, \lambda_n$. Dann ist χ_A die Determinante der oberen Dreiecksmatrix

$$\begin{pmatrix} X-\lambda_1 & & * \\ & \ddots & \\ 0 & & X-\lambda_n \end{pmatrix}.$$

Nach Lemma 6.1.5 (bzw. genauer dessen ebenfalls gültiger Version für Matrizen mit Einträgen in $K[X]$) gilt

$$\begin{aligned}\chi_A(X) &= (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n) \\ &= X^n - (\underbrace{\lambda_1+\lambda_2+\cdots+\lambda_n}_{=\mathrm{tr}(A)})X^{n-1} \pm \cdots + (-1)^n \underbrace{\lambda_1\lambda_2\cdots\lambda_n}_{=\det(A)}.\end{aligned}$$

Speziell für $n = 3$ erhält man

$$\chi_A(X) = X^3 - (\underbrace{\lambda_1+\lambda_2+\lambda_3}_{\mathrm{tr}(A)})X^2 + (\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3)X - \underbrace{\lambda_1\lambda_2\lambda_3}_{\det(A)},$$

was man natürlich auch aus der Formel im vorigen Beispiel bekommt.

Beispiel 7.2.13. Ist $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$ eine Obere-Dreiecks-Blockmatrix, wobei $B \in K^{r\times r}$ und $D \in K^{s\times s}$, so gilt $\chi_A = \chi_B \cdot \chi_D$. Dies folgt aus 7.2.3, denn Korollar 6.5.2 gilt sinngemäß für Matrizen mit Einträgen in $K[X]$.

Satz 7.2.14. *Sei $A \in K^{n\times n}$ eine quadratische Matrix. Dann ist χ_A ein normiertes Polynom vom Grad n, also von der Form*

$$\chi_A(X) = X^n - s_1X^{n-1} + \cdots + (-1)^{n-1}s_{n-1}X + (-1)^n s_n$$

für eindeutig bestimmte Zahlen $s_1, s_2, \ldots, s_n \in K$. (Beachte, dass $(-1)^i s_i$ der Koeffizient von X^{n-i} ist.) Es gelten

$$\begin{aligned}s_1 &= \mathrm{tr}(A), \\ s_n &= \det(A).\end{aligned}$$

Beweis. Nach der Leibniz-Formel gilt

$$\chi_A(X) = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma)(X\delta_{1,\sigma(1)} - a_{1,\sigma(1)})\cdots(X\delta_{n,\sigma(n)} - a_{n,\sigma(n)}).$$

Der Summand für $\sigma = \mathrm{id}$ ist

$$(X-a_{11})(X-a_{22})\cdots(X-a_{nn}) = X^n - \underbrace{(a_{11}+a_{22}+\cdots+a_{nn})}_{\mathrm{tr}(A)}X^{n-1} \pm (\text{Terme von kleinerem Grad}).$$

Alle Summanden für $\sigma \neq \mathrm{id}$ haben Grad $\leq n-2$, denn für jedes $\sigma \neq \mathrm{id}$ gibt es mindestens zwei Indizes $1 \leq i < j \leq n$ mit $i \neq \sigma(i)$ und $j \neq \sigma(j)$.

Dies zeigt, dass $\chi_A(X)$ normiert vom Grad n ist und dass $s_1 = \mathrm{tr}(A)$ gilt, wenn die s_i wie in der Formulierung des Satzes definiert sind.

Es ist $(-1)^n s_n$ der Wert des Polynoms $\chi_A(X)$ bei Null, d. h.

$$(-1)^n s_n = \chi_A(0) = \det(0-A) = \det(-A) = (-1)^n \det(A),$$

wobei die letzte Gleichung daraus folgt, dass die Determinante multilinear ist. □

Satz 7.2.15. *Eine Matrix $A \in K^{n\times n}$ alias lineare Abbildung $A\colon K^n \to K^n$ hat höchstens n verschiedene Eigenwerte.*

7.2.16. Wir kennen diese Aussage bereits aus Korollar 7.1.15, jedoch ist der folgende Beweis mit dem charakteristischen Polynom für mich leichter merkbar.

Beweis. Das charakteristische Polynom $\chi_A \in K[X]$ hat nach Satz 7.2.14 Grad n und somit nach Korollar 2.4.23 höchstens n Nullstellen. Da die Menge der Nullstellen von χ_A in K mit der der Eigenwerte von A übereinstimmt (Satz 7.2.6), folgt der Satz. □

Satz 7.2.17. *Ähnliche Matrizen haben dasselbe charakteristische Polynom: Seien $A, B \in K^{n\times n}$ ähnliche Matrizen (d. h. es gibt eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ mit $A = SBS^{-1}$). Dann gilt*

$$\chi_A = \chi_B.$$

Insbesondere haben A und B haben dieselben Eigenwerte.

Beweis. Zunächst beachte man

$$S(XI_n - B)S^{-1} = S(XI_n)S^{-1} - SBS^{-1} = (XI_n)SS^{-1} - A = XI_n - A$$

in $K[X]^{n\times n}$.[45] Damit berechnen wir

$$\begin{aligned}
\chi_A &= \det(XI_n - A) \\
&= \det(S(XI_n - B)S^{-1}) \\
&= \det(S)\det(XI_n - B)\det(S^{-1}) \\
&= \det(XI_n - B)\det(S)\det(S^{-1}) \\
&= \det(XI_n - B)\det(SS^{-1}) \\
&= \det(XI_n - B)\det(I_n) \\
&= \det(XI_n - B) \\
&= \chi_B.
\end{aligned}$$

[45] Die Aussage von Korollar 6.3.3 gilt sinngemäß auch für Matrizen mit Einträge in $K[X]$: Gegeben $C, D \in K[X]^{n\times n}$, so dass eine invertierbare Matrix $T \in K[X]^{n\times n}$ (vgl. Bemerkung 6.8.9) mit $D = TCT^{-1}$ existiert, so gilt $\det(C) = \det(D)$. In der Situation hier ist $S \in \mathrm{GL}_n(K)$ sicherlich auch invertierbar in $K[X]^{n\times n}$, und somit folgt die Behauptung des Satzes direkt aus der Gleichheit $S(XI_n - B)S^{-1} = XI_n - A$.

Da die Eigenwerte die Nullstellen des charakteristischen Polynoms sind (Satz 7.2.6), folgt die zweite Behauptung. □

Definition 7.2.18. Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums. Dann ist das **charakteristische Polynom von** f durch

$$\chi_f := \chi_{{}_{\mathcal{B}}[f]_{\mathcal{B}}} = \det(XI_n - {}_{\mathcal{B}}[f]_{\mathcal{B}})$$

definiert, wobei $\mathcal{B}$ eine beliebige Basis von V und $n = \dim V$ gilt. Nach Satz 5.12.10 und Satz 7.2.17 hängt χ_f nicht von der Wahl der Basis ab.

7.2.19. Das charakteristische Polynom χ_f ist normiert und hat Grad $\dim(V)$. Dies folgt sofort aus Satz 7.2.14.

7.2.20. Das charakteristische Polynom einer quadratischen Matrix $A \in K^{n\times n}$ stimmt mit dem charakteristischen Polynom des Endomorphismus $A\colon K^n \to K^n$ überein.

Satz 7.2.21. *Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums. Dann sind die Nullstellen von χ_f genau die Eigenwerte von f.*

Beweis. Sei $\mathcal{B}$ eine beliebige Basis von V. Nach Satz 7.2.6 sind die Nullstellen von χ_f genau die Eigenwerte von ${}_{\mathcal{B}}[f]_{\mathcal{B}}$. Wir behaupten, dass diese mit den Eigenwerten von f übereinstimmen.

Um dies zu zeigen, verwenden wir die Gleichheit $\varphi_{\mathcal{B}} \circ {}_{\mathcal{C}}[f]_{\mathcal{B}} = f \circ \varphi_{\mathcal{B}}$, die aus (5.10.2) in Proposition 5.10.5 folgt.

Sei λ ein Eigenwert von f. Sei $v \in V$ ein zugehöriger Eigenvektor, d. h. es gilt $f(v) = \lambda v$. Wir setzen $x := \varphi_{\mathcal{B}}^{-1}(v) \in K^n \setminus \{0\}$ und berechnen

$$\varphi_{\mathcal{B}}({}_{\mathcal{B}}[f]_{\mathcal{B}}x) = f(\varphi_{\mathcal{B}}(x)) = f(v) = \lambda v = \lambda\varphi_{\mathcal{B}}(x) = \varphi_{\mathcal{B}}(\lambda x).$$

Da $\varphi_{\mathcal{B}}$ injektiv ist, folgt ${}_{\mathcal{B}}[f]_{\mathcal{B}}x = \lambda x$, d. h. x ist ein Eigenvektor von ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ zum Eigenwert λ. Also ist λ auch ein Eigenwert von ${}_{\mathcal{B}}[f]_{\mathcal{B}}$.

Analog ist jeder Eigenwert von ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ auch ein Eigenwert von f, denn aus ${}_{\mathcal{B}}[f]_{\mathcal{B}}x = \lambda x$ folgt $f(\varphi_{\mathcal{B}}(x)) = \lambda\varphi_{\mathcal{B}}(x)$ durch Anwenden von $\varphi_{\mathcal{B}}$. □

Beispiel 7.2.22 (Endomorphismen von K^2)**.** Dieses Beispiel motiviert einige der nachfolgenden Begriffsbildungen (Trigonalisierbarkeit, Diagonalisierbarkeit) und Resultate.

Betrachte den Endomorphismus

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : K^2 \to K^2.$$

Nach Satz 7.2.6 ist $\lambda \in K$ genau dann ein Eigenwert von A, wenn λ eine Nullstelle von

$$\chi_A = \det(XI_2 - A) = X^2 - \underbrace{(a+d)}_{\mathrm{tr}(A)}X + \underbrace{ad-bc}_{=\det(A)}$$

ist. Da ein Polynom vom Grad 2 maximal zwei Nullstellen hat (Korollar 2.4.23), tritt genau einer der drei folgenden Fälle ein.

1. Fall. χ_A hat genau zwei (verschiedene) Nullstellen $\lambda_1 \neq \lambda_2$ in K (und ist somit von der Form $(X-\lambda_1)(X-\lambda_2)$):

Sei $v_i \in K^2$ ein Eigenvektor mit $Av_i = \lambda_i v_i$, für $i = 1, 2$. Nach Satz 7.1.13 (oder auch 7.1.6) sind v_1, v_2 linear unabhängig und bilden somit eine Basis $\mathcal{B}$ des K^2. Wir erhalten

$$ {}_{\mathcal{B}}[A]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{7.2.3} $$

Im Sinne der späteren Definition 7.4.4 ist A „diagonalisierbar“ (vgl. Korollar 7.4.8).

2. Fall. χ_A hat genau eine Nullstelle λ_1 in K (und ist somit von der Form $(X - \lambda_1)^2$, d. h. λ_1 ist eine doppelte Nullstelle):

Dann hat A genau einen Eigenwert, nämlich λ_1. Der zugehörige Eigenraum $\mathrm{Eig}(A, \lambda_1)$ ist als Untervektorraum $\neq \{0\}$ von K^2 entweder zwei- oder eindimensional.

(a) $\dim \mathrm{Eig}(A, \lambda_1) = 2$: Dies bedeutet $K^2 = \mathrm{Eig}(A, \lambda_1)$, also $Av = \lambda_1 v$ für alle $v \in K^2$, d. h.

$$ A = \lambda_1 \mathrm{id}_{K^2} = \lambda_1 I_2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}. \tag{7.2.4} $$

Wiederum ist A „diagonalisierbar“ (vgl. Satz 7.4.14). Da A ein Vielfaches der Identität ist, gilt ${}_{\mathcal{B}}[A]_{\mathcal{B}} = \lambda_1 I_2$ für jede Basis $\mathcal{B}$ von K^2.

(b) $\dim \mathrm{Eig}(A, \lambda_1) = 1$: Wir behaupten, dass K^2 eine Basis $\mathcal{B} = (v_1, v_2)$ mit

$$ {}_{\mathcal{B}}[A]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix} \tag{7.2.5} $$

hat. (Im Sinne der späteren Definition 7.3.3 ist A „trigonalisierbar“ (vgl. Satz 7.3.6); A ist aber nicht „diagonalisierbar“ (vgl. Satz 7.4.14).)

Sei v_1 ein Eigenvektor von A zu λ_1. Sei $v_2 \in K^2 \setminus \langle v_1 \rangle$. Dann sind v_1, v_2 linear unabhängig und bilden somit eine Basis von K^2. Es folgt $Av_2 = xv_2 + yv_1$ für geeignete $x, y \in K$. Da v_2 kein Eigenvektor von A ist, folgt $y \neq 0$. Indem wir die Gleichung mit y^{-1} multiplizieren und v_2 durch $y^{-1}v_2$ ersetzen, können wir ohne Einschränkung annehmen, dass $Av_2 = xv_2 + v_1$ gilt.

Wir behaupten, dass $x = \lambda_1$ gilt. Aus der Annahme $x \neq \lambda_1$ würde nämlich folgen, dass $w := v_2 + \frac{1}{x - \lambda_1} v_1$ ein Eigenvektor von A zum Eigenwert x ist, denn

$$ Aw = Av_2 + \frac{1}{x - \lambda_1} Av_1 = xv_2 + v_1 + \frac{1}{x - \lambda_1} \lambda_1 v_1 = xv_2 + \frac{x}{x - \lambda_1} v_1 = xw. $$

Dies widerspricht der Annahme, dass λ_1 der einzige Eigenwert von A ist. Also gilt $x = \lambda_1$ und somit $Av_2 = \lambda_1 v_2 + v_1$. Da v_1, v_2 linear unabhängig sind, ist $\mathcal{B} = (v_1, v_2)$ eine Basis von K^2. Nach Konstruktion gilt dann (7.2.5) wie gewünscht.

3. Fall. χ_A hat keine Nullstelle (und ist somit irreduzibel in $K[X]$):

In diesem Fall hat A keinen Eigenvektor. Nach Lemma 7.1.16 gibt es keine Basis von K^2, bezüglich der A durch eine obere Dreiecksmatrix dargestellt wird. Also ist A nicht „trigonalisierbar“ (vgl. Satz 7.3.6) und erst recht nicht „diagonalisierbar“.

Für algebraisch abgeschlossene Körper K, etwa für $K = \mathbb{C}$, tritt der dritte Fall nicht ein (nach dem Hauptsatz der Algebra 2.4.29). Die darstellenden Matrizen in (7.2.3), (7.2.4), (7.2.5), die wir für geeignete Basen erhalten haben, sind Beispiele für Matrizen in Jordanscher Normalform (siehe die spätere Definition 7.5.3).

Aufgabe 7.2.23. Geben Sie für jeden der drei Fälle in Beispiel 7.2.22 (samt weiterer Fallunterscheidung im zweiten Fall) eine reelle (2×2)-Matrix an, für die dieser Fall eintritt.

7.3. Trigonalisierbare Endomorphismen.

Definition 7.3.1. Eine quadratische Matrix $A \in K^{n\times n}$ heißt genau dann **trigonalisierbar**, wenn sie zu einer oberen Dreiecksmatrix ähnlich ist, wenn es also eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ gibt, so dass SAS^{-1} eine obere Dreiecksmatrix ist.

7.3.2. Trivialerweise ist jede obere Dreiecksmatrix trigonalisierbar.

Definition 7.3.3. Ein Endomorphismus $f\colon V \to V$ eines endlichdimensionalen Vektorraums heißt genau dann **trigonalisierbar**, wenn es eine Basis $\mathcal{B}$ von V gibt, so dass ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine obere Dreiecksmatrix ist, d. h.

$$ {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{7.3.1} $$

für geeignete $\lambda_1, \dots, \lambda_n \in K$.

7.3.4. Für einen Endomorphismus $f\colon V \to V$ eines endlichdimensionalen Vektorraums sind äquivalent:

(a) f ist trigonalisierbar;
(b) für jede Basis $\mathcal{B}$ von V ist ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine trigonalisierbare Matrix.
(c) für eine Basis $\mathcal{B}$ von V ist ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine trigonalisierbare Matrix.

Die Äquivalenz dieser Bedingungen folgt ebenfalls sofort aus Korollar 5.12.13.

Insbesondere ist eine quadratische Matrix $A \in K^{n\times n}$ genau dann trigonalisierbar, wenn der Endomorphismus $A\colon K^n \to K^n$ trigonalisierbar ist.

Satz 7.3.5. *Jede nilpotente Matrix $A \in K^{n\times n}$ ist trigonalisierbar und hat charakteristisches Polynom $\chi_A = X^n$ gilt.*[46]

Beweis. Sei $m \in \mathbb{N}$ so gewählt, dass $A^m = 0$. Offensichtlich haben wir Inklusionen

$$ \{0\} = \ker(I_n) = \ker(A^0) \subset \ker(A) = \ker(A^1) \subset \ker(A^2) \subset \cdots \subset \ker(A^m) = \ker(0) = K^n. $$

Das leere Tupel () ist eine Basis von $\{0\}$. Wir ergänzen es zu einer Basis $(b_1, \dots, b_{i_1})$ von $\ker(A)$. Dieses ergänzen wir zu einer Basis $(b_1, \dots, b_{i_1}, b_{i_1+1} \dots, b_{i_2})$ von $\ker(A^2)$ undsoweiter. Schließlich ist $\mathcal{B} := (b_1, \dots, b_{i_m})$ eine Basis von K^n (und es gilt notwendig $i_m = n$). Wegen $A(\ker(A^s)) \subset \ker(A^{s-1})$ ist ${}_{\mathcal{B}}[A]_{\mathcal{B}}$ eine obere Dreiecksmatrix mit Nullen auf der Diagonalen. Weil A und ${}_{\mathcal{B}}[A]_{\mathcal{B}}$ ähnlich sind (Korollar 5.12.13), ist A trigonalisierbar. Außerdem haben sie dasselbe charakteristische Polynom (Satz 7.2.17). Wir erhalten $\chi_A = \chi_{{}_{\mathcal{B}}[A]_{\mathcal{B}}} = X^n$ nach Beispiel 7.2.12.(c). □

Ende 22. Vorlesung am 09.07.2020

Satz 7.3.6. *Ein Endomorphismus $f\colon V \to V$ eines endlichdimensionaler K-Vektorraums ist genau dann trigonalisierbar, wenn χ_f in $K[X]$ vollständig in Linearfaktoren zerfällt.*

[46] Gilt $\chi_A = X^n$ für eine beliebige quadratische Matrix $A \in K^{n\times n}$, so ist A nilpotent. Dies folgt aus Satz 7.3.6.

Beweis. Sei f trigonalisierbar. Sei $\mathcal{B}$ eine Basis, so dass (7.3.1) für geeignete $\lambda_1, \dots, \lambda_n \in K$ gilt. Dann gilt $\chi_f = (X - \lambda_1)(X - \lambda_2) \dots (X - \lambda_n)$, wie wir bereits in Beispiel 7.2.12.(c) gesehen haben. Dies bedeutet, dass f vollständig in Linearfaktoren zerfällt.

Gelte umgekehrt, dass χ_f vollständig in Linearfaktoren zerfällt.

Wir beweisen per Induktion über $n = \dim V$, dass f trigonalisierbar ist. Für $n = 1$ (und auch für $n = 0$) ist dies trivial.

Sei $n \geq 2$. Gelte $\chi_f = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$ für $\lambda_1, \dots, \lambda_n \in K$. Dann ist λ_1 eine Nullstelle von χ_f und somit ein Eigenwert von f (Satz 7.2.21). Nach Lemma 7.1.16 gibt es eine Basis $\mathcal{B} = (v_1, v_2, \dots, v_n)$ von V mit

$$A := {}_{\mathcal{B}}[f]_{\mathcal{B}} = \left(\begin{array}{c|c} \lambda_1 & z \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right),$$

wobei $B \in K^{(n-1)\times(n-1)}$ und $z \in K^{1\times(n-1)}$. Es gilt $\chi_f(X) = (X - \lambda_1)\chi_B$ (siehe Beispiel 7.2.13). Durch Kürzen im Integritätsbereich $K[X]$ (2.2.6, Satz 2.4.13) folgt $\chi_B = (X - \lambda_2) \cdots (X - \lambda_n)$, d. h. χ_B zerfällt vollständig in Linearfaktoren. Per Induktionsannahme ist der Endomorphismus $B\colon K^{n-1} \to K^{n-1}$ trigonalisierbar, d. h. es gibt ein $S \in \mathrm{GL}_{n-1}(K)$, so dass SBS^{-1} eine obere Dreiecksmatrix ist. Die Matrix

$$T := \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & S & \\ 0 & & & \end{array}\right) \in \mathrm{GL}_n(K)$$

ist invertierbar und

$$TAT^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & S & \\ 0 & & & \end{array}\right) \cdot \left(\begin{array}{c|c} \lambda_1 & z \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right) \cdot \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & S^{-1} & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & zS^{-1} \\ \hline 0 & \\ \vdots & SBS^{-1} \\ 0 & \end{array}\right)$$

ist eine obere Dreiecksmatrix. Also ist die Matrix $A = {}_{\mathcal{B}}[f]_{\mathcal{B}}$ trigonalisierbar. Nach 7.3.4 bedeutet dies, dass f trigonalisierbar ist. □

Korollar 7.3.7. *Jeder Endomorphismus $f\colon V \to V$ eines endlichdimensionalen $\mathbb{C}$-Vektorraums ist trigonalisierbar.*

Beweis. Dies folgt aus Satz 7.3.6, denn nach Satz 2.4.31 zerfällt jedes Polynom in $\mathbb{C}[X] \setminus \{0\}$ vollständig in Linearfaktoren. □

7.4. **Diagonalisierbare Endomorphismen.**

7.4.1. Die folgende Begriffsbildung der Diagonalisierbarkeit ist strukturell vollkommen analog zur Trigonalisierbarkeit (siehe Abschnitt 7.3).

Definition 7.4.2. Eine quadratische Matrix $A \in K^{n\times n}$ heißt genau dann **diagonalisierbar**, wenn sie zu einer Diagonalmatrix ähnlich ist, wenn es also eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ gibt, so dass SAS^{-1} eine Diagonalmatrix ist.

7.4.3. Trivialerweise ist jede Diagonalmatrix diagonalisierbar.

Definition 7.4.4. Ein Endomorphismus $f\colon V \to V$ eines endlichdimensionalen Vektorraums heißt genau dann **diagonalisierbar**, wenn es eine Basis $\mathcal{B}$ von V gibt, so dass ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine Diagonalmatrix ist, d. h.

$$ {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{7.4.1} $$

für geeignete $\lambda_1, \dots, \lambda_n \in K$.

7.4.5. Für einen Endomorphismus $f\colon V \to V$ eines endlichdimensionalen Vektorraums sind äquivalent:

(a) f ist diagonalisierbar;
(b) es gibt eine Basis von V, die aus Eigenvektoren von f besteht. In Formeln: Es gibt eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V und Skalare $\lambda_1, \dots, \lambda_n \in K$ mit $f(v_i) = \lambda_i v_i$ für alle $i \in \{1, \dots, n\}$;
(c) es gibt ein Erzeugendensystem $(v_1, \dots, v_m)$ von V, das aus Eigenvektoren von f besteht;
(d) für jede Basis $\mathcal{B}$ von V ist ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine diagonalisierbare Matrix.
(e) für eine Basis $\mathcal{B}$ von V ist ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine diagonalisierbare Matrix.

Die Implikationen (a) $\Leftrightarrow$ (b) $\Rightarrow$ (c) sind klar. Die Implikation (c) $\Rightarrow$ (b) gilt, da sich jedes (bei uns per Definition endliche) Erzeugendensystem zu einer Basis verkleinern lässt (siehe Beweis von Satz 4.2.14). Die Implikationen (a) $\Leftrightarrow$ (d) $\Leftrightarrow$ (e) folgen sofort aus Korollar 5.12.13.

Insbesondere ist eine quadratische Matrix $A \in K^{n\times n}$ ist genau dann diagonalisierbar, wenn der Endomorphismus $A\colon K^n \to K^n$ diagonalisierbar ist. Dies folgt sofort aus Korollar 5.12.13.

Beispiele 7.4.6. In Beispiel 7.1.7 sind die Punktspiegelung und die Spiegelung an der ersten Winkelhalbierenden diagonalisierbar; die Drehung ist nicht diagonalisierbar.

Diagonalisierbar sind die Abbildungen in den Beispielen 7.1.8, 7.1.12 und 7.2.7.

Nicht diagonalisierbar ist die Abbildung in Beispiel 7.1.10.

Satz 7.4.7. *Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums. Seien $\lambda_1, \dots, \lambda_r$ sämtliche Eigenwerte von f, wobei $\lambda_i \neq \lambda_j$ für alle $i \neq j$ gelte (beachte, dass f nach Korollar 7.1.15 nur endlich viele Eigenwerte hat). Dann gilt*

$$ \sum_{i=1}^{r} \dim(\mathrm{Eig}(f, \lambda_i)) \leq \dim V \tag{7.4.2} $$

mit Gleichheit genau dann, wenn f diagonalisierbar ist.

Beweis. Betrachte die Abbildung

$$ g\colon \mathrm{Eig}(f, \lambda_1) \times \cdots \times \mathrm{Eig}(f, \lambda_r) \to V, $$

$$(v_1, \dots, v_r) \mapsto v_1 + \cdots + v_r.$$

Das Produkt der Eigenräume wird durch komponentenweise Addition und Skalarmultiplikation ein Vektorraum (vgl. Aufgabe 4.1.19); g ist dann offensichtlich eine lineare Abbildung.

Wir behaupten, dass g injektiv ist oder äquivalent, dass $\ker(g) = \{0\}$ gilt. Sei $(v_1, \dots, v_r) \in \ker(g)$, d. h. $v_1 + \cdots + v_r = 0$ für Vektoren $v_i \in \operatorname{Eig}(f, \lambda_i)$. Sei $I := \{i \mid v_i \neq 0\}$. Falls I nicht leer ist, ist $\sum_{i \in I} v_i = 0$ eine Darstellung der Null als nicht-triviale Linearkombination von Eigenvektoren zu paarweise verschiedenen Eigenwerten. Da Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind (Satz 7.1.13), kann es eine solche Darstellung nicht geben. Dies zeigt $I = \emptyset$, d. h. $0 = v_1 = \cdots = v_r$. Dies zeigt $\ker(g) = \{0\}$, also die Injektivität von g.

Wir begründen nun die folgenden Gleichheiten und die Ungleichung.

$$\sum_{i=1}^{r} \dim \operatorname{Eig}(f, \lambda_i) = \dim \Big(\operatorname{Eig}(f, \lambda_1) \times \cdots \times \operatorname{Eig}(f, \lambda_r) \Big) \tag{7.4.3}$$
$$= \dim \Big(\operatorname{Eig}(f, \lambda_1) + \ldots + \operatorname{Eig}(f, \lambda_r) \Big) \leq \dim V$$

Die erste Gleichheit gilt, da die Dimension eines Produkts die Summe der Dimensionen der Faktoren ist (Basen der Vektorräume $\operatorname{Eig}(f, \lambda_i)$ liefern eine Basis ihres Produkts liefern, vgl. Aufgabe 4.3.16). Die zweite Gleichheit folgt aus $\ker(g) = \{0\}$ und $\operatorname{im}(g) = \operatorname{Eig}(f, \lambda_1) + \ldots + \operatorname{Eig}(f, \lambda_r)$ und dem Dimensionssatz für lineare Abbildungen 5.5.3. Die Ungleichung ist offensichtlich (Satz 4.3.8). Damit ist (7.4.2) bewiesen.

Zu zeigen bleibt, dass die Diagonalisierbarkeit von f äquivalent zur Gleichheit in (7.4.3) ist.

Gilt Gleichheit in (7.4.3), so muss $\operatorname{Eig}(f, \lambda_1) + \cdots + \operatorname{Eig}(f, \lambda_r) = V$ gelten (Satz 4.3.8). Somit hat V ein Erzeugendensystem aus Eigenvektoren von f (jedes $\operatorname{Eig}(f, \lambda_i)$ hat ein Erzeugendensystem/eine Basis aus Eigenvektoren von f; schreibt man all diese hintereinander, so liefert dies ein Erzeugendensystem von V aus Eigenvektoren von f). Also ist f diagonalisierbar (siehe 7.4.5).

Sei umgekehrt f diagonalisierbar. Dann gibt es eine Basis $(b_1, \dots, b_n)$ von V aus Eigenvektoren von f. Jedes Basiselement b_j liegt in einem der Eigenräume $\operatorname{Eig}(f, \lambda_i)$, da unsere Liste von Eigenwerten vollständig ist. Also enthält $\operatorname{Eig}(f, \lambda_1) + \cdots + \operatorname{Eig}(f, \lambda_r)$ eine Basis von V und stimmt damit mit V überein. Also gilt Gleichheit in (7.4.3). □

Korollar 7.4.8. *Sei $f \colon V \to V$ ein Endomorphismus eines n-dimensionalen Vektorraums. Hat f genau n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.*

Beweis. Sei $\lambda_1, \dots, \lambda_n$ die Liste paarweise verschiedener Eigenwerte von f. Für jedes i gilt sicherlich $1 \leq \dim \operatorname{Eig}(f, \lambda_i)$. Summieren wir diese Ungleichungen und verwenden Satz 7.4.7, so erhalten wir

$$n \leq \sum_{i=1}^{n} \dim \operatorname{Eig}(f, \lambda_i) \leq \dim(V) = n.$$

Alle Kleiner-gleich-Zeichen sind somit Gleichheiten, so dass f nach Satz 7.4.7 diagonalisierbar ist. □

Beispiel 7.4.9 (Fibonacci-Zahlen; Herleitung einer expliziten Formel aus der induktiven Definition mit Hilfe von Diagonalisierbarkeit)**.** Betrachte die Fibonacci-Folge

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots$$

(vgl. [Wik19, Fibonacci-Folge]). Sie ist induktiv durch $F_0 = 0$, $F_1 = 1$ und $F_n = F_{n-1} + F_{n-2}$, für $n \geq 2$, definiert.

Definieren wir die lineare Abbildung

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} : \mathbb{C}^2 \to \mathbb{C}^2,$$

so gilt

$$A \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} F_{n+1} + F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{n+2} \\ F_{n+1} \end{pmatrix}$$

für alle $n \geq 0$. Wegen $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} F_1 \\ F_0 \end{pmatrix}$ erhalten wir daraus per Induktion sofort die Formel

$$A^n e_1 = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} \tag{7.4.4}$$

für alle $n \geq 0$. Die Eigenwerte von A sind die Nullstellen von

$$\chi_A = \det\left(\begin{pmatrix} X-1 & -1 \\ -1 & X \end{pmatrix}\right) = (X-1)X - 1 = X^2 - X - 1,$$

also (Mitternachtsformel, p-q-Formel) die beiden (verschiedenen reellen) Zahlen

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \qquad \text{(vgl. [Wik19, Goldener Schnitt])},$$
$$\lambda_2 = \frac{1-\sqrt{5}}{2}.$$

Nach Korollar 7.4.8 oder Unterergebnis (7.2.3) in Beispiel 7.2.22 ist A diagonalisierbar. Die Rechnung

$$A \begin{pmatrix} \lambda_i \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_i + 1 \\ \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i^2 \\ \lambda_i \end{pmatrix} = \lambda_i \begin{pmatrix} \lambda_i \\ 1 \end{pmatrix}$$

zeigt, dass $v_i := \begin{pmatrix} \lambda_i \\ 1 \end{pmatrix}$ ein Eigenvektor zu λ_i ist, für $i = 1, 2$. Also ist (v_1, v_2) eine Basis von $\mathbb{C}^2$ (klar oder Satz 7.1.13), die A diagonalisiert.

Eine kleine Rechnung liefert die folgende Darstellung von e_1 in dieser Basis:

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix} - \frac{1}{\sqrt{5}} \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \Big[v_1 - v_2 \Big].$$

Daraus folgt

$$A^n e_1 = \frac{1}{\sqrt{5}} \Big[\lambda_1^n v_1 - \lambda_2^n v_2 \Big]$$

für alle $n \in \mathbb{N}$. Wir erinnern daran, dass die zweite Koordinate von $A^n e_1$ nach (7.4.4) die n-te Fibonacci-Zahl F_n ist. Wir erhalten so die **Formel von Moivre/Binet**

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

Man beachte, dass der Ausdruck auf der rechten Seite für alle $n \in \mathbb{N}$ eine natürliche Zahl liefert, obwohl er einige irrationale Zahlen enthält.

Definition 7.4.10. Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums und $\lambda \in K$. Dann heißt

- die Dimension $\dim \operatorname{Eig}(f,\lambda)$ die **geometrische Vielfachheit von λ bezüglich f** und
- die Nullstellenordnung $\operatorname{ord}_\lambda(\chi_f)$ die **algebraische Vielfachheit von λ bezüglich f**.

7.4.11. Ist eine dieser Vielfachheiten von λ Null, so auch die andere, denn für jedes $\lambda \in K$ sind nach Satz 7.2.21 die folgenden Bedingungen äquivalent:

- $\dim \operatorname{Eig}(f,\lambda) = 0$;
- λ ist kein Eigenwert von f;
- $\chi_f(\lambda) \neq 0$, d. h. λ ist keine Nullstelle von χ_f;
- $\operatorname{ord}_\lambda(\chi_f) = 0$.

Beispiel 7.4.12. Betrachte die nilpotente Matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Wegen $\chi_A = X^2$ ist 2 die algebraische Vielfachheit von 0 (bezüglich A). Die geometrische Vielfachheit von 0 (bezüglich A) ist $1 = \dim \operatorname{Eig}(f,0) = \dim \ker(f)$.

Lemma 7.4.13. *Seien $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen K-Vektorraums und $\lambda \in K$. Die geometrische Vielfachheit von λ ist stets kleiner oder gleich der algebraischen Vielfachheit von λ, in Formeln*

$$\dim \operatorname{Eig}(f,\lambda) \leq \operatorname{ord}_\lambda(\chi_f).$$

Beweis. Sei $(v_1,\dots,v_m)$ eine Basis von $\operatorname{Eig}(f,\lambda)$. Wir ergänzen diese Basis zu einer Basis $\mathcal{B} = (v_1,\dots,v_m,v_{m+1},\dots,v_n)$ von V. Dann gilt

$${}_{\mathcal{B}}[f]_{\mathcal{B}} = \left(\begin{array}{c|c} \lambda I_m & C \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & C \\ 0 & & \lambda & \\ \hline & 0 & & D \\ & & & \end{array}\right).$$

Nach Beispiel 7.2.13 gilt $\chi_f = (X-\lambda)^m \cdot \chi_D$ und damit $\dim \operatorname{Eig}(f,\lambda) = m \leq \operatorname{ord}_\lambda(\chi_f)$. □

Ende 23. Vorlesung am 14.07.2020

Satz 7.4.14. *Ein Endomorphismus $f\colon V \to V$ eines endlichdimensionalen K-Vektorraums ist genau dann diagonalisierbar, wenn die beiden folgenden Bedingungen erfüllt sind:*

- *Das charakteristische Polynom χ_f zerfällt vollständig in Linearfaktoren.*

- *Für jeden Eigenwert $\lambda \in K$ von f stimmen algebraische und geometrische Vielfachheit bezüglich f überein, in Formeln* $\operatorname{ord}_\lambda(\chi_f) = \dim \operatorname{Eig}(f, \lambda)$.[47]

Insbesondere gilt im Falle $K = \mathbb{C}$ wegen Korollar 2.4.31: Die lineare Abbildung f genau dann diagonalisierbar, wenn $\operatorname{ord}_\lambda(\chi_f) = \dim \operatorname{Eig}(f, \lambda)$ *für jeden Eigenwert λ von f gilt.*

Beweis. $\Rightarrow$: Sei f diagonalisierbar. Dann gibt es eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V aus Eigenvektoren von f. Wir können die Basisvektoren so umordnen, dass

$$ {}_{\mathcal{B}}[f]_{\mathcal{B}} = \left(\begin{array}{c|c|c|c} \lambda_1 I_{m_1} & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 I_{m_2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & \lambda_r I_{m_r} \end{array}\right) = \begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & & & & & & & \lambda_r \end{pmatrix} $$

gilt, wobei die Eigenwerte $\lambda_1, \dots, \lambda_r \in K$ paarweise verschieden sind. Offensichtlich ist dann $(v_1, \dots, v_{m_1})$ eine Basis von $\operatorname{Eig}(f, \lambda_1)$ und $(v_{m_1+1}, \dots, v_{m_1+m_2})$ ist eine Basis von $\operatorname{Eig}(f, \lambda_2)$ undsoweiter. Es gilt also $\dim \operatorname{Eig}(f, \lambda_i) = m_i$.

Offensichtlich gilt $\chi_f(X) = (X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \cdots (X - \lambda_r)^{m_r}$, d. h. χ_f zerfällt in Linearfaktoren. Weiter gilt wie gewünscht $\operatorname{ord}_{\lambda_i}(\chi_f) = m_i = \dim \operatorname{Eig}(f, \lambda_i)$ für alle $i = 1, \dots, r$. Schließlich beachte man, dass $\{\lambda_1, \dots, \lambda_r\}$ genau die Menge der Eigenwerte von f ist (vgl. Beispiel 7.1.9).

$\Leftarrow$: Nach Annahme gilt $\chi_f(X) = (X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \cdots (X - \lambda_r)^{m_r}$ für geeignete paarweise verschiedene $\lambda_1, \dots, \lambda_r \in K$ und positive natürliche Zahlen $m_1, \dots, m_r \in \mathbb{N}_{\geq 0}$. Die Menge $\{\lambda_1, \dots, \lambda_r\}$ besteht somit genau aus den Nullstellen von χ_f alias den Eigenwerten von f. Außerdem gilt nach Annahme $\operatorname{ord}_{\lambda_i}(\chi_f) = \dim(\operatorname{Eig}(f, \lambda_i))$ für alle $i = 1, \dots, r$.

Da χ_f Grad $\dim(V)$ hat und da $\operatorname{ord}_{\lambda_i}(\chi_f) = m_i$ gilt, schließen wir

$$ \dim(V) = \deg(\chi_f) = \sum_{i=1}^{r} m_i = \sum_{i=1}^{r} \operatorname{ord}_{\lambda_i}(\chi_f) = \sum_{i=1}^{r} \dim \operatorname{Eig}(f, \lambda_i). $$

Nach Satz 7.4.7 bedeutet dies, dass f diagonalisierbar ist. □

Aufgabe 7.4.15. Ist $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{R}^{2\times 2}$ eine „symmetrische“ reelle Matrix, so ist A diagonalisierbar (*symmetrisch* bedeutet $A = A^{\mathrm{t}}$, siehe Definition 7.6.2).

Bemerkung: Allgemeiner gilt Satz 7.6.4.

7.5. **Jordansche Normalform.**

[47] Für jeden Nicht-Eigenwert λ gilt diese Gleichheit trivialerweise nach 7.4.11.

Definition 7.5.1. Sei $\lambda \in K$ und $d \in \mathbb{N}$. Der **Jordan-Block** oder das **Jordan-Kästchen** der Größe d zu λ ist die quadratische $(d \times d)$-Matrix

$$J_d(\lambda) := \sum_{i=1}^{d} \lambda E_{ii} + \sum_{i=1}^{d-1} E_{i,i+1} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in K^{d\times d}.$$

Alle Diagonaleinträge sind λ, alle Einträge auf der „oberen Nebendiagonalen“ sind Einsen (also alle Einträge an Positionen $(i, i+1)$, für $i = 1, \dots, d-1$), und alle restlichen Einträge sind Null.

Beispiel 7.5.2. Die Matrizen $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ oder $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ sind Jordan-Blöcke der Größe 3.

Definition 7.5.3. Eine **Matrix in Jordanscher Normalform** (kurz JNF) ist eine quadratische Blockdiagonalmatrix, deren Diagonalblöcke Jordan-Blöcke sind, also eine quadratische Matrix der Form

$$\begin{pmatrix} J_{d_1}(\lambda_1) & & & \\ & J_{d_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{d_r}(\lambda_r) \end{pmatrix} \tag{7.5.1}$$

für geeignete $r \in \mathbb{N}$, $d_1, \dots, d_r \in \mathbb{N} \setminus \{0\}$ und $\lambda_1, \dots, \lambda_r \in K$. Auf der Diagonalen dieser Matrix stehen die λ_i. Auf der „Nebendiagonalen“ darüber stehen je nach Größe der Jordan-Kästchen Nullen oder Einsen. Alle anderen Einträge sind Null.

7.5.4. Fassen wir die Matrix (7.5.1) als Endomorphismus eines geeigneten K^n auf (es ist $n = \sum_{i=1}^{r} d_i$), so sind die λ_i seine Eigenwerte. Man überlegt sich leicht, dass für jedes $\lambda \in K$ die geometrische Vielfachheit von λ bezüglich dieses Endomorphismus durch die Anzahl

$$\#\{i \in \{1, \dots, r\} \mid \lambda_i = \lambda\}$$

der Jordan-Blöcke mit λ als Diagonaleintrag gegeben ist und dass die algebraische Vielfachheit von λ bezüglich dieses Endomorphismus durch die Summe

$$\sum_{\substack{i\in\{1,\dots,r\},\\ \lambda_i=\lambda}} d_i$$

der Größen dieser Jordan-Blöcke gegeben ist.

Satz 7.5.5 (Jordansche Normalform - ohne Beweis)**.** *Sei $f\colon V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums. Ist f trigonalisierbar (nach Satz 7.3.6 können wir äquivalent verlangen, dass χ_f vollständig in Linearfaktoren zerfällt), so gibt es eine Basis $\mathcal{B}$*

von V, so dass ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine Matrix in Jordanscher Normalform ist, d. h. es gilt

$$ {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} J_{d_1}(\lambda_1) & & & \\ & J_{d_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{d_r}(\lambda_r) \end{pmatrix} \tag{7.5.2} $$

für geeignete positive natürliche Zahlen $d_1, \dots, d_r \in \mathbb{N} \setminus \{0\}$ und Körperelemente $\lambda_1, \dots, \lambda_r \in K$.

Genauer sind die Anzahl r der vorkommenden Jordan-Blöcke und die Typen der vorkommenden Jordan-Blöcke (also die Folge der Paare $(d_1, \lambda_1), (d_2, \lambda_2), \dots, (d_r, \lambda_r)$) bis auf Reihenfolge eindeutig bestimmt.

Korollar 7.5.6. *Für jeden Endomorphismus $f\colon V \to V$ eines endlichdimensionalen komplexen Vektorraums gibt es eine Basis $\mathcal{B}$ von V, so dass ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ eine Matrix in Jordanscher Normalform ist.*

Beweis. Dies folgt aus Satz 7.5.5, denn über den komplexen Zahlen zerfällt jedes Polynom $\neq 0$ in Linearfaktoren (siehe Korollar 2.4.31). □

7.6. Orthogonale Diagonalisierbarkeit symmetrischer reeller Matrizen (Hauptachsentransformation).

7.6.1. Wir erinnern an das in der Analysis definierte **Skalarprodukt**

$$ \langle -, - \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, $$

$$ (x, y) \mapsto \langle x, y \rangle := x^{\mathrm{t}} \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i. $$

Die **Länge** oder **Norm** von $x \in \mathbb{R}^n$ ist die nicht-negative reelle Zahl $\|x\| = \sqrt{\langle x, x \rangle} \in \mathbb{R}_{\geq 0}$. Man nennt x **normiert**, falls $\|x\| = 1$ gilt.

Man sagt, dass zwei Vektoren $x, y \in \mathbb{R}^n$ **senkrecht** oder **orthogonal aufeinander stehen**, falls $\langle x, y \rangle = 0$ gilt.

Vielleicht ist die folgende Definition aus der Schule bekannt: Der **unorientierte Winkel** $\sphericalangle(x, y)$ für Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ ist die eindeutige Zahl $\alpha \in [0, \pi]$ mit

$$ \cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}. $$

Man beachte, dass der Ausdruck rechts des Gleichheitszeichens auf Grund der Cauchy-Schwarzschen Ungleichung $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$ im Intervall $[-1, 1]$ liegt. Mit anderen Worten gilt $\alpha = \arccos\left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}\right)$.

Zwei Vektoren $x, y \in \mathbb{R}^n$ stehen also genau dann aufeinander senkrecht, wenn $\sphericalangle(x, y) = \frac{\pi}{2}$ (alias 90°) gilt.

Definition 7.6.2. Eine quadratische Matrix $A \in K^{n \times n}$ heißt symmetrisch, falls $A^{\mathrm{t}} = A$ gilt.

Beispiel 7.6.3. Die Matrix (7.6.1) ist symmetrisch. Bezüglich der Spiegelung an der Diagonalen sind die Einträge symmetrisch.

Satz 7.6.4 (Orthonormale Diagonalisierbarkeit symmetrischer reeller Matrizen - ohne Beweis)**.** *Jeder reelle symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ ist diagonalisierbar.*

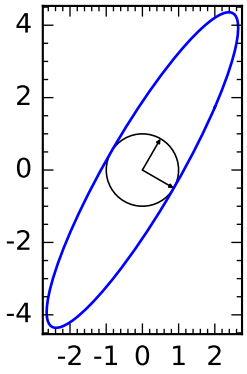
Genauer hat $\mathbb{R}^n$ eine „Orthonormalbasis“ $\mathcal{B} = (b_1, \dots, b_n)$ aus Eigenvektoren von $A\colon \mathbb{R}^n \to \mathbb{R}^n$. Dies bedeutet:

- *(Basis) $\mathcal{B} = (b_1, \dots, b_n)$ ist eine Basis von $\mathbb{R}^n$.*

- *(aus Eigenvektoren) Jedes b_i ist ein Eigenvektor der linearen Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}^n$, in Formeln $Ab_i = \lambda_i b_i$ für geeignete $\lambda_i \in \mathbb{R}$.*
  In anderen Worten ist ${}_{\mathcal{B}}[A]_{\mathcal{B}}$ eine Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \dots, \lambda_n$.
- *(Orthonormalbasis) Jedes b_i ist normiert und verschiedene Basiselemente stehen orthogonal aufeinander, in Formeln gilt $\langle b_i, b_j\rangle = \delta_{ij}$ für alle $i, j = 1, \dots n$.*[48]

7.6.5. Sind in der Notation des obigen Satzes alle λ_i verschieden, so nennt man die von den b_1 aufgespannten 1-dimensionalen Untervektorräume auch die Hauptachsen von A.

Beispiel 7.6.6. Die symmetrische Matrix $A = \left(\begin{smallmatrix} 2 & \sqrt{3} \\ \sqrt{3} & 4 \end{smallmatrix}\right)$ alias lineare Abbildung $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ hat die normierten Eigenvektoren $b_1 = \left(\begin{smallmatrix} \sqrt{3}/2 \\ -1/2 \end{smallmatrix}\right)$ zum Eigenwert 1 und $b_2 = \left(\begin{smallmatrix} 1/2 \\ \sqrt{3}/2 \end{smallmatrix}\right)$ zum Eigenwert 5. In der Illustration sind diese Vektoren, der Einheitskreis in schwarz und sein Bild unter A in blau eingezeichnet.



Die Hauptachsen von A sind die Symmetrieachsen der Ellipse (und diese werden ebenfalls als Hauptachsen bezeichnet).

Beispiel 7.6.7. Betrachte die symmetrische reelle Matrix

$$A = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 8 & -1 \\ 2 & -1 & 8 \end{pmatrix}. \tag{7.6.1}$$

Dann sind die Vektoren

$$v_1 := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \qquad v_2 := \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, v_3 := \qquad \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix}$$

Eigenvektoren von A, denn es gelten

$$Av_1 = 9v_1, \qquad Av_2 = 8v_2, \qquad Av_3 = -v_3.$$

Die Vektoren v_1, v_2, v_3 sind als Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig (Satz 7.1.13) und bilden somit eine Basis des $\mathbb{R}^3$ aus Eigenvektoren von A. Mit anderen Worten ist A diagonalisierbar.

[48] Aus dieser Gleichung folgt bereits, dass die $b_1, \dots, b_n$ linear unabhängig sind (und damit eine Basis des $\mathbb{R}^n$ bilden): Wende $\langle b_i, -\rangle$ auf eine Relation $0 = c_1 b_1 + \cdots + c_n b_n$ an.

Offensichtlich stehen je zwei verschiedene der Vektoren v_1, v_2, v_3 orthogonal aufeinander. Somit bilden die normierten Vektoren $\frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}v_1, \frac{v_2}{\|v_2\|} = \frac{1}{3}v_2, \frac{v_3}{\|v_3\|} = \frac{1}{3\sqrt{2}}v_3$ eine Orthonormalbasis aus Eigenvektoren von A.

Ende 24. Vorlesung am 16.07.2020

8. Euklidische Vektorräume

Kapitel im Sommersemester 2020 weggelassen.

8.0.1. In diesem Kapitel betrachten wir nur Vektorräume über den reellen Zahlen $\mathbb{R}$.

8.1. Skalarprodukte und euklidische Vektorräume.

Definition 8.1.1. Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung[49]

$$\begin{aligned}\langle -, - \rangle \colon V \times V &\to \mathbb{R},\\ (v, w) &\mapsto \langle v, w \rangle,\end{aligned}$$

heißt **Skalarprodukt**, wenn die folgenden Eigenschaften erfüllt sind:

(a) Die Abbildung ist linear in jedem Argument, wenn das andere Argument fixiert ist (*Bilinearität*):
   (i) Linearität im ersten Argument: Für jedes $w \in V$ ist die Abbildung
   $$\begin{aligned}\langle -, w \rangle \colon V &\to \mathbb{R},\\ w &\mapsto \langle v, w \rangle,\end{aligned}$$
   $\mathbb{R}$-linear.
   (ii) Linearität im zweiten Argument[50]: Für jedes $v \in V$ ist die Abbildung
   $$\begin{aligned}\langle v, - \rangle \colon V &\to \mathbb{R},\\ w &\mapsto \langle v, w \rangle,\end{aligned}$$
   $\mathbb{R}$-linear.

(b) Die Abbildung ist *symmetrisch*: Für alle $v, w \in V$ gilt
$$\langle v, w \rangle = \langle w, v \rangle.$$

(c) Für alle $v \in V$ gilt $\langle v, v \rangle \geq 0$ mit Gleichheit genau dann, wenn $v = 0$ gilt (*positive Definitheit*).

Man nennt $\langle v, w \rangle$ das **Skalarprodukt von v und w**.

Definition 8.1.2. Ein **euklidischer Vektorraum** ist ein endlichdimensionaler[51] reeller Vektorraum V zusammen mit einem Skalarprodukt $\langle -, - \rangle$. Wenn man sagt, dass V ein euklidischer Vektorraum ist, so ist ein zugehöriges Skalarprodukt implizit fixiert.

Beispiel 8.1.3. Der n-dimensionale reelle Vektorraum $\mathbb{R}^n$ zusammen mit dem sogenannten **Standardskalarprodukt**

$$\langle -, - \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

[49] Die horzontalen Striche sind keine Minuszeichen, sondern stehen für freie Argumente.

[50] Dies ist äquivalent zur Linearität im ersten Argument, wenn man bereits weiß, dass $\langle -, - \rangle$ symmetrisch ist.

[51] Einige der im Folgenden bewiesenen Aussagen gelten auch ohne diese Annahme.

$$(x,y) \mapsto \langle x,y\rangle := x^{\mathrm{t}} \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i,$$

ist ein euklidischer Vektorraum. Linearität in beiden Argumenten ist klar: Die Abbildung $\langle -, y\rangle \colon \mathbb{R}^n \to \mathbb{R}$ ist Rechtsmultiplikation mit dem Spaltenvektor y. Die Abbildung $\langle x, -\rangle \colon \mathbb{R}^n \to \mathbb{R}$ ist Linksmultiplikation mit dem Zeilenvektor x^{t}. Symmetrie ist auch klar (offensichtlich oder per $x^{\mathrm{t}} \cdot y = (x^{\mathrm{t}} \cdot y)^{\mathrm{t}} = y^{\mathrm{t}} \cdot (x^{\mathrm{t}})^{\mathrm{t}}) = y^{\mathrm{t}} \cdot x$). Zur positiven Definitheit: Da Quadrate reeller Zahlen und deren Summen stets nicht-negativ sind, gilt

$$\langle x,x\rangle = x_1^2 + \cdots + x_n^2 \geq 0.$$

Diese Summe ist genau dann Null, wenn alle Quadrate Null sind, wenn also alle x_i Null sind, wenn also $x = 0$ gilt.

8.1.4. Wenn wir im Folgenden von $\mathbb{R}^n$ als euklidischem Vektorraum sprechen, ist er stets mit dem Standardskalarprodukt versehen.

Beispiel 8.1.5. Untervektorräume euklidischer Vektorräume sind euklidische Vektorräume: Sei V zusammen mit dem Skalarprodukt $\langle -,-\rangle \colon V \times V \to \mathbb{R}$ ein euklidischer Vektorraum (etwa $\mathbb{R}^n$ mit dem Standardskalarprodukt). Sei U ein Untervektorraum von V. Dann ist die Abbildung

$$\begin{aligned} U \times U &\to \mathbb{R}, \\ (u,u') &\mapsto \langle u,u'\rangle \end{aligned}$$

ein Skalarprodukt auf dem Vektorraum U, denn alle definierenden Eigenschaften sind offensichtlich. Dieses Skalarprodukt macht U zu einem euklidischen Vektorraum.

8.2. **Cauchy-Schwarzsche Ungleichung und Dreiecksungleichung.** Im Rest dieses Kapitels sei V stets ein euklidischer Vektorraum. Sein Skalarprodukt wird wie oben als $\langle -,-\rangle$ notiert.

Definition 8.2.1. Zwei Vektoren $v, w \in V$ heißen genau dann **senkrecht zueinander**, notiert $v \perp w$, wenn $\langle v,w\rangle = 0$ gilt.

Man sagt dann auch, dass v und w **orthogonal zueinander** sind oder **senkrecht aufeinander stehen**, oder dass v **senkrecht auf** w **steht** oder dass w **senkrecht auf** v **steht**.

Beispiel 8.2.2. Im $\mathbb{R}^2$ mit dem Standardskalarprodukt stehen die beiden Vektoren $\begin{pmatrix}1\\2\end{pmatrix}$ und $\begin{pmatrix}2\\-1\end{pmatrix}$ senkrecht aufeinander.

Definition 8.2.3. Für $v \in V$ heißt

$$\|v\| := \sqrt{\langle v,v\rangle} \in \mathbb{R}_{\geq 0}$$

die **(euklidische) Norm von** v oder die **Länge von** v. Man beachte, dass $\langle v,v\rangle$ reell und ≥ 0 ist (wegen der positiven Definitheit), weswegen die Wurzel wieder reell ist; mit der Wurzel meinen wir die nicht-negative Wurzel.

Ein Vektor $v \in V$ heißt **normiert**, falls $\|v\| = 1$ gilt.

Anschauung 8.2.4. Unsere Definition der Länge entspricht der Anschauung, wenn wir den euklidischen Vektorraum $\mathbb{R}^2$ (alias die Zeichenebene) betrachten: Nach dem klassischen Satz des Pythagoras ist das Quadrat des Abstands eines Vektors $v = \begin{pmatrix} x \\ y \end{pmatrix}$ vom Nullpunkt gerade $x^2 + y^2$. Mehrmaliges Anwenden des klassischen Satzes von Pythagoras liefert die entsprechende Anschauung für den $\mathbb{R}^n$.

Lemma 8.2.5. *Seien $\lambda \in \mathbb{R}$ und $v, w \in V$. Dann gelten:*

(a) $\|v\| \geq 0$ *mit Gleichheit genau dann, wenn* $v = 0$.

(b) $\|\lambda v\| = |\lambda| \cdot \|v\|$.

(c) Falls v und w senkrecht aufeinander stehen, so gilt der ***Satz des Pythagoras***

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2.$$

(d) ***Parallelogrammgleichung***

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

Beweis. Der erste Punkt folgt aus der positiven Definitheit des Skalarprodukts. Der zweite Punkt folgt aus der Rechnung

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \cdot \|v\|.$$

Der Satz des Pythagoras folgt aus der Rechnung

$$\begin{aligned} \|v - w\|^2 &= \langle v - w, v - w \rangle = \langle v - w, v \rangle - \langle v - w, w \rangle \\ &= \langle v, v \rangle - \langle w, v \rangle - \langle v, w \rangle + \langle w, w \rangle = \langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle = \|v\|^2 - 2\underbrace{\langle v, w \rangle}_{=0} + \|w\|^2. \end{aligned}$$

Die Parallelogrammgleichung folgt aus der Rechnung

$$\begin{aligned} \|v + w\|^2 + \|v - w\|^2 &= \langle v + w, v + w \rangle + \langle v - w, v - w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle + \langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle = 2\|v\|^2 + 2\|w\|^2. \end{aligned}$$

□

Satz 8.2.6 (Cauchy-Schwarzsche Ungleichung)**.** *In einem euklidischen Vektorraum V gilt für alle Vektoren $v, w \in V$ die* ***Cauchy-Schwarzsche Ungleichung***

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|.$$

Es gilt Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis. Die Aussagen stimmen offensichtlich, wenn $w = 0$ gilt.

Gelte im Folgenden $w \neq 0$. Für beliebiges $\lambda \in \mathbb{R}$ gilt

$$0 \leq \langle v - \lambda w, v - \lambda w \rangle = \langle v, v \rangle - 2\lambda \langle v, w \rangle + \lambda^2 \langle w, w \rangle \tag{8.2.1}$$

Wegen $w \neq 0$ gilt $\langle w, w \rangle > 0$. Wir dürfen also $\lambda := \frac{\langle v, w \rangle}{\langle w, w \rangle}$ einsetzen. Multiplizieren wir außerdem mit $\langle w, w \rangle > 0$, so erhalten wir

$$0 \leq \langle v, v \rangle \langle w, w \rangle - 2\langle v, w \rangle^2 + \langle v, w \rangle^2 = \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle^2,$$

also

$$\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle$$

und durch Wurzelziehen (die Wurzelfunktion $\sqrt{-}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ ist (streng) monoton steigend)

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|.$$

Dies ist die behauptete Ungleichung.

Hier gilt genau dann Gleichheit, wenn in (8.2.1) für $\lambda = \frac{\langle v,w\rangle}{\langle w,w\rangle}$ Gleichheit gilt. Dies ist aber wegen positiver Definitheit äquivalent zu

$$v - \frac{\langle v, w\rangle}{\langle w, w\rangle} w = 0.$$

Daraus folgt, dass v, w linear abhängig sind. Seien umgekehrt v, w linear abhängig. Wegen $w \neq 0$ muss $v = \mu w$ für ein $\mu \in \mathbb{R}$ gelten. Daraus folgt die Gleichheit

$$|\langle v, w\rangle| = |\langle \mu w, w\rangle| = |\mu| \cdot |\langle w, w\rangle| = |\mu| \cdot \|w\|^2 = \|\mu w\| \cdot \|w\| = \|v\| \cdot \|w\|.$$

□

Korollar 8.2.7. *In einem euklidischen Vektorraum V gilt die* ***Dreiecksungleichung***

$$\|v + w\| \leq \|v\| + \|w\|$$

für alle Vektoren $v, w \in V$.

Beweis. Es gilt

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w\rangle \\
&= \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \\
&\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \qquad \text{(nach der CSU, siehe Satz 8.2.6)} \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Nun beachte man wieder, dass die Wurzelfunktion monoton wachsend ist. □

Aufgabe 8.2.8. Gegeben Elemente $u, v \in V$ setzen wir $\operatorname{dist}(u, v) := \|u - v\|$ und nennen dies die **Distanz** oder den **Abstand** von u und v. Zeigen Sie für alle $u, v, w \in V$:

- $\operatorname{dist}(u, v) \geq 0$ mit Gleichheit genau dann, wenn $u = v$ (positive Definitheit);
- $\operatorname{dist}(u, v) = \operatorname{dist}(v, u)$ (Symmetrie);
- $\operatorname{dist}(u, w) \leq \operatorname{dist}(u, v) + \operatorname{dist}(v, w)$ (Dreiecksungleichung).

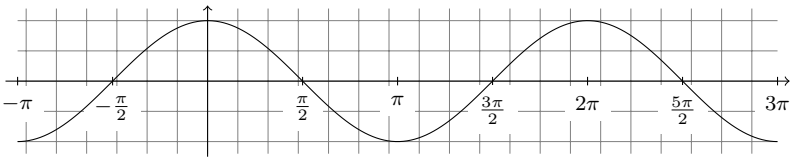
Diese Eigenschaften von $\operatorname{dist}\colon V \times V \to \mathbb{R}$ besagen, dass die Menge V zusammen mit dist einen **metrischen Raum** bildet. Ein metrischer Raum ist eine Menge mit einem sinnvollen Abstandsbegriff.

8.2.9. Seien $v, w \in V \setminus \{0\}$. Dann gilt $\|v\| \neq 0 \neq \|w\|$, und nach der Cauchy-Schwarzschen Ungleichung (siehe Satz 8.2.6) gilt

$$0 \leq \frac{|\langle v, w\rangle|}{\|v\| \cdot \|w\|} \leq 1,$$

also

$$-1 \leq \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \leq 1.$$



8.2.10. Wir erinnern an die hoffentlich aus der Analysis bekannte Aussage, dass die Kosinusfunktion das Intervall $[0, \pi]$ bijektiv auf das Intervall $[-1, 1]$ abbildet.

Definition 8.2.11 (Winkel zwischen zwei Vektoren eines euklidischen Vektorraums)**.** Der (unorientierte) **Winkel** $\sphericalangle(v, w)$ zwischen zwei Vektoren $v, w \in V \setminus \{0\}$ ist definiert als die eindeutige Zahl $\alpha \in [0, \pi]$ mit

$$\cos \alpha = \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}.$$

8.2.12. Der Winkel zwischen den Vektoren v und w ist genau dann der *rechte Winkel* $\frac{\pi}{2}$ (alias $90°$), wenn $\langle v, w \rangle = 0$ gilt, wenn also v und w wie oben definiert senkrecht aufeinander stehen. Dies entspricht der Anschauung.

Man nennt einen Winkel $\alpha \in [0, \pi]$ genau dann *spitz* bzw. *stumpf*, wenn $\alpha < \frac{\pi}{2}$ bzw. $\alpha > \frac{\pi}{2}$ gilt.

Der Winkel zwischen den Vektoren v und w ist also genau dann spitz bzw. stumpf, wenn $\langle v, w \rangle > 0$ bzw. $\langle v, w \rangle < 0$ gilt.

8.2.13. Der Winkel ist also so definiert, dass das Skalarprodukt von v und w das Produkt der Normen von v und w mit dem Kosinus des Winkels zwischen v und w ist, in Formeln

$$\langle v, w \rangle = \|v\| \cdot \|w\| \cdot \cos(\sphericalangle(v, w)).$$

Da der Kosinus nur Werte in $[-1, 1]$ annimmt, ist die rechte Seite $\leq \|v\| \cdot \|w\|$. Dies liefert eine „geometrische“ Interpretation der Cauchy-Schwarzschen Ungleichung.

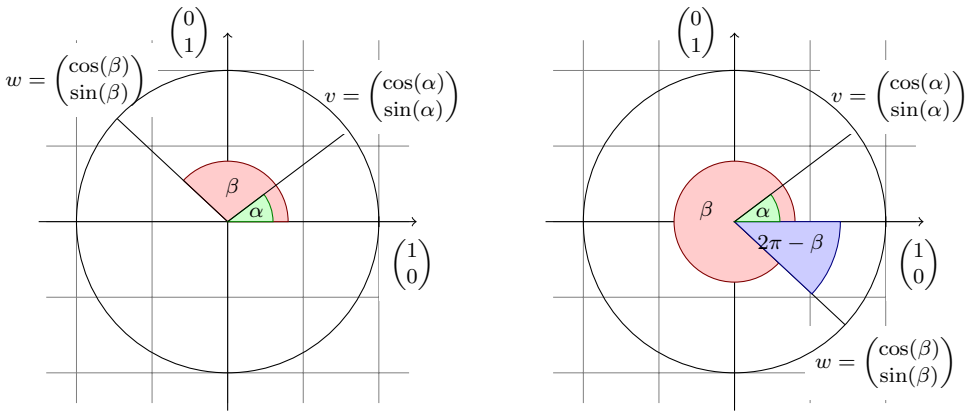
8.2.14. Aus der Definition ist klar, dass

$$\sphericalangle(w, v) = \sphericalangle(v, w) = \sphericalangle(\lambda v, w) = \sphericalangle(v, \lambda w)$$

für alle $v, w \in V \setminus \{0\}$ und alle positiven reellen Zahlen $\lambda > 0$ gilt.

Anschauung 8.2.15. Wir wollen begründen, warum unsere Definition des Winkels der Anschauung entspricht, wenn wir den euklidischen Vektorraum $\mathbb{R}^2$ (alias die Zeichenebene) betrachten.

Seien $v, w \in \mathbb{R}^2$ beliebig. Da sich der Winkel bei Skalierung der Argumente um eine positive reelle Zahl nicht ändert, können wir ohne Einschränkung annehmen (ersetze v durch $\frac{v}{\|v\|}$ und w durch $\frac{w}{\|w\|}$), dass $\|v\| = \|w\| = 1$ gilt, dass also v und w auf dem Kreis mit Radius eins um



den Ursprung liegen. Dann gibt es eindeutige $\alpha, \beta \in [0, 2\pi)$ mit[52]

$$v = \begin{pmatrix}\cos(\alpha)\\\sin(\alpha)\end{pmatrix}, \qquad w = \begin{pmatrix}\cos(\beta)\\\sin(\beta)\end{pmatrix},$$

vgl. die Illustration. Wir können ohne Einschränkung annehmen, dass $\alpha \leq \beta$ gilt. Wir erhalten unter Verwendung des Additionstheorems für den Kosinus und der Symmetrie des Kosinus

$$\langle v, w\rangle = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \cos(\beta - \alpha) = \cos(\alpha - \beta) = \cos(2\pi - \beta + \alpha).$$

Es gilt nach Annahme $0 \leq \beta - \alpha < 2\pi$. Im Fall $\beta - \alpha \leq \pi$ (linkes Bild) folgt also $\sphericalangle(v, w) = \beta - \alpha$, was der Anschauung entspricht. Im Fall $\pi < \beta - \alpha < 2\pi$ (rechtes Bild) gilt $0 < 2\pi - \beta + \alpha < \pi$ und $\sphericalangle(v, w) = 2\pi - \beta + \alpha$, was ebenfalls der Anschauung entspricht.

8.3. **Gram-Schmidtsches Orthonormalisierungsverfahren.**

8.3.1. Weiterhin sei V ein euklidischer Vektorraum.

Definition 8.3.2. Ein Tupel $(v_1, \ldots, v_m)$ von Vektoren in V heißt genau dann

(a) **orthogonal**, wenn $v_i \perp v_j$ für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$ gilt.
(b) **orthonormal**, wenn es orthogonal ist und zusätzlich $\|v_i\| = 1$ für alle $i \in \{1, \ldots, m\}$ gilt. Eine äquivalente Bedingung ist $\langle v_i, v_j\rangle = \delta_{ij}$ für alle $i, j \in \{1, \ldots, m\}$.
(c) **Orthonormalbasis (ONB) von** V, wenn es orthonormal ist und eine Basis von V bildet.

Beispiel 8.3.3. Die Standardbasis ist eine Orthonormalbasis des euklidischen Vektorraums $\mathbb{R}^n$ (mit dem Standardskalarprodukt).

---

[52] Anschaulich ist klar, dass für $\varphi \in [0, 2\pi)$ die Vektoren $\begin{pmatrix}\cos\varphi\\\sin\varphi\end{pmatrix}$ den Kreis mit Radius Eins um den Nullpunkt genau einmal durchlaufen. Formal kann man so vorgehen: Sei $\begin{pmatrix}a\\b\end{pmatrix} \in \mathbb{R}^2$ ein Vektor der Länge Eins, d. h. es gilt $a^2 + b^2 = 1$. Aus $a^2 + b^2 = 1$ folgt $0 \leq a^2 \leq 1$ und $-1 \leq a \leq 1$. Sei $\varphi' \in [0, \pi]$ die eindeutig bestimmte Zahl mit $\cos\varphi' = a$. Wegen $(\cos\varphi')^2 + (\sin\varphi')^2 = 1$ folgt $b^2 = (\sin\varphi')^2$, also $b = \sin\varphi'$ oder $b = -\sin\varphi'$.

Im Fall $b = \sin\varphi'$ setze $\varphi := \varphi'$. Sonst gilt $b = -\sin\varphi' = \sin(-\varphi') = \sin(2\pi - \varphi')$ und wir setzen $\varphi := 2\pi - \varphi'$ (man beachte, dass in diesem Fall $\varphi' \neq 0$ gilt, denn für $\varphi' = 0$ gilt $a = \cos\varphi' = \cos(0) = 1$, und damit $b = 0 = \sin(0) = \sin\varphi'$ und wir sind im ersten Fall).

Lemma 8.3.4. *Jedes orthogonale Tupel $(v_1, \ldots, v_m)$ von Vektoren $\neq 0$ in V ist linear unabhängig.*

Ist insbesondere $(v_1, \ldots, v_m)$ ein orthonormales Tupel von Vektoren in V und gilt $m = \dim V$, so ist $(v_1, \ldots, v_m)$ eine ONB von V.

Beweis. Seien $a_1, \ldots, a_m \in K$ mit

$$0 = a_1 v_1 + \ldots + a_m v_m.$$

Wende auf diese Gleichung die lineare Abbildung $\langle -, v_j \rangle \colon V \to \mathbb{R}$ an, wobei $j \in \{1, \ldots, m\}$ beliebig ist. Dies liefert aufgrund der angenommenen Orthogonalität

$$0 = a_1 \langle v_1, v_j \rangle + \ldots + a_j \langle v_j, v_j \rangle + \cdots + a_m \langle v_m, v_j \rangle = a_j \langle v_j, v_j \rangle = a_j \|v_j\|^2.$$

Wegen $v_j \neq 0$ folgt $\|v_j\| \neq 0$ und damit $a_j = 0$. Dies zeigt die lineare Unabhängigkeit.

Die zweite Behauptung folgt, da ein linear unabhängiges Tupel von $\dim V$ Vektoren automatisch eine Basis von V ist (siehe Satz 4.3.6.(b)). □

Satz 8.3.5 (Gram-Schmidtsches Orthonormalisierungsverfahren)**.** *Seien $v_1, \ldots, v_m$ linear unabhängige Vektoren in einem euklidischen Vektorraum. Dann existiert genau ein orthonormales Tupel $(w_1, \ldots, w_m)$ von Vektoren in V mit*[53]

$$w_i \in \mathbb{R}_{>0} v_i + \langle v_{i-1}, \ldots, v_1 \rangle \qquad \text{für alle } i = 1, \ldots, m.$$

Mit anderen Worten lässt sich jedes w_i als Linearkombination der $v_1, \ldots, v_i$ schreiben, wobei der Koeffizient bei v_i positiv ist.

Beweis der Eindeutigkeit des Tupels nicht in der Vorlesung.

Beweis. Wir beweisen die Aussage per Induktion über m.

Induktionsanfang: Für $m = 0$ ist die Aussage offensichtlich korrekt. Wer mag, kann auch mit $m = 1$ starten und

$$w_1 := \frac{1}{\|v_1\|} v_1$$

definieren. Dieser Vektor ist normiert und ein positives Vielfaches von v_1. Für die Eindeutigkeit sei $\lambda > 0$ so, dass λv_1 normiert ist, d. h. $1 = \|\lambda v_1\| = |\lambda| \|v_1\|$. Es folgt $|\lambda| = \frac{1}{\|v_1\|}$ (beachte $\|v_1\| \neq 0$ auf Grund der linearen Unabhngigkeit von v_1) beziehungsweise genauer $\lambda = \frac{1}{\|v_1\|}$, da λ positiv ist.

Induktionsschritt: Sei $m > 0$. Wir wenden die Induktionsannahme auf die linear unabhängigen Vektoren $v_1, \ldots, v_{m-1}$ an und erhalten genau ein orthonormales Tupel $(w_1, \ldots, w_{m-1})$ in V mit

$$w_i \in \mathbb{R}_{>0} v_i + \langle v_{i-1}, \ldots, v_1 \rangle \qquad \text{für alle } i = 1, \ldots, m-1.$$

Wir behaupten, dass $\langle w_1, \ldots, w_{m-1} \rangle = \langle v_1, \ldots, v_{m-1} \rangle$ gilt: Die Inklusion $\subset$ ist offensichtlich; für die Gleichheit reicht es somit zu zeigen, dass beide Vektorräume dieselbe Dimension haben (siehe Satz 4.3.8). Nach Annahme sind $v_1, \ldots, v_{m-1}$ linear unabhängig, und nach Lemma 8.3.4 sind die Vektoren $w_1, \ldots, w_{m-1}$ der Länge Eins ebenfalls linear unabhängig. Also haben beide Vektorräume Dimension $m - 1$.

[53] In der abgesetzten Formel stehen die spitzen Klammern für den Spann der angegebenen Vektoren, nicht für das Skalarprodukt (das ohnehin nur für zwei Vektoren definiert ist).

Wir behaupten, dass der Vektor

$$w' := v_m - \sum_{i=1}^{m-1} \langle v_m, w_i \rangle w_i \in v_m + \langle w_1, \ldots, w_{m-1} \rangle = v_m + \langle v_1, \ldots, v_{m-1} \rangle$$

senkrecht auf allen Vektoren $w_1, \ldots, w_{m-1}$ steht. In der Tat, es gilt

$$\langle w', w_j \rangle = \langle v_m, w_j \rangle - \sum_{i=1}^{m-1} \langle v_m, w_i \rangle \underbrace{\langle w_i, w_j \rangle}_{=\delta_{ij}} = \langle v_m, w_j \rangle - \langle v_m, w_j \rangle = 0$$

für alle $j = 1, \ldots, m-1$. Außerdem gilt $w' \neq 0$ auf Grund der linearen Unabhängigkeit der $v_1, \ldots, v_m$ (denn sonst wäre $0 \in v_m + \langle v_1, \ldots, v_{m-1} \rangle$).

Da der Vektor w' im Allgemeinen nicht normiert ist, setzen wir

$$w_m := \frac{1}{\|w'\|} w' \in \mathbb{R}_{>0} v_m + \langle v_1, \ldots, v_{m-1} \rangle.$$

Dann ist w_m normiert und steht als skalares Vielfaches von w' ebenfalls auf allen Vektoren $w_1, \ldots, w_{m-1}$ senkrecht.

Also ist $(w_1, \ldots, w_m)$ ein orthonormales Tupel, und nach Konstruktion ist klar, dass es die gewünschten Eigenschaften hat.

Für die Eindeutigkeit genügt es zu zeigen, dass wir bei der Konstruktion von w_m keine andere Wahl hatten. Sei $\widetilde{w}_m \in V$ so, dass das Tupel $(w_1, \ldots, w_{m-1}, \widetilde{w}_m)$ orthonormal ist und

$$\widetilde{w}_m \in \mathbb{R}_{>0} v_m + \langle v_{m-1}, \ldots, v_1 \rangle$$

gilt, sagen wir

$$\widetilde{w}_m \in t v_m + \langle v_{m-1}, \ldots, v_1 \rangle \qquad \text{für ein } t \in \mathbb{R}_{>0}.$$

Es folgt

$$t^{-1} \widetilde{w}_m \in v_m + \langle v_{m-1}, \ldots, v_1 \rangle = v_m + \langle w_{m-1}, \ldots, w_1 \rangle,$$

und dieser Vektor steht auf allen Vektoren $w_1, \ldots, w_{m-1}$ senkrecht. Schreiben wir

$$t^{-1} \widetilde{w}_m = v_m + \sum_{i=1}^{m-1} \alpha_i w_i$$

für geeignete (eindeutig bestimmte) Skalare $\alpha_i \in \mathbb{R}$ und wenden darauf $\langle -, w_j \rangle$ an, für beliebiges $j \in \{1, \ldots, m-1\}$, so liefert dies

$$0 = \langle t^{-1} \widetilde{w}_m, w_j \rangle = \langle v_m, w_j \rangle + \sum_{i=1}^{m-1} \alpha_i \underbrace{\langle w_i, w_j \rangle}_{=\delta_{ij}} = \langle v_m, w_j \rangle + \alpha_j,$$

also $\alpha_j = -\langle v_m, w_j \rangle$. Vergleich mit der Definition von w' liefert $w' = t^{-1} \widetilde{w}_m$. Nehmen wir die Norm beider Seiten, so erhalten wir

$$\|w'\| = t^{-1} \|\widetilde{w}_m\| = t^{-1},$$

und wir schließen $\widetilde{w}_m = t w' = \frac{1}{\|w'\|} w' = w_m$. Dies zeigt die Eindeutigkeit. □

Korollar 8.3.6. *Jeder euklidische Vektorraum besitzt eine ONB.*

Beweis. Sei V ein euklidischer Vektorraum. Da er per Definition endlichdimensional ist, besitzt er eine Basis $v_1, \dots, v_n$. Mit Gram-Schmidt (siehe Satz 8.3.5) finden wir ein orthonormales Tupel $(w_1, \dots, w_n)$ von Vektoren in V. Dieses Tupel ist eine ONB von V, wie direkt aus Lemma 8.3.4 folgt (oder alternativ aus dem Beweis von Satz 8.3.5). □

Satz 8.3.7 (Orthogonales Komplement)**.** *Sei U ein Untervektorraum eines euklidischen Vektorraums V. Dann ist*

$$U^\perp := \{v \in V \mid \langle v, u\rangle = 0 \text{ für alle } u \in U\}$$

ein Untervektorraum von V, der zu U komplementär ist, in Formeln $V = U \oplus U^\perp$. Er heißt ***orthogonales Komplement von*** *U* ***in*** *V.*

Insbesondere gilt $\dim(V) = \dim(U) + \dim(U^\perp)$.

Beweis. Sei $v \in U \cap U^\perp$. Dann gilt $\langle v, v\rangle = 0$, also $v = 0$. Dies zeigt $U \cap U^\perp = \{0\}$.

Nach Beispiel 8.1.5 ist U ein euklidischer Vektorraum. Nach Korollar 8.3.6 hat er eine ONB $(b_1, \dots, b_m)$. Betrachte die Abbildung

$$f\colon V \to \mathbb{R}^m,$$
$$v \mapsto \begin{pmatrix} \langle v, b_1\rangle \\ \vdots \\ \langle v, b_m\rangle \end{pmatrix}.$$

Diese Abbildung ist offensichtlich linear, und ihr Kern ist offensichtlich genau $U^\perp$. Dies zeigt die ohnehin offensichtliche Aussage, dass $U^\perp$ ein Untervektorraum von V ist.

Nach Definition einer ONB ist $f(b_i)$ der i-te Standardbasisvektor e_i des $\mathbb{R}^m$, für jedes $i = 1, \dots, m$. Dies zeigt, dass f surjektiv ist.

Die Dimensionsformel (5.5.1) für lineare Abbildungen zusammen mit der Dimensionsformel für zwei Untervektorräume (siehe Satz 4.3.14) zeigt

$$\begin{aligned}\dim(V) &= \dim(\ker(f)) + \dim(\operatorname{im}(f)) = \dim U^\perp + \dim \mathbb{R}^m = \dim(U^\perp) + m \\ &= \dim(U^\perp) + \dim(U) = \dim(U^\perp + U) + \dim(\underbrace{U^\perp \cap U}_{=\{0\}}) = \dim(U^\perp + U).\end{aligned}$$

Der Unterraum $U + U^\perp$ von V hat also dieselbe Dimension wie V. Daraus folgt $U + U^\perp = V$ (siehe Satz 4.3.8). □

Aufgabe 8.3.8. Jedes orthonormale Tupel von Vektoren in einem euklidischen Vektorraum läßt sich zu einer Orthonormalbasis ergänzen.

8.4. Orthogonale Matrizen.

8.4.1. Wir betrachten in diesem Abschnitt ausschließlich den euklidischen Vektorraum $\mathbb{R}^n$ mit dem Standardskalarprodukt $\langle x, y\rangle = x^{\mathrm{t}} \cdot y$ (siehe Beispiel 8.1.3).

Definition 8.4.2. Eine reelle quadratische Matrix $A \in \mathbb{R}^{n\times n}$ (und auch die zugehörige lineare Abbildung $A\colon \mathbb{R}^n \to R^n$) heißt genau dann **orthogonal**, wenn $A^{\mathrm{t}}A = I_n$ gilt. Sei

$$\mathrm{O}(n) := \{A \in \mathbb{R}^{n\times n} \mid A^{\mathrm{t}}A = I_n\}$$

die Menge aller orthogonalen Matrizen.

Lemma 8.4.3. *Für $A \in \mathbb{R}^{n\times n}$ sind die folgenden vier Bedingungen äquivalent:*

(a) *A ist orthogonal, d. h. $A^{\mathrm{t}}A = I_n$.*

(b) *Die Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}^n$ erhält das Standardskalarprodukt: Für alle $x, y \in \mathbb{R}^n$ gilt*

$$\langle Ax, Ay\rangle = \langle x, y\rangle.$$

(c) *Für alle $i, j \in \{1, \dots, n\}$ gilt*

$$\langle Ae_i, Ae_j\rangle = \delta_{ij}.$$

(d) *Die Spalten $(Ae_1, \dots, Ae_n)$ von A bilden eine ONB des $\mathbb{R}^n$.*

Insbesondere gelten:

- *Orthogonale Matrizen erhalten Längen und Winkel.*
- *Ist $A \in \mathrm{O}(n)$ orthogonal, so ist auch $A^{-1} = A^{\mathrm{t}}$ orthogonal, und es gilt $\det(A) = \pm 1$.*
- *Die Menge $\mathrm{O}(n)$ aller orthogonaler Matrizen ist eine Untergruppe der $\mathrm{GL}_n(\mathbb{R})$, die sogenannte **orthogonale Gruppe**.*

Beweis. Sei A orthogonal. Dann gilt für alle $x, y \in \mathbb{R}^n$

$$\langle Ax, Ay\rangle = (Ax)^{\mathrm{t}}Ay = x^{\mathrm{t}}A^{\mathrm{t}}Ay = x^{\mathrm{t}}I_n y = x^{\mathrm{t}}y = \langle x, y\rangle.$$

Also impliziert die erste Bedingung die zweite.

Dass die zweite Bedingung die dritte impliziert, ist klar.

Die dritte Bedingung impliziert die vierte, denn ein orthonormales Tupel aus n Vektoren ist automatisch eine ONB des $\mathbb{R}^n$, siehe Lemma 8.3.4.

Die vierte Bedingung liefert

$$(I_n)_{ij} = \delta_{ij} = \langle Ae_i, Ae_j\rangle = (Ae_i)^{\mathrm{t}}Ae_j = e_i^{\mathrm{t}}(A^{\mathrm{t}}A)e_j = (A^{\mathrm{t}}A)_{ij},$$

für alle $i, j \in \{1, \dots, n\}$, denn für eine beliebige Matrix B gilt $e_i^{\mathrm{t}}Be_j = B_{ij}$. Dies zeigt $I_n = A^{\mathrm{t}}A$, d. h. die erste Bedingung gilt.

Zu den Zusatzaussagen:

- Die Aussage über Längen und Winkel ist klar, da diese über das Skalarprodukt definiert sind: $\|Ax\| = \sqrt{\langle Ax, Ax\rangle} = \sqrt{\langle x, x\rangle} = \|x\|$ und damit $\cos(\sphericalangle(Ax, Ay)) = \frac{\langle Ax, Ay\rangle}{\|Ax\|\cdot\|Ay\|} = \frac{\langle x, y\rangle}{\|x\|\cdot\|y\|} = \cos(\sphericalangle(x, y))$, also $\sphericalangle(Ax, Ay) = \sphericalangle(x, y)$ wegen der Bijektivität des Kosinus $\cos\colon [0, \pi] \to [-1, 1]$.
- Sei $A \in \mathrm{O}(n)$ orthogonal, d. h. es gilt $A^{\mathrm{t}}A = I_n$. Dies zeigt $A^{-1} = A^{\mathrm{t}}$. Wegen $(A^{-1})^{\mathrm{t}}A^{-1} = (A^{\mathrm{t}})^{\mathrm{t}}A^{-1} = AA^{-1} = I_n$ ist auch A^{-1} orthogonal.

  Aus $A^{\mathrm{t}}A = I_n$ folgt $1 = \det(I_n) = \det(A^{\mathrm{t}}A) = \det(A^{\mathrm{t}})\det(A) = \det(A)\det(A)$, also $\det(A) \in \{\pm 1\}$.
- Die Einheitsmatrix I_n ist offensichtlich orthogonal, und wir haben gerade gesehen, dass mit A auch A^{-1} orthogonal ist. Sind A und B orthogonal, so gilt

  $$(AB)^{\mathrm{t}}AB = B^{\mathrm{t}}(A^{\mathrm{t}}A)B = B^{\mathrm{t}}I_n B = B^{\mathrm{t}}B = I_n,$$

  also $AB \in \mathrm{O}(n)$. Dies zeigt, dass $\mathrm{O}(n)$ eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$ ist.

□

Aufgabe 8.4.4. Sei $A\colon \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung, die im folgenden Sinne Längen erhält: Für alle $x \in \mathbb{R}^n$ gilt

$$\|Ax\| = \|x\|.$$

Dann gilt bereits $A \in \mathrm{O}(n)$.

Aufgabe 8.4.5 (etwas schwieriger, aber mit dem Hinweis hoffentlich gut lösbar)**.** Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ eine Abbildung von Mengen, die im folgenden Sinne abstandserhaltend ist: Für alle $x, y \in \mathbb{R}^n$ gilt

$$\|x - y\| = \|f(x) - f(y)\|.$$

Zusätzlich gelte $f(0) = 0$. Dann ist f bereits eine lineare bijektive Abbildung (also ein Isomorphismus von Vektorräumen), und ihre darstellende Matrix $[f]$ ist orthogonal.

Hinweis: Zeigen Sie zunächst $\langle f(x), f(y)\rangle = \langle x, y\rangle$ für alle $x, y \in \mathbb{R}^n$. Für die Linearität verwende man, dass $f(e_1), \ldots, f(e_n)$ eine ONB von $\mathbb{R}^n$ ist. Bijektivität ist dann offensichtlich.

Beispiel 8.4.6. Sei $n = 2$.

(a) Für $0 \leq \varphi < 2\pi$ (oder auch $\varphi \in \mathbb{R}$ beliebig) setze

$$A_\varphi := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Es ist klar, dass die beiden Spalten dieser Matrix senkrecht aufeinander stehen. Außerdem sind sie normiert, denn es gilt $(\cos(\varphi))^2 + (\sin(\varphi))^2 = 1$. Also bilden die Spalten von A_φ eine ONB von $\mathbb{R}^2$, d. h. $A_\varphi \in \mathrm{O}(2)$ ist orthogonal. Dieselbe trigonometrische Identität zeigt auch $\det(A_\varphi) = 1$.

Wir nennen A_φ die **Drehung um den Winkel** φ.

Wir bestätigen diese Terminologie, indem wir für einen beliebigen Vektor $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ den Winkel zwischen x und $A_\varphi x$ berechnen. Er ist definiert als das eindeutige $\alpha \in [0, \pi]$ mit

$$\begin{aligned} \cos\alpha &= \frac{\langle x, A_\varphi x\rangle}{\|x\| \cdot \|A_\varphi x\|} \\ &= \frac{\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1\cos\varphi - x_2\sin\varphi \\ x_1\sin\varphi + x_2\cos\varphi \end{pmatrix} \right\rangle}{\|x\| \cdot \|x\|} \\ &= \frac{x_1^2\cos\varphi - x_1x_2\sin\varphi + x_1x_2\sin\varphi + x_2^2\cos\varphi}{\|x\|^2} \\ &= \cos\varphi. \end{aligned}$$

Im Fall $0 \leq \varphi \leq \pi$ gilt also $\alpha = \pi$. Im Fall $\pi < \varphi < 2\pi$ gilt $\cos\varphi = \cos(-\varphi) = \cos(2\pi - \varphi)$ und $0 < 2\pi - \varphi < \pi$ und somit $\alpha = 2\pi - \varphi$; da per Definition des Winkels $0 \leq \alpha \leq \pi$ gilt, ist das das erwartete Ergebnis.

(b) Anschaulich ist klar, dass die Drehung um ψ gefolgt von der Drehung um φ die Drehung um $\varphi + \psi$ ist. Die Additionstheoreme für die trigonometrischen Funktionen (die man leicht aus der Eulerschen Formel $e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)$ herleitet) liefern in der Tat

$$A_\varphi \circ A_\psi = A_{\varphi+\psi} \tag{8.4.1}$$

für alle $\varphi, \psi \in \mathbb{R}$.

(c) Setze

$$T := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2.$$

Die Spalten von T bilden offensichtlich eine ONB von $\mathbb{R}^2$. Also gilt $T \in \mathrm{O}(2)$. Beachte $\det(T) = -1$.

Anschaulich ist T die Spiegelung an der x_1-Achse. Offensichtlich gilt $T^2 = I_2$.

Satz 8.4.7. *Sei $A \in \mathrm{O}(2)$ beliebig.*

(a) Ist $\det(A) = 1$, so existiert ein eindeutiges $\varphi \in [0, 2\pi)$ mit $A = A_\varphi$, d. h. A ist die Drehung um den Winkel φ.

(b) Ist $\det(A) = -1$, so existiert ein eindeutiges $\varphi \in [0, 2\pi)$ mit $A = A_\varphi T$. Anschaulich bedeutet dies, dass A die Spiegelung an der Achse ist, die zur x_1-Achse den Winkel $\varphi/2$ hat.

Beweis. Gelte $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Da die erste Spalte von A normiert ist, gilt $a^2 + c^2 = 1$. Sei $\varphi \in [0, 2\pi)$ die eindeutige Zahl mit $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$ (vgl. die Fußnote in Anschauung 8.2.15). Da die beiden Spalten von A stehen aufeinander senkrecht, gilt $ab+cd = 0$. Man folgert sofort $\begin{pmatrix} b \\ d \end{pmatrix} = \lambda \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$ für ein $\lambda \in \mathbb{R}$ (anschaulich ist das hoffentlich eh klar!). Da dieser Vektor ebenfalls normiert ist, gilt $1 = \lambda^2((\cos\varphi)^2 + (\sin\varphi)^2) = \lambda^2$, also $\lambda = \pm 1$. Wir berechnen $\det(A) = \lambda((\cos\varphi)^2 + (\sin\varphi)^2) = \lambda$.

Im Fall $\det(A) = 1$ gilt also $A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = A_\varphi$.

Im Fall $\det(A) = -1$ gilt also $A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = A_\varphi T$. Anschaulich ist hoffentlich klar, dass diese Matrix die angegebene Spiegelung beschreibt. Formal kann man dies wie folgt begründen.

Die Spiegelung an der Achse, die zur x_1-Achse den Winkel $\varphi/2$ hat, ist gegeben durch die Abbildung $A_{\varphi/2} T A_{-\varphi/2}$, denn man sieht unter Verwendung von (8.4.1) und $A_0 = I_2$ sofort, dass diese Abbildung den Vektor $A_{\varphi/2} e_1$, der die Spiegelachse aufspannt, fixiert, und den zu ihr senkrechten Vektor $A_{\varphi/2} e_2$ auf sein Negatives abbildet.

Es bleibt die Gleichheit $A_\varphi T = A_{\varphi/2} T A_{-\varphi/2}$ zu zeigen. Per Linksmultiplikation mit $A_{-\varphi/2}$ ist äquivalent $A_{\varphi/2} T = T A_{-\varphi/2}$ zu zeigen. Dies rechnet man direkt nach. □

Aufgabe 8.4.8. Es ist anschaulich klar, dass die Drehung $A = A(\varphi)$ des $\mathbb{R}^3$ um einen Winkel φ um die Dreh-Achse $\mathbb{R}. \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ eine orthogonale Abbildung ist. Bestimmen Sie die Matrix $A \in \mathrm{O}(3)$.

Satz 8.4.9 (Satz vom Fußball; Euler's rotation theorem)**.** *Sei $A \in \mathrm{O}(3)$ mit $\det(A) = 1$. Dann hat $A\colon \mathbb{R}^3 \to \mathbb{R}^3$ einen Eigenvektor zum Eigenwert 1.*

Anschauung 8.4.10. Der Satz besagt idealisiert, dass sich mindestenz zwei Punkte auf der Oberfläche eines Fußballs zu Beginn der ersten und der zweiten Halbzeit eines Fußballspiels an genau derselben Stelle befinden.

Dies sieht man so: Wir nehmen an, dass der Mittelpunkt des Balles beim Anstoß jeweils im Ursprung liegt, und dass der Ball Radius 1 hat. Die Standardbasis e_1, e_2, e_3 definiert zu Beginn der ersten Halbzeit drei Punkte p_1, p_2, p_3 auf dem Ball. Diese drei Punkte definieren

zu Beginn der zweiten Halbzeit eine ONB des $\mathbb{R}^3$. Sei $A \in \mathrm{O}(3)$ die Matrix, deren Spalten aus den Koordinaten dieser drei Punkte bestehen. Kurz gesagt liefert Anwenden von A auf den Ball zu Beginn der ersten Halbzeit den Ball zu Beginn der zweiten Halbzeit.

Analog kann man zu jedem Zeitpunkt t den Ball gedanklich auf den Anstoßpunkt legen und erhält so eine Matrix $A(t) \in \mathrm{O}(3)$. Insgesamt erhält man eine Abbildung $[0, 90] \to \mathrm{O}(3)$, $t \mapsto A(t)$, wobei wir eventuelle Nachspielzeiten und Pausen vernachlässigen. Da die Abbildung $[0, 90] \to \mathbb{R}$, $t \mapsto \det(A(t))$, stetig ist, nur die Werte ± 1 annehmen kann, und $\det(A(0)) = \det(I_3) = 1$ gilt, folgt $\det(A(t)) = 1$ für alle t (dies verwendet den Zwischenwertsatz aus der Analysis). Insbesondere erhalten wir, dass die Matrix $A = A(45)$ Determinante Eins hat.

Nach Satz 8.4.9 hat A einen Eigenvektor v mit $Av = v$. Ohne Einschränkung können wir annehmen, dass $\|v\| = 1$ gilt. Dann ist v ein Punkt auf der Oberfläche des Fussballs, der sich zu Beginn beider Halbzeiten am selben Ort befindet. Dasselbe gilt für seinen Antipodenpunkt $-v$.

Beweis. Beobachtung: Hat $A \in \mathrm{O}(n)$ einen reellen Eigenwert $\lambda \in \mathbb{R}$, so gilt $\lambda \in \{\pm 1\}$. In der Tat, ist $v \in \mathbb{R}^n$ ein Eigenvektor von A zum Eigenwert λ, so folgt dies sofort aus $\|v\| = \|Av\| = \|\lambda v\| = |\lambda|\|v\|$ und $v \neq 0$.

Sei nun $A \in \mathrm{O}(3)$ mit $\det(A) = 1$. Sei $\chi_A \in \mathbb{R}[X]$ das charakteristische Polynom von A. Als reelles normiertes Polynom ungeraden Grades hat χ_A eine reelle Nullstelle $\lambda \in \mathbb{R}$ (dies folgt aus dem Zwischenwertsatz und der Beobachtung, dass ein solches Polyom als Funktion auf $\mathbb{R}$ für $x \to \infty$ gegen $+\infty$ und für $x \to -\infty$ gegen $-\infty$ strebt, oder aus Aufgabe 2.7.7). Durch Abspalten der Nullstelle erhalten wir $\chi_A = (X - \lambda)q$ für ein quadratisches normiertes Polynom $q \in \mathbb{R}[X]$.

1. Fall: Es hat q eine relle Nullstelle $\mu \in \mathbb{R}$. Durch Abspalten dieser Nullstelle sehen wir, dass q eine weitere reelle Nullstelle $\nu \in \mathbb{R}$ hat. Es gilt also $\chi_A = (X - \lambda)(X - \mu)(X - \nu)$ und somit $\lambda\mu\nu = \det(A) = 1$. Da die drei Nullstellen reelle Eigenwerte von A sind, zeigt die obige Beobachtung $\lambda, \mu, \nu \in \{\pm 1\}$, und somit muss eine dieser drei Zahlen Eins sein. Dies zeigt, dass Eins ein Eigenwert von A ist.

2. Fall: Es hat q keine weitere reelle Nullstelle. Jedoch hat q (als Element von $\mathbb{C}[X]$) zwei Nullstellen $z = a + bi$ und $\overline{z} = a - bi$ (dass diese konjugiert sind, folgt zum Beispiel aus der Lösungsformel für quadratische Gleichungen). In $\mathbb{C}[X]$ gilt also $\chi_A = (X - \lambda)(X - z)(X - \overline{z})$ und somit $1 = \det(A) = \lambda z\overline{z} = \lambda|z|^2$. Wegen $|z|^2 \geq 0$ zeigt unsere Beobachtung $\lambda = 1$. Also ist Eins ein Eigenwert von A. □

Aufgabe 8.4.11. Sei $A \in \mathrm{O}(3)$ mit $\det(A) = 1$. Dann gibt es eine ONB $\mathcal{B} = (v_1, v_2, v_3)$ von V und ein $\varphi \in [0, 2\pi)$ mit

$$ {}_{\mathcal{B}}[A]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}. $$

Es ist also A eine Drehung um den Winkel φ um die Drehachse $\mathbb{R}v_1$.

Definition 8.4.12. Seien Vektoren $v_1, \ldots, v_n \in \mathbb{R}^n$ gegeben. Das von diesen Vektoren **aufgespannte Parallelepiped** ist die Menge

$$ P_{v_1,\ldots,v_n} := \{t_1v_1 + \cdots + t_nv_n \mid t_1, \ldots, t_n \in [0, 1]\}. $$

Das **Volumen** dieser Menge ist definiert als Betrag der Determinante der Matrix, deren Spalten die aufspannenden Vektoren sind, also

$$\operatorname{vol}(P_{v_1,\dots,v_n}) := |\det(v_1 \mid v_2 \mid \cdots \mid v_n)|.$$

8.4.13. Wir listen einige Eigenschaften auf, um den Leser zu überzeugen, dass diese Definition des Volumens sinnvoll ist, vgl. auch Anschauung 6.1.4, wo wir den zweidimensionalen Fall diskutiert haben.

Das Volumen ändert sich nicht, wenn man die Vektoren $v_1, \dots, v_n$ permutiert, wenn man einen der Vektoren mit seinem Negativen ersetzt, wenn man zu einem der Vektoren ein beliebiges Vielfaches eines anderen Vektors hinzuaddiert (*Scherung* des Parallelepipeds). Ersetzt man einen Vektor durch sein λ-faches, so verändert sich das Volumen um den Faktor $|\lambda|$. Das Volumen des von den Standardbasisvektoren aufgespannten Würfels ist Eins. Wie wir sofort zeigen werden (siehe Satz 8.4.14), ist das Volumen eines „Quaders“, also eines Parallelepipeds, dessen aufspannende Vektoren paarweise aufeinander senkrecht stehen, das Produkt der Längen der aufspannenden Vektoren. (Das Volumen ist jedoch nicht linear in v_i, falls die anderen Vektoren $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ fixiert sind; dies liegt daran, dass der Betrag einer Summe reeller Zahlen nicht die Summe der Beträge ist.)

Satz 8.4.14 (Volumen von Quadern)**.** *Sei $(v_1, \dots, v_n)$ ein orthogonales Tupel in $\mathbb{R}^n$. Sei $Q := P_{v_1,\dots,v_n}$ der von den Vektoren $v_1, \dots, v_n$ aufgespannte Quader. Dann gilt*

$$\operatorname{vol}(Q) = \|v_1\| \cdots \|v_2\| \dots \|v_n\|.$$

Beweis. Setze $v_i' := \frac{v_i}{\|v_i\|}$. Dann ist $v_1', \dots, v_n'$ eine ONB von $\mathbb{R}^n$. Sei $A' := (v_1' \mid \cdots \mid v_n') \in \mathbb{R}^{n\times n}$ die entsprechende Matrix. Sie ist orthogonal und es gilt $\det(A) = \pm 1$ (siehe Lemma 8.4.3).

Wir berechnen mit den Rechenregeln für Determinanten

$$\begin{aligned}
\operatorname{vol}(Q) &= |\det(v_1 \mid v_2 \mid \cdots \mid v_n)| \\
&= |\det(\|v_1\|.v_1' \mid \|v_2\|.v_2' \mid \cdots \mid \|v_n\|.v_n')| \\
&= \|v_1\| \dots \|v_n\| \cdot |\det(v_1' \mid v_2' \mid \cdots \mid v_n')| \\
&= \|v_1\| \dots \|v_n\| \cdot |\det(A')| \\
&= \|v_1\| \cdots \|v_n\|.
\end{aligned}$$

□

Proposition 8.4.15 (Betrag der Determinante als Volumenänderungsfaktor)**.** *Sei $A\colon \mathbb{R}^n \to \mathbb{R}^n$ ein Endomorphismus. Seien Vektoren $v_1, \dots, v_n \in \mathbb{R}^n$ gegeben, und seien $w_1 := Av_1, \dots, w_n := Av_n$ ihre Bilder unter der Abbildung A. Dann unterscheiden sich die Volumina der zugehörigen Parallelepipede $A(P_{v_1,\dots,v_n}) = P_{w_1,\dots,w_n}$ und $P_{v_1,\dots,v_n}$ um den Faktor $|\det(A)|$, in Formeln*

$$\operatorname{vol}(P_{w_1,\dots,w_n}) = |\det(A)| \cdot \operatorname{vol}(P_{v_1,\dots,v_n}).$$

Insbesondere erhalten orthogonale Matrizen das Volumen von Parallelepipeden.

Beweis. Das folgt durch Einsetzen der Definitionen unter Verwendeung der Multiplikativität der Determinante (siehe Satz 6.3.1)

$$\begin{aligned}
\operatorname{vol}(P_{w_1,\dots,w_n}) &= |\det(w_1 \mid \cdots \mid w_n)| = |\det(Av_1 \mid \cdots \mid Av_n)| = |\det(A \cdot (v_1 \mid \cdots \mid v_n)| \\
&= |\det(A)\det(v_1 \mid \cdots \mid v_n)| = |\det(A)| \cdot |\det(v_1 \mid \cdots \mid v_n)| = |\det(A)| \cdot \operatorname{vol}(P_{v_1,\dots,v_n}).
\end{aligned}$$

Die letzte Behauptung folgt aus Lemma 8.4.3, da orthogonale Matrizen Determinante ± 1 haben. □

Satz 8.4.16 (Iwasawa-Zerlegung oder KAN[54]-Zerlegung für $\mathrm{GL}_n(\mathbb{R})$, liefert QR[55]-Zerlegung)**.** *Sei $A \subset \mathrm{GL}_n(\mathbb{R})$ die Menge[56] aller Diagonalmatrizen mit positiven Einträgen auf der Diagonalen, und sei $N \subset \mathrm{GL}_n(\mathbb{R})$ die Menge[57] aller oberen Dreiecksmatrizen mit Einsen auf der Diagonalen. Dann ist die Abbildung „Multiplikation“*

$$\begin{aligned} \mathrm{O}(n) \times A \times N &\to \mathrm{GL}_n(\mathbb{R}), \\ (k, t, u) &\mapsto ktu, \end{aligned}$$

eine Bijektion. In Worten lässt sich also jede invertierbare Matrix in eindeutiger Weise als Produkt einer orthogonalen Matrix, einer Diagonalmatrix mit positiven Einträgen auf der Diagonalen, und einer oberen Dreiecksmatrix mit Einsen auf der Diagonalen schreiben.

Beweis. Die Abbildung ist offenbar wohldefiniert.

Injektivität: Gelte $ktu = k't'u'$ für $k, k' \in \mathrm{O}(n)$, $t, t' \in A$ und $u, u' \in N$. Es folgt

$$C := k'^{-1}k = t'u'u^{-1}t^{-1}.$$

Es gilt $k'^{-1}k \in \mathrm{O}(n)$. Da N ein Untergruppe von $\mathrm{GL}_n(\mathbb{R})$ ist, ist klar, dass $t'u'u^{-1}t^{-1}$ eine obere Dreiecksmatrix mit positiven Diagonaleinträgen ist.

Also ist C sowohl orthogonal als auch eine obere Dreiecksmatrix mit positiven Diagonaleinträgen. Da die Spalten von C eine Orthonormalbasis des $\mathbb{R}^n$ (Lemma 8.4.3) bilden, folgt $C = I_n$. Dies liefert $k' = k$ und $t'^{-1}t = u'u^{-1} \in A \cap N = \{I_n\}$, also $t' = t$ und $u' = u$.

Surjektivität: Sei $B \in \mathrm{GL}_n(\mathbb{R})$ gegeben. Seien $s_1, \dots, s_n$ die Spalten von B und sei $\mathcal{B} = (s_1, \dots, s_n)$ die von diesen Spalten gebildete Basis des $\mathbb{R}^n$. Das Gram-Schmidtsches Orthonormalisierungsverfahren (Satz 8.3.5) zusammen mit Lemma 8.3.4 liefert eine (eindeutige) Orthonormalbasis $\mathcal{C} = (c_1, \dots, c_n)$ von V mit

$$c_i \in \mathbb{R}_{>0} s_i + \langle s_{i-1}, \dots, s_1 \rangle \qquad \text{für alle } i = 1, \dots, n.$$

Dies bedeutet, dass die Basiswechselmatrix

$$E := {}_{\mathcal{B}}[\mathrm{id}_{\mathbb{R}^n}]_{\mathcal{C}}$$

eine obere Dreiecksmatrix mit positiven Einträgen auf der Diagonalen ist.

Sei $C \in K^{n \times n}$ die Matrix, deren Spalten die Vektoren $c_1, \dots, c_n$ sind. Es gilt $C \in \mathrm{O}(n)$ wegen Lemma 8.4.3.

Laut Definition der Basiswechselmatrix und wegen $\varphi_{\mathcal{B}} = B \colon \mathbb{R}^n \xrightarrow{\sim} \mathbb{R}^n$ und $\varphi_{\mathcal{C}} = C \colon \mathbb{R}^n \xrightarrow{\sim} \mathbb{R}^n$ gilt

$$BE = C$$

[54] Der Buchstabe K steht für kompakt, denn $\mathrm{O}(n)$ ist eine maximale kompakte Untergruppe von $\mathrm{GL}_n(\mathbb{R})$. Der Buchstabe A steht für abelsch (auf Niveau der Lie-Algebren) und N steht für nilpotent (ebenfalls auf Niveau der Lie-Algebren).

[55] Der Buchstabe R stecht für rechte Dreiecksmatrix = obere Dreiecksmatrix. Der Buchstabe Q scheint recht willkürlich gewählt, er kommt wohl schlicht im Alphabet vor R.

[56] Diese Menge ist offensichtlich eine Untergruppe.

[57] Auch diese Menge ist offensichtlich eine Untergruppe (vgl. Aufgabe 5.13.5): Ist R eine echte obere Dreiecksmatrix, so ist $I_n + R$ invertierbar mit Inversem $I_n - R + R^2 + \cdots \pm R^{n-1}$ (wegen $R^n = 0$), was wiederum in N liegt. Das Produkt zweier Elemente von N ist wieder in N, und $I_n \in N$ ist klar.

oder äquivalent

$$B = CE^{-1}.$$

Es ist C orthogonal, und es ist leicht zu sehen, dass mit E auch E^{-1} eine obere Dreiecksmatrix mit positiven Einträgen auf der Diagonalen ist und damit als Produkt $E^{-1} = tu$ geschrieben werden kann, für $t \in A$ und $u \in N$. Dies liefert $B = Ctu$ und zeigt die Surjektivität. □

8.4.17. Sei $B = ktu$ die Iwasawa-Zerlegung von $B \in \mathrm{GL}_n(K)$. Dann gilt

$$\det(B) = \det(k)\det(t)\det(u) = \pm\det(t).$$

Insbesondere ist das Volumen des von den Spalten von B aufgespannten Parallelepipeds gerade $\prod_{i=1}^{n} \lambda_i$, wenn $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_{>0}$ die (positiven) Diagonaleinträge von B sind.

Dies ist auch anschaulich plausibel: Die Spalten von k spannen einen Würfel mit Seitenlänge Eins auf, die Spalten von kt einen Quader mit Seitenlängen $\lambda_1, \ldots, \lambda_n$, und $B = ktu$ entsteht aus diesem Quader durch Scherung, wobei sich das Volumen nicht ändert.

Literatur

[dO11] Oswaldo Rio Branco de Oliveira. The fundamental theorem of algebra: an elementary and direct proof. *Math. Intelligencer*, 33(2):1–2, 2011.

[dO12] Oswaldo Rio Branco de Oliveira. The fundamental theorem of algebra: from the four basic operations. *Amer. Math. Monthly*, 119(9):753–758, 2012.

[Sch20] Olaf Schnürer. Analysis für informatiker. *Skript, Paderborn*, 2020. math.uni-paderborn.de/ag/arbeitsgruppe-algebra/team/olaf-schnuerer/.

[Sed90] Robert Sedgewick. *Algorithms in C*. Addison-Wesley, 1990.

[Wik19] Contributors Wikipedia. Wikipedia, 2019. www.wikipedia.org.

Institut für Mathematik, Universität Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany

Email address: olaf.schnuerer@math.uni-paderborn.de